Animal Science 2000, 70: 207-220

1357-7298/00/96030207$20·00

Introgression of genes responsible for disease resistance in a cattle population selected for production: genetic and economic consequences

E. H. van der Waaij and J. A. M. van Arendonk

Animal Breeding and Genetics Group, Wageningen Institute of Animal Sciences, Wageningen University, PO Box 338, 6700 AH Wageningen, The Netherlands

Abstract

The genetic and economic consequences of introgression of either one or two genes that explain the complete between-breed difference for disease resistance between donor and recipient breeds were investigated. Four backcross strategies (0, 1, 3 or 7 generations of backcrossing) were compared for four initial breed differences (0·1, 1, 2·5 and 5 phenotypic s.d.) when female reproductive capacity was either high (10 offspring) or lower (four offspring). Selection in donor and recipient populations was for production using a selection index. Genetic comparison was based on production level between the hybrid population, after fixation of the disease resistance alleles, and the donor population. For a large initial breed difference and high female reproductive capacity, application of seven generations of backcrossing resulted in the largest genetic difference between donor and hybrid populations. Introgression of one or two genes made no difference to the genetic results. From an economic point of view, optimal number of generations depends on the number of genes involved in the introgression, on the female reproductive capacity and on the initial breed difference. Seven generations of backcrossing in most cases are too many and none to three generations of backcrossing often is more optimal. Introgression of two genes is economically less attractive, especially in case of low female reproduction capacity.

Keywords: *cattle, disease resistance, introgression, production.*

Introduction

Introgression is a breeding strategy aimed at bringing favourable alleles for a certain trait from the donor into the recipient breed or line. This, for example, is of interest when an allele that explains resistance to a certain disease could be introduced into a more productive but susceptible breed. Two components can be distinguished in the introgression process: one is fixation of the favourable allele into a hybrid population, and the other is reduction or elimination of the proportion of donor alleles at the other loci. Reduction of the proportion of donor genome is achieved by (multiple) generations of backcrossing with the recipient breed, after which the favourable alleles are fixed in an intercross generation. The example of introgressing a small number of loci that explain disease resistance into a more productive breed will be used throughout the paper.

A number of studies have focussed on the second component in the introgression process by applying a breeding scheme including a number of generations of backcrossing to the recipient breed (e.g. Tanksley, 1983; Hospital *et al.*, 1992; Groen and Smith, 1995; Visscher *et al.*, 1996; Visscher and Haley, 1999). They have made use of phenotypic and/or genetic marker information to trace the favourable alleles as well as the remaining background genome. Application of multiple backcross generations decreases the contribution of donor genome. However, an increased number of backcross generations could result in an increased risk of losing the favourable allele, unless the exact position of the gene is known (e.g. Hospital and Charcosset, 1997; van Heelsum *et al.*, 1997a and b). In order to cover this increasing risk of losing the allele, an increasing number of animals and genetic markers is needed to obtain the desired number of animals after

intercrossing (Hospital and Charcosset, 1997; Koudandé *et al.*, 1999). Use of genetic markers in selection against the background genome could reduce the number of backcross generations and thus the risk of losing the favourable allele (e.g. Hospital *et al.*, 1992; Visscher *et al.*, 1996), though also would require a higher number of animals to create some room for selection. Increase in the number of genes to be introgressed has a large influence on the number of animals required during the backcross and intercross generations as well (Hospital and Charcosset, 1997; Koudandé *et al.*, 1999). In the case of introgressing recessive alleles that explain disease resistance, animals are required to be homozygous for the favourable allele (i.e. heterozygous animals are not resistant). In such cases, an increase in the number of backcross generations not only requires an increased number of animals, but also postpones the moment of introduction of hybrid animals in an infected area, which could lead to increasing costs for the introgression programme. Little attention has been paid to balancing genetic and economic consequences of different introgression strategies.

This paper compares different strategies aimed at introgression of disease resistance genes and selection on production, both in genetic and economic terms. Disease resistance is determined by a limited number of genes (one or two), while production is assumed to be under polygenic control. Breeding schemes differed in reproductive capacity of dams, number of backcross generations applied and initial breed difference between donor and recipient. Economic comparison was based on costs incurred in the backcross phase compared with profit made per animal after fixation of the favourable allele. Since this is the first study on optimization of both genetic and economic aspects of introgression schemes, the focus of the current paper will primarily be on identifying the major factors determining the optimum design of introgression schemes in a general context.

Material and methods

Two traits are considered in the analysis: disease resistance and production, which are evaluated in subsequent generations for various breeding schemes using a deterministic approach. Each generation, selection for production is applied in the purebred donor and recipient populations, and after fixation of the allele(s) that explain disease resistance also in the hybrid population. Two environments are considered: one in which no disease pressure is present and one in which only animals that are homozygous for the favourable allele would survive. Genetic comparison is based on differences in population mean between purebred and crossbred populations. Economic comparison is based on costs incurred during introgression compared with profit per animal after fixation of the favourable allele. Trypanotolerance is used as an example for disease resistance, with the N'Dama as donor and the Kenyan Boran as the recipient breed.

Genetic model

Disease resistance. The complete breed difference in disease resistance between donor and recipient populations is explained by either one ($n_g = 1$) or two genes ($n_g = 2$). Genetic markers are used for determining the favourable allele for each of the genes. There is no recombination between genetic marker and gene. Marker alleles are completely breed specific. Only homozygous animals are resistant to the disease. No marker-assisted selection against background genotype was applied during backcrossing and it was assumed that the proportion of the remaining donor genome was halved each generation.

Production. The donor and the recipient breed are assumed to differ in production level. Genetic and phenotypic variances were assumed to be equal in donor and recipient population, but the mean production level was different. Production has a polygenic character and the infinitesimal model was assumed to be applicable. However, breed-specific alleles will segregate after crossing these lines, which will result in the occurrence of segregation variance (p. 226, Falconer, 1989). This segregation variance (σ^2_{seg}) exists in addition to additive genetic variance and the size depends on the size of the initial breed difference and the number of generations since the crossbreeding event. Lande (1981) defines the 'effective number of loci' that determine the initial breed difference for the trait under consideration, in terms of the number of blocks of linked breed-specific loci. In the F_1 these blocks of loci will be of the size of a chromosome, since no recombination has occurred yet. Due to recombination events, the effective number of loci is increased each generation, resulting in a decreasing influence of the segregation variance on the total genetic variance. The speed at which the number of blocks ('loci') increases is dependent on the type of breeding in the generations following the F_1 (i.e. F_2, BC_2, etc.).

The increase in the number of 'loci' can be estimated using a Poisson distribution. In the F_1 generation no segregation occurs, resulting in the absence of segregation variance. The size of the segregation variance in the following generations is dependent on the type of breeding applied and the generation number since the F_1. The problem is that, using conventional quantitative genetics theory, it is

impossible to estimate the size of the segregation variance. In generations following the F_1 it is absorbed in the Mendelian sampling variance ($\sigma^2_{ms} = \frac{1}{2}\sigma^2_{a_0} + \sigma^2_{seg}$ i.e. half the additive genetic variance in the base generation plus an additional term representing the segregation variance). Note that, in the case of backcrossing, the influence of the segregation variance each generation is only half of that in cases where intercrossing would be applied (i.e. creating an F_n).

The size of the segregation variance in each generation could be estimated if the means and variances for each effective (= independent) locus, and thus also the effective number of loci, in the parental breeds, were known. Following Lande (1981), the difference in the mean effects of the alleles at locus i in the parental populations (i.e. μ_{i1} and μ_{i2} for parental population 1 and 2, respectively) can be written as $\delta_i = (\mu_{i2} - \mu_{i1})$, which is assumed to be of the same sign for all loci (i.e. the breed with the higher production level carries the favourable alleles). The segregation variance for an F_n generation is equal to:

$$\sigma^2_{seg} = 0{\cdot}5 \sum_{i=1}^{n} \delta_i^2 = 0{\cdot}5n\,[\sigma^2_\delta + (\bar{\delta})^2] \quad (1)$$

where n is the number of loci involved, σ^2_δ is the variance across loci due to differences in size of effect, and $\bar{\delta}$ is the mean value of δ_i averaged across all relevant loci. It is assumed that all loci have equal additive effect and consequently that $\sigma^2_\delta = 0$, reducing the segregation variance to a relation between the phenotypic difference between parental populations and the effective number of loci involved. Using equation (1), and given the fact that genetic variance was assumed to be equal in both populations, segregation variance can be written as:

$$s^2_{seg} = 0{\cdot}5 \frac{\left(\frac{\mu_2 - \mu_1}{\sigma}\right)^2}{n} \quad (2)$$

where $\mu_2 - \mu_1$ is the difference in the mean effects of the alleles at locus i in the parental populations, σ is the genetic standard deviation for the production trait considered, and n is the total number of loci. The method proposed by Lande (1981) is not very robust, as has been shown by Zeng (1990). However, it is considered to be accurate enough to be used here since the number of effective loci increases rapidly across generations, which reduces the influence of the initial number.

Lande (1981) suggested that the initial effective number of loci for most traits is between 5 and 10, with occasional values up to 20. In this paper the effective number of loci for production is set at eight, assuming that each of these eight independent loci is located on a separate chromosome of 100 cM. On average there will be 100 recombinations per 100 meioses per chromosome, resulting in on average one recombination per generation. Thus each generation, eight new independent loci are formed (i.e. when the intercross immediately follows the F_1, this results in $2 \times 8 = 16$ effective loci, which is in accordance with the results of Zeng (1992)).

Breeding schemes

The breeding scheme used throughout the paper is shown in Figure 1. Animals from the donor and recipient population are crossed to create an F_1. Then n_b generations of backcrossing are applied ($n_b = 0, 1, 3$ or 7, corresponding to recovery of on average 50%, 75%, 87·5% and > 99% of the recipient genome), followed by an intercross to fix the favourable

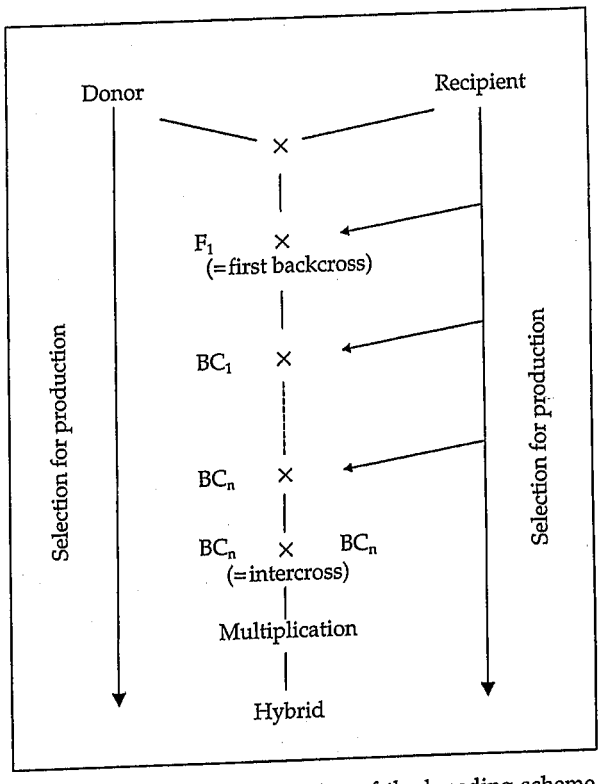

Figure 1 Schematic reproduction of the breeding scheme including terminology used throughout the paper. Selection for genetic markers is applied during the backcross generations.

disease-resistance alleles in a hybrid population. After fixation the number of animals is increased in the multiplication generation in order to create some room for selection for production in subsequent generations. Because introgression usually will occur under experimental conditions, it was assumed that it would be easier to have to buy semen than to buy cows each generation and therefore only females that carry the favourable allele are selected from the backcross population (Gama *et al.*, 1992). The same N_s sires used in the recipient population were also used as sires in the backcross generations.

Especially in the case of few numbers of backcross generations, the difference in mean production between hybrid families after multiplication can be considerable due to Mendelian sampling. Mass selection in that case would result in selection of complete families, which would lead to a large increase in level of inbreeding (ΔF). Therefore, selection for production was within family. In order to keep the situations comparable, within-family selection was applied also in the donor and recipient populations. Selection in the donor population was introduced to make a fair comparison between the results of introgression and selection for production within the donor population itself. Random mating of selected parents was applied in all cases.

Donor, recipient and hybrid populations were selected on production based on a selection index that combines information on an animal's own performance and performance of its sibs (more details in a later section of the paper). Observations for production are available on all animals of both sexes. The parental breeds (i.e. donor and recipient) differ in phenotypic means but are assumed to have equal genetic and phenotypic variances ($\sigma_P^2 = 1$, $h^2 = 0{\cdot}2$). Generations are discrete, with equal generation interval for males and females.

Numbers of animals

The purpose of the introgression programme was to create a new hybrid population that is resistant to the disease and that has a higher production level compared with the donor breed. It is important to start a new population with sufficient numbers of animals in order to keep the increase in inbreeding level limited. It was assumed that the number of females (N_d) used for breeding exceeds the number of males (N_s), so that each male breeds with multiple females and each female only with one male. One sire is selected from each sire family and one dam from each dam family. By using within-family selection, ΔF can now be determined using $\Delta F = 3/32N_s + 1/32N_d$ (Gowe *et al.*, 1959). In this study ΔF is taken to be 0·01. For a mating ratio (d) of five (i.e. five dams per sire), this results in $N_s = 10$ sires and $N_d = 50$ dams. Female reproductive capacity was either high (assuming that modern reproduction techniques were available), i.e. dams have $m_{od} = 10$ offspring per generation, or female reproductive capacity was lower ($m_{od} = 4$). The number of males born is equal to the number of females born and m_{od} represents the number of offspring that survive until reproductive age.

Introgression and multiplication. Starting at the end, the total number of animals to be produced in the intercross is dependent on the required population size after multiplication, which is equal to the eventual size of the hybrid population (m_a). Given that $N_d = 50$ and $m_{od} = 4$ or 10, m_a is equal to either 200 or 500 animals (i.e. 100 or 250 females). In order to accomplish this final number, the number of homozygous dams after fixation should equal $N_d = 50$. The number of animals (m_h) that need to be produced in order to obtain a single animal that is homozygous for the favourable allele is dependent on the number of genes that explain disease resistance ($n_g = 1$ or 2), i.e. $m_h = [(\frac{1}{2})^{2n_g}]^{-1}$. In the case of one gene (with two alleles), one out of four [i.e. $(\frac{1}{2})^{2\times1}$] will be homozygous at the intercross, so one over one out of four, i.e. four, animals are required to produce one homozygous animal. When two genes are to be introgressed, only one out of 16 will be homozygous for the favourable allele of both genes.

The number of animals (m_{ab}) to be produced per backcross generation depends on the number of genes to be introgressed, and on the reproductive capacity of the dam. For $m_{od} = 4$ and $n_g = 1$, the number of animals to be produced each backcross generation is equal to the number of animals to be produced in the intercross. Half of the animals will be female and half of those will carry the favourable allele. For $m_{od} = 10$ and $n_g = 1$, the number of dams can be decreased each previous backcross generation, counting from the intercross. For example, to produce the 400 animals needed in the intercross, only 40 females that carry the favourable allele are necessary, resulting in $2 \times 2 \times 40 = 160$ animals needed in the last backcross generation. Using the same reasoning, this would result in 64 animals in the second-last backcross generation. However, the number of males at birth in reality is not always exactly equal to the number of females. Therefore the minimum number of females carrying the favourable allele is set to 20, as a guarantee that on average the number of female offspring will be sufficient.

For introgressing $n_g = 2$ genes, the situation is different. Half of the animals born in each backcross generation are female, but only one out of four of

those will carry both favourable alleles. For $m_{od} = 10$, the number of animals required each preceding backcross generation can still be decreased, though to a lesser extent than in the case of $n_g = 1$ gene. However, for $m_{od} = 4$, the required eight animals cannot be produced by a single dam and, consequently, each generation preceding the intercross the number of dams, and thus the number of animals born, needs to be doubled.

Genetic parameters

The N'Dama breed is known to possess a high level of tolerance to trypanosomosis and therefore is chosen as an example for the donor breed. The Kenyan Boran lacks such tolerance but has a higher production level and furthermore it has a good ability to produce under local African circumstances and therefore is chosen as the recipient breed. Four breed differences were considered, which cover the range observed in the literature. Breed difference for milk productivity index was small (Murray *et al.*, 1984; International Laboratory for Research into Animal Diseases, 1989; Food and Agriculture Organization (FAO), 1999) and was assumed to be equal to 0·1 phenotypic s.d. Dressing-out percentage was equal for both breeds, and the difference in carcass weight was solely caused by difference in live weight. Average live weight in the Boran was around 600 kg (http: \\ studbook.co.za \ brinfo \ theboran \ boran.html), and in the N'Dama 300 kg (FAO, 1999). Growth rate from birth to 12 months of age in the Boran equalled 0·8 kg/day (http:\\studbook. co. za\brinfo\theboran\boran. html), and was twice as high as in the N'Dama (0·4 kg/day) (Paling and Dwinger, 1993). Assuming a coefficient of variation of 0·20 for both breeds, these two breeds are 2·5 Boran phenotypic standard deviations and 5·0 N'Dama phenotypic standard deviations apart for beef production as well as for growth rate. Considering daily live-weight gain per 100 kg body weight, N'Dama and Boran perform at a similar level (FAO, 1999). Based on these results, it was decided to investigate effects of breed differences (d_{dr}) ranging from 0·1 to 5 phenotypic standard deviations.

Selection index

Two selection paths were considered: a sire and a dam path. A selection index was constructed for each of these paths. Sires were selected within half-sib sire family, based on full-sib family average as a deviation from the half-sib family average, and on their own performance as a deviation from the full-sib family average. Dams were selected within full-sib dam family, based on their own performance as a deviation from their family average. The indices are:

$$I_{sire} = b_1(X_i - \overline{X}_{FS}) + b_2(\overline{X}_{FS} - \overline{X}_{HS}) \quad (3)$$

$$I_{dam} = b_3(X_i - \overline{X}_{FS}) \quad (4)$$

where X_i is the individual's own performance, $\overline{X}_{FS}$ is the average performance of full-sib dam family, including the individual's own performance, $\overline{X}_{HS}$ average performance of the half-sib sire family, and b_1, b_2 and b_3 are the index weights for each information source. Since within full-sib family, each animal had the same sire and dam, differences between animals within family are solely based on expression of the Mendelian sampling variance and the error variance. For calculating genetic response, variance of the selection index (σ^2_I) is determined as:

$$\sigma^2_{I_{(t)}} = \mathbf{b}_{(t)}{}' \, \mathbf{P}_{(t)} \mathbf{b}_{(t)} \quad (5)$$

where $\mathbf{b}_{(t)}$ is a vector with the index weights in generation t and $\mathbf{P}_{(t)}$ is a variance-covariance matrix between information sources in generation t. The index weights are calculated as $\mathbf{b}_{(t)} = \mathbf{P}^{-1}{}_{(t)} \, \mathbf{G}_{(t)}$, where $\mathbf{G}_{(t)}$ is a variance-covariance matrix between information sources and breeding values in generation t. For each generation **P** and **G** in the sires are given as (ignoring the subscript representing generation t):

$$\mathbf{P} = \begin{bmatrix} \mathrm{var}(X_i - \overline{X}_{FS}) & \mathrm{cov}((X_i - \overline{X}_{FS}), (\overline{X}_{FS} - \overline{X}_{HS})) \\ \mathrm{cov}((X_i - \overline{X}_{FS}), (\overline{X}_{FS} - \overline{X}_{HS})) & \mathrm{var}(\overline{X}_{FS} - \overline{X}_{HS}) \end{bmatrix}$$

$$= \begin{bmatrix} (\sigma^2_{ms} + \sigma^2_e) \times (1 - \frac{1}{n}) & 0 \\ 0 & [\sigma^2_d + (\sigma^2_{ms} + \sigma^2_e)/n] \times (1 - \frac{1}{d}) \end{bmatrix}$$

$$\mathbf{G} = \begin{bmatrix} \mathrm{cov}(A_i, (X_i - \overline{X}_{FS})) \\ \mathrm{cov}(A, (\overline{X}_{FS} - \overline{X}_{HS})) \end{bmatrix} = \begin{bmatrix} \sigma^2_{ms} (1 - \frac{1}{n}) \\ (\sigma^2_d + \frac{\sigma^2_{ms}}{n}) \times (1 - \frac{1}{d}) \end{bmatrix}$$

where σ^2_d is the dam variance in generation t, which comprises one quarter of the additive genetic variance in the selected parents of the dam in generation $(t - 1)$. The Mendelian sampling term in generation t is given by $\sigma^2_{ms_t} = \frac{1}{2}\sigma^2_{a0} + \sigma^2_{seg_t}$, where σ^2_{a0} is the additive genetic variance in the base generation (i.e. independent of t). In purebred selection the segregation variance is non-existent. The error variance is σ^2_e (also independent of t), n is the number of full-sib family members and d is the mating ratio.

The b values for the sire index can be expressed as (again ignoring the subscript for generation t):

$$b_1 = \frac{\sigma^2_{ms} (1 - \frac{1}{n})}{(\sigma^2_{ms} + \sigma^2_e) \times (1 - \frac{1}{n})} = h^2_w$$

and

$$b_2 = \frac{\sigma_d^2 + \sigma_{ms}^2/n}{\sigma_d^2 + (\sigma_{ms}^2 + \sigma_e^2)/n}.$$

Note that b_1 is equal to the within full-sib family heritability as mentioned by Falconer (1989) and Hill *et al.* (1996) and b_2 actually is equal to the within half-sib family heritability. In general, accuracy of selection can be represented as the standard deviation of the selection index, divided by the standard deviation of the breeding goal (A): $r_{IA} = \frac{\sigma_I}{\sigma_A}$. Since in this case the breeding goal only contains one trait, the standard deviation of the breeding goal is equal to the genetic standard deviation: $r_{Ia} = \frac{\sigma_I}{\sigma_a}$. Within the sires, accuracy of selection in generation t is given by:

$$r_{Ia_{sire(t)}} = \sqrt{\frac{\sigma^2_{I_{sire(t)}}}{\sigma^2_{a_{sire(t)}}}} \qquad (6)$$

where $\sigma^2_{a_{sire(t)}}$ is the additive genetic variance within a sire family in generation t, which is equal to the total additive genetic variance in generation t minus the additive genetic variance among the selected sires in the previous generation ($\sigma^2_{a_{s(t-1)}}$), since the selection is within half-sib sire family.

Since selection of dams is within full-sib family, both the sire and dam variance are equal for all selection candidates, and are therefore left out of consideration. For each generation

$$\mathbf{P} = \mathrm{var}(X_i - \overline{X}_{FS}) = (\sigma_{ms}^2 + \sigma_e^2) \times (1 - \tfrac{1}{n}),$$

$$\mathbf{G} = \mathrm{cov}(A, (X_i - \overline{X}_{FS})) = \sigma_{ms}^2 (1 - \tfrac{1}{n})$$

and thus

$$b_3 = (\sigma_{ms}^2 (1 - \tfrac{1}{n}))/((\sigma_{ms}^2 + \sigma_e^2) \times (1 - \tfrac{1}{n})) = h_w^2.$$

The accuracy of selection in the dams in generation t is given by:

$$r_{Ia_{dam(t)}} = \sqrt{\frac{\sigma^2_{I_{dam(t)}}}{\sigma^2_{a_{dam(t)}}}} = \sqrt{\frac{\sigma^2_{ms_{(t)}} \times h^4_{w(t)}}{\sigma^2_{ms_{(t)}}}} = \sqrt{h^2_{w(t)}} \qquad (7)$$

where $\sigma^2_{a_{dam(t)}}$ is the additive genetic variance within a dam family in generation t.

Genetic variances in generation t were corrected for reduction due to linkage disequilibrium in previous generations (Bulmer, 1971). Additive genetic variance in generation t is equal to $\sigma^2_{a_{s(t-1)}} + \sigma^2_{a_{d(t-1)}} + \sigma^2_{ms_t}$. Additive genetic variance in selected sires in generation t is given by $\sigma^2_{a_{s(t)}} = \frac{1}{4}\sigma^2_{a_t}$ $(1 - k_s \times r^2_{Ia_{sire(t)}})$ and in selected dams by $\sigma^2_{a_{d(t)}} = \frac{1}{4}\sigma^2_{a_t}$ $(1 - k_d \times r^2_{Ia_{dam(t)}})$, where $k_y = i_y\,(i_y - x_y)$, i_y being the selection intensity in the respective selection path for sex y, and x is the truncation point.

Response to selection

In generation t, the population mean (μ_t) can be represented by the average breeding value in the selected parents:

$$\mu_{(t)} = 0{\cdot}5 \times (\mu_{s(t-1)} + \mu_{d(t-1)}) \qquad (8)$$

where $\mu_{s(t-1)}$ is the average breeding value in the selected sires and $\mu_{d(t-1)}$ is the average breeding value in the selected dams in the previous generation. The average in each of the selected parents is:

$$\mu_{y(t-1)} = i_y \times \sigma_{I_y} + \mu_{(t-2)} \qquad (9)$$

where y is s (sire) or d (dam), i_y is the selection intensity in sex y, and σ_{I_y} is the standard deviation of the selection index in sex y. Selection intensity depends on the family size and thus on the female reproductive capacity. When m_{od} equals 4, selection intensities are $i_s = 1{\cdot}539$ for sires and $i_d = 0{\cdot}564$ for dams. When m_{od} equals 10, selection intensities are $i_s = 1{\cdot}965$ for sires and $i_d = 1{\cdot}163$ for dams. In calculating these selection intensities, the effect of finite population size has been taken into account (Burrows, 1972). Genetic gain in generation t is calculated as the increase in population mean from generation t–1 to generation t (i.e. $\Delta G = \mu_t - \mu_{t-1}$). Since no selection is applied in the backcross generations, genetic gain in the backcross generations is dependent on the genetic gain in the recipient sires.

Environments

Two environments are considered: environment 1 (E1) represents an uninfested area in which non-resistant animals are kept, and environment 2 (E2) an infested area in which only animals that are homozygous for the disease-resistance allele will survive. Medication for heterozygous animals is no option. The recipient breed reaches high production levels in E1 but is not able to survive in E2 because of disease pressure. The donor breed is resistant to the disease present in E2 but is less productive. The crossbred animals in the backcross and intercross generations need to be kept in E1 and will be less productive compared to the purebred recipient breed.

Costs

The cumulative costs consist of the difference in population mean for m_{ab} animals in the t backcross and intercross generations (μ_r) compared with the

purebred recipient population mean (μ_r), cumulated over generations, and are calculated as

$$C = \sum_{i=1}^{t} m_{abi} (\mu_{ri} - \mu_{ci}),$$

expressed in phenotypic standard deviations. Breeding schemes in all populations are equal (i.e. equal number of sires and dams, equal reproductive capacity) and it is assumed that costs (except for genotypings) are entirely due to difference in production level. Costs for production of hybrid animals (i.e. costs for not having recipient animals when having backcross animals) are defined as opportunity costs (= $\mu_r - \mu_c$) (Dijkhuizen and Morris, 1997). The size of the cumulative costs is dependent on the size of the opportunity costs as well as on the number of animals involved in the backcross and intercross generations. Costs are also incurred for collecting DNA samples and for genotyping of the animals in the backcross and intercross generation to trace the favourable allele. All animals need to be genotyped for the first gene. If there is a second gene involved, it is assumed that only those animals that carry the favourable allele of the first gene need to be genotyped for the second, i.e. in the case of two alleles for each gene,

$$T = m_{ab} \sum_{i=1}^{ng} 0{\cdot}5^i,$$

where T is the total number of genotypings. All females are genotyped for $i = 1$. For $i = 2$, half of the females (i.e. those that carry the favourable allele for the first gene) need to be genotyped additionally. In the intercross generation, all animals (males and females) are genotyped for the first gene and half of them for the second gene.

Costs compensation

The hybrid animals are transported to E2 as soon as they are homozygous for the favourable allele. Profit can be made as soon as C is overcome, which is after production of a certain number of hybrid animals (m_{ac}) to compensate for the costs of introgression (C). This m_{ac} is dependent on the opportunity costs (i.e. costs for not having the donor animals when having the hybrid animals, which will be negative when the genetic level of the hybrid animals exceeds the genetic level of the donor animals) and on the size of C (i.e. $m_{ac} = C/(\mu_{ht} - \mu_d)$). Note that m_{ac} is different for each introgression scheme. The time frame for compensating for C is hard to define and depends on the selection and multiplication strategy applied by the breeding company. For example, reproductive capacity of males and/or females could be increased to produce higher number of animals, so that costs are compensated more rapidly. Hybrid animals that are not selected as parents could be sold, so that income is not only dependent on production level anymore.

Genetic and economic comparison

The genetic comparison is based on the difference in population mean for production level of the hybrid compared with the donor population. The economic comparison is based on the number of hybrid animals that need to be produced in order to compensate for the cumulative costs for loss in production during the backcross and intercross generations, and on the number of genotypings that are needed during the introgression process.

Results

Segregation variance

Figure 2 shows the course of the segregation variance relative to the total additive genetic variance over generations for four initial breed differences. When the breed difference is small, the influence of segregation variance on the total variance is negligible. Only when initial breed difference becomes large (2·5 or 5 s.d.), is the influence of segregation variance substantial, especially during the first generations. The influence of segregation variance diminishes rapidly over generations, since the number of effective loci is increased by eight each generation; it remains influential only for an initial breed difference of 5 s.d.

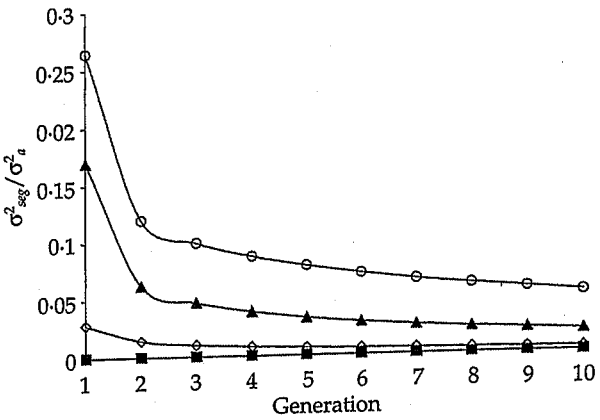

Figure 2 Ratio of segregation variance (σ^2_{seg}) to total additive genetic variance (σ^2_a) across generations for initial breed differences of 0·1 (■), 1·0 (◇), 2·5 (▲) and 5·0 (○) phenotypic s.d.

Table 1 *Difference in production level (in phenotypic standard deviations) between hybrid animals and donor animals during the generation of multiplication for four different backcross strategies (n_b), comparing four initial breed differences (d_{dr}) and high or lower female reproductive rate (m_{od})*

	$n_b = 0$	$n_b = 1$	$n_b = 3$	$n_b = 7$
d_{dr}‡		$m_{od} = 10$†		
0·1 s.d.	−0·670	−0·690	−0·705	−0·711
1·0 s.d.	−0·220	−0·015	0·139	0·185
2·5 s.d.	0·530	1·110	1·545	1·679
5·0 s.d.	1·780	2·985	3·889	4·169
		$m_{od} = 4$†		
0·1 s.d.	−0·461	−0·456	−0·449	−0·445
1·0 s.d.	−0·011	0·219	0·395	0·448
2·5 s.d.	0·739	1·344	1·801	1·942
5·0 s.d.	1·989	3·219	4·145	4·432

† Number of offspring per dam.
‡ Expressed in phenotypic standard deviations.

Genetic comparison

Table 1 shows differences in population mean between the hybrid and the donor population during the generation of multiplication (d_{hd}), expressed in phenotypic standard deviations. Positive values indicate superiority of introgression over selection within the donor population. However, in the case of none or one generation of backcrossing, small negative values can be overcome in the subsequent few generations of selection in the hybrid population, due to a slightly higher genetic variance compared with the donor population. This additional genetic variance is caused by the smaller influence of the Bulmer effect and, to a much lesser extent, by the presence of some segregation variance, though the influence of this additional variance will disappear within a limited number of generations of selection. The smaller reduction of variance due to the Bulmer effect in the hybrid population (i.e. compared with the donor population), is caused by the absence of selection in the dams during the backcross generations, and by the complete absence of selection during fixation and multiplication. Four backcross strategies ($n_b = 0$, 1, 3 or 7) were compared for four initial breed differences ($d_{dr} = 0{\cdot}1$, 1·0, 2·5 or 5·0) and two female reproductive rates ($m_{od} = 4$ or 10). For $d_{dr} = 0{\cdot}1$, and in the case of $n_b = 0$ or 1 for $d_{dr} = 1{\cdot}0$, selection for production within the donor population resulted in a larger increase in population mean compared with introgression in the case of low female reproductive capacity. In the case of high female reproductive capacity, donor animals did not exceed the production level of the hybrid animals when $d_{dr} = 0{\cdot}1$ and $n_b = 1$. In general, mean production level in the hybrid population increases with increasing d_{dr} and increasing n_b. Considering the positive values in Table 1, highest mean production levels in the hybrid population were accomplished when $m_{od} = 4$ for $n_b = 0$ or 1 generations of backcrossing applied, and when $m_{od} = 10$ for $n_b = 3$ or 7.

Figure 3 shows the increase in population mean for production when initial breed difference is 2·5 s.d. and female reproductive capacity is high. The difference in population means between recipient and crossbred individuals decreases during the generations of backcrossing, and subsequently increases again during the generations of fixation and multiplication. Visscher and Haley (1999) have already mentioned such an increase in superiority of the recipient animals compared with the crossbred animals in the later stages of the introgression programme. Another cause of lower mean production level of the crossbred animals is the fact that during introgression, no selection for production is applied in the females. However, especially when $m_{od} = 4$, the selection intensity in females in recipient (and donor) populations is low, and the absence of selection during introgression therefore does not have much influence (not illustrated in the figure).

Genetic gain (ΔG) for production in selected populations (donor, recipient and hybrid populations) was approximately 0·23 for $m_{od} = 10$ and 0·16 when $m_{od} = 4$. The number of generations of selection in the donor population to compensate for the genetic lag with the hybrid population can be

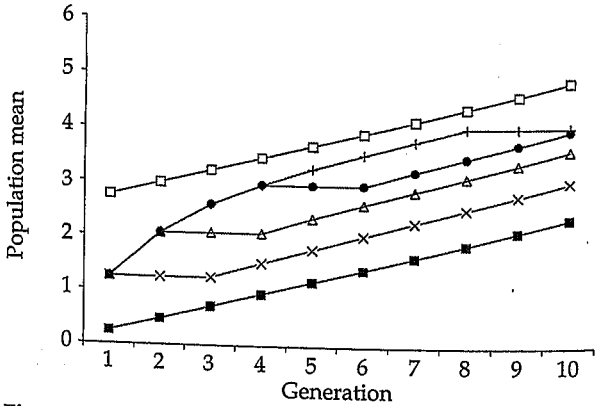

Figure 3 Increase in population mean for production when initial breed difference is 2·5 s.d. and female reproductive rate is high, comparing selected purebred populations (donor (■) and recipient (□)) to animals in the four introgression schemes (n_b = 0 (×), n_b = 1 (△), n_b = 3 (●) or n_b = 7 (+) generations of backcrossing). Selection was applied in each of the purebred populations and after fixation of the favourable alleles also in the hybrid population.

Table 2 *Cumulative costs due to difference in production level between recipient and backcross and intercross populations (C, in phenotypic standard deviations) during introgression of a single gene for disease resistance, numbers of animals to be produced after fixation of the favourable allele to compensate for these costs (m_{ac}) and number of genotypings needed until fixation (T), considering high and lower female reproductive capacity (m_{od}) and various initial breed differences (d_{dr})*

		$m_{od}=10$			$m_{od}=4$		
d_{dr}	n_b	C	m_{ac}†	T	C	m_{ac}†	T
0·1 s.d.	0	261·92	–‡	480	248·00	–	600
	1	299·36	–	520	336·40	–	800
	3	362·00	–	600	512·80	–	1200
	7	477·52	–	760	866·80	–	2000
1·0 s.d.	0	513·92	–	480	608·00	28280	600
	1	461·36	8788	520	696·40	2567	800
	3	456·40	2616	600	872·40	2078	1200
	7	550·24	2718	760	1228·0	2673	2000
2·5 s.d.	0	933·92	1613	480	1208·0	1566	600
	1	731·36	622	520	1296·4	930	800
	3	614·08	389	600	1472·8	807	1200
	7	672·64	397	760	1829·0	936	2000
5·0 s.d.	0	1633·9	894	480	2208·0	1093	600
	1	1181·4	388	520	2296·4	703	800
	3	876·40	224	600	2472·4	593	1200
	7	877·60	210	760	2828·4	637	2000

† Including the animals in the multiplication generation.
‡ No profit will be made due to a higher genetic level of the donor breed.

calculated as $n_c = d_{hd}/\Delta G$. From Table 1 it can thus be concluded that for example $n_b = 0$ for $m_{od} = 10$ results in differences in population mean between donor and hybrid populations that can be overcome in an additional $n_c = 2$ ($d_{dr} = 2{\cdot}5$) to $n_c = 8$ ($d_{dr} = 5$) generations of selection in the donor population. For $m_{od} = 4$, $n_c = 5$ ($d_{dr} = 2{\cdot}5$) to 13 ($d_{dr} = 5$) when there is no backcrossing ($n_b = 0$), and $n_c = 3$ ($d_{dr} = 1$) to 28 ($d_{dr} = 5$) for $n_b = 7$.

Economic comparison

Table 2 shows results of the economic analysis for production of animals that are homozygous for the favourable allele, for the case of one gene ($n_g = 1$) explaining the complete breed difference in disease resistance. The economic comparison is presented in terms of the cumulative costs (C) for production of homozygous animals, the number of animals (m_{ac}) that need to be produced in order to compensate for these costs (i.e. depending on the size C and on the difference in production level between hybrid and donor animals shown in Table 1), and the number of genotypings (T) required through to fixation. For example, the value for C in Table 2 for $m_{od} = 10$, $d_{dr} = 1{\cdot}0$ s.d. and $n_b = 3$ equals 456·40 production units (e.g. kg growth). Difference in production level between hybrid and donor animals is 0·174, so hybrid animals produce 0·174 production unit more than donor animals. Thus the number of hybrid animals needed to compensate for the costs (i.e. m_{ac}) equals 456·4/0·174 = 2616 animals. The choice of the optimal breeding scheme, from an economical point of view, is dependent on the balance between C and m_{ac}. Four backcross strategies ($n_b = 0$, 1, 3 or 7) were compared for four initial breed differences ($d_{dr} = 0{\cdot}1$, 1·0, 2·5 or 5·0) and two female reproductive rates ($m_{od} = 4$ or 10).

An initial breed difference of 0·1 s.d. is too small to enable a profitable introgression scheme. For schemes with a larger initial breed difference and high female reproductive capacity, three generations of backcrossing in most cases results in the lowest values for C and m_{ac}. Only in the case of $d_{dr} = 5{\cdot}0$ s.d., seven generations of backcrossing is more optimal with regards to numbers of animals to be produced. However, this difference in numbers of animals is negligible, especially considering the four generations of backcrossing that are additionally required, and thus, in the case of $n_b = 3$, the four additional generations after multiplication can be used to produce this small difference in number of animals. Low female reproductive capacity produces different results. Immediate fixation of the favourable allele is optimal, considering the size of the opportunity costs. In contrast, three generations of backcrossing is optimal considering the size of m_{ac}. This smaller value of m_{ac} is caused by the higher production level of hybrid animals after $n_b = 3$ compared with $n_b = 0$ or 1. Which scheme is optimum depends on what is more important from an economic point of view: low C or low m_{ac}. The m_{ac}

Table 3 *Cumulative costs due to difference in production level between recipient and backcross and intercross populations (C, in phenotypic standard deviations) during introgression of two genes for disease resistance, numbers of animals to be produced after fixation of the favourable allele to compensate for these costs (m_{ac}) and number of genotypings needed until fixation (T), considering high and lower female reproductive capacity (m_{od}) and various initial breed differences (d_{dr})*

		$m_{od} = 10$			$m_{od} = 4$		
d_{dr}	n_b	C	m_{ac}†	T	C	m_{ac}†	T
0·1 s.d.	0	1237·8	–‡	1920	1352·0	–	2400
	1	1614·5	–	2176	2785·6	–	4000
	3	2172·7	–	2548	11398	–	13600
	7	2771·7	–	2934	183776	–	103200
1·0 s.d.	0	2533·8	–	1920	3512·0	163349	2400
	1	2723·3	51873	2176	6745·6	24846	4000
	3	2933·9	16813	2548	26784	63772	13600
	7	3090·1	15260	2934	429467	934640	103200
2·5 s.d.	0	4693·8	8107	1920	7112·0	9219	2400
	1	4571·3	3883	2176	13346	9567	4000
	3	4203·9	2569	2548	52438	28710	13600
	7	3626·3	2138	2934	839147	429341	103200
5·0 s.d.	0	8293·8	4535	1920	13112	6487	2400
	1	7651·3	2507	2176	24346	6655	4000
	3	6317·8	1610	2548	95184	20873	13600
	7	4524·5	1081	2934	1521749	342190	103200

† Including the animals in the multiplication generation.
‡ No profit will be made due to higher genetic level of the donor breed.

reflects the balance between cost (C) on the one hand and productive advantage of the hybrid animals on the other. Considering the fact that production of animals usually will be more costly than production of one production unit, low values for m_{ac} and thus three generations of backcrossing will be optimal in most situations. However, the two additional generations available for producing the extra animals should be taken into account as well.

Table 3 shows results of the economic analysis in the situation in which two genes ($n_g = 2$) explain the entire breed difference for level of disease resistance. Since genetic gain for production is not influenced by the number of genes to be introgressed, results in Table 1 are also applicable here. An important difference compared to the results in Table 2 is that, for $m_{od} = 10$ and $d_{dr} = 2{\cdot}5$ and 5·0 s.d., the optimal value, when looking at m_{ac}, for n_b has gone from zero, when the favourable allele of a single gene is to be introgressed, to seven, for the situation where two genes are involved. In the case of high female reproductive capacity, the number of animals in the initial generation of the backcross can be decreased by increasing the number of backcross generations. For $n_b = 3$, 664 animals are needed in the F_1, and for $n_b = 7$, this figure has further decreased to 272 animals. This is in contrast to the 1280 animals to be produced in the F_1 for the case of $n_b = 0$. The potential to decrease the number of animals needed in the first few backcross generations can be very profitable, especially in the case of a large initial breed difference ($d_{dr} = 2{\cdot}5$ or 5·0 s.d.), since the difference between crossbred and purebred animals is greatest during the first three generations of backcrossing (see Figure 2 for illustration). One standard deviation initial breed difference is not enough to take advantage of this mechanism.

For $m_{ac} = 4$, immediate fixation of both alleles is optimal in most of the cases, because the number of females (and thus animals) needs to be doubled each additional backcross generation in order to produce the required number of animals carrying both favourable alleles. Only in the case of $d_{dr} = 1{\cdot}0$ s.d., is m_{ac} reduced considerably when a single generation of backcrossing is applied.

T reaches higher values for $n_g = 2$ than for $n_g = 1$, because of both higher number of genes to be genotyped and higher number of animals due to reproductive limitations. Especially for $m_{od} = 4$, the genotype workload can become substantial.

Example. The genotype load is largest when two genes explain the difference in disease resistance. Considering the number of genotypings to be done to create 100 animals that are homozygous for both favourable alleles, costs for collecting DNA and doing the actual genotyping can become substantial.

For the case of $n_b = 7$ and $m_{od} = 4$, in total 409 600 animals need to be genotyped: only females during the generations of introgression and also the males during the intercrossing. Assuming that costs for the collection of DNA for one individual and its genotyping would equal two US dollars, total costs for collection of DNA and genotyping would equal 819,200 US dollars.

Discussion

This paper aimed at determining optimum breeding schemes for introgression of one or two genes that explain the breed difference for disease resistance, considering both genetic and economic aspects. Results in this paper show that, depending on initial breed difference and on female reproductive capacity, it is economically attractive to create a hybrid that is resistant to the disease and more productive than the donor population. Female reproductive capacity affects the number of animals that are needed during the production of the hybrid population.

Introgression of trypanotolerance genes in African cattle is chosen as an example. Even though it is likely that multiple genes are involved in explaining the tolerance level (Kemp *et al.*, 1997), it was assumed that only one or two genes represent all the genetic variance in the level of trypanotolerance between the donor and the recipient breeds. It is also assumed that treatment with drugs was not possible, and thus that only those animals that are homozygous for the favourable alleles explaining the disease resistance could be introduced into infested areas. However, in practice there is still genetic variation for trypanotolerance in the N'Dama (e.g. Trail *et al.*, 1991) and, depending on the severity of infection, many animals will need one or more treatments with drugs. Also it is quite likely that none of the N'Damas is 100 percent resistant to the disease, simply because the alleles conferring such a level of resistance do not exist.

The present study did not aim at giving direct answers to the question of which introgression scheme is best for a specified situation, but it rather aimed at giving support to making such a decision in more general terms, by showing the principles involved in comparing alternative introgression schemes. In doing so, a number of simplifying assumptions have been made, including the use of non-overlapping generations. In cattle schemes, overlapping generations will occur in particular when no reproduction techniques are used to increase the number of offspring per female. With the use of modern reproduction techniques, overlapping generation can be avoided but that, in particular for $m_{od} = 10$, puts high demands on infrastructure as well as financial investments which might make it not practically feasible. However, the use of overlapping generations is not expected to have a large impact on the comparison of alternatives.

With mass selection, in the case of overlapping generations, older animals will not have the advantage of additional information from relatives (i.e. no progeny testing applied), and therefore the youngest animals will be selected for breeding (i.e. they have higher genetic potential), which is equal to the case of discrete generations. During introgression, animals are not selected for production and increase in production level is solely due to the use of selected sires from the recipient line. The difference between purebred and introgression therefore remains the same as with discrete generations. In the present paper, within family selection is applied instead of mass selection, which has some impact on the effect of overlapping generations. Animals from the best families will be used during multiple generations and the poorest families will not be used for breeding at all. This will cause some increase in genetic gain. However, this increase in genetic gain will be passed on to the generations of introgression in the same way as with discrete generations and therefore will have no influence on the difference between lines. The use of overlapping generations in the introgression process will result in fewer animals needed per generation, since animals that carry the favourable allele(s) can be used more then once, which will have a decreasing effect on the costs. However, older animals will have a lower genetic potential for production, and that will result in an increase in costs again. Summarizing, there will be an effect of using overlapping instead of discrete generations, but the size of the effect will be limited.

We have assumed that modern reproduction techniques were applied in the hybrid populations as well as in both purebred populations. By assuming that such techniques have been used in all schemes, differences between the schemes are entirely due to effects of introgression and not a mixture of effects caused by both differences in reproductive capacity and effects of introgression. In practise, however, it might not be realistic to apply the same reproductive techniques in all three populations. In Figure 3, it can be seen that the rate of genetic gain is equal in all populations, which is caused by assuming the same values for m_{od}.

A lower m_{od} in the donor population, for example due to the harsh environment under which the animals are kept, does not affect the cost of the

introgression process (C). However, the lower value results in a larger difference in performance between the hybrid and donor population and consequently in a lower value of m_{ac}. In addition, the difference between both populations would increase with time and thus would favour a larger number of backcross generations. These effects result from differences in reproductive capacity of females, and not from differences in introgression strategy and have, therefore, been avoided in this paper.

Selection in hybrid, donor and recipient populations was on production. The production level of animals is expected to differ between the environment with and without the disease challenge. The difference in production level between the donor and hybrid animals is assumed to be constant in the two environments. In other words, the absolute production level of the hybrid population might differ between an infested and a non-infested environment, but the same difference also holds for the donor population. In the economic calculations, the concept of opportunity costs is applied. Therefore, the difference in production level between the two environments does not affect the economic comparison.

Selection strategies for production are assumed to be equal for the N'Dama and the Boran so that a fair comparison can be made between introgression and selection for production within the N'Dama. The empty cells in Tables 2 and 3 show that selection for production within the N'Dama is preferred over introgression when initial breed difference is small. However, selection in the donor line might be difficult or even impossible, e.g. when animals are kept in a small-holder production system, in which case the difference between donor and hybrid animals will grow larger. However, setting up a breeding scheme for the donor population should be considered as an alternative to an introgression programme. Consequently, the performance of a selected donor is the most appropriate point of comparison.

In small-holder farming systems, often all cattle are retained at the farm since that is considered a more reliable investment than putting money in the bank. Cattle can be used for draught power, calf production (interest), milk and meat production, are not subject to inflation and can be sold when cash is needed. Breed replacement by investing in new animals is therefore often not an option but purchasing semen from hybrid bulls when AI facilities are present, or collectively buying such a bull, is. On larger farms, it is likely that hybrid animals (both cows and bulls, though especially bulls) are used for crossbreeding to upgrade the purebred N'Dama animals. The final result of introgression is thus not only the creation of a 'new breed', but it also provides opportunities for upgrading of the present donor population.

Results in Table 1 suggest that, especially for large d_{dr}, hybrid animals are most superior to donor animals when seven backcross generations are applied. However, from an economical point of view, application of fewer generations of backcrossing can be more favourable, especially for $n_g = 2$ and $d_{dr} = 0{\cdot}1$ or 1 s.d. for $m_{ac} = 10$ and for all d_{dr} in the case of $m_{ac} = 4$ (Table 3). The availability of modern reproduction techniques ($m_{od} = 10$) reduces the number of animals needed during introgression and the number of hybrid animals needed for compensation of costs incurred during introgression (m_{ac}). Especially when $m_{od} = 4$ these costs can become substantial. In the most extreme case, for $m_{od} = 4$, $d_{dr} = 1{\cdot}0$ sd. and $n_b = 7$, it will take a lot of effort to produce m_{ac}. This illustrates that the reproductive capacity of the dam plays an important rôle when deciding to start an introgression programme.

From the economic point of view, application of none or only a few generations of backcrossing seems most favourable (e.g. in the case of $m_{od} = 4$ and $n_g = 1$ or 2). An additional advantage of applying few backcross generations is that animals can be transported to E2 after a small number of generations. This means a saving in costs and a reduction in the risk of losing animals in the backcross phase due to, e.g. disease. Increasing the number of backcross generations results in increased total costs (C) (Tables 2 and 3). The size of the hybrid population after multiplication is determined under the assumption that none of the animals that is selected for breeding will die before the required number of offspring has been produced. However, it is possible that there is, for example, an outbreak of a disease for which none of the animals is resistant (e.g. rinderpest) and new hybrid animals are needed to replace the lost individuals. In that case, costs are also incurred because these hybrid animals are not yet available when required (opportunity costs). An increase in the number of backcross generations then will result in an increase of these opportunity costs. Such extra costs have not been incorporated into the results of the present study.

It was assumed that genotyping is performed after animals are born. For some known genes (e.g. BLAD, kappa casein) it possible to genotype at the embryo level (P. Bredbacka, personal communication). Such developments are expected to seriously reduce costs in introgression programs, since only embryos that

carry the favourable alleles will be implanted. At present, sex determination of embryos using PCR techniques is possible (Bredbacka, 1998) and is used in practice on a small scale. Sex determination would reduce the number of animals required in the backcross generation (i.e. embryos of the other gender do not have to be implanted). Still, use of females in the backcross generations requires a higher number of backcross animals than when using males.

The gene(s) to be introgressed are assumed to account for all genetic variation for disease resistance, though in reality this might not be the case. It is possible that multiple genes influence the level of expression of disease resistance, of which a few genes have a relatively large effect and the rest of the genes all have approximately equal, small effect so that the infinitesimal model can be applied. Introgression of one or two genes then accounts for only a certain percentage of the initial between-breed variation. A low number of backcross generations will result in a relatively large portion of the donor genome that is available for additional polygenic selection for disease resistance in the hybrid population, which would no longer be available after a large number of backcross generations. Furthermore, there might be some interaction between alleles somewhere on the donor genome and the favourable allele that could possibly be preserved when fewer backcross generations are applied.

Acknowledgements

The authors would like to thank WOTRO (Netherlands Foundation for the Advancement of Tropical Research) for financial support. They also would like to thank Leyden Baker for fruitful discussion and Piter Bijma, Dr Soller and Frank Nicholas, for helpful comments on the manuscript.

References

Bredbacka, P. 1998. Recent developments in embryo sexing and its field application. *Reproduction, Nutrition, Development* **38:** 605-613.

Bulmer, M. G. 1971. The effect of selection on genetic variability. *American Naturalist* **105:** 201-211.

Burrows, P. M. 1972. Expected selection differentials for directional selection. *Biometrics* **28:** 1091-1100.

Dijkhuizen, A. A. and Morris, R. S. 1997. *Animal health economics. Principles and applications.* Post Graduate Foundation publication of the University of Sydney, Sydney.

Falconer, D. S. 1989. *Introduction to quantitative genetics, third edition.* Longman Scientific and Technical, New York.

Food and Agriculture Organization. 1999. *Productivity of trypanotolerant livestock.* Animal production and health paper no. 20, pp. 87-97.

Gama, L. T., Smith, C. and Gibson, J. P. 1992. Transgene effects, introgression strategies and testing schemes in pigs. *Animal Production* **54:** 427-440.

Gowe, R. S., Robertson, A. and Latter, B. D. H. 1959. Environment and poultry breeding problems. 5. The design of poultry control strains. *Poultry Science* **38:** 462-471.

Groen, A. F. and Smith, C. 1995. A stochastic simulation study of the efficiency of marker-assisted introgression in livestock. *Journal of Animal Breeding and Genetics* **112:** 161-170.

Heelsum, A. M. van, Visscher, P. M. and Haley, C. S. 1997a. Marker-assisted introgression using non-unique marker alleles. I. Selection on the presence of linked marker alleles. *Animal Genetics* **28:** 181-187.

Heelsum, A. M. van, Visscher, P. M. and Haley, C. S. 1997b. Marker-assisted introgression using non-unique marker alleles. II. Selection on probability of presence of the introgressed allele. *Animal Genetics* **28:** 188-194.

Hill, W. G., Caballero, A. and Dempfle, L. 1996. Prediction of response to selection within families. *Genetics, Selection, Evolution* **28:** 379-383.

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

Hospital, F., Chevalet, C. and Mulsant, P. 1992. Using markers in gene introgression breeding programs. *Genetics* **132:** 1199-1210.

International Laboratory for Research into Animal Diseases. 1989. *N'Dama cattle: managing Africa's genetic resources.* ILRAD report, October 1989. Intern publication, Nairobi.

Kemp, S. J., Iraqi, F., Darvasi, A., Soller, M. and Teale, A. J. 1997. Localization of genes controlling resistance to trypanosomiasis in mice. *Nature Genetics* **16:** 194-196.

Koudandé, O. D., Thomson, P. C. and Arendonk, J. A. M. van. 1999. A model for population growth of laboratory animals subjected to marker-assisted introgression: how many animals do we need? *Heredity* **82:** 16-24.

Lande, R. 1981. The minimum number of genes contributing to quantitative variation between and within populations. *Genetics* **99:** 541-553.

Murray, M., Trail, J. C. M., Davis, C. E. and Black, S. J. 1984. Genetic resistance to African trypanosomiasis. *Journal of Infectious Diseases* **149:** 311-319.

Paling, R. W. and Dwinger, R. H. 1993. Potential of trypanotolerance as a contribution to sustainable livestock production in tsetse affected Africa. *Veterinary Quarterly* **15:** 60-67.

Tanksley, S. D. 1983. Molecular markers in plant breeding. *Plant Molecular Biology* **1:** 3–8.

Trail, J. C. M., d'Ieteren, G. D. M., Maille, J. C. and Yangari, G. 1991. Genetic aspects of anaemia development in trypanotolerant N'Dama cattle. *Acta Tropica* **48:** 285-291.

Visscher, P. M. and Haley, C. S. 1999. On the efficiency of marker-assisted introgression. *Animal Science* **68:** 59-68.

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

Zeng, Z.-B. 1990. How informative is Wright's estimator of the number of genes affecting a quantitative character? *Genetics* **126:** 235-247.

Zeng, Z.-B. 1992. Correcting the bias of Wright's estimates of the number of genes affecting a quantitative character: a further improved method. *Genetics* **131:** 987-1001.

(*Received* 6 *April* 1999—*Accepted* 22 *October* 1999)