Genet. Res., Camb. (1997), **70**, *pp.* 259–265. *With* 3 *figures.* Printed in the United Kingdom © 1997 Cambridge University Press

Multitrait and multipopulation QTL search using selective genotyping

H. MURANTY[1], B. GOFFINET[2] AND F. SANTI[1]*

[1] *Station d'Amélioration des Arbres forestiers, INRA Orléans, 45160 Andon, France*

[2] *Station de Biométrie et Intelligence artificielle, INRA Toulouse, Chemin de Borde Rouge, Auzeville B.P. 27, 31326 Castanet Tolosan, France*

(*Received 11 April 1997 and in revised form 2 September 1997*)

Summary

Selective genotyping, i.e. increasing the size of the population phenotyped and genotyping only individuals from the high and low tails of the population, can considerably improve the efficiency of experiments aimed at detecting and locating quantitative trait loci (QTLs) affecting a single trait. In this paper we study how selective genotyping can increase the efficiency of multitrait QTL experiments. By selecting on an index combining the variables of interest and having the maximum correlation with each variable, the efficiency of QTL detection is increased for each trait. The efficiency of selective genotyping relative to random selection strongly depends on the correlation between the index and each variable. The optimum selection rate that minimizes costs for a given experimental power depends also on this correlation and on the genotyping costs relative to phenotyping costs. When the population segregating for the quantitative traits and the markers is not as simple as a backcross or an F_2 population, but is composed of several connected or unconnected families, selective genotyping can be used to improve the efficiency of the QTL study. In this case, the extreme individuals should be selected within each family. A method is provided to choose the selection rates within each family in order to optimize the global power of the experiment when the family sizes are unequal.

1. Introduction

Most traits on which breeders select show a continuous quantitative distribution. This distribution is thought to result from the action of a variable environment and of several genes, located at loci called quantitative trait loci (QTLs). The QTLs generally cannot be identified or mapped on the basis of quantitative variation only. Consequently, the work of a breeder traditionally consists in manipulating the QTLs collectively without knowing each QTL individually. However, QTLs can be detected and mapped when they co-segregate with marker loci. Since the advent of molecular markers, studies of the association between markers and QTLs have became increasingly numerous.

With a feasible experimental size (about 200 individuals), the obtainable power of QTL detection is often low and the QTLs that can be detected are only those explaining individually more than 10% of the variance (Strauss *et al*., 1992). Moreover, the precision of QTL location and effect estimation is quite low (Hyne *et al*., 1995). Selective genotyping (Lander & Botstein, 1989) is one of the ways to increase the power of QTL detection and to improve precision with constant means. It consists in genotyping only the individuals belonging to the high and low tails of the phenotypic distribution (the phenotypic extremes) after increasing the number of individuals phenotyped. Applying this strategy can substantially increase the power of detection of QTLs of the trait on which the individuals are selected (Lander & Botstein, 1989; Darvasi & Soller, 1992). However, most QTL studies do not deal with one trait but with several. When the traits are not correlated, the extremes for one trait are different from the extremes for another trait and so genotyping the extremes for several traits results in genotyping almost all the population. Recently, we studied the power of detection of QTLs of a trait when selection is done on another trait (Muranty & Goffinet, 1997). Here, we use these results to propose solutions for increasing the power of QTL detection by the use

* Corresponding author.

https://doi.org/10.1017/S0016672397003030 Published online by Cambridge University Press

of selective genotyping in studies focused on several traits.

Until recently the populations proposed and used to detect and map QTLs in plants were essentially based on crosses between two inbred lines, in species where inbred lines can easily be obtained, or between two partly heterozygous plants. In the former case, the QTLs detected are those at which the lines have different alleles; in the latter case, the QTLs detected are those at which the parent plants are heterozygous. Consequently, some QTLs explaining the quantitative variability in the breeding population can be missed just because they do not segregate in the particular population studied (Edwards *et al.*, 1992). A solution is to study simultaneously the offspring of several crosses, e.g. in a diallel mating design (Rebaï *et al.*, 1994; Muranty, 1996). When it seems risky to invest the high resources necessary in a QTL detection study in a population of narrow genetic basis (the cross between two genotypes), a promising strategy is to divide the resources between several connected populations (Charcosset, 1996). In this case, the use of selective genotyping can also greatly increase the power of QTL detection. However, the question of how to select the individuals to genotype or how to distribute the selection pressure among the progenies can be problematic. We propose here methods to solve these problems and illustrate the proposal with the example of a wild cherry (*Prunus avium* L.) five parent half diallel. The wild cherry population is studied at INRA-Orléans to detect QTLs of the cherry leaf spot disease, which is caused by a fungus (*Phloesporella padi*), and also to study vigour traits. This work is a part of an EC project aiming to produce a linkage map for *Prunus* species and to detect QTLs of various traits in various *Prunus* species (Arús *et al.*, 1994).

2. Detection of QTLs of several traits

(i) *Efficiency*

The variance of QTL effect estimation is one of the ways to express the quality of a QTL mapping experiment. It is a measure of QTL effect estimation precision and it is obviously related to detection power. In previous work (Muranty & Goffinet, 1997), we studied the variance of QTL effect estimation when the QTLs affecting the traits Y and Z are located on the marker considered in a backcross population. We found that it can be expressed by the following formulae:

$$\text{var } \hat{q}_Y = \frac{\sigma_Y^2}{S} \times \left\{ E_S\left[\left(\frac{Y-\mu}{\sigma_Y} \right)^2 \right] \right\}^{-1} \qquad (1)$$

for the selected trait (Y) and

$$\text{var } \hat{q}_Z = \frac{\sigma_Z^2}{S} \times \left(1 - \rho^2 \left(1 - \left\{ E_S\left[\left(\frac{Y-\mu}{\sigma_Y} \right)^2 \right] \right\}^{-1} \right) \right) \qquad (2)$$

for another trait (Z).

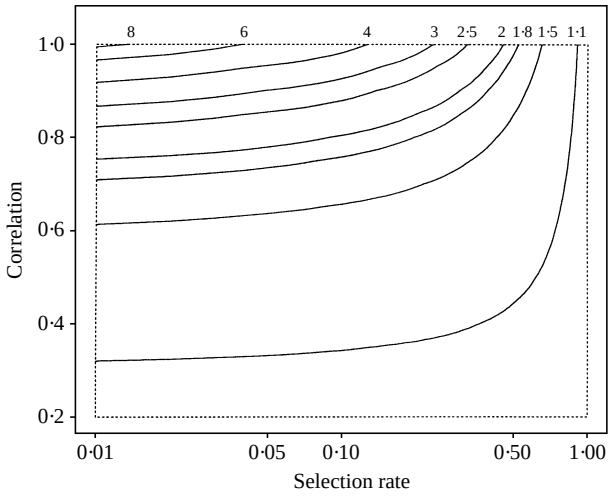

Fig. 1. Efficiency of selective genotyping relative to random selection as a function of selection pressure and correlation between the trait studied and the selected trait.

In these formulae, Y is the phenotype observation for trait Y, μ and σ_Y^2 are the phenotypic mean and variance of the selected trait, σ_Z^2 is the phenotypic variance of the other trait of interest, ρ is the phenotypic correlation between the traits, $\hat{q}_Y$ and $\hat{q}_Z$ are the estimates of the QTL effect on trait Y and Z respectively, S is the number of selected individuals and E_S means expectation for the selected individuals. From this, it is clear that the factors influencing selective genotyping efficiency are $E_S\{[(Y-\mu)/\sigma_Y]^2\}$, which is a function of the selection pressure, and the correlation between the trait for which QTLs are sought and the trait on which the individuals genotyped are selected. Moreover, it can immediately be deduced that selective genotyping never induces a loss of efficiency relative to the genotyping of S randomly chosen individuals, in which case var $\hat{q} = \sigma^2/S$; but of course, it is less efficient than genotyping the whole population.

An expression of $E_S\{[(Y-\mu)/\sigma_Y]^2\}$ as a function of selection pressure, using normal quantile and density functions, was given by Darvasi & Soller (1992) ($E_S\{[(Y-\mu)/\sigma_Y]^2\}$ is equal to what they called γ_p). So we were able to express the efficiency of selective genotyping relative to random selection as a function of the selection pressure and the correlation between traits. In practice, efficiency was calculated as the ratio of QTL effect estimation variances under selective genotyping and random selection for a given number of genotyped individuals. The results are shown in Fig. 1. For example, the efficiency of selective genotyping is more than double that of random selection, with a correlation greater than 0·8 when the selection pressure is 10%. From another point of view, the same relative efficiency is obtained with a selection pressure less than 40% when the correlation is 0·9. Very high relative efficiencies are obtained only with drastic selection pressures and for traits highly correlated with the selected trait.

https://doi.org/10.1017/S0016672397003030 Published online by Cambridge University Press

Table 1. *Relative efficiency of selective genotyping on an index I combining two traits X_1 and X_2*

		Selection rate				
ρ	corr(X_i, I)	0·05	0·10	0·25	0·50	0·75
0	0·71	1·70	1·63	1·49	1·30	1·14
0·25	0·79	2·05	1·93	1·69	1·41	1·18
0·5	0·87	2·60	2·38	1·96	1·53	1·22
0·75	0·935	3·55	3·08	2·34	1·68	1·27
0·9	0·975	4·54	3·76	2·64	1·78	1·30
1	1	5·58	4·39	2·89	1·86	1·32

ρ, correlation between X_1 and X_2.

(ii) *Selection on an index*

How could one use selective genotyping to increase the efficiency of multitrait studies? To avoid the genotyping of the whole population induced by the selection of extreme individuals for each trait, one can select the individuals to genotype on an index of the form $I = \Sigma_i p_i Y_i$, where Y_i is one of the quantitative variables of interest. Then, the correlation of each trait with the selected trait depends on the weights p_i of the different variables in the index and the correlations between the variables. Given this correlation, the coefficient of gain in efficiency of QTL detection can be calculated with formula (2) for each trait. If all traits have the same importance and, consequently, the same gain in efficiency is desired for all traits, the correlation between the index and each trait should be the same. In this case, the weights of the variables in the index should be chosen to fulfil this constraint (same correlation) and to maximize the correlation between the index and each trait. If some traits are more important than others, the weights would be chosen as a function of the interest of the traits and the power desired for them.

To illustrate the efficiency of selective genotyping in the first case, we considered the case of two normally distributed variables of equal importance, of variance 1, with a correlation between the variables denoted ρ. The maximum correlation of each variable with an index of these two variables was obtained with equal weights for the two variables. We calculated the relative efficiency of genotyping extreme individuals for the index for detection of QTLs of each trait. The results, listed in Table 1, show that even with uncorrelated variables, the correlation of each variable with the index is quite high, so that the relative efficiency of selective genotyping is greater than 1·5 for a selection rate lower than 0·25. With moderately correlated variables ($\rho = 0{\cdot}5$), the relative efficiency reaches 1·5 for a selection rate of 0·5.

(iii) *Optimal selection rate*

The optimal selection rate is the selection rate that leads to the best use of the resources; it depends on the cost of completely genotyping an individual (c_g) relative to the cost of growing and measuring an individual (c_{ph}). The total cost of an experiment is $c_g N_g + c_{ph} N_{ph}$, where N_g is the number of individuals genotyped and N_{ph} the number of individuals grown and phenotyped. Following the method of Darvasi & Soller (1992), we obtained a relative cost function $F(s)$ which has the same optimum as the total cost function for the case where all traits have the same importance. It can be written as

$$F(s) = \left[1 - \rho^2\left(1 - \left\{E_S\left[\left(\frac{I-\mu}{\sigma_I}\right)^2\right]\right\}^{-1}\right)\right]\left(\frac{c_g}{c_{ph}} + \frac{1}{s}\right),$$

where s is the selection rate, I is the phenotypic observation for the index, μ and σ_I^2 are the phenotypic mean and variance of the index, ρ is the correlation between the traits of interest and the index, and the other notations as above. The values of s that minimize $F(s)$ were obtained for a wide range of cost ratios, c_g/c_{ph}, and various values of the correlation, through a FORTRAN program using NAG routines. Fig. 2 presents the optimal selection proportion for c_g/c_{ph} varying from 0·01 to 100. It shows that in QTL mapping studies dealing with highly correlated traits so that $\rho \sim 0{\cdot}9$, the use of selective genotyping is attractive even with a cost ratio of 1, in which case the selection rate should be approximately 0·6. On the contrary, if each of the interesting traits is moderately or weakly correlated with the index, selective genotyping should be used only when the cost ratio is high (> 50).

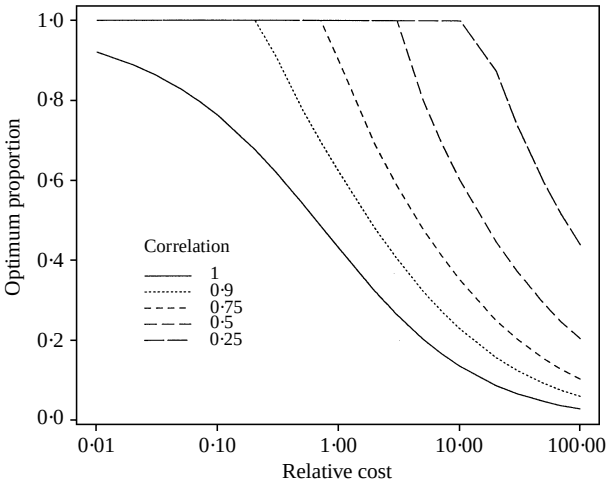

Fig. 2. Optimum selection rate as a function of the genotyping cost relative to the cost of growing and measuring an individual, and of the correlation between the variables of interest and the selection index.

3. Example of the wild cherry half-diallel

The size of the wild cherry population was approximately 600 individuals and could not be increased. The maximum power was required for the trait of main interest, cherry leaf spot susceptibility, so selection was to be based solely on this trait. However, we wanted to choose the selection rate in order to

https://doi.org/10.1017/S0016672397003030 Published online by Cambridge University Press

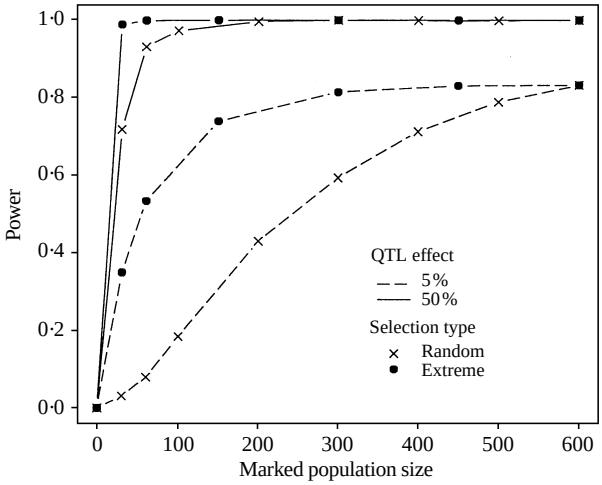

Fig. 3. Power of QTL detection in a five parent half-diallel of 600 individuals as a function of the number of individuals genotyped, for a 5% and a 50% QTL effect and with random or extreme selection.

obtain a reasonable power for the detection of QTLs of other traits, particularly vigour traits, which were moderately correlated with cherry leaf spot susceptibility. So, we calculated the power obtained with the number of individuals actually genotyped for the selected trait and for uncorrelated traits (random selection). A simplification of the wild cherry population studied at INRA-Orléans was considered: a five parent half-diallel where the parents are partly heterozygous and where the families are all the same size. In previous work (Muranty, 1996), the power of QTL detection with such a scheme was calculated as a function of significance level, QTL effect and the number of randomly chosen genotyped individuals. We used the relative efficiency of selective genotyping obtained in the backcross case to extrapolate power obtained in this case because, as is the case in a backcross population, the QTLs are detected on the basis of the differences between the two QT alleles of the parents. In practice, to calculate power under selective genotyping, the ratio of QTL effect estimation variances under selective genotyping and random selection, studied previously, was used to calculate the number of randomly chosen individuals that would give the same power as the number of extreme individuals selected among the whole population.

The results presented here were obtained with a 1% significance level; with other significance levels, the results were similar. Fig. 3 presents the power of QTL detection as a function of the number of individuals genotyped, selected among 600 individuals, for the selected trait (extreme selection) and for a trait uncorrelated with the selected trait (random selection), for QTL effects of 5 and 50%. When the number of genotyped individuals decreases, the power decreases rapidly for an uncorrelated trait, whereas it decreases slowly for the selected trait. The power obtained for the selected trait with a selection rate of 0·5 or even 0·25 is still very near the maximum power obtained with all individuals. The two curves (random vs extreme selection) are much farther from each other for a small QTL effect of 5% than for a large one of 50%. The effect of selective genotyping is particularly important for detecting small effect QTLs: with 100 individuals selected at the extremes for a trait Y among 600, a QTL with a 5% effect acting on trait Y has more than a 60% chance of being detected, but a QTL with a similar effect acting on a trait Z not correlated with Y, so that the individuals are chosen randomly relative to Z, has less than a 20% chance of being detected.

In the wild cherry population, the correlations of the two vigour traits with the trait of main interest, cherry leaf spot susceptibility, were 0·19 and −0·46. Because these correlations are quite low, we decided to genotype extreme individuals selected with a selection rate of 0·5 in order to detect QTL of quite small effects acting on the vigour traits. The resulting relative efficiencies of selective genotyping for these traits are 1·02 and 1·11, respectively, with the chosen selection rate of 0·5, whereas it is 1·86 for the selected trait.

4. Selective genotyping in a complex structured population

The previous paragraphs show that the efficiency of a QTL study is high when the number of plants measured is high (e.g. about 1000) and the number of plants genotyped, selected at the extremes, is also quite high (e.g. about 250). Generally breeders would not dare to invest so much in a population of reduced genetic basis (i.e. the offspring of only two plants), and it would probably be unwise. But such high investments are often involved in testing the breeding population. If this population originates in controlled crosses, the breeder could choose some of the families to obtain an array of connected families, genotype the phenotypic extremes to obtain a map common to all families and locate QTLs on this map (Rebaï & Goffinet, 1993; Muranty, 1996). In this case less QTLs would be missed just because they do not segregate in the population studied and the data obtained in the QTL study can be used directly in marker assisted selection (MAS). The important thing is to have a common map, so the use of dominant markers such as RAPDs or AFLPs is problematic.

The first question, then, is how to select the individuals to genotype. Obviously, the quantitative value of each individual is the result of a family effect, which can be built up in some cases into general and specific combining ability effects, a within-family genotype effect and an environmental effect. If the individuals are chosen at the extremes of the population, all individuals of some families could be chosen and no individuals of some other families would be chosen just because of the family effect. In fact, the quantitative variation that is in linkage disequilibrium

https://doi.org/10.1017/S0016672397003030 Published online by Cambridge University Press

Table 2. *Progeny sizes in the nursery trial and number of extreme individuals to genotype for various selection rates*

Family	Progeny size	Selection rate			
		0·75	0·5	0·25	0·1
143 × 171	46	36	24	12	6
143 × 221	57	42	28	14	6
143 × 226	50	38	26	14	6
143 × 229	50	40	28	16	6
171 × 221	93	66	42	18	8
171 × 226	75	56	36	18	8
171 × 229	45	34	24	12	6
221 × 226	80	58	36	18	6
221 × 229	66	50	34	16	6
226 × 229	40	30	22	12	4
Total size	602	450	300	150	62
Precision	0·0212	0·0214	0·0238	0·0292	0·0472
Relative precision	1	0·99	0·93	0·73	0·45

with the markers because the individuals originate in controlled crosses is essentially the within-family genotypic variation. Consequently, to apply selective genotyping, the individuals should be chosen at the extremes within each family.

The second question is how to distribute the selection pressure among the families. In an ideal situation, all families have the same size and then the same selection rate should be used for all. But it often happens that, for various reasons, the family sizes are unequal. This is the case in the wild-cherry diallel studied at INRA-Orléans (Table 2). To determine the selection rates to use in each family to obtain the highest global efficiency, we looked in detail at the influence of the family size on efficiency. In a half-diallel mating design involving the cross of partly heterozygous parents, without selective genotyping, the estimation variance of the QTL effect of a parent i (q_i), which is a measure of its estimation precision, is proportional to the inverse of the number of offspring of this parent (i.e. the half-sib family size): this can be written as var $\hat{q}_i \propto \{\Sigma_{j \neq i} n_{ij}\}^{-1}$ where n_{ij} is the size of the family originating in the cross of the parents i and j. As a consequence, the global precision (if the QTL effects of all parents have the same size) is proportional to the sum of the inverses of the half-sib family sizes, which can be written as global precision $\propto \Sigma_i \{\Sigma_{j \neq i} n_{ij}\}^{-1}$. As previously, we used the coefficient of gain in efficiency obtained in the backcross case to calculate the equivalent family size n^*_{ij}, i.e. the number of randomly chosen individuals that would give the same power, as a function of the selection rate and the total family size. Using these equivalent family sizes, we calculated equivalent half-sib family sizes and global precision. We wrote a FORTRAN program using NAG routines to obtain the within-family selection rates that give the best global precision when the total number of individuals to genotype is limited. The rates obtained were used to choose the numbers of extreme individuals that should be genotyped in each family for various global selection rates (Table 2); the global precision was calculated as $\Sigma_i \{\Sigma_{j \neq i} n^*_{ij}\}^{-1}$, which is in fact proportional to global precision. We also calculated the global precision relative to the precision obtained when the whole population is genotyped. It can be seen that with a selection rate of 0·25 the relative precision is 0·73, and with a selection rate of 0·5 it is 0·93, which is quite high. It is worth noting that in two families of equal size (143 × 226 and 143 × 229) the numbers of individuals to retain are often different. This is because the half-sib families they belong to have different sizes.

In conclusion, when the family sizes are unequal, the way to determine the selection pressures to apply in each family to obtain the highest global efficiency is to: calculate the global precision without selective genotyping as a function of family sizes, assuming QTL effects of equal size across the families; calculate the equivalent family size as a function of selection rate in selective genotyping; optimize selection rates relative to global precision while fixing a global selection pressure, by using the formulae obtained in the first two steps in a computer program.

5. Discussion

A great deal of information can be derived from QTL studies: for example, the number of genes controlling a quantitative trait, the most important ones for (example the so-called key-enzymes in a metabolic pathway), their distribution in the genome (i.e. dispersal vs clustering) and the origin of genetic correlation between traits (i.e. pleiotropy vs linkage). However, the confidence one will have in the conclusions obtained depends on the global precision of the experiment. This precision will also determine the efficiency of a MAS scheme initiated with a QTL study.

https://doi.org/10.1017/S0016672397003030 Published online by Cambridge University Press

We have shown in this paper that selective genotyping can greatly increase the power of detection of QTLs affecting the trait on which extreme individuals are selected and highly correlated traits, at the expense of an increase in the size of the population phenotypically studied. The precision of QTL effect estimation is increased in the same way, because precision and power are strongly related. We verified with simulations that the precision of QTL location is also increased (Muranty & Goffinet, 1997). Statistical methods have been proposed to detect epistatic QTLs (Haley & Knott, 1992), to detect several QTLs in the same linkage group (Jansen & Stam, 1994; Zeng, 1994) and to test pleiotropy versus linkage of genes affecting correlated traits (Jiang & Zeng, 1995). However, the power of these methods when the effects to be detected are small or when the QTLs are tightly linked would probably not be high enough to answer the questions investigated with the means usually employed (B. Goffinet & B. Mangin, personal communication). This power would probably also be increased by the use of selective genotyping. Thus selective genotyping seems to be the best way to increase the general efficiency of QTL mapping studies, with constant means.

In order to increase the efficiency of multitrait studies by the use of selective genotyping we proposed selecting the individuals to genotype on an index combining the variables of interest. It should always be possible to build up an index that has the same correlation with all variables of interest. As a consequence the increase in power would be the same for all variables. When the traits of interest are highly correlated, and not too numerous, this strategy could be very efficient. On the other hand, when the traits of interest are only moderately or not correlated, the maximum correlation cannot be very high: for example, with two uncorrelated traits having the same variance of 1, the maximum correlation, obtained with equal weights, is $\sqrt{0{\cdot}5} \cong 0{\cdot}71$; with such a correlation the coefficient of gain in efficiency is always less than 1·8 whatever the selection rate. Consequently, to obtain a high power of QTL detection, the number of individuals genotyped should remain quite high.

Sometimes, a higher power or a higher increase of power is desired for one specific trait among the traits studied – for example the trait of main interest in a breeding programme or the less heritable trait, for which the QTL detection power is the lowest because the environmental variance, i.e. the variance that cannot be explained by QTLs, represents a large part of the total phenotypic variance. In this case the individuals could be chosen as the extremes for this trait or for an index that has a higher correlation with this trait than with the other traits. Our results give a basis on which to choose the weights in the index in order to obtain the desired power for the different traits.

The optimal proportion of individuals to select and even the decision to use the selective genotyping strategy, depends on the costs of genotyping an individual relative to the costs of growing and measuring an individual. The optimal proportion was studied considering the selection of the genotyped individuals on an index that has the same correlation with all variables of interest. When a high correlation can be obtained because the traits are highly correlated and not numerous, the optimal selection to apply can be quite drastic, but not too much so as to avoid selection of phenotypic outliers only. On the other hand, if the variables are moderately correlated leading to a correlation of 0·5 between each variable and the index, selective genotyping should be applied only when the cost ratio is quite high (> 5). If the correlation that can be obtained between each variable and the index is lower than 0·25, selective genotyping should be applied only when the cost ratio is very high.

When the purpose of a QTL study is to bring knowledge about the nature of genetic variation of a trait, it is generally sensible to split the trait up into several fine components, which can be difficult to measure. For example, Causse *et al.* (1995) studied the determinism of carbon metabolism traits during early growth in maize with the aid of molecular markers and measured several enzyme activities of each plant. In this case the cost ratio could be less than 10, or even less than 1, depending on the species and on the components measured. Then, selective genotyping should be used only when the components are highly correlated. On the other hand, when the QTL study is the initial part of a MAS scheme, the individuals would be grown and measured anyway because they belong to a breeding population. The cost ratio would consequently be quite high. In this case selective genotyping should be used even with weakly correlated traits.

QTL mapping populations will probably increasingly have a complex structure and so we studied how to apply selective genotyping in such populations. Indeed, a breeding population generally has a large genetic basis and a complex structure. Moreover, to obtain results that have a general meaning in a QTL mapping study, the population should result from crosses between several parents. We considered a particular five parent half-diallel population and suggested a method of choosing the distribution of the selection rate among the families in any complex structured population. We calculated an equivalent family size as a function of selection rate and real family size, using the coefficient of gain in precision obtained in a backcross population. This should be done in each case where, as in a backcross population, the QTLs are detected on the basis of the difference between the two QT alleles of the parents. If the population was composed of F_2 families, the coefficient of gain could be different because the QTLs are

https://doi.org/10.1017/S0016672397003030 Published online by Cambridge University Press

detected on the basis of additivity and dominance. We also suggested that QTL effects of equal size across the families should be assumed: this is because it is not known in advance in which family QTL effects are the largest or the smallest.

6. Conclusion

As long as completely genotyping an individual will be more expensive than growing and measuring it, selective genotyping will be one of the ways to increase the efficiency of a QTL mapping study with constant resources. It should be applied even in multitrait studies, provided the number of traits of interest is not too high or they are not too weakly correlated, and even when the population has a complex structure. Solutions to estimate the QTLs effects correctly have been proposed in the case of a backcross population (Muranty & Goffinet, 1997) and a similar method was used to obtain correct estimators in other types of populations; these solutions will be implemented in QTL detection software developed at INRA-Toulouse (Rebaï, 1996). Another way to increase the efficiency of a QTL mapping study is to replicate the individuals, when possible (clones, doubled haploid or recombinant inbred lines, etc), or to measure their offspring, as in a grand-daughter design (Weller *et al*., 1990) or as F_3 families (Cowen, 1988). However, the consequence of replication is to reduce the environmental variance, so this strategy is only of interest for low heritable traits. The question of how to simultaneously use selective genotyping and replication optimally needs further attention because the interest of both depends on the genotyping cost relative to the phenotyping cost.

References

Arús, P., Messeguer, R., Viruel, M., Tobutt, K., Dirlewanger, E., Santi, F., Quarta, R. & Ritter, E. (1994). The European *Prunus* mapping project. *Euphytica* **77**, 97–100.

Causse, M., Rocher, J. P., Henry, A. M., Charcosset, A., Prioul, J. L. & de Vienne, D. (1995). Genetic dissection of the relationship between carbon metabolism and early growth in maize, with emphasis on key-enzyme loci. *Molecular Breeding* **1**, 259–272.

Charcosset, A. (1996). L'identification de locus affectant des caractères quantitatifs (QTL) à l'aide de marqueurs génétiques est-elle justifiée pour la sélection? *Le Sélectionneur Français* **46**, 35–45.

Cowen, N. M. (1988). The use of replicated progenies in marker-based mapping of QTLs. *Theoretical and Applied Genetics* **75**, 857–862.

Darvasi, A. & Soller, M. (1992). Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. *Theoretical and Applied Genetics* **85**, 353–359.

Edward, M. D., Helentjaris, T., Wright, S. & Stuber, C. W. (1992). Molecular marker facilitated investigations of quantitative trait loci in maize. IV. Analysis based on genome saturation with isozyme and restriction fragment length polymorphism markers. *Theoretical and Applied Genetics* **83**, 765–774.

Haley, C. S. & Knott, S. A. (1992). A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity* **59**, 315–324.

Hyne, V., Kearsey, M. J., Pike, D. J. & Snape, J. W. (1995). QTL analysis: unreliability and bias in estimation procedures. *Molecular Breeding* **1**, 273–282.

Jansen, R. C. & Stam, P. (1994). High resolution mapping of quantitative traits into multiple loci via interval mapping. *Genetics* **136**, 1447–1455.

Jiang, C. & Zeng, Z. B. (1995). Multiple trait analysis of genetic mapping for quantitative trait loci. *Genetics* **140**, 1111–1127.

Lander, E. S. & Botstein, D. (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**, 185–199.

Muranty, H. (1996). Power of tests for quantitative trait loci detection using full-sib families in different schemes. *Heredity* **76**, 156–165.

Muranty, H. & Goffinet, B. (1997). Selective genotyping for the location and the estimation of the effect of a quantitative trait locus. *Biometrics* **53**, 629–643.

NAG: Numerical Algorithms Group (1995). *The NAG FORTRAN LIBRARY Manual*, *Mark* 16. Oxford: NAG Ltd.

Rebaï, A. (1996). Rapport de spécification du logiciel de cartographie de locus à effects quantitatifs (QTL) chez ples plantes: MCQTL. Technical report, Institut National de la Recherche Agronomique, Tolouse, Unité de Biométrie et Intelligence Artificielle.

Rebaï, A. & Goffinet, B. (1993). Power of test for QTL detection using replicated progenies derived from a diallel cross. *Theoretical and Applied Genetics* **86**, 1014–1022.

Rebaï, A., Goffinet, B. & Mangin, B. (1994). Detecting QTLs with diallel schemes. Eucarpia Meeting, Wageningen, July 1994.

Strauss, S. H., Lande, R. & Namkoong, G. (1992). Limitations of molecular-marked-aided selection in forest tree breeding. *Canadian Journal of Forest Research* **22**, 1050–1061.

Weller, J. I., Kashi, Y. & Soller, M. (1990). Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle. *Journal of Dairy Science* **73**, 2525–2537.

Zeng, Z. B. (1994). Precision mapping of quantitative trait loci. *Genetics* **136**, 1457–1468.