HOMOZYGOSITY DUE TO IDENTITY-BY-DESCENT AROUND THE TARGET LOCUS IN GENE INTROGRESSION PROGRAMMES

E.E Wall[1,2,3], P.M. Visscher[1] and J.A. Woolliams[2]

[1] ICAPB, University of Edinburgh, Ashworth Labs, Kings Buildings, Edinburgh EH9 3JT, UK
[2] Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS, UK
[3] Current: SAC, Bush Estate, Penicuik, Midlothian EH26 0PH, UK

INTRODUCTION

Alleles that have descended from a single ancestral allele are said to be identical-by-descent or IBD. The coefficient of inbreeding (F) is the probability that two alleles at a locus within an individual are IBD. Therefore, studying homozygosity due to IBD in gene introgression programmes will provide a measure of inbreeding in the population homozygous for the target locus. Since blocks of linked alleles, rather than individual genes, are transmitted from one generation to the next, blocks of linked alleles rather than single genes become identical as inbreeding proceeds (Stam, 1980). In gene introgression there will be a block of donor genome around the target allele, which may, or may not, trace back to a common ancestor. The decrease in donor allelic diversity (Wall, 2002) in many introgression schemes will increase the possibility of linkage drag segments tracing back to the same ancestor.

It is the aim of this study to quantify the proportion of a carrier chromosome that is homozygous due to identity-by-descent of donor alleles (IBD homozygosity) at the completion of a gene introgression programme. The study examines the effect of: (*i*) the number of backcross generations; (*ii*) the population size; (*iii*) offspring group size; (*iv*) carrier chromosome length; and (*v*) position of the target locus. In this paper, donor IBD homozygosity at a locus is described as a function of distance from the target locus, and is related to the results of loss of allelic diversity described by Wall (2002).

MATERIALS AND METHODS

The introgression of a marker for a desired allele was performed by crossing donor and recipient individuals to create F_1 individuals followed by T generations of backcrossing to new recipient individuals. The initial cross for the introgression scheme was assumed to be between two divergent lines fixed for alternative alleles at each locus. A carrier chromosome of length l Morgan was simulated using Haldane's mapping function (1919). Selection takes places at a target locus s Morgan from the end and N parents are selected at random at each of the T backcross generations from the heterozygous offspring of the previous generation producing n offspring at each backcross generation. All recipient populations were assumed to be inbred. All heterozygous offspring are used for parents of the intercross to produce the homozygous offspring population (IC).

The effect of population structure on IBD homozygosity was studied on the carrier chromosome. Parental population size (N mating pairs) were simulated to have values ranging from N =10 to 100 (with T = 6, n = 4, l = 1 Morgan, s = l/2). The number of offspring per mating (n) had values of 2 to 10 (with T = 6, N = 20, l = 1 Morgan, s = l/2). The number of backcross generations (T) was set at values ranging from 1 to 10 (with N = 20, n = 4, l = 1

Morgan, $s = l/2$). The effect of carrier chromosome length (l) IBD homozygosity was studied simulating values of l ranging from 0.5 to 4 Morgans (representative of chromosome lengths in livestock species) and 10 Morgans. The location of the target locus on the carrier chromosome (s) ranged from the chromosome end to the centre of the chromosome, *i.e.* 0 to $l/2$ Morgans (with $T = 6$, $N = 20$, $n = 4$, $l = 1$ Morgan).

For each set of parameters, 500 replicates were simulated. The following values were recorded from the carrier chromosome of individuals post introgression (IC) (*i*) overall donor IBD homozygosity, (*ii*) IBD homozygosity at a locus x cM away from the target locus and (*iii*) generational homozygosity (homozygosity due to alleles at a locus tracing back to the same ancestral generation) due to donor and recipient alleles at a given locus. The estimates of homozygosity were calculated for individuals homozygous for the target gene.

RESULTS AND DISCUSSION

IBD homozygosity due to donor alleles over the carrier chromosome when averaged over the entire chromosome was low (4.6%). However, the IBD homozygosity at the target locus was much higher (25.6%, Figure 1), showing that the main source of IBD homozygosity on the carrier chromosome is due to the high levels of donor IBD homozygosity at and around the target locus.

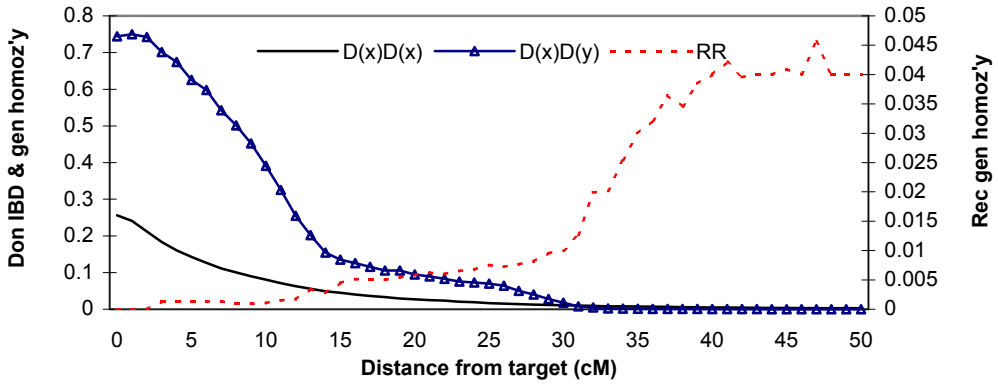

Figure 1. Distribution of donor IBD homozygosity (D(x)D(x)), recipient generational homozygosity (RR) and donor genome homozygosity (D(x)D(y)) at a locus moving away from the target locus (N =20, n = 4, T = 6, l = 1 Morgan, $s = l/2$)

The level of donor IBD homozygosity and donor generation homozygosity at individual loci dropped as distance from the target locus increased (Figure 1, 25% at target locus; 8% at 10 cM away, 0.3% at 50 cM away), halving every 6cM in this case. The dashed line on Figure 1 shows how the homozygosity due to recipient alleles tracing back to the same generation (generational homozygosity) is low at loci close to the target locus. In fact, the generational homozygosity from recipients remains low along the entire carrier chromosome, rarely rising above 4%. The linkage drag segment was approximately 28 cM in the example in Figure 1, corresponding to 14cM either side of the target, and there is a marked inflection point in the

homozygosity arising from donor generations at this point. From this point forward results will focus upon the effect of the parameters on donor IBD and generation homozygosity.
IBD close to the target locus increased as the number of backcross generations increased (Figure 2). The increase slowed down when the number of backcross generations was 6 or higher. However outside the linkage drag, IBD decreased with the number of generations. Therefore the curves shown in Figure 2 are flatter with smaller T. The increase in IBD homozygosity over backcross generations close to the target locus is in accord with the increased loss of alleles as the number of backcross generations increased. The loss of alleles also slowed down when $T > 6$ (Wall, 2002). Increasing the number of backcross generations increased the recombination events on the carrier chromosome, therefore breaking up the linkage and reducing the probability of IBD around the target allele.

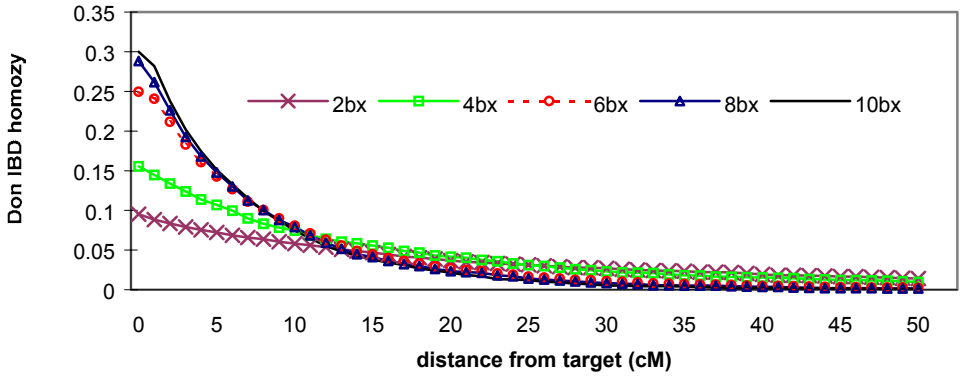

Figure 2. Effect of backcross generation on donor IBD homozygosity

Proportion of donor IBD homozygosity and donor generation homozygosity decreased as the carrier chromosome length (l) increased and as the position of the target locus (s) approached the end of the chromosome (results not shown). However, there was no difference in the IBD homozygosity at a locus x cM away from the target over all chromosome lengths studied, consistent with IBD at x being a function of the number of meioses and the recombination between x and the target locus. There was a slight increase in IBD homozygosity at a locus close to the target locus when the position of the target locus moved to the chromosome end (results not shown). This can be was attributed to the reduced linkage drag length when the target locus is at the chromosome end (when $s = 0$ the linkage drag14 cM versus ~ 28 cM for a more centrally placed target locus).

There was dramatic increase in donor IBD homozygosity at loci close to the target locus when the population size (N) was low (Figure 3) proportional to N^{-1}. The increased IBD homozygosity declined as the distance from the target locus increased to converge at a static IBD homozygosity for all N. The increased IBD homozygosity at a locus close to the target was attributed to fewer donor alleles remaining at a locus when N was low. There was decreased IBD homozygosity at loci close to the target locus when $n = 2$ but little difference for higher values. The decline in IBD homozygosity at a locus was slower for $n = 2$ than for all other n (not shown).

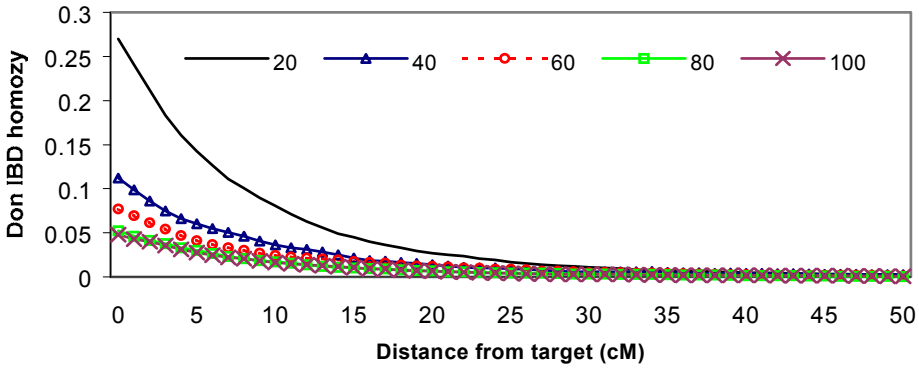

Figure 3. Effect of population size (*N*) on donor IBD homozygosity

CONCLUSION

The results suggest that fewer backcross generations and a population size of at least 50 minimise the risk of IBD homozygosity and maximise the allelic variation across the genome. Even though the IBD homozygosity on the carrier chromosome can be high, this will be diluted when averaged out across the genome. In conflict with this is that increasing the number of backcross generations reduces donor contamination in the genome and therefore genetic lag (Wall *et al.*, 2002). Increasing the population size to reduce the risk of inbreeding in a livestock introgression programme will also increase the cost the scheme due to increased maintenance and genotyping. Inbreeding, genetic lag and monetary cost of an introgression must be balanced to create a viable breeding population post introgression.

ACKNOWLEDGEMENTS

Eileen Wall is grateful for funding from the Ministry of Agriculture, Food and Fisheries Postgraduate Studentship Scheme, Teagasc Walsh Fellowship Scheme and the Irish Cattle Breeding Federation.

REFERENCES

Haldane, J.B.S. (1919) *J. Genet.* **8 :** 299-309.
Stam, P. (1980) *Genet. Res.* **35 :** 131-155.
Wall, E.E. (2002) PhD Thesis, University of Edinburgh, Edinburgh, Scotland.
Wall, E.E., Visscher, P.M. and Woolliams, J.A. (2002) *Genet. Sel. Evol.* (submitted).