STRATEGIES TO IMPROVE BALI CATTLE IN EASTERN INDONESIA

STRATEGIES TO IMPROVE BALI CATTLE IN EASTERN INDONESIA ACIAR

ACIAR PROCEEDINGS No. 110

ACIAR
The Australian Centre for International Agricultural Research

www.aciar.gov.au

Molecular Genetics and Their Place in Breeding Systems

Julius van der Werf and Brian Kinghorn[1]

Abstract

Current estimation of breeding value is based on phenotypic information only. The advent of molecular markers allows determination of actual genotype at specific gene loci, without error due to random and non-random environmental effects. In the ideal situation we can directly identify genotypes at loci containing genes with substantial effects on quantitative traits (QTLs).[2] However, a more likely scenario is where we use genetic markers linked to QTLs to increase the probability of selecting the animals with the desired QTL genotype. With such indirect markers there is no guarantee of QTL genotype. Furthermore, the marker alleles linked to the preferred QTL allele can be different in different families, and information about linkage phase needs to be accumulated based on phenotypic and pedigree information (e.g. a progeny test). Marker-assisted selection is most useful for traits that are hard to measure and have low heritability. Selection of animals based on (most probable) QTL genotype will allow earlier and more accurate selection, increasing short- and medium-term selection response by up to 40%, and may aid in targeting genotypes for specific production environments or markets.

Introduction

DURING the past two decades most livestock industries have successfully developed estimated breeding values (EBVs) to allow identification of the best breeding animals. EBVs are best calculated using BLUP[3], meaning that they are based on pedigree and performance information of several traits from the individual animal and its relatives. BLUP EBVs are the most accurate criteria to identify genetically superior animals based on phenotypic performance recording.

Although the idea of genetic selection is to improve the genes in our breeding animals, we actually never really observe those genes. Selection is based on the final effect of all genes working together, resulting in the performance traits that we observe on production animals. This strategy makes sense, since we select based on what we actually want to improve. However, animal performance is affected not only by genes but also by other factors that we do not control. Selection for the best genes on animal performance alone can never reach 100% accuracy. A large progeny test comes close to such a figure of perfect selection, but this is expensive for some traits (e.g. those related to meat quality), and we have to wait several years before the benefits from a progeny test have an effect. Efficient breeding programs are characterised by selecting animals at a young age, leading to short generation intervals and faster genetic improvement per year. For selecting at younger ages, knowledge about the existence of potentially very good genes could be very helpful.

Selection of Genotypes based on Genetic Markers

The idea behind marker-assisted selection is that there may be genes with significant effects that can be targeted specifically in selection. Some traits are controlled by a single gene (e.g. hair colour) but most traits of economic importance are quantitative traits that are most likely controlled by a fairly large number of genes. However, some of these genes might have a larger effect. Such genes can be called

[1] University of New England, Armidale, NSW 2351, Australia

[2] Quantitative genetics uses phenotypic information to help identify animals with good genes. Extension to use information from techniques of molecular genetics aims to locate and exploit gene loci which have a major effect on quantitative traits (hence QTL — quantitative trait loci).

[3] Best Linear Unbiassed Prediction model

Strategies to Improve Bali Cattle in Eastern Indonesia
edited by K. Entwistle and D.R. Lindsay
ACIAR Proceedings No. 110
(printed version published in 2003)

major genes located at QTL. Although the term QTL strictly applies to genes of any effect, in practice it refers only to major genes, as only these will be large enough to be detected and mapped on the genome. Following the pattern of inheritance QTL might assist in selection.

Figure 1 illustrates that QTL constitute only some of the many genes that affect phenotype. The other relevant genes are termed polygenes. Variation at the polygenes jointly with polymorphism at the QTL determines total genetic variation. Although QTL effects explain only a part of genetic differences between animals, knowledge of the genes located at QTL could greatly assist in estimating an animal's true genotype. Information available at QTL therefore adds to accuracy of estimation of breeding value. If genetic effects at QTL are really large, such genes could be more specifically exploited in breeding programs, e.g. to target specific production circumstances or specific markets.

Figure 1 suggests that the value or the allelic forms at individual QTL could be known. In practice, this is rarely the case. That is, currently there are few examples where QTL effects can be directly determined, but knowledge in this area is rapidly developing. Most QTL known today can be targeted only by genetic marker — 'landmarks' at the genome that can be chosen for their proximity to QTL. We cannot actually observe inheritance at the QTL itself, but we observe inheritance at the marker, which is close to the QTL.

When making selection decisions based on marker genotypes, it is important to know what information can be inferred from those genotypes. Figure 2

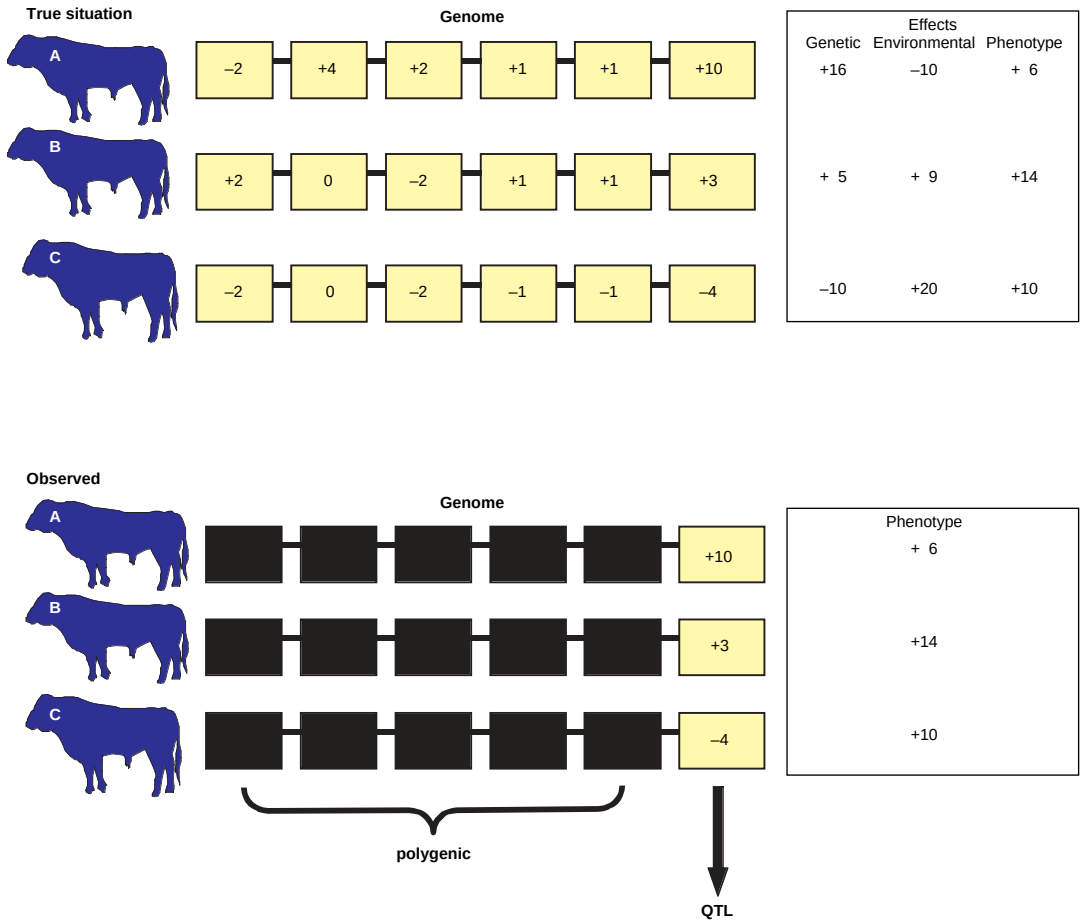

Figure 1. Illustration of three bulls with different phenotypes. The top diagram gives the true allelic values at the different genes affecting body weight; the bottom diagram illustrates what would be observed if QTL could be identified in addition to phenotype, adding significant information about true genotype.

Strategies to Improve Bali Cattle in Eastern Indonesia
edited by K. Entwistle and D.R. Lindsay
ACIAR Proceedings No. 110
(printed version published in 2003)

shows the principle of inheritance of a marker and a linked QTL. We can identify the marker genotype (Mm) but not the QTL genotype (Qq). The last is really what we want to know because of its effect on economically important traits.

Let the Q allele have a positive effect, therefore being the preferred allele. In the example, the M marker allele is linked to the Q in the sire. Progeny that receive the M allele from the sire have a high chance of having also received the Q allele, and are therefore the preferred candidates in selection.

As shown in Figure 2 there are four types of progeny. All progeny will inherit m alleles and q alleles from the mother. The sire will provide them either an M allele or an m allele and either Q or q. In the figure, 90% of the progeny that receive an M allele have also received a Q allele, because M and Q alleles are linked on the same chromosome in the sire. However, in 10% of the cases after the sire reproduced, there has been a recombination between the two loci, and animals that inherited an M allele from their father have received a q allele rather than a Q allele. Therefore, marker alleles do not always provide certainty about the genotype at the linked QTL.

Selection of animals could be based on genetic information only. However, in that case the effect of other genes not covered by the marker would be ignored. Optimal selection should aim for QTL as well as for polygenes. It should be based on information from marker genotypes combined with information on animals' phenotype. The first aims to get the good QTL, the second aims at getting good 'other genes'. Selection with the aid of information at genetic markers is termed marker-assisted selection (MAS).

How Important is the Marker Information?

The question now arises: what value should be given to a particular marker genotype? Should we prefer an animal having a desired marker genotype over one that doesn't, but that performs better phenotypically? The answer is that the information from these different sources should be weighted optimally, as in BLUP. The value of having the right marker genotype (Mm in our example) depends on three things:

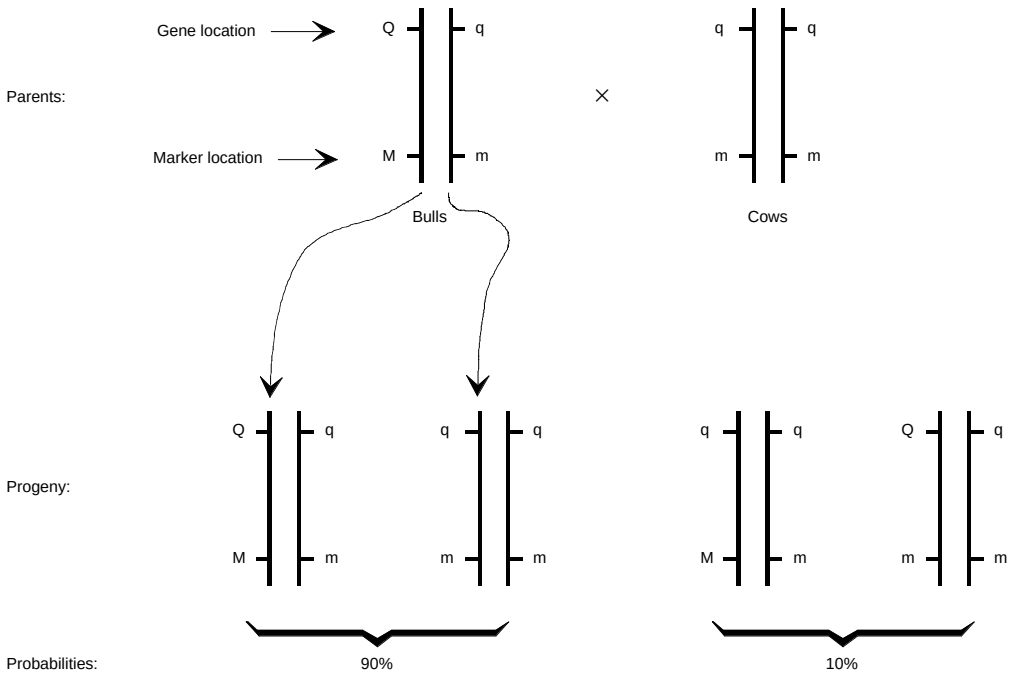

Figure 2. Following the inheritance of a major genotype affecting a quantitative trait (Q locus) with a genetic marker (M locus) closely linked to the Q locus. The sire is heterozygous for either locus and the dam is homozygous. For this example, we can determine for each progeny whether it received M or m allele from the sire. The recombination rate (10%) determines how often Q alleles join M alleles.

Strategies to Improve Bali Cattle in Eastern Indonesia
edited by K. Entwistle and D.R. Lindsay
ACIAR Proceedings No. 110
(printed version published in 2003)

- the size of the QTL effect;
- the frequency of the Q allele. If it is nearly fixed MAS will not add a lot to genetic improvement, whereas the opposite will be true if it is found in low frequency;
- the probability that an Mm animal has indeed inherited the Q allele.

With regard to the probability of how sure we can be that an *M* animal indeed has a Q allele, there is a distinction between direct markers and linked markers. If there is no recombination between marker and QTL, i.e. if the marker exactly identifies the gene, then finding an M implies finding a Q. However, if M is only near Q on the genome, M and Q have a possibility to break up at each meiosis. (Meiosis is the process that produces sperm or eggs and determines the genetic make-up of offspring.) This has the important consequence that finding an M is not a guarantee of finding a Q. Not only is the chance of having a Q allele with a frequency lower than 1, but for a randomly chosen sire we have no idea whether M or m is linked to Q. The implication is that for each family this linkage phase needs to be determined based on data. For example, in the family in Figure 2 we would see that M progeny perform better than m progeny, hence M would most probably be linked to Q (the positive allele). But some progeny will have m linked to Q, so in their offspring m will be the preferred marker genotype.

Direct markers

The easy scenario is when the marker allele M and the QTL are always together. This is the case only if the marker is actually measuring the relevant polymorphism within the gene that causes the effect. Such a direct marker is very convenient, because the marker genotype will directly inform us about the QTL genotype. However, there are currently only a few direct genetic markers for economically important traits. Examples are:

- the Halothane gene in pigs (giving increased proportion of lean meat but also stress susceptibility). This gene has been found — it is the ryanodine receptor gene;
- the double muscling gene in cattle (giving increased muscle mass). This gene has also been found — it is the myostatin gene;
- the marbling gene in cattle. This is related to the thyroglobulin gene.

The first two are typically markers for genes that were known to exist before they were mapped, and had a large effect.

Direct markers are generally much preferred to linked markers, if they are truly markers for major gene effects. Their biggest benefit is that they can be used even without trait measurement or pedigree recording. Despite this there is value in having such information, to monitor the effect of the major gene in different breeds/lines and production systems, and exploit it accordingly.

Possible risks with using direct markers are:

- There can be more than one mutation causing the desired genetic effect. A DNA test for only one of those mutations would not pick up all the animals with the desired effect. An example of this case was the myostatin gene for double muscling, where several mutations within the gene caused the same desired effect. If only some of the single direct markers had been adopted, there could have been false negatives in diagnostic tests.
- There is also some potential to incorrectly identify a candidate gene as a major gene directly affecting the trait of interest, just because it is near the true causative gene. In that case there is a risk of false positives: we pick the 'positive gene' but it turns out to be an indirect marker and recombination might have made it linked to the 'negative allele'.

This highlights the value of re-evaluation of markers in distinctly different stock, and a continuous need for trait recording for monitoring purposes.

Linked markers

Linked markers are only near QTL on the genome, and are not the causative mutation in the gene concerned. For a randomly chosen animal in the population, we have no clue whether one or another marker allele is associated with a preferable QTL allele. If we observe within the progeny of one sire a difference in performance between different marker alleles (as M and m in Fig. 2) we can determine which of the marker alleles is associated with the preferred QTL allele. But this information is useful only for this particular sire and its family! The information on linkage phase is also useful for its progeny, since for them we can determine the probability of whether or not they inherited the preferred QTL allele (depending on the recombination rate). Within the family, the marker and the QTL are said to be in linkage disequilibrium. Note that we can detect linkage only in heterozygous sires. When sires do not show a difference in progeny for the different marker genotypes, they may be homozygous for the QTL.[3]

[3] With linked markers, the information on which marker genotype is linked to the positive QTL allele is family specific. This linkage phase has to be determined by genotyping at least two generations (a sire and its progeny) and using phenotypic information on the progeny. In a number of families we cannot detect linkage, because the sires are homozygous for either the marker or the QTL, or both.

Sires should also be heterozygous for the markers; in addition it will be useful to genotype dams, since otherwise it can be unclear which marker allele an animal received from its sire. However, with markers abundantly available, animals could be genotyped for a panel of markers, thereby reducing the need for genotype information on dams.

It may be obvious that there is a considerable need to gather trait and pedigree information in order to establish links with genetic markers, because for each family the linkage phase between marker and QTL needs to be established. However, many breeding populations already have a performance and pedigree recording system in place. Furthermore, the need for large half-sib families (i.e. offspring by the same sire) is also reduced over time, as marker and trait information is gathered on a deeper pedigree. This is because we now have methods to use information from all relatives to make inferences about which marker variant is linked to the superior gene variants in each animal. Once a linkage phase has been established for a family, as is the case for a tested sire, trait measurement is not required for additional progeny of that sire (Fig. 3).

Besides the need for a lot more genotyping, other disadvantages of linked markers are that it may be more difficult to market the concept that bull X has a 95% chance of carrying the major gene Y, as opposed to a virtual guarantee from a direct marker test. However, the fact that linked markers cover a region of chromosome means that they could be more robust in some ways. A strategy using linked markers will lean on trait recording, and be more likely to track a major gene than relying on a direct marker that turns out to be only closely linked to the causative gene. Moreover, the information gathered in linked marker programs could be of direct benefit in verifying parentage, finding direct markers, and detecting other QTL affecting the measured traits.

In conclusion, direct markers and linked markers may both be useful. They should go hand-in-hand in application, driven by commercial demands, with a natural progression from linked markers to direct markers as more information becomes available for location of QTL.

Selecting for QTL Genotypes

Where a direct marker (DNA test) exists for a QTL, we can use direct MAS (sometimes known as genotype-assisted selection (GAS)). Where only linked markers exist for a QTL, we must use indirect MAS.

In either case, the aim is to determine QTL genotypes to assist selection decisions, either by increasing the frequency of favourable QTL alleles, or targeting their introgression into their lines. The value of this depends on a number of factors:

- Where heritability is low, the value of information on individual QTL tends to be higher because accuracy of breeding is increased by a relatively larger amount.
- Where the trait(s) of interest cannot be measured on one sex, marker information gives a basis to rank animals of that sex. This is particularly useful to determine which males should be progeny tested.
- If the trait is not measurable before sexual maturity, marker information can be used to select at a juvenile stage.
- If a trait is difficult to measure or requires sacrifice (as with many carcass traits) marker information can be used instead.

Most evaluations of MAS have considered short-time horizons. Applications that have been looked at by computer simulation are:

- the value of MAS in young bulls prior to entering a progeny testing scheme. Use of single markers gave increases in the rate of genetic gain of 5–20%, and this was improved by use of groups of markers, pointing at larger parts of the genome (haplotypes);
- the use of MAS in nucleus breeding schemes using new reproductive technologies. Marker-assisted selection is generally more useful when combined with reproductive technologies such as AI and embryo transfer and/or IVF. More progeny per élite parent allows earlier selection at reasonable accuracies, but also relies heavily on the use of family information. MAS is able to utilise more of the information that is available from the rest of the animal's family. MAS can be of value to select candidates within families before they are old enough to have individual records of their own. Improvements in genetic gain of up to 40% have been found in such schemes.

Selecting at QTL will be most useful if the positive QTL allele is found in low frequencies. However, with low frequencies there will be little variation at the QTL. Especially with linked markers, it may then be hard to detect heterozygous parents. These are needed, as illustrated in the examples, because we need both 'haves' and 'have nots' among the progeny to be able to pick out the ones which have indeed inherited the good variant from the sire. Initially we cannot distinguish sires with two desired alleles from the ones that have none of the desired alleles. Again, the combination of pedigree and performance data may allow us to link different families, conferring some to calculate genotype probabilities for (yet unmarked) major genes. This would help to pre-select potential heterozygous sires.

Strategies to Improve Bali Cattle in Eastern Indonesia
edited by K. Entwistle and D.R. Lindsay
ACIAR Proceedings No. 110
(printed version published in 2003)

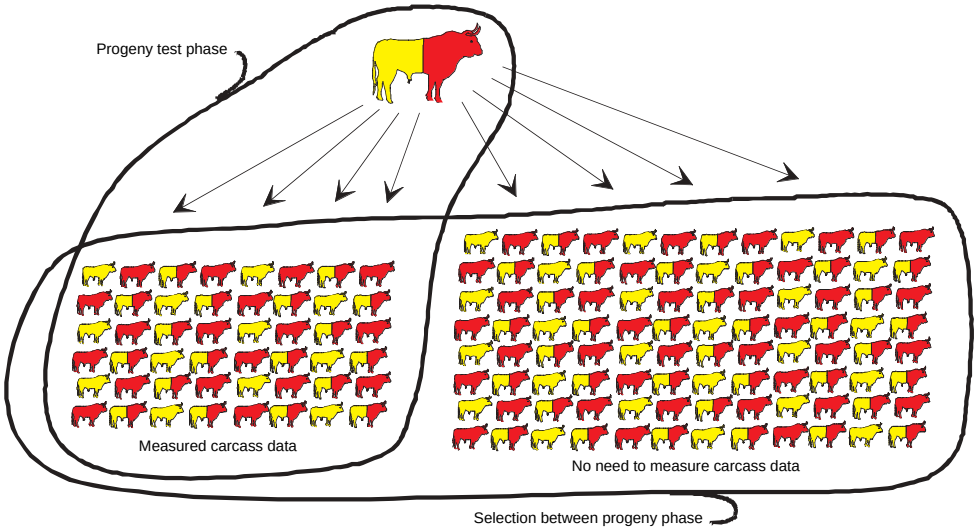

Figure 3. A sire needs to be progeny tested to establish the linkage phase between marker and QTL. Once this is established, more progeny can be generated and these can be selected on the basis of marker genotype only, without the need for trait recording.

Non-additive Genetic Effects and Mate Selection

Genetic value is the value of an animal's genes to itself. Breeding value is the value of an animal's genes to its progeny. These two are not the same when there is dominance. The effect of dominance is illustrated in Table 1. When the Q allele is dominant over the q allele, the genetic value of the heterozygote Qq is close to the effect of the homozygous QQ. The breeding value of QQ is twice the value of giving a Q allele to offspring. The breeding value of Qq is intermediate between the breeding values of QQ and qq, as a heterozygous animal will give a Q in 50% of cases and a q in the other 50%. Therefore, the breeding value of QQ and Qq animals is quite different. Such differences depend on the degree of dominance (the difference between genetic value of Qq and the average of the homozygotes QQ and qq) and on the frequency of Q alleles in the population. Dominance is quite common. It is indicated as a non-additive genetic effect as it causes genetic values of Q and q alleles to be non-additive. Dominance forms the genetic basis of the existence of heterosis in which combinations of heterozygotes are produced that have the same influence as homozygotes.

In general, breeding value has been of much more importance to animal breeders. It reflects the merit which can be transmitted to the next generation. It is the sum of the average effects of alleles carried by the animal, and because of the large number of loci classically assumed, there is no power to capitalise on anything but the average effects of these alleles, as dominance deviations in progeny cannot be predicted under normal circumstances.

However, when dealing with individual QTL we have the power to set up matings designed to exploit favourable non-additive interaction in the progeny. This means that prediction of breeding value at individual QTL will be of only partial value in many circumstances. Accordingly we should consider both prediction of breeding values and prediction of QTL genotypes, and therefore genetic values, at individual QTL.

Continuing the example of Table 1, if a Qq bull were mated consistently to QQ dams, its progeny would on average have nearly the same eye muscle area as the progeny of a QQ bull. However, if the Qq bull were mated to average dams, some of its progeny could have unfavourable qq genotypes.

Of course prediction of QTL genotype of candidates is of real value only in helping to predict genetic values of their progeny — because the object is to improve performance of descendants. This in turn means that the evaluation system should be intimately associated with the mate allocation process, wherever non-additive effects are to be exploited. The combination of animal selection and mate allocation can be termed mate selection. Application of evaluation systems to exploit individual QTL will thus frequently involve mate selection strategies in addition to the simpler ranking processes we are used to with selection.

Strategies to Improve Bali Cattle in Eastern Indonesia
edited by K. Entwistle and D.R. Lindsay
ACIAR Proceedings No. 110
(printed version published in 2003)

Table 1. Hypothetical example of a dominance effect on eye muscle area. Suppose the Q allele has a major positive effect and the difference between QQ and qq genotypes is 20 cm^2. The heterozygote's mean is close to that of the QQ genotype, as the Q allele is dominant over the q allele. The genetic values of QQ and Qq are similar, but breeding values (the value of the alleles that are passed on) are quite different. Breeding values are always additive. The example is for a population frequency of the Q allele of 0.325.

Genotype	Average eye muscle area (cm^2)	Genetic value	Breeding value
QQ	80	+10	+17.3
Qq	78	+ 8	+ 4.5
qq	60	–10	– 8.3

Another consequence of dominance is that the difference between marker genotypes will be very dependent on the dam population. If a marker were tested using a sire on a dam population with only q alleles, a large effect would be found (basically the difference between Qq and qq genotypes). However, if the same sire were tested on a population with predominantly Q alleles, the difference between marker genotypes would be the difference between QQ and Qq genotypes, which could be quite small.

Long-term Response

MAS combines the information at genetic markers with the phenotypic measurements on breeding animals and their relatives. When only parents with the major gene are selected, we will miss some animals with very good polygenic values; hence more emphasis on major genes implies less improvement of the other genes. Ultimately, the QTL will be fixed but the variance at the polygenes remains to be exploited. Therefore when considering genetic improvement after many generations, the selection for 'other genes' should not be ignored. A compromise needs to be found, and the longer the term considered the less emphasis there should be on selection for major genes (Fig. 4). However, it should be noted that the effect shown in Figure 4 is related to simple selection on a trait that can be measured in both sexes, and before selection takes place. It is when these conditions do not exist that MAS shows its real superiority.

Conclusion

Marker-assisted selection can improve selection response. Its value is limited for traits that we can breed for easily by classical methods, especially in the longer term. However, there seems great potential for MAS to generate change in traits such as pigmented fibres, meat quality, milk quality and disease resistance. Biological systems are complex, so interaction between loci should be of importance. The effect of MAS should be assessed with respect to the whole breeding goal, including animals' health. QTL effects on all relevant traits should therefore be somewhat known before MAS begins, as selection based on actual genes is likely to have more impact than selection based on phenotypic characteristics only. Given this, there will be challenging tasks in biological modelling and breeding program design to produce ideal genotypes.

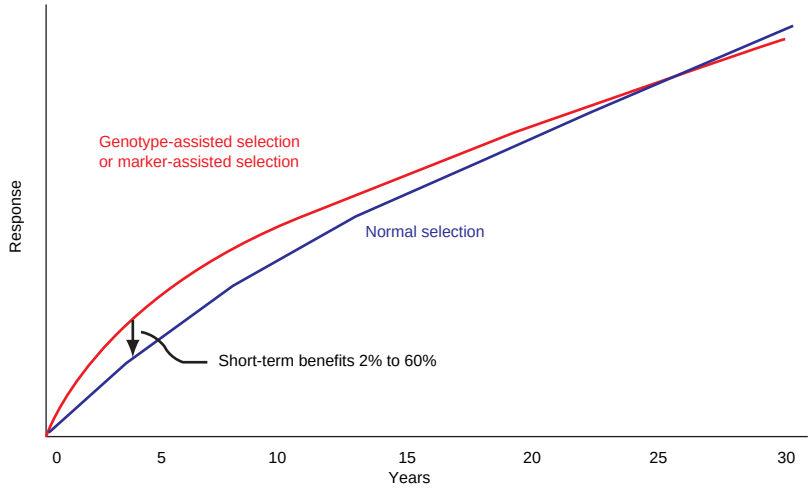

Figure 4. Long-term response compared for GAS/MAS and normal selection.

Strategies to Improve Bali Cattle in Eastern Indonesia
edited by K. Entwistle and D.R. Lindsay
ACIAR Proceedings No. 110
(printed version published in 2003)