NOTE

A Comparison of Tests for QTL Mapping with Introgression Libraries Containing Overlapping and Nonoverlapping Donor Segments

Gregory S. Mahone, Dietrich Borchardt, Thomas Presterl, and Matthias Frisch*

ABSTRACT

Near-isogenic line (NIL) libraries can be used to detect beneficial trait variation in germplasm that is unadapted or has poor agronomic performance. The objectives of our study were to compare the *t* test, Dunnett test, and linear model test with regard to the power and false positive rate of quantitative trait loci (QTL) detection in NIL libraries of different design. We employed computer simulations with maize genome models to investigate nonoverlapping NIL libraries, overlapping NIL libraries, and stepped aligned inbred recombinant strains (STAIRS) libraries for traits with oligogenic inheritance. Quantitative trait loci detection power of the linear model and Dunnett tests were similar for nonoverlapping and STAIRS libraries; for overlapping NIL libraries the Dunnett test was slightly superior. False positives were greatest for the *t* test and lowest for the linear model test. False positive sums with the Dunnett test were generally higher than for the linear model test if the heritability was 0.9 or lower. We conclude that the linear model test is superior to the Dunnett test for nonoverlapping NIL libraries and for overlapping NIL libraries with heritabilities below 0.9, as usually occur. Analysis of a rapeseed (*Brassica napus* L.) library revealed two other major advantages of the linear model test. First, detection of positive and negative QTL effects present in the same line is possible. Second, for NILs with multiple donor segments, observed phenotypic differences can be assigned to individual chromosome segments.

G.S. Mahone and M. Frisch, Institute of Agronomy and Plant Breeding II, Justus Liebig University, Heinrich-Buff-Ring 26-32, Giessen, Germany 35392; D. Borchardt and T. Presterl, KWS SAAT AG, Grimsehlstr. 31, Einbeck, Germany 37555. Received 29 June 2011. *Corresponding author (Matthias.Frisch@agrar.uni-giessen.de).

Abbreviations: BH, Bonferroni-Holm; DH, double haploid; FDR, false discovery rate; NIL, near-isogenic line; STAIRS, stepped aligned recombinant inbred strains; QTL, quantitative trait locus (or loci).

Crop domestication has had a narrowing effect on the genetic variation existing in many species, to the point that harnessing the natural variation prevalent in nonadapted exotic germplasm is increasingly important for improving yield, quality, and resistance (Gur and Zamir, 2004; McCouch, 2004). To uncover and exploit trait variation in exotic by elite crosses, near-isogenic line (NIL) libraries, also referred to as introgression libraries, are a powerful tool in plant breeding. Near-isogenic line libraries have proven useful for investigating yield in rice (*Oryza sativa* L.) (Cheema et al., 2008) and tomato (*Lycopersicon esculentum* Mill.) (Eshed and Zamir, 1995), disease resistance in wheat (*Triticum aestivum* L.) (Leonova et al., 2007) and barley (*Hordeum vulgare* L.) (Schmalenbach et al., 2008), drought tolerance in wheat and barley (review, Nevo and Chen, 2010), metabolites in tomato (Rousseaux et al., 2005) and maize (*Zea mays* L.) (Yang et al., 1995), quality traits in barley (Schmalenbach and Pillen, 2009) and rye (*Secale cereale* L.) (Falke et al., 2009), flowering time in maize (Szalma et al., 2007), and agronomic traits in barley (Schmalenbach et al., 2009) and rye (Falke et al., 2009).

Introgression libraries consist of NILs that contain donor segments in a background of recurrent parent genome. The

Published in Crop Sci. 52:2198–2205 (2012).
doi: 10.2135/cropsci2011.06.0341

introgressed segments are typically short stretches of donor genome, which may overlap in successive NILs depending on the aims of library construction. An alternative to typical NIL libraries is the stepped aligned inbred recombinant strains, or STAIRS, library (Koumproglou et al., 2002). The STAIRS library contains donor segments of increasing size, starting from small donor segments to entire donor chromosomes. This pattern is repeated for each chromosome. The advantage of this design is that it is easier to produce than typical NIL libraries. However, to our knowledge no one has investigated the performance of QTL analysis in STAIRS libraries compared with conventional NIL libraries with smaller targeted introgressions.

Analysis of introgression libraries typically involves a series of pairwise tests between the NILs and the recipient parent for the trait in question (Eshed and Zamir, 1995; Rousseaux et al., 2005; Eduardo et al., 2007; Schmalenbach and Pillen, 2009; Falke et al., 2009). A recent paper by Falke and Frisch (2011) proposed an alternative testing procedure, in which a linear model was used to estimate the effects of the segments directly. The study examined the differences in sums of correctly detected effects and false positive effects in NIL libraries with either nonoverlapping or overlapping segments. Results were based solely on the proposed linear model test but do not include a comparison with the pairwise tests that have been previously used. However, the efficiency of QTL detection might well depend on the statistical test used. While methods employing linear models and regression methods have been previously used to locate QTL in introgression libraries (Wang et al., 2006; Wang et al., 2007; Coles et al., 2011), the effect of the type of test used to identify QTL in NIL libraries has not yet been investigated.

The objectives of this study were to (i) compare the sums of correctly detected and false positive effects for pairwise *t* tests, the Dunnett test, and the linear model test in QTL detection with introgression libraries, (ii) compare the statistical properties of the tests for overlapping and nonoverlapping NIL libraries and STAIRS libraries, (iii) propose suitable tests that may enhance the precision of QTL detection in NIL libraries depending on the heritability and the amount of segment overlap, and (iv) validate our simulation results with experimental data of a rapeseed (*Brassica napus* L.) introgression library.

MATERIALS AND METHODS

Simulations

A model of the maize genome comprising 10 chromosomes of 160 cM length was used for our simulations. Linkage maps with marker distances (*d*) of 20, 10, and 5 cM were investigated for three types of introgression libraries: nonoverlapping libraries, overlapping libraries, and STAIRS libraries (Fig. 1). Nonoverlapping libraries contained donor segments that are contiguous but do not overlap. Overlapping libraries contain segments that are each present in two NIL lines. For STAIRS libraries, each chromosome was divided in parts of equal length. The first of the lines that covered the genome of a chromosome carried one such segment located at the telomere. The second line carried in addition the chromosome segment directly adjacent to the first one. For each subsequent line, a further segment was added, such that the last line contained the donor genome of the entire chromosome. Ten recipient parent plots per replication were included in phenotyping, as justified in Falke and Frisch (2011). The software Plabsoft (Maurer et al., 2008) was used for the simulations. Each simulation run was repeated with heritabilities of 0.5, 0.6, 0.7, 0.8, 0.9, and 0.9999. Therefore each simulation run incorporated the type of introgression library, marker distance *d*, quantitative genetic scenario, and heritability. All simulations were repeated 5000 times to ensure high numerical accuracy and reduce the effects of sampling.

Quantitative Genetic Models

We considered a polygenic trait and assumed that the genotypic value of the donor parent is 100 units superior to that of the recipient parent. The trait was controlled by major genes, minor genes, and genes with small effects. In all scenarios, 10 genes with small effects of size 1 were assumed. The remaining difference between donor and recipient was assigned to major and minor genes in four different scenarios (Table 1). These differed in the number n_a of major and n_i of minor genes, and their effect sizes s_a and s_i, respectively. The genes with small effects were included as background or stochastic noise, as it is unrealistic to assume that all genetic effects underlying a quantitative trait can be modeled and/or detected. The sizes of major and minor effects intended to model oligogenic resistance or quality traits. For each simulation run, genes were assigned to a different set of random locations in the genome.

QTL Detection

We employed pairwise tests and a linear model test to investigate the presence of QTL on donor segments in the NIL libraries. The pairwise testing methods consisted of comparisons between each NIL and the recipient parent. The rationale is that since each NIL contains a single donor segment, differences in phenotype between each NIL and the recipient parent can be attributed to the presence of the donor segment. Pairwise testing consisted of two methods, *t* tests and the Dunnett test (Dunnett, 1955). Pairwise *t* tests were performed with and without adjustment for multiple testing. In unadjusted tests, the per-comparison type I error rate was 0.05. Multiple comparison adjustment of tests followed two procedures: (i) the Bonferroni-Holm (BH) procedure proposed by Holm (1979) for an experiment-wise type I error rate of 0.05 and (ii) the procedure proposed by Benjamini and Hochberg (1995) for a false discovery rate (FDR) of 0.05. For STAIRS libraries, we used the standard error of a difference of treatment means for a Dunnett type comparison of two subsequent lines. Line genotypic values (and by extension the effect of the segment in the line) were calculated from genotypic values of previous lines in the library. To calculate *p* values for the Dunnett test, we used the density function of the multivariate normal distribution provided by the R package *mvtnorm* (Genz et al., 2011).

The linear model test consisted of estimating the effects of donor segments with a linear model. An *F* test was subsequently

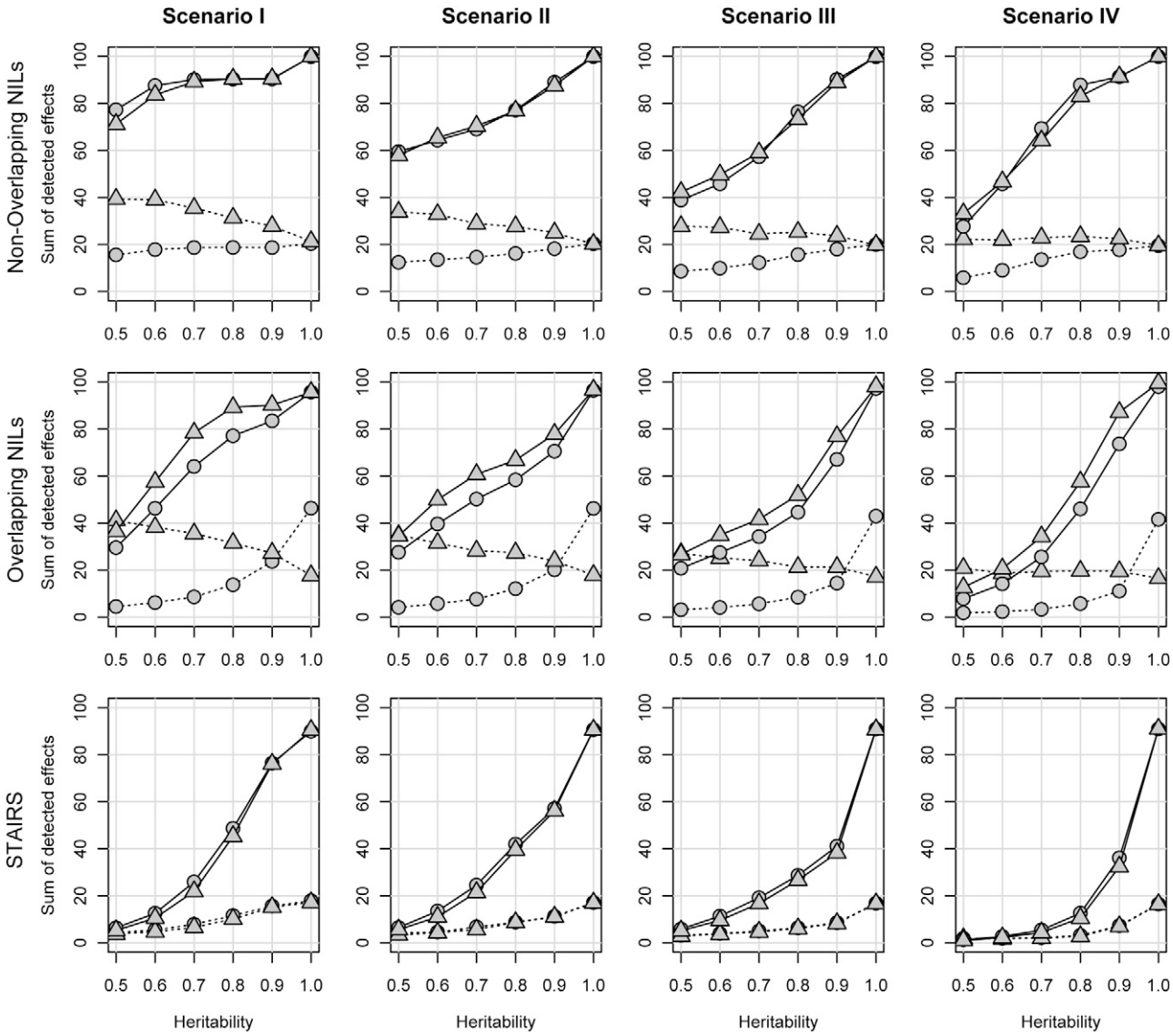

Figure 1. Genomic composition of the different types of near-isogenic lines libraries. Donor segments are indicated with black and recipient parent genome with gray lines. The dashed lines flanking the donor segments are genomic segments of unknown parental origin, located between markers at the end of the donor segment and flanking markers. Marker distance is 5 centiMorgans (*d*).

conducted for every segment to determine significance. Multiple testing adjustments for results from the linear model test were made using the BH procedure. The linear model test procedure was used in total as described by Falke and Frisch (2011).

For our analysis, we determined the sum of correctly detected effects and the sum of false positive effects (false positive rate) of each test to have measures of their efficiency. The sum of correctly detected effects, our measure of test detection power, was calculated by summation of the true QTL effects of segments for which the null hypothesis was correctly rejected and was collected for the total, major, minor, and small effect QTL. True QTL effects were used rather than the estimated effects to avoid bias due to overestimation of QTL effects. The false positive rate was calculated by the summation of all detected QTL effects of segments for which the null hypothesis was incorrectly rejected. A false positive for the Dunnett test was declared when a NIL was found significant despite containing no QTL. We avoided using the terms type I and type II error because not only statistical sampling contributes to these errors. In addition, "genetical sampling" due to QTL located on chromosomal segments adjacent to the target regions is also present in our measurements for success or failure of QTL detection.

Table 1. Quantitative genetic scenarios. Number (n_a and n_i) and effect size (s_a and s_i) of major and minor genes, respectively, for four scenarios.

Scenario	Major Genes		Minor Genes	
	n_a	s_a	n_i	s_i
I	3	30	–	–
II	2	30	3	10
III	1	30	6	10
IV	–	–	9	10

Validation with Experimental Data

To validate our results from simulations, we analyzed an introgression library in rapeseed using the Dunnett test and the linear model. The rapeseed introgression library was contributed by KWS SAAT AG, Einbeck, Germany. The library is a BC_4 double

haploid (DH) population created from the crossing of the winter rapeseed varieties 'Mansholt's Hamburger Raps' and 'Samourai' and was grown in randomized field trials over 3 yr. The same parental cross combination has been previously investigated for various trait QTL (Uzunova et al., 1995; Marwede et al., 2005). Glucosinolate content (μmol/g), measured using near-infrared spectroscopy, was collected from five locations each in 2006 and 2007 and a single location in 2008. The heritability was above 0.9, which is in accordance with previous studies (Marwede et al., 2004). The recipient parent was included repeatedly in field trials. The linkage map consisted of 176 amplified fragment length polymorphic markers and had a length of 1361 cM, resulting in an average marker distance of 6.8 cM. The NIL library contained 127 lines. Each NIL carried at least one donor segment, the average number of donor segments was between two and three. The average segment length was 21.6 cM and the donor genome coverage of the library was 87.7%. Regarding comparison with our simulation libraries, the rapeseed library would most resemble an overlapping library, though it also has aspects of the nonoverlapping (segments present in only a single line) and STAIRS (segments present in multiple lines) libraries.

The model used for the Dunnett analysis was:

$$Y_{ijk} = \mu + G_i + L_j + Y_k + e_{ijk}$$

where Y_{ijk} is the glucosinolate content of genotype G_i at location L_j in year Y_k, with a grand mean of μ and residual error e_{ijk}. The Dunnett test was incorporated using PROC GLM of SAS software version 9.2 (SAS Institute Inc.). A detailed description of the linear model that was used to estimate and test the effects of individual chromosome segments was presented by Falke and Frisch (2011). Calculations were performed with R (R Development Core Team, 2011).

RESULTS

Total detection power was similar for the linear model and Dunnett tests for each of our three simulation sets. The *t* tests, which were included only in the nonoverlapping set, also had similar detection power (results not shown). The Dunnett test generally had a higher power of detection in the overlapping library set. Overall, detection power was directly related to heritability and QTL effect size, an expected result. Power decreased overall as the QTL component of the libraries moved from a few large-effect QTL (major QTL; Scenario I) to many QTL with smaller effects (minor QTL; Scenario IV). Within each scenario, the presence of major QTL lowered the power of both tests at low heritabilities. As the number of major QTL decreased across scenarios, detection of major QTL increased at these low heritabilities. Detection power of minor QTL also increased overall as major QTL number decreased. For both tests, increasing degree of introgression overlap negatively affected power of detection.

In the nonoverlapping library set, false positives decreased across all tests with decreasing marker distances *d* (Table 2). Increasing heritabilities caused consistent decreases in the false positives for the FDR adjusted and the unadjusted pairwise tests across all marker distances and scenarios. At small marker distances (d = 5 cM), the pairwise tests showed decreasing false positives with increasing heritabilities. In contrast, the linear model test showed an increase in false positives as heritabilities increased, though these values were much lower than those of the pairwise tests at low heritabilities. At the highest heritabilities, false positive rates were similar for all tests. The *t* tests were excluded from comparison in the overlapping and STAIRS library sets because of their high false positive rates in the nonoverlapping library set.

Overall, false positives generally decreased with decreasing genetic variance, for example as QTL effect sizes decreased and as heritability increased, with the exception of the linear model test (Fig. 2). The linear model showed generally lower false positive rates than the Dunnett test in the nonoverlapping library and the overlapping library excluding high heritabilities, with similar rates as the Dunnett test found in the STAIRS library. Marker density also affected false positives, as the introgressed segments can be more clearly defined (Table 2). This lowers the chance that a QTL will be outside the marker-defined segment to which the QTL is ascribed.

In the rapeseed library, the Dunnett test detected 26 NILs that had a significantly different glucosinolate content than the recipient parent (Table 3). Eight of those carried a single donor introgression. The remaining carried between two and six introgressions, with the most common number of introgressions being three. All significant lines had glucosinolate contents greater than that of the recurrent parent, with an average difference in means of 22.6. The linear model test found 15 separate significant donor introgressions, varying in length from one to four markers. One to six introgressions were present in 54 NILs. On four occasions, positive and negative QTL located in close proximity were detected with the linear model test. Most of the lines containing these contrasting-effect QTL were not significant in the Dunnett test results.

DISCUSSION

Statistical Tests

Our results confirm that the Dunnett test is better suited for analyzing NIL libraries than pairwise *t* tests. Even with adjustment for multiple testing, the *t* tests had a considerably greater false positive rate (Table 2). A further increase in the precision of QTL detection is expected with the linear model analysis, in particular for libraries with some chromosome regions duplicated in more than one NIL, as in the libraries of previous studies (Eduardo et al., 2007; Falke et al., 2008). The advantage of the linear model test is likely due to a more precise estimation of the residual variance by using the entire library rather than the recipient parent and the introgression line under consideration.

Detection of a QTL depends on the amount of genetic variance that can be attributed to the QTL compared with

Table 2. Sum of false positive effects in maize (*Zea mays* L.) near-isogenic lines libraries with nonoverlapping donor segments for varying marker distances (*d*), heritabilities (h^2), and quantitative genetic scenarios (I–IV). The testing methods are as follows: LM, linear model test; DT, Dunnett test; PW_n, unadjusted pairwise *t* test; PW_{fdr}, pairwise *t* test adjusted using false discovery rate; PW_{bh}, pairwise *t* test adjusted using Bonferroni-Holm. Each sum of false positive effects is a mean value from 5000 simulation runs.

		h^2						h^2					
d	Test	0.5	0.6	0.7	0.8	0.9	1	0.5	0.6	0.7	0.8	0.9	1
		Scenario I						Scenario II					
20	LM	48.0	62.8	71.9	75.5	77.3	85.0	39.0	47.2	52.6	58.5	68.8	81.3
	DT	85.2	87.1	91.3	91.7	87.5	85.7	71.1	72.5	74.1	75.7	78.4	82.7
	PW_n	371.3	325.3	278.0	226.1	179.4	87.5	324.7	275.8	242.0	200.2	161.8	84.7
	PW_{fdr}	215.8	203.7	179.4	160.5	130.5	86.8	191.5	166.5	156.1	141.6	125.1	83.9
	PW_{bh}	92.1	95.6	98.9	94.2	90.8	84.3	82.9	80.3	82.8	81.8	83.7	83.7
10	LM	28.7	34.2	36.7	37.5	37.6	41.3	22.0	25.3	27.7	30.5	34.8	39.7
	DT	55.5	55.0	55.1	49.9	47.0	41.4	47.2	47.1	45.3	42.8	42.6	40.2
	PW_n	304.7	259.1	218.2	171.0	127.9	43.6	265.3	227.3	187.6	151.2	114.8	42.2
	PW_{fdr}	172.1	148.0	128.4	108.3	83.6	42.9	147.6	122.3	106.4	94.5	77.1	41.2
	PW_{bh}	67.9	61.7	59.3	54.5	49.0	41.1	56.9	54.8	50.7	47.4	45.4	40.3
5	LM	15.5	17.8	18.6	18.7	18.6	20.3	12.3	13.5	14.5	16.1	18.1	20.3
	DT	39.4	39.0	35.5	31.3	27.8	21.3	33.8	32.8	28.8	27.6	25.0	20.2
	PW_n	275.2	221.1	182.1	144.1	103.0	23.2	233.5	196.5	160.2	126.1	91.6	22.4
	PW_{fdr}	136.1	112.2	99.6	83.8	63.3	21.9	120.7	100.5	85.6	76.1	56.3	21.6
	PW_{bh}	50.8	45.0	42.8	36.8	30.6	21.4	40.6	39.2	34.7	31.4	27.8	20.4
		Scenario III						Scenario IV					
20	LM	27.9	33.6	41.2	52.9	68.5	80.4	17.6	27.1	40.9	58.6	69.7	78.8
	DT	54.4	57.9	60.9	65.7	74.7	80.6	42.3	48.5	56.5	66.5	74.7	78.6
	PW_n	274.5	233.0	208.4	176.6	143.8	82.9	218.3	194.1	168.9	148.1	123.1	79.9
	PW_{fdr}	152.6	141.5	132.3	124.4	115.0	82.6	124.3	123.6	120.1	115.8	105.4	79.4
	PW_{bh}	65.5	66.9	69.8	72.8	79.2	80.9	51.9	54.8	62.3	69.8	78.4	78.8
10	LM	15.1	17.7	21.6	28.1	33.9	38.3	10.1	15.9	23.8	31.6	34.5	38.3
	DT	38.1	36.8	37.8	38.9	41.0	39.7	29.4	31.0	34.4	37.4	39.1	38.3
	PW_n	221.0	187.1	155.1	130.3	99.0	40.6	173.0	149.3	124.2	106.4	82.3	39.3
	PW_{fdr}	117.4	102.7	93.7	84.5	72.5	40.5	95.3	87.8	83.3	75.1	64.0	38.8
	PW_{bh}	45.6	44.8	42.2	43.6	42.5	39.1	35.9	36.8	37.5	41.3	40.7	38.0
5	LM	8.5	9.8	12.2	15.6	18.0	19.9	5.8	9.0	13.5	16.8	17.6	19.4
	DT	27.7	27.3	24.5	25.3	23.6	19.8	22.3	21.9	22.9	23.4	22.4	19.5
	PW_n	192.6	162.3	135.4	106.2	77.4	21.6	150.7	123.6	103.7	84.8	62.6	20.5
	PW_{fdr}	101.9	84.0	76.9	65.4	49.6	20.5	83.2	71.3	66.1	56.6	44.3	20.1
	PW_{bh}	35.0	32.1	29.1	28.1	26.0	20.1	27.9	26.3	25.7	26.7	24.2	19.5

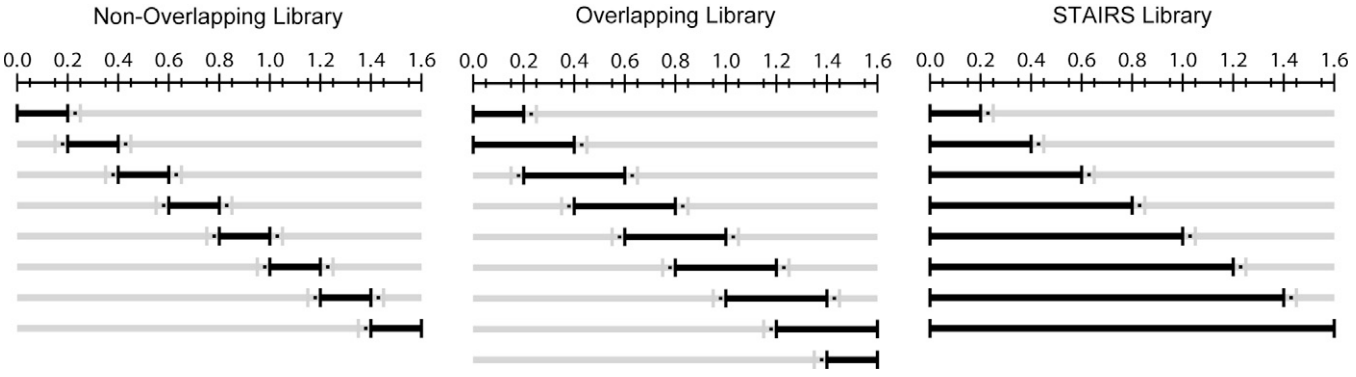

Figure 2. Sums of correctly detected effects (solid line) and false positive effects (dotted line) in different maize (*Zea mays* L.) introgression library types (*d* = 5 cM) and across four genetic scenarios. The graphs compare the linear model test (circle) with the Dunnett test (triangle).

the total variance in the experiment. The four scenarios (Table 1) show a progression from few QTL of large effect to many QTL of smaller effect. As the individual QTL decrease in effect size and increase in number, the variance explained by a single QTL decreases. Likewise, decreasing heritability also decreases the relative variance that a single

Table 3. Lines containing donor segments found to be significant for glucosinolate content in the rapeseed (*Brassica napus* L.) introgression library. Linear model: All lines carrying significant segments are listed and the significant segments are shown. Dunnett test: All significant lines are presented and all donor segments that are contained in those lines are shown.

Linear model		Dunnett test	
Line	Segments	Line	Segments
9	98,119,142	9	98,119–125,142
44	142	–	–
47	128	47	127–128
50	119,142	50	119–127,142
55	98,119,142	55	98,119–128,142
58	162–163,165–166	–	–
59	162,165–166	–	–
117	153,163,165–166	117	34–35,153,163–166
124	82	124	82
172	98	–	–
189	142	–	–
203	119	203	56–57,62,119–125
212	80–82	–	–
227	165–166	227	56–57,62,165–166
257	163,165–166	257	83–85,163–166
258	163,165–166	258	149,163–166
260	163,165–166	260	163–166
261	119,142	–	–
262	98,119,142	–	–
263	98,158	263	98,158
264	119,128	264	119–128
265	119	265	119
280	153	–	–
287	153	–	–
293	48,80–82,165–166	293	48,80–82,112–115,165–166
294	153,163,165–166	294	87–91,132,134,136–137,153,163–166
296	128,142	296	128,142
367	98	–	–
387	165–166	387	53,55–57,62,165–166
430	153,163,165–166	430	32,112,149,153,163–166
496	80–82	–	–
498	80–82,153,163,165–166	498	80–91,153,163–166
499	163,164–165	499	34–35,83–88,90–91,163–166
576	163,164–165	576	163–166
578	38–41	–	–
641	48	–	–
789	38–41	–	–
814	158	–	–
842	38–41	–	–
864	163,165–166	864	83–85,149,163–166
873	38–41	–	–
875	80–82	–	–
877	80–82,128	–	–
1036	48	1036	48,111,113–115
1150	4,39–40,48,60,103–106,108–110	–	–
1155	39–40,82,108–110	–	–
1157	103–106,108–110	–	–
1158	48,103–106,108–110	–	–
1196	48,80–82,128	–	–
1204	142	–	–
1332	48	–	–
1373	48	–	–
–	–	1395	159–161
–	–	1397	159–161
1433	142	–	–
1548	48,103–106,108–110	–	–

QTL explains. This decrease in the variance explained by individual QTL is a contributing reason for the observed decrease in power. Our results indicate that these factors, as well as the number of times a QTL is present in the library, all contribute to the variance and therefore affect detection. For instance, power of detection was highest with nonoverlapping NIL libraries and few major genes (Scenario I) for both the linear model test and the Dunnett test (Fig. 2). Although overlapping NIL libraries and STAIRS libraries may have advantages owing to the reduced efforts for establishing the library, we conclude that these advantages come at the cost of a considerably lower power of QTL detection. This is especially true for minor-effect QTL, which in some cases may be the focus of introgression line population development. Falke and Frisch (2011) reported a considerable lower power of QTL detection with overlapping rather than with nonoverlapping NIL libraries employing the linear model test, and our findings extend those results also to STAIRS libraries.

With increasing heritability, the false positive rate increased for the linear model test and decreased for the Dunnett test in nonoverlapping and overlapping NIL libraries. The increase observed for the linear model test is due in part to a higher power to detect QTL located between the marker at the end of the target segment and the first flanking marker at which selection is performed for the recipient genome, that is, QTL between known donor DNA and known recipient DNA. This trend may also reflect detection of adjoining segments that do not contain QTL but are being declared significant because of low residual variance present at high heritabilities. The decrease observed for the Dunnett test can be explained with the decrease in the residual variance caused by increasing heritability, which reduces spurious QTL detections. For low heritabilities in the nonoverlapping library and the overlapping library, the false positive rate of the Dunnett test was considerably higher than the linear model test. For instance, at marker distance of 0.05 cM, the false positive rate for the Dunnett test was more than twice as high as the linear model test for low heritabilities in the nonoverlapping library. In overlapping NIL libraries, Dunnett test power was slightly greater than the linear model test power, but cannot be exploited because of the inflated false positives. To further investigate false positive rates in the Dunnett test, additional simulations were run in overlapping libraries. In these simulations, false positives were only declared when both lines with non-QTL-containing donor segments were declared significant. While this lowered false positive rates, the linear model test generally still outperformed the Dunnett test regarding false positives.

Evaluating overall test performance by incorporating both the detection power and false positive rate provides a more definitive answer. One way to synthesize the results of type I and type II error rates is to calculate the ratio of

test power to false positive rate. We performed a similar calculation with our values of sum of correctly detected effects vs. the sum of false positive effects. The ratio of major and minor QTL detection power to false positives generally increased for both tests as genetic variance decreased, proceeding from Scenario I to Scenario IV. Within each scenario, major QTL detection ratio of the linear model test peaked at low heritabilities and decreased at high heritabilities while the Dunnett test peaked at high heritabilities. This was true in the nonoverlapping and overlapping libraries, with both tests peaking at high heritabilities in the STAIRS library. The ratio was higher for the linear model overall than for the Dunnett test, as both tests had similar power but the Dunnett test had generally higher false positives. The largest difference between the two tests occurred in the overlapping library. At the lowest heritabilities, the linear model test ratio was over 6:1 for detection power to false positive rate for Scenarios I to III and over 4:1 for Scenario IV. The Dunnett test was below 1:1 for Scenarios I and IV and slightly above 1:1 for Scenarios II and III at those same heritabilities.

To summarize, the sum of correctly detected effects identifies neither the Dunnett test nor the linear model test as the superior method in every case. Lower false positives may be regarded as an advantage of the linear model test in most instances. In overlapping NIL libraries, the Dunnett test is in particular not suitable if heritabilities are low; with heritabilities between 0.9 and 1 it can be a favorable alternative to the linear model test. An additional point to consider is the flexibility allowed by using a linear model approach. Model building is possible, as well as interactions of genetic effects. Using introgression libraries, linear model methods could uncover and investigate epistasis with precision that is hard to achieve in segregating populations. Using mixed models is also possible, as done in a recent publication by Coles et al. (2011).

Rapeseed Introgression Library

A principal difference between the Dunnett test and the linear model test is that the linear model is testing for the presence of QTL on individual chromosome segments, whereas the Dunnett test is testing NILs as a whole. For example, line 203 was found to be significant using the Dunnett test, and it contains three separate introgressed donor segments (Table 3). This includes segment 119, which was found to be significant using the linear model test. The remaining segments, however, were not declared significant with this test. Using the linear model test was able to provide much more information on the location of the QTL than could be determined with the Dunnett test. Similar results were obtained for lines 227, 387, and 1036. We conclude that the linear model test is of great advantage for NIL libraries with lines that carry multiple introgressions, because it can detect those introgressions that are responsible for the differences in the phenotype of the NIL and the recipient parent. Additional simulations support the results, indicating that the linear model test has higher power than the Dunnett test when multiple QTL are on separate introgressions in the same line.

Of the 30 NILs containing segments detected with the linear model but not determined to be significant with the Dunnett test, 17 carried QTL with both positive and negative effects. This includes nearly every NIL containing multiple significant segments detected with the linear model. For example, two QTL with different signs and similar effect size appear in segments 38 to 41. These segments are not present in any NILs detected with the Dunnett test. In conclusion, a second big advantage of the linear model test is that it is able to find QTL in lines that carry more than one QTL with different signs on different chromosome segments.

To investigate the transferability of our results we performed simulations with a model of the barley genome. The detection power and false positive rates differed, but the trends observed for different types of libraries, quantitative genetic scenarios, heritabilities, and the choice of tests were similar. We conclude that our results are robust with respect to the number and length of the chromosomes and should serve as reliable guidelines for introgression libraries in other crops.

Acknowledgments

We would like to thank Frank Breuer for providing experimental data. The financial support from KWS SAAT AG for Gregory Mahone is gratefully acknowledged.

References

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J. R. Stat. Soc., B 57:289–300.

Cheema, K.K., N.S. Bains, G.S. Mangat, A. Das, Y. Vikal, D.S. Brar, G.S. Khush, and K. Singh. 2008. Development of high yielding IR64 × *Oryza rufipogon* (Griff.) introgression lines and identification of introgressed alien chromosome segments using SSR markers. Euphytica 160:401–409. doi:10.1007/s10681-007-9560-x

Coles, N.D., C.T. Zila, and J.B. Holland. 2011. Allelic effect variation at key photoperiod response quantitative trait loci in maize. Crop Sci. 51:1036–1049. doi:10.2135/cropsci2010.08.0488

Dunnett, C.W. 1955. A multiple comparison procedure for comparing several treatments with a control. J. Am. Stat. Assoc. 50:1096–1121.

Eduardo, I., P. Arús, A.J. Monforte, J. Obando, J.P. Fernández-Trujillo, J.A. Martínez, A.L. Alarcón, J.M. Álvarez, and E. van der Knaap. 2007. Estimating the genetic architecture of fruit quality traits in melon using a genomic library of near isogenic lines. J. Am. Soc. Hortic. Sci. 132(1):80–89.

Eshed, Y., and D. Zamir. 1995. An introgression line population of *Lycopersicon pennellii* in the cultivated tomato enables the identification and fine mapping of yield-associated QTL.

Genetics 141:1147–1162.
Falke, K.C., and M. Frisch. 2011. Power and false-positive rate in QTL detection with near-isogenic line libraries. Heredity (in press).
Falke, K.C., Z. Sušić, B. Hackauf, V. Korzun, J. Schondelmaier, P. Wilde, P. Wehling, H. Wortmann, R. Mank, J. Rouppe van der Voort, H.P. Maurer, T. Miedaner, and H.H. Geiger. 2008. Establishment of introgression libraries in hybrid rye (*Secale cereale* L.) from an Iranian primitive accession as a new tool for rye breeding and genomics. Theor. Appl. Genet. 117:641–652. doi:10.1007/s00122-008-0808-1
Falke, K.C., P. Wilde, H. Wortmann, H.H. Geiger, and T. Miedaner. 2009. Identification of genomic regions carrying QTL for agronomic and quality traits in rye (*Secale cereale*) introgression libraries. Plant Breed. 128:615–623. doi:10.1111/j.1439-0523.2009.01644.x
Genz, A., F. Bretz, T. Miwa, X. Mi, F. Leisch, F. Sheipl, and T. Hothorn. 2011. mvtnorm: Multivariate normal and t distributions. R package version 0.9–96. URL: http://CRAN.R-project.org/package = mvtnorm (accessed 14 June 2012).
Gur, A., and D. Zamir. 2004. Unused natural variation can lift yield barriers in plant breeding. PLoS Biol. 2(10):e245. doi:10.1371/journal.pbio.0020245
Holm, S. 1979. A simple sequentially rejective multiple test procedure. Scand. J. Stat. 6:65–70.
Koumproglou, R., T.M. Wilkes, P. Townson, X.Y. Wang, J. Beynon, H.S. Pooni, H.J. Newbury, and M.J. Kearsey. 2002. STAIRS: A new genetic resource for functional genomic studies of *Arabidopsis*. Plant J. 31(3):355–364. doi:10.1046/j.1365-313X.2002.01353.x
Leonova, I.N., L.I. Laikova, O.M. Popova, O. Unger, A. Börner, and M.S. Röder. 2007. Detection of quantitative trait loci for leaf rust resistance in wheat—*T.timopheevii*/*T. tauschii* introgression lines. Euphytica 155:79–86. doi:10.1007/s10681-006-9303-4
Marwede, V., A. Schierholt, C. Möllers, and H.C. Becker. 2004. Genotype × environment interactions and heritability of tocopherol contents in canola. Crop Sci. 44:728–731. doi:10.2135/cropsci2004.0728
Marwede, V., M.K. Gül, H.C. Becker, and W. Ecke. 2005. Mapping of QTL controlling tocopherol content in winter oilseed rape. Plant Breed. 124:20–26. doi:10.1111/j.1439-0523.2004.01050.x
Maurer, H.P., A.E. Melchinger, and M. Frisch. 2008. Population genetic simulation and data analysis with Plabsoft. Euphytica 161:133–139. doi:10.1007/s10681-007-9493-4
McCouch, S. 2004. Diversifying selection in plant breeding. PLoS Biol. 2(10):e347. doi:10.1371/journal.pbio.0020347
Nevo, E., and G. Chen. 2010. Drought and salt tolerances in wild relatives for wheat and barley improvement. Plant Cell Environ. 33:670–685. doi:10.1111/j.1365-3040.2009.02107.x
R Delevopment Core Team. 2011. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. http://www.R-project.org/ (accessed 14 June 2012).
Rousseaux, M.C., C.M. Jones, D. Adams, R. Chetelat, A. Bennett, and A. Powell. 2005. QTL analysis of fruit antioxidants in tomato using *Lycopersicon pennellii* introgression lines. Theor. Appl. Genet. 111:1396–1408. doi:10.1007/s00122-005-0071-7
Schmalenbach, I., N. Körber, and K. Pillen. 2008. Selecting a set of wild barley introgression lines and verification of QTL effects for resistance to powdery mildew and leaf rust. Theor. Appl. Genet. 117:1093–1106. doi:10.1007/s00122-008-0847-7
Schmalenbach, I., J. Léon, and K. Pillen. 2009. Identification and verification of QTLs for agronomic traits using wild barley introgression lines. Theor. Appl. Genet. 118:483–497. doi:10.1007/s00122-008-0915-z
Schmalenbach, I., and K. Pillen. 2009. Detection and verification of malting quality QTLs using wild barley introgression lines. Theor. Appl. Genet. 118:1411–1427. doi:10.1007/s00122-009-0991-8
Szalma, S.J., B.M. Hostert, J.R. LeDeaux, C.W. Stuber, and J.B. Holland. 2007. QTL mapping with near-isogenic lines in maize. Theor. Appl. Genet. 114:1211–1228. doi:10.1007/s00122-007-0512-6
Uzunova, M., W. Ecke, K. Weißleder, and G. Röbbelen. 1995. Mapping the genome of rapeseed (*Brassica napus* L.). I. Construction of an RFLP linkage map and localization of QTL for seed glucosinolate content. Theor. Appl. Genet. 90:194–204. doi:10.1007/BF00222202
Wang, J., X. Wan, H. Li, W.H. Pfeiffer, J. Crouch, and J. Wan. 2007. Application of identified QTL-marker associations in rice quality improvement through a design-breeding approach. Theor. Appl. Genet. 115:87–100. doi:10.1007/s00122-007-0545-x
Wang, J., X. Wan, J. Crossa, J. Crouch, J. Weng, H. Zhai, and J. Wan. 2006. QTL mapping of grain length in rice (*Oryza sativa* L.) using chromosome segment substitution lines. Genet. Res. Camb. 88:93–104. doi:10.1017/S0016672306008408
Yang, W.-J., A. Nadolska-Orczyk, K.V. Wood, D.T. Hahn, P.J. Rich, A.J. Wood, H. Saneoka, G.S. Premachandra, C.C. Bonham, J.C. Rhodes, R.J. Joly, Y. Samaras, P.B. Goldsbrough, and D. Rhodes. 1995. Near-isogenic lines of maize differing for glycinebetaine. Plant Physiol. 107:621–630. doi:10.1104/pp.107.2.621