THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Combined detection and introgression of quantitative trait loci underlying desirable traits

OPEN ACCESS

Combined detection and introgression of quantitative trait loci underlying desirable traits

M. H. Yazdi, A. K. Sonesson, J. A. Woolliams and T. H. E. Meuwissen

The online version of this article, along with updated information and services, is located on the World Wide Web at:

http://www.journalofanimalscience.org/content/86/5/1089

American Society of Animal Science

www.asas.org

Combined detection and introgression of quantitative trait loci underlying desirable traits

M. H. Yazdi,[*][1] A. K. Sonesson,[†] J. A. Woolliams,[*][‡] and T. H. E. Meuwissen[*]

*Department of Animal & Aquacultural Sciences, Norwegian University of Life Sciences, N-1432 Ås, Norway;
†AKVAFORSK, PO Box 5010, 1432 Ås, Norway; and ‡Roslin Institute, Roslin, Midlothian, EH25 9PS, UK

ABSTRACT: This study presents a new method that combines QTL mapping and gene introgression. The effectiveness of this method for simultaneous detection and introgression of a desirable QTL from a donor line into a recipient line was evaluated by simulation. For evaluation, we used the fourth backcross generation of 2 inbred lines. The difference between the 2 lines for the trait of interest was described entirely by 1 QTL, with the donor line carrying the superior allele. Nine scenarios, combinations of 3 heritabilities (h^2 = 0.10, 0.05, or 0.01) and 3 population sizes (N = 100, 500, or 1,000) were considered in the simulation. Selection of parents for the next backcross was based solely upon the estimated probability of carrying the superior allele after a QTL analysis. Estimates of the QTL location and allele substitution effect in most scenarios were comparable to the true values. However (with either small h^2 or N) the QTL allele substitution effect was underestimated, and location was also biased. The SE of the estimates decreased with increasing N. The retained donor chromosome segment and linkage drag were close to the expected values from other published work. In general, combined detection and introgression of genes underlying desirable traits not only saves at least 1 generation, but also it ensures that the desirable QTL is introgressed where its function is simultaneously tested in a planned environment and recipient genome structure.

Key words: backcross design, inbred line, introgression, quantitative trait loci mapping

INTRODUCTION

To date, the most commonly used approach to introgress genes from a donor line into a recipient line is through an introgression program. In the most simple introgression program, only phenotypic differences are known between the 2 lines. For example, the donor line is known to be very high-performing for 1 specific trait, whereas the recipient line is poorly performing for this trait, although the trait is desired for the recipient line. In addition to the phenotypic information, developments in molecular marker technologies have created a great opportunity for geneticists and breeders to utilize the genetic marker information in QTL mapping and gene introgression programs to accelerate the reduction of the remaining donor segment, which is a critical issue in introgression schemes (Tanksley and Rick, 1980; Stam and Zeven, 1981; Tanksley et al., 1981; Soller and Beckmann, 1983; Wall et al., 2005).

Introgression of a detected QTL using genetic markers has been successful in practice (Jefferies et al., 2003; Koudandé et al., 2005). In all of these studies, QTL identification proceeded introgression, and the timescale of improvement was the time taken for QTL detection plus the time taken for introgression, limiting the practicality of the approach.

One approach to shorten the time needed for introgression is to combine the 2 steps, QTL identification and introgression, into a single step. This would combine the strengths of fine-mapping and backcrossing and pave the way for introgression of desirable but unknown QTL into recipient lines in animals and plants simultaneously.

The objective of this study is to present a method that performs QTL mapping and introgression of desirable genes from a donor line into a recipient line simultaneously. The effectiveness of this method will be investigated through simulation considering 2 inbred lines.

MATERIALS AND METHODS

Animal Care and Use Committee approval was not obtained for this study because no animals were used.

1089

Base Populations

Two inbred divergent lines (recipient and donor) were simulated using Monte Carlo simulation. The genome structure of individuals was diploid, with size 100 cM in 1 chromosome (carrier chromosome). The genome was assumed to have 1 QTL affecting the trait of interest and 101 markers. The markers were positioned at equal distances on the chromosome. Each locus, either QTL or marker, was assumed to be biallelic, with additive gene effects for the QTL and no effects for the markers. The 2 lines were fixed for the alternative alleles at each locus, which is consistent with the assumption of inbred lines. The assumption of additivity for the QTL is not essential but simplifies the description, and this will be addressed in the discussion section. The genotypes for the recipient and donor lines were qq and QQ for the QTL and m_1m_1, $\cdots$, $m_{101}m_{101}$ and M_1M_1, $\cdots$, $M_{101}M_{101}$ for the markers, respectively. In this study, the QTL was placed at 76.5 cM from the beginning of the chromosome. This position was chosen because it was away from the center and ends of the chromosome and not located at a marker position. Location at the center was avoided, because this would be the mean location of false positive QTL. The ends of the chromosome were avoided, because it would result in truncated likelihood peaks, which are unsatisfactory for assessing the procedures proposed.

Selection and Mating

The introgression proceeded by crossing the inbred lines to produce an F1 generation and then by recurrent backcrossing of the selected individuals from the crossbred population to the recipient line, to produce generations BC_1, BC_2, BC_3, and BC_4. In this study, BC_4 was the end point, and the selected parents from BC_4 were used to report the results. All generations consisted of N individuals, which varied across different scenarios.

In the F1 generation, all individuals had identical markers and QTL genotypes. Hence, selection of parents was random. In each subsequent generation, selection was based on the probability of the candidate being heterozygous for the QTL, conditional on the markers and the phenotypes observed in that generation. Individuals were selected if the probability of being heterozygous exceeded a predetermined threshold value ($P_{Selection}$). As a consequence, a variable number of candidates were selected and given an opportunity to breed. The calculation of this selection criterion will be described in the section on QTL mapping. The selected F1, BC_1, BC_2, and BC_3 individuals were considered to be nonrecurrent parents.

Mating took place at random to reproduce N offspring (1/2 N males, 1/2 N females). Each offspring was produced by random sampling with replacement of 1 sire and 1 dam from those selected individuals. In each generation, crossing-over events were generated according to the mapping function of Haldane (1919). The chromosome string for an offspring began with either the paternal or maternal chromosome sequence in the parent with equal chance and crossed over to read from the other string when a recombination occurred.

Genetic Models

The difference between the 2 lines for the trait of interest was described entirely by the 1 QTL, with the superiority of the donor line being 2α, where α = the QTL allele substitution effect. In a backcross population, where individuals are either Qq or qq with equal frequency, the segregation of 1 QTL contributes an amount $\alpha^2/4$ to the genetic variance, $\sigma_G^2 = \alpha^2/4$ when the phenotypic difference between lines is 2α (Wright, 1968).

Therefore, the phenotypic value for each individual was simulated based on the following model

$$y_i = \mu + b_i\alpha + e_{ij},$$

where y_i = the phenotypic value of the ith individual ($i = 1, N$); μ = the population mean; b_i = an indicator variable equal to the number of favorable QTL alleles, which will take values of 0 or 1 only; and e_{ij} = a random normal variable, with mean 0.0 and $\sigma_e^2 = 4.95$. Values of α used in the simulations are described in a later section.

QTL Mapping

The single-interval mapping model (Lander and Botstein, 1989) was applied for QTL mapping. In this model, 1 marker interval at a time was used to construct a putative QTL likelihood at the midpoint location of the interval. By using all marker information, denoted by M_i, and the phenotypic value y_i of the recorded trait, the likelihood function for a putative QTL in an interval with midpoint x_i in the backcross program was

$$L(\mu,\alpha,\sigma_e^2) = \prod[\phi_i(y_i,\mu,\sigma_e^2)\pi(qq|M_i)$$
$$+ \phi_i(y_i,\mu + \alpha,\sigma_e^2)\pi(Qq|M_i)],$$

where $\phi_i(y_i,\mu,\sigma_e^2)$ and $\phi_i(y_i,\mu + \alpha,\sigma_e^2)$ = the density functions for a normal distribution, with mean μ and $\mu + \alpha$ and variance σ_e^2, and $\pi(qq|M_i)$ and $\pi(Qq|M_i)$ = the probability of the QTL genotypes conditional on marker genotypes and position of the flanking markers.

The probability of the QTL genotype, $\pi(qq|M_i)$ or $\pi(Qq|M_i)$, was calculated based on the marker genotype of the individual and its nonrecurrent parent at flanking markers in each interval. Nonrecurrent parents were assumed to be heterozygous irrespective of the marker genotypes given that they were selected as heterozygous parents. If a marker locus of the nonrecurrent parent was homozygous (noninformative), then the interval was expanded until the next heterozygous locus. For example, if the genotypes at 4 marker loci of

Table 1. Estimates of the parameters[1] from fourth backcross generation, where the probability of selected animals to be heterozygous at QTL location was $\text{Pr}_i(Qq|y_i, M_i) \geq 0.75$

α^2	Population size[3]	Location[4]	$\hat{\alpha}$[5]	σ_e^2	Donor genome, cM	Linkage drag, cM	Obligate drag, cM	Efficiency of selection	Selected individuals, n
1.4832	100	76.3 ± 0.5	1.3547 ± 0.0238	4.96 ± 0.04	44.67 ± 0.52	40.76 ± 0.54	3.91 ± 0.18	0.88 ± 0.01	50 ± 0
	500	77.1 ± 0.1	1.4792 ± 0.0114	4.95 ± 0.01	44.28 ± 0.24	39.83 ± 0.25	2.49 ± 0.05	0.98 ± 0.00	248 ± 1
	1,000	77.0 ± 0.1	1.4789 ± 0.0075	4.94 ± 0.01	44.47 ± 0.18	39.7 ± 0.18	2.33 ± 0.05	0.99 ± 0.00	496 ± 2
1.0208	100	77.6 ± 0.8	0.9554 ± 0.0299	4.92 ± 0.03	44.17 ± 0.54	41.31 ± 0.62	3.94 ± 0.17	0.81 ± 0.02	49 ± 1
	500	76.9 ± 0.1	1.0155 ± 0.0106	4.98 ± 0.02	44.45 ± 0.23	40.21 ± 0.23	2.58 ± 0.05	0.96 ± 0.01	247 ± 1
	1,000	77.1 ± 0.1	1.0082 ± 0.0076	4.94 ± 0.01	44.54 ± 0.15	39.92 ± 0.15	2.57 ± 0.05	0.98 ± 0.00	495 ± 2
0.4472	100	65.2 ± 2.6	0.4642 ± 0.0347	4.84 ± 0.03	44.52 ± 0.58	42.25 ± 0.75	6.48 ± 1.42	0.58 ± 0.02	48 ± 1
	500	75.0 ± 1.3	0.4175 ± 0.0109	4.92 ± 0.01	44.72 ± 0.35	42.75 ± 0.44	2.54 ± 0.06	0.78 ± 0.02	244 ± 2
	1,000	76.3 ± 0.6	0.4253 ± 0.0076	4.96 ± 0.01	44.40 ± 0.19	41.43 ± 0.28	2.39 ± 0.05	0.86 ± 0.01	495 ± 2

[1]Mean ± SE.
[2]True QTL allele substitution effect.
[3]Population size (N) per generation.
[4]Estimated QTL location, where the true QTL location was 76.5 cM.
[5]Estimated QTL allele substitution effect.

a nonrecurrent parent were $(\cdots, M_{11}m_{11}m_{12}m_{12}m_{13} m_{13}M_{14}m_{14}, \cdots)$, then the left $(M_{11}m_{11})$ and right $(M_{14}m_{14})$ marker loci were taken as flanking markers for this expanded interval for the bracket between markers 11 and 14, which are located at 10 and 13 cM, respectively. The probability of putative QTL locations at 10.5, 11.5, and 12.5 cM was calculated with the corresponding recombination fraction of θ_1 and θ_2 for markers 11 and 14, respectively. For calculation of the probability of putative QTL, an approach for outbreeding populations was used to make extension to outbreeding populations possible. There are other possible genotypes of flanking markers associated with chromosome ends, and these are dealt with as described by Lander and Botstein (1989).

The likelihood function was maximized for each interval using the expectation-maximization algorithm, which deals with the missing genotypes (Lander and Botstein, 1989). This iterative algorithm combines both genetic marker information and phenotypic value for the calculation of the QTL genotype probabilities for each individual and maximizes the likelihood for each interval based on the estimated parameters. The interval with the greatest maximized likelihood value was taken as the estimated location of the QTL, and $\hat{\alpha}$ for this interval was taken as the estimate of the QTL allele substitution effect. For the estimation of the QTL location and effect from the second backcross generation onward, the accumulated information from the previous backcross generations was used by including the families from the previous generations. It was assumed that the selected parents were indeed carriers of the QTL, although there was a possibility that noncarrier parents had been mistakenly selected.

The selection criterion for parents of the next generation was the probability that the individual was a carrier of the donor Q allele at the estimated QTL location. This was calculated conditional on the phenotype and markers and assumed the maximum likelihood estimates for α, μ, and σ^2:

$$\text{Pr}_i(Qq \mid y_i, M_i) =$$

$$\frac{\phi(y_i; \hat{\mu} + \hat{\alpha}, \hat{\sigma}^2_{e\_\max}) \times \pi(Qq \mid M_i)}{\phi(y_i; \hat{\mu} + \hat{\alpha}, \hat{\sigma}^2_{e\_\max}) \times \pi(Qq \mid M_i) + \phi(y_i; \hat{\mu} + \hat{\sigma}^2_{e\_\max}) \times \pi(qq \mid M_i)}.$$

Individuals that were heterozygous at the estimated QTL location with probability $\text{Pr}_i(Qq|y_i, M_i) \geq P_{Selection}$ were selected. The values used in simulations are given in the next section.

Simulations, Parameters, and Summary Statistics

In this study, different values of N, σ_G^2, and $P_{selection}$ were considered. The number of candidates in each generation (N) was 100, 500, and 1,000. The genetic variance of the QTL (σ_G^2) was 0.05, 0.26, or 0.55, corresponding to values of $\alpha = 0.4472$, 1.0208, and 1.4832 and to heritabilities 0.01, 0.05, and 0.10, respectively. The selection criterion $P_{selection}$ was 0.75, 0.95, or 0.99. Schemes were simulated with all combinations of these parameters with each combination replicated 100 times.

For each replicate, the donor genome contribution, linkage, and obligate drags at backcross generations were calculated from direct examination of the marker sequence along the genome of individuals with respect to the estimated QTL location. Figure 1 illustrates these terms in the context of our backcross design, although examples are not from our empirical results. Efficiency of selection was calculated as the ratio of the number of selected individuals who were heterozygous at the true QTL location to the total number of selected individuals. The donor genome contribution is the fraction of the genome that derives from the donor genome and includes all segments of the donor genome regardless of QTL position (multiple segments result from multiple recombinations occurring in the same individuals); the linkage drag is the average length of the intact segment of the donor genome flanking the QTL; and the obligate drag is the minimum segment length of the donor ge-
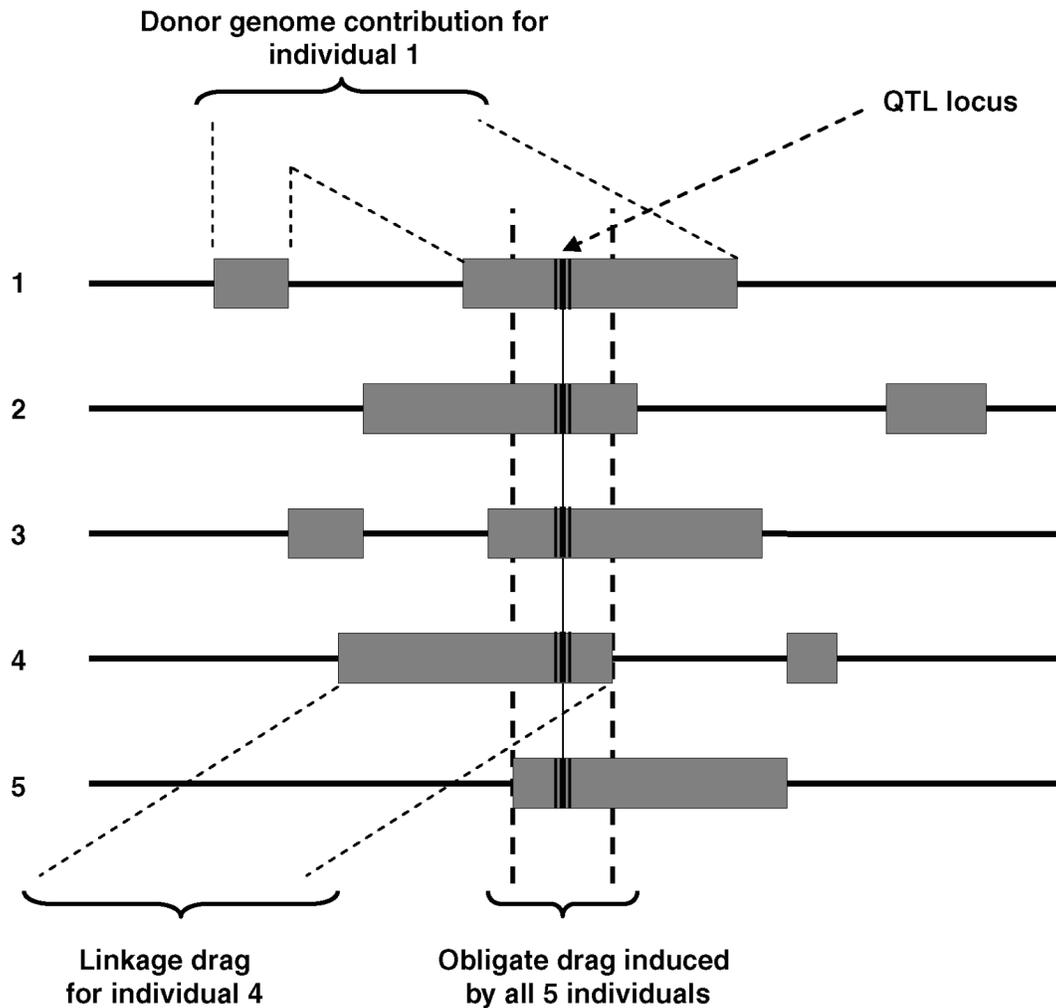
Figure 1. Diagram showing the donor genome contribution, linkage, and obligate drags from the carrier chromosomes of a group of 5 individuals. The diagram is illustrative only and does not represent real data.

nome to the left and to the right of the QTL, which represents that part of the donor genome that cannot be removed from an intercross formed from the final generation.

RESULTS

In Table 1, results are presented for 9 different scenarios, which are combinations of 3 QTL allele substitution effects (α = 0.4472, 1.0208, or 1.4832) and 3 population sizes (N = 100, 500, or 1,000) from the BC_4, in which the selection of parents in BC_1, BC_2, and BC_3 was based on $P_{selection}$ = 0.75. The estimate of the QTL location was unbiased when the QTL allele substitution effect was α = 1.4832,, and the SE of the estimate decreased with increasing population size. There was a trend toward increasing bias with smaller QTL as the population size decreased. Closer examination of these results suggests that the data were insufficient for mapping this small QTL effect for N = 100. The efficiency of selection of 0.58 (Table 1) is only marginally better than random selection (0.5 would be expected by

chance). This explains the downward bias in QTL location, because failure of detecting the true QTL would result in a randomly positioned QTL across the whole chromosome and the average location of false positive QTL over many replicates would be at 50 cM location (center). Thus, the 65.2 cM (Table 1) might be viewed as intermediate between the true finding QTL and false positive QTL over many replicates. The apparent precision of α we considered to be by chance. In general, the SE for location were inversely associated with population size.

The estimate of the QTL allele substitution effect ($\hat{\alpha}$) was unbiased when the population size and allele substitution effect were 500 and 1.0208, respectively, or more, but for the smallest population size of 100, there was a trend to slight underestimation of the substitution effect. The precision of the QTL allele substitution effect increased as the population size increased.

Estimates of the residual variance were close to the true value (σ_e^2 = 4.95) in most cases and ranged from 4.84 to 4.98. Standard errors of the estimated residual variance were very small and ranged from 0.01 to 0.04

Table 2. Estimates of the parameters[1] from fourth backcross generation, where the probability of selected animals to be heterozygous at QTL location was $\text{Pr}_i(Qq|y_i,M_i) \geq 0.95$

α^2	Population size[3]	Location[4]	$\hat{\alpha}$[5]	σ_e^2	Donor genome, cM	Linkage drag, cM	Obligate drag, cM	Efficiency of selection	Selected individuals, n
1.4832	100	77.4 ± 0.4	1.4467 ± 0.0252	4.96 ± 0.04	45.84 ± 0.49	41.39 ± 0.52	3.95 ± 0.16	0.93 ± 0.01	49 ± 1
	500	77.0 ± 0.1	1.4850 ± 0.0112	4.96 ± 0.02	44.85 ± 0.21	39.98 ± 0.21	2.79 ± 0.05	0.99 ± 0.00	247 ± 1
	1,000	77.0 ± 0.0	1.4733 ± 0.0068	4.96 ± 0.01	44.62 ± 0.17	39.79 ± 0.16	2.69 ± 0.05	0.99 ± 0.00	496 ± 2
1.0208	100	76.7 ± 0.8	1.0034 ± 0.0250	4.96 ± 0.04	47.20 ± 0.69	42.57 ± 0.70	4.20 ± 0.22	0.86 ± 0.02	47 ± 1
	500	77.0 ± 0.2	1.0219 ± 0.0094	4.94 ± 0.01	44.81 ± 0.28	40.28 ± 0.28	2.64 ± 0.05	0.96 ± 0.00	246 ± 1
	1,000	77.1 ± 0.1	1.0136 ± 0.0070	4.95 ± 0.01	44.71 ± 0.16	39.87 ± 0.16	2.58 ± 0.05	0.98 ± 0.00	495 ± 1
0.4472	100	65.0 ± 2.8	0.3657 ± 0.0425	4.86 ± 0.03	49.27 ± 0.83	43.64 ± 0.86	8.76 ± 2.03	0.68 ± 0.03	44 ± 1
	500	75.0 ± 1.3	0.4413 ± 0.0118	4.94 ± 0.01	46.99 ± 0.50	42.68 ± 0.48	2.73 ± 0.11	0.86 ± 0.01	233 ± 4
	1,000	76.5 ± 0.4	0.4322 ± 0.0079	4.96 ± 0.01	45.63 ± 0.28	41.51 ± 0.30	2.40 ± 0.05	0.91 ± 0.01	488 ± 2

[1]Mean ± SE.
[2]True QTL allele substitution effect.
[3]Population size (N) per generation.
[4]Estimated QTL location, where the true QTL location was 76.5 cM.
[5]Estimated QTL allele substitution effect.

across all scenarios. The genome contribution of the donor line after 4 backcross generations ranged from 44.17 to 44.72 cM across the different QTL allele substitution effects and population sizes. It should be noted that there was no background selection in this study. Linkage drag ranged from 39.70 to 42.75 cM across all scenarios. With the small QTL allele substitution effect ($\alpha = 0.4472$), the linkage drag across the different population sizes was greater than the corresponding value with the large QTL allele substitution effect ($\alpha = 1.4832$). The obligate drag ranged from 2.33 to 6.48 cM. The SE of the obligate drag was similar across the different QTL allele substitution effects except for the scenario with small allele substitution effect ($\alpha = 0.4472$) and small population size of 100. The efficiency of selection varied across different scenarios. The efficiency increased with increasing population sizes and the QTL allele substitution effect. The number of selected individuals was usually a little less than 50% as expected due to uncertainty about the QTL segregation. The distribution of the calculated $P_{selection}$ values was U-shaped, either close to 0 or 1. Therefore, truncation values of $P_{selection} = 0.75$ or 0.95 had little effect on the numbers of selected parents. The exception to this was when N was small, where the U-shaped distribution of $P_{selection}$ was less extreme, so, when $P_{selection}$ increased, the number of selected parents decreased slightly.

In Table 2, results are shown from BC$_4$, when $P_{selection} = 0.95$ in BC$_1$, BC$_2$, and BC$_3$. These results are consistent with the results presented in Table 1 for $P_{selection} = 0.75$, except that the efficiency of selection was greater, ranging from 0.68 to 0.99. Results for $P_{selection} = 0.99$ are not shown, because they were similar to the results $P_{selection} = 0.95$ (Table 2). However, the estimate of the QTL location was biased at very small QTL allele substitution effects ($\alpha = 0.4472$) regardless of $P_{selection}$.

The size of the total donor genome and linkage drag that remained over the generations when the QTL allele substitution effect was $\alpha = 1.4832$ is illustrated in

Figure 2. The total donor genome was ~79 cM in the BC$_1$ and decreased to ~45 cM in BC$_4$. The linkage drag also decreased from ~74 in BC$_1$ to ~40 cM in BC$_4$. Hence, as expected, the trend was decreasing across generations. Results of these parameters from the other scenarios were very similar to these results.

DISCUSSION

This study presents a new method that combines QTL mapping and gene introgression. The main advantage with this scheme is the reduced time needed compared with a conventional introgression scheme, in which QTL mapping precedes the introgression step. In a conventional scheme, the QTL mapping and estimation of QTL effects use information from a F2 generation. This assumes there was a QTL, whereas in practice, an exit strategy is needed if no significant QTL was found. Thereafter, the QTL is introgressed through a backcross design (i.e., the introgression starts the earliest in the third generation and is consequently 1 generation delayed if the F1 animals are still available and 2 generations delayed if they need to be recreated). Hence, there is a reduction in time of 1 to 2 generations before the introgression starts. This is important from an economic point of view especially for species with long generation intervals and expensive cost of rearing. The presented method also improves upon the advanced backcross QTL analysis method of Tanksley and Nelson (1996), which relies on phenotypic selection to maintain the donor QTL allele within the backcross generations (BC1, BC2, BC3, etc.), whereas the current method uses the markers to achieve this. The latter results in increased power, because the donor QTL allele was not lost by the current method despite low heritabilities, whereas there is a considerable probability of the donor QTL allele being lost by advanced backcross QTL analysis. It may be argued that despite the loss of power in the advanced backcross QTL analysis method, there are cost benefits from only phenotyping in intermediate
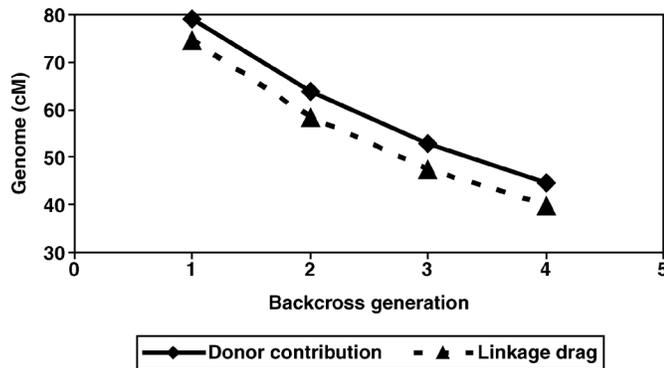
Figure 2. Trend of the estimated donor genome contribution and linkage drag across different backcross generations when $\alpha = 1.4832$ (corresponding to $h^2 = 0.10$) and $N = 1,000$.

generations. However, introgression schemes typically aim to improve traits such as disease resistance, which are difficult to measure and sometimes destructive, making phenotyping expensive.

To investigate the efficiency of retaining the donor favorable QTL allele across several backcross generations, results of BC_4 were presented. The estimates of QTL location and effect were comparable with the true values in most cases considered in this study, particularly with the larger α and increased N. Only with the small QTL allele substitution effect of $\alpha = 0.4472$, corresponding to the heritability of $h^2 = 0.01$, and small N were there indications of inadequate power with insufficient data to provide reliable estimates of QTL location.

The research emphasis on minimizing generations of backcrossing to reach specific tolerance levels for the fraction of donor genome is because the timescale for introgression tends to be long. Yet, before the backcrossing, there is the need to identify the QTL for introgression, which also takes several generations. The use of fore- and background selection (Hospital and Charcosset, 1997) or a dense marker map to get a short distance between the QTL and the marker (Hospital, 2001) accelerate the reduction of remaining donor segment. As expected, the proportional donor contribution to the backcross population decreases as the number of generation increases. The retained donor chromosome segment was close to the expected length (44.09 and 39.06 cM for total donor genome contribution and linkage drag, respectively) calculated from Eq. 1, 3, 4, and 5 of Wall et al. (2005). Hospital (2001) concluded that the most important factor for reducing the donor contribution is distances between the flanking markers and the introgressed gene, and the best way is to use flanking markers that are as close as possible to the QTL. In this study, we used a dense marker map. It is also expected that the donor contribution is more reduced when both foreground and background combined selection is carried out (Visscher et al. 1996; Hospital, 2001).

However, Han et al. (1997) stated that background selection should be applied cautiously on carrier chromosomes when the QTL location is not accurate, otherwise the QTL could be lost. In this study, there was no background selection, and the QTL was not lost in any generation or replicate, probably due to large numbers of selected parents who carried the QTL.

The method of single-interval mapping was used for gene mapping. This method is used for evaluating the association of 1 QTL with a marker interval, which ignores the effects of other segregating QTL in the mapping populations. Because there was no other QTL in this study, the estimates were unbiased, which are verified when there were sufficient genotyped individual records. Using a dense genome map for simultaneous detection and introgression of a desirable QTL or gene was found to be most instructive if the effect of QTL allele substitution and population size were 1.0208 and 500 individuals, respectively, in each generation.

Calculation of the probability of heterozygosity at the estimated QTL location for selection, $\Pr_i(Qq|y_i,M_i)$, was based on the phenotypic and marker information of individuals. Results in Tables 1 and 2 suggest that there is a discrepancy between the $\Pr_i(Qq|y_i,M_i)$ estimate and the efficiency of selection. This discrepancy probably originates from the inaccuracy of the estimate of QTL location and the assumption made in obtaining it, for example, that all parents were heterozygous. As the estimate of QTL location became more accurate, the efficiency of selection was closer to unity.

In this study, we focused on a simple case of only 1 QTL in the whole genome with additive effects as the only source of genetic variation affecting the trait of interest. The model was chosen to provide a proof of principle. If the combined detection and introgression was ineffective in this case, then it would be unlikely to be extended to outbred populations with additional polygenic variation. The inclusion of dominance effect in the model requires no further work when the effect of the QTL shows dominance, the difference between heterozygous (Qq) and homozygous (qq) individuals is $a - d$ instead of a, where a and d = the genotypic values of homozygote (qq) and heterozygote (Qq), respectively. Extension to additional polygenic variation and multiple QTL would require further investigation.

There is evidence of gene × environment interaction, as well as gene × genetic background interaction, on quantitative traits in the literature (Beavis, 1994; Beavis and Keim, 1995; Valdar et al., 2006). Dudley (1993) reported inconsistency in means of marker genotypes among mapping populations and environments. Lecomte et al. (2004) described effects of epistatic interactions between QTL regions and genetic background for fruit quality traits in the half-diallel cross design among 3 tomato lines. When combined approaches of QTL mapping and gene introgression are used, the effects of gene × recipient genome interaction will be implicitly taken into consideration in the analysis, be-

cause the QTL effects are estimated in the presence of the recipient genome.

In general, the combined detection and introgression of genes underlying desirable traits may have several advantages over separate programs of QTL mapping and gene introgression. This scheme ensures that the desirable QTL is introgressed where its function is simultaneously tested in a planned environment and expected recipient genome structure. Hence, it may be preferred, in particular, for species with long generation intervals, expensive cost of rearing, and traits that are difficult to measure. Also, this method saves at least 1 generation of time and related expenses compared with separate QTL mapping and introgression. In this study, we considered inbred lines with 1 QTL, which is mostly applicable in plant species and some animal species that we could have inbred lines, such as mice. However, further research is needed where there are multiple QTL and outbreeding populations.

LITERATURE CITED

Beavis, W. D. 1994. The power and deceit of QTL experiments. Pages 250–266 in Proc. 49th Annu. Corn Sorghum Res. Conf., Chicago, IL. Am. Seed Trade Assoc., Washington, DC.

Beavis, W. D., and P. Keim. 1995. Identification of quantitative trait loci that are affected by environment. Pages 123–149 in Genotype-by-Environment Interaction. M. S. Kang and H. Gaugh, ed. CRC Press, Boca Raton, FL.

Dudley, J. W. 1993. Molecular markers in plant improvement: Manipulation of genes affecting quantitative traits. Crop Sci. 33:660–668.

Haldane, J. B. S. 1919. The combination of linkage values and the calculation of distances between loci of linked factors. J. Genet. 8:299–309.

Han, F., I. Romagosa, S. E. Ullrich, B. L. Jones, P. M. Hayes, and D. M. Wesenberg. 1997. Molecular marker-assisted selection for malting quality traits in barley. Mol. Breed. 3:427–437.

Hospital, F. 2001. Size of donor chromosome segments surrounding introgressions and the reduction of linkage drag in marker-assisted backcross programs. Genetics 158:1363–1379.

Hospital, F., and A. Charcosset. 1997. Marker-assisted introgression of quantitative trait loci. Genetics 147:1469–1485.

Jefferies, S. P., B. J. King, A. R. Barr, P. Warner, S. J. Logue, and P. Langridge. 2003. Marker-assisted backcross introgression of the Yd2 gene conferring resistance to barley yellow dwarf virus in barley. Plant Breed. 122:52–56.

Koudandé, O. D., J. A. M. Van Arendonk, and F. Iraqi. 2005. Marker-assisted introgression of trypanotolerance QTL in mice. Mamm. Genome 16:112–119.

Lander, E. S., and D. Botstein. 1989. Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics 121:185–199.

Lecomte, L., P. Duffé, M. Buret, B. Servin, F. Hospital, and M. Causse. 2004. Marker-assisted introgression of five QTLs controlling fruit quality traits into three tomato lines revealed interactions between QTLs and genetic backgrounds. Theor. Appl. Genet. 109:658–668.

Soller, M., and J. S. Beckmann. 1983. Genetic polymorphism in varietal identification and genetic improvement. Theor. Appl. Genet. 67:25–33.

Stam, P., and A. C. Zeven. 1981. The theoretical proportion of the donor genome in near-isogenic lines of self-fertilizers bred by backcrossing. Euphytica 30:227–238.

Tanksley, S. D., H. Medina-Filho, and C. M. Rick. 1981. The effect of isozyme selection on metric characters in an interspecific backcross of tomato–basis of an early screening procedure. Theor. Appl. Genet. 60:291–296.

Tanksley, S. D., and J. C. Nelson. 1996. Advanced backcross QTL analysis: A method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. Theor. Appl. Genet. 92:191–203.

Tanksley, S. D., and C. M. Rick. 1980. Isozymic gene linkage map of the tomato: Applications in genetics and breeding. Theor. Appl. Genet. 57:161–170.

Valdar, W., L. C. Solberg, D. Gauguier, W. O. Cookson, J. N. P. Rawlins, R. Mott, and J. Flint. 2006. Genetic and environmental effects on complex traits in mice. Genetics 174:959–984.

Visscher, P. M., C. S. Haley, and R. Thompson. 1996. Marker-assisted introgression in backcross breeding programs. Genetics 144:1923–1932.

Wall, E., P. M. Visscher, F. Hospital, and J. A. Woolliams. 2005. Genomic contributions in livestock gene introgression programmes. Genet. Sel. Evol. 37:291–313.

Wright, S. 1968. Evolution and the Genetics of Populations. Volume 1. Genetic and Biometric Foundations. Univ. Chicago Press, Chicago, IL.

References

This article cites 17 articles, 5 of which you can access for free at:
http://www.journalofanimalscience.org/content/86/5/1089#BIBL