Theor Appl Genet (2001) 103:804–816 © Springer-Verlag 2001

J. Wang · D.W. Podlich · M. Cooper · I.H. DeLacy

Power of the joint segregation analysis method for testing mixed major-gene and polygene inheritance models of quantitative traits

Received: 18 August 2000 / Accepted: 9 March 2001

Abstract Understanding the genetic architecture of quantitative traits can greatly assist the design of strategies for their manipulation in plant-breeding programs. For a number of traits, genetic variation can be the result of segregation of a few major genes and many polygenes (minor genes). The joint segregation analysis (JSA) is a maximum-likelihood approach for fitting segregation models through the simultaneous use of phenotypic information from multiple generations. Our objective in this paper was to use computer simulation to quantify the power of the JSA method for testing the mixed-inheritance model for quantitative traits when it was applied to the six basic generations: both parents (P_1 and P_2), F_1, F_2, and both backcross generations (B_1 and B_2) derived from crossing the F_1 to each parent. A total of 1968 genetic model-experiment scenarios were considered in the simulation study to quantify the power of the method. Factors that interacted to influence the power of the JSA method to correctly detect genetic models were: (1) whether there were one or two major genes in combination with polygenes, (2) the heritability of the major genes and polygenes, (3) the level of dispersion of the major genes and polygenes between the two parents, and (4) the number of individuals examined in each generation (population size). The greatest levels of power were observed for the genetic models defined with simple inheritance; e.g., the power was greater than 90% for the one major gene model, regardless of the population size and major-gene heritability. Lower levels of power were observed for the genetic models with complex inheritance (major genes and polygenes), low heritability, small population sizes and a large dispersion of favourable genes among the two parents; e.g., the power was less than 5% for the two major-gene model with a heritability value of 0.3 and population sizes of 100 individuals. The JSA methodology was then applied to a previously studied sorghum data-set to investigate the genetic control of the putative drought resistance-trait osmotic adjustment in three crosses. The previous study concluded that there were two major genes segregating for osmotic adjustment in the three crosses. Application of the JSA method resulted in a change in the proposed genetic model. The presence of the two major genes was confirmed with the addition of an unspecified number of polygenes.

Keywords Joint segregation analysis (JSA) · Mixed inheritance model · Osmotic adjustment · Power analysis · QU-GENE

Communicated by H.C. Becker

J. Wang
Wheat Program, CIMMYT, Adpo. Postal 6-641,
06600 Mexico, D.F., Mexico

D.W. Podlich (✉) · M. Cooper · I.H. DeLacy
School of Land and Food Sciences,
The University of Queensland, Brisbane,
Queensland 4072, Australia
e-mail: d.podlich@mailbox.uq.edu.au
Fax: +61 7-3365-2036

Present address:
M. Cooper, Pioneer Hi-Bred International Incorporated,
7300 N.W. 62nd Avenue, PO Box 1004, Johnston,
Iowa 50131, USA

Introduction

The genetic architecture of the traits possessed by organisms has been the subject of investigation for over a century (Kearsey and Farquhar 1998). For the majority of this period investigations relied predominantly on the use of quantitative methodologies for partitioning phenotypic variation in terms of hypothesised genetic models. Advances over the last two decades have expanded our capacity to measure and characterise genetic variation at the DNA level. This has enabled these investigations of trait architecture to be conducted at both genetic and phenotypic levels. However, determining the pathways from gene to phenotype remains a significant challenge for most quantitative traits. An important component of the analysis of gene to phenotype relationships will be the refinement of the quantitative genetic models that are

used to analyse phenotypic variation. These refinements will need to incorporate those properties of gene effects discovered at the genomic level that influence the phenotypes of individuals and the phenotypic variation observed in populations.

Many of the important traits manipulated in plant-breeding programs are quantitative in nature. Recent evidence from QTL (quantitative trait locus) mapping studies suggests that the genetic effects of QTLs can differ greatly from each other (Paterson 1998; Kearsey and Farquhar 1998). If we extend these observations to the gene level there may be one or a few genes contributing large effects to the phenotype, which are often referred to as major genes. As well as major genes, many minor genes, referred to as polygenes, can also influence traits. The distributions of the estimates of the QTLs detected in mapping studies are typically continuous, rather than discretely or obviously separated into clear groups of major and minor genes. Thus, it can be argued that, based on the results of the QTL mapping studies conducted to-date, any definition of what is a major and minor gene is arbitrary and likely to be context specific. This ambiguity is likely to persist as long as we have an incomplete understanding of the gene to phenotype relationships and the biochemical pathways and physiological processes influenced by the genes. Therefore, here we consider the terms major and minor genes in a relative sense and restrict our considerations to genetic models where a major gene is a gene with a larger relative effect on the genotype and phenotype of an individual within a specific cross or genetic background. This definition is at least consistent with the genetic models we consider in this paper. For situations where major genes are identifiable within a given germplasm pool it is expected that the major genes can be more easily manipulated in breeding and more easily identified and characterised by molecular genetic approaches than polygenes. Equally, the polygene effects can complicate the identification of major genes and hence make classical plant breeding and genetic engineering processes more difficult.

The mixed major-gene and polygene inheritance model was first studied in human genetics and animal breeding (Elston and Steward 1973; Elston 1984; Famula 1986; Hoeschele 1988; Knott et al. 1991; Janss et al. 1995). Recently, Wang (1996) and Gai and Wang (1998) applied the mixed inheritance model to the genetic study of plant quantitative traits. The four genetic models they considered were: the one major-gene inheritance model, the two major-gene inheritance model, the polygene inheritance model and the mixed one major-gene and polygene inheritance model. This has been extended to include the mixed two major-gene and polygene inheritance model (Gai et al. 2001). The method for analyzing and testing the mixed-inheritance models is called joint segregation analysis (JSA), which is a maximum-likelihood approach using mixture distribution models (McLachlan and Basford 1988). The JSA method has been applied to study the inheritance of maturity, tofu quality, cyst nematode resistance and foliar feeding insect resistance in soybean, bacterial blight and wide compatibility in rice, maturity in rapeseed and dwarf mosaic virus resistance in maize (Gai and Wang 1998; Wang et al. 2000; Wang and Gai 2001).

One of the most-commonly asked questions for the genetic analysis of a quantitative trait is “how large should the experimental population size be?” Answering this question requires assessment of the power of the methodology for a range of relevant genetic models (Beavis 1998). To-date there has been no investigation of the power of the JSA methodology for testing mixed-inheritance analysis models. The objective of this study was to use computer simulation to conduct a power analysis of the JSA method for a range of mixed-inheritance models of interest in plant genetics. The properties of the genetic models considered were the number of major genes and polygenes, the heritability values for both the major genes and polygenes, the degree of gene dispersion between the parents, and the population size of the generations included in the experiment. From this study, recommendations on a suitable population size for testing specific genetic models can be given from the power analysis. This information is used to re-interpret a previously published genetic analysis of the inheritance of osmotic adjustment in three grain sorghum crosses (Basnayake et al. 1995).

Materials and methods

The power of the joint segregation analysis (JSA) method was investigated using computer-simulation methodology. A simulation experiment was conducted using the QU-GENE (QUantitative-GENEtics) simulation software (Podlich and Cooper 1998). QU-GENE was used to define different types of genetic models and generate populations of individuals for appropriate generations. For the purposes of this study, the power of the JSA method was considered for the six basic generations, which included the two homogeneous parents (P_1 and P_2), the F_1 cross between the parents, the F_2 generation produced by self-pollination of the F_1, and the two backcross generations (B_1 and B_2) developed by crossing the F_1 to each of the parents. Using the phenotypic information from these generations, the JSA was used to predict the underlying genetic model. The power of the JSA method was evaluated by assessing its ability to select the correct underlying genetic model. The background to the JSA method and the QU-GENE simulation software is given, followed by a detailed description of the parameters used in the simulation experiment.

Joint segregation analysis (JSA) and JSA software

The principle of the JSA is described as follows. Firstly, it is assumed that trait variation in each segregating population is due to the variation in the distribution of major genes modified by polygenes and the environment (see Appendix 1 for more detail). Secondly, the major gene heritability and polygene heritability are defined, and standard curves of the mixture distributions under various genetic conditions are drawn. Thus, comparing practical frequency distributions with the standard curves we can approximate the inheritance model of a quantitative trait. This process is called graphical analysis (Wang 1996; Wang and Gai 2001). Thirdly, likelihood functions under various possible genetic models are established, maximum-likelihood estimates of parameters contained

Table 1 Genetic structure of models for the power analysis. See Appendix 2 for examples on how the genetic models were constructed. h^2: heritability; h^2_{mg}: major-gene heritability; h^2_{mg1}: first major-gene heritability; h^2_{mg2}: second major-gene heritability; h^2_{pg}: polygene heritability; a_1: additive effect of the first major gene; a_2: additive effect of the second major gene; a_i^{10}: additive effect of individual polygenes for genetic models with ten polygenes; a_i^{30}: additive effect of individual polygenes for genetic models with 30 polygenes

Set	Class of genetic model	h^2	$h_{mg}{}^2$	h^2_{mg1}	h^2_{mg2}	h^2_{pg}	a_1	a_2	a_i^{10}	a_i^{30}
1	1MG	0.9	0.9	0.9	0.0	0.0	1.10	0.00	0.00	0.00
2	2MG	0.9	0.9	0.8	0.1	0.0	1.03	0.37	0.00	0.00
3	2MG	0.9	0.9	0.7	0.2	0.0	0.97	0.52	0.00	0.00
4	2MG	0.9	0.9	0.6	0.3	0.0	0.89	0.63	0.00	0.00
5	2MG	0.9	0.9	0.5	0.4	0.0	0.82	0.73	0.00	0.00
6	MX1	0.9	0.7	0.7	0.0	0.2	0.97	0.00	0.20	0.12
7	MX2	0.9	0.7	0.6	0.1	0.2	0.89	0.37	0.20	0.12
8	MX2	0.9	0.7	0.5	0.2	0.2	0.82	0.52	0.20	0.12
9	MX2	0.9	0.7	0.4	0.3	0.2	0.73	0.63	0.20	0.12
10	MX1	0.9	0.5	0.5	0.0	0.4	0.82	0.00	0.28	0.16
11	MX2	0.9	0.5	0.4	0.1	0.4	0.73	0.37	0.28	0.16
12	MX2	0.9	0.5	0.3	0.2	0.4	0.63	0.52	0.28	0.16
13	MX1	0.9	0.3	0.3	0.0	0.6	0.63	0.00	0.35	0.20
14	MX2	0.9	0.3	0.2	0.1	0.6	0.52	0.37	0.35	0.20
15	MX1	0.9	0.1	0.1	0.0	0.8	0.37	0.00	0.40	0.23
16	PG	0.9	0.0	0.0	0.0	0.9	0.00	0.00	0.42	0.25
17	1MG	0.7	0.7	0.7	0.0	0.0	0.97	0.00	0.00	0.00
18	2MG	0.7	0.7	0.6	0.1	0.0	0.89	0.37	0.00	0.00
19	2MG	0.7	0.7	0.5	0.2	0.0	0.82	0.52	0.00	0.00
20	2MG	0.7	0.7	0.4	0.3	0.0	0.73	0.63	0.00	0.00
21	MX1	0.7	0.5	0.5	0.0	0.2	0.82	0.00	0.20	0.12
22	MX2	0.7	0.5	0.4	0.1	0.2	0.73	0.37	0.20	0.12
23	MX2	0.7	0.5	0.3	0.2	0.2	0.63	0.52	0.20	0.12
24	MX1	0.7	0.3	0.3	0.0	0.4	0.63	0.00	0.28	0.16
25	MX2	0.7	0.3	0.2	0.1	0.4	0.52	0.37	0.28	0.16
26	MX1	0.7	0.1	0.1	0.0	0.6	0.37	0.00	0.35	0.20
27	PG	0.7	0.0	0.0	0.0	0.7	0.00	0.00	0.37	0.22
28	1MG	0.5	0.5	0.5	0.0	0.0	0.82	0.00	0.00	0.00
29	2MG	0.5	0.5	0.4	0.1	0.0	0.73	0.37	0.00	0.00
30	2MG	0.5	0.5	0.3	0.2	0.0	0.63	0.52	0.00	0.00
31	MX1	0.5	0.3	0.3	0.0	0.2	0.63	0.00	0.20	0.12
32	MX2	0.5	0.3	0.2	0.1	0.2	0.52	0.37	0.20	0.12
33	MX1	0.5	0.1	0.1	0.0	0.4	0.37	0.00	0.28	0.16
34	PG	0.5	0.0	0.0	0.0	0.5	0.0	0.00	0.32	0.18
35	1MG	0.3	0.3	0.3	0.0	0.0	0.63	0.00	0.00	0.00
36	2MG	0.3	0.3	0.2	0.1	0.0	0.52	0.37	0.00	0.00
37	MX1	0.3	0.1	0.1	0.0	0.2	0.37	0.00	0.20	0.12
38	PG	0.3	0.0	0.0	0.0	0.3	0.0	0.00	0.25	0.14
39	1MG	0.1	0.1	0.1	0.0	0.0	0.37	0.00	0.00	0.00
40	PG	0.1	0.0	0.0	0.0	0.1	0.0	0.00	0.14	0.08
41	NULL	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.00	0.00

in each model are calculated through the EM algorithm (Dempster et al. 1977; McLachlan and Basford 1988), and the best-fitting genetic model and its parameter estimates are chosen by Akaike's Information Criterion (*AIC*) (Akaike 1977), likelihood-ratio test and tests of goodness of fit. From Akaike's Information Criterion, the model with the least *AIC* value is chosen as the best-fitting model. Here, $AIC=-2L_c(\Phi)+2N$, where $L_c(\Phi)$ is the maximized log-likelihood and N is the number of independent parameters in the model. Finally, each individual in the segregating generations is classified into a suitable component distribution using Bayesian posterior probabilities.

Software implementing the JSA method was used in the simulation experiment. The JSA software calculated the maximum-likelihood estimates for a set of parameters (see Appendix 1 for an example) and the *AIC* value for each genetic model. The model with the lowest *AIC* value was selected as the best-fitting genetic model for a given set of phenotypic data.

QU-GENE simulation software

To conduct a power analysis of the JSA method it was necessary to generate phenotypic information of individuals in segregating populations. This was achieved through the use of the GEXP (Genetic EXPeriments) module available within the QU-GENE simulation platform. The QU-GENE software comprises a two-stage architecture (Podlich and Cooper 1998). The first stage is the engine, which is used to specify the properties of the genetic models under investigation. The properties of the genetic models that can be manipulated in the engine include the number of genes influencing a trait, the action and magnitude of genetic effects of individual genes influencing a trait, and the heritability of a trait. The second stage of the QU-GENE software consists of a series of modules that are used to conduct the simulation experiments, using as inputs the properties of the genetic models from the engine. The GEXP module conducts a range of genetic experiments that can be used to study the inheritance of traits. For the current study, the generation means analysis (GMA) option within GEXP was used to create the six basic generations, i.e., both parents (P_1 and P_2), F_1, backcrosses (B_1 and B_2) and F_2. For the purposes of this study it was assumed that the parents were completely inbred.

Genetic models considered

A total of 82 genetic models were considered for the power analysis (Table 1). The 82 genetic models were separated into six classes of inheritance. These were: (1) the null genetic model (NULL), (2) the one major-gene inheritance model (1MG), (3) the two major-gene inheritance model (2MG), (4) the polygene inheritance

model with no major genes (PG), (5) the mixed one-major gene and polygene inheritance model (MX1), and (6) the mixed two-major gene and polygene inheritance model (MX2). For each class of genetic model, a number of heritability values were considered. The heritability values were separated into major and polygene heritabilities (discussed later), resulting in a total of 41 model-heritability combinations (referred to as Sets in Table 1). For each Set, models with either 10 or 30 polygenes were considered. This resulted in the 82 genetic models (41 Sets×2 polygene numbers) considered for the power analysis (Table 1). Each of the genetic models was defined in the engine of the QU-GENE software (see Appendix 2 for examples on how the genetic models were constructed).

For each of the 82 genetic models, six scenarios with different dispersions of genes (GD) between the two parents were considered; GD1: P_1 has the two favourable major genes and all the favourable polygenes; GD2: P_1 has the two favourable major genes and 50% favourable polygenes; GD3: P_1 has the two favourable major genes but no favourable polygenes. For GD4, GD5 and GD6, P_1 has the first favourable major gene and different numbers of the favourable polygenes to be the same as those in GD1, GD2 and GD3, respectively. For each genetic model, it was assumed that all genes were unlinked and there were no epistatic effects. The two major genes were defined to have complete dominance, and the first major gene had a larger effect than the second. All polygenes were defined to be additive, and contributed equally to the phenotypic variation.

Partitioning of heritability values into major and polygene heritabilities

For a mixed-inheritance model, the phenotypic value (p) can be expressed in terms of a linear model as the summation of the population mean (m), the major gene effect (g), the polygene effect (c) and the environmental effect (e), i.e., $p=m+g+c+e$ (Morton and MacLean 1974), where g is different for different major-gene genotypes, and c and e are normally distributed variables. The phenotypic variance (σ_p^2) can be decomposed to consist of the major-gene variance (σ_{mg}^2), the polygene variance (σ_{pg}^2) and the environmental variance (σ_e^2). Therefore, we can define major-gene heritability (h_{mg}^2) and polygene heritability (h_{pg}^2) as $h_{mg}^2=\sigma_{mg}^2/\sigma_p^2$ and $h_{pg}^2=\sigma_{pg}^2/\sigma_p^2$, respectively (Wang and Gai 2001). Because there are two independent major genes, it is possible to further separate the major-gene heritability as the first major-gene heritability (h_{mg1}^2) and the second major-gene heritability (h_{mg2}^2). Here, the heritability is defined in the broad sense, i.e., the ratio of the genetic variation on the phenotypic variation. For this study, the reference population for the definition of heritability is taken to be the F_2 generation.

When the first major-gene heritability, the second major-gene heritability and the polygene heritability are all larger than zero, the genetic model will be the mixed two major-gene and polygene inheritance model (MX2). If the second major gene has a heritability of zero, the model will be the mixed one major-gene and polygene inheritance model (MX1). If both the major-gene heritabilities are zero, the model will be the polygene inheritance model with no major genes (PG). If the polygene heritability is zero, the model becomes the two major-gene inheritance model (2MG). If the polygene heritability and the second (or the first) major-gene heritability are zero, the model becomes the one major-gene inheritance model (1MG). Finally, if all three heritabilities are zero, the model will be the null genetic model (NULL), i.e., there is no gene(s) contributing to the phenotype.

The generations considered in this study were the six basic generations commonly used in quantitative genetics, i.e., two homozygous parents (P_1 and P_2), F_1, backcrosses (B_1 and B_2), and F_2. Without losing generality, we suppose $\sigma_p^2=1$ in the F_2 generation. Hence, the major-gene and polygene heritabilities in the F_2 generation can be represented as:

$$h_{mg1}^2 = \frac{1}{2}a_1^2 + \frac{1}{4}d_1^2,\ h_{mg2}^2 = \frac{1}{2}a_2^2 + \frac{1}{4}d_2^2 \text{ and } h_{pg}^2 = \frac{1}{2}V_A^* + \frac{1}{4}V_D^*,$$

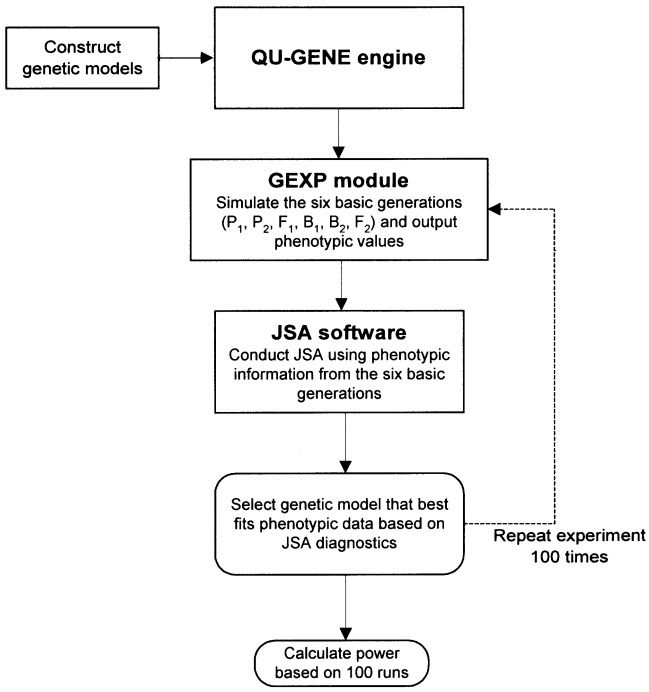

Fig. 1 Flow chart of the procedures used to conduct the simulation experiment and to estimate power

where a_1, d_1, a_2, and d_2 are the additive and dominance effects of the two major genes, and V_A^* and V_D^* are the sums of the additive and dominance variances of the individual polygenes (Kearsey and Pooni 1996). By applying the above specifications, we have

$$a_1 = d_1,\ a_2 = d_2,\ V_A^* = \sum_i a_i^2, \text{ and } V_D^* = \sum_i d_i^2 = 0,$$

where a_i is the additive effect of the individual polygene which has the same value for all polygenes in the simulation study, and d_i is the dominant effect of the individual polygene which is zero for all polygenes in the simulation study. Using the above equations, the 82 genetic models summarised in Table 1 were constructed as input files for use in the QU-GENE engine (see Appendix 2 for examples).

Experimental factor: population size

In combination with the genetic models listed in Table 1, the power of the JSA method for different sizes of populations was examined. Four population sizes (PS) were considered: 30, 50, 100 and 200 individuals. To avoid the number of population-size combinations getting too large, we considered the case where each generation has the same population size, even though the population sizes of the non-segregating generations P_1, P_2 and F_1 would not have much effect on the power.

Procedure for the power-analysis experiment

The procedure for conducting the power-analysis experiment was as follows (Fig. 1).

(1) Create 82 QU-GENE input files for each genetic model summarised in Table 1.
(2) Run the QU-GENE engine to generate 82 output files each representing a reference population of genotypes for one of the genetic models summarised in Table 1. The QU-GENE engine output files provide the input for running the GEXP module.

(3) Run the GEXP module for parents specified to represent the six gene dispersion (GD) scenarios to produce the six basic generations (P_1, P_2, F_1, F_2, B_1 and B_2) with four different population sizes. The phenotypic values for the individuals generated in each of the generations are outputs from the GEXP module and represent the inputs for the JSA.
(4) Run the JSA software for each of the sets of six generations produced by GEXP and for each set select the genetic model that best fits the simulated phenotypic data based on the diagnostics available from the JSA software.
(5) Run each case (82 models×4 population sizes×6 gene dispersion scenarios=1968 total model-experiment combinations) 100 times and count the number of times that each genetic model was selected. The number of times that the correct model was selected out of the 100 runs was taken as a measure of the power of the JSA method for that model-experiment combination.

Example data set: osmotic adjustment

Osmotic adjustment has been identified as a physiological mechanism that contributes to improved adaptation to water stress (Morgan 1980; Ludlow and Muchow 1990). Osmotic adjustment data obtained from a controlled environment experiment were available for the six basic generations from three sorghum crosses (Basnayake et al. 1995). Ten individuals were measured for each of the P_1, P_2 and F_1 generations, and 30 individuals were measured for the F_2, B_1 and B_2 generations for each cross. The crosses were based on three inbred lines selected for high (Tx2813 and TAM422) and low (QL27) levels of osmotic adjustment (Basnayake et al. 1993, 1995). Basnayake et al. (1993, 1995) gave a description of the experimental design, the methodology for screening the plants in the controlled environment facility and techniques for measuring osmotic adjustment. Previous analysis of these data by mixture-model clustering methodology (McLachlan and Basford 1988), which was applied separately to the three segregating generations for each cross, resulted in identification of an inheritance model based on two major genes for high osmotic adjustment. The JSA was applied to these data to evaluate the suitability of the inheritance model proposed by Basnayake et al. (1995) when information from the six generations was used simultaneously.

Results and discussion

Power: the NULL and 1MG models

For the NULL and 1MG models the effects of the polygenes were all zero (Table 1). Consequently, for these cases the number of polygenes (GN) and their dispersion (GD) between the parents did not affect the power of the JSA to detect the 1MG and NULL models. Therefore, the results of the six GD scenarios and the two GN scenarios were combined to estimate the power for these two models. The power of the JSA to correctly select the NULL model was always higher than 80%, and increased with population size (Fig. 2a). The JSA consistently had a power higher than 90% for the simplest genetic model (1MG), regardless of the major-gene heritability and the population size (Fig. 2a).

Power: the 2MG model

The JSA had a high power to correctly identify the 2MG model for scenarios with a total major-gene heritability of 0.9 and 0.7, when the major genes were not dispersed between the parents (GD1, GD2 and GD3) (Fig. 2b). The power exceeded 90% for the total major-gene heritability of 0.9 and it exceeded 60% when the heritability decreased to 0.7. The power increased with population size (Fig. 2b). The power was higher if the two major genes made a similar contribution to the phenotype (e.g., cases 6_3 and 5_4, (Fig. 2b). For the total major-gene heritability of 0.5 (e.g., cases 4_1 and 3_2, Fig. 2b), the power when the population size was 30 and 50 was lower than 40%; however, for a population size of 100 and 200 the power exceeded 60%. For the low heritability of 0.3 (e.g., case 2_1, Fig. 2b), the power to detect the 2MG model was extremely low and did not exceed 40%, even with a population size of 200. The distribution of the two major genes between the parents had a large effect on power (Fig. 2c). For example, for the major-gene heritabilities 0.8 and 0.1 (i.e., case 8_1), the power to detect the 2MG model was greater than 90% (94%, 95%, 99% and 100% for the four population sizes 30, 50, 100 and 200, respectively) for the case where the major genes were not dispersed between the parents (Fig. 2b), but increased in the order 33%, 48%, 76% and 95% for the dispersed case as population size increased from 30 to 200 (Fig. 2c). Therefore, if the major genes are dispersed between the parents, a large population size is required to identify the correct genetic model.

Power: the PG model

There was no effect of the major genes for the PG models considered (Table 1). Therefore, as expected, the gene dispersion of the major genes did not have any effect on power. Fig. 2d–f show the power for three polygene dispersion scenarios for models with 30 polygenes (GN30). For gene dispersion scenarios GD1, GD3, GD4 and GD6, there was no dispersion of the polygenes; hence, one of the parents possessed all of the favourable polygenes. In these cases the power was greater than 50% (Fig. 2d and f). The power to correctly detect the PG model increased with population size and with decreasing heritability. At low heritability the distributions of the phenotypic values for the populations were more continuous and therefore more consistent with the PG model, making correct selection of the PG model easier. Where small population sizes were combined with high heritability, the distributions of phenotypic values were non-normal in some cases making selection of the PG model more difficult. For scenarios where the polygenes were mostly dispersed (scenarios GD2 and GD5), i.e., each parent had 50% of the favourable polygenes, the power to correctly detect the PG model decreased greatly, especially for small population sizes and low heritability (Fig. 2e). For this case, the 1MG or 2MG models were frequently selected for the case of high polygene heritability. However, the NULL model was selected frequently for the case of low polygene heritability. The number of polygenes (GN10 or GN30) did not have much effect on the power to detect the PG model.

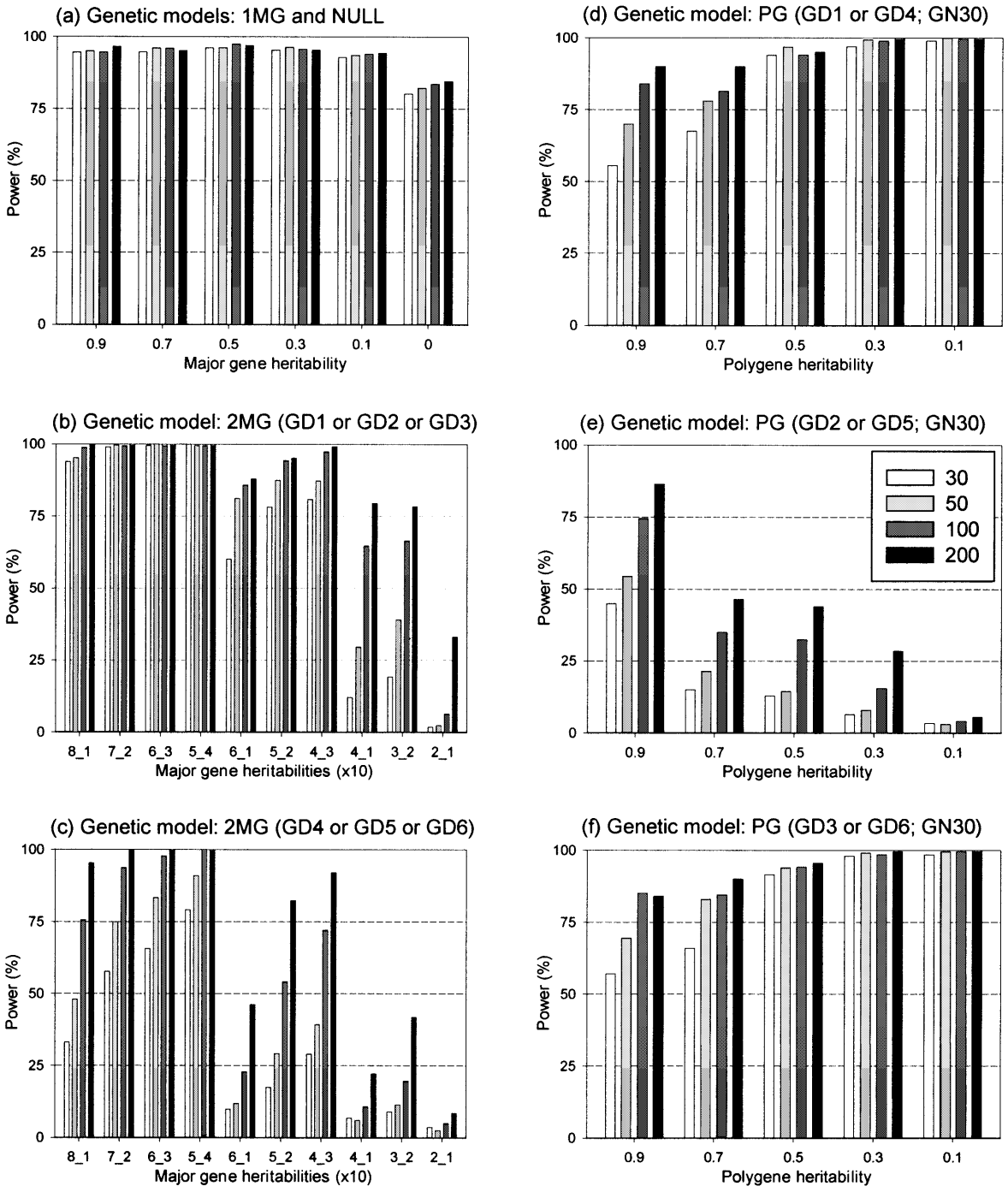

Fig. 2a–f Power (measured as a %) of the joint segregation analysis to correctly identify the NULL, 1MG, 2MG and PG models for a range of major-gene (h^2_{mg}) and polygene (h^2_{pg}) heritability values, and experiments based on four population sizes (PS) of 30, 50, 100 and 200 individuals per generation: (**a**) 1MG (h^2_{mg}= 0.9 to 0.1) and NULL (h^2_{mg}=0) models; (**b**) 2MG model for gene-dispersion scenarios GD1, GD2 or GD3; (**c**) 2MG model for gene-dispersion scenarios GD4, GD5 or GD6; (**d**) PG model for gene-dispersion scenarios GD1 or GD4 and a polygene number of 30 (GN30); (**e**) PG model for gene-dispersion scenarios GD2 or GD5 and a polygene number of 30 (GN30); (**f**) PG model for gene-dispersion scenarios GD3 or GD6 and a polygene number of 30 (GN30). Labels for the heritability values (horizontal axis) for the 2MG models (**b**, **c**) are multiplied by ten

Power: the MX1 model

The change in polygene number (GN) from 10 to 30 did not have a large effect on the power of the JSA to detect the MX1 model (Fig. 3). The power exceeded 50% for the major-gene heritability and polygene-heritability combinations 0.7_0.2, 0.5_0.4 and 0.5_0.2 (cases 7_2, 5_4 and 5_2, respectively) for all the population sizes considered, when the major gene and polygenes were not dispersed between the parents (GD1 and GD4) (Fig. 3a and d). When the major-gene heritability decreased, such as in the combinations 0.3_0.6 (case 3_6) (Set 13,

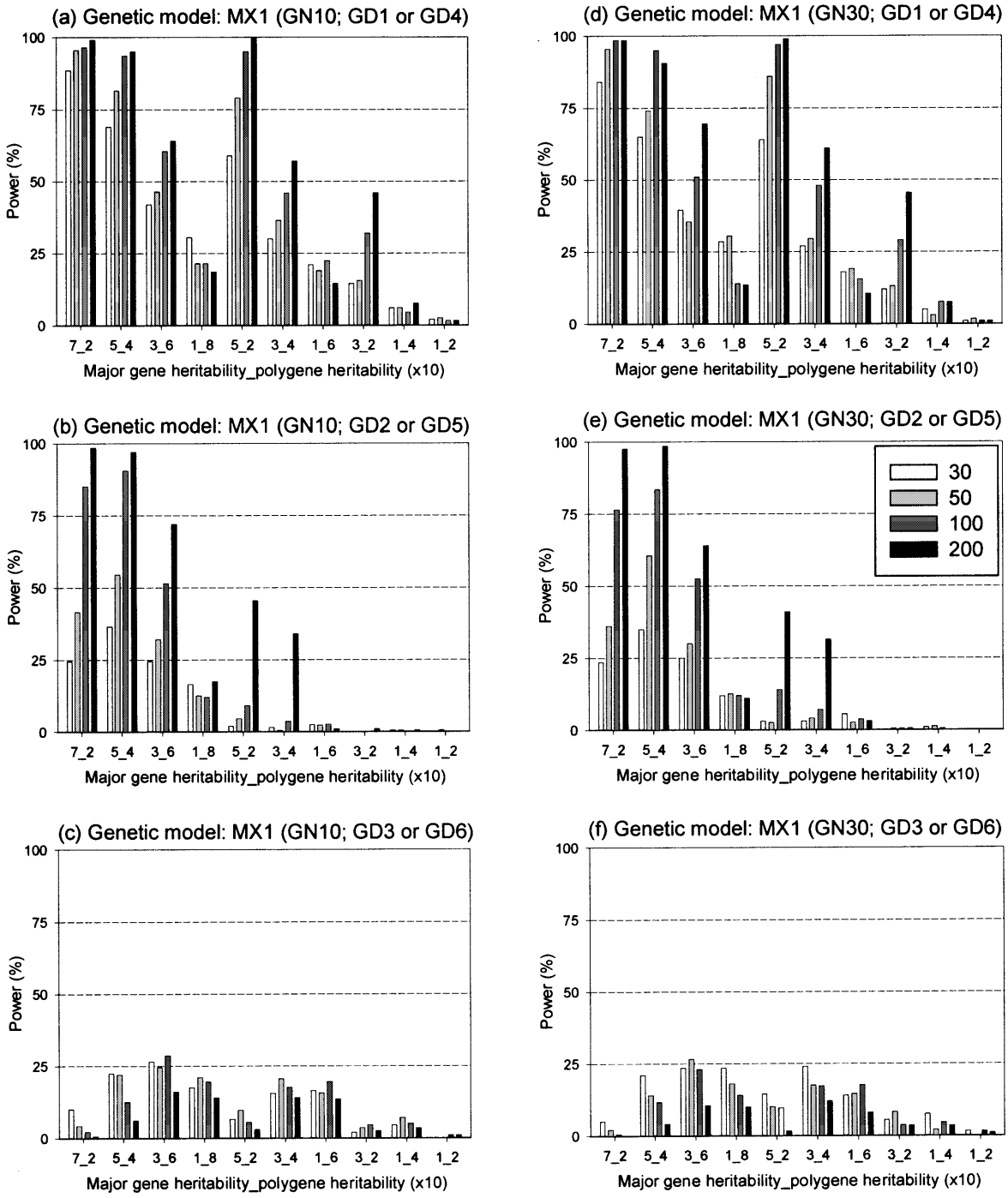

Fig. 3a–f Power (measured as a %) of the joint segregation analysis to correctly identify the MX1 genetic models for a range of major-gene (h^2_{mg}) and polygene (h^2_{pg}) heritability values, different gene-dispersion (GD) scenarios, different polygene numbers (GN10 or GN30) and experiments based on four population sizes (PS) of 30, 50, 100 and 200 individuals per generation: (**a**) polygene number GN10 and gene-dispersion scenarios GD1 or GD4; (**b**) polygene number GN10 and gene-dispersion scenarios GD2 or GD5; (**c**) polygene number GN10 and gene-dispersion scenarios GD3 or GD6; (**d**) Polygene number GN30 and gene-dispersion scenarios GD1 or GD4; (**e**) polygene number GN30 and gene-dispersion scenarios GD2 or GD5; (**f**) polygene number GN30 and gene-dispersion scenarios GD3 or GD6. Labels for the heritability values (horizontal axis) are multiplied by ten

Table 1), 0.1_0.8 (case 1_8) (Set 15, Table 1), 0.1_0.6 (case 1_6) (Set 26, Table 1) and 0.1_0.4 (case 1_4) (Set 33, Table 1), the effect of the individual polygenes is near, or can exceed, that of the major gene. In these cases it is difficult to identify the existence of a major gene. Therefore, the PG model was mostly selected and the power to select the specified MX1 model was very low (Fig. 3a and d). The power to detect the MX1 model did not exceed 50% for combined major-gene and polygene heritability values of 0.5 and lower.

If the polygenes were dispersed and half of the polygenes were in P_1 and the others in P_2 (GD2 and GD5),

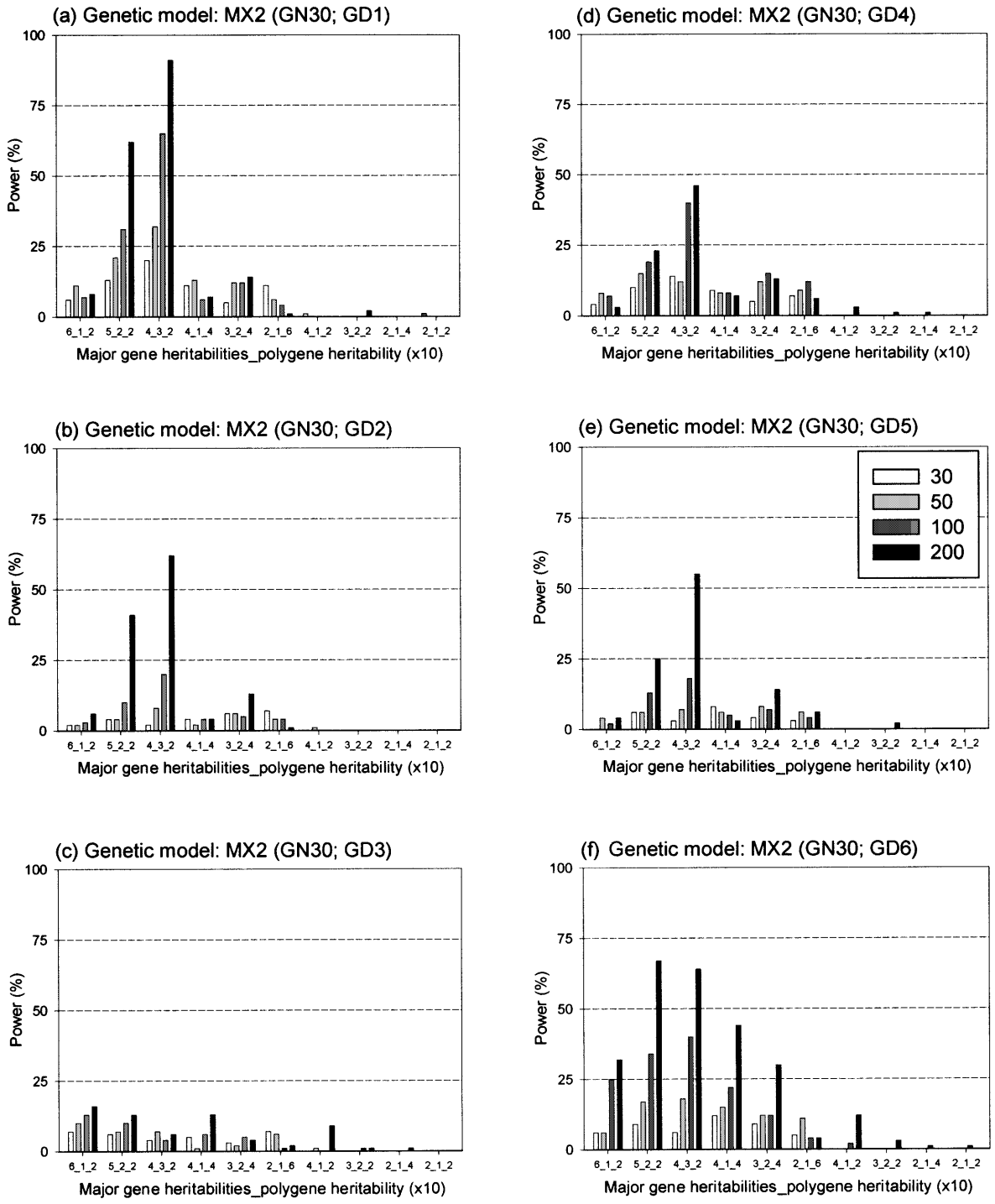

Fig. 4a–f Power (measured as a %) of the joint segregation analysis to correctly identify the MX2 genetic models for a range of major-gene (h^2_{mg}) and polygene (h^2_{pg}) heritability values, different gene-dispersion (GD) scenarios, a polygene number of 30 (GN30) and experiments based on four population sizes (PS) of 30, 50, 100 and 200 individuals per generation: (**a**) Gene-dispersion scenario GD1; (**b**) gene-dispersion scenario GD2; (**c**) Gene-dispersion scenario GD3; (**d**) gene-dispersion scenario GD4; (**e**) gene-dispersion scenario GD5; (**f**) gene-dispersion scenario GD6. Labels for the heritability values (horizontal axis) are multiplied by ten

only for the cases where the population size was 100 or more, and the major-gene heritability was more than 0.3 and the total heritability was 0.9, was the power to detect the MX1 model greater than 50% (Fig. 3b and e). For the case where P_1 had the favourable major gene but none of the favourable polygenes (GD3 and GD6), the power to correctly detect the MX1 model was consistently low (Fig. 3c and f). This suggests that the JSA had difficulty in finding the correct model. Hence, in genetic analysis for a quantitative trait using the JSA method, we should avoid choosing such inbred lines as parents. This may be

of particular significance in studies that aim to identify useful QTLs for quantitative traits from wild relatives by introgressing genomic regions into adapted backgrounds.

Power: the MX2 model

The power of the JSA to detect the MX2 model was the lowest among the six classes of genetic model defined in Table 1 (Fig. 4). There were only a few cases where the power was more than 50%. Among those cases with higher power were the heritability combination 0.5_0.2_0.2 (case 5_2_2) for a population size of 200 and heritability combination 0.4_0.3_0.2 (case 4_3_2) for population sizes of 100 and 200. For gene-dispersion scenario GD1, where the major genes and polygenes are not dispersed between the parents, the MX1 model was always selected in place of the MX2 model when the total heritability was more than 0.7, and the PG model was always selected when the heritability was less than 0.7. Consequently, for these scenarios the power to detect the MX2 model was low (Fig. 4a).

For gene-dispersion scenario GD4, where the major genes were dispersed and the polygenes were not dispersed between the parents, there was no case where the power was more than 50%. The MX1 model was mostly selected when the total heritability was more than 0.7, and the PG model was mostly selected when the heritability was lower (Fig. 4d). For the GD2 (Fig. 4b) and GD5 (Fig. 4e) scenarios, any of the models could be selected by the JSA and the correct MX2 model was only rarely selected. Therefore, it would be difficult for the correct model to be selected in practice when the genes controlling a quantitative trait are mostly dispersed. For the GD3 scenario (Fig. 4c), where P_1 has the two favourable major genes but has none of the favourable polygenes, the PG model was mostly selected. For the GD6 scenario (Fig. 4e), where P_1 has the first favourable major gene but has none of the favourable polygenes, the PG model was mostly selected for a heritability lower than 0.7. The 1MG models or the 2MG models were selected for higher heritabilities.

Estimation of the genetic parameters

The genetic parameters were accurately estimated for the 1MG, 2MG and PG models. For the MX1 model, the distribution of the polygenes had a great effect on the genetic parameter estimation, as it had for the power to detect the model. For gene-dispersion scenarios GD1, GD2, GD4 and GD5, the effect of the major gene was reasonably well estimated (data not shown). However, for the scenarios GD3 and GD6 the major-gene effect was not reliably estimated, even for a high major-gene heritability. The situation for the MX2 model was similar to that for the MX1 model. For model Sets 2, 3, 4 and 5 (Table 1), where the real model was 2MG and the major-gene heritability was 0.9, the power to select the correct model was nearly 100% (Table 2). For these cases the estimated additive effect of the first major gene (a_1) was close to its true value. In contrast, the additive effect of the second major gene (a_2) was always estimated to be lower than its true value (Table 2). For model Sets 18, 19 and 20, where the major-gene heritability was 0.7 and the power was greater than 90%, the additive effect of the first major-gene was always overestimated and the effect of the second was underestimated. For those cases where the MX1 model was selected instead of the 2MG model, the first major-gene effects were accurately estimated; however, the effect of the second major gene was partitioned into the polygene component of the model. The situation was similar for other major-gene heritabilities (data not shown).

For model Sets 7, 8 and 9, where the major-gene heritability was 0.7 and the polygene heritability was 0.2 (Table 1), the additive effects of the first major gene tended to be overestimated and the dominance effects were underestimated (Table 2). Further, the estimates of the second major-gene effects had a larger variance than those of the first gene. For the cases where the MX1 model was selected, only the effects of the first major gene were estimated, and the estimates of the effects were higher than the true value (Table 2). The situation was similar for the cases with higher polygene heritability (data not shown).

Example data set: osmotic adjustment

The JSA was applied to examine the inheritance of osmotic adjustment to water stress in three grain sorghum crosses. Previous results, based on testing Mendelian models using the mixture-model method of clustering applied separately to each segregating generation (Basnayake et al. 1995), indicated one recessive major gene (*oa1* for high osmotic adjustment) in cross C1 (Tx2813/QL27), one additive major gene (*OA2* for high osmotic adjustment) in cross C2 (QL27/TAM422), with both the major genes segregating in cross C3 (Tx2813/TAM422).

Based on the *AIC* values from the JSA for the six classes of genetic model given in Table 1, the models with the least *AIC* were the MX1 model for C1, the MX2 model for C2 and the MX1 model for C3 (Table 3). The *AIC* for these models were considerably lower than those for the models proposed by Basnayake et al. (1995) for crosses C1 and C2 and slightly lower for cross C3. This result indicates that the mixed one major-gene and polygene inheritance model was the most appropriate model for the C1 and C3 crosses, and the mixed two major-gene and polygene inheritance model was the most appropriate model for the C2 cross. Thus, the inclusion of polygene effects in combination with the major genes was a significant change from the previous inheritance models for all three crosses. The dominance effect for the major gene in C1 was negative, implying a recessive major gene for high osmotic adjustment. This result was

Table 2 Estimates of additive (*a*) and dominance (*d*) gene effects for the first and second major genes for the 2MG and MX2 models under gene-dispersion scenarios GD1, GN30 and a population size of 100 from a range of model-experiment combinations (Set). The true genetic-model effects are identified in bold followed in normal case by the estimates from the Joint Segregation Analysis for the correct model. For some Sets, estimates of genetic effects for alternative genetic models are highlighted in italics. The power of alternative models is displayed in parentheses in the Model column

Set[a]	Model[b]	Power[c] (%)	a_1[d]	d_1[d]	a_2[e]	d_2[e]	[*a*][f]	[*d*][f]
2	**2MG**		**1.03**	**1.03**	**0.37**	**0.37**		
	2MG	99	1.11 (0.09)	0.94 (0.11)	0.29 (0.08)	0.35 (0.11)		
3	**2MG**		**0.97**	**0.97**	**0.52**	**0.52**		
	2MG	100	1.09 (0.10)	0.80 (0.12)	0.38 (0.11)	0.44 (0.13)		
4	**2MG**		**0.89**	**0.89**	**0.63**	**0.63**		
	2MG	100	1.09 (0.11)	0.67 (0.12)	0.44 (0.11)	0.50 (0.13)		
5	**2MG**		**0.82**	**0.82**	**0.73**	**0.73**		
	2MG	100	1.06 (0.09)	0.58 (0.12)	0.48 (0.10)	0.55 (0.13)		
18	**2MG**		**0.89**	**0.89**	**0.37**	**0.37**		
	2MG	91	0.94 (0.16)	0.79 (0.18)	0.31 (0.17)	0.39 (0.18)		
	MX1 (*9*)		*0.83* (*0.14*)	*0.96* (*0.14*)			*0.45* (*0.14*)	*0.32* (*0.16*)
19	**2MG**		**0.82**	**0.82**	**0.52**	**0.52**		
	2MG	95	0.90 (0.15)	0.63 (0.21)	0.42 (0.15)	0.52 (0.21)		
	MX1 (*3*)		*0.87* (*0.03*)	*0.89* (*0.20*)			*0.47* (*0.03*)	*0.44* (*0.18*)
20	**2MG**		**0.73**	**0.73**	**0.63**	**0.63**		
	2MG	95	0.89 (0.16)	0.57 (0.18)	0.46 (0.17)	0.58 (0.18)		
	MX1(*3*)		*0.85* (*0.03*)	*0.82* (*0.11*)			*0.50* (*0.03*)	*0.47* (*0.22*)
7	**MX2**		**0.89**	**0.89**	**0.37**	**0.37**	**3.46**	**0.00**
	MX2	7	0.94 (0.09)	0.42 (0.12)	0.54 (0.21)	0.15 (0.30)	1.63 (0.18)	1.23 (0.21)
	MX1 (*93*)		*0.90* (*0.16*)	*0.91* (*0.22*)			*3.83* (*0.17*)	*0.36* (*0.23*)
8	**MX2**		**0.82**	**0.82**	**0.52**	**0.52**	**3.46**	**0.00**
	MX2	31	0.89 (0.13)	0.44 (0.22)	0.62 (0.17)	0.33 (0.21)	1.68 (0.17)	1.06 (0.24)
	MX1 (*67*)		*0.85* (*0.16*)	*0.77* (*0.24*)			*3.95* (*0.16*)	*0.55* (*0.25*)
9	**MX2**		**0.73**	**0.73**	**0.63**	**0.63**	**3.46**	**0.00**
	MX2	65	0.86 (0.12)	0.38 (0.23)	0.60 (0.21)	0.30 (0.25)	1.75 (0.19)	1.18 (0.20)
	MX1 (*29*)		*0.86* (*0.18*)	*0.67* (*0.28*)			*3.96* (*0.18*)	*0.70* (*0.26*)

[a] Set: experiment set (see Table 1)
[b] Model: first model in each set indicates the actual model used for the QU-GENE simulation. The following models in each set indicate the models selected by the JSA
[c] Power indicates the percentage of times the model was selected by the JSA
[d] a_1, d_1: additive and dominance effects of the major gene in MX1 or the first major gene in 2MG or MX2. Values in brackets are standard deviations of the parameter estimates
[e] a_2, d_2: additive and dominance effects of the second major gene in 2MG or MX2
[f] [*a*], [*d*]: additive and dominance effect of the polygene system, (see Appendix 1 and 2)

consistent with the proposed major-gene effect for the *oa1* gene reported by Basnayake et al. (1995). The dominance effect of the first major gene in C2 was also negative, and suggested a partial recessive major gene. The additive effect of the major gene in C3 was negative and the dominance effect was small compared with the additive effect, implying an additive major gene, but the female parent did not possess the favourable allele. This result was different from that reported by Basnayake et al. (1995), who proposed a single additive major gene *OA2* possessed by TAM422 and segregating in C2, and two major genes segregating in C3.

Based on the results of the JSA we considered the possible major-gene genotypes for the three parents. If *A-a* and *B-b* are used to represent the alleles of the major genes in C1 and C3, where the upper-case allele has the effect of increasing osmotic adjustment, the proposed genotypes for Tx2813, TAM422 and QL27 were *AAbb*, *AABB* and *aabb*, respectively. The average phenotypic osmotic adjustment values for these genotypes were 1.58 MPa, 1.67 MPa and 0.77 MPa, respectively. The second major gene in C2 should be the same as the major gene segregating in C3, but its effect was not well estimated in the analysis of C2, where the two genes were segregating. This was consistent with the results from the power analysis, i.e., the effect of the second major gene in the MX2 model was always underestimated.

The results enable a revision of the osmotic-adjustment gene notation and the major-gene inheritance models proposed by Basnayake et al. (1995). Since there is a consistent trend for the high osmotic-adjustment alleles to behave in a recessive manner we propose that *oa1* and *oa2* represent the high osmotic-adjustment alleles for both major genes and that *OA1* and *OA2* represent the alleles for low levels of osmotic adjustment. Here we use the *oa1-OA1* allele combination in place of the *A-a* allele

Table 3 Selected genetic models (bold text) and previously reported genetic models given by Basnayake et al. (1995) (normal text) for osmotic adjustment in three grain sorghum crosses: C1=Tx2813/QL27, C2=QL27/TAM422, C3=Tx2813/TAM422. *AIC* and estimates of genetic effects for the genetic models

Cross	Model	*AIC*	a_1[a]	d_1[a]	a_2[b]	d_2[b]	*aa*[c]	*ad*[c]	*da*[c]	*dd*[c]	[*a*][d]	[*d*][d]
C1	**MX1-A-AD**	**–92.19**	**0.40**	**–0.13**							**–0.03**	**–0.24**
	1MG	–38.64	0.39	–0.20								
C2	MX2-ADI-AD	**–87.46**	**0.45**	**–0.19**	**–0.02**	**–0.11**	**0.08**	**–0.05**	**–0.18**	**0.18**	**2.78**	**0.29**
	1MG	–36.44	0.43	–0.11								
C3	**MX1-A-AD**	**–142.99**	**–0.34**	**0.02**							**0.31**	**–0.39**
	2MG	–131.05	–0.09	–0.03	0.04	–0.12	–0.03	0.08	0.05	–0.30		

[a] a_1, d_1: additive and dominance effects of the major gene for the MX1 model or the first major gene for the MX2 model
[b] a_2, d_2: additive and dominance effects of the second major gene for the MX2 model
[c] *aa*, *ad*, *da*, *dd*: additive×additive, additive×dominance, dominance×additive, dominance×dominance epistatic effects between the two major genes
[d] [*a*], [*d*]: additive and dominance effects of the polygene system, see Appendix 1 and 2

combination used above for the first major gene, and the *oa2-OA2* combination in place of the *B-b* allele combination used above for the second major gene. Therefore, the results of the JSA suggest that the osmotic-adjustment major-gene genotypes of the three parents should be designated as: Tx2813=*oa1oa1OA2OA2*; TAM422=*oa1oa1oa2oa2*; and QL27=*OA1OA1OA2OA2*.

Considering the effects of the polygenes, the estimates of the genetic parameters for C1 (Table 3) corresponded to the gene-dispersion scenario GD1 of Set 6, and those for C3 corresponded to scenario GD3 of Set 6. The estimates of power for GD1 and GD3 of Set 6 for a population size of 30 and a polygene number of 10 were 89% and 10%, respectively. Therefore, the results from C1 might be more reliable than those from C3, due to the possible influence of polygene dispersion in C3. The situation for C2 was consistent with scenario GD1 of Set 9, i.e., the two major genes are not dispersed between the parents and the major genes have similar effects (Table 3) and low polygene heritability. The power for this case for a population size of 30 and polygene number of 10 was 34%. Therefore, the power for the proposed inheritance models for C2 and C3 was low in both cases, which suggests the need for further experimentation with a larger population size to obtain a more reliable test of the inheritance models for osmotic adjustment in these three crosses.

Conclusions

The major results from the power analysis were: (1) The power for the one major-gene model (1MG) was over 90% regardless of the population size and the major-gene heritability. (2) The power for the two major gene model (2MG) was over 60% for total major-gene heritabilities of 0.9 and 0.7 when the major genes were not dispersed between the parents and was higher if the two major genes made a similar contribution. For a major-gene heritability of 0.5 and population sizes of 30 and 50, the power was lower than 40%; however, for population sizes of 100 and 200 the power exceeded 60%. For the lower heritability of 0.3, the power did not exceed 50% even for large population sizes. If major genes are dispersed, large population sizes will be required to reach the same power as the comparable scenario where the genes are not dispersed. (3) The number of polygenes (10 or 30) made little difference to the power of the test. (4) The power for the polygene model (PG) was larger than 50% in most cases if the polygenes were not dispersed between the parents. If the polygenes were highly dispersed, the power decreased greatly, especially for small population sizes and low levels of heritability. (5) For the mixed one major-gene and polygene (MX1) inheritance model, the power exceeded 50% for polygene heritability less than 0.4 when the major genes and polygenes were not dispersed between the parents. When the polygene heritability exceeded 0.4, the PG model was mostly selected in preference to the MX1 model and the power of the test was low. The power was greater than 50% only for the cases where the population size was 100 or more and the major-gene heritability was more than 0.3. For purposes of genetic analysis, the case where the first parent has the favourable major gene but has none of the favourable polygenes should be avoided, as the power to detect the correct model will always be low. (6) In general, the power for the mixed two major-genes and polygene inheritance model (MX2) was low. Either the MX1 or PG models were selected in preference to the MX2 model. (7) Genetic parameters for the 1MG, MX1 and PG models were accurately estimated. For the MX2 model, the genetic parameters were accurately estimated where heritability was high. For low heritability, the effect of the first major gene was over estimated, and the effect of the second major gene was underestimated.

In this paper we have concentrated on the common genetic experiment structure based on the six basic generations that include the two parents, the F_1 cross between the parents, the F_2 generation produced by self-pollination of the F_1, and the two backcross generations developed by crossing the F_1 to each of the parents. A

similar approach can also be applied to genetic experiments based on other combinations of generations. While a limited set of genetic models was considered in this paper, the JSA approach can also be applied to a wider range of more-complex genetic models. Application of the JSA methodology to a data set that was generated to investigate the inheritance of osmotic adjustment in three grain sorghum crosses, enabled a useful revision of the proposed inheritance models in the three crosses studied.

Appendix 1: parameters and assumptions used in the JSA method

Supposing the genetic data are from two homogeneous and homozygous parents (P_1 and P_2), the F_1 hybrid between the parents, the F_2 hybrid by self-pollination of the F_1, and the two backcross generations (B_1 and B_2), then some underlying assumptions are: (1) the variations within P_1, F_1 and P_2 are the same and can be used to measure the environmental variation σ_e^2, (2) polygene effects and environmental effects are all normally distributed, and (3) there is no linkage between the major gene and the polygene and no epistasis. Under these assumptions, sets of parameters can be defined for the different inheritance models. For example, with the mixed one major-gene and polygene inheritance model (MX1), P_1 (*AA*), F_1 (*Aa*) and P_2 (*aa*) are all normally distributed with different means but the same variance, and B_1, B_2 and F_2 are normal mixtures, which can be represented as:

$$P_1 \sim N(\mu_1, \sigma_e^2),\ F_1 \sim N(\mu_2, \sigma_e^2),\ P_2 \sim N(\mu_3, \sigma_e^2),$$

$$B_1 \sim (1/2)N(\mu_{41}, \sigma_4^2)+(1/2)N(\mu_{42}, \sigma_4^2),$$

$$B_2 \sim (1/2)N(\mu_{51}, \sigma_5^2)+(1/2)N(\mu_{52}, \sigma_5^2),$$

$$F_2 \sim (1/4)N(\mu_{61}, \sigma_6^2)+(1/2)N(\mu_{62}, \sigma_6^2)+(1/4)N(\mu_{63}, \sigma_6^2),$$

where μ_1, μ_2, and μ_3 are the means of P_1, F_1 and P_2, respectively, μ_{41} and μ_{42} are the means of the two components in B_1, μ_{51} and μ_{52} are the means of the two components in B_2, and μ_{61}, μ_{62} and μ_{63} are the means of the three components in F_2, σ_e^2 is the environmental variance, σ_4^2, σ_5^2 and σ_6^2 (both polygene variation and environmental variation are included in these variances) are the common variances of components in B_1, B_2 and F_2, respectively. The likelihood function can therefore be built and maximum-likelihood estimates can be found through the EM algorithm (McLachlan and Basford 1988). It should be pointed out that the different models, i.e., NULL, 1MG, 2MG, MX1 and MX2, have different sets of parameters. Here, only an example for MX1 is presented. Major-gene effects, i.e., a_1, d_1, a_2, and d_2, can therefore be estimated from the equations between component means and gene effects (Gai and Wang 1998). JSA cannot estimate effects of individual polygenes; however, additive and dominance effects ([*a*] and [*d*]) of the polygene system can still be estimated (Tables 2 and 3)

Table A1 The *m*, *a*, *d* (midpoint, additive dominance) values and the genetic values of the three possible allelic combinations (*aa*, *Aa*, *AA*; assuming *a* and *A* are the two alleles) for each gene defined in six classes of the genetic model. Genes 1 and 2 refers to the two major genes and genes 3–12 refers to the polygenes. Set numbers correspond to those used in Table 1

Set	Model	Gene	*m*, *a*, *d* Parameters			Allelic combination		
			m_i	a_i	d_i	*aa*	*Aa*	*AA*
1	1MG	1	0.00	1.10	1.10	−1.10	1.10	1.10
		2	0.00	0.00	0.00	0.00	0.00	0.00
		3–12	0.00	0.00	0.00	0.00	0.00	0.00
2	2MG	1	0.00	1.03	1.03	−1.03	1.03	1.03
		2	0.00	0.37	0.37	−0.37	0.37	0.37
		3–12	0.00	0.00	0.00	0.00	0.00	0.00
6	MX1	1	0.00	0.97	0.97	−0.97	0.97	0.97
		2	0.00	0.00	0.00	0.00	0.00	0.00
		3–12	0.00	0.20	0.00	−0.20	0.00	0.20
7	MX2	1	0.00	0.89	0.89	−0.89	0.89	0.89
		2	0.00	0.37	0.37	−0.37	0.37	0.37
		3–12	0.00	0.20	0.00	−0.20	0.00	0.20
16	PG	1	0.00	0.00	0.00	0.00	0.00	0.00
		2	0.00	0.00	0.00	0.00	0.00	0.00
		3–12	0.00	0.42	0.42	−0.42	0.00	0.42
41	NULL	1	0.00	0.00	0.00	0.00	0.00	0.00
		2	0.00	0.00	0.00	0.00	0.00	0.00
		3–12	0.00	0.00	0.00	0.00	0.00	0.00

Appendix 2: construction of genetic effects in the simulation experiment

Construction of genetic models in the QU-GENE engine requires the effects of each individual gene to be specified. For each class of genetic model considered in the experiment (1MG, 2MG, MX1, MX2, PG, NULL), individual gene effects were constructed, noting that the numbers of genes with a positive value differs among the six classes of the genetic model. Table A1 shows the individual gene effects for the six classes of the genetic model for some example sets. Values for each gene are given in terms of the *m*, *a*, *d* (midpoint, additive, dominance) model and for the three possible allelic combinations of each gene (*aa*, *Aa*, *AA*; assuming *a* and *A* are the two alleles). For example, for the 1MG model and a heritability value of 0.9, we have a single positive gene with *m*=0.00, *a*=1.10 and *d*=1.10 (in terms of the *m*, *a*, *d* model). This results in the values −1.10, 1.10 and 1.10 for allelic combinations *aa*, *Aa* and *AA*, respectively. The *m*, *a*, *d* values were obtained using the appropriate equations presented in the Materials and methods section. That is, for the 1MG model, we have

$$h^2 = h_{mg1}^2 = 0.9 = \frac{1}{2}a_1^2 + \frac{1}{4}d_1^2, \text{ where } a_1 = 1.10 \text{ and } d_1 = 1.10.$$

Because JSA can only estimate the additive and dominance effects of the polygene system, [*a*] and [*d*] in Table 2 are the confounding effects from all polygenes. Polygene dispersion affects the values of [*a*] and [*d*]. For example, for Set 16, [*a*] are 4.2, 0, and −4.2 for GD1, GD2 and GD3, respectively

References

Akaike H (1977) On entropy maximization principle. In: Krishnaiah PR (ed) Applications of statistics. North-Holland Publishing Company, Amsterdam, pp 27–41

Basnayake J, Ludlow MM, Cooper M, Henzell RG (1993) Genotypic variation of osmotic adjustment and desiccation tolerance in contrasting sorghum inbred lines. Field Crops Res 35: 51–62

Basnayake J, Cooper M, Ludlow MM, Henzell RG, Snell PJ (1995) Inheritance of osmotic adjustment to water stress in three grain sorghum crosses. Theor Appl Genet 90:675–682

Beavis WD (1998) QTL analyses: power, precision and accuracy. In: Paterson AH (ed) Molecular dissection of complex traits. CRC Press, Boca Raton, Florida, pp 145–162

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm (with discussion). J R Statist Soc 39:1–38

Elston RC (1984) The genetic analysis of quantitative trait differences between two homozygous lines. Genetics 108:733–744

Elston RC, Steward J (1973) The analysis of quantitative traits for simple genetic models from parental, F_1 and backcross data. Genetics 73:695–711

Famula TR (1986) Identifying single genes of large effect in quantitative traits using best linear unbiased prediction. J Anim Sci 63:68–76

Gai J, Wang J (1998) Identification and estimation of a QTL model and its effects. Theor Appl Genet 97:1162–1168

Gai J, Zhang Y, Wang J (2001) Plant quantitative loci system. Academic Press, Beijing

Hoeschele I (1988) Genetic evaluation with data presenting evidence of mixed major-gene and polygenic inheritance. Theor Appl Genet 76:81–92

Janss LLG, Thompson R, van Ardendonk JAM (1995) Application of Gibbs sampling for inference in a mixed major gene-polygene inheritance model in animal populations. Theor Appl Genet 91:1137–1147

Kearsey MJ, Farquhar AGL (1998) QTL analysis in plants; where are we now? Heredity 80:137–142

Kearsey MJ, Pooni HS (1996) The genetical analysis of quantitative traits. Chapman and Hall, London

Knott SA, Haley CS, Thompson R (1991) Methods of segregation analysis for animal breeding data: a comparison of power. Heredity 68:299–311

Ludlow MM, Muchow RC (1990) A critical evaluation of traits for improving crop yield in water-limited environments. Adv Agron 43:107–153

McLachlan GJ, Basford KE (1988) Mixture models: inference and application to clustering. Marcel Dekker, Inc, New York

Morgan JM (1980) Osmotic adjustment in the spikelet and leaves of wheat. J Exp Bot 31:655–665

Morton ME, MacLean CJ (1974) Analysis of family resemblance. III. Complex segregation analysis of quantitative traits. Am J Hum Genet 26:489–503

Paterson AH (1998) Molecular dissection of complex traits. CRC Press, Boca Raton, Florida

Podlich DW, Cooper M (1998) QU-GENE: a platform for quantitative analysis of genetic models. Bioinformatics 14:632–653

Wang J (1996) Studies on identification of mixed major-gene and polygene inheritance of quantitative traits and estimation of genetic parameters. Doctoral dissertation, Nanjing Agricultural University

Wang J, Gai J (2001) Mixed inheritance model for resistance to agromyzid beanfly (*Melanagromyza sojae* Zehntner) in soybean. Euphytica (in press)

Wang J, Wang J, Zhu L, Gai J (2000) Major-polygene effect analysis of resistance to bacterial blight (*Xanthomonas campestris* pv *oryzae*) in rice. Acta Genet Sinica 27:34–38