Animal Genetics, 1997, **28**, 188–194

Marker-assisted introgression using non-unique marker alleles II: selection on probability of presence of the introgressed allele

A M van Heelsum, C S Haley, P M Visscher

Summary

If marker alleles that identify a gene for introgression are not completely unique to the different base populations, the trait allele can be lost quickly during the process of backcrossing. This study considers ways to deal with incompletely informative markers in order to retain the desired allele. Selection was based on the probability of the presence of the desired (introgressed) trait allele, which was calculated for each marker genotype, using a single marker or a diallelic or triallelic marker bracket. The percentage of individuals retaining the introgressed allele was calculated over five generations of backcrossing, for selected fractions between 0 and 1, for marker alleles that could occur in both base populations. The best results were obtained with a rather large selected fraction, when all individuals, heterozygous and homozygous for the most desirable allele at the marker loci, were selected. Additional selection against marker homozygotes (which might have the highest probability of carrying the desired-trait allele, but produce uninformative gametes) altered the optimum selected fraction, making the selected fraction more consistently inversely related to a better retention of the desired-trait allele. A marker bracket was found to give a better retention of the desired-trait allele than a single marker and triallelic markers were better than diallelic markers, giving a retention of almost 50%. The earlier that preselection of parents (on informativeness) took place the better the overall result; preselection should occur preferably in the base populations. Preselection could make marker alleles unique to alternative base populations and markers would effectively become fully informative. Selection in the base populations might not be possible or not desirable, for example, because of the available number of individuals. This is unlikely to be a problem when parents are paired up to exclude any common marker alleles.

A M van Heelsum
C S Haley
Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS, UK
P M Visscher
Institute of Ecology and Resource Management, University of Edinburgh, Edinburgh EH9 3JG, UK

Correspondence: A M van Heelsum.

Accepted 17 February 1997

Keywords: animal breeding, marker-assisted introgression, outbred populations, quantitative trait loci

Introduction

To carry out effective introgression of a valuable gene from a donor population into a recipient population the different alleles of this gene should be directly identifiable. Currently, much work is directed at finding the position of functional genes (e.g. in pigs: Ellegren *et al.* 1994; Rohrer *et al.* 1994; Archibald *et al.* 1995) but, often, direct identification cannot be achieved and markers have to be used to identify the alleles of interest. Markers can also be used to determine the source (donor or recipient population) of the background genotype.

A number of authors have looked into the possible benefits of MAI (marker-assisted introgression). Most authors assumed complete identification of the introgressed allele and only used markers to identify the source of the rest of the genome (Gama *et al.* 1992; Hospital *et al.* 1992; Hillel *et al.* 1993). The markers they used were fully informative. Groen & Smith (1995) touched upon the problem of markers that were not completely informative by simulating a situation where a marker allele, which is fixed in the recipient population, also occurs in the donor population at a low frequency. However, there was still an allele, unique to the donor population, which could be used to reliably identify the introgressed trait allele. Visscher *et al.* (1996) also considered identification of the introgressed gene using markers. They looked at a QTL (quantitative trait locus) to be introgressed at a known as well as at an estimated position, where selection was on nearby markers. They concluded that, as long as selection was on marker haplotypes covering the likely position of the QTL, the introgressed allele would not be lost. Marker alleles were unique to alternative base populations.

In Van Heelsum *et al.* (1997) the approach of Visscher *et al.* (1996) has been extended to non-unique markers for identification of the

introgressed major gene and the background genotype. It was concluded that using markers with alleles that are not completely unique to the donor population (to identify the desirable trait allele), as if they were unique, can be very ineffective. They suggested that distinguishing between all available marker genotypes, rather than selecting all individuals with one or two desired marker alleles per locus, could be a better option. In this paper this distinction is made. The probability of presence of the trait allele of interest was calculated for all different marker genotypes, and ranking and selection took place accordingly. The aim of this work was to discover if this approach gained a better retention of the desired trait allele when marker alleles were not unique to the base populations. The efficiency of this criterion in a backcrossing scheme was assessed for a single marker and a marker bracket and for diallelic and triallelic markers.

Van Heelsum *et al.* (1997) also looked at selection against marker homozygotes because they produce gametes with identical marker genotypes, which makes identification of the desired major gene allele impossible. This approach slightly improved the results, so in this paper a similar approach was investigated, but in combination with selection on probability, to study if the retention of the introgressed trait allele could be further improved. This paper focused solely on the effectiveness of the introgression itself (during the backcrossing phase), and ignored selection to recover the recipient (background) genotype more quickly. This seems a reasonable starting point because a proper identification of the introgressed allele has to be achieved first, before one can think about more complicated, although perhaps more realistic introgression programmes. Furthermore, Van Heelsum *et al.* (1997) showed that the background genotype does not influence the introgression result when no selection on background genotype takes place.

Models and methods

Introgresson

Introgression of the trait allele takes place by backcrossing first-cross (F_1) individuals (the product of crossing a donor and a recipient line or breed) to the recipient population, producing five subsequent backcross generations. The intercrossing phase, which would normally follow the backcrossing phase of an introgression programme, was not investigated in this study. The introgressed gene was assumed to be a gene with a major influence on a valuable trait, which is not phenotypically identifiable at the moment of selection (e.g. disease resistance, litter size). The major gene was assumed to have two alleles, T1 and T2: the donor population was fixed for allele T1 and the recipient population for allele T2. The population sizes were assumed to be infinite. From the starting frequencies of marker and major gene alleles in the base populations, the recombination rate and a selection strategy, the frequencies of all possible genotypes in the following crossbred populations were calculated.

Selection

Selection started in the first backcross (BC_1). Selection decisions were based on the individual marker genotypes; neither the major-gene genotypes nor the parental genotypes were assumed to be known and so the linkage phase between two markers was also unknown. This situation is similar to the one described by Van Heelsum *et al.* (1997), where all genotypes were selected that had one or two marker alleles thought to be linked with the desirable trait allele, so the selected fraction was not a parameter.

In this study, selection was on probability of carrying the desired-trait allele. The probability was calculated as the expected fraction of individuals having a copy of the desired-trait allele within a group of individuals with the same marker genotype. Individuals with the highest probabilities of carrying the desired-trait allele were selected; these were of one or more genotypes depending on their frequencies in the unselected population and the selected fraction. Marker genotypes were ranked according to the probability and the next genotype in rank was not selected until all animals with the higher scoring genotype were used. If several genotypes ranked the same, equal proportions of these genotypes were selected.

From BC_1 onwards, the selected fraction predicted which genotypes were selected, which alleles were produced, how high genotype frequencies were and which marker genotypes had the highest probability of having T1 in the next generation. The probabilities declined over generations and the ranking of the marker genotypes according to their probabilities of carrying the desired-trait allele did not necessarily stay the same.

Three situations were studied: selection using a single diallelic marker, selection using a diallelic marker bracket and selection using a triallelic marker bracket.

Single marker. The recombination rate r between the marker and the major gene was set

to 0. The allele frequencies for allele 1 were 0·9 and 0·1 (and therefore 0·1 and 0·9 for allele 2) in the donor and recipient populations respectively. Selection on probability was compared with selection in two steps: first, marker genotypes were ordered according to marker heterozygosity (individuals heterozygous on a marker locus were selected in preference to homozygotes); and second, further ordering, based on probabilities, took place within heterozygosity class. Van Heelsum *et al.* (1997) showed that by adding selection against marker homozygotes a somewhat better retention of the introgressed allele could be achieved. A backcross individual that is homozygous for the thought-to-be favourable marker allele will produce two types of gametes that cannot be distinguished: one type will contain the favourable T1-allele but the other type will have the unfavourable T2-allele. Tracing the marker alleles to the correct base population will only be possible in individuals heterozygous for the marker.

Marker bracket. When a marker bracket was used, selection was always on heterozygosity and probability. The major gene was placed 4 centimorgans (cM) from marker A and 6 cM from marker B, giving recombination rates r_A and r_B between major gene and marker A and B, respectively, of 0·0384 and 0·0565, assuming Haldane's mapping function. Starting frequencies of marker allele 1 in donor and recipient populations respectively were 1·0/0·0 (completely informative markers), 0·9/0·1, 0·7/0·3, 0·5/0·5 (completely uninformative markers), 0·2/1·0 (comparable with Groen & Smith 1995) and 1·0/0·2 for both markers.

Triallelic markers. Using triallelic markers the frequencies for A1, A2, A3, B1, B2 and B3 were 0·7, 0·2, 0·1, 0·1, 0·2 and 0·7 in the donor population and 0·1, 0·3, 0·6, 0·6, 0·3 and 0·1 in the recipient population. Selection was on heterozygosity and probability.

Results

Single diallelic marker

Figure 1, A & B, shows the percentage of animals with the desired-trait allele over five generations of backcrossing, where selection is on probability of presence of the introgressed allele (Fig. 1A) or on heterozygosity as well as probability (Fig. 1B), using a single diallelic marker, for selected fractions between 0·001 and 1·000. The striking features of Fig. 1A are the peak in the percentage of animals with the desired-trait allele, showing an optimum selected fraction of 0·57 (in BC_5), and the steep drops for only small changes in selected fraction (for a selected fraction just over 0·05 in BC_4 and 0·08 in BC_5). Both graphs show horizontal lines for the most intense selection.

These observations can be explained as follows. For one marker with two alleles three different marker genotypes occur: A1A1, A1A2 and A2A2, with frequencies in BC_1 of 0·05, 0·50 and 0·45, respectively. The probability of the presence of the T1-allele is always highest for A1A1 individuals (in BC_1) because they always have one A1-allele that comes from the crossbred parent, whereas in heterozygotes there is always a chance that allele A1 stems from the recipient parent. However, in later generations for high selection intensities, probabilities can sometimes be equal for A1A1 and A1A2 individuals. For all selected fractions smaller than 0·05 the same genotypes will be selected in each generation (A1A1 only in BC_1, BC_3 and BC_5; both A1A1 and A1A2 in BC_2 and BC_3), resulting in the horizontal lines (equal percentages of animals with the desired-trait allele) shown in both Fig. 1A and Fig. 1B.

Increasing the selected fraction above 0·05 implies the selection of A1A2 individuals as well, which gives improved retention of the introgressed allele over more than two backcross generations, until all A1A1 and A1A2 individuals are selected. The optimum selected fraction of 0·57 over five backcross generations is somewhat higher than 0·55 (the frequency of A1A1 and A1A2 in BC_1), because the frequencies have increased owing to selection, and will increase slightly more over further generations of backcrossing.

If selection is based also on heterozygosity the priority of the A1A1 and A1A2 genotypes is reversed and horizontal lines occur for selected fractions up to 0·50. A higher selected fraction means the additional selection of A1A1 individuals, which have a higher probability of having T1 in BC_1 than A1A2 individuals, resulting in a slightly better retention of the introgressed allele in BC_2. However, A1A1 animals produce uninformative (indistinguishable) gametes, so the percentage of animals with T1 will halve from their offspring to grand offspring. Heterozygotes for the marker do produce distinguishable gametes, accounting for the better retention of T1 using selection on heterozygosity, shown in Fig. 1B.

The drop in percentage of animals with the desired-trait allele for a selected fraction just over 0·05 (BC_4) and 0·08 (BC_5), when selection is on probability only (Fig. 1a), is caused by the addition of another genotype to the selected group in

the great-grandparent generation, which alters the probabilities in the grandparent generation. For example, when in BC_1, apart from A1A1 individuals (with the highest probability), a very small fraction of A1A2 individuals is also selected, A1A1 and A1A2 individuals in BC_2 no longer have equal probabilities. A1A1 individuals score slightly higher, so only they will be selected instead of equal fractions of both genotypes (and A1A2 has a much higher frequency in the unselected population). This results in uninformative offspring (BC_3), and the percentage of animals with the desired-trait allele will halve in the grand offspring (BC_4) instead of dropping slightly.

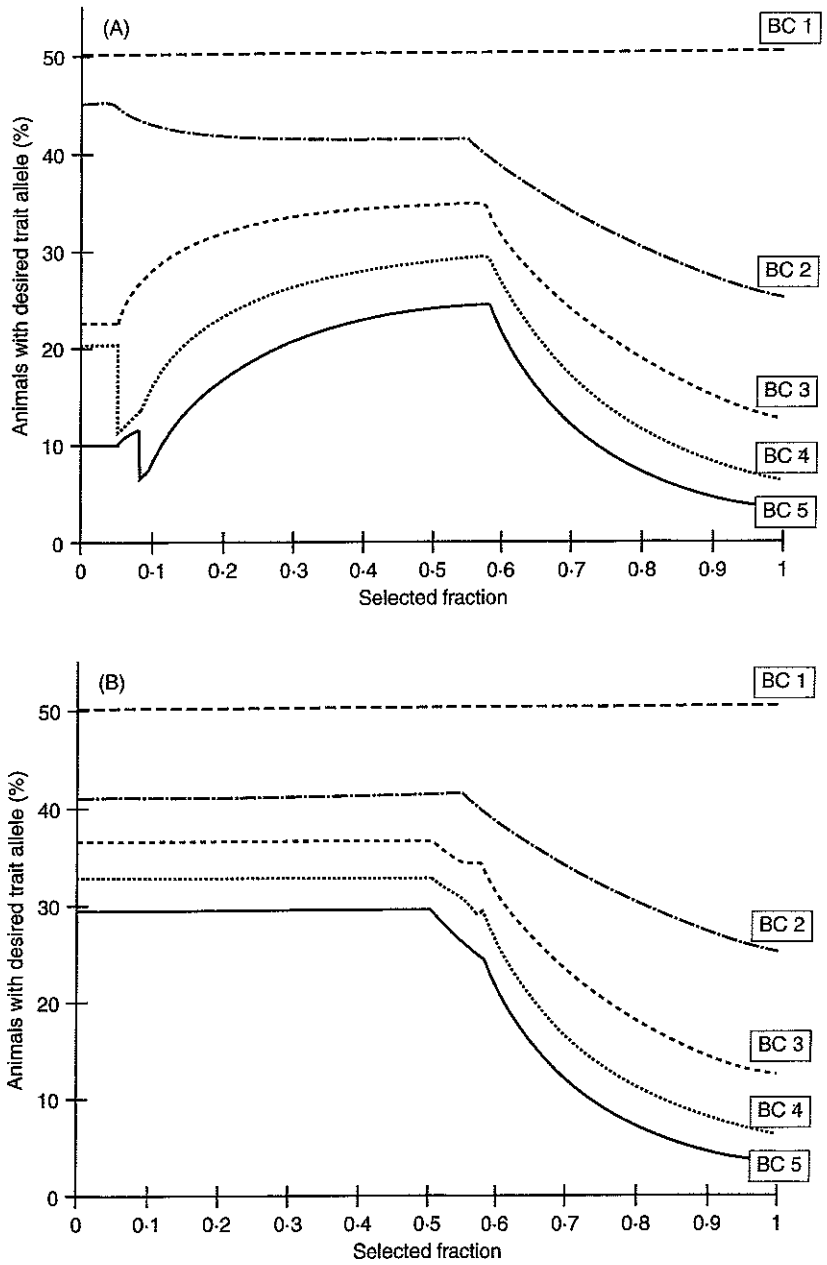

Fig. 1. Percentage of animals with the desired-trait allele, over five generations of backcrossing, for selected fractions between 0·001 and 1·000. Selection on the trait allele was carried out by using one diallelic marker with starting frequencies of alleles in donor and recipient populations of 0·9 and 0·1, and 0·1 and 0·9, respectively. There was assumed to be no recombination between marker and major gene. **A**: selection on probability of presence of the desired trait allele. **B**: selection on heterozygosity on the marker locus (starting in the first backcross generation) and probability of presence of the desired-trait allele.

Diallelic marker bracket

Using a marker bracket rather than a single marker further increases the effectiveness of the introgression programme. The percentage of individuals with the desired trait allele in BC_5 for six different starting frequencies and selection on heterozygosity and probability is shown in Fig. 2. Fully informative markers (1·0/0·0) give a retention of almost 50%, as long as only doubly heterozygous individuals are selected. For starting frequencies of 0·9/0·1 the decline starts at a lower selected fraction and even with intense selection only 45% of the individuals still possess the valuable trait allele in BC_5. Completely uninformative markers (0·5/0·5) always give the same low result, whatever the selected fraction. Having marker alleles unique to the donor population (0·2/1·0, so alleles A2 and B2 only occur in the donor with frequencies of 0·8) gives a result equal to that of fully informative markers (for selected fractions smaller than 0·4). Conversely, for starting frequencies of 1·0 and 0·2, results are much poorer.

Two markers with three alleles per marker

Figure 3 shows the results of using two triallelic markers for selection on heterozygosity on the marker loci and probability of presence of the desired-trait allele. Figure 3 shows an improvement over diallelic markers of about 1% at BC_5, although heterozygosity in the F_1 is comparable with that of diallelic markers (81% for two alleles, 82% for three alleles; the same for both markers).

In Fig. 3 the horizontal stretch for the most intense selection is a lot shorter than in Fig. 2 (starting frequencies 0·9/0·1) or Fig. 1b because the frequency of the genotype with the highest probability is much lower (as there are many more different genotypes present). Selection seems very effective: for some selected fractions there is even an improvement of the retention of the desired trait allele over generations (lines cross over). The point where the line makes a sharp downwards turn shifts to the right over generations, reflecting an increasing number of doubly heterozygous marker genotypes with high probability. The number of double heterozygotes with low probability decreases because the next sharp turning point (upwards this time) shifts to the left over generations. This

point marks the start of adding single heterozygotes (first with very high probability) to the selected group. The high probabilities make the line climb again, before it drops owing to addition of genotypes with (very) low probabilities. In BC_2 there is a slight revival for a selected fraction just over 0·88. At that point addition of doubly homozygous marker genotypes starts, beginning with A1A1B3B3, which has the very highest probability of all marker genotypes in BC_1.

Discussion

If marker alleles are not unique to the alternative base populations, the source of alleles in crossbred individuals cannot always be identified. One can think of two ways to deal with this problem. First, by making full use of the limited information available, retention of the introgressed trait allele can be improved compared with using the markers as if they were fully informative. In this study we propose the 'probability approach' to do this. Second, marker alleles can be made unique to the base populations by selection. This approach will be discussed below. Presumably, retention of the introgressed trait allele is essential for an introgression programme because the reason for such a programme would be to fix the allele of great importance in the newly formed synthetic population. This allele might not have a clear, additive economic value, but might give the synthetic a unique trait that fills a niche in the market, or might make a difference between life and death of an individual (e.g. disease-resistance genes). Therefore, eventually one would like to end up with individuals not just carrying the desired-trait allele, but being homozygous for that allele.

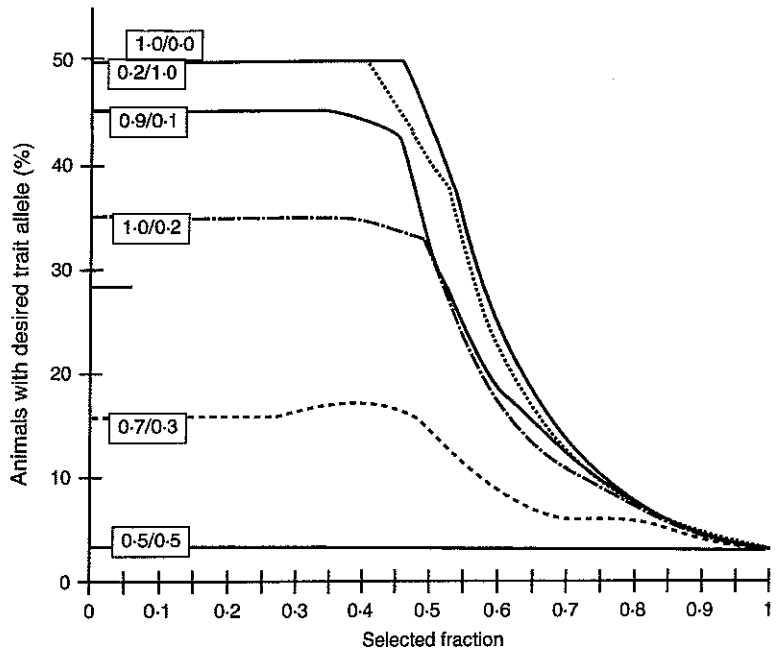

Fig. 2. Percentage of animals with the desired-trait allele in backcross generation five for selected fractions between 0·001 and 1·000. Selection on the trait allele was carried out by using two diallelic markers on 4 and 6 cM from the major gene. Starting frequencies of allele 1 for both markers, in donor and recipient populations, respectively, were 1·0/0·0, 0·9/0·1, 0·7/0·3, 0·5/0·5, 0·2/1·0 and 1·0/0·2. Selection was on heterozygosity on the marker loci (starting in the first backcross generation) and probability of presence of the desired trait allele.

The probability approach, combined with selection against homozygotes, was shown to be a fairly efficient method for retaining the desired-trait allele over several generations of backcrossing, particularly if a marker bracket was used that consisted of multiallelic markers. Achieving 50% of the animals with the introgressed allele over several generations of backcrossing is not possible because of recombination.

Retention of the introgressed allele could be further improved by using fewer generations of backcrossing than the five generations used here, but the recovery of the background genotype would be lower (a problem ignored in this study). According to Hospital *et al.* (1992), usually at least four generations of backcrossing would be required to recover the background genotype with use of markers. Van Heelsum *et al.* (1997) showed that selection on background genotype can severely impede the retention of the desired trait allele. Although, for the same starting frequencies, the results in this study are better than any of the results in the earlier paper, this is by no means a guarantee that the severe loss of the introgressed allele after the first generation (with added selection on background genotype on the same chromosome) will not occur.

If the same selected fractions are considered, results from Van Heelsum *et al.* (1997), for optimal selected fractions, are almost identical to the results of this study. In the simplest case (using a single, diallelic marker) small differences are caused by the slight variations in selected fraction (0·55 in BC_1 and 0·59 in later generations in the work of Van Heelsum *et al.* (1997); 0·57 in this study) and a difference in recombination rate (0·0099 in the work of Van Heelsum *et al.* (1997); 0 in this study). Heterozygosity/probability selection gives always better results when the selected fraction is decreased, whereas Van Heelsum *et al.* (1997) found that more intense selection implied random selection within the group of animals with one or two copies of the marker allele. This will not improve results and is potentially risky because, by chance, one might select only homozygotes and lose the introgressed allele more quickly than expected.

Triallelic markers give a better retention of the introgressed allele than diallelic markers because triallelic markers allow for a greater number of different genotypes, which enables more discrimination between individuals. More markers will not necessarily give better retention, because a marker outside a marker bracket (with the introgressed gene known to be within the marker bracket) does not add any information about the introgressed gene. An increased number of markers could be useful for decreasing the length of the donor genome around a QTL (Hospital *et al.* 1992), or when the exact position of the QTL is not known (Visscher *et al.* 1996). More markers could also be useful if informativeness is not required of all markers; e.g. if the closest marker is not informative but the next marker is, then this one can be used for selection. This might be particularly useful if allele frequencies are similar in the alternative base populations, and when numbers of animals are severely limiting.

The other way to deal with non-unique markers was preselection of parents on uniqueness of marker alleles. This will be most effective when carried out in the base populations, so individuals selected from the first population have no alleles in common with individuals selected from the second population. For diallelic markers the starting frequencies are then, effectively, 1·0/0·0. Preselecting donor individuals might not always be possible because the individuals are no longer available. The next best solution might be to preselect F_1 individuals on heterozygosity (although F_1 animals would normally not be typed because all are known to have one copy of the desired major-gene allele), and/or to preselect only recipient individuals on unique alleles.

Selecting recipient (or donor) animals on marker genotype might not be desirable because it can severely restrict the number of available individuals. Even if sufficient animals were available, selection on background genotype will be less intense, so the crossbred population will lag even more behind the selected recipient nucleus for traits not influenced by the introgressed allele (Gama *et al.* 1992).

However, the effect on numbers available can be restricted by pairing-up parents in a manner that excludes common marker alleles, rather than the whole population not having any alleles in common. For a diallelic marker this would mean that both mates need to be homozygous for the alternative alleles to give only informative offspring; if one of the parents is heterozygous, 50% of the offspring will be homozygous and thus uninformative. If populations have to be unique, and, for example, a single diallelic marker is used for which allele 1 has starting frequencies of 0·9 and 0·1 in the donor and recipient population, respectively, 81% of the individuals of both populations could be used (if 100% informative offspring is required); 99% and 81% of the respective populations could be used if the result of 91% informative (heterozygous) offspring is sufficient. If only pairs have to be unique, 82% of both populations could be used. The difference is much greater if starting frequencies are 0·5/0·5: only 25% of both populations could be used for 100% informative offspring when populations need to be unique, against 50% of both populations when uniqueness is only required within mating pairs.

When the number of markers that need to be unique increases, numbers will be more limiting. On the other hand, if only informativeness is required from one of the available markers, and the pairing up is carried out carefully, it will nearly always be possible to produce a high percentage of informative offspring.

More alleles per marker will widen the choice for selection only on pair level. Rohrer *et al.* (1994) report that microsatellite markers in pigs have, on average, 5·8 alleles, so a problem with finding a suitable mate might not easily occur in a practical situation. They also give values for

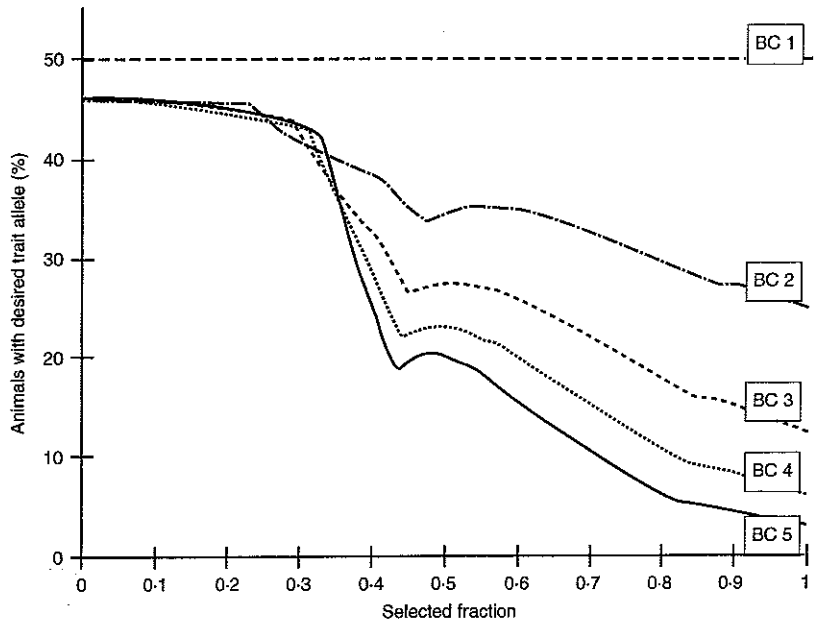

Fig. 3. Percentage of animals with the desired-trait allele, over five generations of backcrossing, for selected fractions between 0·001 and 1·000. Selection on the trait allele was carried out by using two triallelic markers on 4 and 6 cM from the major gene. Starting frequencies of alleles of marker A were 0·7, 0·2, 0·1 and 0·1, 0·3, 0·6; of marker B they were 0·1, 0·2, 0·7 and 0·6, 0·3, 0·1 in donor and recipient populations, respectively. Selection was on heterozygosity on the marker loci (starting in the first backcross generation) and probability of presence of the desired-trait allele.

the average heterozygosity levels for microsatellite markers in several crosses between pig breeds. For a cross between White Composite boars and Chinese sows they report an average heterozygosity of 81·4%, which is close to the values assumed in this study.

If preselection of (base) parents is not possible or less desirable, the probability approach seems to offer a good alternative. Although the retention of the desired trait allele is good if selection is purely on the introgressed allele, the reliability has not been proven for application in combination with selection on background genotype. It is not straightforward to calculate the probabilities for larger, and with that more diverse, genotypes by using the current method, but this could be subject of further study. Additionally, it would be interesting to look at an introgressed gene that has a known, quantifiable effect, rather than a major gene with a practically infinite value. In combination with the background genotype, one could determine the size of the major gene effect needed to make an introgression programme feasible, and to enable a comparison between introgression and other breeding schemes.

The backcrossing phase of an introgression programme would normally be followed by an intercrossing phase, which is not investigated in this study. The aim of the intercrossing phase is to fix the introgressed allele in the synthetic population, therefore producing animals homozygous for the desired allele. Groen & Smith (1995) found that with a marker allele unique to the donor population and selection on genomic similarity, 100% of the animals in the second intercross generation had at least one copy of the desired-trait allele. Although the informativeness of the markers they used will be reduced by recombination, identification of homozygotes for the desired trait allele was good enough to quickly fix the desired allele in the synthetic population. Presumably, in this study, the problems in identifying the desired trait allele occurring in the backcrossing phase, when there is no marker allele unique to the donor population, will be extended to the intercrossing phase. The unreliable identification of the trait alleles also affects the distinction between homozygotes and heterozygotes for the trait allele, making complete fixation more difficult.

A point of further research could be the potential use of phenotypic information, gathered in the process of backcrossing. Although phenotypic information has a number of disadvantages compared with marker information (often only available later in life, sex dependent, influenced by environment, etc.), it could be used to update the estimates of major gene size and position in order to make selection on the major gene, by using markers, more effective in following generations.

Acknowledgements

We wish to thank Bill Hill for helpful comments. This research was supported by the Marker Assisted Selection Consortium of the UK pig industry (Cotswold Pig Development Company Ltd, J.S.R. Farms Ltd, Pig Improvement Company/National Pig Development Company, Newsham Hybrid Pigs Ltd, Rattlerow Farms Ltd. and the Meat and Livestock Commission) and by the Ministry of Agriculture, Fisheries and Food (MAFF), the Department of Trade and Industry (DTI) and the Biotechnology and Biological Sciences Research Council (BBSRC).

References

Archibald A.L., Haley C.S., Brown J.F. *et al.* (1995) The PiGMaP consortium linkage map of the pig (Sus scrofa). *Mammalian Genome* **6**, 157–75.

Ellegren H., Chowdhary B.P., Johansson M. *et al.* (1994) A primary linkage map of the porcine genome reveals a low rate of genetic recombination. *Genetics* **137**, 1089–100.

Gama L.T., Smith C. & Gibson J.P. (1992) Transgene effects, introgression strategies and testing schemes in pigs. *Animal Production* **54**, 427–40.

Groen A.F. & Smith C. (1995) A stochastic simulation study on the efficiency of marker-assisted introgression in livestock. *Journal of Animal Breeding and Genetics* **112**, 161–70.

Hillel J., Schaap T., Haberfield A. *et al.* (1990) DNA fingerprints applied to gene introgression in breeding programs. *Genetics* **124**, 783–9.

Hospital F., Chevalet D. & Mulsant P. (1992) Using markers in gene introgression breeding programs. *Genetics* **132**, 1199–210.

Rohrer G.A., Alexander L.J., Keele J.W., Smith T.P. & Beattie C.W. (1994) A microsatellite linkage map of the porcine genome. *Genetics* **136**, 231–45.

Van Heelsum A.M., Visscher P.M. & Haley C.S. (1997) Marker assisted introgression using non-unique marker alleles I: selection on presence of linked marker alleles. *Animal Genetics* **28**, 181–7.

Visscher P.M., Haley C.S. & Thompson R. (1996) Marker assisted introgression in backcross breeding programs. *Genetics* **144**, 1921–30.