

**Breeding for Trypanotolerance
in African Cattle**

Promotor: **Prof. Dr. ir. J.A.M. van Arendonk**
Persoonlijk hoogleraar bij de leerstoelgroep Fokkerij en Genetica,
Wageningen Universiteit

Co-promotoren: **Prof. Dr. S.C. Bishop**
Roslin Institute, Edinburgh, UK.
Prof. Dr. J.B. Gibson
International Livestock Research Institute, Nairobi, Kenya

Samenstelling Promotie Commissie:
Prof. Dr. ir. A.J. van der Zijpp (Wageningen Universiteit)
Prof. Dr. ir. H.F. Savelkoul (Wageningen Universiteit)
Prof. Dr. J.C.M. Dekkers (Iowa State University, USA)
Prof. Dr. M.C.M. de Jong (Wageningen Universiteit en ID-Lelystad)

NN08201,3022.

Breeding for Trypanotolerance in African Cattle

Elisabeth Hillechien van der Waaij

Proefschrift
ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit
prof. dr. ir. L. Speelman
in het openbaar te verdedigen
op vrijdag 7 september 2001
des namiddags te twee uur in de Aula.

15N 162b09b

Breeding for Trypanotolerance in African cattle

Doctoral thesis

E.H. van der Waaij, 2001

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

Het in dit proefschrift beschreven onderzoek is gefinancierd door de Stichting voor Wetenschappelijk Onderzoek van de Tropen (Wotro), onderdeel van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO).

ISBN: 90-5808-458-2

Printed by Universal Press, Veenendaal, The Netherlands

Abstract

Trypanosomosis, or sleeping sickness, is one of the most important livestock diseases in Africa. Some West African cattle breeds show a degree of resistance to a trypanosome infection: they are trypanotolerant. At the International Livestock Research Institute (ILRI) in Nairobi, Kenya, an F2 experiment has been established to unravel the genetic background of trypanotolerance. This thesis had two main aims: First to determine the genetic background of trypanotolerance, and second to investigate opportunities to incorporate this information in a breeding scheme to increase performance of cattle in tsetse-infested areas. Based on the results from the F2 experiment, several traits were defined, which reflected features of trypanotolerant cattle. Subsequently, based on preliminary results from an analysis to determine chromosome fractions containing genes (QTL) involved in trypanotolerance performed at ILRI, the mode of expression of these QTL was investigated and one of the QTL was found to be maternally imprinted. These QTL could be utilised in an introgression scheme, but also for within breed selection. Both options were investigated. When introgressing QTL for disease resistance the optimal number of backcross generations from genetic or economic point of view was found to be different. The number of animals required is increasing very rapidly with increasing number of QTL to be introgressed. Within breed selection to increase production under constant infection pressure can be applied with or without aid of QTL for disease resistance. Mass selection on production under infection can be applied if no QTL information is available. A non-linear selection response is achieved in both potential production and disease resistance. Important advantage of QTL information for disease resistance is that animals can be selected outside the infected environment. In implementing a breeding scheme it is important to take into account that social-economic values and environments are very different in large parts of Africa as compared to Western countries. This thesis has demonstrated that there are good opportunities for using selection to improve the results of local farming systems.

Table of Contents

Chapter I	General Introduction	1
Chapter II	An F2 Experiment to Determine Properties of Trypanotolerance in African Cattle	9
Chapter III	Mode of expression of some QTL related to trypanotolerance in African cattle.	33
Chapter IV	Introgression of genes responsible for disease resistance in a cattle population selected for production: Genetic and economic consequences	57
Chapter V	Modeling selection for production traits under constant infection pressure	85
Chapter VI	Utilising genetic markers for disease resistance to improve production under constant infection pressure	113
Chapter VII	General Discussion	131
	Summary	153
	Samenvatting	159
	Nawoord	169
	Curriculum Vitae	170

Stellingen

1. Bij constante infectiedruk en onvolledige resistentie resulteert selectie op productie in verbetering van zowel productie als resistentie (*dit proefschrift*).
2. Gezien het grote aantal genen dat resistentie tegen slaapziekte bij rundvee bepaalt, is introgressie niet de optimale fokkerijstrategie om prestatie te verhogen in tseetsee-geïnfesteerd gebied (*dit proefschrift*).
3. Selectie van landbouwhuisdieren in een 'Specific Pathogen Free' omgeving is strijdig met de vraag vanuit de praktische veehouderij naar ontwikkeling van robuustere dieren.
4. Ten gevolge van de sterke urbanisatie zal in veel Afrikaanse landen de sociaal-economische waardering van vee verschuiven van status- naar inkomenverhogend.
5. Interdisciplinair onderzoek leidt tot een toename in het aantal wielen dat echt wordt uitgevonden.
6. Om de acceptatie van allochtonen te bevorderen zouden Nederlandse jongeren moeten worden aangemoedigd om tijdens hun opleiding een periode in het buitenland door te brengen.
7. Met het onderwijs en de gezondheidszorg in Nederland is het net als met het Nederlandse voetbalelftal: de waan van het aan de top staan blijft langer voortduren dan de werkelijkheid.

Stellingen bij het proefschrift van E.H. van der Waaij: Selection for trypanotolerance in African cattle. Wageningen, 7 september 2001.

I General Introduction

Africa is a continent with a very strong increase in human population and a dramatic urbanisation, and as a consequence, the demand for meat and dairy products is also increasing rapidly. Some countries import large amounts of dairy products (Dempfle and Jaitner, 1999). Increase in African livestock production, therefore, would be desirable. However, diseases like trypanosomosis have a seriously limiting influence on production, as well as on animal welfare, in a substantial part of sub-Saharan Africa. Trypanosomosis is the most important livestock disease in Africa and it is transmitted to cattle, as well as to many other types of livestock as for example sheep, goat (e.g. Murray et al., 1984), fish (e.g. Mukherjee and Haldar, 1982), horses, donkeys, camels and also to humans (Murray et al., 1990). On the other hand, many types of wildlife, like African buffalo, oryxes, eland, and waterbuck are highly resistant to the disease (Murray et al., 1984). Trypanosomes are mostly transmitted by the tsetse fly, which currently infests around 11 million km² of Africa, which is 37% of the continent, involving 40 countries (Murray et al., 1991). Figure 1 shows a map of Africa with the tsetse distribution and cattle densities.

Kristjanson et al. (1999) estimated that at present over 46 million cattle are kept in tsetse infested environment, of which 17 million are treated with medication, at an annual cost of \$35 million. Including additional costs due to production loss, mortality, reduced fertility, etc., annual costs for producers and consumers are estimated to exceed \$1 billion (Kristjanson et al., 1999). The trypanosusceptible *Bos indicus* breeds, such as the Kenyan Boran, are not capable of handling the infection. They become anaemic, lose weight, show reduced milk production and reduced capacity to work, lose fertility, often abort the foetus, and, unless treated with medication, frequently die of the infection (Murray et al., 1991). About 24% of the total African cattle population is related to trypanotolerant breeds, but only 6% of the cattle is of pure trypanotolerant breeds, *Bos taurus* breeds such as the N'Dama and the West African Shorthorn breeds. Members of those breeds are able to gain weight and show a normal oestrus cycle and thus maintain reproductive capacity in areas with moderate infection pressure (Murray et al., 1991).

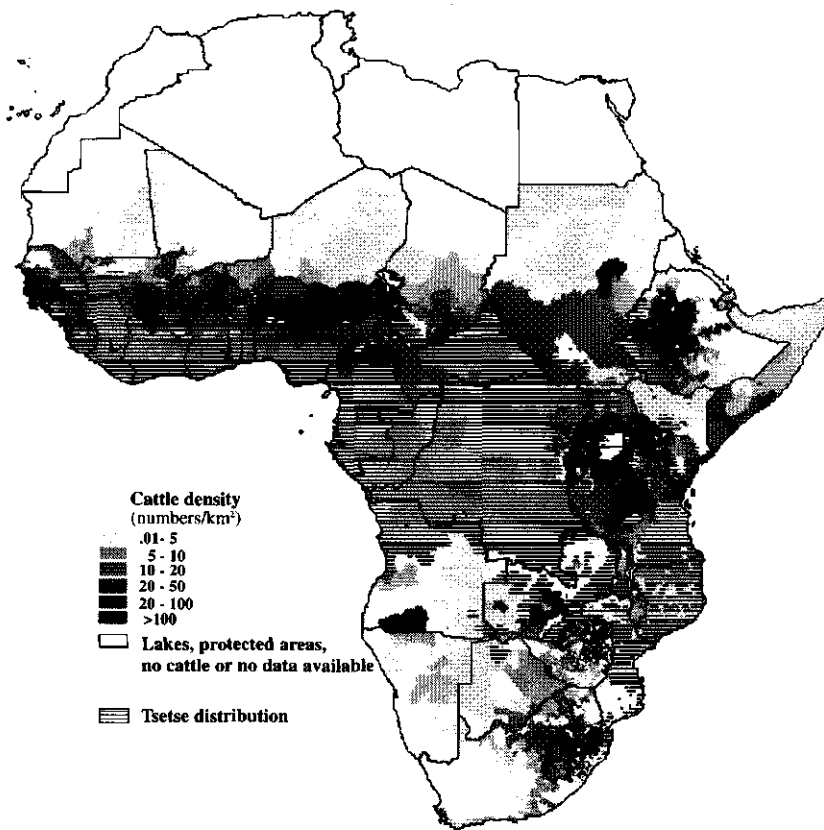


Figure 1. Map of Africa with tsetse distribution and cattle densities (Bron: ILRI, Nairobi)

Reason for the difference in trypanotolerance between the *Bos taurus* and *Bos indicus* breeds is believed to be that the *Bos taurus* longhorn (the ancestor of the N'Dama) may have had time to adapt to trypanosomes infections as this species has arrived into the Nile Delta about 5000 B.C. From there they followed the Mediterranean and subsequently travelled south into West Africa and the tsetse belt. The shorthorn breeds have arrived in the Nile Delta between 2,750 and 2,500 B.C, and from there they travelled south into East Africa (Hanotte et al., 2000). The tsetse fly arrived before the wild bovidea, which came into the area 20-40 million years ago (Murray et al., 1991). Wild bovidea, like the African Buffalo, are resistant to trypanosomosis (d'Ieteren et al., 1998). The *Bos indicus* did not become numerous in Africa until after the Arabian invasion (A.D. 699) and, therefore, had less time to adjust to the local circumstances such as the high infection pressure with trypanosomosis (Murray et al., 1984). Furthermore, because they were less well adapted to tsetse infested

environment and there was an alternative breed (the *Bos taurus* breeds), the *Bos indicus* breeds were most likely not heavily introduced into such environment, resulting in a lower selection pressure to increase trypanotolerance.

Since the introduction, the trypanotolerant taurine West African breeds have adapted themselves to the local circumstances such as distinct seasons with feed shortage, and exposure to a diverse range of infections. Murray et al., (1981) reviewed that the "N'Dama can show a considerable range in rectal temperature during the course of a day, behaving more in the heterothermic manner of wild animals". This would suggest that they have developed an increased capacity for water conservation, which is an important feature to survive in dry savannah zones (Murray et al., 1981), and partly explains why the N'Dama is able to survive a severe dry season of 7 months under constant tsetse challenge (Murray et al., 1991). Apart from trypanotolerant, the N'Dama is also resistant to several tick born diseases such as dermatophilosis, heartwater, bovine anaplasmosis, and bovine babesiosis. The breed shows a lower prevalence to strongyle worm infestations, and when infected they have lower egg outputs than zebu Gobra cattle (reviewed by d'Ieteren et al., 1998). It is likely that the N'Dama is small and does not show a high absolute production level as a consequence of the challenging environment, under which the breed has been kept.

As reviewed by Taylor (1999) it was surgeon/captain David Bruce who in 1894 was ordered to Zululand in South Africa to unravel the mysterious cause of death of hundreds of cattle in that region. He became the first to show that protozoa in the blood of sick animals were causing the disease, that wild herbivores did not seem to be affected by infection with these protozoa and thus served as reservoirs of infection, and that tsetse flies (*Glossina* species) transmitted the disease. The type of protozoa David Bruce studied was later named *Trypanosoma Brucei* (*brucei*). There are several types of trypanosomes, not all of them infective in each type of host. In domestic livestock the three most important tsetse-transmitted types of trypanosomes are: *Trypanosoma congolense*, *Trypanosoma brucei* and *Trypanosoma vivax*. These are transmitted by 36 different species of tsetse fly, each adapted to different ecological and climatic conditions (Murray et al., 1991).

There is no vaccine available to use in the field due to the ability of the trypanosome to change its antigenic composition during infection in the host. Consequently, trypanosomosis control often consists of controlling the tsetse fly population. Until recently, tsetse control was mainly achieved by spraying of insecticides. Problems encountered in that type of control are the high costs, the lack of trained personnel to implement insecticide control programmes, the need to defend sprayed areas and prevent re-invasion, the need for

constant surveillance since early detection is necessary, and the increasing demand for restricted use of insecticides (Murray et al., 1991). Alternative strategies, like use of traps and targets for tsetse flies, and apply odour-baited control (dipwash, spray or pour-on cattle with synthetic pyrethroids) have been investigated (Murray et al., 1991). The exploitation of trypanotolerant cattle will not take away the necessity of tsetse control programmes. It will reduce costs of medication and production loss, but the tsetse population will remain unaffected. Wildlife acts as a constant source of re-infection of the flies, so that infection pressure for other livestock and man will remain present. However, a consequence of introduction of livestock in tsetse infested areas is that there will be some bush clearance. Bush is the natural habitat of the tsetse fly. Clearance of bush may help to decrease the tsetse fly density, and thus decrease infection pressure. Bush clearance, on the other hand, has an ecological impact on the environment.

Trypanosomosis results in anaemia, which is the primary cause of death of infected cattle (Taylor, 1999). Trypanotolerant animals show a greater ability to control infection and a greater resistance to the pathological effects of the parasite (Tizard, 2000). They also seem to compensate more efficiently for anaemia by increasing their red blood cell production (reviewed by Taylor, 1999). The genetic background of trypanotolerance in cattle has not been studied in detail yet because of lack of data. Under practical circumstances, the trypanotolerant breeds such as the N'Dama are kept under local farm management, which often implies small holder farming systems. In most cases no records are kept of the pedigree of the individual animals or the performance of the individual animals with regards to the traits that could be selected on. Yet, some heritability estimates are available for different traits related to trypanotolerance from experiments that were designed to investigate the immunological and epidemiological background of trypanotolerance (e.g. Trail et al. 1991). Due to insufficient number of animals in the data, these estimates are not very accurate, but they demonstrate that there are genetic differences with regards to trypanotolerance within the population. This offers opportunities to set up breeding schemes to exploit these differences.

In 1987, Soller and Beckman proposed to set up an F2 experiment in cattle in an FAO report. This experiment was to take place at the International Trypanotolerance Centre (ITC) in Banjul, the Gambia, but could not be initiated due to insufficient resources. Teale and Kemp subsequently designed an F2 experiment for cattle at the International Livestock Research Institute (ILRI) in Nairobi, Kenya, which should lead to identification of Quantitative Trait Loci (QTL) related to trypanotolerance (Teale, 1993). In 1992-1993 Teale

and Kemp designed a parallel experiment in mice, also aiming at identifying QTL related to resistance to trypanosomosis, and as a test case. The initial intention of the cattle experiment was that QTL subsequently could be used in an introgression program to upgrade the production level of cattle in tsetse-infested areas of Africa. Since the generation interval in cattle is much longer than in mice, the aim was to accurately map QTL related to trypanotolerance in mice, and, using these QTL, design an introgression experiment. Kemp et al. (1997) first identified QTL related to trypanotolerance in mice. Subsequently Clapcott et al. (2000) revealed evidence for genomic imprinting of one of those QTL. QTL related to trypanotolerance in cattle have first been detected by Hanotte et al (in preparation).

Results and experiences from the mice experiment have taught us much about the prospects and problems that may be encountered when setting up an introgression scheme for the much more expensive cattle. Goals of the mice experiment were for example: would the QTL come to expression outside the background genome of the donor breed; would introgression of the three QTL that were detected in the mice experiment be sufficient to create tolerant animals, and how many animals and backcross generations would be required to create high productive trypanotolerant cattle. Koudandé et al (1999, 2000, and 2000a) have addressed most of these issues in mice. An important conclusion from their work related to utilising QTL in introgression schemes in cattle in the future is that a large number of animals is required. This number is rapidly increasing with the number of QTL to be introgressed, and with the number of backcross generations.

This thesis had two main objectives: 1. to define traits related to trypanotolerance and test those on the data from the F2 cattle experiment in order to determine the genetic factors (heritability and QTL) involved in trypanotolerance, and 2. to explore alternative breeding schemes to improve both production and disease resistance.

Before performing a QTL analysis it is important to have a definition at one's disposal that accurately describes trypanotolerance. This enables definition of traits, representing different aspects of trypanotolerance. In Chapter 2 results are presented of a series of traits, based on the observations that are available from the F2 cattle experiment. Once the traits have been defined, they can be combined with the genotypic information in order to detect and characterise QTL affecting these traits (Chapter 3). The resulting QTL could be used in a breeding scheme, for which there are several options. One option is to use the QTL for introgression into a higher productive breed, so that trypanotolerance from one breed can be combined with production potential of another (Chapter 4). Furthermore,

improving production level may be achieved by within breed selection in the trypanotolerant breeds. This can be achieved by phenotypic selection (Chapter 5), or with additional aid of QTL in a marker assisted breeding scheme (Chapter 6). Whichever breeding scheme is chosen, it is important to make sure that the result, i.e. improved animals, successfully finds its way to the farmers. Different steps involved in the implementation of breeding schemes are described and discussed in Chapter 7.

References

- Clapcott, S.J., Teale, A.J. and S.J. Kemp. 2000. Evidence for genomic imprinting of the major QTL controlling susceptibility to trypanosomiasis in mice. *Parasite Immunology* 22: 259-263.
- Dempfle, L., and Jaitner, J. 1999. Case study about the N'Dama breeding programme at the International Trypanotolerance Centre (ITC) in The Gambia (West Africa). FAO workshop on developing breeding strategies for lower input animal production environment. Rome, 22-25 Sep., 1999.
- d'Ieteren, G.D.M., Authié, E., Wissocq, N., and Murray, M. 1998. Trypanotolerance, an option for sustainable livestock production in areas at risk from trypanosomiasis. *Rev. sci. tech. Off. int. Epiz.* 17: 154-175.
- Hanotte, O., Tawah, C.L., Bradley, D.G., Okomo, M., Verjee, Y., Ochieng, J., and Rege, J.E.O. 2000. Geographic distribution and frequency of a taurine *Bos taurus* and an indicine *Bos indicus* Y specific allele amongst sub-Saharan African cattle breeds. *Molecular Ecology* 9: 387-396.
- Kemp, S.J., Iraqi, F.I., Darvasi, A., Soller, M. and Teale, A.J. (1997). Localisation of genes controlling resistance to trypanosomiasis in mice. *Nature Genetics* 16: 194-196.
- Koudandé, O.D., Thomson, P.C., and Van Arendonk, J.A.M. 1999. A model for population growth of laboratory animals subjected to marker-assisted introgression: how many animals do we need? *Heredity* 82: 16-24.
- Koudandé, O.D., Iraqi, F., Thomson, P.C., Teale, A.J., and Van Arendonk, J.A.M. 2000. Strategies to optimize marker-assisted introgression of multiple unlinked QTL. *Mamm. Genome* 11: 145-150.
- Koudandé, O.D., Van Arendonk, J.A.M., Bovenhuis, H., Gibson, J.P., and Iraqi, F. Trypanotolerance QTL introgression in mice: experimental results. In: Koudandé, O.D.

- Introgression of trypanotolerance genes. Doctoral thesis, Wageningen, 2000a. pp. 76-94.
- Kristjanson, P.M., Swallow, B.M., Rowlands, G.J., Kruska, R.L., and de Leeuw, P.N. 1999. Measuring the costs of African animal trypanosomosis: the potential benefits of control and returns to research. *Agricultural Systems* 59: 79-89.
- Mukherjee, M. and Halder, D.P. 1982. Host susceptibility experiments on a new fish trypanosome in India. *Angew. Parasitol.* 23:69-73.
- Murray, M., Clifford, D.J., and Gettinby, G. 1981. Susceptibility to African trypanosomiasis of N'Dama and Zebu cattle in an area of *Glossina morsitans submorsitans* challenge. *Vet. rec.* 109: 503-510.
- Murray, M., Trail, J.C.M., Davis, C.E. and Black, S.J. 1984. Genetic resistance to African trypanosomiasis. *J. Infect. Dis.* 149: 311-319.
- Murray, M., Trail, J.C.M., d'Ieteren, G.D.M. 1990. Trypanotolerance in cattle and prospects for the control of trypanosomiasis by selective breeding. *Rev. sci. tech. Off. int. Epiz.* 9:369-386.
- Murray, M., Stear, M.J., Trail, J.C.M., d'Ieteren, G.D.M., Agyemang, K., and Dwingler, R.H. 1991. Trypanosomiasis in cattle: prospects for control. In: Breeding for disease resistance in farm animals. Edited by J.B. Owen and R.F.E. Axford. CAB International, Wallingford UK. 1991. 499 pp.
- Soller, M. & Beckman, J.S. 1987. Toward an understanding of the genetic basis for trypanotolerance in the N'Dama cattle of West Africa. Consultation Report FAO, Rome.
- Taylor, K.A.. 1999. Immune responses of cattle to African trypanosomes: protective or pathogenic? *Int. J. Parasitol.* 28:219-240.
- Teale, A.J. 1993. Improving control of livestock diseases: Animal Biotechnology in the consultative group on international agricultural research. *Bioscience* 43:475-483.
- Tizard, I.R. 2000. Veterinary immunology: an introduction. Sixth edition. Philadelphia, USA, 482 pp.
- Trail, J.C.M., d'Ieteren, G.D.M., Maille, J.C. and Yangari, G. 1991. Genetic aspects of control of anaemia development in trypanotolerant N'Dama cattle. *Acta Tropica* 48:285-291.

II An F2 Experiment to Determine Properties of Trypanotolerance in African Cattle.

Some West African Bos taurus cattle breeds (N'Dama and West African Shorthorn) are tolerant of trypanosomosis. In literature, changes in PCV and growth rate following infection are considered to be indicators for trypanotolerance. The current paper focuses on description of changes in, and relations between Packed red Cell Volume percent (PCV), body weight and parasite count following an infection with a single clone of Trypanosoma congolense IL1180. In an experiment 214 F2 cattle (Gambian N'Dama x Kenyan Boran) were infected by the bites of infected tsetse flies. Body weight, PCV, and parasite counts were recorded on a weekly basis for 150 days post infection. Seventeen derived traits were defined based on the data recorded. The average of F2 animals were intermediate between the N'Dama and Boran for all traits, and the highest and lowest responders in the F2, when selected on maximum drop in PCV or on maximum drop in body weight, were equal to the average of the Boran and N'Dama breeds. The most trypanotolerant animals follow similar courses to the pure-bred N'Dama animals and the most susceptible animals follow the pure-bred Boran. There are moderate to low phenotypic correlations (0.00 to 0.32) between average log(parasite count) or number of times an animal was detected parasitaemic, and the PCV and body weight derived traits. There are low to moderate phenotypic correlations (0.02 to 0.74) between and (0.01 to 0.96) within PCV and body weight derived traits. Most of the traits defined in this study are heritable. Heritabilities ranged from 0.01 for PCV recovery, to 0.88, for initial PCV. Some F2 animals seem to be able to control anaemia and have a higher average body weight and body weight gain than the pure-bred N'Dama. Body weight gain following infection seems an appropriate and easy to measure indicator of trypanotolerance.

Keywords: trypanotolerance, N'Dama, Boran, F2 experiment, trait definition, *Trypanosoma congolense*

Introduction

Tsetse fly transmitted trypanosomosis affects around 11 million km² (37%) of the African continent, affecting human and their livestock in 40 countries. In susceptible animals a trypanosomosis infection results in poor growth, weight loss, low milk yield, reduced capacity to work, infertility, abortion, and death (Murray et al., 1991). There is no vaccine for trypanosomosis, and currently the most important means of controlling the

This chapter is a paper in preparation for submission

EH van der Waaij, O Hanotte, JAM Van Arendonk, SJ Kemp, JP Gibson, and A Teale. An F2 experiment to determine properties of Trypanotolerance in African cattle.

disease is prophylaxis and treatment with trypanocidal drugs. Some West African *Bos taurus* cattle breeds such as the N'Dama and several of the West African Shorthorn breeds, are able to remain productive under natural challenge without treatment with medication. These so-called trypanotolerant cattle are small by comparison with many susceptible African zebu and European breeds, though under trypanosome challenge their productivity is greater (ILRAD, 1989; Murray et al., 1990). The trypanotolerant taurine breeds account for 5% of the cattle inhabiting the infested areas of Africa (ILRAD, 1989). Research to date suggests that trypanotolerance is a complex of traits, likely to consist of a component describing the clearance of parasites and a component describing the ability of an animal to control anaemia and to remain in good condition, despite infection. Efficiency in clearance of parasites is difficult to describe because available techniques for their quantification are imperfect and, additionally, they give no indication of parasite turnover. The number of times an animal is detected parasitaemic over a series of observations, also has obvious limitations, but it is a more robust measure than parasite counts in blood samples, and has an influence on PCV-related traits (Trail et al., 1991).

Trail et al. (1993) and Trail et al. (1994) have tested some criteria for trypanotolerance and their effect on reproductive performance in N'Dama cattle under field conditions. There is evidence that Packed red Cell Volume percent (PCV), a measure for anaemia, is a good indicator for trypanotolerance (Trail et al., 1992, Dwinger et al., 1994). According to Murray et al. (1991), live weight gain and change in PCV after infection are good indicators for trypanotolerance. One of the most commonly used definitions is that infected trypanotolerant animals are capable of efficiently recovering their PCV after infection and to gain live weight at the same rate as uninfected animals (Paling & Dwinger, 1993; Trail et al., 1992; Murray et al., 1991). The heritability of trypanotolerance related traits has not been studied in detail yet, mainly due to lack of pedigree information, or limited size of data sets. However, Trail et al. (1991) have estimated some heritabilities on data obtained from N'Dama under field conditions, and Rowlands et al. (1995) have estimated some heritabilities for trypanotolerance related traits in zebu cattle.

An F2 population of cattle was established at the International Livestock Research Institute (ILRI), Nairobi, Kenya, to follow segregating alleles related to trypanotolerance in a cross between trypanotolerant and susceptible cattle (Soller and Beckman, 1987; Teale, 1993; Teale, 1997). The purpose of the present study is threefold: first to describe the consequences of a first infection with *Trypanosoma congolense* in controlled experimental conditions and using naïve animals on parasitaemia, anaemia and growth in N'Dama ×

Boran F2 animals and in pure bred N'Dama and Boran animals; second, using this information, to define a series of traits describing these changes in parasitaemia, anaemia and growth; third to estimate heritabilities and phenotypic correlations between these traits.

Material and Methods

The study population

An F2 population was created consisting of a cross between N'Dama (*Bos taurus*) × Boran (*Bos indicus*). The F2-animals, were produced at the International Livestock Research Institute (ILRI), Kenya. Four N'Dama sires (origin ITC, the Gambia, transferred to ILRI as embryos (Jordt et al., 1986)) were mated to four "improved" Kenyan Boran dams (ILRI, Kapiti farm), resulting in four full-sib F1-families. The N'Dama bulls were typically trypanotolerant (Paling et al., 1991) and the Boran cows were typically susceptible (Teale, Kemp and Kennedy, unpublished). Eleven males and 15 females were selected from the F1-animals to produce the F2-generation. Matings between full-sibs were avoided, resulting in seven major full-sib F2 families (22-41 animals per family) and seven minor families (21 F2 animals in total). One of the F1 sires of a minor family (one offspring) was also the sire of one of the major full-sib families. Two of the F1 sires of the minor families were sires of three families each, varying in size from one to three offspring. The minor families consisted of five F1 males, mated to eight F1 females, resulting in families of one to three offspring each. For producing the seven major families, offspring of one pair of grandparents were mated to the offspring of another pair of grandparents, resulting in four

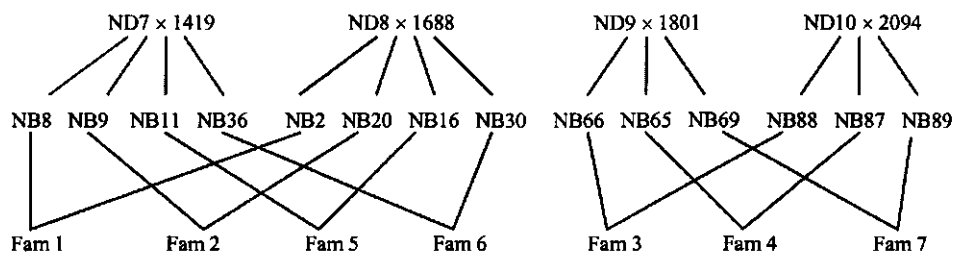


Figure 1. Schematic overview of the F2 family structure. All N'Dama's (ND) are males and all Boran (e.g. 1419) are females. ND8 and ND10 are full sibs. In the F1, NB8, NB9, NB16, NB30, NB66, NB65 and NB89 are females. Number of animals per F2 family is 40 in fam 1, 26 in fam 2, 22 in fam 3, 23 in fam 4, 22 in fam 5, 23 in fam 6 and 25 in fam 7.

major families descending from two pairs of grandparents and three major families from the two other pairs of grandparents. Figure 1 provides an illustration of the family structure of the seven major families.

Both F1- and F2-families were produced by multiple ovulation and embryo transfer. Embryo's were collected in one to 14 flushings per donor. On average 4.2 embryo's were collected per flushing, ranging from 0 to 13. F2-embryos were transferred to randomly assigned Boran recipients in groups per flushing over a three-day period, but because of differences in gestation length calves were born over a period of three weeks. Calves were born all year round on the ILRI ranch (Kapiti Plains Estate), which is a trypanosomosis free environment and were weaned at eight months of age. They were moved from the ranch to the central facilities at ILRI, Nairobi when they were 10 months of age. In total 23 groups were formed, containing 3 to 18 animals each. Group composition remained constant throughout the experiment at ILRI. The F2 animals, both at Kapiti and in Nairobi, were on a complete diet, i.e. no restriction, artificial starving or any food deficiency was applied.

Challenge and data collection

After two months adaptation to the experimental facility, all 12 months old animals (214 F2-animals plus 6 Boran and 6 N'Dama for control) were challenged in the same groups as in which they were weaned. The challenge was performed by bites of eight infected tsetse flies (*Glossina morsitans centralis*), with a different fly for each animal. The flies were carrying a with a single clone (IL 1180 (Geigy et al., 1973)) *Trypanosoma congolense*, mimicking a light to moderate infection pressure. Body weight was measured twice each recording day, and the average recorded for that day. If the difference between both measurements was more than 2kg, a third measurement was taken and the two closest measurements were used to calculate the average. PCV was measured by centrifugal techniques (Dargie et al., 1979). Parasites were visualised in jugular venous blood samples by dark ground phase contrast examination of buffy coats (Murray et al., 1977). Unit of measurement was the number of parasites counted in the blood sample. Body weight, PCV, and parasitaemia were recorded weekly. Measurements were taken on the same day within group, but between groups the number of days between measurements was not always equal. The recordings started three weeks preceding infection, with the last recording 150 days post infection. Therefore, there were on average 25 (ranging from 23 to 26 between groups) observations on body weight, PCV, and parasite levels per animal that finished the experiment. Those animals, in which clinical examination suggested the infection had

become life threatening, and/or PCV fell below 12, were treated with medication and recording for those animals was stopped. The observations that were subsequently not available as a result of treatment were treated as missing. In total 35 of the F2 animals needed medical treatment prior to the end of the recording period, with day of treatment ranging from 44 to 145.

Trait definition

A commonly used definition of trypanotolerance is that trypanotolerant animals are capable of efficiently recovering their PCV after infection and of gaining weight at the same rate as uninfected animals (Paling & Dwinger, 1993; Trail *et al.*, 1992; Murray *et al.*, 1991). Apart from the control of anaemia and the ability to gain weight following infection, the control of parasites is also assumed to be a feature of trypanotolerance. However, there is evidence that parasite control is different from the other aspects of trypanotolerance (Trail *et al.*, 1991, Paling *et al.*, 1991). Traits have been derived from phenotypic observations for PCV, body weight and parasitaemia in F2-animals. Table 1 contains a list of all traits defined and their abbreviations.

“Average Value” is the average level for PCV or body weight over the 150 days post infection. “Starting Value” is an average of three or four pre-challenge observations, the observation on the day of infection not taken into account. The average was taken in order to reduce the measurement error. “Lowest Value” is the absolute lowest level recorded after infection. Some animals did not loose any weight following infection and for those the lowest value often was equal to one of the first observations. “Maximum Decrease” is the absolute decrease in PCV or body weight during challenge and is calculated as “lowest value” minus “starting value”. “Slope Decrease” is the linear trend in change in body weight or PCV until minimum level is reached and is calculated as “maximum decrease” divided by number of days until the minimum level was reached. Both in case of maximum decrease and slope decrease of body weight, the trait value was set to zero for those animals that did not show any decrease following infection. “Final minus Minimum” is the difference in PCV or body weight between minimum and final level, and is calculated as the recording at day 150 post infection minus the “minimum level”. For animals that were treated with medication before day 150, day of treatment was considered the final observation. “Relative Change total period” is the relative change in body weight or PCV compared to the “starting level” and is calculated as “change total period” divided by “starting level”. Survival is the ability of an animal to survive the infection during the recording period without requiring

treatment with drugs. All animals got value 150, unless they were treated with trypanocidal drugs, in which case day of treatment was assigned as value. "Parasitaemia" is the average of the $10\log(\text{parasite count}+1)$ in the blood samples during the 150 days post infection. Finally, "Parasitaemia detected" represents the number of observation during which an animal is detected as parasitaemic. The proportion of times an animal was detected parasitaemic was also investigated, though the trait was highly correlated to parasitaemia (0.95), and is, therefore, not mentioned in the results.

Table 1. Traits and their abbreviations for traits related to PCV, body weight and parasitaemia, following infection

Trait	PCV	Body Weight
1. Average Value	AV-PCV	AV-BW
2. Starting value	ST-PCV	ST-BW
3. Lowest Value	LV-PCV	LV-BW
4. Maximum Decrease	MD-PCV	MD-BW
5. Slope Decrease	SD-PCV	SD-BW
6. Final minus Minimum	FM-PCV	FM-BW
7. Relative Change total period	RC-PCV	RC-BW
8. Survival		SURV
9. mean $\log(\text{Parasite count} + 1)$		LOGPAR
10. number of times Parasitaemia Detected		PARADET

Model

Due to the data structure, genetic parameters were estimated based on within and between family variances. Because embryos were randomly assigned to recipient dams, it was assumed there was no maternal effect within offspring of one pair of F1 parents. Dominance variance and variance due to common environment were assumed to be zero. A generalised linear model (PROC-GLM, SAS, 1997) was used to determine which effects to include in the model to estimate variance components. The fixed effect of group was significant for both body weight and PCV related traits ($p < 0.001$ in least squares analysis). The fitting of group effect adjusts the data for seasonal changes (Dwinger *et al*, 1994) due to

differences in feed quality, average daily temperature, etc. It also corrects for possible differences in challenge level that may have occurred over time. The fixed effect of sex was only significant for body weight related traits ($p < 0.001$ in least squares analysis). However, it was also included in the model for PCV and parasitaemia related traits in order to keep the model identical for all traits. The model used for variance component estimation for all traits, therefore, becomes:

$$Y_{ijk} = \mu + g_i + s_j + a_k + e_{ijk} \quad (1)$$

where g_i is the fixed effect of group i , s_j is the fixed effect of sex j , a_k is the random additive genetic component of animal k , and e_{ijk} is the random error term. An animal model was applied and the traits were analysed using ASREML (Gilmour et al., 1998). This program uses restricted maximum likelihood techniques including the average information matrix as second derivatives in a quasi-Newton procedure (Gilmour, et al., 1998) with a convergence criterion of $1 * 10^{-8}$. For checking consistency of heritability estimates, model (1) was also applied in a least squares analysis using SAS (1997), but as a sire model. Heritability estimates obtained with the least squares analysis were similar to the maximum likelihood estimates and are therefore not presented in the results section.

Because phenotypic observations were only available on the F2 animals, the F1 should be considered the base generation. The estimated genetic variance in the F2 population thus is equal to the genetic variance among the F1 parents. Animals in the F1 generation are heterozygous for more alleles than expected in a normal outbred population, due to linkage disequilibrium between loci. The linkage disequilibrium variance appears in the within family (i.e. Mendelian segregation) variance. This leads to underestimation of the genetic variance in the F1 and subsequent generations. It also leads to a slight underestimation of the average genetic variance in the pure bred parental breeds, due to inflation of residual variance.

Multivariate analysis

Subsequent to the estimation of variance components and heritabilities, a multivariate analysis using maximum likelihood techniques was performed (also using ASREML) to estimate phenotypic correlations between all traits considered. Genetic correlations could

not be determined in many cases due to the limited size of the data set, and are therefore not presented in the results.

Results

Phenotypic description of the data

There is substantial variation in the degree to which F2 animals are affected by trypanosomosis. To illustrate this, 12 F2 animals were selected based on their performance related to body weight change (RC-BW), anaemia control (MD-PCV) and parasite clearance (LOGPAR) following infection: the six highest, and the six poorest performing animals for each trait. For each of the categories of animals (i.e. F2 high performing, F2 poor performing, N'Dama, and Boran), the average performance was determined for each of the traits identified and reproduced in Figures 3a to i. The first column of figures (Figures 2a, d, and e) represents the change in body weight, PCV, and log parasite count for F2 animals that were selected on their performance for RC-BW. The second column of figures (Figures b, e, and h) show performance of F2 animals selected on MD-PCV, and the last column of figures (Figures c, f, and i) show performance of F2 animals selected on the average of the log of their parasite count. The performance of the N'Dama and Boran are in each column of figures. In the figures, the last recording day of F2-animals that needed medical treatment before the end of the recording period is indicated below the horizontal axis. Five of the Boran needed medical treatment before the end of the recording period, one at day 64, one at day 78, one at day 92, and two at day 127, and were subsequently removed from the experiment. These days of treatment are not indicated in the figures. Because not all animals within a category (e.g. N'Dama, poor performing F2) were challenged in the same group, not all observations were taken on the same days post infection. Within category the averages were taken for those days where at least five out of the six animals (should have) had an observation (days after treatment were considered as days as well). This has resulted in difference in day at which the average is calculated across categories, but this will not affect the differences between categories.

Figure 2a shows the average course of change in body weight following infection. The F2 animals in this figure were selected on their (dis)ability to grow following infection (RC-BW). Four of the poor performing F2 animals needed medical treatment before the end of the recording period. The initial body weight of the high performing F2 animals is in

between those of the N'Dama and the Boran, where that of the poor performing F2 animals is just above that of the Boran. Interestingly, the high performing F2 animals are able to continue growing following infection with the largest increase in body weight during the first three weeks following infection, whereas the N'Dama do not start to grow until around day 45. During the entire recording period, the high performing F2 animals show a larger growth rate than the N'Dama, and at the final recording day their body weight has reached the same level as that of the last Boran that managed to survive without treatment with drugs. The body weight of the poor performing F2 animals starts to drop around day 25 post infection and becomes lower than that of the high performing F2 animals around day 50. The Boran animals continue to grow until around day 30, and subsequently start to lose weight. The course of the change of body weight in the Boran is a bit hard to follow because of the removal of five animals during the recording period, which affects the average.

Difference with F2 animals that were selected on the lowest/highest values of maximum decrease in PCV following infection in Figure 2b are that the high performing F2 animals show a lower growth rate compared to figure 2a, now comparable to that of the N'Dama. Still the average body weight of these F2 animals is about 15 kg higher than in the N'Dama. The poor performing F2 animals in this figure again have initial body weight at comparable level to that of the Boran and about 20kg higher than the of the high performing F2 animals. The poor performing F2 animals are able to continue growing until day 50, after which they start to lose weight and have lower body weight than the high performing F2 animals from approximately day 75 onwards. In contrast to Figure 2a, only two poor performing F2 animals needed medical treatment before the end of the recording period. In Figure 2c the F2 animals were selected on the average parasite count. Difference in body weight at onset of the challenge period between high and poor performing F2 animals is minimal. The average initial body weight for the high performing F2 animals is about the same in all three Figures. Both high and poor performing F2 animals lose weight following infection. The body weight of the high performing F2 animals stabilises around day 100, whereas that of the poor performing F2 animals continues to decrease. Three poor performing animals needed medical treatment prior to the end of the recording period.

Figure 2d shows the course of change in PCV following infection for N'Dama, Boran, and F2 animals that were selected on RC-BW. For all four categories PCV started to drop almost immediately after infection. The drop in the N'Dama and high performing F2 animals seems to be a bit earlier and more severe at first than in the Boran and poor performing F2 animals. However, from approximately day 25 onwards, the N'Dama and

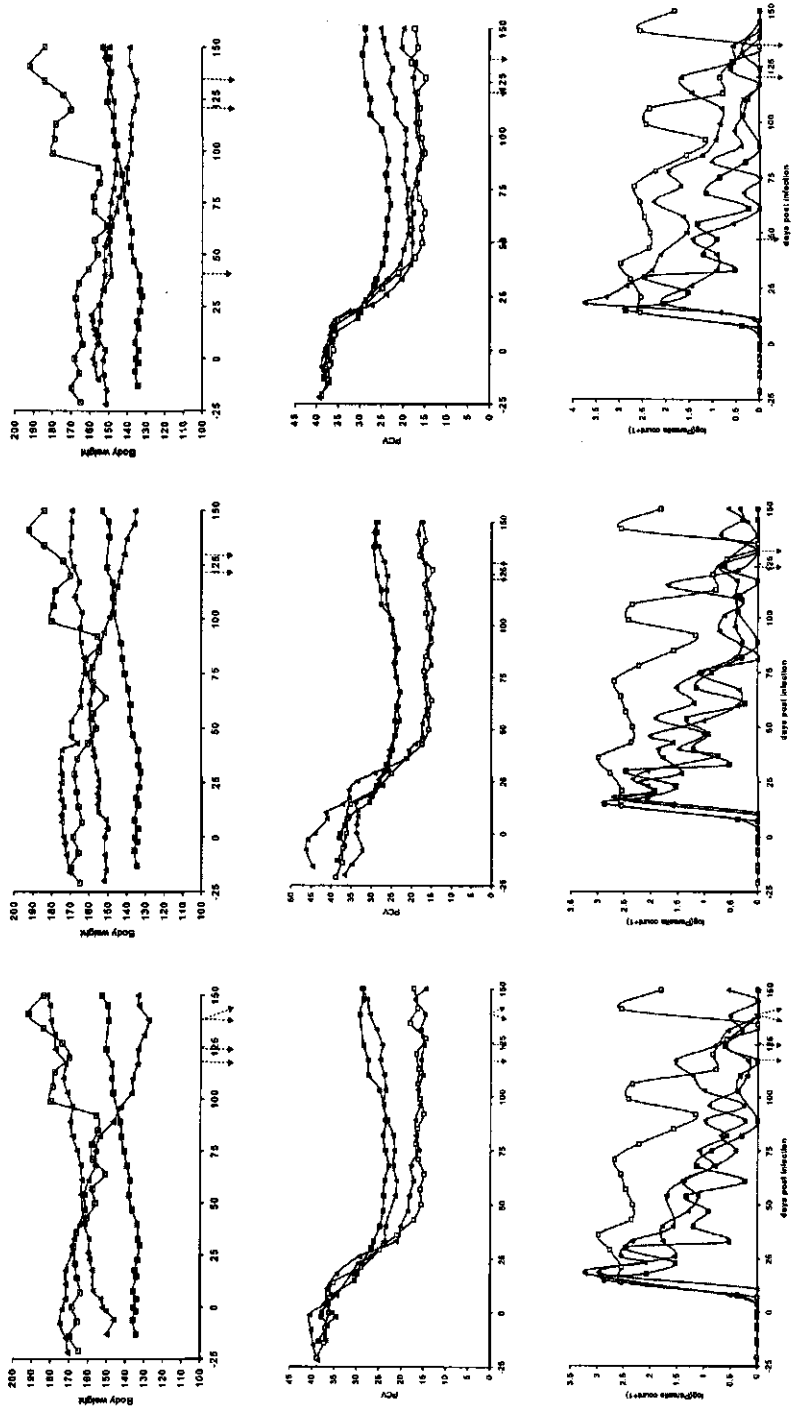


Figure 2. Average course of body weight change (a, b, c), change in PCV (d, e, f), and the log of the number of parasites in the blood samples (g, h, i) following an infection with *Trypanosoma congolense* for F2 animals that were selected on their performance for growth following infection (RC-BW, fig. a, d, g), maximum decrease in PCV (MD-PCV, fig. b, e, h) and number of parasites (LOGPAR, fig. c, f, i). Plots represent the average of respectively six high performing F2 animals (\blacktriangle), six poor performing F2 animals (Δ), six N'Dama (\blacksquare) and six Boran (\square) animals. The arrows in the x-axis indicate the last observation of F2 animals that have been treated with medication and subsequently removed from the experiment. Most Boran animals were treated with medication: one on day 71, 85, 99, and two animals on day 134, so only one Boran animal survived throughout the entire recording period.

high performing F2 animals start to stabilise their PCV values, and manage to do so from approximately day 50 onwards, whereas the poor performing F2 animals and the Boran continue to drop in PCV until day 50, after which they also manage to stabilise their PCV. The stabilised PCV in the poor performing F2 animals and the Boran is much lower (around 15) compared to the high performing F2 animals (around 20) and the N'Dama (around 23), and very close to the critical value of 12, below which the anaemia becomes life threatening and the animals were removed from the experiment. The N'Dama and high performing F2 animals start to recover their PCV from approximately day 70 (F2) and day 85 (N'Dama) onwards. None of the animals is able to fully recover their initial PCV within the recording period.

Differences with F2 animals that were selected on their MD-PCV level in Figure 2e after day 50 are small. High performing F2 animals now show an even more similar course of change in PCV following infection as the N'Dama, whereas poor performing F2 animals show a very similar course of change in PCV as the Boran. However, prior to day 50 there are some distinct differences in Figure 2e compared to 2d. The poor performing F2 animals start at a much higher PCV level than in figure 2d. However, that is what they were selected on. The minimum PCV level remains the same, otherwise they would have been removed from the experiment, so an increased difference in initial versus minimum PCV can only exist by a higher value for initial PCV. More interesting is that high performing F2 animals only show a slightly lower initial PCV level compared to Figure 2d, and that they subsequently show a very similar pattern to that of the N'Dama. The PCV pattern of the N'Dama following infection may be an upper limit in this F2 cross. In Figure 2f are PCV patterns of N'Dama, Boran, and F2 animals that were selected on high or low average parasite count. Until day 25 following infection PCV patterns are very similar. Then the N'Dama starts to stabilise their PCV level, whereas both groups of F2 animals and the Boran continue to drop in PCV. The poor performing F2 animals show a drop in PCV slightly before the Boran and the high performing F2 animals. The Boran and the high performing F2 animals also start to stabilise their PCV just before the poor performing F2 animals, around day 40. From approximately day 50 onwards the PCV of the Boran is slightly lower than that of the poor and high performing F2 animals, the last being able to start recovering PCV from approximately day 100 onwards, at the same speed, though at lower level than the N'Dama.

Figure 2g shows the average course of change in log (parasite count) during the recording period. The F2 animals were selected on their performance with regards to RC-

BW. Most important feature of parasitaemia in case of trypanosomosis in general are the repeated peaks in the parasite count, which occur due to the ability of the parasite to change antigen within the host, and thus induce a new parasitaemia peak unless the previous is cleared completely. All animals showed parasitaemia during second week following infection. The difference in day of onset of parasitaemia in the figure is caused by the difference in recording day across groups in which the animals were challenged. The course of $\log(\text{parasite count})$ of the Boran is influenced by the fact that animals were removed from the experiment during the recording period, but it is still obvious that on average they have a problem controlling the parasite number before the next peak presents itself. Until day 64 all Boran are still present. The N'Dama and the high performing F2 animals are able to considerably decrease their parasite load during the recording period. The poor performing F2 animals are also better able to control their parasite load compared to the Boran, though they still show a much increased peak around day 120, around which four of the six animals needed medical treatment.

Differences with F2 animals that were selected on their MD-PCV level in Figure 2h are that both poor and high performing F2 animals are less able to control parasite numbers. However, the high performing F2 animals show a more drop in parasite count after the second parasite wave. Both high and poor performing F2 animals show a distinct decrease in size of peaks across the recording period. Also in this figure the poor performing F2 animals show a relatively large peak around day 120, though more distinct than in Figure 2g, where is was a rather broad peak. Again, around this peak two a poor performing F2 animals needed medical treatment. The F2 animals in Figure 2i were selected on their average parasite count. The poor performing F2 animals show a high initial peak, and are very poorly able to control parasite burden across the recording period. One poor performing F2 animals needed medical treatment at day 48, where two others again needed medical treatment around day 120, where a distinct peak again is visible. The high performing F2 animals show parasitaemia patterns that are comparable to that of the N'Dama, apart from the fact that their initial peak was lower, and the second peak occurred approximately 20 days after the second peak in the N'Dama.

Descriptive statistics

Descriptive statistics for all traits considered in the univariate analysis are presented in Table 2. Some of the traits are not normally distributed but skewed to the right (e.g. MD-BW) or to the left (e.g. SURV). On average minimum body weight was reached 61 days

post infection while some of the animals already reached the minimum weight after 7.4 days. On average minimum PCV was reached 72 days post infection. The animal with the largest decrease in PCV reached lowest PCV at day 59, while the animal with the smallest decrease in PCV reached its lowest PCV at day 68. Starting bodyweight

Table 2 Minima, maxima, means, adjusted phenotypic standard deviations, and R² for the model including the fixed effects for all traits on 214 F2 animals

Trait*	Min.	Mean	Max.	Phen. sd.	R ²
AV-PCV	17.7	23.9	32.7	3.00	0.35
ST-PCV	29.5	38.5	51.5	3.36	0.29
LV-PCV	10.9	16.4	27.1	2.92	0.26
MD-PCV	-36.0	-20.9	-8.13	3.82	0.32
SD-PCV	-0.64	-0.30	-0.12	0.11	0.13
FM-PCV	0.00	4.97	15.6	3.87	0.29
RC-PCV	-0.73	-0.44	-0.04	0.16	0.24
AV-BW	80.6	150	230	26.21	0.39
ST-BW	91.8	154	242	24.54	0.41
LV-BW	74.0	140	218	25.26	0.38
MD-BW	-60.7	-14.3	0.00	12.6	0.30
SD-BW	-0.79	-0.15	0.00	0.13	0.22
FM-BW	-1.33	8.10	47.3	8.87	0.17
RC-BW	-0.30	-0.04	0.29	0.13	0.30
SURV	44.0	144	150	19.7	0.17
LOGPAR	0.19	1.02	1.93	0.34	0.53
PARADET	3.00	11.5	22.00	3.79	0.51

* for abbreviations see Table 1.

was higher than average body weight during the recording period, indicating that on average the F2 animals lost, rather than gained weight following infection. The same but more severely can be observed in PCV. Growth rate ranged from 30% loss, to 29% gain over the

Table 3 Phenotypic correlations for traits related to PCV following infection (above diagonal) and for traits related to bodyweight (below diagonal), and for survival and parasitaemia, with standard errors between brackets, estimated through multivariate analyses for 214 F₇-animals

	AV	ST	LV	MD	SD	FM	RC	SURV	LOGPAR	PARADET
AV	-	0.53 (0.09)	0.84 (0.03)	0.12 (0.09)	-0.19 (0.09)	0.12 (0.03)	0.68 (0.04)	0.20 (0.09)	-0.22 (0.07)	-0.07 (0.07)
ST	0.92 (0.01)	-	0.29 (0.09)	-0.62 (0.06)	-0.28 (0.09)	0.11 (0.02)	0.02 (0.09)	0.08 (0.09)	-0.12 (0.07)	0.02 (0.07)
LV	0.96 (0.01)	0.85 (0.02)	-	0.43 (0.08)	0.06 (0.09)	0.06 (0.04)	0.74 (0.04)	0.45 (0.06)	-0.20 (0.07)	0.09 (0.07)
MD	0.04 (0.11)	0.00 (0.11)	0.19 (0.10)	-	-0.11 (0.02)	-0.11 (0.02)	0.49 (0.07)	0.14 (0.09)	-0.06 (0.07)	0.05 (0.07)
SD	0.07 (0.11)	0.05 (0.11)	0.07 (0.11)	0.06 (0.11)	-	-0.30 (0.06)	-0.08 (0.08)	0.25 (0.08)	0.04 (0.07)	0.25 (0.07)
FM	0.41 (0.09)	0.23 (0.10)	0.41 (0.09)	0.45 (0.10)	0.03 (0.11)	-	-0.47 (0.09)	-	-0.22 (0.07)	0.00 (0.12)
RC	0.34 (0.07)	-0.02 (0.08)	0.46 (0.06)	0.17 (0.10)	-0.36 (0.09)	0.07 (0.11)	-	0.44 (0.07)	-0.32 (0.07)	-0.03 (0.07)
SURV	0.32 (0.07)	0.18 (0.08)	0.24 (0.08)	0.00 (0.11)	-0.13 (0.10)	-	0.23 (0.08)	-	-0.23 (0.07)	0.31 (0.07)
LOGPAR	-0.16 (0.07)	-0.07 (0.07)	-0.15 (0.07)	-0.10 (0.11)	0.29 (0.10)	-0.09 (0.11)	-0.23 (0.07)	-0.23 (0.07)	-	0.71 (0.04)
PARADET	0.06 (0.07)	0.06 (0.07)	0.01 (0.07)	-0.02 (0.11)	-0.27 (0.10)	0.00 (0.12)	-0.09 (0.07)	0.31 (0.07)	0.00 (0.11)	-

† for abbreviations see Table 1

entire recording period, with an average of 4% weight loss. None of the animals was able to completely recover their initial PCV, which is illustrated by the difference in the maximum decrease in PCV (MD-PCV) and the PCV recovery (FM-PCV). For PCV the relative change ranged from 73% to 4% PCV loss, with an average of 44%. During the 150 days post infection 13.5% of the F2-animals were treated with drugs, the first on day 44 and the last on day 145. On average an animal was detected parasitaemic in 45% of the observations, ranging from 10 to 80%. The model (1) accounted for 13% for the slope of decrease in PCV (SD-PCV), to 53% of the phenotypic variation for the mean of the log(parasite count +1).

Phenotypic correlations between all traits considered are presented in Tables 3 and 4. In Table 3 are the phenotypic correlations within the group of traits related to body weight or PCV following infection. Also in this table are the correlations between survival, average log(parasite count +1), and the number of times an animal was detected as parasitaemic, and the rest of the traits. Phenotypic correlations range from -0.62 (between ST-PCV and MD-PCV) to 0.96 (between AV-BW and LV-BW). Average PCV is highly correlated with the lowest value (0.84), but not with the maximum decrease or the slope of the decrease (-0.19). The initial value for PCV is moderately correlated with the maximum decrease in PCV following infection (-0.62). This correlation is negative, indicating that a higher initial value for PCV is associated with a greater decrease in PCV following infection. This is due to the nature of the trait: there is a lower threshold for PCV to which an animal can drop, below that it will be treated with medication and removed from the experiment. The relative change in PCV following infection is highly correlated with the lowest PCV reached (0.74) and not to initial PCV (0.02). Survival is moderately correlated with the lowest value of PCV (0.45). There is no correlation between survival and PCV recovery due to the fact that all animals that were treated before the end of the recording period automatically got a value of 0.0 for PCV-FM, because in those cases the final value was equal to the minimum value. The log of average number of parasites is uncorrelated with maximum decrease (-0.06) or slope of the decrease in PCV (0.04), but moderately correlated with the average PCV (-0.22), the PCV recovery (-0.22), the relative change in PCV (-0.32), and survival (-0.23). Correlations between the PCV related traits and the number of times an animal is detected parasitaemic generally are low, except for the correlation with the slope in decrease of PCV following infection (0.25) and with survival (0.31). Both parasitaemia traits are highly correlated (0.71).

The traits RC-BW and ST-BW are uncorrelated (-0.02), indicating that growth following infection is independent of the starting weight of the animal. There is a high

correlation between the average body weight during recording period and the lowest value reached (0.96), and between the starting weight of an animal and the lowest value reached (0.85). Also the correlation between growth following infection (RC-BW) and the maximum decrease in body weight is high (0.88). The correlations with the slope of the decrease in body weight are low for some traits (average body weight, starting weight, lowest value, and maximum decrease following infection). However, for other traits the correlations with the slope of the decrease in body weight are moderate (the relative change during the recording period, survival, parasite count, or number of times detected parasitaemic). The relative change in body weight following infection is moderately correlated with average body weight (0.34), the lowest value reached (0.46), and the slope of decrease in body weight (-0.36), but uncorrelated with the number of times detected parasitaemic (0.00). Correlations between survival and the recovery in PCV and body weight following infection do not exist because of the fact that there is no variation for survival in combination with observations of these PCV and body weight related traits that are different from zero.

Table 4 Phenotypic correlations (standard errors) between Body Weight related traits¹ and PCV related traits¹, estimated through multivariate analyses for 214 F₂-animals

	AV-BW	ST-BW	LV-BW	MD-BW	SD-BW	FM-BW	RC-BW
AV-PCV	0.34 (0.08)	0.11 (0.08)	0.39 (0.07)	0.22 (0.10)	-0.14 (0.11)	0.33 (0.09)	0.66 (0.05)
ST-PCV	0.20 (0.09)	0.12 (0.09)	0.18 (0.08)	0.00 (0.11)	0.00 (0.11)	0.10 (0.11)	0.16 (0.09)
LV-PCV	0.36 (0.07)	0.13 (0.08)	0.39 (0.07)	0.26 (0.10)	0.08 (0.11)	0.43 (0.09)	0.64 (0.06)
MD-PCV	0.05 (0.08)	0.00 (0.08)	0.07 (0.08)	0.27 (0.11)	0.17 (0.11)	0.21 (0.10)	0.27 (0.09)
SD-PCV	0.07 (0.09)	0.16 (0.08)	-0.05 (0.08)	0.07 (0.11)	0.59 (0.07)	0.03 (0.11)	-0.31 (0.08)
FM-PCV	0.49 (0.08)	0.02 (0.01)	0.02 (0.01)	0.46 (0.09)	-0.02 (0.11)	0.00 (0.11)	-0.18 (0.07)
RC-PCV	0.34 (0.07)	0.09 (0.08)	0.39 (0.07)	0.25 (0.10)	-0.25 (0.10)	0.12 (0.11)	0.74 (0.04)

¹ for abbreviations see Table

In Table 4 are the phenotypic correlations between the PCV and the body weight related traits. The average value for PCV following infection is moderately correlated with the relative change in body weight following infection (0.66). The starting level of PCV is only lowly to moderately correlated with the traits relating to the change in body weight following infection (0.00 with SD-BW and MD-BW, to 0.20 with AV-PCV). The correlation between the lowest value reached for PCV and the relative change in body

weight following infection is reasonably high (0.64). The maximum decrease in PCV following infection is only lowly to moderately correlated with the body weight related traits (0.00 with ST-BW to 0.27 with MD-BW and RC-BW). The slope of the decrease in PCV following infection is moderately correlated with the slope of the decrease in body weight (0.59), and with RC-BW (-0.31). PCV recovery (FM-PCV) is moderately correlated with average body weight (0.49), and with the maximum decrease in body weight (0.46), but uncorrelated with the recovery in body weight. Relative change in PCV following infection is highly correlated with the relative change in body weight (0.74).

Genetic analyses

In Table 5 are the heritability estimates with standard errors for all traits considered in the analysis. The heritability for ST-PCV is very high (0.88), in contrast to the heritability for ST-BW (0.08). Low heritabilities are also found for most body weight related traits and for FM-PCV, RC-PCV, and number of times detected parasitaemic. AV-PCV, LV-PCV and SD-PCV, survival, and LOGPAR have moderate heritabilities. Standard errors of the estimates are high, which is due to the limited size of the data set.

Table 5 Heritability estimates for all traits (standard errors) on 214 F2 animals

Traits *	h^2 (s.e.)	Traits †	h^2 (s.e.)
AV-PCV	0.31 (0.19)	AV-BW	0.11 (0.11)
ST-PCV	0.88 (0.24)	ST-BW	0.08 (0.11)
LV-PCV	0.26 (0.17)	LV-BW	0.07 (0.09)
MD-PCV	0.14 (0.13)	MD-BW	0.12 (0.11)
SD-PCV	0.18 (0.15)	SD-BW	0.09 (0.09)
FM-PCV	0.01 (0.08)	FM-BW	0.06 (0.09)
RC-PCV	0.09 (0.10)	RC-BW	0.14 (0.12)
SURV	0.19 (0.14)	LOGPAR	0.18 (0.15)
PARADET	0.09 (0.11)		

* for abbreviations see Table 1

Discussion

The current paper presents properties of traits related to of trypanotolerance of 214 F2 cattle (N'Dama × Boran cross) that were challenged with *Trypanosoma congolense* in an experimental setting. Results from the present study are unique in that an experimental data set of the current size, with such a wide range of observations related to trypanotolerance has not been presented before. The results indicate that there is a moderate correlation between level of PCV and growth following infection. But even under these highly controlled conditions the best R2 between PCV and body weight traits indicated that only 55% of the variation in RC-BW was predicted by RC-PCV (and vice versa). The parasitaemia related traits are low to moderately correlated with the PCV and body weight related traits. Results also indicate that there is large variation within the traits considered in the F2 population.

The performance of F2 animals ranges between the average levels of the N'Dama to those of the Boran animals. Table 2 shows that none of the F2-animals were able to maintain or to recover their original PCV level during the 150 day recording period nor the pure-bred N'Dama. This is in contrast to results of Paling et al. (1991), where the N'Dama recovered their original PCV levels within two to four months following an infection with *Trypanosoma congolense* under experimental settings. Murray et al., (1981) reported a severe drop in PCV in both Zebu and N'Dama cows under natural field challenge, and so did Roberts and Gray (1973) for young first infection N'Dama and Zebu animals, though both studies were based on very limited numbers of animals. In the present study PCV started to decrease in all groups of animals (i.e. F2, N'Dama and Boran) approximately at the day the first parasite peak occurred. This feature was already reported by Murray et al., (1981). The lowest PCV level in the F2 animals ranged from 10.7 to 27.1 with an average of 16.4, which is lower than results by Trail et al., (1992) who reported an average lowest PCV value of 23.0 for pure bred N'Dama cattle that were detected parasitaemic. The population of N'Dama referred to by Trail et al., (1992) was exposed to a first infection, being a natural challenge of medium infection pressure, in Gabon. Because the F2 animals are a cross between trypanotolerant N'Dama and trypanosusceptible Boran, it is according to expectation that the average of the F2 is lower than of pure bred N'Dama cattle.

Similar differences can be observed from the average PCV values, which is equal to 29.1 in Trail et al., (1992), and are ranging from 17.7 to 32.7, with an average of 23.7 in this study. It is reported that N'Dama that are infected with trypanosomes are able to grow at the

same rate as uninfected animals, provided they can maintain their PCV values (Trail et al., 1992a). However, in the present study no uninfected animals were included for comparison, so no conclusions with that regard can be drawn. Paling et al. (1987, reviewed by Trail et al., 1991) and Trail et al., (1991) could not demonstrate the presence of a correlation between mean parasitaemia and PCV values in the N'Dama, where Trail et al., (1991a) demonstrated that parasitaemia had a significant effect on growth. However, in the present study there is a significant correlation between both average log(parasite count) and proportion of times an animal in detected parasitaemic, and average PCV, the difference between minimum and final PCV, the relative change in PCV, and survival. This correlation is determined in an F2 population and it may, therefore, be possible that the correlation is established due to gene effects originating from the Boran.

In the present study it was assumed that the ability of an animal to continue gaining weight despite being infected is a good indicator of the ability of an animal to control the trypanosome infection. The correlations between growth following infection (RC-BW) and average, lowest, and relative change in PCV were moderate to high (0.64 to 0.74); Table 4). However, it is important that growth following infection is measured over a long enough period, because the poor performing F2 animals in Figure 2b (selected on large difference between starting value and minimum value for PCV) continue to gain weight following infection until approximately day 50. After this period of growth there is an almost linear decrease in body weight until the end of the recording period. A trait that accurately describes the clearance of parasites has not been described yet, but it is likely that also for that trait the correlation with growth following infection will be high. The correlation of RC-BW with survival time was only low to moderate (0.23). The reason for this is that survival time in the current experiment is primarily determined by achieving a defined minimum PCV and that only 13.5% of animals were removed for treatment (deemed not surviving). Moreover, the fact that an animal survives provides little indication of how well an animal is coping with the infection. Comparing Figure 2a and 2d to Figures 2b and 2e, it becomes clear that the F2 animals that were selected on growth performance following infection show a higher relative growth rate (2a) than F2 animals that were selected for poor growth following infection. A higher growth rate results in a control and recover PCV that is close to that of the N'Dama (2d). Selection for control of PCV following infection results in lower weight gain (2b), and little better control in PCV (2e). However, the economic profit will be obtained from the growth of the animal (and the probability that those animals will also be able to continue to produce milk, remain reproductive, are more useful for draught power,

etc.), rather than from the ability to control PCV slightly better. Also Trail et al., (1991) have shown that growth following infection with *Trypanosoma congolense* is affected by both parasitaemia and low PCV levels. Van der Waaij et al. (2000) showed by simulation that selection on production under continuous infection pressure automatically would result in increase in resistance to that infection, provided the lack of resistance is limiting the production level. In conclusion, growth following infection seems a good and practically implacable indicator of multiple aspects of trypanotolerance.

One purpose of this paper was to define heritable traits that are related to trypanotolerance. Five traits with heritabilities of 0.2 or higher have been defined (Table 5). Trail *et al.* (1991a) estimated heritabilities on 148 one-year old N'Damas, using least squares estimates and a sire model, under practical production circumstances, and with medium infection rate. They estimated heritabilities for growth (0.39), lowest PCV reached (0.50) and average PCV (0.64). Rowlands *et al.* (1995) estimated a heritability of 0.34 for average PCV, 0.15 for growth (RC-BW) and 0.25 for lowest PCV reached (LV-PCV) on 936 zebu animals of 21 months to three years of age under practical production circumstances with high infection pressure. An animal model was applied using maximum likelihood techniques. The heritabilities estimated in the current study are comparable to the ones found by Rowlands et al. (1995) on zebu cattle.

In conclusion, there is substantial variation in the F2 generation for the traits defined in this study. If the results for the N'Dama in this paper could be generalised to a larger population of N'Dama, then creating an F2 and breeding from the most tolerant animals may be an option to quickly improve body weight and growth rate, while maintaining a good trypanotolerance level.

Acknowledgement

The authors would like to thank WOTRO (Netherlands Foundation for the Advancement of Tropical Research) for financial support. They also would like to thank Henrie gaThuo, Joel Mwakava, Moses Ogugo and Daniel Mwangi for their help in collecting and processing the data, all the personnel of the ILRI farm at Kapiti and Nairobi for taking care of the animals.

References

- Agyemang, K., Dwinger, R.H., Little, D.A. and Rowlands, G.J. 1997. Village N'Dama cattle production in West Africa. Six years of research in the Gambia. International Livestock Research Institute, Nairobi, Kenya, and International Trypanotolerance Centre, Banjul, The Gambia. 131 pp.
- Dargie J.D., P.K. Murray, Murray M., Grimshaw W.R.I. and McIntyre W.I.M. 1979. Bovine trypanosomiasis: the red cell kinetics of N'Dama and Zebu cattle infected with *Trypanosoma congolense*. *Parasitology* 78: 271-286.
- Dwinger, R.H., Agyemang, K., Snow, W.F., Rawlings, P., Leperre, P. and Bah, M.L. 1994. Productivity of trypanotolerant cattle kept under traditional management conditions in the Gambia. *Veterinary Quarterly* 16:81-86.
- Falconer, D.S. 1989. *Introduction to Quantitative Genetics*. Third edition. pp. 233-235. Longman Scientific & Technical, Harlow, England.
- d'Ieteren, G.D.M., Authié, E., Wissocq, N. and Murray, M. 1998. Trypanotolerance, an option for sustainable livestock production in areas at risk from trypanosomosis. *Rev. sci. tech. Off. int. Epiz.* 17: 154-175.
- Gilmour, A.R., Cullis, B.R., Welham, S.J. and Thompson, R. ASREML. 1998. Rothamsted Experiment Station Harpendem, UK.
- ILRAD Reports, October 1989: N'Dama Cattle: Managing Africa's Genetic Recourses
- Jordt, T., Mahon, G.D., Touray, B.N., Ngulo, W.K., Morrison, W.I., Rawle, J. and Murray, M. 1986. Successful transfer of frozen embryos from the Gambia to Kenya. *Trop. Anim. Hlth. Prod.* 18: 65-75.
- Kraaijeveld, A.R., Limentani, E.C. and H.C.J. Godfray. 2000. Basis of the trade-off between parasitoid resistance and larval competitive ability in *Drosophila melanogaster*. *Proc. R. Soc. Lond.* 268: 259-261.
- Murray, C., Murray, M., Murray, P.K. Morrison, W.I., Pyne, C., and McIntyre, W.I.M. 1977. Diagnosis of African trypanosomiasis in cattle. Improved parasitological and serological techniques. In: International Scientific Council for Trypanosomiasis Research and Control 15th Meet., the Gambia. OAU/STRC. Publication No. 110, p. 247-254.

- Murray, M., Clifford, D.J., and Gettinby, G. 1981. Susceptibility to African trypanosomiasis of N'Dama and Zebu cattle in an area of *Glossina morsitans submorsitans* challenge. *Vet. rec.* 109: 503-510.
- Murray, M., Trail, J.C.M., Davis, C.E. and Black, S.J. 1984. Genetic resistance to African Trypanosomosis. *The Journal of Infectious Diseases* 149:311-319.
- Murray, M., Trail, J.C.M. and d'Ieteren, G.D.M. 1990. Trypanotolerance in cattle and prospects for the control of trypanosomiasis by selective breeding. *Rev. sci.tech. Off. Int. Epiz.*9: 369-386.
- Murray, M., Stear, M.J., Trail, J.C.M., d'Ieteren, G.D.M., Agyemang, K. and Dwinger, R.H. 1991. Trypanosomosis in Cattle: Prospects for Control. in: *Breeding for Disease Resistance in Farm Animals* (ed. J.B. Owen and R.F.E. Axford), pp. 203-223. Cab International, Wallingford, UK.
- Paling, R.W., Moloo, S.K., Scott, J.R., Gettinby, G., McOdimba, F.A. and Murray, M. (1991) Susceptibility of N'Dama and Boran cattle to sequential challenges with tsetse-transmitted clones of *Trypanosoma congolense*. *Parasite Immunology* 13, 427-445.
- Paling, R.W. and Dwinger, R.H. 1993. Potential of trypanotolerance as a contribution to sustainable livestock production in tsetse affected Africa. *The Veterinary Quarterly* 15:60-67.
- Roberts, C.J., and Gray, A.R. 1973. Studies on trypanosome-resistant cattle. II. The effect of trypanosomiasis on N'Dama, Muturu and Zeby cattle. *Trop. Anim. Hlth.* 5: 220-233.
- Rowlands, G.J., Woudyalew, M., Nagda, S.M., Dolan, R.B., d'Ieteren, G.D.M. 1995. Genetic variation in packed red cell volume and frequency of parasitaemia in East African Zebu cattle exposed to drug-resistant trypanosomes. *Livestock Production Science* 43:75-84.
- SAS version 6.12. 1997. SAS Institute Inc., Cary, USA.
- Soller, M. & Beckman, J.S. 1987. *Toward an understanding of the genetic basis for trypanotolerance in the N'Dama cattle of West Africa*. Consultation Report FAO, Rome.
- Taylor, K.A. 2000. Immune response of cattle to African trypanosomes: protective or pathogenic? *Int. J. Parasitol.* 28: 219-240.
- Teale, A.J. 1993. Improving control of livestock diseases: Animal Biotechnology in the consultative group on international agricultural research. *Bioscience* 43:475-483.
- Teale, A.J. 1997. Biotechnology: a key element in the CGIAR's livestock research programme. *Outlook on Agriculture* 26: 217-225.

- Trail, J.C.M., d'Ieteren, G.D.M., Feron, A., Kakiese, O., Mulungo, M., and Pelo, M. 1991. Effect of trypanosome infection, control of parasitaemia and control of anaemia development on productivity of N'Dama cattle. *Acta Tropica* 48:37-45.
- Trail, J.C.M., d'Ieteren, G.D.M., Maille, J.C. and Yangari, G. 1991a. Genetic aspects of control of anaemia development in trypanotolerant N'Dama cattle. *Acta Tropica* 48:285-291.
- Trail, J.C.M., d'Ieteren, G.D.M., Maille, J.C., Yangari, G. and Nantulya, V.M. 1992. Use of antigen-detection enzyme immunoassays in assessment of trypanotolerance in N'Dama cattle. *Acta Tropica* 50:11-18.
- Trail, J.C.M., d'Ieteren, G.D.M., Viviani, P., Yangari, G., and Nantulya, V.M. 1992a. Relationships between trypanosome infection measured by antigen detection enzyme immunoassays, anaemia and growth in trypanotolerant N'Dama cattle. *Veterinary Parasit.* 42: 213-223.
- Trail, J.C.M., d'Ieteren, G.D.M., Murray, M., Ordner, G., Yangari, G., Collardelle, C., Sauveroché, B., Maille, J.C. and Viviani, P. 1993. Measurement of trypanotolerance criteria and their effect on reproductive performance of N'Dama cattle. *Veterinary Parasitology* 45: 241-255.
- Trail, J.C.M., Wissocq, N., d'Ieteren, G.D.M., Kakiese, O. and Murray, M. 1994. Quantitative phenotyping of N'Dama cattle for aspects of trypanotolerance under field tsetse challenge. *Veterinary Parasitology* 55: 185-195.
- Van der Waaij, E.H., Bijma, P., Bishop, S.C., and Van Arendonk, J.A.M. 2000. Modeling selection for production traits under constant infection pressure. *J. Anim. Sci.* 78: 2809-2820.
- Wilks, S.S. 1938. The large sample distribution of the likelihood ratio test for testing composite hypotheses. *Ann. Math. Stat.* 9:60-62.

III Mode of expression of some QTL related to trypanotolerance in African cattle.

*An F2 experiment, consisting of a cross between N'Dama and Boran cattle, was set up at the International Livestock Research Institute (ILRI) in Nairobi, Kenya, for the detection of QTL related to trypanotolerance in cattle. An initial analysis revealed a number of chromosomes harbouring significant QTL. In mice there is evidence that the gene *Tir1*, relating to survival following a trypanosome infection, is genomically imprinted. In this paper these chromosomes are further analysed to determine the mode of expression of these QTL. QTL were detected for traits related to anaemia, body weight change, and parasitaemia following a trypanosome infection on nine different chromosomes. In total eight QTL have been detected, most of them with Mendelian expression. One QTL (for the recovery of PCV following infection on BTA19) showed evidence of genomic imprinting, with maternal expression. Size of effect of the QTL ranged from 5.8 percent of the phenotypic variance for initial body weight prior to infection on BTA2 and the recovery in PCV following infection on BTA7, to 10.4 percent for the lowest value of PCV reached following infection, on BTA2.*

Introduction

Results from a backcross experiment with a susceptible and a resistant strain to trypanosomosis in mice suggest genomic imprinting of a gene related to survival, located on BTA17 (Clapcott et al., 2000). Offspring from F1 fathers survived significantly longer compared to offspring from F1 mothers. The difference in survival seemed related to *Tir1*, which is located close to a region with known imprinted genes in mice. The fact that there seems to be evidence for genomic imprinting of a gene related to resistance to trypanosomosis in mice, was reason to check whether the comparative DNA region in cattle would also show imprinting in relation to trypanotolerance. In addition it could be investigated whether there would be more genes related to trypanotolerance in cattle that are subject to imprinting. Over the past few years, genomic imprinting is detected related to genes on several regions of the genome, mainly in mice and human (e.g. Schmidt, et al., 2000; Li et al., 1999; reviewed by Reik and Walter, 2001), but also in pigs (De Koning et al., 2000). Imprinted genes detected so far (in human, mice and other species) seem to be

This chapter has been submitted to the Journal of Animal Science in an adjusted version

E.H van der Waaij, J.A.M. van Arendonk, O. Hanotte, D.-J. de Koning and J. B. Gibson. Evidence for imprinting of QTL related to trypanotolerance in African cattle. Submitted to the Journal of Animal Science

especially related to growth and development, and related to some medical syndromes, but not so much to disease resistance.

Based on preliminary results (Hanotte et al., in preparation), several chromosomes were selected where evidence for presence of QTL was found, and genotypes for these chromosomes were subsequently re-coded in order to incorporate information on the breed of origin alleles. This paper focuses on determining the mode of expression of QTL related to trypanotolerance on several chromosomes in an F2 population of two African cattle breeds.

Material and Methods

Data and experiment.

Four Gambian N'Dama sires were mated to four Kenyan Boran dams resulting in multiple offspring per mating (with use of Multiple Ovulation and Embryo Transfer techniques). The N'Dama sires were transported to Kenya as embryo's (Jordt et al. 1986) and like the Boran dams, they were never infected with trypanosomes. Two N'Dama (ND8 and ND10) sires were full sibs. The other two N'Dama sires (ND7 and ND9) both came from different regions of the Gambia, and are therefore unlikely to be related. Three or four offspring (male and female) were selected per couple and mated to offspring of another couple to create an F2-population, again using MOET techniques. Recipient dams for embryos were assigned randomly. The F2-population, consisting of 181 animals, can be divided into 7 full sib families, varying in size from 22 to 40 animals. For more details on family structure, see Figure 1.

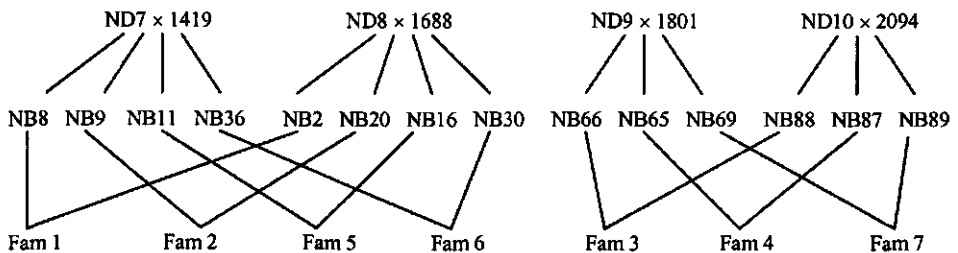


Figure 1. Schematic overview of the F2 family structure. All N'Dama's (ND) are males and all Boran (e.g. 1419) are females. ND8 and ND10 are full sibs. In the F1, NB8, NB9, NB16, NB30, NB66, NB65 and NB89 are females.

The F2-calves were born at Kapiti, which is an experimental farm outside Nairobi and a tsetse-free environment, and weaned in 23 groups at approximately eight months of age. These groups were transported to (tsetse-free) experimental facilities at ILRI headquarters in Nairobi in these groups at approximately ten months of age. At about 12 months of age animals, still in the same groups, were artificially challenged with *Trypanosoma congolense* through bites from eight infected tsetse flies with different flies for each animal. During 150 days post infection, body weight and Packed red blood Cell Volume (PCV) were measured at weekly basis on all infected animals, and the number of parasites was counted in blood samples using the dark ground phase buffy coat technique (Murray et al., 1977). Animals were not treated with medication unless their PCV dropped below 15, which was considered a life-threatening situation. Those animals were treated and observations on PCV, body weight and parasitaemia were no longer recorded on these animals. For a more extended description of the experiment, see Van der Waaij et al. (2001).

Traits

One of the most commonly used definitions of trypanotolerance is that infected trypanotolerant animals are capable of efficiently recovering from anaemia after infection and of gaining live weight at the same rate as uninfected animals (Paling & Dwinger, 1993; Trail et al., 1992; Murray et al., 1991). Another mechanism involved in resistance to trypanosomosis is the parasite clearance, which seems to be very little related to the control of anaemia and is therefore considered an independent mechanism (Trail et al., 1991). Van der Waaij et al. (2001) defined 17 traits based on these two mechanisms from the observations on the infected F2 cattle. In Table 1 is a list with abbreviations and brief descriptions for each of the traits.

Seven traits were related to PCV (a measure for anaemia), seven to body weight gain, two to parasitaemia, and one to survival. The traits related to PCV and body weight can roughly be divided into traits describing the decrease in PCV or body weight following infection, and traits describing the extent to which animals subsequently recovered. In addition, also PCV (and body weight) prior to infection were included. In case of premature termination of the recording because of life threatening low PCV level, values last recorded were considered as final observations for that animal when determining trait values.

The same definitions were used for PCV, as for body weight related traits. "Average Value" is the average level for PCV or body weight over the 150 days post infection, or less in the case that recording for that animal was stopped. "Starting value" is an average of the

three or four pre-challenge observations, the average was taken in order to reduce the measurement error, and was excluding the observation on day 0. "Lowest Value" is the absolute lowest level after infection during the recording period. "Maximum Decrease" is the absolute decrease in PCV or body weight after challenge and is calculated as "Lowest Value" minus "Starting value". "Slope Decrease" is the linear trend in daily change in body weight or PCV until minimum level is reached and is calculated as "Maximum Decrease" divided by number of days until the minimum level was reached.

Table 1. Trait definitions and their abbreviations for traits related to PCV, body weight and parasitaemia, following infection

Trait definition	PCV (%)	Body Weight (kg)
1. Average Value	AV-PCV	AV-BW
2. Starting value	ST-PCV	ST-BW
3. Lowest Value	LV-PCV	LV-BW
4. Maximum Decrease	MD-PCV	MD-BW
5. Slope Decrease	SD-PCV	SD-BW
6. Final minus Minimum	FM-PCV	FM-BW
7. Relative Change total period	RC-PCV	RC-BW
8. Survival		SURV
9. mean ln(Parasite count)		LNPAR
10. number times Parasitaemia Detected		PARADET

"Final minus Minimum" is the difference in PCV or body weight between minimum and final level, and is calculated as the final recording minus the "Minimum Level". Animals without observation on day 150 were treated as missing for this trait. The difference between initial and final trait value, "Change Total period", is calculated as the final value minus "Starting value". "Relative Change total period" is the relative change in body weight or PCV compared to the "Starting Level" and is calculated as "Change Total period" divided by "Starting level". Survival is the ability of an animal to survive the infection during the recording period without requiring treatment with drugs. All animals got value 150, unless they were treated with drugs in which case day of treatment was assigned as value. "ln(Parasitaemia)" is the average of the natural logarithm of the parasite

count in the blood samples during the 150 days post infection. Finally, "Parasitaemia detected" represents the number of times an animal is detected as parasitaemic. Except for survival, which is not normally distributed, outliers were detected and removed in case their trait values were beyond four SD from the trait average. In total six animals were excluded from the data for that reason, resulting in 175 F2 animals. All traits were analysed with the same number of animals in the data.

Model for fixed effect correction

Prior to the QTL analyses, phenotypes were pre-adjusted for fixed effects of group in which the animal was weaned and challenged, sex of the animal, and F1 family (1).

$$y_{ijkl} = s_i + g_j + f_k + e_{ijkl} \quad (1)$$

where s_i is the effect of sex ($i = 2$), g_j is the effect of group ($j = 12$), f_k is the effect of family ($k = 7$), and e_{ijkl} is the error component. Groups ranged in size from three to 15 animals. To prevent confounding with family effects, groups with less than 10 animals were combined with the group that was challenged next because groups were often challenged during overlapping periods. The effects and their significance were determined by generalised linear model procedure of SAS (SAS, 1997). Even though the effects were not significant in all cases, the same model was applied for pre-correcting all traits.

Genotypes and Map

Nine chromosomes (BTA1, BTA2, BTA3, BTA5, BTA7, BTA19, BTA20, BTA23, and BTA26) were selected for further analysis based on significant QTL detected in a line-cross analysis (Hanotte et al., in preparation), containing 7 to 15 genetic markers each. Genotypes of all animals needed to be traced back to the genotypes of the parent of origin, rather than the genotypes referring to their breed origin (NN, NB, BB) for these chromosomes in order to be able to do an imprinting analysis. Markers were generally not informative in all seven families. For each chromosome a linkage map was constructed using Cri-map (Green *et al.* 1990). The results were transformed into Haldane mapping distance (Haldane, 1919). An overview of markers used, distances between markers and information content at these marker positions are in Appendix 1. The map of BTA5 has a large gap (106 cM) for which no markers were typed. The chromosome can therefore be

considered as two linkage groups, however, it was analysed as one. The map for BTA26 is not shown because on that chromosome no QTL were detected.

QTL analysis

For all traits, interval mapping using regression methods was applied, following the line-cross approach proposed by Haley et al. (1994). Under the line-cross approach it is assumed that the two founder lines are fixed for alternative alleles at the QTL affecting the traits of interest, although they may share alleles at the marker loci. Using multi-marker information, four probabilities are calculated at 1 cM intervals along the genome. P_{11} is the probability are that an F2 animal inherited two N'Dama alleles, p_{22} that it inherited two Boran alleles, and p_{12} or p_{21} that it inherited one from each line (different subscripts according to parental origin; first subscript is paternally inherited allele). At every cM across the investigated chromosomes, an additive effect (a) and a dominance effect (d) are estimated using the regression of the phenotypes on a linear combination of the line origin probabilities:

$$y_j = m + ap_{aj} + dp_{dj} + e_j \quad (2)$$

where y_j is the trait score of animal j (adjusted for systematic effects), m is the population mean, a and d are the estimated additive and dominance effect of a putative QTL at the given location, p_{aj} is the conditional probability of animal j to carry two N'Dama alleles, rather than two Boran alleles (p_{11} - p_{22}), p_{dj} the conditional probability of animal j to be heterozygous (p_{12} + p_{21}), and e_j is the residual error. A detailed description of these methods is given by Haley *et al.* (1994).

The presence of a QTL where the allele from only one of the parents is expressed was detected following the procedures presented by (De Koning et al. (2000). The contribution of the parents was separated using the probability that the individual inherited an N'Dama allele from its father ($p_{pat} = [p_{11}+p_{12}]-[p_{22}+p_{21}]$) or from its mother ($p_{mat} = [p_{11}+p_{21}]-[p_{22}+p_{12}]$). The presence of imprinting was assessed by a saturated model that contained a paternal, maternal and dominance component to the positions of the best Mendelian and imprinted QTL

$$y_j = m + a_{pat}P_{pat,j} + a_{mat}P_{mat,j} + dp_{d,j} + e_j \quad (3)$$

The re-parameterisation allowed additional models to be fitted with exclusive paternal or maternal expression (single parent model):

$$y_j = m + a_{parent} P_{parentj} + e_j \quad (4)$$

Imprinting was tested in two ways: I) Following De Koning et al (2000), imprinting was inferred if only one of the parental contributions was significant and no dominance was present, using F ratios for the individual components of model (4). II) At the position of the best imprinted QTL, it was tested whether the full model (4) explained significantly more variance than a Mendelian model (2) at that position..

Chromosome-wise significance levels for the presence of a QTL against the H0 of no QTL were determined by a permutation test by running 10,000 permutations per trait of interest per chromosome (both for the Mendelian and the single parent model). For the test of a full model (3) against a Mendelian model (2), a permutation test is not very straightforward. Therefore, the significance level for this test for presence imprinting was approximated by a permutation test on results that were achieved by fitting a single parental effect (the effect that came to expression) in a model, when testing for the presence of a QTL. This could be considered to result in significance levels for a test with one degree of freedom given the data, and multiple testing on the chromosome and the distribution of the trait are accounted for. The thresholds obtained this way were equal to a threshold that would be obtained by fitting an additive and no dominance component (results not shown). A QTL was considered to be of interest when it was detected with a $p \leq 0.05$ chromosome-wise significance level. Non-significant results ($0.1 > p \geq 0.05$) pointing to the same region were added to the table if it points towards the same region as QTL that did reach significance level on that chromosome. Also when more than one such suggestive QTL were pointing to the same region on the chromosome these results were added, because, even though these results on their own provide no significant evidence, they add to the probability that there is indeed a QTL in that region.

Results

Traits

In Table 2, means, minima, maxima, phenotypic standard deviations, and the proportion of phenotypic variance explained by fixed effects model (1) (R^2) are presented for all 17 traits considered in this analysis on 175 F2 animals. All traits were approximately normally distributed, except for survival, which was skewed towards higher values. For most traits the fixed effects model explains considerable amount of variance (26% for SD-PCV to 58% for LNPAR). For SD-BW the model explains only 9% of the variance, partly because a number of animals did not show a decrease in body weight following infection. The trait was included in the analysis to remain consistent with earlier work. For more details see Van der Waaij et al., (2001).

Table 2. Minima, maxima, means, adjusted phenotypic standard deviations, and proportion of variance explained by the fixed effect model (R^2) for all traits on 175 F₂-animals

Trait*	Min.	Mean	Max.	Phen. sd.	R^2
AV-PCV	17.7	23.8	32.7	2.94	0.43
ST-PCV	29.5	38.4	48.5	3.36	0.53
LV-PCV	10.9	16.3	27.1	2.87	0.35
MD-PCV	-34.6	-22.1	-9.70	3.69	0.42
SD-PCV	-0.64	-0.30	-0.12	0.12	0.26
FM-PCV	-0.80	5.24	17.5	4.16	0.29
TC-PCV	-32.33	-16.8	-1.40	6.11	0.29
AV-BW	80.6	149.4	219	25.1	0.44
ST-BW	92.0	153.2	215	23.8	0.49
LV-BW	74.0	139	210	24.2	0.39
MD-BW	-60.7	-14.1	4.33	12.7	0.39
SD-BW	-4.00	-0.15	3.67	0.54	0.09
FM-BW	-8.00	5.24	47.0	9.42	0.28
TC-BW	-58.7	-6.01	46.0	19.6	0.37
SURV	44.0	144	150	20.2	0.27
LNPAR	2.77	5.05	6.94	0.73	0.58
PARADET	3.0	11.5	22.0	3.81	0.52

*for abbreviations see Table 1.

QTL analysis

No QTL were detected on BTA26. Results for the QTL detected on the other chromosomes are in Tables 3 and 4. In Table 3 are the estimated paternal, maternal and dominance effects obtained from a full model (3) and also the estimates of the additive and dominance component obtained from a Mendelian model (2). Estimated effects with a positive sign indicate that the allele inherited from the N'Dama has the largest effect. Traits that are mentioned in between brackets point towards a possible QTL position, but the estimates were not significant ($0.1 > p \geq 0.5$). In total eight QTL have been detected. For the QTL for FM-PCV on BTA19 there is evidence for imprinting. Five QTL influence anaemia control, two of which also influence body weight change following infection. One QTL influences the PCV level prior to infection, and one is solely related to growth following infection. One QTL influences the number of times an animal is detected as parasitaemic (PARADET). Different QTL seem to be of influence on the decrease in PCV (BTA2, BTA5), as compared to the recovering of the PCV levels (BTA7, BTA19). The size of the QTL varies from 5.8% of the total phenotypic variance for starting value body weight (ST-BW) on BTA2 and PCV recovery (FM-PCV) on BTA7, to 10.4% for the lowest PCV level reached (LV-PCV) on BTA2 (Table 3). The QTL for the lowest value for PCV (LV-PCV) on BTA2 is the QTL with the highest significance.

Table 3. Estimated QTL effects, expressed in units of the trait, for all QTL detected, applying both a full model, including paternal, maternal and dominance effect, and a Mendelian model, including an additive and a dominance effect[#], for the eight chromosomes considered.

Chr	Trait	Pos. (cM)	QTL effect*	Full model			Mendelian	
				Paternal effect	Maternal effect	Dominance	Additive effect	Dominance
1	AV-BW	63	0.073	3.14 (1.50)	2.63 (1.51)	-9.28 (3.16)	5.77 (2.20)	-9.28 (3.15)
	ST-BW	63	0.058	1.85 (1.34)	2.81 (1.35)	-7.22 (2.83)	4.65 (1.96)	-7.22 (2.82)
	LV-BW	66	0.066	3.33 (1.55)	1.37 (1.55)	-10.43 (3.40)	4.70 (2.26)	-10.47 (3.40)
2	AV-BW	93	0.073	2.45 (1.53)	3.55 (1.51)	-7.18 (3.26)	6.02 (2.11)	-7.21 (3.25)
	ST-BW	96	0.058	2.34 (1.39)	1.66 (1.36)	-7.39 (3.00)	3.99 (1.92)	-7.33 (2.99)
	LV-BW	89	0.072	1.20 (1.43)	3.94 (1.43)	-7.17 (2.88)	5.14 (1.97)	-7.05 (2.89)
	AV-PCV	88	0.084	0.28 (0.17)	0.56 (0.17)	-0.52 (0.33)	0.83 (0.23)	-0.51 (0.33)
	LV-PCV	88	0.104	0.36 (0.18)	0.61 (0.18)	-0.65 (0.35)	0.97 (0.24)	-0.63 (0.35)
	MD-PCV	92	0.089	0.54 (0.23)	0.53 (0.23)	-0.99 (0.48)	1.07 (0.31)	-0.99 (0.48)
	TC-PCV	91	0.094	1.03 (0.41)	1.17 (0.40)	-1.31 (0.85)	2.20 (0.56)	-1.13 (0.85)
3	PARADET	9	0.063	-0.49 (0.23)	-0.53 (0.23)	-0.89 (0.53)	-1.02 (0.34)	-0.89 (0.53)
5	MD-BW	124	0.061	3.83 (1.68)	1.93 (1.17)	18.55 (7.86)	5.82 (2.39)	17.93 (7.81)
7	(TC-PCV)	96	0.061	0.00 (0.44)	0.95 (0.44)	-2.87 (0.95)	0.95 (0.64)	-2.86 (0.95)
	(FM-PCV)	95	0.058	0.05 (0.31)	0.21 (0.30)	-2.18 (0.68)	0.26 (0.44)	-2.17 (0.68)
19	FMPCV	49	0.077	-0.26 (0.32)	1.16 (0.31)	0.13 (0.71)	0.95 (0.46)	0.48 (0.75)
20	SURV	23	0.061	-3.11 (1.66)	-1.86 (1.91)	-10.56 (4.34)	-5.15 (2.59)	-10.95 (4.26)
23	STPCV	5	0.086	-0.37 (0.23)	-0.52 (0.23)	-1.51 (0.52)	-0.90 (0.33)	-1.49 (0.51)

[#] the position of the QTL mentioned in the table is equal to the estimated location of the Mendelian QTL position, which on BTA19 was equal to the estimated position of the imprinted QTL. * = proportion of the phenotypic variance

Table 4 presents the p-values for the presence of a Mendelian QTL, p-values for the presence of a QTL that only comes to expression if inherited from either one of the parents (single parent model), and p-values obtained from the test whether there is significant evidence for imprinting (model 1 vs model 3). The QTL for PCV recovery (FM-PCV) on BTA19 showed clear evidence of imprinting and is maternally expressed. Some of the most significant QTL with Mendelian expression also reached significance when testing for the presence of an imprinted QTL using a single parent model. However, there was no significant evidence for imprinting at these QTL. Large values for FM-PCV are desirable. All animals show a decrease in PCV (Van der Waaij et al., 2001), so large values for FM-PCV represent good recovery. The favourable allele for this QTL comes from the N'Dama. The single parent analysis and the test for imprinting lead to the same conclusions on the mode of expression of this QTL.

Table 4. The chromosome wise p-values for testing the presence of a QTL using a Mendelian model (including dominance), and a model with a single parent effect (paternal or maternal, the one with most significant effect), and for determining whether there is imprinting by testing a full model (paternal, maternal and dominance effect) versus a Mendelian model.

BTA	Trait	Position (cM)	Chromosome wise p-values				
			Size QTL effect*	Mendelian model	Single parent (expr) [†]	Imprinting?	
1	AV-BW	63	0.073	0.017	>0.1	>0.1	
	ST-BW	63	0.058	0.059	>0.1	>0.1	
	LV-BW	66	0.066	0.030	>0.1	>0.1	
2	AV-BW	93	0.073	0.023	>0.1	>0.1	
	ST-BW	96	0.058	0.082	>0.1	>0.1	
	LV-BW	89	0.072	0.028	>0.1	>0.1	
	AV-PCV	88	0.084	0.010	0.029	M	>0.1
	LV-PCV	88	0.104	0.002	0.014	M	>0.1
	MD-PCV	92	0.089	0.005	0.007	M	>0.1
	TC-PCV	91	0.094	0.003	>0.1		>0.1
3	PARADET	9	0.063	0.037	>0.1		>0.1
5	MD-BW	124	0.061	0.035	0.039	P	>0.1
7	(TC-PCV)	96	0.061	0.085	>0.1		>0.1
	(FM-PCV)	95	0.058	0.086	>0.1		>0.1
19	FMPCV	49	0.077	>0.1	0.002	M	0.012
20	SURV	23	0.061	0.038	0.084	P	>0.1
23	STPCV	5	0.086	0.004	>0.1		>0.1

* the position of the QTL mentioned in the table is equal to the estimated location of the Mendelian QTL position, which on BTA19 was equal to the estimated position of the imprinted QTL; * = proportion of the phenotypic variance; † expression of the inherited paternal allele: P = paternal expression, M = maternal expression

Profiles for chromosome-wise test statistics of the imprinting analysis for FM-PCV on BTA19 are presented in Figure 2. The figure present test statistics for the single parent model and for the Mendelian model. The figure clearly shows the difference in test statistic between the maternal effect that did, and the paternal that did not come to expression. No dominance was present (Table 4), which is also a prerequisite for the evidence of imprinting.

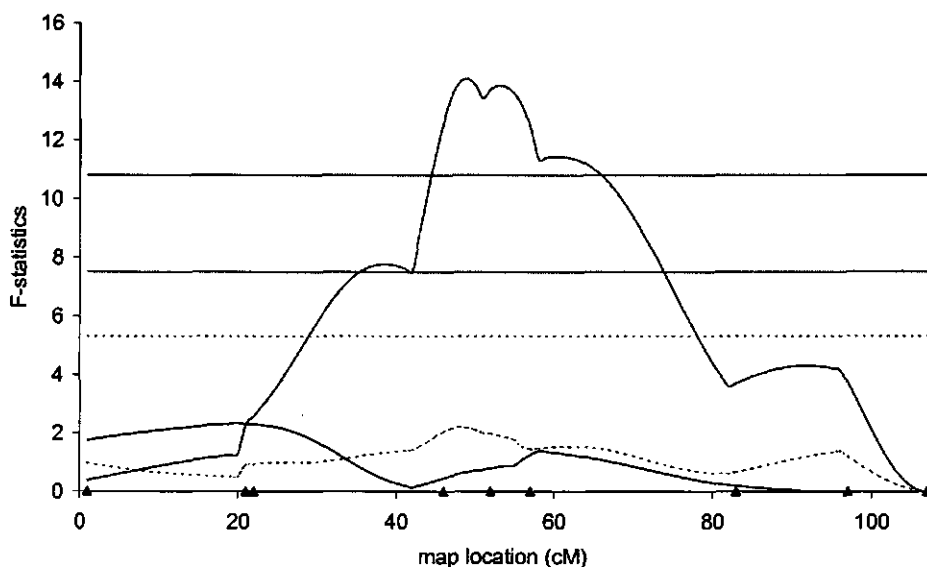


Figure 2. Distribution of the F-statistic for PCV recovery (FM-PCV) on chromosome 19, for a model with a paternal component (lower solid line), a maternal component (high solid line), or a Mendelian model (dotted line). The horizontal lines represent thresholds, the dotted line is the 5% threshold for the Mendelian model and the two solid horizontal lines are the 5% (lower line) and the 1% (upper line) threshold for a single component model. Marker positions are indicated with \blacktriangle on the x-axis.

Discussion

Based on preliminary results of an F2 experiment, nine chromosomes were selected, for which the mode of expression of the QTL was determined. Of the eight QTL detected, favourable alleles for six QTL originated from the N'Dama and two from the Boran. QTL were related to anaemia control, body weight change, and parasitaemia, both prior to and following infection with *Trypanosoma congolense*. For one of the QTL descending from the N'Dama there is evidence for imprinting. Also the favourable allele with largest effect originated from the N'Dama and explained 10.4% of the total phenotypic variance for the lowest value for PCV on BTA2. There is not evidence for an imprinted QTL in the comparative region of *Tir1* on BTA23 in cattle.

The results presented in this paper are obtained by pre-correcting the data for fixed effects for group in which the animals were weaned and challenged, sex of the animal, and

also for the family effect. Inclusion of the family effect can be questioned. In a line-cross analysis it is assumed that the QTL are not segregating in the grandparents, and thus that all F1 animals have equal genotype for the QTL. If this would hold for all genes, there would be no family differences. However, this F2-experiment was designed for a cross between two cattle breeds, which were not necessarily inbred for all QTL considered. Correction for family is expected to remove variance caused by polygenic effects (and by environmental effects other than included in the group effect), which would result in reduction in residual variance. All F2 families are expected to have the same genotypic distribution for the QTL. If family size is large enough, correction for the family effect should not cause any problems, but in smaller families correction for the family effect might remove part of the QTL effect. Correction for family effect might have resulted in the fact that some QTL may not have been detected, but the chance of false positive QTL will be reduced as well. In deciding on the inclusion of family effects, one needs to balance the advantage if a reduced residual variance (and reduced risk of false positives), against the risk of eliminating QTL effects. Given the relative large size of the families, we decided to include the family effect in our analysis.

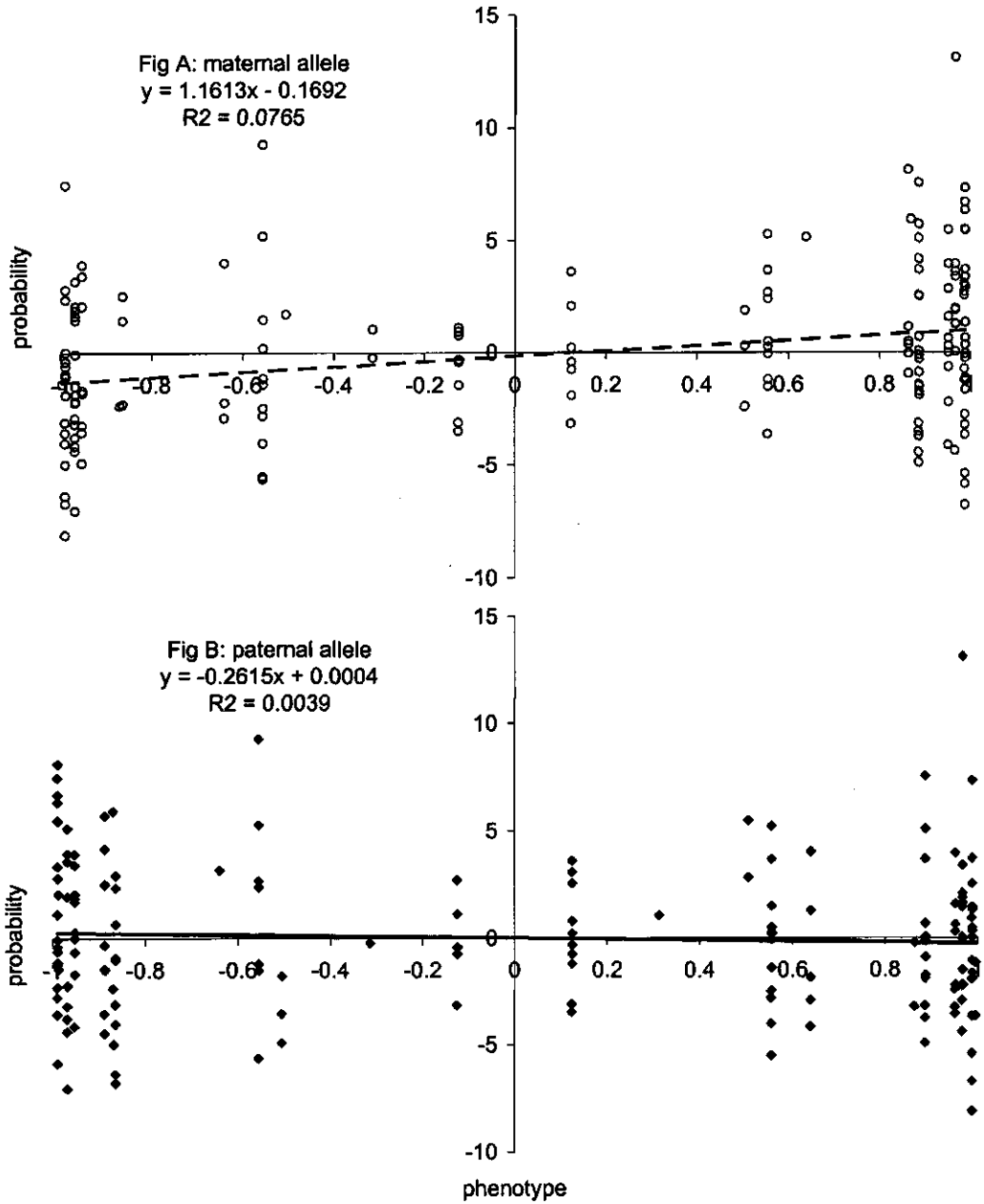
Table 5 shows the differences in results of a QTL analysis when also pre-correcting for the family effect, as compared to no pre-correction for family effect. Group and sex effect were always pre-corrected for. The significance level for most traits most changed when including family effect in the model. However, only those cases where the significance level changed from $p > 0.05$ to $p < 0.05$ or vice versa are presented in the table and will be discussed here. First thing to notice for all traits, is that the estimated QTL effects (especially the additive genetic effect) remain approximately of the same size after excluding the family effect. In some cases (BTA2, BTA5 and BTA20) the estimated QTL position changed as well, but at the significant QTL position estimated effects were very much comparable for both models. In these cases inclusion of the family effect in the model most likely reduced the error variance, resulting in more significant QTL estimates. The size of the estimated QTL effects did not change considerably, and neither did the estimated QTL position. Also, the significance dropped only to just below significance level. One explanation for the reduction in significance of the QTL estimates for these traits is that there may be confounding between genotypes and families, and/or confounding between genetic markers and family.

Table 5 Difference in number of QTL detected when including pre-correction for family effect, as compared to no pre-correction for family effect, on the basis of additive and dominance effect (standard error), the QTL location, and the chromosome-wise p-values (5% threshold for that QTL position).

Additional QTL, after including family effect						
BTA	Trait	Model	Additive (se)	Dominance (se)	cM*	p-value (5% thresh.)
1	AV-BW	Without family	5.53 (2.35)	-7.70 (3.37)	63	p > 0.1
		With family	5.77 (2.20)	-9.28 (3.15)	63	p = 0.017
	LV-BW	Without family	4.61 (2.35)	-8.78 (3.49)	65	p > 0.1
		With family	4.70 (2.26)	-10.47 (3.40)	66	p = 0.030
2	AV-PCV	Without family	0.84 (...)	-0.08 (...)**	88/94	p = 0.062
		With family	0.83 (0.23)	-0.51 (0.33)	88	p = 0.010
5	MD-BW	Without family	4.10 (...)	14.10 (...)**	124/65	p = 0.060
		With family	5.82 (2.39)	17.93 (7.81)	124	p = 0.036
20	SURV	Without family	-3.18 (...)	-8.84 (...)**	23/2	p > 0.1
		With family	-5.15 (2.59)	-10.95 (4.26)	23	p = 0.038
QTL that became non-significant after including family effect						
1	AV-PCV	Without family	-0.42 (0.30)	-1.30 (0.46)	71	p = 0.034
		With family	-0.30 (0.28)	-1.09 (0.42)	70	p > 0.1
2	FM-PCV	Without family	1.45 (.043)	0.83 (0.68)	95	p = 0.036
		with family	1.14 (0.38)	-0.17 (0.56)	90	p > 0.1
7	TC-PCV	without family	1.02 (0.65)	-3.13 (0.97)	96	p = 0.035
		with family	0.95 (0.64)	-2.86 (0.95)	96	p = 0.070
	FM-PCV	without family	0.49 (0.47)	-2.67 (0.74)	95	p = 0.011
		with family	0.26 (0.44)	-2.17 (0.68)	95	p = 0.086

*Estimated QTL position, where the first is the position of the significant QTL, and, if very different from the first, the most significant position of the QTL given the model including family effect, which is where the F-statistics are based on; ** no standard errors on the estimates were available for positions other than the most likely QTL position given the model.

Figure 3. Probability of inheriting the maternal allele (A) or the paternal allele (B) from the N'Dama (positive probabilities) or from the Boran (negative probabilities), versus the phenotypic performance of these F2 animals.



Most important objective of this paper was to determine the mode of expression of the QTL. The homologue of the imprinted QTL related to trypanotolerance in mice, *Tir1*, would have been located on cattle BTA23, but given the present data, no evidence for imprinting was detected on BTA23. However, on BTA19 there is evidence for imprinting (maternal expression) related to PCV recovery (FM-PCV). The region in question on BTA19 (see Figure 2) is not known as an imprinted region in either mouse or human. However, plotting the probability that an animal received an allele originating from the N'Dama (positive probabilities) or from the Boran (negative probabilities) from its mother, to the phenotypic performance of the animal, and subsequently fitting a regression line through these points (Figure 3a) illustrates the results for this QTL position in Tables 3 and 4 suggesting there is genomic imprinting. The figure shows a clear relation between the origin of the allele inherited from the mother and the phenotypic performance.

The regression coefficient of the line is significantly different from zero ($p=0.0002$). Mean phenotypic observation for the probabilities that the inherited allele came from the N'Dama larger than 0.85, was 0.81 with a SD of 4.11. The mean phenotypic observation for the probabilities that the inherited allele came from the Boran smaller than -0.85, was -1.37 with a SD of 3.41. For comparison, Figure 3b shows the same relation for the paternally inherited allele. This figure clearly shows the absence of a relation between inherited allele and phenotypic performance, and the regression coefficient of the line is not significantly different from zero ($p=0.417$).

Comparative mapping

For comparative mapping we have used the USDA cattle genome site (<http://www.marc.usda.gov/genome/genome.html>), the Bovine Genome Database (<http://bos.cvm.tazmu.edu/bovgbase.html>), and the NCBI *Homo Sapiens* (http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/hum_srch?chr=hum_chr.inf&query) map view. For information on function of the candidate genes we have used the Online Mendelian Inheritance (OMIM) (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM>). Information on imprinted genes was obtained from the Catalogue of Imprinted Genes and Parent-of-origin Effects in Human and Animals (<http://cancer.otago.ac.nz/IGC/Web/home.html>).

Results in this study suggest that the decrease in PCV and body weight following infection is caused by a different mechanism with regard to the possible subsequent recovery of PCV. Also Van der Waaij et al. (2001) have shown that decrease and recovery

are not strongly correlated. Based on these findings one can conclude that there are two (partly overlapping) waves of gene expression, relating to these two mechanisms. The first wave may involve the complement system, which is rapidly activated by high numbers of parasites occur, and results in very efficient in parasite clearance. If sufficient complement is available, this first wave may be sufficient to control the infestation. In many cases however, parasites may trigger the adaptive part of the immune system, and cause production of specific antibodies. This at the one hand will improve phagocytosis and clearance but at the other hand includes the risk of lysis of erythrocytes coated with parasitic antigen-antibody complexes (Taylor, 1999). In cases in which the infection outnumbers the innate defence capacity, it is important that the second mechanism, the specific immune response, takes over in time. This mechanism takes longer to be activated, usually about 14 days following infection (Taylor, 1999), and in the case of parasite infection it may take a little longer since the parasites need to multiply first before the immune system is seriously triggered.

A third mechanism involved in trypanotolerance is related to parasite clearance. Parasite kill can occur extra-cellular by complement activation as outlined but also intra-cellular following phagocytosis and cidal action of either cytotoxic NK-cells or cytotoxic T-cells. This clearance seems uncorrelated with body weight or PCV following infection (Trail et al., 1991; reviewed by Taylor, 1999). However, it is clear that animals that cannot control the parasite burden will seriously suffer from the infection, which will affect both their PCV level and bodyweight (e.g. Agyemang et al, 1997). A prominent characteristic of trypanotolerant animals, therefore, is their ability to control parasite numbers without adverse effects to themselves (Taylor, 1999).

One reason for not being able to find a correlation between parasitaemia and PCV may be the lack of a proper definition and/or observation on the level of parasitaemia. *T. congolense* is able to attach to the endothelial cells lining the blood vessels. This implies that parasites circulating in the peripheral blood are not representative of the entire population (reviewed by Taylor, 1999). Also, it is not the absolute number of parasites that seems to be of major influence, rather than the ability to control the number before the PCV drops below the critical level. Wildlife is a carrier of trypanosomes and thus a constant source of re-infection for the tsetse fly.

The first mechanism that comes to expression following infection is the complement system. The complement system is activated by a combination of antigen with a single molecule of IgM or two closely spaced IgG molecules (Tizard, 2000). IgM production is stimulated by B-cells. The role of B-cells in the response to trypanosomosis infection in the

N'Dama is different from trypanosusceptible breeds (Taylor et al., 1996). Also, the parasite may directly activate complement, as *T. congolense* releases factors that directly activate the bovine complement cascade (reviewed by Taylor, 1999). QTLs that represent the decrease in PCV and body weight following infection are located on BTA2 and BTA5. The QTL on BTA2 may include TNFAIP6 (HSA2), the tumour necrosis factor alpha-induced protein 6. TNF α is secreted by macrophages, T cell, B cells and fibroblasts and can act on almost all nucleated cells. It mediates many immune functions and regulates the growth of many cell types (Tizard, 2000). The QTL may also include ATP5G3 (ATP synthesis H⁺ transporting mitochondrial F0 complex (sub unit 9), isoform 3, in human on HSA2). This gene catalyses the synthesis of ATP during oxidative phosphorylation. The complement system requires a lot of energy, so it is likely that ATP synthesis is associated with complement activation since ATP-synthesis is of importance in times of extra activity.

The QTL on BTA5 includes many genes that may be related to trypanotolerance. For example C1s (in human on HSA12p13.1), of which the complement components C2 and C4 are natural substrates, and together form C42. One molecule of C42 can act on several hundreds molecules of C3, which is the key complement factor (Tizard, 2000). Also, the B7 protein (in human on HSA 12p13) assists T helper cells to help B-cells by cell to cell contact. The N'Dama have more circulating B-cells than Boran and these B-cells are in a higher state of activation (reviewed by Taylor, 1999). There are also several natural killer (NK) cell (receptor) related genes in the region of the QTL: KLRD1, ILT1, KLRC4, KLRF1 (all in HSA 12p13 region). NK cells can bind to target cells loaded with parasites by means of specific antibody, resulting in a cytotoxic complex. NK-cells can act without previous sensitisation (Tizard, 2000).

Anaemia is regarded the most important cause of death following a trypanosome infection (e.g. Taylor, 1999). Early in infection, anaemia is likely to be a result of extra-vascular haemolysis. This occurs predominantly in the spleen, liver and lymph nodes. Insufficient haemopoiesis for substitution of lost erythrocytes, may contribute to anaemia later in infection (reviewed by Taylor, 1999). Failure to control the anaemia will result in a life-threatening situation. Survival implies successful control of such lethal drop in PCV, i.e. severe anaemia. QTL for survival were detected on BTA20. The QTL may include complement components C6 and C9 are mapped in cattle. C6 binds to a C5b and C7, resulting in a C5b67 complex, which again binds to C8. Then, twelve to eighteen C9 molecules are required to form a tubular transmembrane pore called the membrane attack complex. If sufficient numbers of these complexes are formed, osmotic lysis and killing of

the organism occurs (Tizard, 2000). So especially C9 may add to efficient parasite clearance and prevent excessive anaemia. The QTL may also include IL7R (HSA 5p13), the interleukin-7 receptor. IL7 acts on thymocytes, T-cells, B-cells, monocytes, and lymphoid stem cells, and supports generation of cytotoxic T-cells from thymocytes (Tizard, 2000).

QTL related to PCV recovery are located on BTA7, and BTA19. These QTL should be involved in the second wave of gene expression, and therefore related to the T-cell response. The QTL on BTA 7 may include ITK (HSA 5q31-q32), which codes for IL2-inducible T-cell kinase. IL2 is secreted when T-helper 1 cells encounter macrophage-bound antigen and it activates T cells, B cells, NK cells and macrophages (Tizard, 2000). IL2 affects the proliferation of activated T-cells and is essential for their long-term growth. Another gene in that region is IL12B (HSA 5q31), which is natural killer cell stimulator factor 2, cytotoxic lymphocyte maturation factor 2. IL12 is produced by all the major antigen-presenting cells. It stimulates T-helper 1 cells to produce IL2 and IFN γ , and enhances T and NK cell proliferation and cytotoxicity (Tizard, 2000). A third possible candidate gene for the QTL on BTA7 is IL5, which is produced by T-helper 2 cells and stimulates the growth and activation of eosinophils, and growth of B-cells (Tizrad, 2000). Major role of the eosinophil is the destruction of invading parasites. Another important gene in that same region is IL9 (HSA 5q31.1), which is a haematopoietic growth factor (Tizard, 2000), and therefore plays an important role in the control of anaemia. The QTL on BTA19 is maternally expressed and no real candidate gene that is also subject to imprinting could be found in the comparative regions of either human or mice.

Finally, the third mechanism involved in trypanotolerance is parasite clearance. The gene density is high in the homologue region in human, but probably the best candidate gene for the QTL for the efficiency in parasite clearance (PARADET) on cattle BTA3 is CD64 (HSA q21.2-q21.3), which is a high affinity IgG receptor and is expressed on monocytes and interferon-stimulated granulocytes. It plays a key-role in antibody-dependent cellular cytotoxicity (Tizard, 2000). Other possible candidate genes are CD1B, CD1C, and CD1D (all HSA q22-q23), which are antigen presenting molecules, or CD244 (HSA p36.13-q23.3) which is a natural killer cell receptor.

Conclusion

In conclusion, there seem to be many QTL involved in trypanotolerance, most with large effect. Fine mapping of these QTL, and subsequent gene-expression studies on candidate genes in the region, would increase the understanding of the immunological background of trypanotolerance. A selection of QTL related to trypanotolerance that are closely linked to genetic markers could then be implemented in a breeding scheme to improve trypanotolerance, with less necessity to infect the animal or withhold it from medical treatment.

Acknowledgements

The authors would like to thank WOTRO (Netherlands Foundation for the Advancement of Tropical Research) and ILRI (International Livestock Research Institute) for financial support. They also would like to thank Alan Teale and M. Soller for initiating the project, the people in the molecular genetics lab at ILRI for genotyping, Huub Savelkoul for some useful discussion, and Raoul Meuldijk for assistance in data processing.

References

- Clapcott, S.J., Teale, A.J. and S.J. Kemp. 2000. Evidence for genomic imprinting of the major QTL controlling susceptibility to trypanosomiasis in mice. *Parasite Immunology* 22: 259-263.
- De Koning, D.J., Rattink, A.P., Harlizius, B., Van Arendonk, J.A.M., Brascamp, E.W., Groenen, M.A.M., 2000. Genome-wide scan for body composition in pigs reveals important role of imprinting. *Proc. Natl. Acad. Sci. U. S. A.* 97: 7947-7950.
- Green, P., Falls, K. and Crooks, S. (1990): Cri-map Version 2.4, Washington University School of Medicine, St Louis.
- Haldane, J.B.S. 1919. The combination of linkage values and the calculation of distances between the loci of linked factors. *J. Genet.* 8:299-309.
- Haley, C. S., Knott, S.A., and Elsen, J.M. 1994 Mapping Quantitative Trait Loci in crosses between outbred lines using least squares. *Genetics* 136:1195-1207.
- d'Ieteren, G.D.M., Authié, E., Wissocq, N. and Murray, M. 2000. Exploitation of resistance to trypanosomes and trypanosomosis. In: Breeding for disease resistance in farm

- animals. 2nd edition, edited by R.F. E. Axford, S.C. Bishop, F.W. Nicholas and J.B. Owen. Cab International, Wallingford, UK.
- ILRAD, 1989. N'Dama cattle: managing Africa's genetic resources. Report October 1989.
- Jordt, T., Mahon, G.D., Touray, B.N., Ngulo, W.K., Morrison, W.I., Rawle, J. and Murray, M. 1986. Successful transfer of frozen embryos from the Gambia to Kenya. *Trop. Anim. Hlth. Prod.* 18: 65-75.
- Knott, S.A., Marklund, L., Haley, C.S., Andersson, K., Davies, W., Ellegren, H., Fredholm, M., Hansson, I., Hoyheim, B., Lundstrom, K., Moller, M., Andersson, L., 1998. Multiple marker mapping of quantitative trait loci in a cross between outbred wild boar and large white pigs. *Genetics* 149: 1069-1080.
- Li, L.-L., Keverne, E.B., Aparicio, S.A., Ishino, F., Barton, S.C. and Surani, M.A. 1999. Regulation of maternal behaviour and offspring growth by paternally expressed Peg3. *Science* 284: 330-333.
- Murray, C., Murray, M., Murray, P.K. Morrison, W.I., Pyne, C., and McIntyre, W.I.M. 1977. Diagnosis of African trypanosomiasis in cattle. Improved parasitological and serological techniques. In: International Scientific Council for Trypanosomiasis Research and Control 15th Meet., the Gambia. OAU/STRC. Publication No. 110, p. 247-254.
- Murray, M., Trail, J.C.M., Davis, C.E. and Black, S.J. 1984. Genetic resistance to African Trypanosomosis. *The 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. *The Veterinary Quarterly* 15:60-67.
- Reik, W. and Walter, J. 2001. Genomic imprinting: parental influence on the genome. *Nature Reviews/ Genetics*. 2: 21-32.
- SAS version 6.12. 1997. SAS Institute Inc., Cary, USA.
- Schmidt, J.V., Matteson, P.G., Jones, B.K., Guan, X.-J. and Tilghman, S.M. 2000. The Dlk1 and Dtl2 genes are linked and reciprocally imprinted. *Genes and Development* 14: 1997-2002.
- Taylor, K.A., Lutje, V., Kennedy, D., Authié, E., Boulangé, A., Logan-Henfrey, L., Gichuki, B. and Gettinby, G. 1996. Trypanosoma congolense: B-lymphocyte responses differ between trypanotolerant and trypanosusceptible cattle. *Exp. parasitol* 83: 106-116.

- Taylor, K. 1999. Immune responses of cattle to African trypanosomes: protective or pathogenic? *Int. J. Parasitol.* 28: 219-240.
- Trail, J.C.M., d'Ieteren, G.D.M., Feron, A., Kakiese, O., Mulungo, M. and Pelo, M. 1991. Effect of trypanosome infection, control of parasitaemia and control anaemia development on productivity of N'Dama cattle. *Acta Tropica* 48: 37-45.
- Trail, J.C.M., d'Ieteren, G.D.M., Maille, J.C., Yangari, G. and Nantulya, V.M. 1992. Use of antigen-detection enzyme immunoassays in assessment of trypanotolerance in N'Dama cattle. *Acta Tropica* 50: 11-18.
- Van der Waaij, E.H., Hanotte, O., Van Arendonk, J.A.M., S.J. Kemp, D. Kennedy, J.P. Gibson, A. Teale. 2001. An F2 experiment to determine properties of trypanotolerance in African cattle. *In preparation.*

Appendix

Table 1. Marker names, cumulative distances between markers (in this data) and information content at the marker positions for BTA 1, 2 3 and 5.

Chr 1			Chr 2			Chr 3			Chr 5		
Marker	Dist*	Info	Marker	Dist	Info	Marker	Dist	Info	Marker	Dist	Info
DIK070	1	0.39	BM3627	1	0.86	INRA006	1	0.53	BP1	1	0.81
RM095	17	0.64	ILST026	5	0.72	ILST096	9	0.53	Int2	44	0.53
ILSTS104	21	0.79	ILST025	26	0.63	BM0723	28	0.76	BL37	49	0.56
TGLA057	52	0.58	CSFM50	36	0.43	INRA003	39	0.72	BR2936	65	0.82
DIK024	63	0.73	CSSM42	64	0.81	INRA123	58	0.46	ETH10 ¹	69	0.92
INRA128	75	0.79	BM003	73	0.88	HUI1177	91	0.74	ETH152 ¹	175	0.65
BM6506	78	0.90	BMS803	78	0.82	INRA88	110	0.84	BM049	183	0.88
URB038	92	0.89	BMC9002	88	0.90	IDVGA35	121	0.78	BM2830	184	0.89
DIK104	96	0.83	ILSTS098	104	0.79	HAUT31	127	0.86	ETH2	185	0.87
BM864	100	0.92	BM4440	106	0.79	BM2924	137	0.83			
TGLA130	112	0.82	ILSTS082	117	0.41	IDVGA27	151	0.43			
BM1824	128	0.66	ARO28	141	0.53						
CSSM19	149	0.44	BM6444	141	0.53						
BM3205	154	0.36	BM4117	154	0.76						
			OarFcB11	175	0.82						

¹ info content drops to 0.10 in between these markers

* Distance in cM

Table 2. Marker names, cumulative distances between markers (in this data) and information content at the marker positions for BTA 7, 19, 20 and 23.

Chr 7			Chr 19			Chr 20			Chr 23		
Marker	Dist.	Info	Marker	Dist.	Info	Marker	Dist.	Info	Marker	Dist.	Info
BP41	1	0.88	BM9202	1	0.31	BMN3517	1	0.74	INRA64	1	0.40
BMS713	3	0.71	HEL10	20	0.74	BM1225	10	0.70	INRA132	1	0.40
RM006	11	0.59	AFLP361	21	0.76	BMS1282	34	0.82	CSSM005	27	0.60
BM116	17	0.67	BMS1920	42	0.62	TGLA126	71	0.74	RM033	46	0.57
BM7242	48	0.80	UWCA33	51	0.68	BMS2361	86	0.80	UWCA1	50	0.53
CSSM29	58	0.82	BP20	55	0.76	BM5004	110	0.87	DRB3	72	0.67
DIK050	69	0.84	UWCA40	58	0.78	UWCA26	121	0.87	TNFA	74	0.68
BMS2258	70	0.86	DIK42	82	0.68				MTGT007	76	0.73
INRA112	74	0.63	BMS2842	96	0.54				BORADRBPI	80	0.81
BM1853	84	0.56	BMC1013	107	0.48				RM185	85	0.83
BM9065	96	0.61							BM7233	95	0.65
ILST006	114	0.71							CSS24	105	0.81
BMS522	119	0.88							BM1905	111	0.83
INRA53	122	0.85									
BL1043	140	0.83									

IV

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

The genetic and economic consequences of introgression of either one or two genes that explain the complete difference for disease resistance between donor and recipient breed 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 (4 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 population. Introgression of one or two genes made no difference for 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 with low female reproduction capacity.

Keywords: introgression, disease resistance, cattle, 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. For example, this 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.

This chapter has been published and is reprinted with permission

E.H. Van der Waaij, J.A.M. Van Arendonk. 1999. Introgression of genes responsible for disease resistance in a cattle population selected for production: genetic and economic consequences. *Animal Science* 70: 207-220.

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 the 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, 1997b). In order to cover this increasing risk of losing the allele, an increased 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 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 it would also 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 that case, an increase in 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 increased costs for the introgression program. 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 (1 or 2), 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 to profit made per animal after fixation of the favourable allele. Since this is the first study on optimisation 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 to 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 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_{seg}^2) exists in addition to additive genetic variance and its 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 determines 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 proceeding 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 proceeding 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 ($\sigma_{ms}^2 = \frac{1}{2} \sigma_{a_0}^2 + \sigma_{seg}^2$, 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:

$$\sigma_{seg}^2 = 0.5 \sum_{i=1}^n \delta_i^2 = 0.5n \left[\sigma_{\delta}^2 + (\bar{\delta})^2 \right] \quad (1)$$

where n is the number of loci involved, σ_{δ}^2 is the variance across loci due to difference 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 $\sigma_{\delta}^2 = 0$, 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:

$$\sigma_{seg}^2 = 0.5 \frac{\left(\frac{\mu_2 - \mu_1}{\sigma} \right)^2}{n} \quad (2)$$

where $\mu_1 - \mu_2$ 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, with each of these eight independent loci located on a separate chromosome of length 100 cM. On average there will be 100 recombinations per 100 meioses per chromosome, resulting in on average 1 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 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.

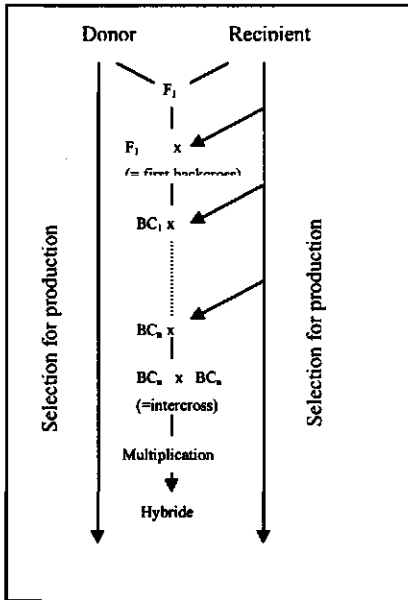


Figure 1. Schematic reproduction of the breeding scheme including terminology used throughout the paper. Selection for genetic markers is applied during the backcross generations

Because introgression usually will occur under experimental conditions, it was assumed that it would be easier 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 also was applied in the donor and recipient populations. Selection in the donor population was introduced to make a fair comparison

between 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 own performance and performance of 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 ($\sigma_p^2 = 1$, $h^2 = 0.2$). Generations are discrete, with equal generation interval for males and females.

Numbers of animals.

The purpose of the introgression program was to create a new hybrid population that is resistant to the disease and that has a higher production level compared to 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, Robertson and Latter, 1959). In this study ΔF is allowed 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_h). Given that $N_d = 50$ and $m_{od} = 4$ or 10, m_h 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. $m_h = \left[\left(\frac{1}{2} \right)^{2n_g} \right]^{-1}$. In the case of one gene (with two alleles), one out of four [i.e. $(\frac{1}{2})^{2 \times 1}$] 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 sixteen 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; ILRAD, 1989; FAO, 1999) and was assumed to be equal to 0.1 phenotypic sd. 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 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 liveweight gain per 100kg 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_d) 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 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:

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

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

where X_i is the individual's own performance, \bar{X}_{FS} is the average performance of the full-sib dam family, including the individual's own performance, \bar{X}_{HS} is the 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 ($\sigma_{I_i}^2$) is determined as:

$$\sigma_{I_i}^2 = \mathbf{b}_{(i)}' \mathbf{P}_{(i)} \mathbf{b}_{(i)} \quad (5)$$

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

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

$$= \begin{bmatrix} (\sigma_{ms}^2 + \sigma_e^2) \times (1 - \frac{1}{n}) & 0 \\ 0 & [\sigma_d^2 + (\sigma_{ms}^2 + \sigma_e^2)/n] \times (1 - \frac{1}{d}) \end{bmatrix}$$

$$\mathbf{G} = \begin{bmatrix} \text{cov}(A_i, (X_i - \bar{X}_{FS})) \\ \text{cov}(A, (\bar{X}_{FS} - \bar{X}_{HS})) \end{bmatrix} = \begin{bmatrix} \sigma_{ms}^2 (1 - \frac{1}{n}) \\ (\sigma_d^2 + \frac{\sigma_{ms}^2}{n}) \times (1 - \frac{1}{d}) \end{bmatrix}$$

where σ_d^2 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 $\sigma_{ms}^2 = \frac{1}{2} \sigma_{a_0}^2 + \sigma_{seg}^2$, where $\sigma_{a_0}^2$ is the additive genetic variance in the base generation (i.e. independent of t). In purebred selection there is no segregation variance. The error variance is σ_e^2 (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): $b_1 = \frac{\sigma_{ms}^2 (1 - \frac{1}{n})}{(\sigma_{ms}^2 + \sigma_e^2) \times (1 - \frac{1}{n})} = h_w^2$ 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:

$$r_{Ia_{sire(t)}} = \sqrt{\frac{\sigma_{I_{sire(t)}}^2}{\sigma_{a_{sire(t)}}^2}} \quad (6)$$

where $\sigma_{a_{sire(t)}}^2$ 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 ($\sigma_{a_{s(t-1)}}^2$), 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

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

$$\mathbf{G} = \text{cov}(A, (X_i - \bar{X}_{FS})) = \sigma_{ms}^2 (1 - \frac{1}{n})$$

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

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

$$r_{Ia_{dam(t)}} = \sqrt{\frac{\sigma_{I_{dam(t)}}^2}{\sigma_{a_{dam(t)}}^2}} = \sqrt{\frac{\sigma_{ms(t)}^2 \times h_{w(t)}^4}{\sigma_{ms(t)}^2}} = \sqrt{h_{w(t)}^2} \quad (7)$$

where $\sigma_{a_{dam(t)}}^2$ 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 $\sigma_{a_{s(t-1)}}^2 + \sigma_{a_{d(t-1)}}^2 + \sigma_{ms_t}^2$. Additive genetic variance in selected sires in generation t is given by $\sigma_{a_{s(t)}}^2 = \frac{1}{4} \sigma_{a_t}^2 (1 - k_s \times r_{Ia_{sire(t)}}^2)$ and in selected dams by $\sigma_{a_{d(t)}}^2 = \frac{1}{4} \sigma_{a_t}^2 (1 - k_d \times r_{Ia_{dam(t)}}^2)$, 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:

$$\mu_{(t)} = 0.5 * (\mu_{s(t-1)} + \mu_{d(t-1)}) \tag{8}$$

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:

$$\mu_{x(t-1)} = i_x * \sigma_{I_x} + \mu_{(t-2)} \tag{9}$$

where x is s (sire) or d (dam), i_x is the selection intensity in sex x, and σ_{I_x} is the standard deviation of the selection index in sex x. Selection intensity depends on the family size and thus on the female reproductive capacity. When m_{od} equals 4, selection intensities are $i_s=1.539$ for sires and $i_d=0.564$ for dams. When m_{od} equals 10, selection intensities are $i_s=1.965$ for sires and $i_d=1.163$ 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_c) compared to the purebred recipient population

mean (μ_r), cumulated over generations, and are calculated as
$$C = \sum_{i=1}^t m_{ab_i} (\mu_{r_i} - \mu_{c_i}),$$

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 & 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, $T = m_{ab} \sum_{i=1}^{n_g} 0.5^i$, 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. $m_{ac} = C / (\mu_h - \mu_d)$). 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 to 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 genetic variance over generations for four initial breed differences. When the breed difference is small,

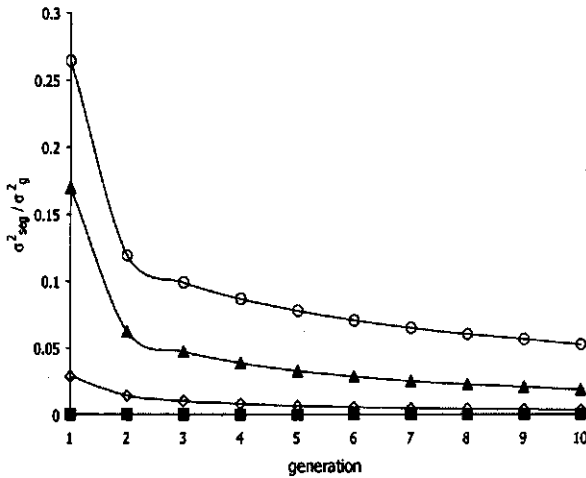


Figure 2 Ratio segregation variance (σ_{seg}^2) to total additive genetic variance (σ_a^2) across generations for initial breed differences of 0.1 (■), 1.0 (◇), 2.5 (▲) and 5.0 (○) phenotypic s.d.

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..

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 to 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 to 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 ($d_{dr} = 0.1, 1.0, 2.5$ or 5.0) and two female reproductive rates ($m_{od} = 4$ or 10). For $d_{dr} = 0.1$, and in the case

of $n_b = 0$ or 1 for $d_{dr} = 1.0$, selection for production within the donor population resulted in a larger increase in population mean compared to 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 $d_{dr} = 0.1$ 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.

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 low female reproductive rate (m_{od})

d_{dr}^{\ddagger}	$n_b = 0$	$n_b = 1$	$n_b = 3$	$n_b = 7$
	$m_{od} = 10^{\dagger}$			
0.1 sd	-0.670	-0.690	-0.705	-0.711
1.0 sd	-0.220	-0.015	0.139	0.185
2.5 sd	0.530	1.110	1.545	1.679
5.0 sd	1.780	2.985	3.889	4.169
$m_{od} = 4^{\dagger}$				
0.1 sd	-0.461	-0.456	-0.449	-0.445
1.0 sd	-0.011	0.219	0.395	0.448
2.5 sd	0.739	1.344	1.801	1.942
5.0 sd	1.989	3.219	4.145	4.432

[‡] expressed in phenotypic standard deviations; [†] = number of offspring per dam

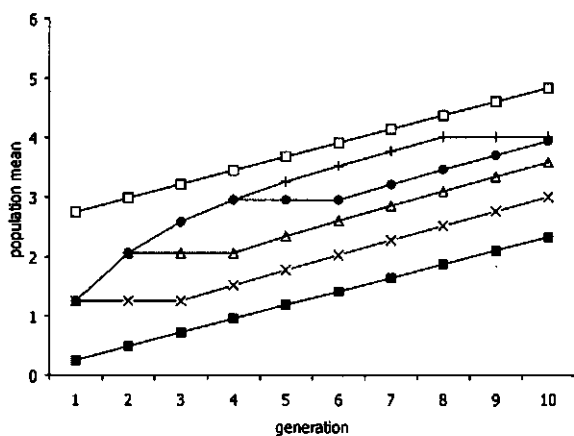


Figure 3 Increase in population mean for production when initial breed difference is 2.5 s.d. and female reproductive rate is high, comparing selected purebred populations (donor (■) and recipient (□)) to animals in the four introgression schemes ($nb = 0$ (×), $nb = 1$ (Δ), $nb = 3$ (●) or $nb = 7$ (+) generations of backcrossing). Selection was applied in each of the purebred populations and after fixation of the favourable alleles also in the hybrid population.

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 to the crossbred animals in

the later stages of the introgression program. 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 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 $n_c = 2$ ($d_{dr} = 2.5$) to $n_c = 8$ ($d_{dr} = 5$)

generations of selection in the donor population. For $m_{od} = 4$, $n_c = 5$ ($d_{dr} = 2.5$) 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 $m_{od} = 10$, $d_{dr} = 1.0$ 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 ($d_{dr} = 0.1, 1.0, 2.5$ 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 result in the lowest values for C and m_{ac} . Only in the case of $d_{dr} = 5.0$ 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 to $n_b = 0$ 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} reflects

Introgression of genes for disease resistance in cattle

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})

d_{dr}	n_b	$m_{od} = 10$				$m_{od} = 4$	
		C	m_{ac}^\dagger	T	C	m_{ac}^\dagger	T
0.1 sd.	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 sd.	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 sd.	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 sd.	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

† = incl. the animals in the multiplication generation; ‡ - no profit will be made due to higher genetic level of the donor breed.

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 $d_{dr} = 2.5$ 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 ($d_{dr} = 2.5$ or 5.0 s.d.), since the difference between crossbred and purebred animals is biggest 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 $d_{dr} = 1.0$ s.d., m_{ac} is 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.

Introgression of genes for disease resistance in cattle

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_{od}) and number of genotypings needed until fixation (T), considering high and lower female reproductive capacity (m_{od}) and various initial breed differences (d_{dr})

d_{dr}	n_b	$m_{od} = 10$			$m_{od} = 4$		
		C	m_{ac}^\dagger	T	C	m_{ac}^\dagger	T
0.1 sd.	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 sd.	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 sd.	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 sd.	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

[†] = incl. the animals in the multiplication generation; [‡] - = no profit will be made due to higher genetic level of the donor breed.

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 still is 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 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 generations can be avoided but that, in particular for $m_{od} = 10$, puts high demands on infrastructure as well as financial investments which might not make it 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 being needed per generation, since animals that carry the favourable allele(s) can be used more than 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. Summarising, 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 infested and 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 for an introgression program. 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 therefore often is 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 thus not only is 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 $d_{dr} = 0.1$ 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 $m_{od} = 4$, $d_{dr} = 1.0$ 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 role when deciding to start an introgression program.

From the economical 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 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 are 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 is possible to genotype at the embryo level (P. Bredbacka pers. comm.). 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.

Acknowledgement

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. *Reprod. Nutr. Dev.* 38: 605-613.
- Bulmer, M.G. 1971. The effect of selection on genetic variability. *The 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, 1997.
- FAO Animal Production and Health Paper. 1999. Productivity of trypanotolerant livestock 20:87-97.
- Falconer, D.S. 1989. *Introduction to quantitative genetics*. Third edition, Longman Scientific Technical, New York.
- Gama, L.T., Smith, C. and Gibson, J.P. 1992. Transgene effects, introgression strategies and testing schemes in pigs. *Anim. Prod.* 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. *J. Anim. Breed. Genet.* 112:161-170.
- Hill, W.G., Caballero, A. and Dempfle, L. 1996. Prediction of response to selection within families. *Genet. Sel. Evol.* 28:379-383.
- Hillel, J., Verrinder Gibbins, A.M., Etcjes, R.J. and McQ. Shaver, D. 1993. Strategies for the rapid introgression of a specific gene modification into a commercial poultry flock from a single carrier. *Poultry Science* 72:1197-1211.
- Hospital, F., Chevalet, C. and Mulsant, P. 1992. Using markers in gene introgression breeding programs. *Genetics* 132: 1199-1210.
- Hospital, F. and Charcosset, A. 1997. Marker-assisted introgression of quantitative trait loci. *Genetics* 147:1469-1485.
- ILRAD report October 1989. N'Dama cattle: Managing Africa's genetic resources. 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 Van Arendonk, J.A.M. 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. *The Veterinary Quarterly* 15:60-67.
- Tanksley, S.D. 1983. Molecular markers in plant breeding. *Plant. Mol. Biol.* 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.
- Van Heelsum, A.M., 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.
- Van Heelsum, A.M., 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.
- Visscher, P.M., Haley, C.S. and Thompson, R. 1996. Marker-assisted introgression in backcross breeding programs. *Genetics* 144:1923-1932.
- Visscher, P.M. and Haley, C.S. 1999. On the efficiency of marker-assisted introgression. *Animal Science* 68: 59-68.
- 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.

V Modeling selection for production traits under constant infection pressure

This paper presents a model describing the relationship between level of disease resistance and production under constant infection pressure. The model assumes that given a certain infection pressure, there is a threshold for resistance below which animals will stop producing, and that there is also a threshold for resistance above which animals produce at production potential. In between both thresholds animals will show a decrease in production, the size of decrease depending on the severity of infection, and the level of resistance. The dynamic relationship between production and resistance when level of resistance changes, such as due to infection is modeled both stochastically and deterministically. Selection started in a population with very poor level of resistance introduced in an environment with constant infection pressure. Mass selection on observed production was applied, which resulted in a non-linear selection response for all three traits considered. When resistance is poor, selection for observed production resulted in increased level of resistance. With increasing level of resistance, selection response shifts to production potential and eventually selection for observed production is equivalent to selection for production potential. The rate at which resistance is improved depends on its heritability, the difference between both thresholds, and selection intensity. The model also revealed that when a zero correlation between resistance and production potential is assumed, the phenotypic correlation between resistance and observed production level increases for low levels of resistance and subsequently asymptotes to zero, whereas the phenotypic correlation between production potential and observed production asymptotes to one. For most breeding schemes investigated, the deterministic model performed well in relation to the stochastic simulation results. Experimental results reported in literature support the model predictions.

Keywords: Genetic Epidemiology, Disease Resistance, Production, Selection, Non-linear Models, Modelling

Introduction

Animals that are not fully resistant to a certain infectious disease will suffer from the infection and consequently show a decrease in production. The size of this decrease in production depends on the degree of resistance of the animal and on the severity of infection. It is economically important to increase the level of disease resistance in the population to a level where the influence of infection on production is negligible and the number of animals that need medical treatment is minimal (e.g., Bennett et al., 1999). This can for example be achieved by vaccination, when available. Another option is to increase

This Chapter has been published and is reprinted with permission

E.H. van der Waaij, P. Bijma, S.C. Bishop, and J.A.M. Van Arendonk. 2000. Modeling selection for production traits under constant infection pressure. *Journal of Animal Science* 78: 2809-2820.

the genetic potential for disease resistance in the population (e.g., Gavora and Spencer, 1978).

Before an effective selection scheme for disease resistance can be established, it is important to investigate the relationship between production and level of disease resistance. Broadly two types of infections can be distinguished: sporadic infections (e.g., viral infections) and non-sporadic infections (e.g., gastro-intestinal parasites in sheep [Bishop and Stear, 1999] or trypanosomiasis [e.g., Murray et al., 1990]), the latter type having a constant influence on production. The influence of level of resistance to this last type of infection on production will be addressed in this paper, for the case where the host animal resistance level has no impact on the overall level of disease challenge, hence the epidemiology of the infection.

The objective of this research was to model and give insight in the underlying relationship between level of disease resistance and productivity under continuous infection pressure when selection is on observed production. The model will account for the dynamic relationship between production and resistance when level of production changes, e.g., due to infection.

Methods

The effect of level of disease resistance on the production level is defined as 'resilience' (Bisset and Morris, 1996; Woolaston and Baker, 1996). We use the term 'resistance' throughout this paper as we consider some of the traits underlying resilience, that is, resistance per se and productivity.

Model

Under infection pressure, the level of resistance of an animal to the infection affects the production that can be achieved. Individuals that are less resistant to the infection will be less productive. When a high infection pressure occurs in combination with a low level of resistance, production can drop to, or even below (e.g., in case of growth) zero. On the other hand, the production of animals that have a high enough level of resistance will not be reduced in the presence of the disease.

Under the infinitesimal model, level of resistance (**R**) and production potential (**Pp**) may be defined as continuous normally distributed traits. The level of production that would

have been achieved if the animal were completely resistant is denoted as P_p . Under infection pressure, however, production can be influenced by the level of disease resistance. In other words, observed production (P_o) is not always equal to the underlying production potential, but is influenced by the level of disease resistance. Observed production only exists at phenotypic level.

In each of the sub populations, the interaction between production potential and disease resistance can be defined as:

$$P_o = P_p \times f(R), \quad (1)$$

where $f(R)$ depends on the value of resistance in relation to the thresholds, and may represent any function that is suitable for this relation, for example a linear, exponential, or logistic function. It is assumed that there is no correlation between production potential and disease resistance and values for resistance should be non-negative. Although in reality the function of resistance is likely to describe some s-curve, with asymptotes at zero and one, in this paper it is assumed that two thresholds for resistance level divide the population into three sub populations. There is a certain lower threshold (L) for disease resistance, below which production drops to a minimum ($P_o = 0$). There also is an upper threshold (U) for disease resistance, above which the animal is fully resistant and production equals production potential (i.e., observed production = P_p). Between the lower and upper threshold, observed production will depend on both production potential and resistance. The upper and lower thresholds are considered fixed but they will depend on the type of infection and other environmental factors.

The influence of resistance on observed production differs in each sub population and observed production can, therefore, be considered as consisting of three separate distributions I, II, and III (see also Figure 1):

$$\begin{array}{ll} \text{I} : P_o = 0, & \text{for } R \leq L \\ \text{II} : P_o = P_p \times f(R) & \text{for } L < R < U \\ \text{III} : P_o = P_p & \text{for } R \geq U \end{array} \quad (2)$$

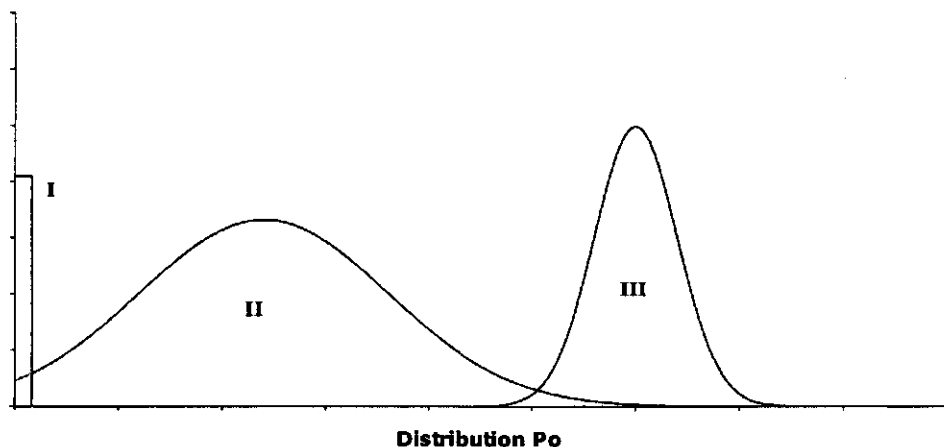


Figure 1. Schematic overview of what the distribution of observed production is assumed to look like in a situation where the population contains animals that are resistant as well as animals that have zero production: divided into the three sub-distributions, two of them normally distributed with different mean and variances, and one with mean and variance equal to zero.

This model is general and can be applied to a wide variety of disease scenarios. The infection pressure modeled may differ within as well as between diseases. These range from scenarios where the mean level of resistance of the host population can mediate the disease pressure of the pathogen (the disease epidemiology), as for gastrointestinal parasites in sheep, to cases where there is a large environmental reservoir on which the host genotype has no measurable effect on the disease pressure. This for example is the case with trypanosomiasis in many African countries, where the large numbers of wildlife, grazing the same area as the cattle population, provide a constant reservoir of parasites. This paper investigates the latter scenario.

Stochastic Implementation

The model described above is implemented using stochastically simulation (500 replicates) for a population of animals, with both resistance and production potential modeled as heritable traits. As first approximation $f(R)$ was assumed to be a linear function of R :

$$f(R) = (R - L) / (U - L), \quad (3)$$

where $f(R)$ has bounds of 0 and 1. Initial genotypes for production potential and resistance are randomly drawn from independent normal distributions with means μ_{Pp} and μ_R and additive genetic variances σ_{APp}^2 and σ_{AR}^2 . Environmental effects are randomly drawn from independent normal distributions with mean zero and variance σ_{EPp}^2 and σ_{ER}^2 . Heritabilities for production potential (h^2_{Pp}) and resistance (h^2_R) in the base population are: $h^2 = \sigma_A^2 / \sigma_P^2$, where σ_A^2 and σ_P^2 represent the respective additive genetic and phenotypic variance ($\sigma_P^2 = \sigma_A^2 + \sigma_E^2$). Infection pressure and other environmental conditions are assumed constant over time. Dominance and epistatic effects are ignored. Phenotypic values for observed production are determined for each individual using equations 2 and 3.

Population Structure. A population with a constant size of N individuals is considered, consisting of n_f females and n_m males. Generations are discrete and n_{msel} males and n_{fsel} females are selected for reproduction each generation, resulting in selected fractions $p_m = n_{msel} / n_m$ and selection intensity i_m in the males and $p_f = n_{fsel} / n_f$ and i_f in the females. Each sire is mated to $n_{fm} = n_{fsel} / n_{msel}$ dams. Each dam is mated to one sire and has n_o offspring of which 50% are male and 50% female. All n_o individuals survive to reproduction. The number of generations considered is denoted as n_{gen} .

Selection. Animals are ranked on their phenotype for observed production after which mass selection is applied. When the number of animals with an observation for observed production is less than either n_{msel} or n_{fsel} , (depending on the sex under consideration), additional animals will be randomly drawn from I (i.e., assuming that these animals are capable of reproduction). Genotypes for production potential and resistance in the offspring of selected parents are determined as the average of their parental genotypes plus a Mendelian sampling term. Across generations, the population mean of resistance shifts relative to the two thresholds due to selection. This will change the correlations between production potential and observed production and between resistance and observed production. These correlations are calculated each generation for illustration (i.e., $r_{p_o,z} = \text{cov}(P_o,z) / (\sigma_{P_o} \times \sigma_z)$, where z is either production potential or R).

Deterministic Implementation

The model can also be implemented deterministically. Advantage of deterministic models they can provide more insight in the nature of the problem. By describing the process in terms of means and (reduction in) (co)variances and correlations, the reader gets

more insight in what is actually happening during selection. Deterministic models also lend themselves to future developments such as breeding scheme optimization. Population structure and selection policy is equal to that described for the stochastic simulation. Observed production levels are determined using equations 2 and 3.

General Procedure. Determination of the genetic parameters for each of the three components production potential, resistance, and observed production is an iterative process. It is assumed that means and variances for production potential and resistance are known in the base generation. Also the thresholds for resistance are known so that means and variances for resistance in I, II, and III can be determined. The mean and variances for observed production are only unknown in II (i.e., where the linear function of resistance is applicable) and, therefore, only the mean and variances for observed production in II need to be determined each generation. Mass selection is on observed production and selection differentials are determined in II or III or a combination of both (no selection differential in I since mean and variances are 0). By regression of production potential and resistance on these selection differentials, selection differentials, and genetic response can be determined in production potential and resistance separately. Genetic responses in II and III are combined to an overall response both for production potential and R. New means and variances for production potential and resistance can now be determined, which result in new mean and variances for observed production in II, etc.

Sub-distributions for R and Po. Given the distribution of resistance and the thresholds, the fractions (p_I , p_{II} and p_{III}) of the population present in I, II, and III can be determined, and subsequently means and variances for resistance in each of the population fractions (equations 1 and 2 in Appendix 1). Now the mean and variance of observed production in II (μ^2_{PoII} and σ^2_{PoII}) can be determined (equations 4 and 5 in Appendix 1). Note that the variance for observed production in II will become larger than the variance for production potential, since it includes almost the entire range of possible productions from zero to close to maximum production. The overall mean and phenotypic variance for observed production can be determined from the means and variances in I, II, and III, weighted by the respective fraction (i.e., p_I , p_{II} , or p_{III}) (equations 5 and 6 in Appendix 1).

Selection. Truncation selection is on observed production, resulting in truncation in sub populations II, III, or a combination of II and III, depending on the size of p_{sel} (i.e., p_{sel} is

either p_m or p_f) and the distribution of resistance. Sub populations II and III are overlapping because animals with equal observed production can either have a very high value for production potential and are not fully resistant (II), or have a less high value for P_p , but because they are fully resistant (III), their production is comparable to or higher than animals in II. In case animals are selected from both II and III, the separate selected fractions in II and III, $p_{sel_{II}}$ and $p_{sel_{III}}$ need to be determined. Given p_{sel} and p_{II} and p_{III} , the selected fractions in II and III is obtained by iterating $p_{sel} = p_{sel_{II}}(n_{II}/n_{tot}) + p_{sel_{III}}(n_{III}/n_{tot})$ using subroutine RIDDR_ROOT (Press et al., 1992). As in the stochastic simulation, when the number of animals in II and III is insufficient (i.e., less than either n_{msel} or $n_{f sel}$), additional animals will be randomly drawn from I (i.e., assuming that these animals are capable of reproduction).

Selected proportions differ between males and females. Therefore, reduction in phenotypic and genetic variance in the selected parents due to selection (Bulmer, 1971) as well as selection differentials are determined separately for males and females. Below the general selection procedure is described, which, apart from the input values, is equal for males and females. Since both production potential and resistance are normally distributed, observed production in II consists of a multiplication of two normally distributed traits and, therefore, is not normally distributed itself. However, observed production is assumed to be normally distributed so that selection intensities can be obtained from tables for the normal distribution. Because the selected fraction in the sub-distributions can become very small, the selection intensities are corrected for finite population size (Burrows, 1972). Once selected fractions are known in each sub-distribution, selection differentials for observed production in II and III can be determined.

Selection differential for observed production in II and III is determined for males and females as $S = i \times \sigma_{p_o}$. Selection differentials for production potential and resistance are approximated by regression on observed production (see Appendix 2). Note that these regression coefficients are different for distribution II and III. Genetic selection differential (G) in I, II, and III are calculated by multiplying S with the accuracy of selection squared (heritability in case of mass selection). Total genetic selection differentials per sex (i.e., I, II, and III together) for production potential and for resistance are calculated by weighting the separate responses to the proportion of individuals present in I, II, and III (i.e., p_I , p_{II} , and p_{III}): $G = G_I \times p_I + G_{II} \times p_{II} + G_{III} \times p_{III}$. Overall genetic response is determined by averaging selection differential in males and females. Population means for production

potential and resistance in the next generation are calculated as the sum of the previous population mean and the accumulated genetic selection differential. Population mean for observed production is then determined as outlined in equation 6 (Appendix 1).

Variances. Because the relation between production potential and observed production and between resistance and observed production changes across generations, the mean and variance for observed production need to be calculated each generation from means and variances of resistance and production potential (equations 4 and 5 in Appendix 1). Since selection is on Po, selection intensities (i) in production potential and resistance cannot be determined directly but can be approximated from the respective genetic selection differentials, phenotypic variances and accuracies of selection (i.e., heritabilities in case of mass selection) in generation t (ignoring the subscript for t): $i = G / \sqrt{h^2 \times \sigma_p^2}$, where σ_p^2 is the phenotypic variance of the respective trait. Reduced additive genetic variance due to selection is then calculated as $\sigma_{y_t}^2 = \sigma_{a_t}^2 (1 - k_y h^2)$, where y represents the sex, $\sigma_{a_t}^2$ is the additive genetic variance in generation t , k is the reduction factor due to selection and is calculated as $k = i (i-x)$, where x is the standardized truncation point. The total additive genetic variance in the next generation can now be calculated as $\sigma_{t+1}^2 = 0.25\sigma_{s_t}^2 + 0.25\sigma_{d_t}^2 + \sigma_{ms}^2$, where σ_{t+1}^2 is the additive genetic variance in generation $t+1$, $\sigma_{s_t}^2$ is the additive genetic variance in the selected sires, $\sigma_{d_t}^2$ is the additive genetic variance in the selected dams and σ_{ms}^2 is Mendelian sampling variance, which is equal to half the additive genetic variance in the base population. The phenotypic variance in the next generation is calculated as the genetic variance in the next generation plus the constant environmental variance. Phenotypic correlations between observed production and resistance and between observed production and production potential are calculated each generation for illustration (Appendix 2).

Input Parameters

Population parameters and thresholds were equal for the deterministic and stochastic approach. Initial phenotypic population mean for production potential was 100 with a SD of 15. The heritability for production potential was set to 0.3, resulting in an environmental variance of $225 - (0.3 \times 225) = 157.5$. The phenotypic population mean for resistance was

30, with a SD of 6. The heritability for resistance was varied as described in Table 1, resulting in a different environmental variance for each scheme of $\sigma_p^2 - (h^2 \times \sigma_p^2) = \sigma_e^2$. The lower threshold for resistance was chosen such that in generation one, approximately 17.5% of the individuals were in II and III (i.e., had phenotypic observations for observed production in the first generation), which was less than the proportion of animals needed to produce the next generation. As a consequence, in all cases additional parents were randomly drawn from I, which mimics an extreme situation to start selection. The upper threshold was varied as given in Table 1. Selected proportions for sires and dams were varied as described in Table 1. The total population size $n_{tot} = 384$ remained constant across generations. This population size was chosen such that all selected fractions considered resulted in discrete numbers of animals selected, which was necessary for the stochastic simulation. Production traits considered, may include growth, milk production, egg production, etc. The model would also be suitable to describe the effect of resistance on continuous fertility traits, like number of days between births. However, for those traits it is important that sufficient numbers of animals are available in II and III, because animals in I will be infertile, contrary to our assumption. The results presented, therefore, can be applied to fertility traits, however not from generation one but some generations later onwards.

Results

General Features of the Model

Figure 2 shows the change in population mean for observed production, production potential, and resistance, during 50 generations of mass selection on observed production. The figure shows that if starting with a population with a low level of resistance, then by selecting on observed production, emphasis in the first 15-20 generations mainly is on improving the level of resistance. Increase in observed production in generation zero to six is small because there are not enough animals in II and III yet, which restricts intensity of selection. Observed production is increasing most rapidly during generation 6 to 15, which is due to the increasing level of resistance. Given the heritability for resistance presented in the figure, the increase during this period is more rapid than when selection is purely on underlying production (i.e., resistance level above the higher threshold). This is because the variance for observed production is largest in II (i.e., in II individuals can have values for observed production between zero and production potential). Therefore, selection response

for observed production will increase with more individuals coming into II and is highest when most are in II. The variance for observed production will decrease again when the population moves into III. Resistance level increases almost linearly until generation 15, after which the increase flattens off and observed production approximates production potential.

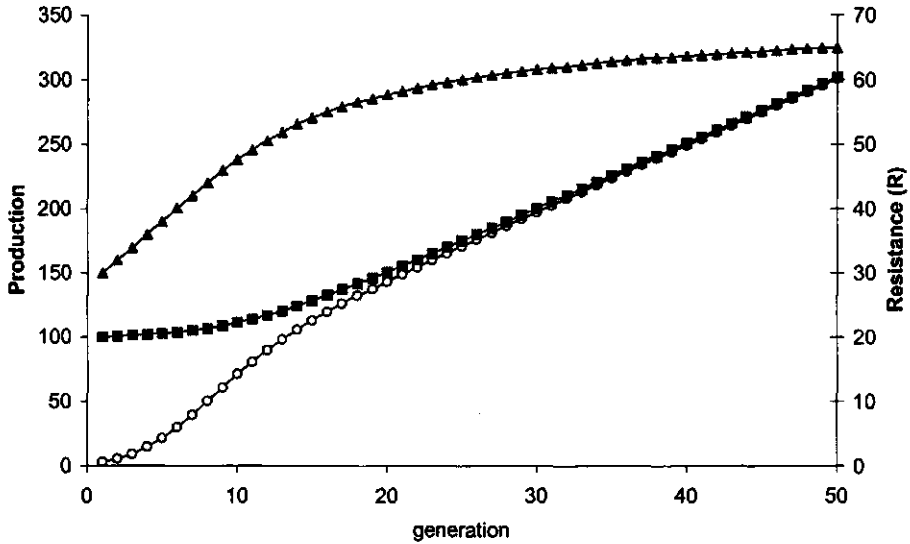


Figure 2. Stochastic simulation results for the change in phenotypic population mean for observed production (o), resistance (\blacktriangle), and production potential (\blacksquare) during 50 generations of phenotypic selection on observed production, where $h^2_R = 0.2$, difference between lower and upper threshold = 3.0 SD, and $p_{sire} = 0.0625$.

Figure 3 shows the change in number of individuals per sub population for the base situation. In generation one, about 85% of the individuals have resistance level lower than the lower threshold value (L). Selection on observed production results in a correlated response in resistance. Between approximately generation 6 and 15, most animals are in II. This period coincides with the period with the strongest increase in observed production (Figure 2). From generation 10 onwards, the fraction in distribution III is increasing, resulting in an increase in selection pressure on Pp. From generation 15 onwards, more animals are present in III compared to II, resulting in a decreasing selection response in resistance (start of the plateau in Figure 2) and a decrease in selection response in Po.

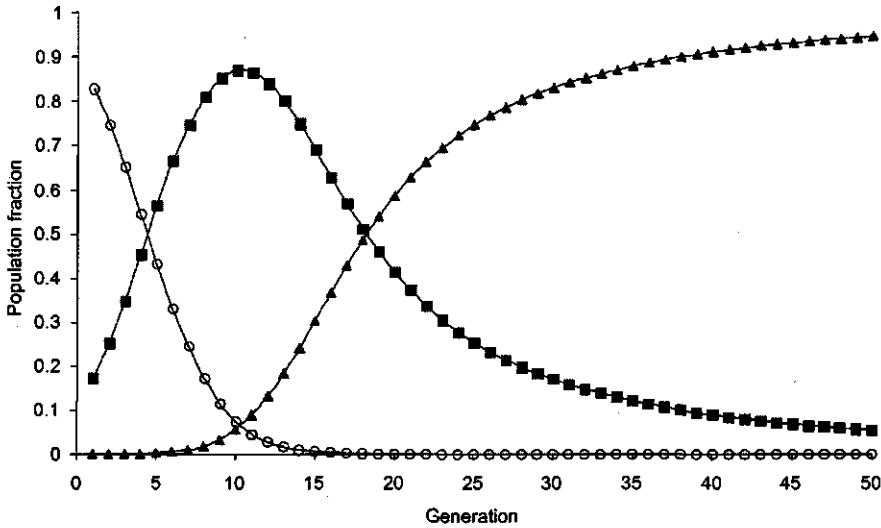


Figure 3. Stochastic simulation results for the change in number of animals in I (o), II (■), and III (▲) during 50 generations of phenotypic selection on observed production, where $h^2_R = 0.2$, difference between lower and upper threshold = 3.0 SD, and $p_{sire} = 0.0625$.

Figure 4 shows the change in phenotypic correlation between observed production and production potential ($r_{Po,Pp}$) and between observed production and resistance ($r_{Po,R}$), again for the base situation. The change in correlation between observed production and resistance is informative. It starts at about 0.6, then goes up with increasing number of individuals in II with the highest correlation in generation 8 and 9 (i.e., highest number of animals in II). Subsequently it decreases again. It does not reach zero, because there are still animals in II at generation 50. The correlation between observed production and production potential increases with increasing number of animals in III (observed production is approaching production potential). This is as would be expected from the change in $f(R)$, when going from I to II and III. The correlation does not reach 1.0 in generation 50, since not all animals are in III (see Figure 3). This at first sight is unexpected, but is caused by the large variance in observed production, resulting from the small population fraction in II. For clarification we will give a numerical example here, representing generation 50 in the figure. The $cov(Po,Pp) = E(Po \times Pp) - E(Po) \times E(Pp)$, and $E(Po \times Pp)$ in I (=0), II (= 7,1887.2) and III (=83,619.5) weighted by p_1 (= 0), p_2 (=0.054) and p_3 (=0.946), resulting in $E(Po \times Pp) =$

82,986.0. In generation 50 $\mu_{Pp} = 288.8$ with (phenotypic) $\sigma^2_{Pp} = 214.0$ and $\mu_{Po} = 286.6$ with $\sigma^2_{Po} = 367.6$. Combining this information gives $r_{Po,Pp} = (82,986.0 - (288.8 \times 286.6)) / \sqrt{(214.0 \times 367.6)} = 215.92 / 280.48 = 0.77$, which is given in the figure.

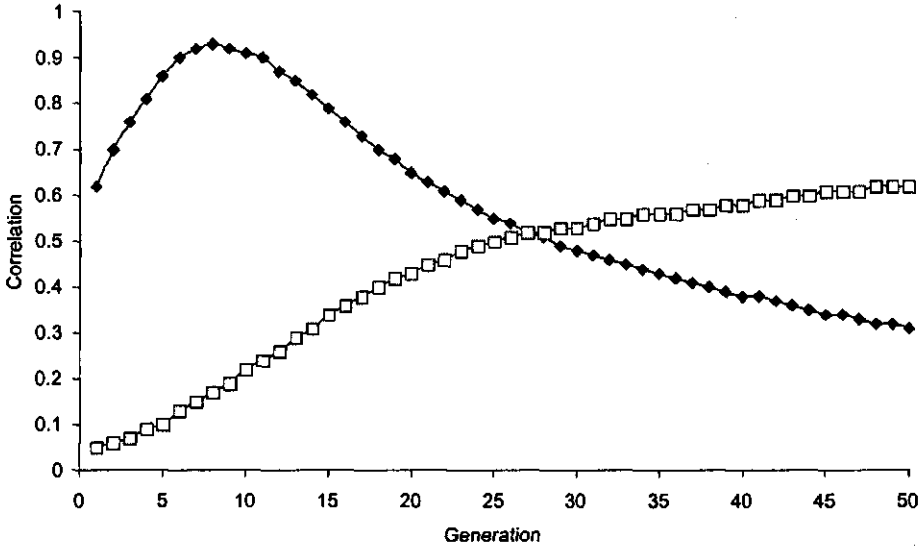


Figure 4. Stochastic simulation results for the phenotypic correlations between observed production and resistance (◆) and between observed and potential production (□) during 50 generations of phenotypic selection, where $h^2_R = 0.2$, difference between lower and upper threshold = 3.0 SD. and $v_{inv} = 0.0625$.

The correlation between observed production and resistance will not become equal to 1.0 due to the fact that the genetic and phenotypic correlation between production potential and resistance is equal to zero. Once they are in sub population III, animals with the highest production can have any value for resistance above the upper threshold. Parents selected from III with values for resistance just above the upper threshold, will have offspring of which some have values for resistance below the upper threshold due to Mendelian sampling and environmental influences. Because animals in II have lesser chance of being selected, offspring of parents with resistance just above U with positive Mendelian sampling are favored. So, indirectly there still is some selection pressure on resistance and the correlation between observed production and production potential will approximate 1.0 asymptotically. The speed at which the asymptote is approximated is dependent on factors such as the heritability for resistance, difference between the lower and upper threshold, and intensity of selection.

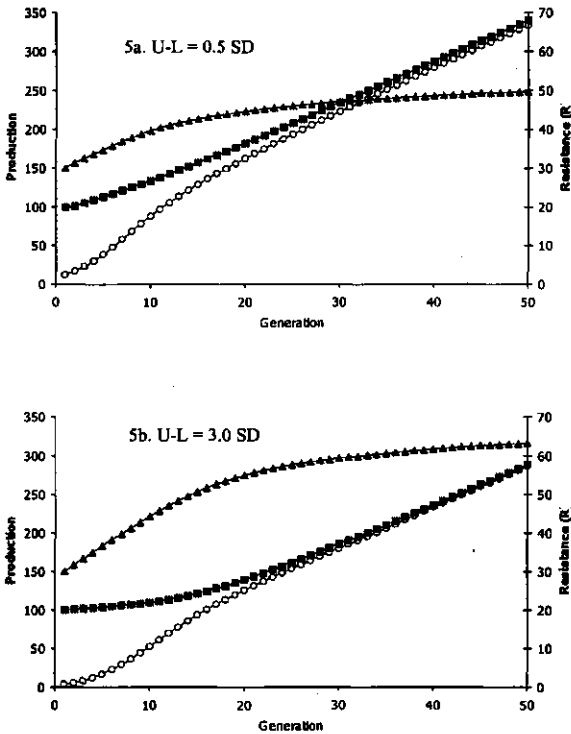


Figure 5 a+b. Stochastic simulation results for the change in response in observed production (o), resistance (\blacktriangle), and production potential (\blacksquare) due to increase in difference between lower and upper threshold from 0.5 phenotypic standard deviation for resistance (a), to 3.0 SD (b), while $h^2_R = 0.2$ and one out of 16 males and one out of two females is selected for breeding.

Consequences of Changing Parameters. The consequences of changing the difference between lower and upper threshold have been investigated. The lower threshold was kept constant in all cases. In Figure 5, consequences of changing the difference between lower and upper threshold from 0.5 phenotypic SD. (5a) to 3.0 phenotypic SD (5b) are given. Increasing the difference between the thresholds results in an increase in the level for resistance at which the leveling off towards the plateau starts. The generation number at which the plateau starts therefore shifts from about 15 for 0.5 SD to about 20 for 3.0 SD. For 0.5 SD difference, observed production does not become equal to production potential for reasons mentioned earlier, even though the plateau in resistance has been reached at an earlier generation than for 3.0 SD, where observed

production does become approximately equal to production potential from generation 40 onwards. Animals that are in II are influenced by the infection to an extent that is described by $f(R)$. A small difference between lower and upper threshold (e.g., 0.5 SD), would result in a steep slope of $f(R)$ in II, so that the few animals in the tail of the distribution can be largely influenced by the infection. Whereas in case of a large (e.g., 3.0 SD) difference between lower and upper threshold, the few animals in II would hardly be influenced at all, since $f(R)$ is close to one. In spite of the fact that observed production does not become equal to production potential when difference between lower and upper threshold is small,

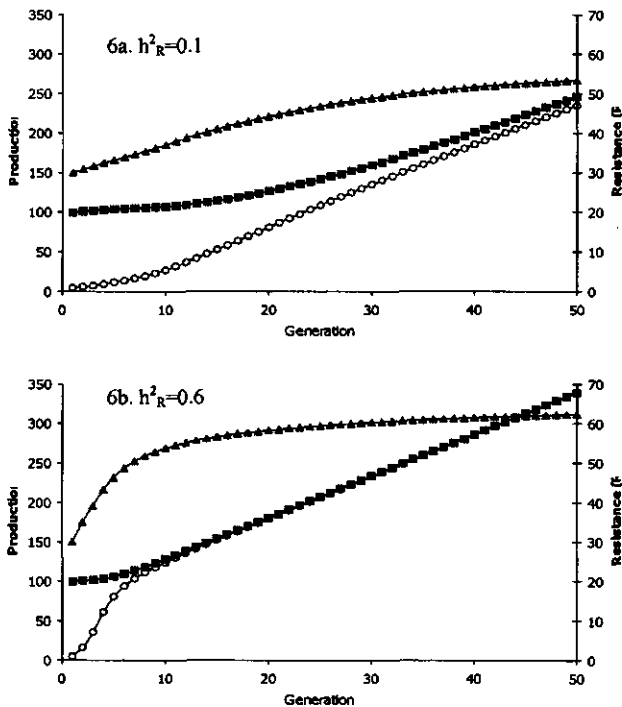


Figure 6 a+b. Stochastic simulation results for the change in response in observed production (○), resistance (▲) and production potential (■) due to increase in heritability for resistance from $h^2_R = 0.1$ (a) to 0.6 (b), while difference between lower and upper threshold = 2.0 SD and one out of 16 males and one out of two females is selected for breeding.

0.61 for Marek's infection [Gavora et al., 1974]). Figure 6 shows the consequences of changing the size of the heritability for resistance (R) from 0.1 (6a) to 0.6 (6b). As expected selection response increases with increasing heritability, so with a high heritability for resistance, the difference between lower and upper threshold can be overcome more efficiently and selection weight shifts from resistance to production potential in an earlier generation. With a low heritability for resistance, the population-mean for resistance shifts more slowly, resulting in a larger portion of animals in I and II after 50 generations of selection on observed production (Figure 6a).

Selection intensity in the population depends on the type of breeding scheme and is largely influenced by the number of offspring per parent. For example, in the case of natural

such a small difference is easier to overcome than a large one. Result is that in case of a small difference, selection weight shifts from resistance to production potential at an earlier stage, and more genetic progress in observed production and production potential can be made during the same number of generations.

Selection response depends on the size of the heritability. For some infections the heritability will be low (e.g., 0.044 for mastitis or 0.016 for reproductive infections in cattle [Nielsen et al., 1996]), while for others the heritability will be much higher (e.g., 0.39 for tick resistance in cattle [Mackinnon et al., 1991] or

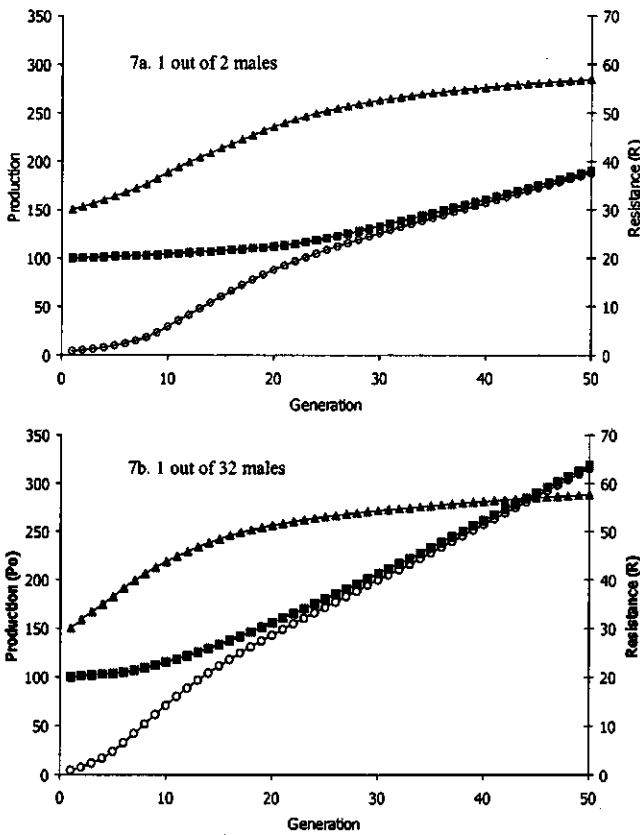


Figure 7 a+b. Stochastic simulation results for the change in response in observed production (o), resistance (▲) and production potential (■) due to increase in the selected proportion in the males from one out of two (a, above) to one out of 32 (b, below), while $h^2_R = 0.2$ and the difference between lower and upper threshold = 2.0 SD.

mating, the number of offspring per parent is limited. The possibility of using AI would allow for a large number of offspring per male, and thus for a higher selection intensity in males. When modern reproduction techniques (e.g., MOET, ovum pickup) are available, selection intensity in females could be increased as well. In Figure 7, consequences of decreasing the selected proportion in males from one out of two (7a) to one out of 32 (7b) are shown. Main difference is that selection is more efficient, and thus selection response is increased, with decreased selected proportion. Selection response for observed production increases and selection pressure shifts from

resistance to production potential at an earlier stage. The population-mean for resistance does not change with decreasing selected proportion since the level for resistance at which the plateau is reached, remains equal for both selection intensities.

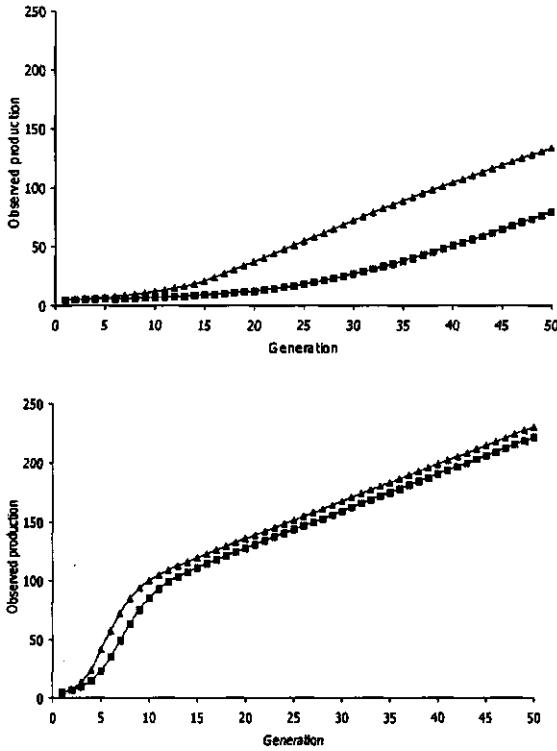


Figure 8 a+b. Example of most extreme difference between deterministic and stochastic simulation results, represented by response in P_0 using the deterministic (\blacktriangle) and the stochastic (\blacksquare) approach, for heritability of resistance of 0.1 (a) and 0.6 (b) and the selected proportion in the males and females equals one out of two and difference between lower and upper threshold = 2.0 SD.

Comparison of Stochastic and Deterministic Prediction

Predictions obtained from stochastic simulation are compared to predictions using the deterministic model. Figure 8 shows the difference in change in population mean in observed production using the stochastic or the deterministic approach for equal selected proportion in males and females ($p_d = p_s = 0.5$), $U - L = 2.0$ s.d. and heritability for resistance is either 0.1 (a) or 0.6 (b). To give a clear example, Figure 8a represents the combination of breeding scheme and parameters resulting in the most extreme difference between deterministic and stochastic prediction. Figure 8b represents a situation with an average difference between stochastic and deterministic simulation results (see Table 1). This figure clearly shows that the difference between

both approaches is mainly caused during the generations in which many animals are in II. As soon as most animals are in III, the increase in difference across generations is negligible. A low heritability for resistance (Figure 8a) results in a low selection response in observed production (and thus in resistance) and thus in a prolonged number of generations with a large proportion of the population in II. The non-normality in II is causing an overestimation of increase in number of animals in II and III when using the deterministic approach. A distribution that is built up from a multiplication of two normally distributed variables has thicker tails than a normal distribution. Applying selection theory that has been

developed for normally distributed traits, therefore, would result in selection of a too large number of animals from this non-normal distribution. So the longer the route to get the animals in III (e.g., due to increase in difference between lower and upper threshold, low selection intensity or low heritability for resistance), the larger the difference in prediction between the stochastic and the deterministic model.

Table 1 shows mean differences in stochastic and deterministic predictions in population mean across generations. Four mating schemes are compared (selected proportion in males equal 0.5 (A), 0.125 (B), 0.0625 (C), and 0.03125 (D), respectively,

*Table 1. Mean errors in population mean for observed production, comparing results using the deterministic model versus the stochastic simulation, relative to production potential, across 50 generations (lower threshold is constant)**

Mating scheme	h^2_R	U = 0.5 SD	U = 1 SD	U = 2 SD	U = 3 SD
A	0.1	0.112	0.126	0.306	0.299
	0.2	0.105	0.162	0.217	0.254
	0.4	0.045	0.080	0.115	0.138
	0.6	0.019	0.045	0.070	0.086
B	0.1	0.029	0.028	0.079	0.104
	0.2	0.031	0.020	0.051	0.070
	0.4	0.022	0.017	0.036	0.050
	0.6	0.021	0.013	0.025	0.037
C	0.1	0.120	0.016	0.052	0.080
	0.2	0.080	0.019	0.031	0.057
	0.4	0.048	0.021	0.027	0.041
	0.6	0.034	0.020	0.026	0.034
D	0.1	0.139	0.091	0.044	0.073
	0.2	0.107	0.114	0.026	0.057
	0.4	0.078	0.072	0.040	0.048
	0.6	0.064	0.059	0.039	0.047

* In mating scheme A selected proportion in males was 0.5, in mating scheme B it was 0.125, in mating scheme C it was 0.0625, and in mating scheme D it was 0.03125. Selected proportion in females was 0.5 in all mating schemes. The upper threshold U is expressed in phenotypic standard deviations of resistance.

selected proportion in females equals 0.5), each for four different heritabilities for resistance ($h^2_R = 0.1, 0.2, 0.4, \text{ or } 0.6$) and four differences between lower and upper threshold (0.5, 1.0, 2.0, or 3.0 s.d.). The lower threshold was kept constant in all cases. The mean error (across 50 generations) is the average across generations of (ignoring the subscript for generation t): $(\mu_{Po(d)} - \mu_{Po(s)}) / \mu_{Pp(s)}$, where $\mu_{Po(d)}$ is the population mean in generation t for observed production, determined using the deterministic model, $\mu_{Po(s)}$ using the stochastic model and $\mu_{Pp(s)}$ is the population mean in generation t for production potential determined using the stochastic model. The differences between deterministic and stochastic predictions have been related to production potential and not to observed production, because observed production takes very low values during the first few generations. Small deviations from a small value can result in very large relative differences, which would give a distorted view on the situation. Small values in the table indicate that on average the prediction given by the deterministic model matched the prediction by the stochastic model very well (e.g., mating scheme A, $U = 0.5$ s.d. and $h^2 = 0.6$). Given this information, general trend in Table 1 is that differences increase with increasing difference between lower and upper threshold (i.e., longer route through II to III) and decrease with increasing heritability (i.e., increase in selection response in resistance, and thus increase in efficiency to get animals in III). Trends in differences between stochastic and deterministic models across mating schemes are less clear.

Differences between stochastic and deterministic predictions especially occur in situations where large proportions of males are selected (Table 1). However, such schemes usually are not applied in practical breeding (i.e., selection is on production, which in most breeding schemes gets highest priority). Selection intensities in mating schemes C and D, where differences are small, can be met. That is, in practical dairy cattle breeding, selected proportion for production of 1 out of 32 in males is not unusually small. Influence of the size of the heritability for resistance on the selection response and thus on the number of generations needed to get the largest proportion of the population in III, is infection type (and population) dependent. Increase in selection response due to decrease in difference between lower and upper threshold depends, apart from the infection trait, also on the severity of infection. When the size of the differences increase, e.g., with low heritability for resistance and small difference between lower and upper threshold, stochastic simulation would be the method of choice.

Discussion

General Features of the Model.

In the present paper a model has been presented to describe the relation between level of disease resistance and production. The model allows for the production to increase proportionally with increasing level of disease resistance, until the production potential is reached. The model illustrates that in practice level of disease resistance will increase, provided the population is not fully resistant yet, also when selection is on observed production only. The lower threshold for disease resistance was introduced to mimic a point of resistance below which animals do not produce anymore. An important property of the model is that the result of selection on observed production results in a shift from an increase in level of disease resistance to an increase in production potential across generations. As a consequence, correlations between observed production and both production potential and level of disease resistance change across generations under the infinitesimal model. The model allows for non-linear selection response in observed and underlying production and disease resistance. Furthermore, for selection schemes that are realistic in practical animal breeding, the deterministic approach accurately predicts genetic progress under the present model and allows insight into the nature of the responses.

Support of the Model by Literature Results.

Experimental data supports the results obtained in this study. Spencer *et al.* (1979) detected less Lymphoid leukosis virus and less group specific antigens in albumen of eggs in strains of laying hens that were selected for high production, as compared to the random bred control lines (both under normally occurring constant infection pressure). In concept, resistance to infectious diseases with continuous infection pressure is comparable to resistance to other environmental stresses like excessive heat. Frisch (1980) has presented results from an extensive long-term selection experiment in cattle, where the animals were selected for observed production. His results show that increase in growth rate due to selection for observed growth rate can be primarily attributed to improvement in resistance to environmental stresses. It is likely that his cattle had low level of resistance when he started selection for growth rate. Growth rate improved by indirectly improving level of resistance and that is what the proposed model would predict. Frisch (1980) also states that selection for growth rate under stressful environment eventually would eliminate loci responsible for the lack of resistance, and therefore, selection moves to genes affecting

production potential. Within our model, the effect of alleles that explain poorer resistance asymptotically decreases, and eventually, probably due to genetic drift, these alleles are likely to disappear from the population (correlation between production potential and observed production becomes 1.0) due to the absence of selection pressure in favor of those alleles. Results from Mackinnon *et al* (1991) and Burrow (unpublished results) show that heat resistance increased with multiple generations selection for increased observed production. Burrow (unpublished results) also found an improved level of tick resistance in the more tick-susceptible line when selection was for increased growth level.

Other Models.

The relation between disease resistance and observed production under infection pressure has been studied to some extent both through data analysis (e.g., Spencer, 1979; Burrow, unpublished results) and simulation (e.g., Bishop and Stear, 1999). In simulation studies, the relation between disease resistance and production has been acknowledged, however, assumed to be linear (bivariate normal distribution). It does not allow for the presence of a plateau above which there is no noticeable effect on production anymore. Assuming a linear relation between production and resistance, re-estimation of covariances between resistance and observed production and between observed production and production potential each generation would account for such a plateau. However, in practice it is impossible to disentangle the relation between production potential and level of disease resistance from information on observed production only. Information on production potential is difficult to obtain under infection pressure and often no accurate parameters are available to measure resistance. The model presented here could be used without having to re-estimate the genetic parameters at regular basis, since the model already accounts for that.

Other Applications.

The application of the model presented here is not limited to the relation between production and disease resistance. As already mentioned the model might also fit to the influence of environmental stress on production, as discussed by Frisch (1980). Beilharz *et al.* (1993) argue there is a plateau in the total amount of energy an animal can utilise for all processes. When more energy is needed to overcome infections, resist leg problems, fertility problems, or metabolic infections, less energy can be spend on production. There appears to be genetic variation in the efficiency with which an animal can respond to stresses like environment, parasites, or pathogens. The model presented in this paper could be of help to

try to find the optimum combination of production potential and resistance to various types of stresses, so that animals will be able to produce at higher level under less optimal circumstances.

Implications

This paper has two important conclusions. Firstly, it is not necessary to measure the level of resistance in order to increase it. Because observed production is determined by production potential as well as by level of resistance, observed production can be considered a selection index in which the weighting of the components is already taken care of. Secondly, it is important to realize that observed production and production potential are not the same trait. Imagine one half of a partially resistant population is selected for production in an infected environment and one half in an uninfected environment. When afterwards both populations are compared in an uninfected environment, animals that were in the infected environment will show lower production. This is not because of a negative correlation between resistance and production, but because indirectly part of the selection weight in the infected environment was on resistance.

Acknowledgement

The authors would like to thank WOTRO (Netherlands Foundation for the Advancement of Tropical Research) for financial support. SCB was funded by the Biotechnology and Biological Sciences Research Council.

References

- Beilharz, R. G., B. G. Luxford, and J. L. Wilkinson. 1993. Quantitative genetics and evolution: Is our understanding of genetics sufficient to explain evolution? *J. Anim. Breed. Genet.* 110:161-170.
- Bennett, R.M., K. Christansen, and R. S. Clifton-Hadley. 1999. Modelling the impact of livestock infection on production: case studies of non-notifiable infections of farm animals in Great Britain. *Anim. Sci.* 68:681-689.
- Bishop, S.C., and M. J. Stear. 1999. Genetic and epidemiological relationships between productivity and disease resistance: gastrointestinal parasite infection in growing lambs. *Anim. Sci.* 69:515-524.

- Bisset, S.A., and C.A. Morris. 1996. Feasibility and implications of breeding sheep for resilience to nematode challenge. *Int. J. for Parasit.* 26:857-868.
- Bulmer, M.G. 1971. The effect of selection on genetic variability. *The American Naturalist* 105:201-211.
- Burrows, P.M. 1972. Expected selection differentials for directional selection. *Biometrics* 28:1091-1100.
- Frisch, J.E. 1980. Changes occurring in cattle as a consequence of selection for growth rate in a stressful environment. *J. Agric. Sci. Camb.* 96:23-38.
- Gavora, J.S., A. A.Grunder, J. L. Spencer, R. S. Gowe, A. Robertson, and G. W. Speckmann. 1974. An assessment of effects of vaccination on genetic resistance to Marek's infection. *Poultry Science* 53:889-897.
- Gavora, J.S., and J. L. Spencer. 1978. Breeding for genetic resistance to infection: specific or general? *World's Poultry Science J.* 34:137-148.
- Lynch, M. and B. Walsh. 1998. Genetics and analysis of quantitative traits. first ed. Sinauer Associates, Sunderland, MA.
- Mackinnon, M. J., K. Meyer, and D. J. S. Hetzel. 1991. Genetic variation and covariation for growth, parasite resistance and heat tolerance in tropical cattle. *Livest. Prod. Sci.* 27:105-122.
- Murray, M., J. C. M. Trail, and G. D. M. D'Ieteren. 1990. Trypanotolerance in cattle and prospects for the control of trypanosomiasis by selective breeding. *Rev. sci. tech. Off. Int. Epiz.* 9:369-386.
- Nielsen, U. S., G. A. Pedersen, and J. Pedersen. 1996. Indices for resistance against infections. Proc. Of the open session of the Interbull meeting, Veldhoven, the Netherlands, June 23-24, 1996. Pp.161-166.
- Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. 1992. Numerical Recipes in Fortran; The art of scientific computing. 2nd edition. Cambridge University Press, Cambridge, MA.
- Spencer, J. L., J. S. Gavora, and R. S. Gowe. 1979. Effect of selection for high egg production in chickens on shedding of lymphoid leukosis virus and gs antigen into eggs. *Poultry Science* 58:279-284.
- Woolaston, R. R., and R. L. Baker. 1996. Prospects of breeding small ruminants for resistance to internal parasites. *Int. J. for Paras.* 26:845-855.

Appendix 1

Determining the mean and phenotypic variance of resistance and observed production in part II of the distribution.

Mean and phenotypic variance of resistance in II

The distribution of resistance is divided into three parts, separated by the lower (L) and upper (U) thresholds. The overall mean can be determined by summing the weighed expectations of each of the three parts:

$$\mu_R = p_I(\mu_{R_I}) + p_{II}(\mu_{R_{II}}) + p_{III}(\mu_{R_{III}})$$

The value for the expectations of each of these parts can be determined by treating them as the sum of selection response after phenotypic selection and overall mean. Selection intensities i_I and i_{III} are derived from the fractions of the population in each part of the distribution, p_I , p_{II} and p_{III} .

$$\mu_{R_{II}} = [(\mu_R - p_I(\mu_R - i_I\sigma_R) - p_{III}(\mu_R + i_{III}\sigma_R))] / p_{II} \quad (1)$$

The phenotypic variances in each of the separate parts are corrected for the shifting mean. The phenotypic variance for resistance in II can be determined by subtracting the weighed phenotypic variances in I and III and correction for II from the overall variance of resistance. Reduction of variance due to selection is taken into account by introducing the k -values in I and III, k_I and k_{III} , (i.e., $k = i(i - x)$) (Bulmer, 1971). The phenotypic variance is a combination of the variances in I, II and III:

$$\sigma_R^2 = \{p_I[(\mu_{R_I} - \mu_R)^2 + (1 - k_I)\sigma_R^2] + p_{II}[(\mu_{R_{II}} - \mu_R)^2 + \sigma_{R_{II}}^2] + p_{III}[(\mu_{R_{III}} - \mu_R)^2 + (1 - k_{III})\sigma_R^2]\}$$

So the phenotypic variance of resistance in II can now be determined as (2):

$$\sigma_{R_{II}}^2 = \frac{\{(\sigma_R^2 - p_I[(\mu_{R_I} - \mu_R)^2 + (1 - k_I)\sigma_R^2] - p_{III}[(\mu_{R_{III}} - \mu_R)^2 + (1 - k_{III})\sigma_R^2]\}}{p_{II} - (\mu_{R_{II}} - \mu_R)^2}$$

Mean and phenotypic variance of observed production in II

The mean of observed production in I = 0 and in III = mean production potential. The mean of observed production in II can be determined following the definition of P_o , as a function of resistance and production potential:

$$\mu_{P_{oII}} = \left[\frac{\mu_{R_{II}}}{(U-L)} - \frac{L}{(U-L)} \right] \times \mu_{P_u} \quad (3)$$

Also, the phenotypic variance of observed production in I = 0 and in III = σ^2_{Pp} . The variance of observed production in II can be determined using the definition (as mentioned by Lynch and Walsh, 1998, p. 118) and ignoring the function of R:

$$\sigma^2_{P_{oII}} = \sigma^2_{P_u} \times \mu^2_{R_{II}} + \sigma^2_{P_u} \times \sigma^2_{R_{II}} + \mu^2_{P_u} \times \sigma^2_{R_{II}} \quad (4)$$

Including the function of R this results in:

$$\sigma^2_{P_{oII}} = \frac{1}{(U-L)^2} \times (\sigma^2_{P_u} \times (\mu_{R_{II}} - L)^2 + \sigma^2_{P_u} \times \sigma^2_{R_{II}} + \mu^2_{P_u} \times \sigma^2_{R_{II}})$$

Note that observed production is not normally distributed in II, since it is obtained by multiplying two normally distributed traits.

Overall mean and variance for observed production can now be determined as follows:

$$\mu_{P_o} = p_I \mu_{P_{oI}} + p_{II} \mu_{P_{oII}} + p_{III} \mu_{P_{oIII}} \quad (5)$$

$$\sigma^2_{P_o} = p_I [(\mu_{P_{oI}} - \mu_{P_o})^2 + \sigma^2_{P_{oI}}] + p_{II} [(\mu_{P_{oII}} - \mu_{P_o})^2 + \sigma^2_{P_{oII}}] + p_{III} [(\mu_{P_{oIII}} - \mu_{P_o})^2 + \sigma^2_{P_{oIII}}] \quad (6)$$

Appendix 2

Regression coefficients and phenotypic correlations for observed production on resistance and for observed production on production potential

Regression coefficients

Selection response of production potential and resistance in II can be determined by regressing observed production in II on resistance in II and on production potential in II, respectively. In general regression coefficient can be determined as:

$$b_{x,y} = \frac{\text{cov}(x, y)}{\text{var}(y)} = \frac{E(x \times y) - E(x)E(y)}{\text{var}(y)}$$

Using this, the regression coefficient of observed production on resistance can be determined as follows:

$$\text{cov}(P_o, R) = E(P_o \times R) - E(P_o) \times E(R)$$

$$E(P_o \times R) = E\left(\frac{R-L}{U-L} \times P_u \times R\right)$$

$$= [\mu_{P_u} / (U-L)] \times (\mu_{R_{II}}^2 + \sigma_{R_{II}}^2 - L \times \mu_{R_{II}})$$

$$E(P_o) \times E(R) = E\left(\frac{R-L}{U-L} \times P_u\right) \times E(R) = \frac{\mu_{R_{II}}^2 - L \times \mu_{R_{II}}}{U-L} \times \mu_{P_u}^2$$

$$\text{cov}(P_o, R) = (\mu_{P_u} / (U-L)) \sigma_{R_{II}}^2$$

$$b_{P_o, R} = \text{cov}(P_o, R) / \sigma_{P_o_{II}}^2$$

Using the same reasoning as for the regression of observed production on R, the regression coefficient of observed production on production potential can be written as:

$$b_{Po, Pp} = \text{cov}(Po, Pp) / \sigma_{PoII}^2$$

where $\text{cov}(Po, Pu) = (\sigma_{Pu}^2 / U - L) \times (\mu_{RII} - L)$

Phenotypic correlations between observed and potential production and between observed production and resistance

Phenotypic correlations between observed production and production potential and between observed production and resistance are determined across I, II, and III. This is in contrast to the regression components, which are used for determining the relation between observed production and production potential and between observed production and resistance in II alone.

The correlation between observed production and resistance is determined as:

$$r_{Po, R} = \text{cov}(Po, R) / (\sigma_{Po} \times \sigma_R)$$

where σ_{Po} is the phenotypic standard deviation for observed production for I, II, and III together. The total phenotypic variance for observed production is determined as:

$$\sigma_{Po}^2 = p_I ((\mu_{PoI} - \mu_{Po})^2 + \sigma_{PoI}^2) + p_{II} ((\mu_{PoII} - \mu_{Po})^2 + \sigma_{PoII}^2) + p_{III} ((\mu_{PoIII} - \mu_{Po})^2 + \sigma_{PoIII}^2)$$

where μ_{PoI} is the mean for observed production in I, σ_{PoI}^2 is the variance for observed production in I, p_I is the fraction of the total population that is in I and μ_{Po} is determined as the means in I, II and III weighted by the fractions in I, II and III.

$$\text{cov}(Po, R) = (p_1 \times \text{cov}(Po_I, R) + p_2 \times \text{cov}(Po_{II}, R) + p_3 \times \text{cov}(Po_{III}, R)) - \mu_R \times \mu_{Pu}$$

cov in I = 0

cov in II = as in regression

cov in III = $\mu_{Pp} \times \mu_{RIII}$

The correlation between observed production and production potential is determined as:

$$R_{P_o, P_p} = \text{cov}(P_o, P_p) / \sigma_{P_o} \times \sigma_{P_p}$$

Where the cov is determined similar as the cov (Po,R) as described above.

cov in I = 0

cov in II = as in regression

cov in III = $\mu_{P_p}^2 + \sigma_{P_p}^2$

VI Using genetic markers for disease resistance to improve production under constant infection pressure

Animals will show reduced production when exposed to a constant infection pressure unless they are fully resistant, the size of the reduction depending on the degree of resistance and the severity of infection. In this article, the use of QTL for disease resistance for improving productivity under constant infection pressure is investigated using stochastic simulation. A previously published model was used with two thresholds for resistance: a threshold below which production is not possible and a threshold above which production is not affected by the infection. In between thresholds, observed production under constant infection is a multiplicative function of underlying potential production and level of resistance. Some simplifications of reality were adopted in the model, such as no genetic correlation between potential production and resistance, the absence of influence of lack of resistance on reproductive capacity, and the availability of phenotypes in both sexes. Marker assisted selection was incorporated by assuming a proportion of the genetic variance to be explained by the QTL, which thus is defined as a continuous trait. Phenotypes were available for production, not for resistance. The infection pressure may vary across time. Results were compared to mass selection on production under constant, as well as intermittent infection pressure, where the infection pressure varied between but not within years. Selection started in a population with a very poor level of resistance. Incorporation of QTL information is valuable (i.e. the increase in observed production relative to mass selection) when a large proportion of the additive genetic variance is explained by the QTL (50% genetic variance explained) and when the heritability for resistance is low ($h^2_R = 0.1$). Under constant infection pressure, incorporating QTL information does not increase selection responses in observed production when the QTL effect explains less than 25% of the genetic variance. Under intermittent selection pressure, the use of QTL information gives a slightly greater increase in observed production in early generations, relative to mass selection on observed production, but still only when the QTL effect is large or the heritability for resistance is low. The additional advantage of incorporating QTL information is that use of (preventive) medical treatment is possible, or animals may be evaluated in non-infected environments.

Key Words: Disease Resistance, Genetics, Production, QTL, Selection, Stochastic Model

This Chapter has been accepted for publication

E.H. van der Waaij, P. Bijma, S.C. Bishop, and J.A.M. Van Arendonk. Using genetic markers for disease resistance to improve production under constant infection pressure. Accepted by the Journal of Animal Science

Introduction

Animals that are not fully resistant to a certain disease may show a decreased production when infected. A high infection pressure combined with a low level of resistance may cause production to drop to, or even below (in case of growth) zero. When the level of resistance is high, infection may have little or no influence on performance. Van der Waaij et al. (2000) presented a model describing the relation between production under constant infection pressure and level of resistance. Their results suggest that selection for production under constant infection pressure results in an increase in potential production (the production that would have been achieved in the absence of infection), as well as in resistance. Trypanosomosis is an example of an important disease with constant infection pressure. Trypanotolerant local cattle breeds do exist, but both their resistance and production may be further improved.

A disadvantage of the model presented by Van der Waaij et al. (2000) is that exposure to the pathogen is required in order to get phenotypic observations for observed production. An alternative option would be to keep the animals in an uninfected environment, or under medication, and combine potential production with QTL for disease resistance, to predict production under constant infection pressure. Also, in Van der Waaij et al. (2000) it was assumed that infection pressure was constant. In reality, infection pressure may vary across time, resulting in intermittent (indirect) selection pressure on resistance when mass selection on observed production is applied. The use of QTL would enable selection pressure on resistance, irrespective of the presence of infection pressure.

The objective of this article is to investigate the use of resistance QTL for improving productivity under constant, as well as intermittent infection pressure, avoiding exposing the animals to infection, where infection pressure is assumed constant within a season but different between seasons.

Material and Methods

Model

The model of Van der Waaij et al. (2000) was used to describe the relationship between level of disease resistance (R) and observed production (P_o) under infection pressure. In that model, the level of production that would have been achieved if the animal were completely resistant is denoted as production potential (P_p). Under infection pressure,

the level of resistance of an animal to the infection affects observed production, so that observed production is a function of both production potential and resistance:

$$P_o = f(R) \times P_p, \quad (1)$$

where $f(R)$ is a function of resistance, which describes the effect of resistance on observed production. Van der Waaij et al. (2000) distinguished three categories of resistance, which are separated by thresholds: 1) Animals are fully susceptible when their level of resistance is below the lower threshold (L), in which case production is no longer possible ($P_o = 0$). 2) Animals are fully resistant when their level of resistance is above the upper threshold (U), in which case observed production equals production potential ($P_o = P_p$). 3) Animals between both thresholds are partly resistant and their observed production will be lower than their production potential, the size of the decrease depending linearly on the level of resistance. Thus the effect of resistance on observed production is summarized by:

$$\begin{aligned} f(R) &= 0 && \text{for } R < L \\ f(R) &= (R - L) / (U - L) && \text{for } L < R < U \\ f(R) &= 1 && \text{for } R > U \end{aligned} \quad (2)$$

The thresholds were assumed to be fixed and their values depend on the type of infection and other environmental factors. It is assumed that resistance and production potential are normally distributed, that there is no correlation between production potential and resistance, and that values for resistance are non-negative.

Selection

Two selection strategies were considered. Selection was either based on the phenotype for observed production under infected circumstances or on a combination of QTL for disease resistance and phenotypic information on production potential under non-infected circumstances.

Mass Selection Under Constant Infection. Animals were ranked according to their phenotype for observed production under infected circumstances, after which truncation selection was applied. When animals were treated with medication, it was assumed that this treatment took place after observed production was recorded. When the number of animals with an observation for observed production is smaller than the number to be selected, additional animals were taken randomly from the group of animals with resistance below the lower threshold. Results of mass selection under constant infection pressure will be used as

point of reference, against which results obtained using other selection criteria or intermittent infection pressure (described below) will be evaluated.

QTL Selection. In the absence of infection, phenotypic information on observed production is not available. In this situation, therefore, selection was based on an estimated breeding value for observed production, which was estimated using QTL information on resistance and phenotypic information on production potential. It was assumed that a number of QTL for resistance had been identified, which explain a fraction p_m of the total additive genetic variance. The total QTL effect was assumed to be approximately normal distributed, with mean μ_{QTL} and a heritability of $h^2_{QTL} = 1.0$, because the QTL were supposed to be known without error. The proportion of variance explained by QTL remained constant over time, i.e., it was assumed that fixation of QTL due to selection was completely compensated for by identification of new QTL during the experiment.

In the strict sense, estimated breeding value of an individual is defined as "twice the phenotypic performance of its offspring measured as a deviation from the population mean." In the usual mixed linear models, this definition is equivalent to the additive genetic merit of the individual itself, measured as a deviation from the population mean. In the present nonlinear model, however, expected offspring performance does not only depend on genetic merit of the parent, but also on the variance of the offspring performance around its mean. For example, when the mid-parent value for resistance is just above the upper threshold, phenotypic resistance of the offspring may still be below the upper threshold due to Mendelian sampling and environmental variance in resistance, which reduces observed production of the offspring. Thus, estimated breeding value for observed production cannot simply be based on estimated genetic merit of parents, but the expected observed production of offspring need to be predicted. Because the breeding goal is to improve observed production under infection, offspring performance will be predicted for infected circumstances.

Observed production of an offspring was predicted from the mean of its parents and the variance of the offspring performance around this mean. Given the estimated breeding value for resistance of the parent, the phenotypic resistance of an offspring is normally distributed with mean $\frac{1}{2}ebv_p + \frac{1}{2}\overline{ebv_m}$ and variance $\sigma_{PR}^2 - \frac{1}{4}p_m\sigma_{AR}^2$;

$$R_o \sim N\left(\frac{1}{2}ebv_p + \frac{1}{2}\overline{ebv_m}, \sigma_{PR}^2 - \frac{1}{4}p_m\sigma_{AR}^2\right),$$

where ebv_p is the estimated breeding value

for resistance of the parent and \overline{ebv}_m is the mate average. The term $\sigma_{P_R}^2 - \frac{1}{4} p_m \sigma_{A_R}^2$ represents the phenotypic variance for resistance in the offspring given the parent, i.e. selection of the parent based on QTL for resistance explains an amount of $\frac{1}{4} p_m \sigma_{A_R}^2$ of the phenotypic variance of the offspring (note that the squared correlation between QTL-effect and breeding value for resistance equals p_m).

Depending on the mean and variance of R_o , a proportion p_L of the offspring will have values for R_o below the lower threshold, a proportion p_U will have values above the upper threshold, and a proportion p_B will have values between both thresholds. These proportions were determined as:

$$p_L = \Phi \left(\frac{L - \frac{1}{2} ebv_p - \frac{1}{2} \overline{ebv}_m}{\sqrt{\sigma_{P_R}^2 - \frac{1}{4} \rho^2 \sigma_{A_R}^2}} \right), p_U = 1 - \Phi \left(\frac{U - \frac{1}{2} ebv_p - \frac{1}{2} \overline{ebv}_m}{\sqrt{\sigma_{P_R}^2 - \frac{1}{4} \rho^2 \sigma_{A_R}^2}} \right)$$

and $p_B = 1 - p_U - p_L$, where Φ is the lower tail proportion of the standardised normal distribution. Potentially, one or two of the fractions p_L , p_B , or p_U can be zero. Given the distribution of resistance in the offspring, observed offspring production was predicted by weighting observed production in each part of the distribution of R_o by the appropriate proportion. Because $P_o = 0$ for $R < L$, and $P_o = P_p$ for $R > U$, expected offspring production equals:

$$SC_{P_o} = p_B \times P_{o_B} + p_U \times P_p, \tag{3}$$

where $P_{o_B} = [(\mu_{R_M} - L)/(U - L)] \times Pp_o$, which is the expected observed production for offspring with $L < R < U$, and $Pp_o = \frac{1}{2} Pp_{sire} + \frac{1}{2} Pp_{mate}$, which is the expected potential production of the offspring. Next, μ_{R_B} , the expected phenotypic resistance of offspring with $L < R < U$ was calculated from $\mu_{R_B} = (\mu_{R_o} - p_U \times \mu_{R_U} - p_L \times \mu_{R_L}) / p_B$, where $\mu_{R_U} = \mu_{R_o} + i_U \sigma_{P_{R_o}}$ is the expected phenotypic resistance of offspring with $R > U$ and $\mu_{R_L} = \mu_{R_o} - i_L \sigma_{P_{R_o}}$ for offspring with $R < L$, and i_L (i_U) is the "selection intensity" corresponding to p_L (p_U).

Animals were ranked and selected by truncation according to their SC_{P_o} . In the remainder of this paper, predicted observed production of the offspring (SC_{P_o}) will be referred to as "predicted production."

Environment: Infectious vs Noninfectious

The model of Van der Waaij et al. (2000) was developed assuming constant infection pressure, which in reality often will not be the case. It is possible that, for example, even though infection pressure often is present, during performance recording the infection pressure is absent and observed production actually has become equal to production potential. So during some generations selection may occur under infection pressure, while during other generations the infection pressure at the time of selection is absent. Mass selection on observed production in the absence of infection pressure will result in selection on production potential, neglecting resistance. Selection on predicted production puts selection pressure directly on production potential and indirectly on resistance via genetic marker information, independent of whether or not there is infection pressure.

In this paper three schemes with intermittent infection pressure were compared: one generation with infection pressure followed by one generation without (1-1), three generations with infection followed by one generation without (3-1) and three generations with infection pressure followed by three generations without (3-3).

Stochastic Implementation

A population was stochastically simulated (200 replicates), with both resistance and production potential modeled as heritable traits. Initial genotypes for production potential and resistance were randomly sampled from independent normal distributions with means μ_{Pp} and μ_R and additive genetic variances σ^2_{APp} and σ^2_{AR} . Environmental effects were randomly drawn from independent normal distributions with mean zero and variance σ^2_{EPp} and σ^2_{ER} . Heritabilities for production potential (h^2_{Pp}) and resistance (h^2_R) in the base population were: $h^2 = \sigma^2_A / \sigma^2_P$, where σ^2_A and σ^2_P represent the respective additive genetic and phenotypic variance ($\sigma^2_P = \sigma^2_A + \sigma^2_E$). Environmental variance remained constant across generations. Phenotypic values for observed production were determined for each individual using equation [1] and [2]. Genotypes for production potential and resistance in the offspring of selected parents were determined as the average of their parental genotypes plus a Mendelian sampling term.

The additive genetic variance explained by the QTL represents a fraction p_m of the total additive genetic variance for resistance, σ^2_{AR} , under infectious circumstances. The variance of the total QTL effect in the population (σ^2_{QTL}), therefore, represents the same fraction p_m of the total additive genetic variance. The QTL were assumed to have been accurately mapped, so there was no recombination between genetic marker and the QTL.

The QTL effects sampled from σ^2_{QTL} and genotypes for resistance were defined as $(1-p_m) \times$ the QTL effect.

Population and Parameters

The population consisted of 384 animals, of which 50% were female. Each generation one out of 16 sires were selected, as were one out of two dams. Each sire was mated to eight dams and each dam had four offspring: two male and two female who survived until reproduction. Selection was performed for 50 generations.

The heritability for production potential was assumed to be 0.3, with initial mean 100 and phenotypic variance of 225. The initial mean for resistance was 30, with a phenotypic variance of 36. The heritability for resistance was varied: 0.1, 0.3, or 0.5. The fraction of additive genetic variance explained by the markers was also varied: 10%, 25% or 50%. The thresholds were three phenotypic S.D. for R apart, and chosen such that in the first generation approximately 17.5% of the animals had an observed production greater than zero. In the first generation, 83.5 % of the population had a level of resistance below the lower threshold, so that extreme situations could be explored.

Results

Comparing Selection Strategies

Figure 1 shows population means for observed production, production potential and resistance for 50 generations of mass selection on observed production, as presented by Van der Waaij et al. (2000). When resistance is low (in the initial generations), selection for observed production initially results in an increase in resistance, rather than in production potential. With increasing level of resistance, selection on observed production results in an increase in both resistance and production potential. As soon as resistance has reached values above the upper threshold, selection on observed production becomes equal to selection on production potential.

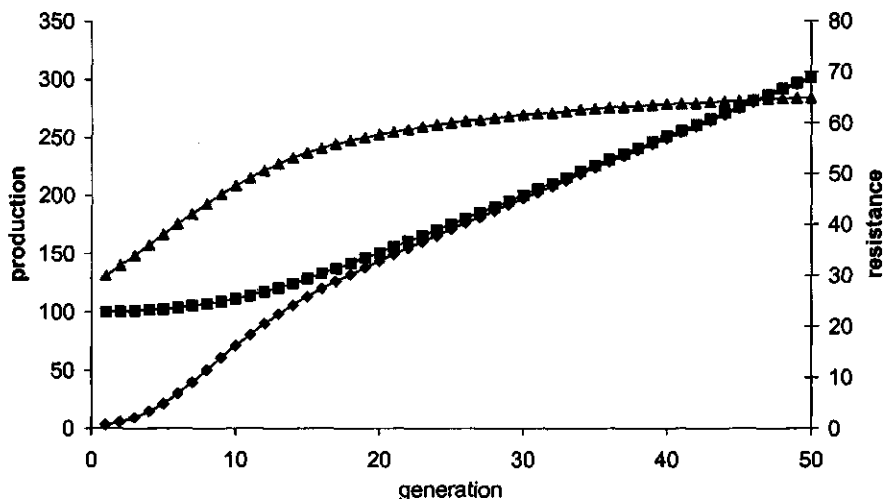


Figure 1. Population means for Pp (■), R (▲) and Po (◆) during 50 generations of mass selection on Po, under continuous infection pressure and for $h^2_R = 0.3$.

Figure 2 shows results for observed production (closed symbols) and resistance (open symbols) during 50 generations of selection on predicted production when the total QTL effect for resistance explained 10, 25, or 50% of the additive genetic variance. The results are expressed as deviations from results for mass selection on observed production under constant infection pressure, where the heritability for resistance was equal to 0.1 (2a), 0.3 (2b) or 0.5 (2c).

A general finding from this study is that marker assisted selection is most successful when the heritability for the trait is low, in agreement with previous studies of marker assisted selection (e.g., Smith and Simpson, 1986, Moreau et al., 1998). Figure 2a shows that for $h^2_R = 0.1$, when 10% of the additive genetic variance for resistance is explained by QTL, selection on predicted production results in almost the same genetic gain as mass selection. When 25 or 50% of the additive genetic variance explained by the QTL, considerably higher genetic gain for both resistance and observed production is achieved when selection is on predicted production compared to mass selection. However, Figure 2b shows that for observed production and $h^2_R = 0.3$, marker assisted selection only remains the method of choice when 50% of the additive genetic variance is explained by the QTL.

In terms of resistance, mass selection is initially superior to selection on predicted production, for $h^2_R = 0.3$ and 0.5. When resistance approaches the upper threshold, the difference between the two selection strategies decreases again, and in the case where 25 or

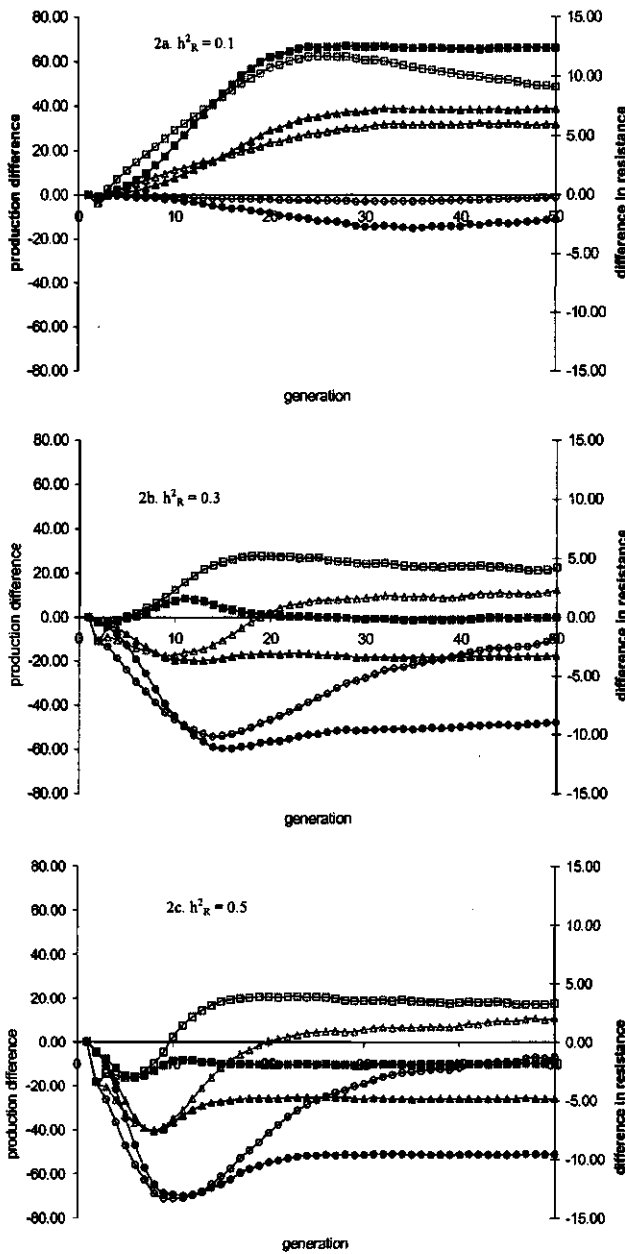


Figure 2. R (open symbols) and Po (closed symbols) expressing the difference between 50 generations of mass selection and selection on predicted Po in the offspring using genetic markers for resistance where 10% (●), 25% (▲), or 50% (■) of the additive genetic variance for resistance was explained by the QTL, for $h^2_R = 0.1$, $h^2_R = 0.3$, or $h^2_R = 0.5$.

50% of the additive genetic variance is explained by the QTL, selection on predicted production becomes the method of choice. The reason for this can be found in Van der Waaij et al. (2000). Selection pressure on resistance under mass selection is ceased as soon as resistance has passed the upper threshold. The result is that animals with values for resistance just above the threshold with high level of production will have the same chance of being selected as animals with the same production capacity, but with much higher level of resistance. Therefore, a proportion of the offspring of selected animals will have a level of resistance just below the threshold. However, during the generations in which most of the population has reached values of resistance above the upper threshold, this falling back into the less resistant category has no influence on the genetic gain anymore (these animals will

not be selected as parents). This situation does not occur with selection on predicted production because selection pressure on resistance remains as long as there is a possibility that a fraction of offspring will have a level of resistance below the upper threshold.

When the heritability for resistance increases further to 0.5, mass selection becomes the method of choice with regard to increase in observed production. The initial increase in the difference in resistance between mass selection and selection on predicted production becomes more distinct and only when 50% of the genetic variance is explained by the QTL does resistance reach comparable level to the responses obtained in resistance with mass selection.

Intermittent Infection Pressure

During generations without infection pressure, mass selection on observed production gives equivalent responses to mass selection for production potential. Thus, when mass selection is applied under intermittent infection pressure, selection pressure alternates between production and a combination of resistance and production, for as long as the selection candidate is not fully resistant. Selection for predicted production, however, is not affected by a change in infection pressure, provided the assumption that production level after vaccination/medication is equal to production level under noninfectious circumstances is valid. Figure 3 shows population means for resistance and observed production (as if there were infection pressure) during 50 generations of intermittent infection pressure, expressed as deviations from mass selection under constant infection pressure for $h^2_R = 0.1, 0.3, \text{ or } 0.5$. Mass selection under constant infection is compared to alternated infection schemes 1-1 and 3-1 and to selection for predicted production when the QTL explained 10, 25, or 50% of the additive genetic variance.

As a result of changing selection pressure due to intermittent infection pressure, differences in gain in resistance and observed production do not follow a smooth trajectory for the intermittent infection pressure scenario when mass selection is applied, especially when the heritability for resistance is high (Figure 3). Conversely, when selection is on predicted performance, the selection pressure is not influenced by the absence of infection pressure. For all three heritabilities for resistance considered here, intermittent infection pressure always decreases the genetic gain for resistance because selection pressure is ceased during generations without infection pressure. When infection pressure is absent during half of the generations, selection on predicted production becomes the method of choice for all heritabilities for resistance considered here, when 25 or 50% of the genetic

variance is explained by the QTL. In the long-term, and when three generations with infection is alternated with one without infection pressure, selection on predicted production becomes the method of choice for QTL that explain 25 or 50% of the genetic variation.

The general pattern for responses in observed production in the short term is comparable to that for resistance. Important differences arise in the long-term, where an increase in resistance no longer has a significant influence on observed production and mass selection, both under constant and intermittent infection pressure, becomes the method of choice. The reason for this is that in the case of mass selection all selection pressure is placed on production potential, whereas with selection for predicted production part of the selection pressure remains on resistance until all offspring have resistance above the upper threshold.

Discussion

General Features and Assumptions

In the present article, the use of QTL for disease resistance in selection for increased production under infection pressure is evaluated and compared to mass selection, for constant as well as intermittent infection pressure. Results show that selection on predicted production, making use of genetic markers that are linked to multiple QTL affecting resistance can, under some circumstances, be a good alternative to mass selection for increasing production and resistance simultaneously. The advantages of MAS are that it is no longer necessary to withhold animals from vaccination or treatment with medication after infection until measurements are taken for observed production, or even to keep them in an infected environment. We have demonstrated that, in general, the level of resistance is increased more by using QTL for resistance than it is under mass selection on performance. However, under the assumptions made in the model, genetic gains in observed production are only comparable, or superior to, mass selection when the heritability for resistance is low. If the selection pressure is intermittent, then mass selection to improve production under constant infection pressure becomes less efficient.

When comparing mass selection on observed production to index selection on the predicted production, it was assumed that treatment with medication would increase observed production to production potential in cases where the animal was not fully resistant. In reality, in most situations this assumption is too optimistic and the infection, for example a trypanosomosis infection, may still result in an increased requirement of energy

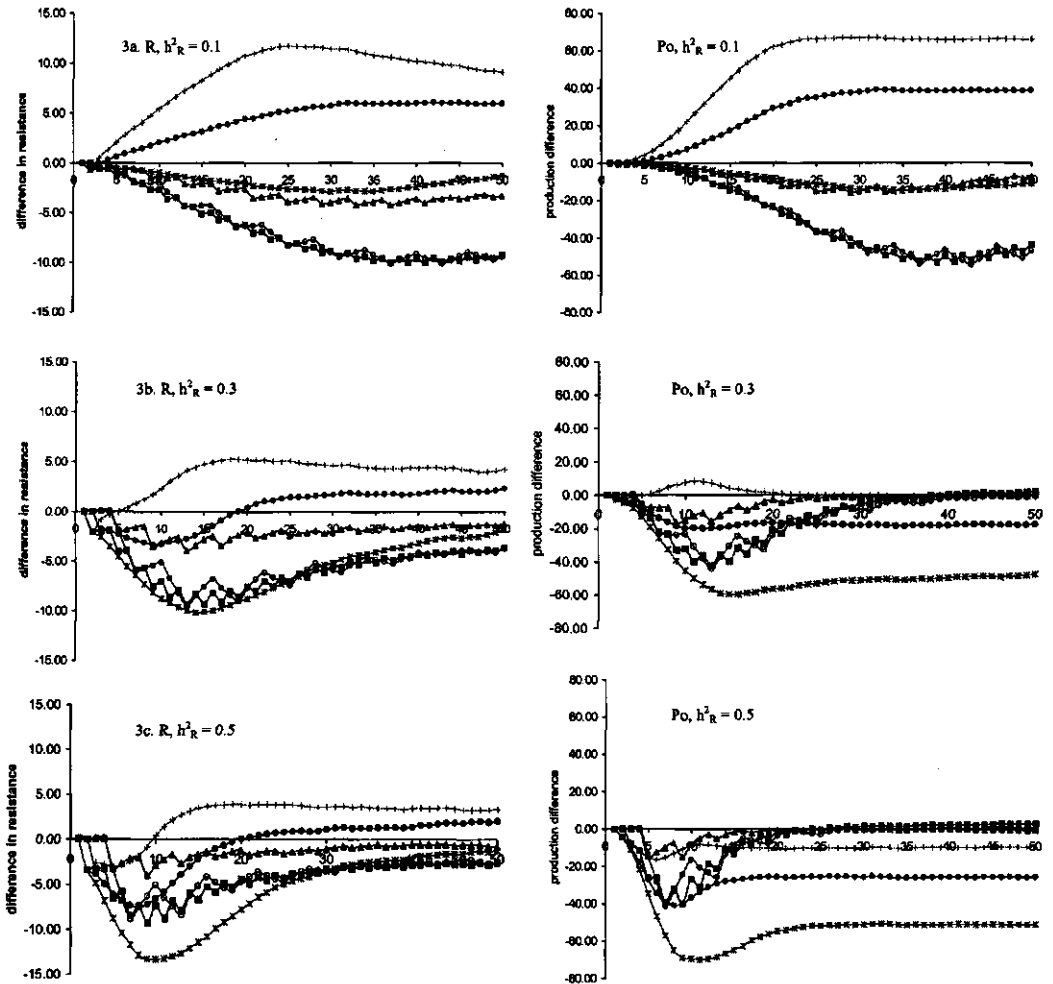


Figure 3. Resistance (left) and observed production (right) as deviations from mass selection on Po under continuous infection pressure, for alternative intermittent schemes 1-1 (■), 3-1 (▲) or 3-3 (○), or selection on predicted production when 10% (*), 25% (●), or 50% (+) of the additive genetic variance is explained by the QTL. Results shown for resistance heritabilities of 0.1, 0.3, or 0.5.

and a decreased feed intake (Van Dam, 1996), at the expense of production. Also, energy may be required to repair the damage done during periods when the animals were infected. So medication will restore production loss to some extent, but the resulting production level may still be lower than production potential. The consequence of the fact that the actual

production level may be lower than production potential is that, depending on the population level of resistance, indirectly more selection pressure will be put on resistance (part of the production still is influenced by lack of resistance, i.e., observed production is still not equivalent to production potential).

When the animals are kept in an uninfected environment, the model may be a closer approximation to reality. However, it should be realized that the model not only describes the effect of resistance to the infection under consideration, but actually a combination of effects of all environmental influences that are continuously present, including other infections, husbandry system, climatic aspects, and nutritional state. Therefore, depending on the recording of the phenotype, QTL for resistance may not only influence disease resistance, but also the adaptability of the animal to the environment the QTL are mapped in. An environment, in which the infection under consideration is absent, may also differ in some other constant environmental factors compared to the infected environment. Therefore, it is important that the QTL for resistance are mapped in the environment for which the animals are selected.

Unlike mass selection, with selection on predicted production it is necessary to know which values of resistance coincide with the lower and with the upper thresholds. This information is needed to determine the correct predicted fractions of offspring in each of the three regions divided by the thresholds. The distance between thresholds is, amongst other factors, determined by the infection pressure. Intermittent selection pressure may be thought of as representing a situation where threshold values are occasionally assigned wrongly (in the absence of infection pressure $f(R) = 1$). Results in this article show that the model is reasonably robust to incorrect threshold values, as the results between constant and intermittent selection pressure only differ slightly.

QTL Information

In order to successfully select on predicted production, multiple QTL are required. The accuracy will decrease with increasing distance between QTL and genetic markers (increasing occurrence of recombination events) and with decreasing frequency of the QTL in the population. It was assumed that fixation and detection of QTL balance each other out so that sufficient QTL remain available to give an approximate normal distribution. In published literature, essentially two types of assumptions can be found with regard to number of QTL under selection. Some assume there are only a limited number of QTL available and once those are fixed there will be no new ones detected (e.g., Dekkers and van

Arendonk, 1998). Others (e.g., Meuwissen and Goddard, 1996) assume that new QTL will be detected continually. This reasoning is based on the fact that in populations that have been under strong selection for many generations, QTL for production traits are still being detected (e.g., a QTL for milk yield in dairy cattle [Coppieters et al., 1998]). One important reason to adhere to either of the previous theories, apart from mutation rate of the QTL (Falconer, 1989), is the presence or absence of interaction between QTL. The QTL detection studies carried out so far assumed that the effect of each QTL stands alone (is additive and sometimes dominant), and possible epistatic effects with other QTL are ignored (e.g., Coppieters et al., 1998). Result of this assumption is that marker assisted selection is assumed to lead to fixation of the QTL, and thus to serious decrease in variance explained by the QTL. However, when assuming a very simple interaction between two QTL, as suggested by Nijhout and Paulsen (1997), a small change in genetic background (e.g., by fixation of QTL through selection) may have large effect on the expression of (new) QTL. They show that not only is it likely that new QTL will be detected at regular basis (with a limitation to the number of genes involved in the trait expression of course), but also that it is important to continue QTL detection because new QTL occur due to change in expression during the selection process.

Which ever theory is correct, in the present study QTL that are detected first are most influential in the selection process, because selection pressure on the QTL effect (resistance) decreases with increasing level of resistance. When resistance is approaching the upper threshold, an increase in resistance is of less importance. This effect is strengthened by the fact that it is likely that the QTL with large effect will be detected first. Only a few QTL are expected to have large effects, with many having small effects (Bost et al., 1999). Thus, the QTL with smaller effects that will be detected later will have a smaller influence not only because of smaller size, but also due to the fact that population level of resistance is already at such a level that more emphasis is put on production. Therefore, if a change in variance due to fixation of a QTL with a large effect may occur, it is likely to be of less influence when the threshold model is applied than when a linear model is used.

The size of the total QTL effects considered in this study was varied from 10 to 50% of the additive genetic variance, a wide range of effects. Phenotypic values for resistance as defined in this study often cannot be measured accurately, but results for several indicator traits describing resistance have been reported. Dumas et al. (2000) reported an effect explaining 46% of the phenotypic variance for stress response in rats. Stear et al. (1996) found that allelic variation at the MHC-DRB1 locus accounted for 33% of the additive

genetic variance for faecal egg count in sheep. Zhang et al. (1998) found several QTL for somatic cell count in dairy cows, the largest explaining more than 27% of the additive genetic variance. The sum of QTL effects they detected for somatic cell count explained approximately 50% of the additive genetic variance, although they indicate that their estimates might be overestimated. Thus, our assumed QTL effects are within the range of values reported in the literature.

Concluding Remarks

This article has presented methodologies for utilizing disease resistance QTL, under certain assumptions with regard to infection pressure. It is important to realize that while such QTL are often important and beneficial, there are also circumstances where their use can not be justified either in terms of potential genetic progress or cost-effectiveness. As described by Van der Waaij et al. (2000), this model and general methodology has applications which are wider than disease resistance *per se*. Other important applications include environmental or metabolic stresses. A further application of this model may be to the challenge faced by pig and chicken breeding companies which select animals under high health status conditions for production in more challenging commercial environments.

Implications

Simultaneous genetic gain is considered for both production and resistance to a disease characterized by constant infection pressure, where genetic markers (i.e. QTL) for resistance aid the selection of animals with enhanced resistance. Selection responses are quantified by stochastic simulation. The model assumes some simplifications compared to reality. Results suggest that only when the heritability for resistance is low does selection on predicted production using resistance QTL result in greater gain for resistance and observed production than mass selection on observed production under constant infection. Intermittent infection pressure reduces the effectiveness of mass selection, increasing the relative value of QTL for resistance. Selection on predicted production requires the availability of multiple, accurately mapped QTL, which should soon be technically possible. Mapping QTL requires phenotypic observations, so (some) animals will need to be infected to map resistance QTL.

Acknowledgement

The authors would like to thank WOTRO (Netherlands Foundation for the Advancement of Tropical Research) for financial support. SCB was funded by the Biotechnology and Biological Sciences Research Council.

References

- Bost, B., C. Dillmann, and D. de Vienne. 1999. Fluxes and metabolic pools as model traits for quantitative genetics. I. The L-shaped distribution of gene effects. *Genetics* 153:2001-2012.
- Coppieters, W., J. Riquet, J.-J. Arranz, P. Berzi, N. Cambisano, B. Grisart, L. Karim, F. Marq, L. Moreau, C. Nezer, P. Simon, P. Vammanshoven, D. Wagenaar, and M. Georges. 1998. A QTL with major effect on milk yield and composition maps to bovine Chromosome 14. *Mammalian Genome* 9:540-544.
- Dekkers, J. C. M., and J. A. M. Van Arendonk. 1998. Optimizing selection for quantitative traits with information on an identified locus in outbred populations. *Genet. Res. Camb.* 71:257-275.
- Dumas, P., Z. Pausová, V. Kren, D. Krenová, M. Pravenec, M. Dumont, D. Ely, M. Turner, Y. Sun, J. Tremblay, and P. Hannot. 2000. Contribution of autosomal loci and the Y chromosome to the stress response in rats. *Hypertension* 35:568-573.
- Falconer, D. S. 1989. Introduction to quantitative genetics. 3rd ed. Longman Scientific & Technical, Harlow, UK.
- Meuwissen, T. H. E., and M. E. Goddard. 1996. The use of marker haplotypes in animal breeding schemes. *Genet. Sel. Evol.* 28:161-176.
- Moreau, L., A. Charcosset, F. Hospital, and A. Gallais. 1998. Marker-assisted selection efficiency in populations of finite size. *Genetics* 148:1353-1365.
- Nijhout, H. F., and S. M. Paulsen. 1997. Developmental models and polygenic characters. *The American Naturalist* 149:394-405.
- Smith, C. and S.P. Simpson. 1986. The use of polymorphisms in livestock improvement. *J. Anim. Breeding Genetics* 103:205-217.

- Stear, M. J., M. Park, and S. C. Bishop. 1996. The key components of resistance to *Ostertagia circumcincta* in lambs. *Parasitology Today* 12:438-441.
- Van Dam, J. T. P. 1996. The interaction between nutrition and metabolism in West African dwarf goats, infected with trypanosomes. PhD thesis, Wageningen University, The Netherlands.
- Van der Waaij, E. H., P. Bijma, S. C. Bishop, and J. A. M. van Arendonk. 2000. Modelling selection for production traits under constant infection pressure. *J. Anim. Sci.* 78:2809-2820.
- Zhang, Q., D. Boichard, I. Hoeschele, C. Ernst, A. Eggen, B. Murkve, M. Pfister-Genskow, L. A. Witte, F. E. Grignola, P. Uimari, G. Thaller, and M. D. Bishop. 1998. Mapping quantitative trait loci for milk production and health of dairy cattle in a large outbred population. *Genetics* 149:1959-1973.

VII General Discussion

Introduction

Trypanosomosis is a disease resulting from infection with trypanosomes, parasites that usually live intra-vascular and extra-cellular (e.g. in human, cattle, goat, sheep, horse), and is most often transmitted by the tsetse fly. Consequences of infection with trypanosomes in cattle often are production loss, abortion, and eventually, unless treated with medication, death (Murray et al., 1991). Vaccination is not effective due to the ability of the parasite to 'adapt' to the current immune response and repeatedly changing of antigen within the host (Murray et al., 1984). Some cattle breeds, like the N'Dama and some West African shorthorn breeds, remain productive without medical treatment while infected with trypanosomes: they are trypanotolerant. However, even though these animals can cope with infection, they still are affected to some extent and would perform better without infection pressure (ILRAD, 1989). Generally these breeds are not considered to be productive because of their small size (referred to as 'dwarf breeds', e.g. Roberts and Gray, 1973). However, they are only kept in infected environment and, consequently, their production potential outside these areas is not well recorded. As a result of their natural habitat these breeds have not been recognised as potential breeds for production. It was a report by Stewart in 1951 that revealed the benefits of these breeds under harsh conditions, which raised the interest in these small animals. Under light to moderate infection pressure, field studies have shown that the N'Dama on average has a birth weight of 17.9 kg, a growth rate of 0.15 kg/day, an average weaning weight of 85.4 kg, age at first calving at about 49.5 months, a calving interval of about 625 days, a lactation length of about 463 days, and a lactation off take of 492 kg (Agyemang et al., 1997).

This thesis is part of a much larger research, aiming at improving the level of resistance to a trypanosome infection and the production potential of animals in tsetse infested areas. The project aimed at detecting Quantitative Trait Loci (QTL) that determine trypanotolerance, which could subsequently be used as a selection tool for producing higher productive animals in tsetse-infested parts of the continent. This general discussion will first discuss the possibility that there still is genetic variation for trypanotolerance in trypanotolerant breeds such as the N'Dama. It will then focus on how to utilise this genetic

diversity in a breeding program. To set up a breeding program, it is essential to describe the production system so that it becomes clear what circumstances the breeding program should be designed for. Also it is essential to define a breeding goal and accompanying selection criteria. Finally breeding goal, selection criteria and production system will be combined into suggestions on how to set up a breeding program under the various circumstances.

Historical background of trypanotolerance in the N'Dama

There are several theories about how and when the Longhorn cattle were introduced into West Africa. One theory is that the *Bos taurus* Hamitic Longhorn cattle breed is ancestor to for example the N'Dama, and arrived in the Nile Delta from the Near East around 5000 BC. The taurine Shorthorn cattle were only introduced into the same area between 2750 and 2500 BC (reviewed by Murray et al., 1990). Other people suggest that there have been two centres of domestication of cattle: one is Western Asia, but also one in North Africa, where taurine short horn cattle were domesticated some two thousand years later (Payne and Hodges, 1997). Loftus et al., (1999) and Bradley et al., (1996) believe that the domestication process of the *Bos taurus* cattle occurred in the Near East, independently from the domestication of the *Bos indicus* further to the west. From their origins of domestication the cattle spread across Africa, the *Bos taurus* cattle first (Hanotte et al., 2000). There are several theories on the (number of) route(s). During the time of domestication, North Africa and the Sahara were savannah type of environments The cattle that were portrayed in rock paintings in the Sahara suggested they were predominantly of the longhorn type, but there were also paintings of cattle with other size horns. This would suggest that Saharan cattle around 5500 BP were partly crossbred, presumably a product of the Hamitic Longhorn with the local short-horn type of cattle (Payne and Hodges, 1997). The rock paintings from that time also suggest there was a widespread pastoral industry and that cattle were milked. Possibly because of increased drought conditions in the central Sahara, people and cattle travelled westwards and southwards, towards the forests of West Africa. It is likely that N'Dama type of cattle were introduced into West Africa at an earlier period than shorthorn cattle, and by a different route. Today, latitudes 12-14°N separate the regions in which the more susceptible zebu cattle can be utilised from those grazed by taurine and sanga cattle, where sanga is a historical cross between *Bos taurus* and *Bos indicus* breeds (Payne and Hodges, 1997).

Since the introduction into West Africa there has been considerable natural selection for resistance to trypanosomosis and other types of infections (other animals most likely

showed at least a poorer reproductive performance, and probably died at younger age). Increase in resistance may have occurred through fixation of alleles that were already present in the population, or by mutation on base pair (base substitution like transition or transversion), or chromosomal level (e.g. inversion or translocation) (Hartl and Clark, 1988). One would need a separate study to be able to distinguish between these two types of mutation as a cause of increased trypanotolerance, but it is possible to speculate a bit on the probability that there still is genetic variation for the genes involved in increased trypanotolerance in the N'Dama.

Based on the fact that the taurine cattle were introduced into the Nile Delta by human settlers around 5000 BC, it is likely that population size was limited. Because the cattle were travelling with humans, and because sleeping sickness also is a great problem in human, it is likely that in that time, the cattle were not exposed to high infection pressure. Selection pressure, therefore, was not likely to have been extremely high, but high enough not to have all cattle surviving, and also high enough so that only part of the cattle remained reproductive. Favourable mutations or fixation of favourable alleles with regards to trypanotolerance would have created a selective advantage for animals possessing this genomic build up compared to other animals. Even though there has been a continuous selection pressure for trypanotolerance since the introduction of cattle in West Africa, it is likely that there is genetic variation for traits such as trypanotolerance within and between the West African cattle breeds, such as the N'Dama. The situation is somewhat comparable to the *Bos taurus* breeds in Europe, where substantial genetic differences for milk production traits were detected within and between different Friesian strains (British Friesian as compared to Dutch Friesian regions in The Netherlands) (e.g. Politiek et al., 1981). Another reason why it is likely that there is genetic diversity with regards to trypanotolerance within breed is that the breed is not resistant to trypanosome infection (yet), even though it is living in trypanosome infested environment. This indicates that the process of developing such resistance is still taking place. Mutations that occur in one herd will only slowly find their way to other herds, and possibly may not even reach some herds.

Results in Chapter 3 indicate that there are several genes involved in expression of trypanotolerance related traits originating from the N'Dama. There is also evidence, at least for the four N'Dama bulls that were grandparents to the F2 population, that the animals were not homozygous for all of these genes. This proves that these genes that have a positive influence on trypanotolerance and are not fixed in the population, which may indicate that the mutation only occurred relatively recently. It is also possible that the

favourable allele has a dominant effect so that fixation is an extremely slow process and mainly dependent on the presence of genetic drift. It would be interesting to find out whether there is difference in background of trypanotolerance between the different N'Dama populations. If so, then "crossbreeding" within breed, in the sense of utilising N'Dama cattle from different origins, could further improve the trypanotolerance level, and probably also the production level in the N'Dama. This is comparable to what occurred in Europe with regards to utilising the different strains of Friesians. A proportion of those breeds were exported to the USA, where such "crossbreeding" was performed and the (more productive) result (the Holstein Friesian) was subsequently imported again by the European countries to upgrade the local Friesian breeds.

Local circumstances

Farmer's choice. An African cattle farmer would include different selection criteria in his wish list as compared to his Western colleague. In both cases production would be an important criterion, since that generates income. However, apart from milk and meat, in some areas in Africa animals are used for traction power to work on the land, so that size and strength of the animal may be a breeding objective. Also often manure may be of economic importance, for example in Ethiopia up to one third of the annual income of some farmers results from the sale of dung for fuel (Wilson, 1995). The health of the animal will be very important, both for the African and the Western farmer, because healthy livestock saves medication costs and sick animals will have a reduced production, and in the worst case, may even die. Unlike for the Western farmer, who can usually adjust the environment to the animal, for African farmers in some areas it is important that the animals are able to cope with the more harsh local circumstances. For example in areas with two distinct seasons, during one of which there may be severe shortage of feed, and possibly even water, animals that cannot cope with such environment will stop producing, and may even die. Reproductive capacity of cows is economically of importance for replacement, both for the African and the Western farmer, but for the African farmer livestock often serves as an alternative savings account. If money is needed, for example for school fees, for an urgent medical treatment of a relative, or for buying seed for next years crops, an animal can be sold, and newborn animals can be considered interest. Most often goats and sheep are sold when money is needed, but if larger sums are required cattle may be sold as well. Animals can be considered a type of insurance. In some situations, therefore, a high number of animals may be more important than fewer, but more productive animals. In addition,

reproduction for the African farmer often also is of social importance because a higher number of animals results in a higher social status.

Production system. Which traits are most important for a farmer depends on the production system and on the geographical region. The production system is often tailored to the region. The climate, for example, will be a factor that differs across regions. In a dry region, crop farming will be more complicated, or even impossible, as compared to a more temperate climate. In a dry region feed supply for livestock will not be of constant and high quality, or even constantly available. Under these circumstances farmers will travel around with their animals in search of feed. In these areas it is more important to have animals that are able to survive in a harsh environment, than it is to have animals that have a high production potential but cannot bring that potential to expression due to insufficient nutrition and/or water supply. Farmers in the more temperate climate regions will, apart from livestock, often also farm crops. In such, usually small, types of farms, the cattle will often be part of a larger system including some crops, perhaps some sheep or goats, some chicken, or even fish. In general the ecological conditions like type of soil, climate, altitude, and vegetation are an important factor in determining what type of farming system can be applied. A simple classification of ecological conditions is into humid, sub-humid, semi-arid, arid, very arid, and highland areas (Udo, 1996, electronic publication on www.zod.wau.nl/dps/publikat.html). In the highland areas, like for example around Nairobi, Kenya, livestock originating from elsewhere, like for example Friesian cattle, can relatively easy adapt, whereas in very arid, arid, semi-arid, and humid ecological zones, improvement of local stock may be the only option.

Depending on the ecological zone, there is also difference in availability of feed resources for the livestock. In most ecological zones, cattle is usually fed on grazing lower quality land, crop residues, grass collected from road-sides and forages. Often the quality of these feeds is not sufficient to support high milk production or meat production levels. Grazing areas are decreasing in size in favour of crop farming, which makes it increasingly difficult, especially for the pastoralist farmers, to keep livestock in such areas. Disease pressure is dependent on the ecological zone. Not only climate, but also the density of the human population influences the severity of infection pressure. In the case of trypanosomosis, for example, the infection pressure is lower close to urban areas because of the lower tsetse fly density in these areas. The location of a farm in relation to such urban areas will influence not only the infection pressure, but also the infrastructure. The

infrastructure determines largely the market the production is aimed at, and the availability of supplies such as medication. A high milk production is desirable when farming close to a city, where there is a large and still rapidly increasing demand for fresh milk (Dempfle and Jaitner, 1999). However, further away from the city not only the demand will decrease because most people will have their own animals, but also because the transportation of the milk to the city is not guaranteed due to for example changes in road quality across seasons. Also, a good road network would enable the transportation of goods, and AI. Infrastructure also relates to a good communication system so that training opportunities become available, to the presence of farmers organisations and co-operations so that breeding policies and feed supply can be co-ordinated, and also to government policies, so that improved farming is encouraged and possibly structured across the country.

Breeding goal

All factors mentioned above indicate that more than one breeding goal is required for the whole of Africa. Farmers in different regions will farm under different circumstances, which result in different traits they would like to improve. Farmers that are more closely located to an urban centre most likely are more interested in increased production because there is a secure opportunity to market their products. Apart from the market opportunities, disease pressure is often lower, so there will be a decreased risk of losing animals than with high disease pressure, and additional feed requirements of the cattle often can be met by buying in feed using the money obtained from selling the animal products. In the highlands around Nairobi, Kenya, the farms are very small, often only one to three cows, usually black and white ('Friesian') or red and white ('Ayrshire'), and the farmers are often use AI, or buy pregnant cows. With a large market as Nairobi in the neighbourhood, milk prices are sufficiently high to buy supplementary feed to increase milk production. Other breeding goals for these farmers will be an increased health status of the animals (though of less need than in other areas with higher infection pressure), and an increased reproduction rate.

In more remote areas, the ecological zone usually is different and such high productive cattle would not survive. Some farming systems in the humid and sub-humid climate regions often consist of a combination of livestock and crop farming. Such farms should be considered as an integrated system, since improving the production level of the cattle may have a strong influence on the rest of the system. The improved cattle will have an increased nutritional demand in order to bring the increased production potential to expression. This implies that somewhere else in the system sacrifices should be made. This

will involve an increased pressure on the land (Wilson, 1995), but also appointing more labour to collecting the feed, an increased time spend on milking and feeding the higher productive cows, and possibly increased labour in processing the milk, if done on the farm. This labour is not automatically available and should be generated by neglecting other tasks. The increase in income, therefore, should counterbalance the sacrifices that need to be made in order to generate the extra income. Health status and reproduction are also of importance in this type of farming system.

In many farming systems, but especially in the pastoralist systems, a higher number of cattle is required to serve the social purposes as mentioned previously. The pastoralists also often live in region where there are distinct periods of shortage of feed and possibly even water. Farming higher productive animals with increased nutritional demands would mean that the total number of animals should be lower in order to profit from their production potential. However, because of their way of living, pastoralists often do not live in the vicinity of an urban centre, which means that possibility to market their products is not very large. A decrease in number of animals would decrease the social status of the farmer, decrease the financing function, and the insurance function. It, therefore, is likely that these farmers are not interested in higher productive animals, but more in higher reproduction rate and increased health status. Also, even though it has no genetic background itself, it is also possible to select for number of animals in the herd. This can be considered a production trait and the economic value to it can be determined from a combination of market price and socio-economical value. Kosgey et al. (2001) have looked into determining a socio-economic value in sheep in a Kenyan pastoralist system, and conclude that it is necessary to include the non-productive roles of sheep in a breeding program.

Selection criteria

In this thesis special attention has been paid to definition of some animal health and production traits, under trypanosome infection, i.e. how to measure trypanotolerance. Trypanotolerant animals are better capable of controlling anaemia (e.g. Trail et al., 1993; Chapter 2 of this thesis), and to continue growing following infection (e.g. Murray et al., 1984; Chapter 2 of this thesis). For selection on a combination of both resistance to a trypanosome infection and production, growth following infection is a good and easy to measure indicator of trypanotolerance (Chapter 2). Even though PCV can be measured accurately, the precise use of PCV as a trypanotolerance criterion in the field to a large

extent depends on the accuracy of measurement of trypanosome infection status (d'Ieteren et al., 1998), and also largely on the condition of the animals (ILRAD, 1989).

Growth (or an other (re)production trait) following infection under field conditions can be considered a trait that combines all factors that is influencing growth rate. This is not restricted to trypanotolerance, but it also combines the presence of different types of infection that are continuously present, continuous limitations to nutrition and even availability of water (Chapter 5). Selection for production under field conditions, therefore, will not just improve resistance to one disease, running the risk of losing resistance to another important type of infection, as may occur when selecting under controlled environment, but will result in animals that are adjusted to a combination of infections or other permanent environmental constrains that are present in the environment in which the animals are maintained. Animals are often selected for breeding at a relatively young age, so that observations for growth can still be obtained. Body weight of an animal is easy and relatively accurately to measure, even without a scale, which makes it very suitable for field conditions. If no scale is available, an option is to use body measurements to estimate the body weight of the animal. The genetic correlation between heart girth and body weight in lactating Holstein heifers, for example, is 0.77 (phenotypic correlation is 0.74) (Koenen and Groen, 1998). It is conceivable that this relation will be different from breed to breed, and also for younger animals, since the Holstein is a relatively early maturing breed. Such a relation could be quite easily be determined in some field research in the N'Dama. Growth is not only influenced by permanent environmental factors, but also by factors such as illness due to non-permanent infection like a virus infection, temporary absence of sufficient amount of feed, etc., which will be a disadvantage when using growth as a selection criterion.

With regards to trypanotolerance as specified breeding goal, it is not clear whether apart from growth under field conditions, also some measurement of parasitaemia should be included as a possible selection criterion. The low correlation between the results using the detection method as currently mostly used (the dark ground phase buffy coat method (Murray et al., 1977)) and anaemia and body weight related traits is not reason enough to include it in a breeding goal. It is clear that animals that are not infected with trypanosomes will also not show any degree of anaemia caused by the infection, so parasitaemia does have some effect on animal performance. However, this effect may already be sufficiently included in production traits like, for example, growth following infection. Also, measuring parasitaemia would require taking blood samples, which in turn requires trained personnel.

It will depend on the infrastructure whether this is available. It is, therefore, more likely that such a selection criterion will be implemented in a breeding scheme in urban areas, as compared to the more remote areas. If an accurate method to measure the degree of parasitaemia may become available, which is also cheap and easy to use, and shows a high correlation with the animal's performance, and provided the infrastructure allows for it, then including such measurement as a selection criterion may be an efficient way to increase trypanotolerance.

The exact way of how such an additional trait describing trypanotolerance, rather than production affected by trypanosome infection, should be included as selection criterion is not straightforward (Chapter 5). The relationship between observed production and parasitaemia would need to be incorporated when determining the ranking of the animals for breeding, which is very complicated. Additional complicating factor is that the model in Chapter 5 assumed that lack of resistance to one disease is the only production limiting. It would be very difficult to assign how much influence lack of trypanotolerance exactly has on the decrease in production under field conditions. Not only are not all animals equally infected, but also other constant environmental factors, especially other continuous infections, are not likely to have equal influence on the production of all animals. Ignoring the fact that lack of resistance to trypanosomosis influences the observed production when ranking the animals would result in sub-optimal selection. The decreased production level due to lack of resistance will be considered to equal the production potential, and will be assumed to decrease further, depending on the level of resistance. In a way the lack of resistance is counted double that way.

Other selection criteria may involve reproduction related traits such as age at first calving, calving interval, number of pregnancies per cow, number of calves per cow, number of pregnancies per bull, etc.

Selection strategies

There are large number of selection strategies (e.g. F1 production, rotational cross breeding schemes, etc), but we will concentrate on two selection strategies: a. within breed selection or b. creating a hybrid from crossing a trypanotolerant breed (e.g. the N'Dama) with a higher productive breed. Selection will be aimed at improving trypanotolerance and production level.

- a. *Within breed selection* would exploit the variation within a breed for making genetic progress. Selection within for example the N'Dama can be practised

under harsh environment, which would reduce the risk of losing trypanotolerance, and would make use of the other qualities of the breed, like resistance to other diseases, heat tolerance and the capability to survive with restricted water supply. Because of the small size of the N'Dama, animal production traits like growth and traction power will not improve rapidly to the same absolute level as for example that of the Kenyan Boran in tsetse free environment. However, for the more extreme ecological regions with a low input system, the unique features of the indigenous breeds are absolutely necessary because of the periodically extreme local environment with regards to scarcity of feed and presence of infectious diseases like trypanosomiasis and some tick-borne diseases (Dempfle and Jaitner, 1999). It is possible to obtain substantial genetic gain by phenotypic selection within breed in small populations (Chapter 5). If pedigree information would be available, in combination with a computer, phenotypic selection could be extended to BLUP selection. In addition, if there would be a molecular genetic lab, and QTL are available, an option would be to apply marker assisted selection utilising QTL explaining trypanotolerance in combination with phenotypic observations on production outside the tsetse infested environment (Chapter 6).

- b. *Creation of a hybrid* by crossing two breeds may create additional genetic variance and offer opportunities to improve the potential production level more rapidly. That way, positive feature of one breed (e.g. trypanotolerance in the N'Dama) can be combined with positive features of another breed (e.g. production level in the Boran). The resulting F1 animals would be inter-mated, to fix genes responsible for both trypanotolerance and production (i.e. creating an synthetic). It is possible that during the F1 generation, animals that are heterozygous for trypanotolerance genes are less resistant to infection with trypanosomes than pure bred N'Dama. Reproduction of these animals under infection pressure may become problematic due to decreased fertility. Treatment with medication may help to some extent, depending on the infection pressure. Creation of a hybrid as described above can be achieved without the help of molecular genetic marker (Chapter 2). Phenotypic selection in an F2 generation under infection pressure may already result in a substantial increase in production, while maintaining similar trypanotolerance level as

compared to the pure bred N'Dama. If a molecular lab is available and QTL information can be used in selecting those animals that carry favourable alleles for the QTL considered (Marker Assisted Introgression, MAI), infection of the animals is no longer required for selection on observed production and selection may occur outside the tsetse infested areas. Disadvantage of selection outside these areas is that other permanent environmental influences that influence observed production in that case might differ from the environment the animals are selected for (Chapter 6). MAI also introduces a risk of losing alleles originating from the positive mutations (positive alleles) that are not identified as QTL, as also discussed in Chapter 4.

The choice between these two methods of selection is crucial and will be discussed in greater detail:

a. Within breed selection: Advantage of selection within a trypanotolerant, as described in Chapter 5 and 6, is that animals at onset of the selection scheme are already trypanotolerant. This is an important feature for breeding schemes in tsetse infested areas. Selection in that case can be on observed production (i.e. production under infection pressure), which will not only result in increased production, but also in increased resistance to trypanosomiasis and to other permanent environmental factors. Selection pressure on these factors automatically remains until all parents (phenotypic selection, Chapter 5), or even all offspring of those parents (Marker Assisted Selection (MAS), Chapter 6), are fully resistant to the infection (e.g. trypanosomes infection) and other permanent environmental factors. This approach may reduce the risk of decreasing resistance to a disease that is not under direct selection pressure, while resistance to trypanosomiasis increases.

Within breed selection can be performed at different levels: from pure phenotypic selection without any knowledge of genetic relations between animals, to MAS, where apart from the genetic relation, also a prediction of the genetic make up of an animal for certain traits is incorporated. A positive feature of phenotypic selection on production is that it is relatively easy to perform under local conditions. Milk production, body weight, traction, and reproduction are all traits that can be measured without the need of sophisticated technical features, or the assistance of specially educated and trained lab personnel. However, a possibly negative feature of phenotypic selection on observed production, is that in order to have a proper phenotype for 'observed production', the animals may not be treated with medication until the observations have been recorded. In many cases this will

not be a problem because animals will only be treated with medication when they are really sick and run the risk of dying. Because medication is expensive, animals that are influenced by the infection but remain productive, though at lower level, will not be treated and those are the animals that should be selected from, rather than from the treated animals. Phenotypic selection within one breeding scheme should be performed under the same circumstances, so that for example management differences between herds do not influence the selection decisions. If selection would be across herds, BLUP could be applied to account for these differences, but this would require pedigree registration, and also computer facilities. If computer facilities and pedigree registration are available, BLUP would also increase accuracy compared to mass selection.

When a molecular lab would be available, and QTL are determined for resistance, animals could be selected outside the tsetse-infested area. Observed production could be predicted from the production potential (assumed equal to observed production after medical treatment, or in an environment without infection pressure) and marker information regarding the QTL for resistance. Selection could then be on the 'observed production' that would be predicted for the offspring of selection candidates (Chapter 6). Advantage of this type of marker assisted selection is that all animals will be assigned the same degree of infection when determining the predicted observed production. In that way there will be no difference in severity of infection between animals, and also, if sufficient number of animals are trypanotolerant, the 'assigned intensity of infection' can be adjusted so that selection is as if the infection is more severe (Chapter 6). Disadvantage is that selection is only on the QTL and not on other genes that may be involved in trypanotolerance. This illustrates the need for continuous QTL detection simultaneously to the MAS scheme. Also, selection will not include resistance to other environmental factors that are present in the area the animals are bred for, but absent in the selection environment, like for example infection pressure of another type of disease.

There is a potential problem when deciding to apply a MAS scheme outside the area the animals are selected for. This involves the risk that the QTL is not segregating (in the case the QTL was detected in another population and not validated in the population under selection), or the phase between genetic marker and QTL changed due to recombination. An obvious solution to the absence of the QTL at onset of selection, and requirement for successful MAS, is the verification of the presence of the QTL and the frequency of the favourable allele, which could be done simultaneously to running the MAS scheme. Requirement for verification of the QTL is the availability of phenotypic information on the

animals, and thus need to be done under infection pressure. Additional checks on the presence of the QTL during generations of selection will add to the probability that indeed the favourable alleles are inherited to the next generation. Advised frequency of such checks will depend on the distance between genetic marker and QTL, and on the required efficiency of the MAS. It would be enough to have part of the population infected every so many generations to check whether the QTL is still present. MAS under infection pressure would help to overcome the differences in the degree of infection between herds, but also between animals.

b. Creation of a hybrid: Within breed selection for improved production within trypanotolerant breeds may be considered a too slow process due to lack of sufficient genetic variation. At the moment, even though it is considered the best performing breed in a tsetse infested environment, in the absolute sense the N'Dama is small, produces little milk, and, because of their size, they are not very suitable to use for traction. A possibly more efficient way to improve production in tsetse infested areas is considered to be introgression of the QTL responsible for trypanotolerance into higher productive cattle (Soller and Beckman, 1987). Like in the case of within breed selection, the QTL should be accurately mapped and genetic markers should be tightly linked to the QTL. However, when looking at results from Chapters 3 and 4, introgression as generally defined in animal breeding literature, i.e. the introgression of certain genes, while eliminating all other background genome (e.g. Hospital et al., 1997), does not seem the best option to combine features of two breeds. If all QTL that are related to trypanotolerance, and which are not all reported in Chapter 3 would need to be introgressed in order to create trypanotolerant animals, an enormous number of animals would be required (Chapter 4). Also other breeding schemes, as suggested by Koudandé et al., (2000), would require a large number of animals. Given the number of QTL detected, introgressing them all at the same time is, therefore, not a realistic option. Marker assisted selection using BLUP could be applied for a more gradually introduction of the QTL into the recipient breed. This, however, would require a larger number of generations to the 'final' product, even though the difference in trypanotolerance with the donor breed will decrease across generations. By infecting the animals during some back cross generations, it is possible to get some idea whether there is interaction between QTL and genes that are not detected, though not a total picture because the recessive QTL will not come to expression until the intercross generation. Introgression may in that case lead to poorer results than would be expected from the estimated QTL

effects, though results from the introgression experiment of trypanotolerance QTL in mice (Koudandé, 2000b) suggest this not a great danger.

Results in Chapter 2 suggest that by creating an F2 population (so introgression with immediate fixation), some of the F2 animals combine trypanotolerance (perform at N'Dama level) and production (considerably higher than the N'Dama). The N'Dama used for comparison in Chapter 2 may not be representative for the breed average. It needs to be further investigated what proportion of an F2 population is expected to be significantly more productive than the higher productive pure bred N'Dama that are available in West Africa, and trypanotolerant at the same time. It also should be investigated whether the production level in these high performing F2 animals could not also have been achieved in pure bred N'Dama in the same number of generations (two, plus an additional generation needed to multiply these animals) by 'within breed selection'. However, from Chapter 2 it seems a rather effective and safe option in the sense that production is improved, while trypanotolerance is maintained.

An advantage of creating an F2 is that the N'Dama possesses more features that makes them adaptive to the local environment (Tizard, 2000; Chapter 1). For selection in the F2 population it is of importance to have either aid of QTL or presence of (controlled) infection pressure. Absence of infection pressure or QTL may result in selection of the wrong animals. Expanding towards one backcross generation, followed by an intercross, may work when sufficient numbers of animals are produced. It would include an additional generation of animals that are heterozygous for trypanotolerance genes, which may result in higher costs for medication or production loss under infected, or lower production compared to indigenous breeds in uninfected environment. It is important to realise that, as pointed out in Chapter 4, if hybrid animals are created without infection pressure, these animals should be introduced into infected environment as soon as possible. Apart from the better environment to select on observed production, the production level of the hybrid animals will be lower than that of breeds in uninfected environment, but higher than that of breeds in infected environment. From economic point of view delaying introduction into infected environment would increase costs.

An alternative option for introgression or MAS, is to use only those QTL that are related to production following infection (i.e. 'observed production'). This type of production is reasonably highly correlated to anaemia control following infection (see earlier and also Chapter 2), which is the most important cause of death following a trypanosome infection (Taylor, 1999), and fewer QTL need to be introgressed or applied in

a MAS scheme. It is not clear whether introgression of only those QTL that are related to observed production following a trypanosome infection would be sufficient to create high productive trypanotolerant hybrid cattle breeds.

Organisation of the breeding

A distinction should be drawn between the breeding of improved animals, which should occur in organised breeding programs, and the subsequent dissemination of the improved animals to the farmers. The design of the breeding program will depend on the ecological region and on the production system the breeding program is aimed at. As already explained previously, different ecological regions will require different types of animals. In the highland the higher productive breeds may survive, whereas more harsh environments, indigenous breeds are required. In tsetse infested areas these will mainly include the taurine trypanotolerant breeds, whereas outside these areas, the zebu breeds are in favour. Since this thesis is focusing mainly on the genetic background of trypanotolerance, we will concentrate on that trait when presenting options on the design of a breeding scheme. Depending on the selection strategy and the ecological region the breeding program is set up in, breeding programs may differ in size, relating to both the number of animals and to the number of herds involved. Also, complexity of the selection strategy that can be applied will depend on the local infrastructure. Use of MAS, on the one hand, will require a much more organised infrastructure as compared to phenotypic selection on the other.

There is evidence for genetic variation for trypanotolerance in the N'Dama (Chapters 2 and 3). This genetic variation can be utilised by setting up a breeding scheme. General aim of a breeding scheme is to generate genetic gain. The traits this genetic gain is generated in are determined by the breeding goal. Genetic gain is generated by selecting the best animals and use those as parents of the subsequent generation. The actual selection of these parents is usually organised by a breeding organisation (e.g. farmers co-operative). Such breeding organisation may own a breeding herd, from which the future parents are selected. The genetic gain in that case will heavily depend on the size of the breeding herd. A breeding organisation can also decide to buy part of the breeding material (e.g. part of the sperm) from elsewhere. This is a normal approach in the Holstein breeding associations in the developed countries.

The number of animals a breeding organisation should have at their disposal also depends on the final product they are planning to sell. When the product is sperm for AI use

(or even embryos), then a relatively low number of top cows, and some top bulls are sufficient to continue producing new bulls for sperm production. The breeding organisation can maintain a nucleus herd, where new material is brought in, either by buying in sperm and/or new cows from outside. However, the infrastructure does not always allow for AI, especially in the more remote areas. Also, it is relatively expensive. Another product a breeding organisation may aim at selling is live animals. This may involve bulls, but also (pregnant) cows or weaned calves. In order to be able to offer such product, the breeding organisation will require a much larger herd to serve the same number of farmers. A bull can serve much fewer cows by natural mating than by AI. Of course a combination of the two breeding structures is also possible, because those animals that are not selected for breeding in the "AI breeding program" may be sold to farmers.

Whether the main emphasis will be on production of live animals or on sperm production, depends on the potential market. Close to urban areas the infrastructure is good, and electricity (e.g. for computer, AI facilities like storing semen, etc), water, and, optionally, modern molecular genetic facilities are at one's disposal, and there is a large demand for animal products. Emphasis of a breeding organisation will be on sperm production, and to a lesser extent on production of live animals. In such a breeding scheme, facilities to further increase genetic gain such as utilisation of genetic markers and QTL information, computerised selection strategies such as BLUP, and modern reproduction techniques can be implemented. Pedigree and trait registration would be readily available, and breeding can be organised as an open (Dempfle and Jaitner, 1999) or closed (Syrstad and Ruane, 1998) nucleus scheme.

At a larger distance from the urban areas, infrastructure will most likely be less organised. Roads may not be available during certain seasons, and electricity may not always be available throughout the entire day, which makes it more difficult to for example store and distribute semen (liquid nitrogen required). In those areas there are two options: purchase genetic material from outside (e.g. from the breeding scheme in the urban area), or organise a breeding scheme within the own region. Purchasing genetic material from the breeding organisation close to the city as described above involves the risk that these animals are not able to cope with the usually more harsh environment outside the urban areas. A separate breeding scheme may be set up in the urban area, using genetic material from the more remote areas. Selection in that case will not involve further adaptation to the local environment (e.g. disease resistance). In addition also the delivery of the animals to the area they were selected for may be problematic.

For setting up a breeding program in the region the program is aimed at, distinction may be made between a breeding program aimed at improving cattle of farmers that have a combined farming system that also includes crop farming, and cattle of farmers that travel around with their cattle in the pastoralist system. The breeding goal of these two types of farmers may differ, but, because of their travelling around, it is more difficult for the pastoralist farmers to get organised and set up their own breeding program. For these farmers it would be easier to purchase improved cattle for upgrading their own. The farmers in a combined farming system may want to get organised in a co-operative and set up their own breeding program. A co-operative may reach further than just within the village and include multiple villages in the same area. Such organised breeding can vary from a very basic selection scheme where the best bulls are selected for breeding using phenotypic selection and bulls are circulated across herds (and villages), up to the same level as in the urban areas. In the most basic setting, pedigree registration most often is not very well organised or completely absent (Dempfle and Jaitner, 1999). If there are people that can read and write, a pedigree registration can be set up. Utilising such pedigree registration will only be possible in the case one can have a large enough computer at one's disposal to estimate breeding values. Otherwise all pedigree registration can be used for is avoidance of sib-matings.

Delivery

The possibilities of delivery of the improved genetic material to the farmers can be discussed based on the three basic production systems: 1. highland farming close to an urban area, 2. the mixed farming system in areas further away from urban areas, in (sub)-humid climate, and 3. the pastoralist farming system. In all production systems it applies that whenever a farmer cannot afford to purchase genetic material (semen, calves, cows, etc) from the breeding organisation, they may be able to buy material from those farmers that can. This will create a genetic lag between breeding organisation, 'first buy farmers', i.e. those farmers that can afford to purchase their genetic material directly from the breeding organisation, and those farmers that buy genetic material indirectly (i.e. from for example the 'first buy farmers'). If breeding programs successfully introduce improved animals into a certain area, it will reach all farmers eventually, though with a delay depending on the availability (and thus monetary value) of improved animals and the financial situation of the farmer. Below we will consider the situation of the 'first buy farmers'.

The highland farming area is the area where good facilities are available and for example AI could be implemented. Farmers can buy semen and the breeding organisation may also provide other services like an advisory service on animal husbandry or nutrition. Such additional services can easily be provided by the technician because he comes to the farms for insemination at regular basis. An example of such a system is an AI centre, which sells semen and inseminates cows in the highlands close to Nairobi. Farmers that cannot afford AI, or that prefer to upgrade their cattle in an alternative way, often buy pregnant cows. These cows may have been inseminated with semen of bulls from the AI organisation, or by bulls related to these bulls. In this type of production system there will be a gradual increase in genetic potential.

In the more remote areas, an alternative way to deliver the improved genetic material to the farmer is, for example, by establishing a type of heifer entrust scheme, where a monetary or alternative fee will be established in exchange for which the farmer receives a heifer. Such alternative fee may for example include the first one or two calves that heifer produces, or one or two calves of choice from the already existing herd. A bull entrust scheme would work in a similar way, but in a larger context, for example one bull for an entire village. A 'fee' again will be established in advance, possibly consisting of first choice of the calves or money. The bull could also be leased from the breeding organisation, again with a certain fee in return. Such bull entrust scheme may also work in the pastoralist setting, where the bull can be picked up at the breeding farm, and, for example after half a year, returned to the farm. Another option is to buy weaned calves that were selected out by the breeding organisation.

One last alternative is selection within the own village herd, without the interference of a breeding organisation. In this situation, breeding and delivery are closest together and organised by the same people. In fact, the village is the breeding organisation. The cows animals may be selected from all cows in the village, and those can be mated to the best bulls to breed the bulls for the next generation. More bulls are selected to mate with the rest of the cows and the remaining bulls should be castrated, sold, or kept in a separate herd so that they cannot mate with the cows. Inbreeding in the village herd can be prevented by exchange of bulls between villages. A sort of bull-circulation system could be established.

Role of genetic improvement

Providing improved genetic material does not automatically mean that the production of the genetically superior animals is actually higher compared to the original cattle.

Nutrition and husbandry should be adjusted to the needs of the genetically improved animals. Because small holder farmers only have limited facilities for feed production, the food may have to come from elsewhere. Feeding of local by-products has proven to be very effective in improving the production level of cattle (e.g. Murray et al., 1990). Also, the output of the improved cow, and the profit made from that output, needs to be that much higher to justify the extra costs feed from outside the farm. Whether this leads to extra profits will depend on the local cost-benefit balance; i.e. on both the benefit that can be achieved by selling animal products and the costs of purchasing feed. In some cases total farm output may improve by reducing the number of animals on the farm, so that all available food and attention can be assigned to the best cow(s). Even though this may result in higher output, it will also increase the risk of great loss. If on a small farm high productive animal(s) loose production for some reason (for example illness), or even die, the economic losses are considerable.

The genetically improved animals play an important role in improving the living standard of small holder farmers in large parts of Africa, but it certainly is not automatically achieved by improving the genetic production potential of their cattle alone. Improving the living standard of these people should involve a systems approach. Apart from genetic improvement, also animal nutrition and husbandry should be improved simultaneously to fully appreciate the genetically improved animal. This is a gradual process, and large sudden changes are likely to disrupt the production system, and, therefore, fail to improve the living standard of the small holder farmer.

References

- Agyemang, K., Dwinger, R.H., Little, D.A. and Rowlands, G.J. 1997. Village N'Dama cattle production in West Africa. Six years of research in the Gambia. International Livestock Research Institute, Nairobi, Kenya, and International Trypanotolerance Centre, Banjul, The Gambia. 131 pp.
- Dempfle, L., and Jaitner, J. 1999. Case study about the N'Dama breeding programme at the International Trypanotolerance Centre (ITC) in The Gambia (West Africa). FAO workshop on developing breeding strategies for lower input animal production environment. Rome, 22-25 Sep., 1999.
- Hospital, F., and Charcosset, A., 1997. Marker-assisted introgression of quantitative trait loci. *Genetics* 147: 1469-1485.

- d'Ieteren, G.D.M., Authié, E., Wissocq, N., and Murray, M. 1998. Trypanotolerance, an option for sustainable livestock production in areas at risk from trypanosomosis. *Rev. sci. tech. Off. int. Epiz.* 17: 154-175.
- ILRAD, 1989. N'Dama cattle: managing Africa's genetic resources. ILRAD publication, Nairobi, Kenya, October 1989.
- Koenen, E.P.C. and Groen, A.F. 1998. Genetic evaluation of body weight of lactating Holstein heifers using body measurements and conformation traits. *J. Dairy Sci.* 81: 1709-1713.
- Kosgey, I.S., Van Arendonk, J.A.M., and Baker, R.L. 2001. Breeding objectives for meat sheep in the tropics: inclusion of non-productive roles. submitted to *Livest. Prod. Sci.*
- Koudandé, O.D., Iraqi, F., Thomson, P.C., Teale, A.J., and Van Arendonk, J.A.M. 2000. Strategies to optimize marker-assisted introgression of multiple unlinked QTL. *Mamm. Genome* 11: 145-150.
- Koudandé, O.D. 2000. Introgression of trypanotolerance genes. Wageningen, The Netherlands, PhD thesis, pp 127.
- Murray, C., Murray, M., Murray, P.K. Morrison, W.I., Pyne, C., and McIntyre, W.I.M. 1977. Diagnosis of African trypanosomiasis in cattle. Improved parasitological and serological techniques. In: International Scientific Council for Trypanosomiasis Research and Control 15th Meet., the Gambia. OAU/STRC. Publication No. 110, p. 247-254.
- Murray, M., Clifford, D.J., and Gettinby, G. 1981. Susceptibility to African trypanosomiasis of N'Dama and Zebu cattle in an area of *Glossina morsitans submorsitans* challenge. *Vet. Rec.* 109: 503-510.
- Murray, M., Trail, J.C.M., Davis, C.E. and Black, S.J. 1984. Genetic resistance African to trypanosomiasis. *J. Infect. Dis.* 149: 311-319.
- Murray, M., Trail, J.C.M. and d'Ieteren, G.D.M. 1990. Trypanotolerance in cattle and prospects for the control of trypanosomiasis by selective breeding. *Rev. sci. tech. Off. Int. Epiz.* 9:369-386.
- Murray, M., Stear, M.J., Trail, J.C.M., d'Ieteren, G.D.M., Agyemang, K., and Dwinger, R.H. 1991. Trypanosomiasis in cattle: prospects for control. In: Breeding for disease resistance in farm animals. Edited by J.B. Owen and R.F.E. Axford. CAB International, Wallingford UK. 1991. 499 pp.
- Roberts, C.J., and Gray, A.R. 1973. Studies on trypanosome-resistant cattle. II. The effect of trypanosomiasis on N'Dama, Muturu and Zebu cattle. *Trop. Anim. Hlth.* 5: 220-233.

- Soller, M. & Beckman, J.S. 1987. Toward an understanding of the genetic basis for trypanotolerance in the N'Dama cattle of West Africa. Consultation Report FAO, Rome.
- Stewart, J.L. 1951. The West African shorthorn cattle; their value to Africa as trypanosomiasis-resistant animals. *Vet. Rec.* 63:454-457.
- Syrstad, O, and Ruane, J. 1998. Prospects and strategies for genetic improvement of the dairy potential of tropical cattle by selection. *Trop. Anim. Health and Prod.* 30: 257-268.
- Taylor, K.A.. 1999. Immune responses of cattle to African trypanosomes: protective or pathogenic? *Int. J. Parasitol.* 28:219-240.
- Tizard, I.R. 2000. Veterinary immunology: an introduction. sixth edition. Philadelphia, USA, 482 pp.
- Trail, J.C.M., d'Ieteren, G.D.M., Murray, M., Ordner, G., Yangari, G., Collardelle, C., Sauveroche, B., Maille, J.C. and Viviani, P. 1993. Measurement of trypanotolerance criteria and their effect on reproductive performance of N'Dama cattle. *Veterinary Parasitology* 45: 241-255.
- Weller, J.I., Kashi, Y., and Soller, M. 1990. Power of daughter and granddaughter designs for determining linkage between marker loci and quantitative trait loci in dairy cattle. *J. Dairy Sci.* 73 : 2525-2537.
- Wilson, R.T. 1995. Livestock production systems. Edited by R. Coste and A.J. Smith. The tropical agriculturalist. Wageningen, The Netherlands, pp. 141.

Summary

Africa is a continent with a very rapid increase in human population and a dramatic urbanisation, and as a consequence, the demand for meat and dairy products is also increasing rapidly. To meet his growing demand and increase in African livestock production is desirable. However, diseases like trypanosomosis have a seriously limiting influence on production, as well as on animal welfare, in a substantial part of sub-Saharan Africa. The trypanosusceptible *Bos indicus* breeds, such as the Kenyan Boran, are not well capable of handling the infection. They become anaemic, lose weight, show reduced milk production and reduced capacity to work, lose fertility, often abort the foetus, and, unless treated with medication, frequently die of the infection. Some West African *Bos taurus* cattle breeds (N'Dama and West African Shorthorn) are tolerant of trypanosomosis. Members of those breeds are able to gain weight and show a normal oestrus cycle and thus maintain reproductive capacity in areas with moderate infection pressure. An F2 experiment was set up in cattle at the International Livestock Research Institute (ILRI) in Nairobi, Kenya, to determine the genetic background of trypanotolerance in cattle.

This thesis basically had two main aims: 1. to define traits related to trypanotolerance and test those on the data from the F2 cattle experiment in order to determine the genetic factors (heritability and QTL) involved in trypanotolerance, and 2. to explore alternative breeding schemes to improve both production and disease resistance.

In literature, changes in PCV and growth rate following infection are considered to be indicators for trypanotolerance. **Chapter 2** focuses on description of changes in, and relations between Packed red Cell Volume percent (PCV), body weight and parasite count following an infection with a single clone of *Trypanosoma congolense* IL1180. In total 214 F2 cattle (Gambian N'Dama x Kenyan Boran) were infected by the bites of infected tsetse flies. Body weight, PCV, and parasite counts were recorded on a weekly basis for 150 days post infection. Seventeen derived traits were defined based on the data recorded. The average of F2 animals were intermediate between the pure-bred N'Dama and Boran for all traits. The highest and lowest responders in the F2, when selected on maximum drop in PCV or on maximum drop in body weight, were equal to the average of the Boran and N'Dama breeds, respectively. There were moderate to low phenotypic correlations (0.00 to

Summary

0.32) between average log(parasite count) or number of times an animal was detected parasitaemic, and the PCV and body weight derived traits. There were low to moderate phenotypic correlations (0.02 to 0.74) between and (0.01 to 0.96) within PCV and body weight derived traits. Most of the traits defined in this study are heritable. Heritabilities ranged from 0.01 for PCV recovery, to 0.88, for initial PCV. Some F2 animals seem to be able to control anaemia and have a higher average body weight and body weight gain than the pure-bred N'Dama. Body weight gain following infection seems an appropriate and easy to measure indicator of trypanotolerance.

The F2 experiment was designed to detect QTL related to trypanotolerance in cattle, and QTL were detected based on data generated from this experiment, results of which are presented in **Chapter 3**. An initial analysis revealed a number of chromosomes harbouring significant QTL. In mice there is evidence that the gene *Tir1*, relating to survival following a trypanosome infection, is genomically imprinted. In this paper these chromosomes are further analysed to determine the mode of expression of these QTL. In this chapter QTL were detected for traits related to anaemia, body weight change, and parasitaemia following a trypanosome infection on nine different chromosomes. In total eight QTL have been detected, most of them with Mendelian expression. One QTL (for the recovery of PCV following infection on BTA19) showed evidence of genomic imprinting, with maternal expression. Size of effect of the QTL ranged from 5.8 percent of the phenotypic variance for initial body weight prior to infection on BTA2 and the recovery in PCV following infection on BTA7, to 10.4 percent for the lowest value of PCV reached following infection, on BTA2.

One alternative of utilising QTL in a breeding scheme is through marker assisted introgression. The genetic and economic consequences of introgression of either one or two genes that explain the complete difference for disease resistance between donor and recipient breed were investigated in **Chapter 4**. 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 (4 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 population. Introgression of one or two genes made no difference for the genetic results. From an economic point of view, optimal number of generations depended 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 with low female reproduction capacity.

In **Chapter 5** a model is presented describing the relationship between level of disease resistance and production under constant infection pressure. The model assumes that given a certain infection pressure, there is a threshold for resistance below which animals will stop producing, and that there is also a threshold for resistance above which animals produce at production potential. In between both thresholds animals will show a decrease in production, the size of decrease depending on the severity of infection, and the level of resistance. The dynamic relationship between production and resistance when level of resistance changes, such as due to infection is modeled both stochastically and deterministically. Phenotypic selection started in a population with very poor level of resistance introduced in an environment with constant infection pressure. Mass selection on observed production was applied, which resulted in a non-linear selection response for all three traits considered. When resistance is poor, selection for observed production resulted in increased level of resistance. With increasing level of resistance, selection response shifts to production potential and eventually selection for observed production is equivalent to selection for production potential. The rate at which resistance is improved depends on its heritability, the difference between both thresholds, and selection intensity. The model also revealed that when a zero correlation between resistance and production potential is assumed, the phenotypic correlation between resistance and observed production level increases for low levels of resistance and subsequently asymptotes to zero, whereas the phenotypic correlation between production potential and observed production asymptotes to one. For most breeding schemes investigated, the deterministic model performed well in relation to the stochastic simulation results. Experimental results reported in literature support the model predictions.

Animals will show reduced production when exposed to a constant infection pressure unless they are fully resistant, the size of the reduction depending on the degree of resistance

and the severity of infection. In **Chapter 6**, the use of QTL for disease resistance for improving productivity under constant infection pressure is investigated using stochastic simulation. Marker assisted selection was incorporated in the model described in **Chapter 5** by assuming a certain proportion of the genetic variance to be explained by the QTL. Phenotypes were available for production, not for resistance. In practice the infection pressure may vary across time, thus, results were compared to mass selection on production under constant, as well as intermittent infection pressure, where the infection pressure varied between but not within years. Selection started in a population with a very poor level of resistance. Incorporation of QTL information is valuable (i.e. the increase in observed production relative to mass selection) when a large proportion of the additive genetic variance is explained by the QTL (50% genetic variance explained) and when the heritability for resistance is low ($h^2_R = 0.1$). Under constant infection pressure, incorporating QTL information does not increase selection responses in observed production when the QTL effect explains less 25% of the genetic variance. Under intermittent selection pressure, the use of QTL information gives a slightly greater increase in observed production in early generations, relative to mass selection on observed production, but still only when the QTL effect is large or the heritability for resistance is low. The additional advantage of incorporating QTL information, is that (preventive) medical treatment is possible, or productivity of animals may be evaluated in non-infected environments.

In **Chapter 7** the results presented in chapters 2 to 6 are used to discuss possibilities to improve the genetic potential of cattle in Africa, with special emphasis on small holder farmers in tsetse-infested environments. There is no one breeding structure of choice for the African continent, but the design of a breeding program depends on the production system and, in connection to that, on the local environment and climate. Breeding goals, for example, will differ in urban areas (milk production) and more remote areas (animal health and reproduction). Infection pressure and climatic conditions have a large influence on the production system, and, therewith, also on the type of animals required. Close to an urban centre the infection pressure usually is low and there is a large demand for milk and meat. Infrastructure is good and breeding programs may include non-indigenous breeds and make use of facilities like AI, pedigree registration, and maybe even molecular genetic techniques so that marker assisted selection can be applied. Outside urban areas, the infrastructure most likely is poorer, infection pressure higher, and possibly the climate more demanding. In these circumstances indigenous breeds are required, which may be improved applying

phenotypic selection, or, if the facilities are available, using BLUP. It is less likely that marker assisted selection techniques would be applicable here in the near future.

In general it can be concluded that there is potential of genetic improvement of cattle in Africa and the organisation should be adjusted to the local circumstances. However, it is very important to realise that not only the genetic potential, but also the nutrition and the husbandry should be improved.

Samenvatting

Inleiding

Trypanosomosis, ook wel slaapziekte genoemd, is een parasitaire aandoening, die voornamelijk wordt overgedragen door de tsetsee-vlieg, die al miljoenen jaren in Afrika voorkomt. Trypanosomosis is de belangrijkste veeziekte in Afrika, en is, naast het rund, ook besmettelijk voor veel andere typen vee zoals schapen, geiten, ezels, paarden, kamelen, vissen, en voor de mens. Geïnfecteerde dieren worden anemisch, verliezen gewicht, hebben een verlaagde melkproductie en trekvermogen, zijn minder vruchtbaar en aborteren vaak, en uiteindelijk sterven veel dieren aan de gevolgen van de infectie, tenzij ze behandeld worden met medicijnen. Er zijn verschillende soorten trypanosomen, waarvan sommige ook kunnen worden overgedragen op de mens. De tsetsee-vlieg komt voor in een gebied dat 11 miljoen km² beslaat, wat gelijk staat aan 37% van het continent en waar 40 landen in liggen. Er is geen vaccin beschikbaar tegen trypanosomosis. Momenteel worden er naar schatting 46 miljoen koeien in geïnfecteerd gebied gehouden, waarvan er 17 miljoen worden behandeld met medicijnen. De jaarlijkse kosten ten gevolge van de infectie worden geschat op \$35 000 000. Wanneer daar de verliezen als gevolg van productiederving, sterfte, verminderde vruchtbaarheid, etc. worden bijgeteld, dan komen deze kosten op meer dan \$100 000 000 per jaar.

In West Afrika komen koeien rassen van het *Bos taurus* type voor die gedeeltelijk resistent zijn tegen een trypanosoom-infectie: ze zijn trypanotolerant. Dat wil zeggen, deze koeien zijn in staat om onder infectiedruk te blijven produceren en de reproductie in stand te houden. Zonder infectie is het productieniveau van deze koeien hoger, wat erop wijst dat het resistentie niveau niet volledig is. Het grootste trypanotolerante ras, uitgedrukt in aantallen dieren, is de N'Dama. Kenmerkend voor dieren van dit ras (en andere trypanotolerante rassen) is de kleine afmeting, welke gepaard gaat met een lagere absolute melk en vlees productie in vergelijking met verschillende Afrikaanse runderrassen van het *Bos indicus* type (zebu). Een andere belangrijke reden voor het lagere productieniveau van de N'Dama, in vergelijking met verschillende zebu rassen, is dat het productieniveau van de N'Dama nagenoeg uitsluitend gemeten is onder geïnfecteerde omstandigheden. Runderen van het *Bos indicus* type zijn niet resistent tegen een trypanosoom-infectie en moeten worden behandeld met medicijnen wanneer ze in geïnfecteerd gebied worden gehouden. En zelfs dan hangt het van de infectiedruk af of ze kunnen blijven (re)produceren, of het zelfs

overleven. In geïnfecteerd gebied kan het dus zijn dat het productieniveau van een N'Dama zonder medicijn behandeling hoger is dan dat van een zebu die behandeld is met medicijnen. De N'Dama is, naast trypanotolerant, ook resistent tegen verschillende andere infecties zoals een aantal ziektes die door teken worden overgebracht als dermatophilosis, heartwater, bovine anaplasmosis en bovine babesiosis. De N'Dama is beter bestand tegen strongloïde worm infecties en wanneer ze geïnfecteerd zijn hebben ze een lagere eier-uitscheiding dan zebu rassen. Ook kan de N'Dama beter tegen langere perioden van droogte en voertekort. Met andere woorden: de N'Dama heeft zich goed aangepast aan de lokale omstandigheden. Het productie niveau van de N'Dama zou door selectie verbeterd kunnen worden, hetgeen tot nu toe om verschillende redenen nog niet serieus is geprobeerd.

Afrika is een continent met een sterke bevolkingsgroei en een bijzonder snelle urbanisatie. Consequentie is een groeiende vraag naar melk en vlees, met name rond de steden. Om aan deze vraag te kunnen voldoen is het noodzakelijk om de productiecapaciteit van de veestapel te verhogen. Probleem hierbij is dat verschillende ziekten, zoals trypanosomosis, een productie- en dieren welzijn verlagende invloed hebben in grote delen van Afrika. Met fokkerijmaatregelen kunnen de prestaties van dieren onder lokale omstandigheden verbeterd worden, maar goede handvaten en inzichten hiervoor ontbreken op dit moment.

Doel van dit proefschrift was om de genetische achtergrond van trypanotolerantie bij koeien in kaart te brengen en om plannen te maken om de daarbij vrijkomende informatie te gebruiken bij de het ontwerpen van fokkerijstrategieën om de productie in tsetsee-geïnfesteerde gebieden in Afrika te verhogen.

Trypanotolerantie

De immunologische achtergrond van trypanotolerantie is nog niet helemaal duidelijk. Het lijkt erop dat het aangeboren deel van het immuunsysteem (inate immune system) een belangrijkere rol speelt dan het deel van het immuunsysteem dat een geheugen opbouwt volgend op een eerdere infectie. Over de precieze werking van de inate immune response is nog weinig bekend. Het is daarom ook niet duidelijk wat er gemeten zou moeten worden om de mate van resistentie van een dier tegen een infectie met trypanosomen te bepalen. Wel is bekend dat dieren na infectie anemisch worden, waarvan de mate wordt uitgedrukt in het PCV (Packed Cell Volume, een maat voor het aantal rode bloedcellen). Ook is bekend dat dieren na infectie een verlaagde productie laten zien (bijvoorbeeld verminderde groei of

zelfs gewichtsverlies). Nog niet helemaal duidelijk is of het aantal parasieten (trypanosomen) in het bloed een indicatie geeft van de mate van trypanotolerantie. Het is bekend dat sommige wilde dieren, zoals de Afrikaanse buffel, weinig door een trypanosoom infectie worden beïnvloed, terwijl er wel parasieten in de bloedbaan voorkomen.

Op het International Livestock Research Institute (ILRI) in Nairobi, Kenia, is een F2-kruisingsexperiment (een kruising tussen twee rassen, waarvan de nakomelingen (de F1) onderling weer gekruist zijn (de F2)) opgezet met vijf N'Dama stieren en vijf Boran koeien, om stukken van een chromosoom op te sporen waar genen liggen die verantwoordelijk zijn voor trypanotolerantie in het rund (Quantitative Trait Loci QTL). Rondom Nairobi komen geen tsetsee-vliegen voor. Alle F2 dieren zijn gecontroleerd besmet door beten van een geïnfecteerde tsetsee-vlieg en vervolgens zijn wekelijks, voor een periode van 150 dagen, het verloop van het PCV, lichaamsgewicht, en aantal parasieten in het bloed na infectie gemeten bij alle F2 dieren. Dit maakte het mogelijk een beeld te vormen van wat het verloop voor deze kenmerken is van een typisch trypanotolerant dier en van een typisch vatbaar dier (Hoofdstuk 2). Ter vergelijking zijn, tegelijk met de F2 dieren, ook zes zuivere N'Dama en zes zuivere Boran dieren geïnfecteerd en gevolgd.

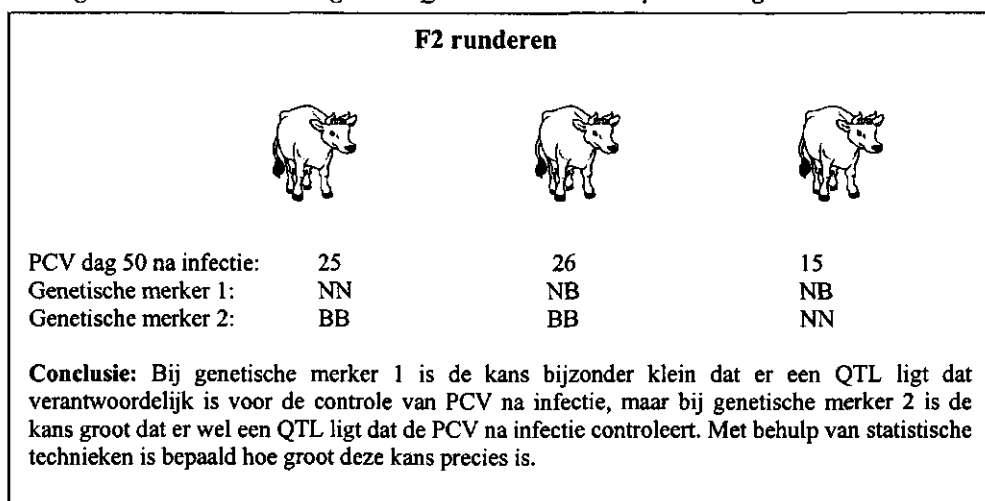
Aan de hand van het verschil in verloop in de kenmerken tussen tolerante en vatbare dieren zijn een zeventiental afgeleide kenmerken gedefinieerd die gerelateerd zijn aan trypanotolerantie. De correlaties tussen deze kenmerken zijn bepaald. Een hoge correlatie geeft aan dat een verandering in het ene kenmerk samen gaat met een verandering in het andere kenmerk. Het blijkt dat (de logaritme van) het aantal parasieten in het bloed niet tot middelmatig (0.00 tot 0.32) is gecorreleerd met PCV en lichaamsgewicht gerelateerde kenmerken. De PCV gerelateerde kenmerken waren niet tot sterk gecorreleerd (0.02 tot 0.74) met de lichaamsgewicht gerelateerde kenmerken. Binnen de categorieën PCV en lichaamsgewicht gerelateerde kenmerken waren de kenmerken ook niet tot sterk gecorreleerd (0.01 tot 0.96). Op basis van de familierelaties tussen de F2 dieren was het mogelijk om de variantie in de kenmerken uit te splitsen in een deel dat is bepaald door het verschil in genetische aanleg tussen de dieren en een deel dat door invloeden van buitenaf (bv verschil in voeding, seizoen, etc) wordt bepaald. Deze variantiecomponenten worden gebruikt om de erfelijkheidsgraad (de ratio van de genetische variantie en de totale variantie) te bepalen. De erfelijkheidsgraad (aangegeven met h^2) varieerde van 0.01 voor het herstel in PCV na infectie, tot 0.88 voor PCV voor infectie. Er zijn F2 dieren die hun verval in PCV na infectie op dezelfde manier kunnen controleren als de N'Dama, maar met een hoger gemiddeld lichaamsgewicht en met een hogere toename in lichaamsgewicht na

infectie. In het algemeen lijkt groei na infectie een goede en makkelijk te meten indicator van trypanotolerantie in nog niet volwassen dieren.

QTL voor trypanotolerantie

Van de F2 dieren zijn bloedmonsters genomen, waarna van vooraf bepaalde stukjes DNA de code is bepaald. Dit zijn zogenaamde genetische merkers. Op elk van de 29 chromosomen van het rund zijn zeven tot 15 van zulke merkers bij alle dieren in kaart gebracht. De informatie over deze, voor de meeste dieren verschillende, stukjes DNA code wordt gecombineerd met de aan PCV, lichaamsgewicht en aantal parasieten gerelateerde kenmerken, waarvan de waarde ook per dier verschillend is. Gegeven de genetische merkers kan nu voor elk stukje van een chromosoom bekeken worden wat de kans is dat daar een QTL ligt dat verantwoordelijk is voor de verschillen tussen dieren voor dat kenmerk. Bijvoorbeeld, als er een aantal dieren dezelfde code hebben voor een genetische merker, en alleen deze dieren kunnen hun verval in PCV na infectie beperken, dan is de kans groot dat op dat stukje chromosoom een QTL ligt dat ervoor zorgt dat de PCV na infectie niet te veel zakt (Figuur 1).

Figuur 1. Schematische uitleg van de QTL detectie met behulp van twee genetische merkers.



Bij de mens en de muis (en in minder detail bij enkele andere diersoorten) is het bekend dat er voor sommige genen sprake is van een "parent of origin" effect, wat wil zeggen dat uitsluitend het allel tot expressie komt wat bijvoorbeeld via de moeder en niet wat via de vader wordt overgeërfd (of omgekeerd). Dit fenomeen heet genetische

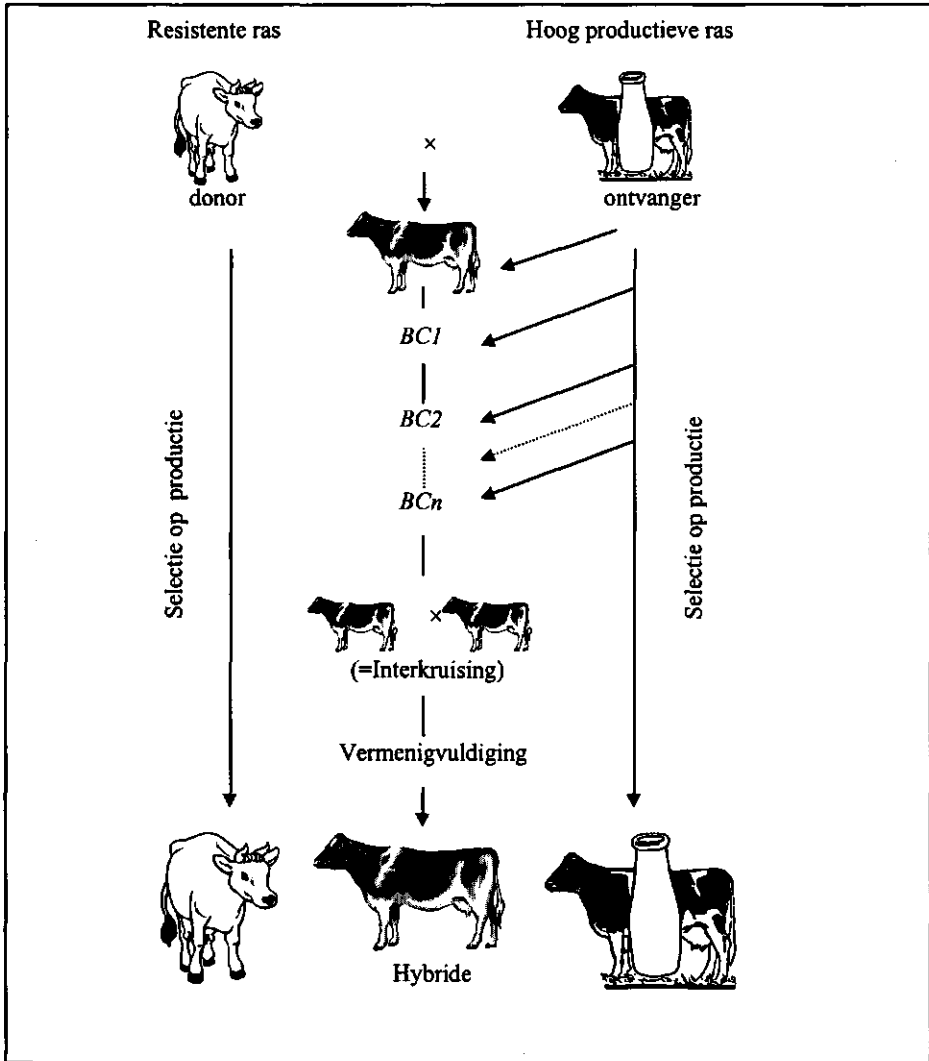
inprenting. In de muis is er bewijs gevonden voor inprenting van een QTL, *Tir 1*, dat gerelateerd is aan overlevingskans na een trypanosoom infectie. Dit was de aanleiding om na te gaan of er ook sprake is van inprenting van een QTL wat gerelateerd is aan trypanotolerantie bij het rund. Voor F2 dieren die voor een bepaalde genetische merker zowel een allel afkomstig van de N'Dama, als een allel afkomstig van de Boran bezitten, is het van belang terug te traceren of de vader, danwel de moeder het N'Dama allel heeft doorgegeven. Hiervoor moesten de genotypen terug gecodeerd worden naar de oorspronkelijke genotypen. Gezien de hoeveelheid werk dat dit met zich meebracht is, op basis van resultaten van een QTL analyse die op ILRI is uitgevoerd, een selectie van acht chromosomen gemaakt. Voor elk van de kenmerken is er op deze acht chromosomen gekeken of er een QTL aanwezig is en of er sprake was van inprenting (Hoofdstuk 3).

In totaal is er bewijs gevonden voor de aanwezigheid van acht QTL op deze acht chromosomen, waarvan er bij één bewijs is dat het alleen tot expressie komt wanneer het wordt overgeërfd via de moeder. Dit QTL ligt op chromosoom 19 en bepaalt een deel van het herstel van PCV na infectie. De meeste QTL beïnvloedden meer dan één kenmerk. De grootte van het effect van de acht QTL varieerde van 5,8% van de totale variantie voor lichaamsgewicht voor infectie op chromosoom 2 en het herstel van PCV na infectie op chromosoom 7, tot 10,4% van de totale variantie voor de minimum waarde voor PCV na infectie op chromosoom 2 (zelfde QTL), wat een aanzienlijk effect is.

Gebruik van QTL bij introgressie

QTL die een deel van resistentie tegen een bepaalde infectie verklaren kunnen worden gebruikt in een introgressie schema om de resistentie van een ander, bijvoorbeeld hoger productief, ras te verbeteren. Hierbij worden beide rassen gekruist, waarna er terug gekruist wordt met het hoger productieve ras, waarbij in de gaten gehouden wordt dat het gewenste QTL wordt overgeërfd (Introgressie, zie Figuur 2). Met elke generatie terugkruising wordt de fractie (ongunstige) genen van het donor ras gehalveerd. Het aantal generaties terugkruisen kan variëren en hangt ondermeer af van hoe nauwkeurig de positie van het QTL bekend is en hoe groot het verschil in productie is tussen de beide uitgangsrassen. Na een bepaald aantal terugkruisingen worden de kruisingsdieren niet langer met het onvanger ras, maar onderling gepaard. Dit heet de interkruising en heeft ten doel de gunstige QTL allelen te fixeren. In het geval van introgressie van een enkel gen zal een vierde deel van de nakomelingen uit deze paringen homozygoot zijn voor het gunstige allel (de hybride dieren). Vervolgens is vaak nog één of meerdere generaties nodig om het aantal

Figuur 2. Schematische weergave van gebruik van introgressie voor vorming van een hybride. Binnen de uitgangsrassen wordt geselecteerd op productie. BC staat voor de eventuele Back Cross (terug kruising)



hybride dieren te vergroten zodat deze ingezet kunnen worden als afzonderlijk ras, of om andere rassen op te waarderen.

In Hoofdstuk 4 is onderzocht wat de genetische en economische consequenties zouden zijn van introgressie van één of twee genen die de totale resistentie tegen een levensbedreigende infectie bepalen. Dit gen of deze genen moeten homozygoot in een dier aanwezig zijn om voldoende resistentie te verkrijgen om in geïnfecteerd gebied te kunnen

overleven. De verschillen in productie tussen de uitgangsrassen is gevarieerd, alsook het aantal generaties terugkruisen met het ontvanger ras (0, 1, 3 of 7). Voor de vrouwelijke reproductie capaciteit is gekeken naar een situatie met natuurlijke voortplanting en een met toepassing van embryo transplantatie. De resultaten van introgressie zijn in Hoofdstuk 4 vergeleken met selectie op productie binnen het donorras (zie ook figuur 2).

Het blijkt dat een groot aantal generaties terugkruising vanuit genetisch oogpunt gunstig is. Een groot verschil tussen de uitgangsrassen, in combinatie met embryo transplantatie en zeven generaties terugkruisen, resulteerde in het grootste verschil in productieniveau tussen het donor ras en de hybride. Introgressie van één of twee genen maakte geen verschil vanuit genetisch oogpunt. Echter, vanuit economisch oogpunt bepaalt het aantal genen, de reproductiecapaciteit van de vrouwelijke dieren, en het productieverschil in de uitgangsrassen het optimale aantal generaties terugkruisen bij introgressie. Zeven generaties terugkruising is in de meeste gevallen niet rendabel en nul tot drie generaties is meer optimaal. Introgressie van twee genen is economisch minder aantrekkelijk, met name wanneer geen gebruik gemaakt kan worden van embryo transplantatie.

Selectie op productie binnen ras onder constante infectiedruk

In Hoofdstuk 5 wordt gekeken naar de efficiëntie van verschillende selectie maatregelen om genetische aanleg voor productie en resistentie te verbeteren. Daarvoor is een model ontwikkeld waarmee relatie tussen waargenomen productie onder constante infectiedruk enerzijds, en productie zonder infectiedruk en resistentie anderzijds wordt beschreven. In het model wordt aangenomen dat bij een gegeven infectiedruk er een resistentieniveau bestaat waaronder productie niet meer mogelijk is (de lage drempelwaarde), en dat er ook een resistentieniveau is waarboven de productie niet meer wordt beïnvloed door de infectie (de hoge drempelwaarde). Voor elke infectie die continue aanwezig is, zal de productie tussen deze beide resistentieniveaus worden verlaagd door het gebrek aan resistentie, waarbij de mate van verlaging afhangt van de infectiedruk en van de mate van resistentie. Dit dynamische model is zowel stochastisch als deterministisch gesimuleerd. Uitgangspunt was een populatie met een heel laag resistentie niveau, in een omgeving met constante infectiedruk.

Selectie op waargenomen productie niveau resulteerde in een niet-lineaire response in de waargenomen productie, resistentie, en de potentiële productie wanneer er geen infectie was geweest. Bij een lage mate van resistentie resulteert selectie op waargenomen

productie met name in een toename in aanleg voor resistentie. Bij een toenemende resistentie in de populatie verschuift de selectie respons meer in de richting van potentiële productie en uiteindelijk is selectie op waargenomen productie gelijk aan selectie op potentiële productie.

De mate waarin resistentie wordt verbeterd door selectie op waargenomen productie is afhankelijk van de erfelijkheidsgraad, het verschil tussen beide drempelwaarden, en de selectieintensiteit. Het model laat op overtuigende wijze zien dat de relatie tussen waargenomen productie en resistentie onder infectiedruk verandert met de verandering in resistentie. In een situatie waarbij resistentie en potentiële productie genetische ongecorreleerd zijn, zal bij een laag resistentie niveau de waargenomen productie heel laag tot afwezig zijn. Als gevolg van selectie op waargenomen productie zal de correlatie tussen resistentie en waargenomen productie aanvankelijk stijgen om vervolgens, wanneer de populatie een hoog resistentie niveau heeft bereikt, weer af te nemen. De correlatie wordt uiteindelijk nul wanneer alle dieren in de populatie een resistentiewaarde hebben die boven de bovenste drempelwaarde ligt. Voor de meeste selectieschema's die zijn onderzocht voldoet het deterministische model goed in relatie tot de stochastische resultaten. Resultaten van experimenten die beschreven zijn in de literatuur lijken het model te onderschrijven.

Gebruik van QTL voor om productie onder infectiedruk te verhogen

Zolang dieren niet volledig resistent zijn, zal hun productie worden verlaagd wanneer ze blootgesteld worden aan een infectie. In Hoofdstuk 6 wordt het gebruik van QTL voor resistentie om de productiviteit onder infectiedruk te verhogen onderzocht met behulp van stochastische simulatie. Gebruik van QTL informatie is ingebouwd in het model dat beschreven wordt in Hoofdstuk 5. Hierbij wordt aangenomen, dat een bepaald deel van de genetische variantie wordt verklaard door het QTL. Selectie beslissingen worden gebaseerd op QTL informatie en waarnemingen aan potentiële productie. Er is aangenomen dat selectie plaats vond buiten geïnfecteerd gebied, of dat medicijnbehandeling afdoende was om het waargenomen productieniveau gelijk te maken aan het potentiële productieniveau. Resultaten zijn vergeleken met massa selectie onder constante en onder niet-constante infectiedruk waarbij de infectiedruk varieerde tussen, en niet binnen jaren. Uitgangspopulatie was gelijk aan die van Hoofdstuk 5.

Het blijkt dat opnemen van QTL informatie zinvol is wanneer een groot deel van de additief genetische variantie (50%) wordt verklaard door het QTL, en wanneer de

erfelijkheidsgraad voor resistentie laag is ($h^2_R = 0.1$). Bij een constante infectiedruk zorgt opname van QTL informatie niet voor een toename in respons in waargenomen productie wanneer de QTL minder dan 25% van de genetische variantie verklaard. In de praktijk zal de infectiedruk over het algemeen variëren over de tijd. Bij niet-continue infectiedruk resulteert gebruik van QTL informatie bij selectie, in vergelijking met massa selectie op waargenomen productie, in een iets grotere toename in waargenomen productie gedurende de eerste generaties. De toegevoegde waarde van opname van QTL informatie is met name dat (preventieve) behandeling met medicijnen mogelijk is, en ook dat de productie van dieren bepaald kan worden buiten geïnfecteerd gebied.

Algemene discussie

In hoofdstuk 7 worden de resultaten uit de hoofdstukken 2 tot en met 6 gebruikt om mogelijkheden tot verbetering van de genetische aanleg voor trypanotolerantie en daarmee het productie potentieel van rundvee in Afrika te bediscussiëren in een meer praktische context. Er bestaat niet zoiets als dé selectie strategie voor heel Afrika. Het ontwerp van het fokprogramma hangt af van het productiesysteem en, in verband daarmee, het lokale milieu en klimaat. De gewenste richting en mate van verandering in de genetische aanleg (fokdoel), bijvoorbeeld, zal in stedelijk gebied verschillen van de meer afgelegen gebieden. Infectiedruk en klimatologische condities hebben een grote invloed op het productiesysteem en, daarmee, ook op het type dieren dat daarvoor nodig is. In de nabijheid van een stedelijk gebied zal de infectiedruk normaliter lager zijn, is de vraag naar melk en vlees groter, en is de infrastructuur vaak beter. Fokprogramma's kunnen daarom vaak niet-lokale rassen opnemen en gebruik maken van voorzieningen als afstamming registratie, kunstmatige inseminatie, en misschien zelfs van moleculair genetische technieken zodat QTL informatie kan worden benut. Buiten de stedelijke gebieden is de infrastructuur vaak minder goed, de infectiedruk hoger, en de productieomstandigheden meer veeleisend. Onder deze omstandigheden genieten lokale rassen de voorkeur. Verbetering van deze rassen is mogelijk door selectie van dieren op basis van eigen prestatie of, wanneer de faciliteiten dat toelaten, op basis van fokwaarden waarbij bij het schatten rekening gehouden wordt met de prestaties gemeten aan verwante dieren. Het is minder waarschijnlijk dat hierbij in de nabije toekomst genetische merker informatie zal worden gebruikt.

In het algemeen kan worden geconcludeerd dat het genetisch potentieel van rundvee in Afrika verbeterd kan worden en de organisatie daarvan zal moeten worden toegespitst op

Samenvatting

de lokale omstandigheden. Voor een verbetering van de prestaties van dieren onder lokale omstandigheden is het belangrijk dat, naast de genetische aanleg, ook andere elementen van het productiesysteem zoals voeding en huisvesting in ogenschouw worden genomen.

Nawoord

Tijdens een project van vier jaar zijn er altijd punten die je liever anders had gezien, maar de afgelopen vier jaar heb ik voor het overgrote deel met heel veel plezier aan mijn promotie onderzoek gewerkt. De goede sfeer binnen de leerstoelgroep fokkerij en genetica, maar ook binnen Lab7 op ILRI, heeft daar zeker toe bijgedragen. Goede collega's zijn minstens zo belangrijk voor het plezier in je werk als werk dat je ligt. Een aantal mensen zou ik hier speciaal willen bedanken voor hun bijdrage aan het verloop van mijn onderzoek.

Allereerst Johan van Arendonk, mijn co-promoter die later mijn promotor werd. Johan, het is niet altijd een makkelijk project geweest, maar jouw gevoel voor politiek heeft er toe bijgedragen dat het altijd weer op zijn pootjes terecht is gekomen. Wat ik bijzonder in je waardeer is dat je je promovendi in principe de vrijheid laat om een eigen pad te banen, maar dat je er altijd bent als ze je nodig hebben. Voor mij een ideale formule. Then my overseas co-workers, who became my co-promoters: Steve Bishop and John Gibson. Steve, it has been a great pleasure to work with you. Also, your interest and support during the rest of the project have been very much appreciated. John, I have really enjoyed our meetings. You and Johan are a perfect combination for reviewing my papers. You both give very useful comments, but from different focus points.

Another person who has been closely involved in the ILRI part of my project is Olivier Hanotte. Olivier, I know I have been hassling you quite a bit with difficult data questions, thanks for remaining enthusiastic all the same. Dan wil ik graag mijn begeleidingscommissie bedanken: Henk Bovenhuis, Pim Brascamp, Luc Janss en Henk Udo: bedankt voor de kritische input op mijn papers en het meedenken over de verdere invulling van het project.

Langs deze weg wil ik ook graag mijn ouders bedanken voor de steun en interesse die ze al die jaren getoond hebben. Handig als je dezelfde dingen leuk vindt. En natuurlijk Piter, bedankt....

Crisbeth

Curriculum Vitae

Ik, Elisabeth Hillechien van der Waaij, ben op 22 april 1968 in Voorburg geboren. Op zevenjarige leeftijd ben ik verhuisd naar Haren (Gr). In 1987 behaalde ik mijn atheneum diploma aan het St Maartenscollege in Haren.

In 1988 begon ik mijn studie Zoötechniek aan de toenmalige Landbouwwuniversiteit in Wageningen. In januari 1993 ben ik voor zes maanden naar de Sveriges Lantbruksuniversitet in Uppsala, Zweden vertrokken voor een stage waarbij ik het keuringssysteem voor rijaardhengsten in Zweden heb vergeleken met dat in enkele andere Europese landen. Bij terugkomst ben ik begonnen aan een afstudeervak bij de leerstoelgroep Fokkerij en Genetica met als onderwerp het, met behulp van Gibbs sampling, schatten van genetische parameters voor groeikenmerken bij schapen. Daarna heb ik bij de leerstoelgroep Kwantitatieve Epidemiologie een afstudeervak gedaan met als onderwerp een op literatuur gebaseerde evaluatie van het EpiMan systeem. Dat is een systeem dat, gebruik makend van oa Geografische Informatie Systemen, voorschrijft hoe op een uitbraak van een besmettelijke dierziekte gereageerd zou moeten worden.

In januari 1995 ben ik afgestudeerd, waarna ik in mei 1995 voor zes maanden werkzaam was op het ID-DLO in Zeist. Daar heb ik me bezig gehouden met de ontwikkeling van een vleesindex voor melkvee. In November 1995 ben ik voor 9.5 maanden naar Massey University in Palmerston North, New Zealand vertrokken, waar ik aan verschillende projecten heb meegewerkt, oa aan het onderzoeken van de consequenties van het uitbesteden van opfok van melkvaarzen en aan de ontwikkeling van een selectieindex voor vleesvee. Van augustus tot en met oktober 1996 heb ik als fokkerijdeskundige bij het Dienstencentrum Schapen en Geitenhouderij in Lelystad gewerkt. Van november 1996 tot en met juli 1997 heb ik als toegevoegd onderzoeker bij de leerstoelgroep Fokkerij en Genetica gewerkt, oa aan een literatuuronderzoek naar de consequenties van mating in de fokkerij en aan een vergelijkend onderzoek naar de economische gevolgen van proefstierinzet door de verschillende fokkerijorganisaties. Op 1 augustus 1997 ben ik aan dit promotieonderzoek begonnen. Per 1 oktober 2001 ben ik als wetenschappelijk medewerker, deels bij de faculteit Diergeneeskunde in Utrecht en deels bij de leerstoelgroep Fokkerij en Genetica in Wageningen, betrokken bij de opzet van onderzoek naar genetische aspecten van diergezondheid en bij de ontwikkeling en uitvoering van genetica onderwijs voor diergeneeskunde studenten.