Indirect Genetic Effects for Group-housed Animals

Setegn Worku Alemu

Thesis committee

Promotor
Prof. Dr J.A.M. van Arendonk
Professor of Animal Breeding and Genetics
Wageningen University, The Netherlands

Co-promotors
Dr P. Bijma
Assistant professor, Animal Breeding and Genetics Group,
Wageningen University, The Netherlands

Dr P. Berg
Section Leader NordGen Farm Animals,
Nordic Genetic Resource Center, Ås, Norway

Main Supervisor at Aarhus University
Dr L.G. Janss
Senior researcher Center for Quantitative Genetics and Genomics
Aarhus University, Tjele, Denmark

Other members
Prof. Dr M. Naguib - Professor of Behavioural Ecology, Wageningen University
Prof. Dr E. Strandberg - Professor in Animal Breeding and Genetics, Swedish University of Agricultural Sciences, Uppsala, Sweden
Prof. Dr J. Jensen - Professor of Genetics and Genomics, Aarhus University, Tjele, Denmark
Dr G.A.A. Albers - Director Hendrix Genetics (is this the right title of Gerard?)

This research was conducted jointly under the auspices of the Graduate School of Science and Technology (GSST), Aarhus University and Wageningen Institute of Animal Sciences (WIAS), Wageningen University.

Indirect Genetic Effects for Group-housed Animals

Setegn Worku Alemu

Thesis

Submitted in fulfillment of the requirements for the joint degree of doctor between
Aarhus University
by the authority of Head of Graduate School
Prof.dr. Jes Madsen
and
Wageningen University
by the authority of the Rector Magnificus
Prof.dr. M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Academic Board at Wageningen University
and the Head of The Graduate School of Science and Technology at Aarhus University
to be defended in public
on Wednesday June 3, 2015
at 11 a.m. in the Aula, Wageningen University

Alemu, S.W

Indirect Genetic Effects for Group-housed Animals,

210 pages

Joint PhD thesis, Aarhus University, Aarhus, DK and

Wageningen University, Wageningen, NL (2015)

With references, with summaries in English, Danish and Dutch

ISBN 978-87-93176-71-3

ABSTRACT

Alemu, SW(2015) Indirect Genetic effects for Group-housed Animals. Joint PhD thesis between Aarhus University, Denmark and Wageningen University, the Netherlands**.**

Social interactions among individuals are common both in plants and animals. With social interactions, the trait value of an individual may be influenced by the genes of its interacting partners, a phenomenon known as indirect genetic effects (IGE). An IGE is heritable effect of an individual on trait values of another individual. A large body of literature has shown that social interactions can create addition heritable variation in both plants and animals, for both behavioural and production traits.

When IGE are estimated it is usually assumed that an individual interacts equally with all its group mates, irrespective of genetic relatedness. This assumption may not be true in mixed groups of kin and non-kin, where an individual may interact systematically different with kin and non-kin. Current IGE models ignore such systematically different interactions between kin and non-kin. Thus, the main aim of this thesis was to develop and apply statistical methods to estimate IGE when interactions differ between kin and non-kin.

Social interactions are important in mink that are kept in groups for the production of fur. Group housing of mink increases aggression behaviours, which is reflected by an increase in the number of bite marks on the pelts, and reduces the welfare of the animals. We estimated the genetic parameter for bite mark traits in group-housed mink, to investigate the prospects for genetic improvement of bite mark traits. We found that there are good prospects to produce mink that have a low level of biting. Finally, we further concluded that genetic parameter estimation for bite mark score should take into account systematic interactions due to sex or kin.

In this thesis we also investigated genomic selection for socially affected traits, considering survival time in two lines of brown egg layers showing cannibalistic behaviour. Despite the limited reference population of ~234 progeny tested sires, the accuracy of estimated breeding values (EBV) was ~35% higher for genomic selection compared with the parent average-EBV. We found that the response to genomic selection per year for line B1 was substantially higher than for the traditional breeding scheme, whereas for line BD response was slightly higher than for the traditional breeding scheme. In conclusion, genetic selection with IGE combined with marker information can substantially reduce detrimental social behaviours such as cannibalism in layers and biting in group-housed mink.

Dedicated to

My late father Worku Alemu Dilie who raised me with love and immensely helped me in my education.

My late sister kibret Mihretie Tegnegn her words of inspiration and encouragement never cease.

Contents

1

General introduction

General introduction

1.1 Social interaction

Social interactions among individuals are a common phenomenon in both domestic and wild populations, and in plants and animals (Frank, 2007). Social interaction can be either cooperative or competitive. Cooperative interactions originate from cooperative behaviour among individuals, for example, cooperative hunting in African wild dogs (Clutton-Brock, 2009), or cooperative behaviour between mother and offspring (Willham, 1963). Cooperative social interactions have a positive effect on production and welfare of livestock.

Competitive social interactions originate from competition for limited resources such as feed or space. It can also originate from aggressive behaviours such as social dominance. Competitive social interactions have a negative effect on productivity and welfare of livestock. For example, mixing of unfamiliar pig is common practise in pig production systems (Tan *et al.*, 1991). After mixing there will be intensive fighting and this aggressive behaviour results skin lesion (Stookey and Gonyou, 1998; Turner *et al.*, 2010). Also in group-housed layers, there are social interactions, and these interactions lead to feather pecking and mortality (Muir, 1996; Blokhuis and Wiepkema, 1998). The same is true for group-housed mink. Group housing of mink increases social interactions and these interactions result in more bite marks (Pedersen and Jeppesen, 2001; Moller *et al.*, 2003; Hansen *et al.*, 2014). Competitive social interactions can also affect growth rate and feed intake in pigs (Arango *et al.*, 2005; Chen *et al.*, 2008), and growth rate in forest trees (Cappa and Cantet, 2008; Brotherstone *et al.*, 2011; Silva *et al.*, 2013). Other than maternal genetic effect models, traditional genetic models ignore these social interactions and may result in suboptimal response to selection. Therefore, social interactions are an

important factor when designing artificial breeding programmes in domestic animals for which group housing is common practise (Muir, 2005).

Because of social interactions, the trait value of the individual may be influenced by the interacting partners' phenotype. The effect of the interacting partners' phenotype on the focal individual may be heritable and this heritable effect is termed Indirect Genetic Effect (Moore *et al.*, 1997; Wolf *et al.*, 1998), social effect or associative effect (Griffing, 1967; Muir, 2005; Bijma *et al.*, 2007a). A well-known example is the maternal genetic effect of a mother on pre-weaning growth rate of her offspring (Willham, 1963).

Apart from maternal genetic effects, the traditional quantitative genetic model ignores IGE created by social interactions. Griffing (1967) showed theoretically that ignoring indirect genetic effects will result in suboptimal response to selection, or even negative response to selection. This theoretical prediction was later proven empirically, using selection experiments both in laboratory populations and domestic populations (Goodnight, 1985; Craig and Muir, 1996). For example, individual selection to increase (decrease) leaf area of *Arabidopsis thaliana* decreased (increased) leaf area, and individual selection to decrease mortality in Japanese quail increased the mortality, though the increase was non-significant (Muir, 2005). Thus, inclusion of IGE in quantitative genetic models is essential to get an optimal response to artificial selection for socially affected traits (Bijma *et al.*, 2007b).

Group selection is selection among groups. In group selection, either the entire group is selected or culled based on the performance of the group (Muir, 2005). Group selection is one way of including IGE in breeding programmes. Group selection is effective compared with individual selection when the trait is influenced by IGEs. Group selection was effective compared with individual mass selection in decreasing mortality of laying hens,

mainly from aggression, from 68 % in generation 2 to 9 % in generation 6 (Muir, 1996). Leaf area of *Arabidopsis thaliana* responded positively for both low leaf and high leaf area with group selection, but responded negatively with individual selection (Goodnight, 1985). Group selection was also effective in improving longevity of layers (Craig and Muir, 1996). The reason for the effectiveness of group selection is that it accounts for part of the IGEs (Griffing, 1976a).

Though group selection was effective compared with individual selection, it has two important limitations. First, group selection is not efficient especially when the relationship in the group is lower. For example, when the group consists of unrelated individual, group selection utilizes only part of the total genetic variance since genetic variation between unrelated groups is small. On the other hand when the group consists of cloned individuals, for example, group selection fully takes into account both DGEs and IGEs, resulting in high accuracy (Griffing, 1976a; Griffing, 1976b; Bijma, 2011). However, though relatedness increases the efficiency of group selection, it also increases the rate of inbreeding. Second, though group selection can be used to generate response to selection, it does not explain the relative importance of direct *vs*. indirect genetic effects. It is vital to understand the genetic parameters underlying the interaction. This allows us to quantify the potential contribution of IGEs to response to selection, to estimate breeding values for both direct and indirect genetic effects, and to optimize breeding programmes (Bijma *et al.*, 2007a). To achieve optimal response to selection an index that weigh both the direct and indirect genetic effect optimally is required (Griffing, 1977). This can be achieved by a BLUP (best linear unbiased prediction) model that separates the direct breeding value weighted by 1 and indirect genetic effects weighted by *n-1*. (Muir, 2005; Bijma *et al.*, 2007a; Bijma *et al.*, 2007a).

1.2 Quantitative genetic model

In traditional quantitative genetics, the phenotypic value of individual i is the sum of a heritable component, A_i, known as breeding value, and a non-heritable residual, E_i (*Falconer and Mackay, 1996)*

$$P_i = A_i + E_i \tag{1}$$

With the presence of IGEs, the observed phenotype of an individual originates from two unobserved effects: A direct effect originating from the individual itself, and the sum of indirect effects originating from each of its $n-1$ group mates (Griffing, 1967),

$$P_i = A_{D,i} + E_{D,i} + \sum_{j=1}^{n-1} A_{S,j} + \sum_{j=1}^{n-1} E_{S,j} \ , \tag{2}$$

where i denotes the focal individual, j a group mate, $A_{D,i}$ the direct genetic effect (DGE) of i, $E_{D,i}$ the corresponding non-heritable direct effect, $A_{S,j}$ the IGE of group mate j, and $E_{S,j}$ the corresponding non-heritable indirect effect. Equation 2 contains two genetic effects, direct effects, A_D, and indirect effects, A_S.

The phenotypic variance is given as (Bergsma *et al.*, 2008)

$$\sigma_P^2 = \sigma_{A_D}^2 + \sigma_{E_D}^2 + (n-1)\left(\sigma_{A_S}^2 + \sigma_{ES}^2\right) + (n-1)r\left(2\sigma_{A_{Ds}} + (n-2)\sigma_{A_S}^2\right) \tag{3}$$

where $\sigma^2_{A_D}$ refers to the direct genetic variance, $\sigma^2_{A_D}$ to the indirect genetic variance , $\sigma_{A_{DS}}$ to the covariance between direct genetic effects and indirect genetic effects, and *r* refers to the average genetic relatedness in a group.

The total breeding value of an individual (A_T) is the heritable impact of an individual on the population mean and is given as:

$$A_{Ti} = A_{Di} + (n-1)A_{Si} \quad (4a)$$

The total breeding value entirely originates from the focal individual *i,* it is purely the focal individual's heritable effect on the population mean. However, the phenotype of an individual is the direct genetic effect originating from the focal individual *i* and the social genetic effect originating from group mates *j* (Equation 2).

The variance of total breeding values is the potential heritable variation available for response to selection

$$\sigma^2_{A_T} = \sigma^2_{A_D} + 2(n-1)\sigma_{ADS} + (n-1)^2\sigma^2_{A_S} \quad (4b)$$

Expressing total heritable variation as a proportion of phenotypic variance, $T^2 = \frac{\sigma^2_{A_T}}{\sigma^2_P}$,

which is an analogy of classical heritability, $h^2 = \frac{\sigma^2_A}{\sigma^2_P}$, helps to judge the contribution of social effects to heritable variance. Note that T^2 is not a classical heritability, *i.e.*, not the regression coefficient of breeding value on phenotype; it just represents total heritable variance among individuals expressed on the scale of phenotypic variance among individuals.

Using the model in Equation 2, there is a large body of literature that shows that IGE may contribute a substantial amount of heritable variation (Griffing, 1967; Muir, 2005; Bijma *et al.*, 2007a; Bergsma *et al.*, 2008; Ellen *et al.*, 2008; Wilson *et al.*, 2009; Wilson *et al.*, 2011; Silva *et al.*, 2013). For survival days in layer chicken, Bijma *et al.*(2007a) showed that 67% of the heritable variation was due to interactions among individuals. Alemu *et al.* (2014b) found that heritable variation from social interaction for total bite mark trait in mink accounts for about 85% of the total heritable variation. Wilson *et al.* (2009) found that more than 80% of heritable variation for aggressive behavioural traits such as rearing rate and reciprocal latency to fight in deer mice was due to interactions among individuals. Bergsma *et al.* (2013) found that more than 30% of heritable variation for growth rate and feed intake in a population of domestic pigs originated from social interaction. Therefore, for socially affected traits, IGE can be a significant source of heritable variation in both animals and plants, for both behavioural traits and production traits (Moore *et al.*, 1997; Muir, 2005; Bijma *et al.*, 2007b; Wilson *et al.*, 2011).

The estimation of genetic parameters in the above mentioned empirical studies relies on an important assumption. The assumption is that an individual express the same IGE on each of its social partners, irrespective of whether this partner is its relative or a non-relative. This assumption is at odds with kin selection theory (Hamilton, 1964). Kin selection theory predicts that individuals behave differently towards relatives *vs.* non-relatives (Hamilton, 1964). Hence, IGEs expressed on relatives may differ systematically from those expressed on non-relatives.

Empirical evidence indeed suggests that individuals show different behaviours towards relatives *vs.* non-relatives. Social insects, for example, such as honey bees, sweat bees, and some ants can recognise their kin and selectively care for related individuals (Hepper, 1986). Blanding's ground squirrel and Richardson's ground squirrel are less aggressive to

their relatives than to unrelated individuals (Sheppard and Yoshida, 1971; Holmes and Sherman, 1982). Kin recognition also occurs in fish (Olsen, 1989; Brown and Brown, 1996; Olsen *et al.*, 1998). When fish are reared in kin groups, individuals weigh more and differ less in size compared with individuals reared in non-kin groups. Aggressive interactions are also more common among strangers than in kin groups (Brown and Brown, 1993; Brown and Brown, 1996; Brown *et al.*, 1996). Also plants can recognize other plants in their surroundings based on relatedness and identity (Biedrzycki and Bais, 2010). The annual plant *Cakile edentula*, for example, generates more root when grown with strangers than when grown with family members (Dudley and File, 2007). Kin recognition and preferential behaviour towards kin, therefore, appear to be wide-spread in both plants and animals.

The current model ignores this systematic interaction due to kin recognition, which may result in biased estimates of genetic parameters and suboptimal response to selection. Thus, this thesis fills this gap by developing statistical methods to estimate IGE when interactions differ between kin and non-kin. The model is important, for instance, for pigs, fish, and group-housed mink.

1.3 Group housing of mink

Naturally, mink is a solitary and territorial species. Traditionally mink are housed in a pair of full sibs, one male and one female. This type of housing limits the aggressive interactions among cage mates. However, this housing has some limitations, such as limited space for the individuals, which makes the mink stressful. Group housing of mink was recommended by the Council of Europe (European Commission, 1999). This is because it may improve welfare from 'social enrichment' as outlined in (European

Commission, 2001), and it increases the stocking density in the cages and thereby decreases housing investments. Though group housing offers some advantages, it has still some limitations such as increased competition for food and aggressive behaviour (Pedersen and Jeppesen, 2001; Moller *et al.*, 2003).

The aggressive behaviour in group-housed mink is higher than in pair-wise housing and it is reflected by increased bite marks and bite wounds (Hansen and Damgaard, 1991; Pedersen and Jeppesen, 2001; Moller *et al.*, 2003).

Thus, for continuity of group housing, the welfare of mink should be improved. For example, mink in group housing should have a lower level of biting. One solution to improve the welfare in group-housed mink is to improve the management by the use of environmental enrichment, such as plastic tubes (Hansen, 2012). However, this still does not improve the welfare sufficiently (Hansen, 2012).

Another promising solution is genetic selection. Genetic selection can reduce bite marks in group-housed mink. Thus, this thesis addresses genetic selection with indirect genetic effect models, and investigates whether this can be a solution to reduce aggression behaviour in group-housed mink.

Bite mark traits in mink are recorded after life, after the mink are killed and pelleted. Thus, for breeding against bite marks, we will not have own performance records. We, therefore, have to use sib or progeny information. If we use sib information, the accuracy of estimating the breeding value is lower and the rate of inbreeding is higher, and if we use progeny testing the generation interval is increased. Consequently, using sib or progeny information may yield limited response to selection. We hypothesize that genomic selection can offers a solution for this problem.

Genomic selection is a type of marker assisted selection in which genetic markers covering the entire genome are used so that all quantitative trait loci are in linkage disequilibrium at least with one single nucleotide polymorphism (Meuwissen *et al.*, 2001). Genomic selection estimates the breeding value more accurately than pedigree- BLUP. Genomic selection is therefore particularly promising for low heritable traits and for traits that are difficult to record or recorded later in life for example, such as carcase quality, bite marks, and survival time (Calus *et al.*, 2008). Thus, we planned to test whether genomic selection increases the response to selection compared with traditional selection for socially affected traits. Due to unavailability of genomic data for mink, we focussed on survival time in brown layers showing cannibalism. Thus, we investigated whether genomic selection can increase the accuracy of estimating breeding values and increase the response to selection compared with pedigree BLUP for survival time in layers.

1.4 Outline of the thesis

The primary aim of this thesis was to develop statistical methods to estimate IGE when interactions differ systematically between kin and strangers. This thesis also investigates IGE for bite marks in group-housed mink, and the prospects of genomic selection for socially-affected traits

Chapter two develops statistical methods to estimate IGE when interactions differ systematically between kin and non-kin. There is a lot of empirical evidence that individuals interact differently with kin and non-kin. The current IGE model ignores such systematic differences in interaction between kin and non-kin. Thus, Chapter two develops statistical methods that estimate IGE when interactions differ between kin and non-kin. I further investigate the bias in the estimated genetic parameters when IGEs differ between

kin and non-kin while this is ignored in the statistical analysis (using the traditional IGE model).

Chapter three investigates IGE for bite mark trait in group-housed mink. Group housing of mink increases aggressive behaviour and this behaviour leads to an increased number of bite marks. Chapter three shows that genetic selection with IGE can reduce bite marks and possibly aggressive behaviour. In the estimation method the non-genetic social environment was accounted for by fitting cage*sex effect. Cage*sex effects of the same sex was equal and different sex was zero.

In Chapter four I studied the non-genetic systematic interactions in more detail. The non-genetic social interaction was accounted by fitting a cage random effect plus a cage*sex random interaction effect. The cage*sex random effect had a separate variance for males and for females. Accounting non-genetic social environment using cage plus cage*sex improved the fit of the model. Finally, it is concluded from Chapter four that ignoring the systematic interactions due to sex or kin results biased estimates of all the genetic parameters.

Bite mark traits in mink are recorded after the life of the animal. This will increase the generation interval and/or reduce the accuracy of estimating the breeding values (see above). With genomic information we can estimate the breeding value of an individual immediately after birth. Due to unavailability of genomic data for mink, we studied a similar trait which is survival time in a population of brown layers showing cannibalistic behaviour. Thus in Chapter five we investigated genomic selection for survival time in two lines of brown layers. The accuracy of estimating the breeding value was higher for B1 and BD line when we used genomic information than the pedigree-BLUP, and the predicted

response to selection was higher for the B1 line when we used genomic information than progeny testing.

Chapter six is the general discussion. I discuss five main topics that put the thesis is a broader perspective. These are: kin recognition mechanisms and consequences for the estimation of genetic parameters when we ignore kin recognition, genome-wide marker to estimate all genetic parameter when IGE differ between kin and non-kin, trait-based models when IGE differ between kin and non-kin, accuracy of estimating the breeding value using individual selection when IGE differ between kin and non-kin, and prospects of reducing the number of bite mark in mink using genetic selection.

References

Alemu SW, Bijma P, Moller SH, Janss L, Berg P (2014b). Indirect genetic effects contribute substantially to heritable variation in aggression-related traits in group-housed mink (Neovison vison). *Genet Sel Evol* **46:30**.

Arango J, Misztal I, Tsuruta S, Culbertson M, Herring W (2005). Estimation of variance components including competitive effects of Large White growing gilts. *J Anim Sci* **83:** 1241-1246.

Bergsma R, Kanis E, Knol EF, Bijma P (2008). The contribution of social effects to heritable variation in finishing traits of domestic pigs (*Sus scrofa)*. *Genetics* **178:** 1559-1570.

Bergsma R, Mathur PK, Kanis E, Verstegen MWA, Knol EF, Van Arendonk JAM (2013). Genetic correlations between lactation performance and growing-finishing traits in pigs. *J Anim Sci* **91:** 3601-3611.

Biedrzycki ML, Bais HP (2010). Kin recognition in plants: a mysterious behaviour unsolved. *J Exp Bot* **61:** 4123-4128.

Bijma P (2011). A general definition of the heritable variation that determines the potential of a population to respond to selection. *Genetics* **189:** 1347-1359.

Bijma P, Muir WM, Ellen ED, Wolf JB, Van Arendonk JAM (2007a). Multilevel selection 2: Estimating the genetic parameters determining inheritance and response to selection. *Genetics* **175:** 289-299.

Bijma P, Muir WM, Van Arendonk JAM (2007b). Multilevel selection 1: quantitative genetics of inheritance and response to selection. *Genetics* **175:** 277-288.

Blokhuis HJ, Wiepkema PR (1998). Studies of feather pecking in poultry. *Vet Quart* **20:** 6-9.

Brotherstone S, White IMS, Sykes R, Thompson R, Connolly T, Lee S *et al.* (2011). Competition Effects in a Young Sitka Spruce (Picea sitchensis, Bong. Carr) Clonal Trial. *Silvae Genet* **60:** 149-155.

Brown GE, Brown JA (1993). Social Dynamics in Salmonid Fishes - Do Kin Make Better Neighbors. *Anim Behav* **45:** 863-871.

Brown GE, Brown JA (1996). Does kin-biased territorial behavior increase

kin-biased foraging in juvenile salmonids? *Behav Ecol* **7:** 24-29.

Brown GE, Brown JA, Wilson WR (1996). The effects of kinship on the growth of juvenile Arctic charr. *J Fish Biol* **48:** 313-320.

Calus MPL, Meuwissen THE, de Roos APW, Veerkamp RF (2008). Accuracy of genomic selection using different methods to define haplotypes. *Genetics* **178:** 553-561.

Cappa EP, Cantet RJC (2008). Direct and competition additive effects in tree breeding: Bayesian estimation from an individual tree mixed model. *Silvae Genet* **57:** 45-56.

Chen CY, Kachman SD, Johnson RK, Newman S, Van Vleck LD (2008). Estimation of genetic parameters for average daily gain using models with competition effects. *J Anim Sci* **86:** 2525-2530.

Clutton-Brock T (2009). Cooperation between non-kin in animal societies. *Nature* **462:** 51-57.

Craig JV, Muir WM (1996). Group selection for adaptation to multiple-hen cages: Beak-related mortality, feathering, and body weight responses. *Poultry Science* **75:** 294-302.

Dudley SA, File AL (2007). Kin recognition in an annual plant. *Biology Lett* **3:** 435-438.

Ellen ED, Visscher J, van Arendonk JAM, Bijma P (2008). Survival of laying hens: Genetic parameters for direct and associative effects in three purebred layer lines. *Poultry Science* **87:** 233-239.

European Commission (1999). Standing committee of the European convention for the protection of animals kept for farming purposes: Recommendation concerning fur animals. European Commission; 1999. http://.www.coe.int/.

European Commission (2001). Scientific Committee on Animal Health and Animal Welfare: The welfare of animals kept for furproduction.European Commission

.http://ec.europa.eu/food/animal/welfare/international/out67_en.pdf

Frank SA (2007). All of life is social. *Current Biology* **17:** R648-R650.

Goodnight CJ (1985). The influence of environmental variation on group and individual selection in a cress. *Evolution* **39:** 545-558.

Griffing B (1967). Selection in reference to biological groups .I. Individual and group selection applied to populations of unordered groups. *Australian Journal of Biological Sciences* **20:** 127-139.

Griffing B (1976a). Selection in reference to biological groups. V. Analysis of full-sib groups. *Genetics* **82:** 703-722.

Griffing B (1976b). Selection in reference to biological groups. VI. Use of extreme forms of nonrandom groups to increase selection efficiency. *Genetics* **82:** 723-731.

Griffing, B., 1977 Selection for populations of interacting genotypes. In: Proc Int Conf Quant Genet pp. 413–434. Iowa State University Press: Ames IA., pp

Hamilton WD (1964). The genetical evolution of social behavior. *J Theor Biol* **7:** 1–16.

Hansen S (2012). Plastrør og halmbriketter reducerer pelsgnav. (Ed Steffen W.hansen og birthe M Damgaard) Temadag om aktuel Minkforsking - DCA Rapport nr. 010, spetember 2012. Aarhus university.p. 13-19.

Hansen SW, Damgaard BM (1991). Stress Physiological, Hematological and Clinical-Chemical Status of Farm Mink Placed in Groups or Singly. *Acta Agr Scand* **41:** 355-366.

Hansen SW, Moller SH, Damgaard BM (2014). Bite marks in mink-Induced experimentally and as reflection of aggressive encounters between mink. *Appl Anim Behav Sci* **158:** 76-85.

Hepper PG (1986). Kin recognition: functions and mechanisms. A review. *Biol Rev Camb Philos Soc* **61:** 63-93.

Holmes WG, Sherman PW (1982). The Ontogeny of Kin Recognition in 2 Species of Ground-Squirrels. *Am Zool* **22:** 491-517.

Meuwissen THE, Hayes BJ, Goddard ME (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157:** 1819-1829.

Moller SH, Hansen SW, Sorensen JT (2003). Assessing animal welfare in a strictly synchronous production system: The mink case. *Animal Welfare* **12:** 699-703.

Moore AJ, Brodie ED, Wolf JB (1997). Interacting phenotypes and the evolutionary process .1. Direct and indirect genetic effects of social interactions. *Evolution* **51:** 1352-1362.

Muir WM (1996). Group selection for adaptation to multiple-hen cages: selection program and direct responses. *Poultry Science* **75:** 447-458.

Muir WM (2005). Incorporation of competitive effects in forest tree or animal breeding programs. *Genetics* **170:** 1247-1259.

Olsen KH (1989). Sibling Recognition in Juvenile Arctic Charr, Salvelinus-Alpinus (L). *J Fish Biol* **34:** 571-581.

Olsen KH, Grahn M, Lohm J, Langefors A (1998). MHC and kin discrimination in juvenile Arctic charr, Salvelinus alpinus (L.). *Anim Behav* **56:** 319-327.

Pedersen V, Jeppesen LL (2001). Effects of family housing on behaviour, plasma cortisol and performance in adult female mink (Mustela vison). *Acta Agriculturae Scandinavica Section a-Animal Science* **51:** 77-88.

Sheppard DH, Yoshida SM (1971). Social behavior in captive Richardson´s ground squirrels. *J Mammal* **52:** 793-799.

Silva JCE, Potts BM, Bijma P, Kerr RJ, Pilbeam DJ (2013). Genetic control of interactions among individuals: contrasting outcomes of indirect genetic effects arising from neighbour disease infection and competition in a forest tree. *New Phytol* **197:** 631-641.

Stookey JM, Gonyou HW (1998). Recognition in swine: recognition through familiarity or genetic relatedness? *Appl Anim Behav Sci* **55:** 291-305.

Tan SSL, Shackleton DM, Beames RM (1991). The Effect of Mixing Unfamiliar Individuals on the Growth and Production of Finishing Pigs. *Anim Prod* **52:** 201-206.

Turner SP, D'Eath RB, Roehe R, Lawrence AB (2010). Selection against aggressiveness in pigs at re-grouping: practical application and implications for long-term behavioural patterns. *Animal Welfare* **19:** 123-132.

Willham RL (1963). Covariance between Relatives for Characters Composed of Components Contributed by Related Individuals. *Biometrics* **19:** 18-27.

Wilson AJ, Gelin U, Perron MC, Reale D (2009). Indirect genetic effects and the evolution of aggression in a vertebrate system. *P Roy Soc B-Biol Sci* **276:** 533-541.

Wilson AJ, Morrissey MB, Adams MJ, Walling CA, Guinness FE, Pemberton JM *et al.* (2011). Indirect genetics effects and evolutionary constraint: an analysis of social dominance in red deer, *Cervus elaphus*. *J Evol Biol* **24:** 772-783.

Wolf JB, Brodie ED, Cheverud JM, Moore AJ, Wade MJ (1998). Evolutionary consequences of indirect genetic effects. *Trends Ecol Evol* **13:** 64-69.

2

Indirect genetic effects and kin recognition: estimating IGEs when interactions differ between kin and strangers

S.W. Alemu[1,3], P. Berg[1,2], Luc Janss[1], and P. Bijma[3]

[1]Department of Molecular Biology and Genetics, Aarhus University, Tjele DK-8830, Denmark, [2]NordGen, Nordic Genetic Resource Center, P.O.Box 115, 1431 Ås, Norway, [3]Animal Breeding and Genomics Centre, Wageningen University, 6700AH, Wageningen, The Netherlands

Heredity (2014) **112,** 197–206

ABSTRACT

Social interactions among individuals are wide-spread, both in natural and domestic populations. As a result, trait values of individuals may be affected by genes in other individuals, a phenomenon known as Indirect Genetic Effects (IGEs). IGEs can be estimated using linear mixed models. The traditional IGE-model assumes that an individual interacts equally with all its partners, whether kin or strangers. There is abundant evidence, however, that individuals behave differently towards kin compared to strangers, which agrees with predictions from kin-selection theory. With a mix of kin and strangers, therefore, IGEs estimated from a traditional model may be incorrect, and selection based on those estimates will be suboptimal. Here we investigate whether genetic parameters for IGEs are statistically identifiable in group-structured populations when IGEs differ between kin and strangers, and develop models to estimate such parameters. First, we extend the definition of total breeding value and total heritable variance to cases where IGEs depend on relatedness. Next, we show that the full set of genetic parameters is not identifiable when IGEs differ between kin vs. strangers. Subsequently, we present a reduced model that yields estimates of the total heritable effects on kin, on non-kin, and on all social partners of an individual, as well as the total heritable variance for response to selection. Finally we discuss the consequences of analysing data in which IGE depend on relatedness using a traditional IGE-model, and investigate group structures that may allow estimation of the full set of genetic parameters when IGEs depend on kin.

Key words: Social interactions, indirect genetic effects, kin recognition, kin, stranger

2.1 INTRODUCTION

Social interactions among individuals are common, both in wild and domestic populations, and in animals, plants and microorganisms (Frank, 2007). With social interactions, the trait value of an individual may be affected by genes in other individuals, a phenomenon that works out as Indirect Genetic Effects (IGEs; Griffing, 1967; Griffing, 1976; Moore *et al.*, 1997; Wolf *et al.*, 1998). An IGE is a heritable effect of one individual on the trait value of another individual (reviewed in Wolf *et al.*, 1998; Bijma, 2011a). A well-known example is the maternal genetic effect of a mother on pre-weaning growth rate of her offspring (Willham, 1963; Falconer, 1965; Kirkpatrick and Lande, 1989).

IGEs may have significant effects on the rate and direction of response to selection, and can substantially increase or decrease heritable variation in a trait (Griffing, 1967; Moore *et al.*, 1997; Bijma and Wade, 2008; McGlothlin and Brodie III, 2009; Bijma, 2011b; Wilson *et al.*, 2011). Thus, knowledge of IGEs is essential for understanding response to selection in socially affected traits. The magnitude of IGEs can be estimated using linear mixed models that include a direct genetic effect for the individual producing the record, and an IGE for each of its social partners (Arango *et al.*, 2005; Muir, 2005; Bijma *et al.*, 2007b). This approach has been used both in agricultural populations of animals and plants (*e.g.*, Muir 2005; Silva *et al.*, 2013), and in natural populations (*e.g.,* Wilson *et al.*, 2011). Bijma (2010a) showed that estimation of genetic parameters for IGEs in group-structured populations can be optimized by placing two families in each group. Such schemes are an attractive breeding design, because they also yield a relatively high response to selection (Odegard and Olesen, 2011).

In the linear mixed model commonly used to estimate IGEs (Muir, 2005), it is assumed that an individual expresses the same IGE on each of its social partners, irrespective of

whether a partner is its family member or an unrelated individual. Kin selection theory, however, predicts that individuals behave more cooperatively towards their relatives, because this increases their inclusive fitness (Hamilton, 1964). Hence, IGEs expressed on kin may differ systematically from those expressed on strangers; they may differ not only in average level, but also show incomplete correlation. Empirical evidence indeed suggests that kin recognition and preferential behaviour towards kin are wide-spread in both animals and plants (e.g. Holmes and Sherman, 1982; Hepper, 1986; Olsen, 1989; Dudley and File, 2007; Biedrzycki and Bais, 2010), and at least four mechanisms for kin recognition have been described (Tang-Martinez, 2001; Mateo, 2004; Mateo and Holmes, 2004; Coffin *et al.*, 2011).

When individuals express a different IGE on kin versus strangers, estimated breeding values for direct and indirect effects from the common linear mixed model are incorrect, and selection based on those estimates will yield suboptimal response. Moreover, when IGEs are estimated from groups composed of strangers (*e.g.*, Ellen *et al.* 2008), the resulting estimates may not accurately reflect the IGEs that occur in the relevant natural or domestic populations, which may consist of kin groups. In natural populations, limited dispersal often leads to interactions among relatives (Hamilton, 1964), while in livestock populations such as domestic pigs, groups often contain a number of family members (Chen *et al.*, 2008). Thus, a potential difference between IGEs on kin *vs.* strangers is relevant for both livestock and natural populations. The current statistical methods for estimating IGEs, however, ignore the dependency of IGEs on relatedness.

Here we propose a model for traits affected by IGEs that differ between kin and strangers, investigate whether genetic parameters of that model are statistically identifiable, and develop statistical models to estimate those parameters. First we show that the full set of

genetic parameters is not identifiable when IGEs differ between kin *vs.* strangers. Subsequently, we developed a reduced model, and show that the reduced model can estimate meaningful linear combinations of the genetic parameters. In the Discussion, we consider population structures that may allow estimating the full set of genetic parameters.

2.2 QUANTITATIVE GENETIC MODEL

2.2.1 Trait Model

This section introduces the trait-model when IGEs differ between kin *vs.* strangers. We consider here a population stratified in to groups of n members each, where interactions occur within groups. We consider the scheme that is optimal for the estimation of IGEs in the absence of kin recognition (Bijma, 2010a). In this scheme, each group is composed of members of two families, each family contributing $n/2$ individuals. Generalisation of results to other group structures is addressed in the Discussion.

In traditional quantitative genetics, the phenotypic value of individual i is the sum of a heritable component, A_i, known as breeding value, and a non-heritable residual, E_i (Falconer and Mackay, 1996; see Table 1 for a notation key),

$$P_i = A_i + E_i \,. \qquad (1)$$

With IGEs that do not depend on relatedness, the phenotype of an individual stems from two components: a direct effect originating from the individual itself, and the sum of indirect effects originating from each of its $n-1$ group mates (Griffing, 1967),

$$P_i = A_{D,i} + E_{D,i} + \sum_{j=1}^{n-1} A_{S,j} + \sum_{j=1}^{n-1} E_{S,j}\,, \tag{2}$$

where *i* denotes the focal individual, *j* a group mate, $A_{D,i}$ the direct genetic effect (DGE) of *i*, $E_{D,i}$ the corresponding non-heritable direct effect, $A_{S,j}$ the IGE of group mate *j*, and $E_{S,j}$ the corresponding non-heritable indirect effect (subscript *S*, suggesting "social", is used to denote indirect effects instead of a subscript *I*, to avoid confusion of *i* with *I*; Equation 2 is known as a variance component model of IGEs, as opposed to a trait-based model. See McGlothlin and Brodie III, 2009 for a comparison of models). Equation 2 contains two kinds of genetic effects, direct effects, A_D, and indirect effects, A_S. Hence, fitting Equation 2 involves the estimation of three genetic variance components; $\sigma^2_{A_D}$, $\sigma_{A_{DS}}$ and $\sigma^2_{A_S}$. (Throughout, σ^2 denotes a variance and σ a covariance).

With different interactions among kin versus strangers, two types of IGEs may be distinguished: IGEs on kin versus IGEs on strangers. In our population structure, where *n/2* members of each family make up a group, the trait model becomes:

$$P_i = A_{D,i} + E_{D,i} + \sum_{j=1}^{\frac{n}{2}-1} A_{S_f,j} + \sum_{j=1}^{\frac{n}{2}-1} E_{S_f,j} + \sum_{k=1}^{n/2} A_{S_u,k} + \sum_{k=1}^{n/2} E_{S_u,k} \tag{3}$$

where *j* denotes a family member of *i*, *k* a member of the other family in the group, $\frac{n}{2}-1$ the number of group mates of *i* from its own family, *n*/2 the number of group mates of *i* from the other family, subscript "S_f" denotes IGEs on family members, and subscript "S_u" denotes IGEs on members of the other, unrelated, family (*u* indicating "unrelated").

Equation 3 contains three genetic effects, direct effects, A_D, IGEs on family members, A_{S_f}, and IGEs on strangers, A_{S_u}. Hence, fitting Equation 3 involves the estimation of six genetic variance components; three variances: $\sigma^2_{A_D}$, $\sigma^2_{A_{S_f}}$ and $\sigma^2_{A_{S_u}}$, and three covariances: $\sigma_{A_{D,S_f}}$, $\sigma_{A_{D,S_u}}$ and $\sigma_{A_{S_f},S_u}$. The genetic correlation between an individual's IGE on kin and its IGE on strangers, $r_{A_{S_f}A_{S_u}} = \sigma_{A_{S_f},A_{S_u}} /(\sigma_{A_{S_f}}\sigma_{A_{S_u}})$, reflects the difference between IGEs on kin *vs*. strangers. Equation 3 does not explicitly include a potential difference in the mean value of the IGE on kin *vs*. strangers, because this has little consequences for the estimation of genetic parameters. Nevertheless, such a difference is relevant in statistical data analysis, and can be accommodated easily in the fixed-effects part of the model (see Discussion).

Table 1 Notation Key[1]

Symbol	Meaning
i,j,k,x,l,m	Subscript to denote an individual.
P_i	Observed trait value of an individual.
$P_{D,i}, P_{S_{f,i}}, P_{S_{u,i}}$	Direct effect of *i*, indirect effect of i to kin, indirect effect of *i* to stranger.
$\sigma^2_{P,}\sigma^2_{P_D}, \sigma^2_{P_{S_f}}, \sigma^2_{P_{S_u}}$	Phenotype variance among individuals, unobserved phenotype variance on self, on kin, on strangers.
$\sigma^2_{A_D}, \sigma^2_{A_{S_f}}, \sigma^2_{A_{S_u}}$	Variance of DGEs among individuals, variance of IGEs on kin among individuals, variance of IGEs on strangers among individuals.
$\sigma^2_{A_{T_f}}, \sigma^2_{A_T}$	Variance of family breeding value among individuals, variance of total breeding value among individuals.
$\sigma^2_{E_D}, \sigma^2_{E_{S_f}}, \sigma^2_{E_{S_u}}$	Variance of direct environment among individuals, variance of indirect environment on kin among individual, variance of indirect environment on strangers among individual.
$\sigma_{A_{D,S_f}}, r_{A_{D,S_f}}, \sigma_{A_{D,S_u}}, r_{A_{D,S_u}}$	Covariance between DGEs and IGEs to kin, correlation between DGEs and IGEs to kin, covariance between DGEs and IGEs to strangers, correlation between DGEs and IGEs to strangers.
$\sigma_{A_{S_f},s_u}\ r_{A_{S_f},s_f}$	Covariance between IGEs to kin and IGEs to strangers, correlation between IGEs to kin and IGEs to strangers.
$\sigma_{E_{D,S_f}}, r_{E_{D,S_f}}$	Covariance and correlation between non-genetic direct and non-genetic indirect on kin.
$\sigma_{E_{D,S_u}}, r_{E_{D,S_u}}$	Covariance and correlation between non-genetic direct and non-genetic indirect on strangers.
$\sigma_{E_{S_f},s_u}, r_{E_{S_f},s_u}$	Covariance and correlation between non-genetic in direct on kin and non-genetic indirect on strangers.
r, ρ, n, σ^2_g	Relatedness among individual in a group, residual correlation of family member in a group, group size, variance of non-family member in a group.

Abbreviations: DGE, direct genetic effect; IGE, indirect genetic effect.

[a]Throughout, hats (^) denote estimates, whereas symbols without hats refer to true value.

2.2.1 Total breeding value and heritable variation

This section presents the heritable variation available for response to selection in a trait when IGEs differ between kin and strangers.

Irrespective of the trait model, response to selection in any trait can be expressed as

$$R = \Delta\bar{A}_T = \imath\rho\sigma_{A_T} \quad (4)$$

,

where R is the genetic change in mean trait level from one generation to the next due to selection, $\Delta\bar{A}_T$ the change in mean total breeding value (A_T) of the population, ι the intensity of selection, ρ the accuracy of selection, and σ_{A_T} the standard deviation in total breeding value (Bijma, 2011a); an equivalent expression in terms of a selection gradient can also be found there, and may be more appropriate for natural populations). In the context of Equation 4, the accuracy of selection is the correlation between an individual's value for the selection criterion and its total breeding value. (This definition applies to any selection criterion; see Falconer and Mackay 1996 for further explanation of the "accuracy of selection"). The total breeding value represents the average impact of an individual's genes on the mean trait value of the population, and is a generalization of the traditional breeding value to account for IGEs and to allow modelling of so-called emergent traits (Bijma, 2011b). Thus, analogous to the classical breeding value, the total breeding value represents an individual's value for response to selection. As illustrated in Equation 4, in which $\imath$ and ρ are standardized parameters, the standard deviation in total breeding value represents the intrinsic potential of a population to respond to selection.

For any trait model, the total breeding values follows from the genetic mean of the population (Bijma, 2011b). From Equation 3, the genetic mean of the trait value for our population structure equals

$\bar{P}_A = \overline{A_D} + (\tfrac{1}{2}n-1)\overline{A_{S_f}} + \tfrac{1}{2}n\,\overline{A_{S_u}}$.

Therefore, following Bijma (2011b), an individual's total breeding value is the sum of its DGE, ½n−1 times its IGE on family members, and ½n times its IGE on strangers,

$$A_{T,i} = A_{D,i} + (\tfrac{1}{2}n-1)A_{S_f,i} + \tfrac{1}{2}n\,A_{S_u,i} \qquad (5)$$

.

Taking the variance of the total breeding value yields an expression for the heritable variation available for response to selection,

$$\begin{aligned}\sigma^2_{A_T} \;=\; & \sigma^2_{A_D} \\ & + (n-2)\sigma_{A_D,S_f} \;+\; (\tfrac{1}{2}n-1)^2\sigma^2_{A_{S_f}} \\ & + n\sigma_{A_D,S_u} \;+\; n(\tfrac{1}{2}n-1)\sigma_{A_{S_f},S_u} \;+\; \tfrac{1}{4}n^2\sigma^2_{A_{S_u}}\,.\end{aligned} \qquad (6)$$

Note that $\sigma^2_{A_T}$ does not reflect the additive genetic component of phenotypic variance, but the heritable variation that determines the potential of a population to respond to selection (see Equation 4 and Bijma, 2011b).

An individual's total breeding value can be partitioned into a family component, A_{T_f} , which summarizes all its heritable effects on family members (including the direct effect on itself) and is considered the family breeding value here, and a non-family component, A_{T_u} , the non-family breeding value. This partitioning will be used below, where the family components of the total breeding value will be grouped for reasons of statistical identifiability. With each family contributing ½n group members

$$A_{T,i} = A_{T_f,i} + A_{T_u,i}, \tag{7a}$$

$$A_{T_f,i} = A_{D,i} + (\tfrac{1}{2}n - 1)A_{S_f,i}, \tag{7b}$$

$$A_{T_u,i} = \tfrac{1}{2}nA_{S_u,i}. \tag{7c}$$

Taking the variances of Equation 7 yields

$$\sigma^2_{A_T} = \sigma^2_{A_{T_f}} + 2\sigma_{A_{T_f T_u}} + \sigma^2_{A_{T_u}}, \tag{8a}$$

$$\sigma^2_{A_{T_f}} = \sigma^2_{A_D} + (n-2)\sigma_{A_{D,S_f}} + (\tfrac{1}{2}n-1)^2\sigma^2_{A_{S_f}}, \tag{8b}$$

$$\sigma^2_{A_{T_u}} = \tfrac{1}{4}n^2\sigma^2_{A_{S_u}} \tag{8c}$$

$$\sigma_{A_{T_f T_u}} = \tfrac{1}{2}n\sigma_{A_{D,S_u}} + \tfrac{1}{2}n\left(\tfrac{1}{2}n-1\right)\sigma_{A_{S_f,S_u}} \tag{8d}$$

2.2.2 VARIANCE COMPONENT ESTIMATION

Genetic parameters can be estimated using a linear mixed model including correlated random genetic effects, the so-called animal model (Henderson, 1953; Henderson, 1975; Lynch and Walsh, 1998). The classical animal model includes DGEs only, but can be extended with IGEs (Muir, 2005).

Full model

The full model includes DGEs, IGEs on family members, and IGEs on strangers,

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_D\mathbf{a}_D + \mathbf{Z}_{S_f}\mathbf{a}_{S_f} + \mathbf{Z}_{S_u}\mathbf{a}_{S_u} + \mathbf{Wg} + \mathbf{e} \quad (9)$$

where $\mathbf{b}$ is a vector of fixed effects with incidence matrix $\mathbf{X}$, $\mathbf{a}_D$ is a vector of DGEs with incidence matrix $\mathbf{Z}_D$ linking observation on individuals to their own DGE, $\mathbf{a}_{S_f}$ is vector of IGEs on family members with incidence matrix $\mathbf{Z}_{S_f}$ linking observations on individuals to the IGEs of their group mates belonging to the same family, and $\mathbf{a}_{S_u}$ is vector of IGEs on strangers with incidence matrix $\mathbf{Z}_{S_u}$ linking observations on individuals to the IGEs of their group mates belonging to the other family, $\mathbf{g}$ is a vector of random group effects, with $\mathbf{g} \sim N(\mathbf{0}, \mathbf{I}_g\sigma_g^2)$ and incidence matrix $\mathbf{W}$ linking records to groups, and $\mathbf{e}$ is a vector of residuals with $\mathbf{e} \sim N(\mathbf{0}, \mathbf{I}_e\sigma_e^2)$, where $\mathbf{I}$ is an identity matrix. The covariance structure of the genetic terms is

$$\begin{bmatrix} \mathbf{a}_D \\ \mathbf{a}_{S_f} \\ \mathbf{a}_{S_u} \end{bmatrix} \sim MVN(\mathbf{0},\ \mathbf{C} \otimes \mathbf{A}),$$

where $\mathbf{C} = \begin{bmatrix} \sigma^2_{A_D} & \sigma_{A_{D,S_f}} & \sigma_{A_{D,S_u}} \\ \sigma_{A_{D,S_f}} & \sigma^2_{A_{S_f}} & \sigma_{A_{S_f,S_u}} \\ \sigma_{A_{D,S_u}} & \sigma_{A_{S_f,S_u}} & \sigma^2_{A_{S_u}} \end{bmatrix}$,

$\otimes$ indicates the Kronecker product of matrices, and $\mathbf{A}$ is a matrix of additive genetic relationships between individuals, the so-called numerator relationship matrix (Henderson, 1985).

When fitting the full model, the results showed that there are multiple parameter combinations that give the same likelihood. Hence, using this model, the genetic parameters are statistically non-identifiable. In particular, results showed that the variance of IGEs on strangers, $\sigma^2_{A_{S_u}}$, is identifiable, but that the variance components referring to interactions between family members, $\sigma^2_{A_D}$, $\sigma_{A_D,A_{S_f}}$, and $\sigma^2_{A_{S_f}}$, are fully confounded. We investigated why this occurs and found that there are only five informative genetic covariances in the data, but six genetic parameters to estimate (Appendix A). Thus, when IGEs differ between kin *vs.* strangers, it is not possible to estimate all six genetic parameters from group-structured data. This is not a problem of the estimation method, but a property of the data structure and occurs when group composition with respect to family is the same for all groups (See Discussion and Appendix A).Thus the data structure that is optimal for estimating the variance of IGEs that do not depend on kin renders the estimation of kin-dependent IGEs impossible. In the Discussion, we consider alternative schemes that may allow estimating all parameters of the full model. Note that the variance structure given above for the residual of Equation 9 ignores the distinction between indirect effects on kin *vs.* strangers. However, as the full model is non-identifiable, we did not further investigate this issue.

2.2.2 Reduced Model

Because the full model was not identifiable, we investigated a reduced model, aiming to estimate part of the genetic parameters or meaningful linear combinations. Since the full model indicated that the effects due to the focal family were fully confounded, we fitted only a single term for the family of the focal individual. Therefore, the reduced model was

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_D\mathbf{a}_F + \mathbf{Z}_{S_u}\mathbf{a}_{S_u} + \mathbf{Wg} + \mathbf{e}\,, \qquad (10)$$

where $\mathbf{a}_F$ is a vector of genetic effects due to the family of the focal individual, and $\mathbf{Z}_D$ is the incidence matrix for direct genetic effects as in the full model (Equation 9). Hence, with respect to the genetic terms, the only difference between the full and reduced model is that the $\mathbf{Z}_{S_f}\mathbf{a}_{S_f}$ term is omitted in Equation 10; the other genetic terms are the same. However, as omitting the $\mathbf{Z}_{S_f}\mathbf{a}_{S_f}$ will change both the estimates and the interpretation of the "direct" genetic effects, we write $\mathbf{Z}_D\mathbf{a}_F$ in Equation 10, where subscript *F* suggests "family", rather than $\mathbf{Z}_D\mathbf{a}_D$. The covariance structure of the genetic terms in Equation 10 is

$$\begin{bmatrix} \mathbf{a}_F \\ \mathbf{a}_{S_u} \end{bmatrix} \sim MVN\left(\mathbf{0}, \mathbf{C}_r \otimes \mathbf{A}\right),$$

where $\mathbf{C}_r = \begin{bmatrix} \sigma^2_{A_F} & \sigma_{A_F,S_u} \\ \sigma_{A_F,S_u} & \sigma^2_{A_{S_u}} \end{bmatrix}$.

The **Wg** term is as in Equation 9. The covariance structure for the residual term is

$$\text{var}(\mathbf{e}) = \mathbf{R}\sigma^2_e, \tag{11}$$

where $\mathbf{R}_{ii} = 1$, $\mathbf{R}_{ij} = \rho$ when *i* and *j* are group mates from the same family, and $\mathbf{R}_{ij} = 0$ otherwise. Hence, this structure allows for a covariance between residuals of group mates belonging to the same family. Thus, when individuals are ordered by group and by family within group, then **R** is block-diagonal, with blocks of size $n/2$, diagonal elements equal to 1, off-diagonals of blocks equal to ρ, all other off-diagonals equal to zero, and two blocks per group, one for each family. Appendix B shows that this residual variance structure together with the random group effect corresponds to the non-genetic variance structure generated by the assumed true model (Equation 3). Thus, the $\mathbf{Wg} + \mathbf{e}$ in Equation 10

accounts for the variance structure generated by the term $E_{D,i} + \sum E_{S_{f,j}} + \sum E_{S_{u,k}}$ in Equation 3.

Investigation of Equation 10 showed that there are five informative genetic covariances in the data to estimate three genetic parameters, indicating that the model in Equation 10 is identifiable. To investigate the interpretation of the genetic estimates from the reduced model, we derived their expectation, assuming that the data is generated by the model given in Equation 3 (Appendix A). With

$$A_{F,i} = A_{D,i} + (\tfrac{1}{2}n - 1)A_{S_f,i}, \tag{12a}$$

and

$$A_{S_u,i} = A_{S_u,i}, \tag{12b}$$

it follows that

$$\mathrm{E}(\hat{\sigma}^2_{A_F}) = \sigma^2_{A_D} + (n-2)\sigma_{A_{D,S_f}} + (\tfrac{1}{2}n-1)^2\sigma^2_{A_{S_f}}, \tag{12c}$$

$$\mathrm{E}(\hat{\sigma}_{A_{F,S_u}}) = \sigma_{A_{D,S_u}} + (\tfrac{1}{2}n-1)\,\sigma_{A_{S_f,S_u}}, \tag{12d}$$

$$\mathrm{E}(\hat{\sigma}^2_{A_{S_u}}) = \sigma^2_{A_{S_u}}. \tag{12e}$$

Equation 12c-e sum up all the variance components considered in the true model. Equation 12e shows that the reduced model yields an estimate of the variance of IGEs on strangers. Moreover, combining Equations 12c-e with the decomposition of the total breeding value into a family and a non-family component given in Equations 7 & 8 above shows that the reduced model yields estimates of the family and non-family genetic parameters,

$$\hat{\sigma}^2_{A_{T_f}} = \hat{\sigma}^2_{A_F}, \tag{13a}$$

$$\hat{\sigma}_{A_{T_f} A_{T_u}} = \tfrac{1}{2} n \hat{\sigma}_{A_F, S_u} \tag{13b}$$

$$\hat{\sigma}^2_{A_{T_u}} = \tfrac{1}{4} n^2 \hat{\sigma}^2_{A_{S_u}} \ . \tag{13c}$$

Thus, the variance of the total breeding value can be obtained from the reduced model as

$$\hat{\sigma}^2_{A_T} = \hat{\sigma}^2_{A_F} + n \hat{\sigma}_{A_F, S_u} + \tfrac{1}{4} n^2 \hat{\sigma}^2_{A_{S_u}} \tag{13d}$$

Thus, the reduced model allows the estimation of the total heritable variation, even though not all the underlying parameters are identifiable.

Equation 13b refers to the covariance between the family breeding value and the non-family breeding value. This is a meaningful linear combination, as it expresses the covariance between genetic effects on kin (including self) versus those on strangers. If this covariance is positive, members from different families are cooperative, whereas a negative value indicates competition between families.

Appendix B shows that the expectations of the non-genetic variance components in Equation 10 are given by

$$E(\hat{\sigma}^2_g) = 2\sigma_{E_{D,S_u}} + (n-2)\sigma_{E_{S_f, S_u}} \tag{14a}$$

$$\mathrm{E}(\hat{\sigma}^2_e) = \sigma^2_{E_D} + (\tfrac{1}{2} n - 1)\sigma^2_{E_{S_f}} + \tfrac{1}{2} n \sigma^2_{E_{S_u}} - \sigma^2_g \tag{14b}$$

$$\mathrm{E}(\hat{\rho}) = \frac{2\sigma_{E_D E_{S_f}} + (\tfrac{1}{2} n - 2)\sigma^2_{E_{S_f}} + \tfrac{1}{2} n \sigma^2_{E_{S_u}} - \sigma^2_g}{\sigma^2_e} \tag{14c}$$

Equations 14a-c shows that the underlying non-genetic parameters are not uniquely identifiable, because there are only three estimable parameters (σ^2_g , σ^2_e and ρ) which are a

function of six unknowns. This was expected, as it is also the case for models not distinguishing between IGEs on kin *vs*. strangers (Bijma *et al.*, 2007b).

2.2.3 Consequences of ignoring kin-dependent IGEs

This section investigates the bias in the estimated genetic parameters when IGEs differ between kin and strangers while this is ignored in the statistical analysis. Thus, it is assumed that the true model generating the trait values is given by Equation 3 above, which distinguishes between IGEs on kin *vs*. strangers, while the statistical model used to estimate genetic parameters is the traditional direct-indirect mixed linear model (Muir, 2005),

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_D\mathbf{a}_D + \mathbf{Z}_S\mathbf{a}_S + \mathbf{Wg} + \mathbf{e}, \qquad (15)$$

where $\mathbf{a}_S$ is vector of IGEs on group mates, not distinguishing between kin and strangers, and $\mathbf{Z}_S$ an incidence matrix linking observations on individuals to the IGEs of all their group mates. The $\mathbf{Z}_D\mathbf{a}_D$ and $\mathbf{Wg}$ are as in the full model (Equation 9), whereas the residual variance structure is as in the reduced model (Equations 11; see discussion below). Note that Equation 15 differs from the reduced model (Equation 10), because the term $\mathbf{Z}_S\mathbf{a}_S$ includes the IGEs of all $n-1$ group mates; not only those belonging to the other family making up the group.

To investigate the bias resulting from fitting a conventional IGE model (Equation 15) to data in which IGE differ between kin and strangers, we derived the expectations of the estimated breeding values and variance components produced by Equation 15 when data are generated by Equation 3. Those expectations follow from the informative covariances

in the data that Equation 15 utilizes to estimate the genetic parameters, and can be obtained using the method of Bijma (2010a; see Appendix C). Results showed that

$$E(\hat{\sigma}^2_{A_D})_{Eqn.15} = \sigma^2_{A_D} + (½n-1)^2\left(\sigma^2_{A_{S_f}} - 2\sigma_{A_{S_f}A_{S_u}} + \sigma^2_{A_{S_u}}\right) + (n-2)(\sigma_{A_D A_{S_f}} - \sigma_{A_D A_{S_u}}) \quad (16a)$$

$$E(\hat{\sigma}_{A_{DS}})_{Eqn.15} = \sigma_{A_D A_{S_u}} + (\tfrac{1}{2}n-1)(\sigma_{A_{S_f}A_{S_u}} - \sigma^2_{A_{S_u}}) \quad (16b)$$

$$E(\hat{\sigma}^2_{A_S})_{Eqn.15} = \sigma^2_{A_{S_u}} \quad (16c)$$

$$E(\hat{\sigma}^2_{A_T})_{Eqn.15} = \sigma^2_{A_T} \quad (16d)$$

These results show that the direct genetic variance and the direct-indirect genetic covariance estimated with the conventional linear model for IGEs (Equation 15) are biased when IGEs depend on relatedness. In other words, the estimate of $\sigma^2_{A_D}$ is biased because the right-hand side of Equation 16a differs from. Similarly, difference of the right-hand side of Equation 16b from $\sigma_{A_{DS}}$ indicates bias of $\sigma_{A_{DS}}$. Moreover, the estimated indirect genetic variance from Equation 15 refers to the magnitude of IGEs expressed on strangers (Equation 16c). Surprisingly, despite the incorrect model assumptions, the traditional direct-indirect model yields an unbiased estimate of the total heritable variance (Equation 16d). Beware that results in Equations 16a-d are correct only if the residual co-variance structure accounts for differences between indirect effects on kin *vs*. strangers, as given by Equation 11. Therefore, when the aim is to estimate TBVs using the traditional direct-indirect mixed model of Muir (2005), this model should be implemented including a random group effect and the residual variance structure given in Equation 11 above.

We did not attempt to derive the expectations of estimated genetic parameters from the traditional direct-indirect model (Equation 15) when the residual variance structure is

incorrect (*i.e.,* different from that given in Equation 11). The reason is that those expectations will depend not only on the assumed true genetic model (Equation 3), but also on the data structure. For example, in data consisting of many groups, the covariance between relatives in different groups will dominate the estimates, and incorrect covariances within groups may have little effect. In that case, estimates may be close to values given in Equation 16. On the other hand, when groups are fewer, information from the within-group (co)variances will become more important and results may deviate more from Equation 16.

2.2.4 Simulation

Methods. We used Monte Carlo simulation to validate the theoretical relationships between the true model, the reduced model and the traditional model presented above (Equations 12, 14 and 16). Data were generated under the model in Equation 3, and analysed using either the reduced model in Equation 10 or the traditional model in Equation 15, using the ASReml software (Gilmour *et al.*, 2006). A population of two discrete generations was simulated using R (R Development Core Team, 2011). No fixed effects were simulated. The base generation consisted of 100 sires and 1000 dams, which were unrelated. To produce the second generation, sires and dams of the first generation were mated at random, each sire being mated to 10 dams, and each dam producing 10 full sib offspring. Individuals of the second generation were kept in 2,500 groups of 4 individuals each, and each group consisted of two full sib families, each family contributing two individuals. Table 2 shows the range of genetic parameters simulated. For each set of genetic parameters, estimates were averaged over one hundred replicates. Details of the simulation are given in Appendix D.

2.3 Results

Table 3 shows a comparison between simulated and estimated values for $\sigma^2_{A_{T_f}}$, $\sigma_{A_{T_f} A_{S_u}}$, $\sigma^2_{A_{S_u}}$ and $\sigma^2_{A_T}$ from the reduced model, for different magnitudes of IGEs. We used $\sigma^2_{P_{S_f}} = \sigma^2_{P_{S_u}}$ of either 50% or 25% of $\sigma^2_{P_D}$, to represent high or low indirect effects, and $\sigma^2_{A_{S_f}}$, $\sigma^2_{A_{S_u}}$ of either 10%,12.5%, 20% and 25% of $\sigma^2_{P_{S_u}}$, $\sigma^2_{P_{S_f}}$, to represent high or low heritability of IGE, and a range of genetic correlations between direct effects, indirect effects on kin and indirect effects on strangers (Table 2). Results show close agreement between simulated and estimated values as proven by the relative error which is less than or equal to 5% in all cases. (Those small errors originate from stochasticity among replicates, and do not indicate systematic bias). These results confirm the theoretical relationships between the full and reduced model presented in Equations 12 and 13. Thus, the reduced model yields unbiased genetic parameters of the family and non-family breeding values, and of the total breeding value. We also compared the estimated non-genetic components to their expectations given in Equation 14, showing close agreement (results not shown).

Table 2. Parameter values used for validation of reduced and traditional model

Scheme	Deviation from basic scheme[1]
Alt.1	$\sigma^2_{A_{S_f}} = 0.125, \sigma^2_{A_{S_u}} = 0.100$
Alt.2	$\sigma^2_{P_{S_f}} = \sigma^2_{P_{S_u}} = 0.25,\ \sigma^2_{A_{S_f}} = 0.050, \sigma^2_{A_{S_u}} = 0.063$
Alt.3	$\sigma^2_{P_{S_f}} = \sigma^2_{P_{S_u}} = 0.25,\ \sigma^2_{A_{S_f}} = 0.063, \sigma^2_{A_{S_u}} = 0.050$
Alt.4	$r_{A_D,A_{S_f}} = r_{E_D,E_{S_f}} = 0.5$
Alt.5	$r_{A_D,A_{S_f}} = r_{E_D,E_{S_f}} = 0.5, \sigma^2_{A_{S_f}} = 0.125, \sigma^2_{A_{S_u}} = 0.1$
Alt.6	$r_{A_D,A_{S_f}} = r_{E_D,E_{S_f}} = 0.5\ , \sigma^2_{A_{S_f}} = 0.050, \sigma^2_{A_{S_u}} = 0.063, \sigma^2_{P_{S_f}} = \sigma^2_{P_{S_u}} = 0.25$
Alt.7	$r_{A_D,A_{S_f}} = r_{E_D,E_{S_f}} = 0.5\ , \sigma^2_{A_{S_f}} = 0.125, \sigma^2_{A_{S_u}} = 0.05, \sigma^2_{P_{S_f}} = \sigma^2_{P_{S_u}} = 0.25$
Alt.8	$r_{A_D,A_{S_f}} = r_{E_D,E_{S_f}} = -0.5$
Alt.9	$r_{A_D,A_{S_u}} = r_{A_{S_f},A_{S_u}} = r_{E_D,E_{S_u}} = r_{E_{S_f},E_{S_u}} = 0.5$
Alt.10	$r_{A_D,A_{S_u}} = r_{A_{S_f},A_{S_u}} = r_{E_D,E_{S_u}} = r_{E_{S_f},E_{S_u}} = 0.5,\ \sigma^2_{A_{S_f}} = 0.125, \sigma^2_{A_{S_u}} = 0.100$
Alt.11	$r_{A_D,A_{S_u}} = r_{A_{S_f},A_{S_u}} = r_{E_D,E_{S_u}} = r_{E_{S_f},E_{S_u}} = 0.5,$ $\sigma^2_{A_{S_f}} = 0.050, \sigma^2_{A_{S_u}} = 0.063, \sigma^2_{P_{S_f}} = \sigma^2_{P_{S_u}} = 0.25$
Alt.12	$r_{A_D,A_{S_u}} = r_{A_{S_f},A_{S_u}} = r_{E_D,E_{S_u}} = r_{E_{S_f},E_{S_u}} = 0.5,\ \sigma^2_{A_{S_f}} = 0.0625, \sigma^2_{A_{S_u}} = 0.050,$ $\sigma^2_{P_{S_f}} = \sigma^2_{P_{S_u}} = 0.25$
Alt.13	$r_{A_D,A_{S_f}} = r_{A_D,A_{S_u}} = r_{A_{S_f},A_{Su}} = r_{E_D,E_{S_f}} = r_{E_D,E_{S_u}} = r_{E_{S_f},E_{S_u}} = 0.5$
Alt.14	$r_{A_D,A_{S_f}} = r_{A_D,A_{S_u}} = r_{A_{S_f},A_{Su}} = r_{E_D,E_{S_f}} = r_{E_D,E_{S_u}} = r_{E_{S_f},E_{S_u}} = -0.1$

[1]The basic scheme has $\sigma^2_{P_D} = 1$, $\sigma^2_{A_D} = 0.5$, $\sigma^2_{P_{S_f}} = \sigma^2_{P_{S_u}} = 0.5$, $\sigma^2_{A_{S_f}} = 0.125$, $\sigma^2_{A_{S_u}} = 0.100$ and all correlations are zero. Alternative schemes only show parameters that deviate from the basic scheme.

Table 3. Errors in estimates for the reduced model

Scheme	Error%			
	$\hat{\sigma}^2_{A_{T_f}}$	$\hat{\sigma}_{A_{T_f},A_{S_u}}$	$\hat{\sigma}^2_{A_{S_u}}$	$\hat{\sigma}^2_{A_T}$
Basic	0	0	-1	0
Alt.1	1	0	0	1
Alt.2	-2	0	2	-1
Alt.3	-1	0	-2	-1
Alt.4	2	0	-1	-1
Alt.5	0	0	0	-1
Alt.6	-2	0	0	-2
Alt.7	-1	0	2	0
Alt.8	-1	0	1	0
Alt.9	1	1	2	1
Alt.10	-1	-2	-2	-2
Alt.11	0	1	2	1
Alt.12	1	1	0	1
Alt.13	-2	-4	-3	-3
Alt.14	1	5	-1	1

See Table 2 for a description of schemes. Error %=100% $\times$(estimated -simulated)/simulated. When the prediction equals the true value $E[\text{error\%}] \approx 0$. The expected absolute error equals E [|error%|] ~2.5%, and E|error %|> 5% implies significant bias ($p<0.05$; two sided).

Table 4. Comparison of the expected (Equation 16) and empirical estimates for the traditional model

Scheme	Error%			
	$\hat{\sigma}^2_{A_D}$	$\hat{\sigma}_{A_{DS}}$	$\hat{\sigma}^2_{A_S}$	$\hat{\sigma}^2_{A_{TBV}}$
Basic	-1	- 2	-2	-1
Alt.1	1	-1	0	1
Alt.2	0	1	0	-1
Alt.3	0	4	-2	0
Alt.4	0	-4	0	2
Alt.5	-2	0	-1	-2
Alt.6	-1	2	1	-1
Alt.7	0	0	0	0
Alt.8	1	2	-1	-3
Alt.9	0	2	2	1
Alt.10	-3	-4	0	-2
Alt.11	-1	-2	0	-1
Alt.12	0	2	0	1
Alt.13	-2	-5	-3	-3
Alt.14	1	3	0	-2

Table 4 shows a comparison between the theoretically expected values of the estimated variance components from the traditional model (Equation 16) and the empirical values estimated from the simulated data using the traditional model (Equation 15). Results confirm the theoretical expectation that the traditional direct-indirect model yields biased

estimates of the direct genetic variance and the direct-indirect genetic covariance, but unbiased estimates of the genetic variance of IGEs on strangers and of the total genetic variance.

2.4 Discussion

We have proposed a quantitative genetic model and investigated methodology to estimate the genetic parameters of traits affected by IGEs when those IGEs differ systematically between kin vs. strangers. Results show that the full set of genetic parameters for the full model is not statistically identifiable. We also presented a reduced model that yields unbiased estimates of meaningful linear combinations of genetic parameters: the variance of the family breeding value, the covariance between family breeding value and IGEs on strangers, and the variance of IGEs on strangers. The reduced model also provides estimates of the variance in total breeding value, and predictions of the total breeding values of individuals.

An interesting question is whether experimental designs exist that allow estimating all six genetic parameters of the full model (Equation 3). Our results show that this is not possible when pairs of individuals can be categorized into either kin or unrelated, each category shows a different IGE, and group-composition is the same for all groups. As long as group composition with respect to family is the same for all groups, this situation results in full confounding of the direct effect and the IGE on kin, irrespective of the composition of the groups (i.e., 50/50, 25/75, etc.; Appendix A).

When differences in IGE originate from factors that usually go together with relatedness such as familiarity, rather than from relatedness *per se*, experimental designs that disconnect relatedness from those factors may allow estimation of the full set of genetic

parameters. For example, when individuals recognize each other due to prior association (see Introduction), relatives that grow up together will recognize each other and adjust their behaviour, whereas relatives that grow up separately will interact similarly to unrelated individuals. This may, for example, occur in mammals such grey mouse lemur (Kessler *et al.*, 2012) or rats (Hepper, 1983; Hepper, 1986), where full siblings often grow up in the same litter, while paternal half siblings grow up in different environments. Our preliminary investigations show that all six genetic parameters are statistically identifiable in this situation when groups consist of a mix of full sibs, half sibs and unrelated individuals. A statistically more powerful approach may come from cross-fostering designs, where full siblings that grow up in different litters may interact as if they were unrelated. When cross-fostering is impossible and a mix of full and half siblings is unavailable, a solution may come from utilizing the variation in relatedness among pairs of full siblings, estimated using genome-wide genetic markers (Hill, 1993; Visscher *et al.*, 2006). However, as variation in relatedness among full siblings is limited, this approach will require large sample sizes.

When relatedness itself (as opposed to, *e.g.*, familiarity) is the causal factor underlying a difference in IGE, it would seem unlikely that the full set of genetic parameters can be identified. When individuals adjust their behaviour according to their relatedness to the recipient of the behaviour, as predicted by kin selection theory (Hamilton, 1964), any covariance between trait values of individuals is a function of relatedness and of genetic parameters of interest, which depends on this relatedness. This would seem to suggest full confounding.

However, variation in group-composition seems to offer a solution. For example, having three different group compositions in a population may allow estimating all six genetic parameters. The first composition may have unrelated individuals only, the second may

have two family members supplemented with unrelated individuals, and the third may have three family members supplemented with unrelated individuals. From the first composition, $\sigma^2_{A_D}$, $\sigma_{A_{D,}A_{S_u}}$ and $\sigma^2_{A_{S_u}}$ can be estimated using the traditional direct-indirect mixed model (Muir, 2005). Then, using the reduced model, $\sigma^2_{A_F}(n_f = 2) = \sigma^2_{A_D} + 2\sigma_{A_{D,}A_{S_f}} + \sigma^2_{A_{S_f}}$ can be estimated from the second composition, and $\sigma^2_{A_F}(n_f = 3) = \sigma^2_{A_D} + 4\sigma_{A_{D,}A_{S_f}} + 4\sigma^2_{A_{S_f}}$ can be estimated from the third composition, as well as $\sigma_{A_{S_f,}A_{S_u}}$ and again $\sigma^2_{A_{S_u}}$. Then, since $\sigma^2_{A_D}$ is known from the first composition, this yields two equations with two unknowns, and thus can be solved yielding estimates of $\sigma_{A_D A_{S_f}}$ and $\sigma^2_{A_{S_f}}$. Moreover, the estimate of $\sigma_{A_{F,}A_{S_u}}$ from either the second or third composition can be used to obtain $\sigma_{A_{S_f,}A_{S_u}}$, because $\sigma_{A_{D,}A_{S_u}}$ is known from the first composition (see Equation 12b). Then all six genetic parameters are estimated. Thus, variation in group composition with respect to family seems to allow estimating all six genetic parameters. Statistical power, however, may be very limited, and further complications may arise when IGEs depend on group size (Hadfield and Wilson, 2007; Bijma, 2010b), which we did not investigate here.

When IGEs depend on relatedness, the traditional direct-indirect mixed model that ignores this dependency yields biased estimates of the direct genetic variance and the direct-indirect genetic covariance, but an unbiased estimate of the variance in total breeding value. Thus, even though the full set of genetic parameters is not statistically identifiable, the total heritable variance and total breeding values can be estimated, either using the reduced model or the traditional model. This is an important result, because kin-

dependent IGEs appear to be widespread in natural and domestic populations of both animals and plants (see Introduction).

The reduced model and traditional model are statistically equivalent, *i.e.*, yield the same maximum likelihood, but represent different linear combinations of the underlying parameters. The main difference is that the estimates of the reduced model are biologically meaningful in the context of kin-selection theory (Hamilton, 1964), as they separate the effects on kin (the family breeding value) from those on unrelated individuals. The correlation between the family breeding value and IGE on strangers, for example, measures the degree of competition or cooperation between families. With the exception of the IGEs on strangers and the total breeding value, the estimates of the traditional model do not seem to have a clear biological meaning (Equation 16). Thus, the reduced model is preferable in terms of interpretation.

In this study, we have considered only the random effects; consequences of kin-dependent IGEs on the fixed effects to be included in the **Xb** term of the models have been ignored. When IGEs depend on relatedness, IGEs on kin *vs.* strangers probably not only show incomplete correlation, but also differ systematically in level. In other words, individuals interacting primarily with kin probably receive more favourable IGEs than those interacting primarily with strangers, which creates a systematic difference in trait level between individuals interacting with different numbers of kin. This is not accounted for by the random effects in the model, because those are zero on average by construction. Hence, a fixed effect for the number of relatives an individual interacts with should be included in the model. This is similar to the inclusion of a fixed effect for the number of group mates when group size varies. Because estimation of a fixed effect with a few degrees of freedom is straight forward, we did not investigate this in detail. In our simulations, there was no

need to account for such a fixed effect, because all individuals had the same number of kin and strangers among their group mates.

In animal and plant breeding, the focus is on improving the mean trait value of the population in the next generations. Theoretical studies have shown that group and kin selection methods utilize the total heritable variation for response to selection (Muir, 2005; Bijma *et al.*, 2007b; Ellen *et al.*, 2008; McGlothlin *et al.*, 2010). This theoretical expectation is supported by results from selection experiments that have used group and/or kin selection without explicit reference to the total breeding value (Wade, 1976; Wade, 1977; Goodnight, 1985; Muir, 1996). Whether or not this result extends to the situation where IGEs differ between kin and strangers is interesting, but has not been investigated to our knowledge.

To optimize selection for traits affected by interactions among individuals, the ideal selection criterion is the TBV of selection candidates estimated using all available information. This is because response to selection equals the change in mean TBV from one generation to the next, so that maximizing the accuracy of estimated TBVs also maximizes response to selection. Because Equation 4 is generally valid, this result holds irrespective of whether or not IGEs depend on relatedness (Bijma, 2011b). Hence, the availability of kin and group selection methods does not make estimated TBVs superfluous. Moreover, knowledge of the total heritable variance quantifies the intrinsic potential of a population to respond to selection, and therefore provides a measure of efficiency for breeding schemes (Bijma, 2011b). The variance in TBV, therefore, is an important parameter both for optimizing individual selection decisions and for the evaluation of breeding schemes. This work has shown how the definition and estimation of the variance in TBV can be extended to schemes where IGEs differ between kin and strangers. This

extension of variance in TBV to schemes where IGEs differ between kin and strangers may contribute to breeding plan design and application.

Acknowledgments

We acknowledge financial support for PB by the Dutch Technology Foundation (STW), which is part of the Netherlands Organisation for Scientific Research (NWO), and support from the Danish Ministry of Food, Agriculture and Fisheries, Program Innovationsloven.

SWA thanks Lei Zhou for his good inputs.

Conflict of interest

The authors declare no conflict of interest.

References

Arango J, Misztal I, Tsuruta S, Culbertson M, Herring W (2005). Estimation of variance components including competitive effects of large white growing gilts. *J Anim Sci* **83:** 1241-1246.

Bijma P (2010a). Estimating indirect genetic effects: precision of estimates and optimum designs. *Genetics* **186:** 1013-1028.

Bijma P (2010b). Multilevel selection 4: modeling the relationship of indirect genetic effects and group size. *Genetics* **186:** 1029-1031.

Bijma P (2011a). Breeding for Social Interaction, for Animal Welfare, in : Meyers RA (ed) *Encyclopedia of sustainability science and technology, S*pringer Science and Business Media LLC: New York, pp 1-40.

Bijma P (2011b). A general definition of the heritable variation that determines the potential of a population to respond to selection. *Genetics* **189:** 1347-1359.

Bijma P, Muir WM, Ellen ED, Wolf JB, Van Arendonk JAM (2007b). Multilevel selection 2: estimating the genetic parameters determining inheritance and response to selection. *Genetics* **175:** 289-299.

Bijma P, Wade MJ (2008). The joint effects of kin, multilevel selection and indirect genetic effects on response to genetic selection. *J Evol Biol* **21:** 1175-1188.

Chen CY, Kachman SD, Johnson RK, Newman S, Van Vleck LD (2008). Estimation of genetic parameters for average daily gain using models with competition effects. *J Anim Sci* **86:** 2525-2530.

Coffin HR, Watters JV, Mateo JM (2011). Odor-based recognition of familiar and related conspecifics: a first test conducted on captive Humboldt penguins (*Spheniscus humboldti). PLoS One* **6:** e25002.

Ellen ED, Visscher J, van Arendonk JA, Bijma P (2008). Survival of laying hens: genetic parameters for direct and associative effects in three purebred layer lines. *Poult Sci* **87:** 233-239.

Falconer DS (1965). Maternal effects and selection response. In S.J.Geerts(ed) *Genetics today: Proceedings of the XI International congress on genetics*. Pergamon Press : New york. Vol 3, pp 763–774.

Falconer DS, Mackay TFC (1996). *Introduction to Quantitative Genetics*, Ed 4. Longmans Green, Harlow, Essex, UK.

Frank SA (2007). All of life is social. *Current Biology* **17:** R648-R650.

Gilmour AR, Gogel BJ, Cullis BR, Thomson R (2006). *ASReml user guide release 2.0*. VSN International Ltd. Hemel Hempstead,UK.

Goodnight CJ (1985). The influence of environmental variation on group and individual selection in a cress. *Evolution* **39:** 545-558.

Griffing B (1967). Selection in reference to biological groups .I. Individual and group selection applied to populations of unordered groups. *Australian Journal of Biological Sciences* **20:** 127-139.

Griffing B (1976). Selection in reference to biological groups. V. Analysis of full-sib groups. *Genetics* **82:** 703-722.

Hadfield JD, Wilson AJ (2007). Multilevel selection 3: modeling the effects of interacting individuals as a function of group size. *Genetics* **177:** 667-668.

Hamilton WD (1964). The genetical evolution of social behavior. *J Theor Biol* **7:** 1–16.

Henderson CR (1953). Estimation of Variance and Covariance Components. *Biometrics* **9:** 226-252.

Henderson CR (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics* **31:** 423-447.

Henderson CR (1985). Best Linear Unbiased Prediction Using Relationship Matrices Derived from Selected Base-Populations. *Journal of Dairy Science* **68:** 443-448.

Hepper PG (1983). Sibling Recognition in the Rat. *Animal Behaviour* **31:** 1177-1191.

Hepper PG (1986). Kin recognition: functions and mechanisms. A review. *Biol Rev Camb Philos Soc* **61:** 63-93.

Hill WG (1993). Variation in genetic identity within kinships. *Heredity* **71:** 652-653.

Kessler SE, Scheumann M, Nash LT, Zimmermann E (2012). Paternal kin recognition in the high frequency/ultrasonic range in a solitary foraging mammal. *Bmc Ecology* **12**.

Kirkpatrick M, Lande R (1989). The evolution of maternal characters. *Evolution* **43:** 485-503.

Lynch M, Walsh B (1998). *Genetics and analysis of quantitative traits*. Sinauer Associates, Sunderland, MA.

Mateo JM (2004). Recognition systems and biological organization: The perception component of social recognition. *Ann Zool Fenn* **41:** 729-745.

Mateo JM, Holmes WG (2004). Cross-fostering as a means to study kin recognition. *Animal Behaviour* **68:** 1451-1459.

McGlothlin JW, Brodie III ED (2009). How to measure indirect genetic effects: the congruence of trait-based and variance-partitioning approaches. *Evolution* **63:** 1785-1795.

McGlothlin JW, Moore AJ, Wolf JB, Brodie III ED (2010). Interacting phenotypes and the evolutionary process. III. Social evolution. *Evolution* **64:** 2558-2574.

Moore AJ, Brodie III ED, Wolf JB (1997). Interacting phenotypes and the evolutionary process : 1. direct and indirect genetic effects of social interactions. *Evolution* **51:** 1352-1362.

Muir WM (1996). Group selection for adaptation to multiple-hen cages: selection program and direct responses. *Poultry Science* **75:** 447-458.

Muir WM (2005). Incorporation of competitive effects in forest tree or animal breeding programs. *Genetics* **170:** 1247-1259.

R Development Core Team. (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Tang-Martinez Z (2001). The mechanisms of kin discrimination and the evolution of kin recognition in vertebrates: a critical re-evaluation. *Behav Process* **53:** 21-40.

Visscher PM, Medland SE, Ferreira MAR, Morley KI, Zhu G, Cornes BK *et al.* (2006). Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. *Plos Genet* **2:** 316-325.

Wade MJ (1976). Group selection among laboratory populations of Tribolium. *P Natl Acad Sci USA* **73:** 4604-4607.

Wade MJ (1977). Experimental-study of group selection. *Evolution* **31:** 134-153.

Willham RL (1963). The covariance between relatives for characters composed of components contributed by related individuals. *Biometrics* **19:** 18-27.

Wilson AJ, Morrissey MB, Adams MJ, Walling CA, Guinness FE, Pemberton JM *et al.* (2011). Indirect genetics effects and evolutionary constraint: an analysis of social dominance in red deer, *Cervus elaphus*. *J Evol Biol* **24:** 772-783.

Wolf JB, Brodie III ED, Cheverud JM, Moore AJ, Wade MJ (1998). Evolutionary consequences of indirect genetic effects. *Trends Ecol Evol* **13:** 64-69.

Appendix A

This appendix shows that the full model is not statistically identifiable, whereas the reduced model is identifiable. Estimation of genetic parameters for direct and indirect genetic effects rests on covariances between phenotypes of relatives and of their social partners (Lynch and Walsh, 1998). Only covariances between relatives (or social partners) present in different groups contribute to the estimation of genetic parameters, because within-group covariances are fully confounded with the non-genetic direct and indirect effects. The following, therefore, considers between-group covariances only.

Each group consists of members of two families. There are no genetic covariances between groups not sharing a family; hence those group combinations can be ignored. Then, when considering two groups having one family in common, there are three families in total; the common family, denoted F_1, and its partner family in each group, denoted F_2 and F_3. Before we derive covariance between individual, the individual's total breeding value, which is the total heritable impact of an individual's genes on the mean trait value of the population when interaction differ between kin *vs*. strangers, is given as:

$$A_{T,i} = A_{Di} + 2(n_f - 1)A_{S_f} + n_f A_{S_u} \tag{A1}$$

Taking the variance of the total breeding value yields

$$\sigma^2_{A_T} = \sigma^2_{A_D} + 2(n_f - 1)\sigma_{A_D A_{S_f}} + (n_f - 1)^2 \sigma^2_{A_{S_f}} + 2n_f(n_f - 1)\sigma_{A_{S_f} A_{S_u}} + n^2 \sigma^2_{A_{S_u}} \tag{A2}$$

When we have one family common in two groups there are only three informative covariances. First, the covariance between the phenotypes of a member of F_1 in each group,

$$Cov(P_i, P_j \mid i, j \in F_1) = r\left[\sigma^2_{A_D} + 2(n_f - 1)\sigma_{A_D A_{S_f}} + (n_f - 1)^2 \sigma^2_{A_{S_f}}\right] \quad \text{(A3)}$$

where r denotes relatedness between members of the same family, and n_f the number of members of F_1 in each group (assumed to be the same in both groups). Second, the covariance between a member of the common family (F_1) in the one group, and a member of a partner family in the other group (F_2 is considered here, but the result for F_3 is identical),

$$Cov(P_i, P_j \mid i \in F_1, j \in F_2) = r\left[n_f \sigma_{A_D A_{S_u}} + n_f (n_f - 1)\sigma_{A_{S_f} A_{S_u}}\right] \quad \text{(A4)}$$

Third, the covariance between two members of the partner families (F_2 and F_3) in different groups is

$$Cov(P_i, P_j \mid i \in F_2, j \in F_3) = r\, n_f^2\, \sigma^2_{A_{S_u}} \quad \text{(A5)}$$

This equation shows that the variance of IGE on strangers is estimable. In total, however, these three equations contain six unknowns (the six genetic parameters to be estimated) and cannot be solved. Thus, the full model is not identifiable.

Equations A1 through A3 also show that the reduced model is identifiable, since they represent the informative covariances and there are only three genetic parameters to estimate. Moreover, Equations A1 through A3 imply that the expected values of the estimated genetic parameters of the reduced model are given by Equation 12a-c when $n_f = \frac{1}{2}n$.

Appendix B

This appendix shows the derivation of the non-genetic covariance structure generated by Equation 3 (Equations 11 and 14) and refers to the reduced model. Non-genetic covariances occur only among individuals within the same group. Since the genetic model terms fully account for genetic covariances within groups, those can be ignored here.

There are three non-genetic parameters of interest: the covariance between group members of different families, the covariance between group members of the same family, and the residual variance. Because all groups have the same composition, these parameters are the same for all groups. This leads to the block-diagonal residual variance structure given by Equation 11, which has a single residual variance, a covariance between group mates of the same family, and a second covariance between group mates of different families.

Consider two group mates, say *i* and *k*. The group mates of *i* of its own family are denoted *j*, and those of the other family *j'*. Analogously, the group mates of *k* are denoted *l* and *l'*. Note that *k* is one of the individuals included in *j* and *j'*, whereas *i* is one of the individuals included in *l* and *l'*. Then the non-genetic covariance between the phenotypes of *i* and *k* is given by

$$\operatorname{cov}(P_i, P_k)_E = \operatorname{cov}\left(E_{D_i} + \sum_{j=1}^{\frac{n}{2}-1} E_{S_{f,j}} + \sum_{j'=1}^{\frac{n}{2}} E_{S_{u,j'}}, E_{D_k} + \sum_{l=1}^{\frac{n}{2}-1} E_{S_{f,l}} + \sum_{l'=1}^{\frac{n}{2}} E_{S_{u,l'}} \right). \quad \text{(B1)}$$

First consider this covariance when *i* and *k* are group members of different families, giving

$$\operatorname{cov}(P_i, P_k)_E = 2\sigma_{E_{D,S_u}} + (n-2)\sigma_{E_{S_f,S_u}} . \quad \text{(B2)}$$

The first term arises because *i* affects *k* and vice versa, whereas the second term arises because *i* and *k* have $n-2$ group mates in common. In Equation 10, the non-genetic covariance between unrelated group mates equals the variance of the random group effect. Hence,

$$E(\hat{\sigma}_g^2) = 2\sigma_{E_{D,S_u}} + (n-2)\sigma_{E_{S_f,S_u}} , \tag{B3}$$

which is Equation 14a.

Next, consider the full non-genetic variance. From Equation B1, it follows that

$$\text{var}(P)_E = \sigma_{E_D}^2 + (\tfrac{1}{2}n-1)\sigma_{E_{S_f}}^2 + \tfrac{1}{2}n\sigma_{E_{S_u}}^2 \tag{B4}$$

In Equation 10, the full non-genetic variance is the sum of the group variance and the residual variance. Hence, the residual variance in Equation 10 follows from subtracting the group variance from Equation B4, giving

$$\text{E}(\hat{\sigma}_e^2) = \sigma_{E_D}^2 + (\tfrac{1}{2}n-1)\sigma_{E_{S_f}}^2 + \tfrac{1}{2}n\sigma_{E_{S_u}}^2 - \sigma_g^2$$

which is Equation 14b.

Finally, consider the covariance when *i* and *k* are group members of the same family,

$$\text{cov}(P_i, P_k)_E = 2\sigma_{E_{D,S_f}} + (\tfrac{1}{2}n-2)\sigma_{E_{S_f}}^2 + \tfrac{1}{2}n\sigma_{E_{S_u}}^2 \tag{B5}$$

The first term arises because *i* affects *k* and vice versa, the second term arises because *i* and *k* have $\frac{1}{2}n-2$ group mates of their own family in common, and the third term arises because *i* and *k* have $\frac{1}{2}n$ group mates of the other family in common. In Equation 10 and 11, the covariance between unrelated group mates is the sum of the group variance and the residual covariance, $\sigma_g^2+\rho\sigma_e^2$. Hence, the residual correlation follows from subtracting the group variance from Equation B5 and dividing by the residual variance, giving

which is Equation 14b.

$$\mathrm{E}(\hat{\rho})=\frac{2\sigma_{E_D E_{S_f}}+\left(\frac{1}{2}n-2\right)\sigma^2_{E_{S_f}}+\frac{1}{2}n\sigma^2_{E_{S_u}}-\sigma_g^2}{\sigma_e^2}$$

,

which is Equation 14c.

Appendix C

This appendix shows the derivation of Equations 16a-d, being the expectations of genetic parameters when the traditional direct-indirect model (Equation 15) is applied to data generated by the model in Equation 3. The derivation uses the method of Bijma (2010a).

Direct genetic variance: With two families per group, the information for estimating the direct genetic variance using Equation (15) comes from the variable $z_{kl} = \bar{P}_{kl} - \varphi\, \bar{P}_{k'l}$, in which $\varphi = \frac{½n-1}{½n}$, $\bar{P}_{kl}$ is the mean phenotype of the family of interest k in group *l*, and $\bar{P}_{k'l}$ is the mean phenotype of the other family *k´* in group *l* (Equations B15 and B16 in Bijma, 2010a; the z_{kl} is referred to as the "effective record") and *n* is the group size. When the data are generated by Equation 3, the expectation of z_{kl} conditional on the family of interest *k* in group *l* equals

$$E[z_{kl} \mid k] = A_{D,k} + (½n-1)A_{S_f,k} - (½n-1)A_{S_u,k} \qquad \text{(C1)}$$

which depends not only on the DGE of family *k*, but also on the IGEs of family *k* on kin and strangers. The expected value of the estimated direct genetic variance follows from the variance of *z*, giving

$$E(\hat{\sigma}^2_{A_D})_{Eqn.15} = \sigma^2_{A_D} + (½n-1)^2\left(\sigma^2_{A_{S_f}} - 2\sigma_{A_{S_f}A_{S_u}} + \sigma^2_{A_{S_u}}\right) + (n-2)(\sigma_{A_D A_{S_f}} - \sigma_{A_D A_{S_u}}),$$

which is Equation 16a.

Indirect genetic variance: The information for estimating the indirect genetic variance using Equation (15) comes from the variable $z_{kl} = \overline{P}_{k'l} / (½ n)$ (Equation B18 in Bijma, 2010a). When the data are generated by Equation 3, the expectation of z_{kl} conditional on the family of interest, equals

$$E[z_{kl} \mid k] = A_{S_u,k} \qquad \text{(C2)}$$

Thus the expected value of the estimated indirect genetic variance equals

$$E(\hat{\sigma}^2_{A_S})_{Eqn.15} = \sigma^2_{A_{S_u}}$$

, which is Equation 16c.

Direct-indirect genetic covariance: From Equations C1 and C2, it follows that

$$E(\hat{\sigma}_{A_{DS}})_{Eqn.15} = \sigma_{A_D A_{S_u}} + (½ n - 1)(\sigma_{A_{S_f} A_{S_u}} - \sigma^2_{A_{S_u}}),$$

which is Equation 16b. Thus, when the IGE on kin is identical to the IGE on strangers, the second term becomes zero, and $E(\hat{\sigma}_{A_{DS}})_{Eqn.15} = \sigma_{A_{DS}}$.

Total heritable variation: The information for estimating the total heritable variance using Equation (15) comes from the variable $z_{kl} = \sum_{j=1}^{n} P_{kl,j} / (½ n)$ (Equation B20 in Bijma, 2010a). When the data are generated by Equation 3, the expectation of z_{kl} conditional on the family of interest equals

$$E[z_{kl} \mid k] = ½ n \left[A_D + (½ n - 1) A_{S_f} + ½ n A_{S_u} \right] / (½ n) = A_{T,k} . \qquad \text{(C3)}$$

Thus the expected value of the estimated total heritable variance equals

$$E(\hat{\sigma}^2_{A_T})_{Eqn.15} = \sigma^2_{A_T}$$

,

which is Equation 16d.

Appendix D

This appendix shows details of the stochastic simulation. Breeding values of individuals in the base generation were simulated from the multivariate normal distribution

$$\begin{bmatrix} A_D \\ A_{S_f} \\ A_{S_u} \end{bmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{A_D} & r_{A_{D,S_f}}\sigma_{A_D}\sigma_{A_{S_f}} & r_{A_{D,S_u}}\sigma_{A_D}\sigma_{A_{S_u}} \\ r_{A_{D,S_f}}\sigma_{A_D}\sigma_{A_{S_f}} & \sigma^2_{A_{S_f}} & r_{A_{S_f,S_u}}\sigma_{A_{S_f}}\sigma_{A_{S_u}} \\ r_{A_{D,S_u}}\sigma_{A_D}\sigma_{A_{S_u}} & r_{A_{S_f,S_u}}\sigma_{A_{S_f}}\sigma_{A_{S_u}} & \sigma^2_{A_{S_u}} \end{pmatrix}\right].$$

To produce the second generation, sires and dams of the first generation were mated at random, each sire being mated to 10 dams, and each dam producing 10 full sib offspring. Second generation breeding values for all three genetic effects were simulated as $A_i = \frac{1}{2}A_{sire} + \frac{1}{2}A_{dam} + MS_i$, where MS denotes the Mendelian sampling term. The *MS* were simulated from the multivariate normal distribution

$$\begin{bmatrix} MS_D \\ MS_{S_f} \\ MS_{S_u} \end{bmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{2}\sigma^2_{A_D} & \frac{1}{2}r_{A_{D,S_f}}\sigma_{A_D}\sigma_{A_{S_f}} & \frac{1}{2}r_{A_{D,S_u}}\sigma_{A_D}\sigma_{A_{S_u}} \\ \frac{1}{2}r_{A_{D,S_f}}\sigma_{A_D}\sigma_{A_{S_f}} & \frac{1}{2}\sigma^2_{A_{S_f}} & \frac{1}{2}r_{A_{S_f,S_u}}\sigma_{A_{S_f}}\sigma_{A_{S_u}} \\ \frac{1}{2}r_{A_{D,S_u}}\sigma_{A_D}\sigma_{A_{S_u}} & \frac{1}{2}r_{A_{S_f,S_u}}\sigma_{A_{S_f}}\sigma_{A_{S_u}} & \frac{1}{2}\sigma^2_{A_{S_u}} \end{pmatrix}\right].$$

Non-genetic effect were simulated only for the second generation, from

$$\begin{bmatrix} E_D \\ E_{S_f} \\ E_{S_u} \end{bmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{E_D} & r_{E_{D,S_f}}\sigma_{E_D}\sigma_{E_{S_f}} & r_{E_{D,S_u}}\sigma_{E_D}\sigma_{E_{S_u}} \\ r_{E_{D,S_f}}\sigma_{E_D}\sigma_{E_{S_f}} & \sigma^2_{E_{S_f}} & r_{E_{S_f,S_u}}\sigma_{E_{S_f}}\sigma_{E_{S_u}} \\ r_{E_{D,S_u}}\sigma_{E_D}\sigma_{E_{S_u}} & r_{E_{S_f,S_u}}\sigma_{E_{S_f}}\sigma_{E_{S_u}} & \sigma^2_{E_{S_u}} \end{pmatrix}\right]$$

Then we calculated the phenotypes of the individual from second generation using the full model (Equation 3). Those phenotypic values were used to estimate the variance components. We simulated hundred replicates for each set of genetic parameter and the estimates were averaged over replicates

3

Indirect genetic effects contribute substantially to heritable variation in aggression-related traits in group-housed mink (*Neovison vison*)

Setegn Worku Alemu[1,2], Piter Bijma[2], Steen Henrik Møller[3], Luc Janss[1], Peer Berg[1,4]

[1] Department of Molecular Biology and Genetics, Aarhus University, Tjele DK-8830, Denmark, [2] Animal Breeding and Genomics Centre, Wageningen University, 6700AH Wageningen, the Netherlands, [3] Department of Animal Science - Epidemiology and management, Aarhus University, Tjele DK-8830, Denmark, [4] NordGen, Nordic Genetic Resource Center, P.O.Box 115, 1431 Ås, Norway

Genetics Selection Evolution 2014, 46:30

Abstract

Since the recommendations on group housing of mink (*Neovison vison*) were adopted by the Council of Europe in 1999, it has become common in mink production in Europe. Group housing is advantageous from a production perspective, but can lead to aggression between animals and thus raises a welfare issue. Bite marks on the animals are an indicator of this aggressive behaviour and thus selection against frequency of bite marks should reduce aggression and improve animal welfare. Bite marks on one individual reflect the aggression of its group members, which means that the number of bite marks carried by one individual depends on the behaviour of other individuals and that it may have a genetic basis. Thus, for a successful breeding strategy it could be crucial to consider both direct (DGE) and indirect (IGE) genetic effects on this trait. However, to date no study has investigated the genetic basis of bite marks in mink. A model that included DGE and IGE fitted the data significantly better than a model with DGE only, and IGE contributed a substantial proportion of the heritable variation available for response to selection. In the model with IGE, the total heritable variation expressed as the proportion of phenotypic variance (T^2) was six times greater than classical heritability (h^2). For instance, for total bite marks, T^2 was equal to 0.61, while h^2 was equal to 0.10. The genetic correlation between direct and indirect effects ranged from 0.55 for neck bite marks to 0.99 for tail bite marks. This positive correlation suggests that mink have a tendency to fight in a reciprocal way (giving and receiving bites) and thus, a genotype that confers a tendency to bite other individuals can also cause its bearer to receive more bites. Both direct and indirect genetic effects contribute to variation in number of bite marks in group-housed mink. Thus, a genetic selection design that includes both direct genetic and indirect genetic

effects could reduce the frequency of bite marks and probably aggression behaviour in group-housed mink.

3.1 Introduction

Social interactions among individuals are common both in plants and animals (Frank, 2007) and can have significant effects on production and welfare traits. For example, social interactions can affect feed intake and growth rate in domestic pigs (Arango *et al.*, 2005; Chen *et al.*, 2008), lead to mortality due to cannibalism in laying hens (Muir, 1996) result in aggression and tail biting if mixing is carried out in pigs (Turner *et al.*, 2010) , increase competition in fish (Moav and Wohlfart.Gw, 1974) , affect growth rate and disease traits in forestry (Cappa and Cantet, 2008; Brotherstone *et al.*, 2011; Silva *et al.*, 2013), and result in bite marks in mink (Hansen and Damgaard, 1991; Damgaard and Hansen, 1996; Pedersen and Jeppesen, 2001; Moller *et al.*, 2003). Because social interactions may have a heritable component, selection acting on these interactions may affect significantly response to artificial selection (Griffing, 1967; Muir, 2005; Bijma *et al.*, 2007a; Bijma *et al.*, 2007b). Therefore, social interactions are a key factor when designing artificial breeding programmes in domestic animals for which group housing is common practise (Muir, 2005).

Results have shown that social interactions among individuals may create additional heritable variation(Bijma *et al.*, 2007b) .Ellen *et al.* (2008) found that, in laying hens, total heritable variation in survival days, expressed as the proportion of phenotypic variance, was 1.5 to 3-fold greater than the variance of the direct genetic effect (DGE). Wilson *et al.* (2009) reported that indirect genetic effects (IGE) increased total heritable variation, expressed as a proportion of phenotypic variance, from 0.01 to 0.6 for rearing rate and 0.05 to 0.56 for reciprocal latency rate. These results indicate that more than 80% of the heritable variation of these behavioural traits is due to social interactions (Wilson *et al.*, 2009). Therefore, for socially affected traits, the heritable variation due to social

interactions can be a significant source of heritable variation in domestic, natural, and laboratory populations, for both behavioural traits and production traits (Moore *et al.*, 1997; Muir, 2005; Bijma *et al.*, 2007b; Wilson *et al.*, 2011) and taking such interactions into account may reveal that their genetic variation is significantly greater than previously thought. However, if these interactions are competitive, the heritable variation may be significantly reduced, even to a value of zero when the direct-indirect genetic correlation equals -1 (Bijma, 2011; Wilson *et al.*, 2011). The negative covariance between direct and indirect genetic effect cancels both the direct and indirect genetic effects (Bijma, 2011; Wilson *et al.*, 2011).

With the exception of maternal genetic effects, breeders have focused on improving the direct effect of the genotype of the individual on its own phenotype (Falconer, 1960). Hence, the traditional genetic model does not include the social effect of an individual on the phenotypes of its group mates, the so-called Indirect Genetic Effect (IGE; (Griffing, 1967; Moore *et al.*, 1997) . Ignoring IGE may result in a suboptimal response to selection and even a negative response to selection for socially affected traits (Griffing, 1967). For example, individual selection to increase the size of flour beetles populations (*Tribolium castaneum*) decreased the population size in the next generations (Wade *et al.*, 2010). Similarly, in non-beak-trimmed laying hens, selection of the survivors decreased survival rate in the next generations (Muir, 2005). Thus, inclusion of IGE is vital to obtain an optimal response to selection for socially affected traits, which means that the traditional quantitative genetic model should be extended to include the heritable effect of an individual on the phenotypes of its group mates (Griffing, 1967; Muir, 2005; Bijma *et al.*, 2007b; Ellen *et al.*, 2008).

One way of using IGE for response to selection is group selection. It was shown that group selection was effective compared to individual mass selection in decreasing the mortality rate of laying hens, mainly due to aggression, from 68% in generation 2 to 9% in generation 6 (Muir, 1996) and in improving longevity of layers (Craig and Muir, 1996). Another example is the positive response for low leaf and high leaf area in *Arabidopsis thaliana* obtained with group selection versus the negative response with individual selection (Goodnight, 1985). The reason for the effectiveness of group selection is that it takes into account part of the IGE.

Although group selection is effective in reducing mortality in chickens and increasing growth in *Arabidopsis thaliana*, it uses only the between-group genetic variance and completely ignores the within-group variance. Thus, group selection is efficient only when group members are sufficiently related (Griffing, 1976b; Griffing, 1976a; Bijma, 2011). Moreover, using group selection does not provide any insight into the relative importance of direct *vs*. indirect genetic effects. It is important to understand the genetic parameters that underlie the interactions because it would help to quantify the potential contribution of IGE to response to selection, to estimate breeding values for both direct and indirect genetic effects, and to optimize breeding programmes (Bijma *et al.*, 2007a). This can be achieved by a BLUP (best linear unbiased prediction) model that separates DGE and IGE and gives weights to each of them according to the variance covariance structure of the genetic parameters (Arango *et al.*, 2005; Muir, 2005; Bijma *et al.*, 2007a).

IGE are increasingly important in European mink production because of changes in the housing system from pair-wise to group housing that is becoming more and more frequent. In the wild, juvenile mink leave the mother's territory at the age of three to four months in order to find their own territory (Birks, 1986; Dunstone, 1993) and by the end of the

growth season, their territorial behaviour is fully developed. This process of dispersal involves increased aggression between the dam and the juveniles as well as between juveniles. The male territory may overlap that of several females but is defended against mink of the same sex (Gerell, 1970; Dunstone, 1993). Therefore, in Europe during the growth season, juvenile mink are traditionally housed in pairs of one male and one female per cage. In spite of their territorial nature, recommendations on cage sizes for group housing of mink were adopted by the Council of Europe in 1999 (European Commission, 1999), probably because welfare improvements were expected from 'social enrichment' as discussed in (European Commission, 2001). Group housing has become more and more common because it increases the stocking density in the cages and thereby decreases housing investments. Group housing also increases the social dynamics of the environment which could be a potential disadvantage, since studies on animal welfare in group housing report increased aggression resulting in more bite wounds and bite marks (Pedersen and Jeppesen, 2001; Moller *et al.*, 2003; Hanninen *et al.*, 2008a; Hanninen *et al.*, 2008b).

Direct observation of aggression is time-consuming and it is difficult to distinguish between aggressions and play in mink (Hanninen *et al.*, 2008a; Hansen and Malmkvist, 2011). Thus, it is not a feasible option for collecting the required data for breeding against aggressive behaviours in mink. An alternative solution could be to record the consequences of aggressive behaviours, such as bite marks. Bite marks are the result of a hard pressure to the skin, e.g. a bite, during the 7-week growth phase of the winter coat (Hansen *et al.*, 2014) and, as such, are an excellent indicator of aggression accumulated over this period, and of reduced animal welfare (European Commission, 2001; Moller *et al.*, 2003; Hansen and Houbak, 2005). In mink, bite marks can occur anywhere on the body and are often scored on the neck, tail and all the body without neck and tail (referred to as "body" in the following) , in order to quantify different types of aggressive interactions (Damgaard and

Hansen, 1996; Hansen and Houbak, 2005). If inflicting bite marks is a genetically inherited behaviour, then genetic selection that includes both DGE and IGE may be an efficient way to reduce bite marks in group-housed mink. To date, no studies have quantified direct and indirect genetic variation for the number of bite marks in mink.

In this study, we tested the hypothesis that the number of bite marks on different parts of the body is affected by both DGE and IGE. Towards this aim, we estimated the direct and indirect additive genetic (co)variances for the number of bite marks on different regions of the body. Genetic correlations between the numbers of bite marks on different parts of the body were also estimated. Furthermore, we tested whether DGE and IGE on the bite marks in different parts of the body were related to the individual's body weight, since body weight can be an indicator of social dominance. For instance, a positive genetic correlation between body weight and IGE on bite mark number could indicate that individuals with a dominant genotype for higher weight inflict more bite marks on group mates.

3.2 Methods

3.2.1 Materials

The consequences of aggressive behaviour in mink (*Neovison vison*) can be recorded by visual observation of injuries i.e. scars on the skin of live animals or dead bodies at pelting, or by the number of bite marks on the flesh side of the skin just after fleshing during the pelting process. The number of bite marks gives an indication of the number of aggressions received by the individual over a period of time prior to pelting.

We used bite marks recorded at pelting as an indirect measure of aggressive behaviour. Bite marks were recorded just after fleshing and after scraping and brushing off sawdust. In 2009, a selection experiment was initiated to select for reduced number of bite marks at

pelting, at the mink farm at the Research Centre Foulum in Denmark. We analysed data from the first three generations of that selection experiment. A total of 1985 mink descending from 136 sires and 349 dams were used. Two weeks after weaning i.e. at around 10 weeks post-partum, the juveniles were separated into groups composed of four juvenile mink. Each group of two male siblings and two female siblings was placed in a two storey cage. These procedures were applied in 2009 and repeated in 2010 and 2011. The female siblings were unrelated to the male siblings within a cage except for the 2009 data set, but most individuals had siblings present in other cages. In some cases, data from only three or two mink was obtained mainly because of lack of pedigree information or loss of ID tags during the pelting procedure, and in few cases because of injury or death. Overall, useful data was recorded for two mink from 208 pens, for three mink from 87 pens and all four mink from 327 pens. Individuals were pelted in November 2009, December 2010 and December 2011. At pelting, the number of bite marks on the skin side of the pelt was recorded. The number of bite marks was subjectively measured based on the scale described in Table 1, and expressed as a bite mark score (BMS). From each litter, siblings of the group-housed juveniles were kept in pairs and were the selection candidates. Parents for the next generation were selected from the candidates based on the number of bite marks in their group-housed litter mates. Each individual was selected based on the performance of the mean phenotype of the litter mates' pen. Thus, the selection method takes into account both DGE and IGE.

Table 1 Bite mark score (BMS) used for subjectively measuring the number of bite marks at pelting

BMS	Number of bite marks
0	0
1	1-5
2	6-10
3	11-15
4	16-20
5	21-25
6	26-30
7	31-35
8	36-45
9	More than 45

The number of bite marks was scored on the **Neck** (from the nose tip to the shoulder/front leg), body (from the shoulder, including the front legs, to 10 cm above the base of the tail) and **Tail** (from 10 cm above the base of the tail, including the hind legs). A total score was computed as the sum of these three scores. As shown on the histogram in Figure 1, the data were not normally distributed. Log transformation after adding 100 to each observation improved the normality slightly, as illustrated by skewness and kurtosis before and after transformation (Figure 2). Table 2 summarizes BMS per sex for the different parts of the body.

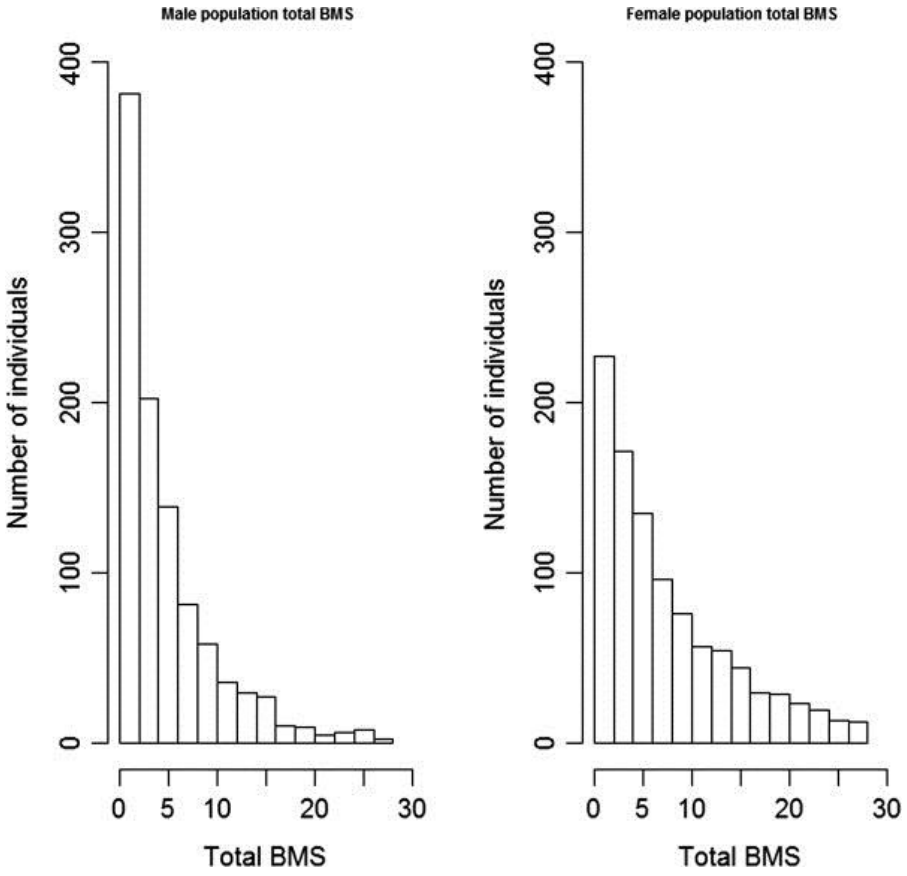

Figure 1 Histogram of residuals[1] of raw data on total BMS[2] before transformation[3]. [1]Residuals come from a model **y** = **Xb** + **e**, where fixed effects in **Xb** are identical to those used in the mixed model that is explained in the text; [2]since total BMS is the sum of BMS on the three body regions, it ranges from 0 to 27 (see Table 1); [3]for the male and female populations, skewness for total BMS corrected for fixed effects was equal to 1.67 and 1.12, respectively and kurtosis was equal to 4.14 and 1.43, respectively

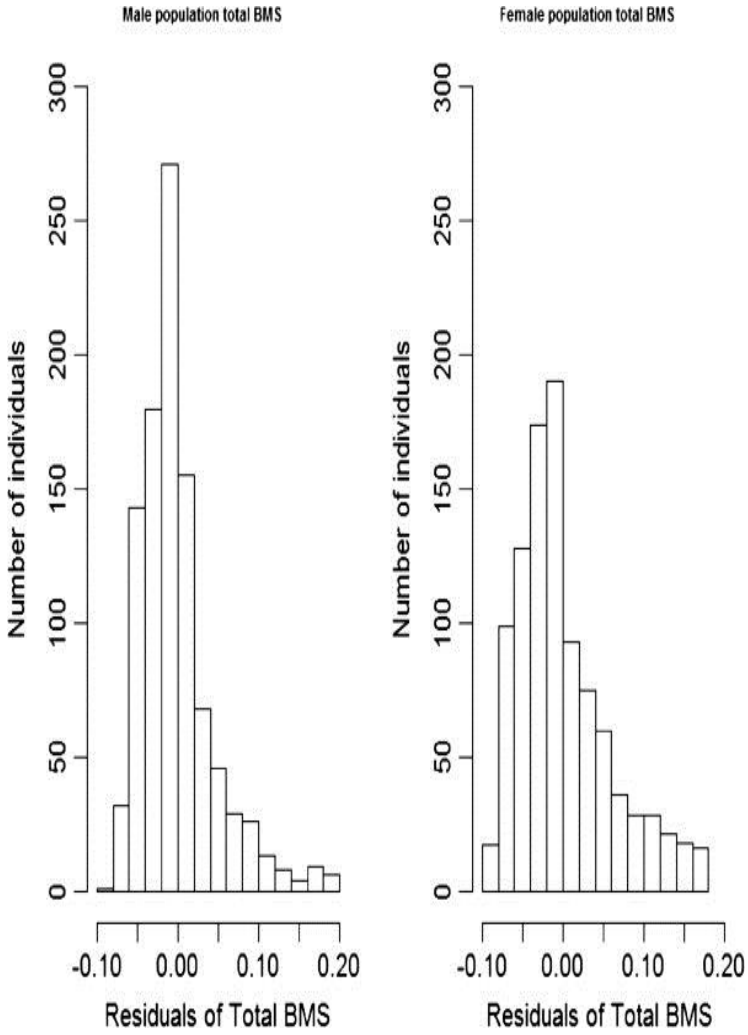

Figure 2 Histogram of residuals[1] for total BMS[2] after transformation[3]. [1]Residuals come from the model **y** = **Xb** + **e**, where fixed effects in **Xb** are identical to those used in the mixed model explained in the text; [2]since total BMS is the sum of BMS on the three body regions, it ranges from 0 to 27 (see Table 1), [3]for the male and female populations, skewness for total BMS corrected for fixed effects was equal to 1.54 and 0.96, respectively andkurtosis was equal to 3.1 and 0.43, respectively; [2]the transformation was $y_t = \ln(y + 100)$.

Table 2 Mean (standard deviation) of BMS and body weight per sex

Trait	Male	Nb individuals	Female	Nb individuals
Neck BMS	1.35 (1.62)	996	2.72 (2.33)	986
Body BMS	2.18 (2.53)	992	2.25 (2.34)	984
Tail BMS	1.54 (1.95)	992	2.91 (2.98)	984
Total BMS	5.06 (5.13)	983	7.87 (6.65)	991
Body weight (kg)	2.87 (0.41)	965	1.46 (0.24)	964

Data on BMS and weight were analysed using the GLM procedure in R (R Development Core Team). This programme was used to decide which fixed effects should be included in the model to estimate the genetic parameters. The following fixed effects i.e. year, sex, number of individuals in a cage (group size; fitted as a factor), and the linear regression on the proportion of male mates per cage (i.e., a covariate, referred to as social sex ratio) were included in the model.

Genetic parameters were estimated using residual maximum likelihood with an animal model (Henderson, 1975; Kruuk, 2004). Six models were compared with different combinations of random effects. All six models included the DGE of the individual on which the BMS was recorded. The first three models did not include IGE. The first model fitted cage as a random effect, while the second model fitted sex within cage (cage*sex) as a random effect. The reason for fitting cage*sex as a non-genetic random effect, was to test whether social interactions in mink depend on sex. This hypothesis is based on the fact that male mink are usually larger than female mink and thus aggression could occur

mainly between cage mates of the same sex. The third model included a cage plus cage*sex random effect. Each of these three models was extended with IGE, giving a total of six models. The best model was selected based on its Akaike information criterion (AIC). In all six models, we used the same fixed effects (see above). Non-genetic maternal effects (common litter effects) were not significant for BMS, and thus were not included in the models. Based on AIC, non-genetic maternal effects were included in the model for body weight. In this section, we present only the most complete model; in the simpler models the relevant terms were omitted. However, we will present results for the two models that had the highest likelihood i.e. one in which IGE were ignored and one in which IGE were included.

The most complete model (referred to as Model 6; see Table 3) was:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_D\mathbf{a}_D + \mathbf{Z}_S\mathbf{a}_S + \mathbf{Wg} + \mathbf{Vg}^*\mathbf{s} + \mathbf{e},$$

where $\mathbf{y}$ is a vector of observed BMS; $\mathbf{b}$ is a vector of fixed effects, with the incidence matrix $\mathbf{X}$ linking observations to fixed effects, $\mathbf{Z}_D$ and $\mathbf{Z}_S$ are known incidence matrices for direct DGE and IGE, and $\mathbf{a}_D$ and $\mathbf{a}_S$ are vectors of random DGE and IGE, with

$$\begin{bmatrix} \mathbf{a}_D \\ \mathbf{a}_S \end{bmatrix} \sim MVN\left(\mathbf{0}, \mathbf{G} \otimes \mathbf{A}\right), \text{in which}$$

$$\mathbf{G} = \begin{bmatrix} \sigma^2_{A_D} & \sigma_{A_{DS}} \\ \sigma_{A_{DS}} & \sigma^2_{A_S} \end{bmatrix},$$

$\otimes$ indicates the Kronecker product, and $\mathbf{A}$ is the numerator relationship matrix (Falconer, 1960; Lynch and Walsh, 1998). $\mathbf{g}$ is a vector of random cage effects and $\mathbf{W}$ the incidence

matrix linking records to cages, with $\mathbf{g} \sim N\left(0, \mathbf{I}_g \sigma_g^2\right)$, where $\mathbf{I}_g$ is an identity matrix of appropriate dimension and σ_g^2 is the cage variance, **g** * **s** a vector of random cage*sex effects and **V** an incidence matrix, with $\mathbf{g}*\mathbf{s} \sim N\left(0, \mathbf{I}_{g*s}\, \sigma_{g*s}^2\right)$, where $\mathbf{I}_{g*s}$ is an identity matrix of appropriate dimension and σ_{g*s}^2 is the variance of the cage*sex effect, and **e** is a vector of residuals. We fitted different residual variances for male and female individuals,

$$\mathbf{e} = \begin{bmatrix} \mathbf{e}_m \\ \mathbf{e}_f \end{bmatrix} \sim MVN\left(\mathbf{0}, \mathbf{C} \otimes \mathbf{I}_\mathbf{e}\right),$$

$$\mathbf{C} = \begin{bmatrix} \sigma_{e_m}^2 & 0 \\ 0 & \sigma_{e_f}^2 \end{bmatrix},$$

$$\mathbf{I}_\mathbf{e} = \begin{bmatrix} \mathbf{I}_m & 0 \\ 0 & \mathbf{I}_f \end{bmatrix},$$

where $\mathbf{e_m}$ is the vector of residuals for males, and $\mathbf{e_f}$ the vector of residuals for females, and $\sigma_{e_m}^2$ and $\sigma_{e_f}^2$ are the corresponding variances. In Model 6, the $\mathbf{Z}_S\mathbf{a}_S$ accounts for heritable indirect effects, and **Wg** and **Vg** * **s** account for covariances among cage mates due to non-heritable indirect effects (Bijma *et al.*, 2007a; Alemu *et al.*, 2014a).

A was calculated using information on five generations of pedigree that included 2806 animals. Bivariate analysis was also used to estimate the genetic correlation between bite

marks on each part of the body, and to estimate the genetic correlation between bite marks and body weight.

Table 3 Model comparisons using AIC[1]

		Neck BMS		Body BMS		Tail BMS		Total BMS	
Model	**# Param.**	**Log L**	**AIC**	**Log L**	**AIC**	**Log L**	**AIC**	**Log L**	**AIC**
1. cage	10	-24.9	47.9	-35.3	68.6	-28.0	54.0	-34.1	66.3
2. cage*sex	10	-45.8	89.6	-57.3	112.6	-36.5	70.9	-69.8	137.6
3. cage + cage*sex	11	-24.5	49.0	-33.4	66.8	-16.7	33.4	-31.1	64.0
4. IGE + cage	11	-1.4	2.7	0	0	-3.2	6.4	-0.1	0.2
5. IGE + cage*sex	11	0	0	-0.2	0.4	0	0	0	0
6.IGE + cage + cage*sex	12	0.06	1.9	0.2	1.5	-1	1.3	0.1	1.9

[1]Akaike's information criterion (AIC) and likelihood value AIC were set to zero as reference for the best model; AIC = 2× # parameters – 2 × log-likelihood; thus lower values indicate a better model. Genetic parameters for BMS were estimated by implementing the above-mentioned linear animal models in the ASReml software (Gilmour *et al.*, 2002). The matrix of additive genetic relationships

3.2.2 Heritable variation

The above model yields estimates of three genetic parameters, $\sigma^2_{A_D}$, $\sigma_{A_{DS}}$ and $\sigma^2_{A_S}$. Following Bijma (Bijma, 2011), these three parameters can be combined into a measure of the total heritable variance that determines the potential of the trait to respond to selection. Since each individual interacts with n-1 group mates, the total heritable impact of an individual's genes on trait values in the population equals:

$$A_{T_i} = A_{D_i} + (n-1)A_{S_i}, \quad (1)$$

where A_T represents the total breeding value, which is a generalization of the traditional breeding value to account for IGE. The total heritable variance is the variance of the total breeding values among individuals,

$$\sigma^2_{A_T} = \sigma^2_{A_D} + 2(n-1)\sigma_{A_{DS}} + (n-1)\sigma^2_{A_S}. \quad (2)$$

The $\sigma^2_{A_T}$ expresses the heritable variance in absolute units as the additive genetic variance in classical models. The interpretation of $\sigma^2_{A_T}$ becomes easier by expressing heritable variance relative to phenotypic variance, similarly to the classical heritability (Bergsma *et al.*, 2008):

$$T^2 = \frac{\sigma^2_{A_T}}{\sigma^2_P}.$$

A comparison of h^2 *versus* T^2 reveals the proportion of the contribution of IGE to the heritable variance that determined the potential of the population to respond to selection.

3.3 Results

Table 4 shows the estimated fixed effects and their statistical significance. The fixed effect year was significant for bite marks on all regions of the body i.e. neck, body, and tail. Sex and social sex ratio were also significant for bite marks on the neck and tail but not on body and group size was significant only for bite marks on the neck.

Table 4 Estimated fixed effects and their significance

Trait	Year	Sex[1]	Social sex ratio[2]	Group size[3]
Neck	***	1.44**	-0.20**	(0.63,0.71) *
Body	***	0.01NS	0.17NS	(1.80, 1.93) NS
Tail	***	1.81***	-0.97***	(1.89,1.67) NS
Total	***	1.2***	-0.50**	(1.53,1.54) NS
Body weight	NS	2.7***	-0.10NS	(-0.33 ,-0.31) *

*, **, *** significant at P ≤ 0.05, 0.01, 0.001, respectively; [1]the estimate for sex refers to females minus males; [2]social sex ratio represents the proportion of male group mates of an individual, fitted as a covariable, thus the estimate is the regression coefficient of bite marks on proportion of male group mates in the cage; [3]the two group size estimates refer to the difference between group size three minus group size two and group size four minus group size two, respectively.

Table 3 (see above) shows the log-likelihood values and AIC for all Models 1 through 6. Based on AIC, the best model among the six tested is Model 5 for bite marks on all regions except body for which Model 4 is slightly better. AIC values show that, in spite of the relatively small dataset, models that included IGE were substantially better than those that did not (Models 1 to 3 *vs*. 4 to 6). Thus, IGE contribute to the heritable variation of BMS on all locations of the body. Models with a random cage*sex effect were the best based on AIC, but differences in AIC between Models 4 to 6 were very small.

Table 5 shows the estimated variance components obtained with the classical model that included direct genetic effects only, but accounted for both cage and cage*sex as non-genetic random effects (Model 3). The genetic parameters were assumed to be the same in both sexes. The additive genetic variance ranged from 0.78 for neck bite marks to 1.15 for tail bite marks. Heritability ranged from 0.18 to 0.23, and differed significantly from zero. We found no common maternal effects for BMS. Non-genetic variances of BMS were

higher in females than in males, which agrees with the observation that the mean BMS for females was closer to the middle of the scale used for BMS (Tables 1 and 2).

Table 5 Estimated variance components (±SE) from a traditional animal model ignoring IGE (model 3)[1]

Parameter	Neck BMS	Body BMS	Tail BMS	Total BMS	Weight (Kg)
$\hat{\sigma}^2_A$	0.62 ± 0.15	1.06 ± 0.22	0.95 ± 0.19	7.26 ± 1.38	0.06 ± 0.015
[2] $\hat{\rho}$	0.28 ± 0.047	0.26 ± 0.04	0.17 ± 0.028	0.26 ± 0.04	-0.15 ± 0.09
[3] $\hat{\rho}_s$	0.05 ± 0.054	-0.09 ± 0.05	-0.17 ± 0.03	-0.09 ± 0.05	0.40 ± 0.19
$\hat{\sigma}^2_{e_m}$	1.18 ± 0.12	2.74 ± 0.22	2.31 ± 0.20	11.4 ± 1.14	0.026 ± 0.008
$\hat{\sigma}^2_{e_f}$	2.93 ± 0.22	3.53 ± 0.27	5.98 ± 0.34	22.4 ± 1.72	0.03 ± 0.009
[4] $\hat{\sigma}^2_P$	3.54 ± 0.11	4.95 ± 0.24	5.31 ± 0.18	31.09 ± 1.00	0.011 ± 0.005
$\hat{h}^2$	0.18 ± 0.04	0.21 ± 0.08	0.18 ± 0.036	0.23 ± 0.04	0.57 ± 0.13
$\hat{c}^2$	-	-	-	-	0.07 ± 0.05

[1]Model 3 was **y** = **Xb** + $\mathbf{Z}_D\mathbf{a}_D$ + **Wg** + **Vg** * **s** + **e**; [2]although cage and cage*sex covariances were fitted, the result is expressed as the non-genetic correlation between phenotypes of cage mates, $^2\hat{\rho} = \frac{\hat{\sigma}^2_g}{\hat{\sigma}^2_{g*s} + \hat{\sigma}^2_g + 0.5(\hat{\sigma}^2_{e_m} + \hat{\sigma}^2_{e_f})}$, and as the non-genetic correlation between phenotypes of cage mates of the same sex, $^3\hat{\rho}_s = \frac{\hat{\sigma}^2_{g*s}}{\hat{\sigma}^2_{g*s} + \hat{\sigma}^2_g + 0.5(\hat{\sigma}^2_{e_m} + \hat{\sigma}^2_{e_f})}$; [4]for BMS, phenotypic variance was estimated from a separate analysis using the model **y** = **Xb** + **e**, this was done because our objective was to present a single number for phenotypic variance and heritability, covering both sexes, since a single genetic variance was fitted covering both sexes; however, since our aim was to estimate the other model terms with the best fitting model, a separate analysis for phenotypic variance was performed; the standard errors of heritability

estimates were calculated from the full model, averaging the residual variances for both sexes; $\hat{c}^2 = \frac{\hat{\sigma}^2_{nd}}{\hat{\sigma}^2_p}$, $\hat{\sigma}^2_{n_d}$ refers to the non-genetic dam variance; $\hat{c}^2$ refers to the non-genetic maternal effect.

The estimated heritability for body weight was equal to 0.58. We found a non-significant (σ_c^2 / σ_P^2 = 0.07) common maternal environment effect for body weight. Although the effect is non-significant, it improved the fit of the model for body weight. The cage variance was significantly different from zero for bite marks on all regions of the body. (This conclusion is based on the ratio of the estimate and its SE (standard error), which was much greater than 2). Although the cage*sex-effect was not significantly different from zero for all regions of the body, it was included in the model because it improved the AIC (Table 3). Thus, both cage and cage*sex effects improved the AIC when IGE were ignored.

Table 6 shows the estimated variance components obtained with Model 5 that includes both DGE and IGE and the cage*sex effect. The standard errors on the estimated genetic variances show that both DGE and IGE contributed to variation in BMS. IGE were significantly different from zero for bite marks on all regions of the body and variance of IGE ranged from 0.14 for tail bite marks to 0.27 for body bite marks. The total heritable variation for BMS ranged from 1.65 to 19.13, and was significantly higher than the additive genetic variance obtained with the traditional model. The total heritable variation expressed as the proportion of phenotypic variance ranged from 0.41 to 0.61, and was ~6 times greater than the direct heritability. The correlation between DGE and IGE of BMS ranged from 0.55 to 0.99 in all parts of the body. Comparison of heritability estimates in Tables 5 and 6 (h^2 and h_D^2) indicates that ordinary heritability estimated with the traditional model overestimates the importance of direct effects by a factor of ~2. In the traditional model, presence of IGE biases the estimate of additive genetic variance upwards. This occurs because cage mates are partly related and thus, an individual receives an IGE from its cage mates that is similar to its own IGE. This in turn increases the

covariance between relatives in different cages, which biases heritability upwards (Peeters *et al.*, 2013).

Table 6 Estimated variance components (±SE) for both direct effect and indirect effects using Model 5[1]

Parameter	Neck BMS	Body BMS[5]	Tail BMS	Total BMS
$\hat{\sigma}^2_{A_D}$	0.26 ± 0.11	0.37 ± 0.14	0.34 ± 0.13	2.95 ± 0.90
$\hat{\sigma}_{A_{D,S}}$	0.12 ± 0.04	0.27 ± 0.05	0.21 ± 0.04	1.97 ± 0.30
$\hat{\sigma}^2_{A_S}$	0.18 ± 0.04	0.27 ± 0.06	0.14 ± 0.04	1.6 ± 0.32
[2] $\hat{\sigma}^2_{AT}$	1.65 ± 0.25	2.56 ± 0.56	2.19 ± 0.30	19.13 ± 2.40
$\hat{r}_{A_{DS}}$	0.55 ± 0.22	0.67 ± 0.21	0.99 ± 0.23	0.90 ± 0.15
[3] $\hat{\rho}_s$	0.09 ± 0.05	-0.04 ± 0.04	-0.09 ± 0.03	-0.02 ± 0.04
$\hat{\sigma}^2_{e_m}$	1.40 ± 0.12	3.15 ± 0.21	2.80 ± 0.18	14.8 ± 1.01
$\hat{\sigma}^2_{e_f}$	3.07 ± 0.20	3.90 ± 0.25	6.10 ± 0.32	24.77 ± 1.54
[4] $\hat{\sigma}^2_P$	3.54 ± 0.11	4.95 ± 0.14	5.31 ± 0.16	31.09 ± 1.00
[5] $\hat{h}^2_D$	0.07 ± 0.10	0.07 ± 0.03	0.06 ± 0.02	0.10 ± 0.03
[6] $\hat{T}^2$	0.47 ± 0.08	0.52 ± 0.21	0.41 ± 0.06	0.61 ± 0.08

[1]Model 5 was **y** = **Xb** + $\mathbf{Z}_D\mathbf{a}_D$ + $\mathbf{Z}_S\mathbf{a}_S$ + **Vg** * **s** + **e**; [2]from Equation 2 using a pen size of 3.18; [3] $\hat{\rho}_s = \frac{\hat{\sigma}^2_{g*s}}{\hat{\sigma}^2_{g*s} + 0.5(\hat{\sigma}^2_{e_m} + \hat{\sigma}^2_{e_f})}$ is the non-genetic correlation between phenotypes of cage mates of the same sex; [4] for BMS, phenotypic variance was estimated from a separate analysis using the model **y** = **Xb** + **e**, this was done because our objective was to present a single number for phenotypic variance and heritability, covering both sexes since a single genetic variance was fitted covering both sexes; however, since our aim was to estimate the other model terms with the best fitting model, a separate analysis for phenotypic variance was performed; the standard errors of heritability estimates were calculated from the full model, averaging the residual variances for both sexes; [5]although Model 4 was slightly better, we presented estimates obtained with Model 5 for reasons of consistency; [5] $\hat{h}^2_D = \frac{\hat{\sigma}^2_A}{\hat{\sigma}^2_p}$ [6] $\hat{T}^2 = \frac{\hat{\sigma}^2_A}{\hat{\sigma}^2_p}$

Table 7 shows the genetic correlations between BMS on different regions of the body and body weight. Genetic correlations were positive for all bite mark correlations (direct-direct, direct-indirect, and indirect-indirect). Since the bivariate analysis of total BMS with BMS at specific regions of the body failed to converge, total BMS was removed from Table 7.

However, there were small negative genetic correlations between direct effects on BMS and body weight, and between indirect effects on BMS and body weight, some of which were significantly different from zero. Hence, there is a weak indication that heavier individuals are less likely to get involved in aggressive interactions.

Table 7 Genetic correlation estimates (±SE) between bite mark scores[1] at different parts of the body and with body weight

			Direct			Indirect	
		Weight[2]	**Neck BMS**	**Body BMS**	**Tail BMS**	**Neck BMS**	**Body BMS**
Direct	Neck BMS	-0.29 ± 0.17					
	Body BMS	-0.08 ± 0.17	0.48 ± 0.22				
	Tail BMS	0.21 ± 0.16	0.57 ± 0.22	0.57 ± 0.22			
Indirect	Neck BMS	-0.05 ± 0.10	0.55 ± 0.22	0.78 ± 0.19	0.89 ± 0.16		
	Body BMS	-0.10 ± 0.10	0.52 ± 0.19	0.67 ± 0.21	0.68 ± 0.22	0.78 ± 0.19	
	Tail[3] BMS	-0.17 ± 0.10	0.60 ± 0.23	0.85 ± 0.24	0.99 ± 0.23	0.96 ± 0.21	0.99 ± 0.27

[1]The analysis for total BMS did not converge and was thus omitted from this Table; [2] there was no evidence for IGE on body weight, thus, weight refers to the direct effect only; genetic correlation of direct total BMS vs. direct weight was equal to -0.28 ± 0.13 and indirect total BMS vs. direct weight was equal to -0.15 ± 0.08; correlation of TBV of total BMS with body weight was - 0.21 ± 0.08.

3.4 Discussion

We have provided evidence that BMS is a heritable trait, and thus can be changed by selective breeding. We found that both DGE and IGE contribute to genetic variation of BMS on all regions of the body. IGE contributed a significant proportion of the heritable variation available for response to selection ($\sigma^2_{A_T}$). The contribution of IGE variance to total heritable variation, measured by the ratio $(n-1)^2\sigma^2_{A_S} / \sigma^2_{A_T}$, ranged from 30% for tail bite

marks to 52% for neck bite marks, while that of DGE variance was about 16% for all regions of the body. Moreover, there was a strong positive correlation between DGE and IGE, which further increased total heritable variance. Thus, most of the heritable variation in BMS relates to IGE. For instance, for total BMS, the variance in IGE and the direct-indirect genetic covariance together contributed 85% of the heritable variation. Estimated genetic correlations between direct and indirect genetic effects were strong and positive and ranged from 0.55 to 0.99, i.e. significantly different from zero, except for bite marks in the neck region. Thus, these results suggest that if a genotype causes an individual to bite more, it also leads the individual to be more bitten, which, in turn, suggests that an individual benefits from not harming others.

Regarding the non-genetic random effects, the cage*sex effect fitted the data better than the cage effect (except for Body BMS). Ignoring cage*sex effects may cause bias in the estimates of the genetic parameter, which has been reported in previous studies using both simulated (Van Vleck and Cassady, 2005) and real data (Bijma *et al.*, 2007a). Without fitting cage*sex effects, the estimated variance in both the DGE and the IGE was about 7% lower in our data, indicating a minor effect. This makes sense since the cage*sex effect was not very significantly different from zero.

Both cage and cage*sex effects improved the AIC of the traditional model (Model 3). In contrast, when IGE were included in the model, cage effect did not improve the fit of the model. This suggests that IGE are included in the cage variance when they are not accounted for. We included a cage*sex random effect to allow for stronger interactions between individuals of the same sex within a cage (this was expected based on knowledge of behaviour in mink) (Gerell, 1970; Birks, 1986; Dunstone, 1993). Such within-sex interactions might lead to systematic similarities or dissimilarities between cage mates of

the same sex. Although we fitted cage and cage*sex as covariances, the results are presented as non-genetic correlations between cage mates and between cage mates of the same sex, for ease of interpretation. The cage*sex correlation was close to zero for all parts of the body (Table 6). This result indicates that the non-genetic direct-indirect correlation is close to zero, since the expected value of ρ_s is calculated as:

$$\rho_s = \frac{2\sigma_{E_{DS}} + (n_{\text{sex}} - 2)\sigma^2_{E_S}}{\sigma^2_{E_D} + (n-1)\sigma^2_{E_S}} = \frac{2\sigma_{E_{DS}}}{\sigma^2_{E_D} + 2.18\sigma^2_{E_S}},$$

where $\sigma_{E_{DS}}$ is the covariance between direct and indirect non-genetic effects, $\sigma^2_{E_D}$ is the direct environmental variance, $\sigma^2_{E_S}$ is the indirect environmental variance, n is the number of individuals in a cage, and n_{sex} is the number of individuals of the same sex in a cage, which on average was equal to 2 in our data. Thus, in contrast to the clearly positive direct-indirect genetic correlation ($r_{A_{DS}}$, Table 6), the non-genetic direct-indirect correlation was practically zero.

Given the strong positive direct-indirect genetic correlation, it is surprising that the non-genetic direct-indirect correlation is near zero. However, in our data, group mates of the same sex were full sibs. Thus, the cage*sex correlation not only represents the non-genetic correlation between group mates of the same sex, but also between full sibs and those correlations are fully confounded in our data. The kin selection theory predicts that sibs show less competitive interactions (Hamilton, 1964), which agrees with observations reported for pigs, where members of the same family fight less compared to unrelated individuals (Stookey and Gonyou, 1998; Li and Johnston, 2009). Hence, the apparent difference between the genetic and non-genetic correlations between direct and indirect effects may be due to the fact that information on the non-genetic correlation depends

completely on interactions between siblings in our data. The estimated direct-indirect genetic correlation, in contrast, includes interactions among non-kin.

By including the cage*sex correlation, we have, at least partly, accounted for non-genetic-indirect effects that depend on relatedness. However, the indirect genetic effects may also differ between kin and non-kin. Hence, estimated parameters for DGE and IGE may depend on group composition with respect to relatedness. This has proven to be a complex issue that we will explore in a future study.

The direct-direct genetic correlations for BMS on different regions of the body were positive (Table 7), which suggests that an individual that is less bitten on one part of its body is also likely to be less bitten on the other parts of its body. The direct-indirect genetic correlations for BMS on different regions of the body were also positive, which indicates that an individual that is less bitten on one part of its body is less likely to bite other parts of the body of its cage mates. Finally, the indirect-indirect genetic correlations for BMS were also positive, which implies that an individual that bites more or less one part of the body of its cage mates will also bite more or less the other parts of the body of its cage mates. We also investigated the genetic correlations between weight and direct and indirect effects on BMS, but found no significant correlations. Thus, selecting for increased size (larger pelts) animals, which implies increased weight, is not expected to lead to more biting.

Our findings suggest that it is possible to select mink that have a considerably lower level of biting. Irrespective of the selection strategy, response to selection is always equal to the product of the intensity of selection, the accuracy of selection, and the standard deviation of total heritable variation, $R = i\rho\sigma_{A_T}$ (Bijma, 2011). For instance, for total BMS, σ_{A_T} is equal to 4.36 and the mean of total BMS is equal to 6.47, which means that the current

total level is only 1.48 genetic standard deviation away from zero. Even with a low accuracy and a moderate intensity, we can produce mink that have a significantly lower level of biting. For instance, with mass selection for total bite marks, which would require recording BMS on live animals, the accuracy is (Ellen *et al.*, 2007)

$$\rho_{T,IS} = \frac{r\sigma^2_{A_T} + (1-r)\left[\sigma^2_{A_D} + (n-1)\sigma_{A_{DS}}\right]}{\sigma_{A_T}\sigma_P},$$

which equals ~0.4 based on our estimates. Then, if 10% of the population is used for breeding to have an intensity of selection equal to 1.76, the predicted response to selection will be equal to ~3.07 and the total BMS is predicted to reduce from ~6.47 to ~ 3.4, which is a very substantial reduction in a single generation of selection. When using group selection for groups of four sibs, two males and two females that all belong to the same family, it is possible to reach an even higher accuracy i.e. ~ 0.65, and thus the predicted response to selection will be ~5. Using sib selection, which is more appropriate for bite marks since they are recorded on the pelts of dead individuals, the predicted accuracy will be equal to ~ 0.54 and the response to selection to ~4.14. Thus, total BMS will be reduced from ~6.47 to ~ 3.33, again a very substantial reduction in a single generation of selection. In 2011, on average, the difference in total BMS between the selected and control lines was 4.5 in both sexes, which is in reasonable agreement with the range of predicted responses. Thus, although in practice response to selection is usually lower than the theoretical predicted value, our results indicate that it is possible to select mink that have a considerably lower level of biting in a few generations.

3.5 Conclusion

In summary, we confirm the hypothesis that both DGE and IGE contribute to variation in number of bite marks in group-housed mink. Since IGE contribute a substantial amount of heritable variation, genetic selection can reduce bite marks and possibly aggressive behaviour in group-housed minks. Including IGE in selection designs would ensure a more efficient selection against bite marks.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

PBe initiated the research. SWA analysed the data, wrote the draft and final paper. PBi provided guidance to analyse the data, interpret the output, and helped to write the paper. PBe and SHM collected the data and helped to write the paper. LJ helped to write the paper. All authors read and approved the final manuscript.

Acknowledgments

We acknowledge financial support for PBi by the Dutch Technology Foundation (STW), which is part of the Netherlands Organization for Scientific Research (NWO), and support from the Danish Ministry of Food, Agriculture and Fisheries, Program Innovationsloven. We also acknowledge Alastair Wilson for his insightful comments.

References

Alemu SW, Berg P, Janss L, Bijma P (2014a). Indirect genetic effects and kin recognition: estimating IGEs when interactions differ between kin and strangers. *Heredity* **112:** 197-206.

Arango J, Misztal I, Tsuruta S, Culbertson M, Herring W (2005). Estimation of variance components including competitive effects of Large White growing gilts. *J Anim Sci* **83:** 1241-1246.

Bergsma R, Kanis E, Knol EF, Bijma P (2008). The contribution of social effects to heritable variation in finishing traits of domestic pigs (*Sus scrofa)*. *Genetics* **178:** 1559-1570.

Bijma P (2011). A general definition of the heritable variation that determines the potential of a population to respond to selection. *Genetics* **189:** 1347-1359.

Bijma P, Muir WA, Van Arendonk JAM (2007b). Multilevel selection 1: Quantitative genetics of inheritance and response to selection. *Genetics* **175:** 277-288.

Bijma P, Muir WM, Ellen ED, Wolf JB, Van Arendonk JAM (2007a). Multilevel selection 2: Estimating the genetic parameters determining inheritance and response to selection. *Genetics* **175:** 289-299.

Birks J (1986). *The mammals Society: Oswestry Shropshire*.

Brotherstone S, White IMS, Sykes R, Thompson R, Connolly T, Lee S *et al.* (2011). Competition Effects in a Young Sitka Spruce (Picea sitchensis, Bong. Carr) Clonal Trial. *Silvae Genet* **60:** 149-155.

Cappa EP, Cantet RJC (2008). Direct and competition additive effects in tree breeding: Bayesian estimation from an individual tree mixed model. *Silvae Genet* **57:** 45-56.

Chen CY, Kachman SD, Johnson RK, Newman S, Van Vleck LD (2008). Estimation of genetic parameters for average daily gain using models with competition effects. *J Anim Sci* **86:** 2525-2530.

Craig JV, Muir WM (1996). Group selection for adaptation to multiple-hen cages: Beak-related mortality, feathering, and body weight responses. *Poultry Science* **75:** 294-302.

Damgaard BM, Hansen SW (1996). Stress physiological status and fur properties in farm mink placed in pairs or singly. *Acta Agriculturae Scandinavica Section a-Animal Science* **46:** 253-259.

Dunstone N (1993). *The mink. T &A.D. Poyster, London*.

Ellen ED, Muir WM, Teuscher F, Bijma P (2007). Genetic improvement of traits affected by interactions among individuals: sib selection schemes. *Genetics* **176:** 489-499.

Ellen ED, Visscher J, van Arendonk JAM, Bijma P (2008). Survival of laying hens: Genetic parameters for direct and associative effects in three purebred layer lines. *Poultry Science* **87:** 233-239.

European Commission (1999). Standing committee of the European convention for the protection of animals kept for farming purposes: Recommendation concerning fur animals. European Commission; 1999. http://.www.coe.int/.

European Commission (2001). Scientific Committee on Animal Health and Animal Welfare: The welfare of animals kept for fur production. European Commission;. http://ec.europa.eu/food/animal/welfare/international/out67_en.pdf.

Falconer DS, Mackay TFC (1996). *Introduction to Quantitative Genetics*, Ed 4. Longmans Green, Harlow, Essex, UK.

Frank SA (2007). All of life is social. *Current Biology* **17:** R648-R650.

Gerell R (1970). Home ranges and movements of the mink mustela vison Shreber in Southern Sweden. *Oikos* **21:** 160–173.

Gilmour AR, Gogel BJ, Cullis BR, Thomson R (2006). *ASReml user guide release 2.0*. VSN International Ltd. Hemel Hempstead,UK.

Goodnight CJ (1985). The influence of environmental variation on group and individual selection in a cress. *Evolution* **39:** 545-558.

Griffing B (1967). Selection in reference to biological groups .I. Individual and group selection applied to populations of unordered groups. *Australian Journal of Biological Sciences* **20:** 127-139.

Griffing B (1976a). Selection in Reference to Biological Groups .4. Use of Extreme Forms of Nonrandom Groups to Increase Selection Efficiency. *Genetics* **82:** 723-731.

Griffing B (1976b). Selection in Reference to Biological Groups .5. Analysis of Full-Sib Groups. *Genetics* **82:** 703-722.

Hamilton WD (1964). Genetical Evolution of Social Behaviour *J Theor Biol* **7:** 1-16.

Hanninen S, Ahola L, Pyykonen T, Korhonen HT, Mononen J (2008a). Group housing in row cages: an alternative housing system for juvenile mink. *Animal* **2:** 1809-1817.

Hanninen S, Mononen J, Harjunpaa S, Pyykonen T, Sepponen J, Ahola L (2008b). Effects of family housing on some behavioural and physiological parameters of juvenile farmed mink (Mustela vison). *Applied Animal Behaviour Science* **109:** 384-395.

Hansen, S., and B. Houbak, 2005 Hansen SW, Houbak BT: To skridt frem og tre tilbage – gruppeind-husning af mink. In Annual 2004 Report of the Danish Fur Breeders Research Center. Edited by Sandbøl P. ; 2005:39–4

Hansen, S., and J. Malmkvist, 2011 Fodringsmæssige tiltag til begrænsning af bidmærker hos mink holdt i grupper - foreløbige resultater. In Temadag om aktuel minkforskning. Intern rapport nr. 109, september 2011. Aarhus Universitet. Edited by Peer B. ; 2011:19–34. ISBN 978-87-91949-92-0
http://dca.au.dk/fileadmin/DJF/Arrangementer/temadage_mink/temadag_mink_2011.pdf

Hansen SW, Damgaard B (1991). Effect of Environmental-Stress and Immobilization on Stress Physiological Variables in Farmed Mink. *Behav Process* **25:** 191-204.

Hansen SW, Moller SH, Damgaard BM (2014). Bite marks in mink-Induced experimentally and as reflection of aggressive encounters between mink. *Appl Anim Behav Sci* **158:** 76-85.

Henderson CR (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics* **31:** 423-447.

Kruuk LEB (2004). Estimating genetic parameters in natural populations using the 'animal model'. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **359:** 873-890.

Li YZ, Johnston LJ (2009). Behavior and performance of pigs previously housed in large groups. *Journal of Animal Science* **87:** 1472-1478.

Lynch M, Walsh B (1998). *Genetics and analysis of quantitative traits*. Sinauer: Sunderland, MA.

Moav R, Wohlfart.Gw (1974). Magnification through Competition of Genetic Differences in Yield Capacity in Carp. *Heredity* **33:** 181-202.

Moller SH, Hansen SW, Sorensen JT (2003). Assessing animal welfare in a strictly synchronous production system: The mink case. *Animal Welfare* **12:** 699-703.

Moore AJ, Brodie ED, Wolf JB (1997). Interacting phenotypes and the evolutionary process .1. Direct and indirect genetic effects of social interactions. *Evolution* **51:** 1352-1362.

Muir WM (1996). Group selection for adaptation to multiple-hen cages: selection program and direct responses. *Poultry Science* **75:** 447-458.

Muir WM (2005). Incorporation of competitive effects in forest tree or animal breeding programs. *Genetics* **170:** 1247-1259.

Pedersen V, Jeppesen LL (2001). Effects of family housing on behaviour, plasma cortisol and performance in adult female mink (Mustela vison). *Acta Agriculturae Scandinavica Section a-Animal Science* **51:** 77-88.

Peeters K, Ellen ED, Bijma P (2013). Using pooled data to estimate variance components and breeding values for traits affected by social interactions. *Genet Sel Evol* **45:27**.

R Development Core Team. (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Silva JCE, Potts BM, Bijma P, Kerr RJ, Pilbeam DJ (2013). Genetic control of interactions among individuals: contrasting outcomes of indirect genetic effects arising from neighbour disease infection and competition in a forest tree. *New Phytol* **197:** 631-641.

Stookey JM, Gonyou HW (1998). Recognition in swine: recognition through familiarity or genetic relatedness? *Appl Anim Behav Sci* **55:** 291-305.

Turner SP, D'Eath RB, Roehe R, Lawrence AB (2010). Selection against aggressiveness in pigs at re-grouping: practical application and implications for long-term behavioural patterns. *Animal Welfare* **19:** 123-132.

Van Vleck LD, Cassady JP (2005). Unexpected estimates of variance components with a true model containing genetic competition effects. *Journal of Animal Science* **83:** 68-74.

Wade MJ, Bijma P, Ellen ED, Muir W (2010). Group selection and social evolution in domesticated animals. *Evol Appl* **3:** 453-465.

Wilson AJ, Gelin U, Perron MC, Reale D (2009). Indirect genetic effects and the evolution of aggression in a vertebrate system. *P Roy Soc B-Biol Sci* **276:** 533-541.

Wilson AJ, Morrissey MB, Adams MJ, Walling CA, Guinness FE, Pemberton JM *et al.* (2011). Indirect genetics effects and evolutionary constraint: an analysis of social dominance in red deer, Cervus elaphus. *J Evolution Biol* **24:** 772-783.

4

Estimation of indirect genetic effects in group-housed mink (*Neovison vison*) should account for systematic interactions either due to kin or sex.

S.W. Alemu[1,3], P. Berg[1,2], Luc Janss[1], and P. Bijma[3]

[1]Department of Molecular Biology and Genetics, Aarhus University, Tjele DK-8830, Denmark, [2]NordGen, Nordic Genetic Resource Center, P.O.Box 115, 1431 Ås, Norway, [3]Animal Breeding and Genomics Centre, Wageningen University, 6700AH, Wageningen, The Netherlands

J.Anim.Breed.Genet.(2015)1–8

Abstract

Social interactions among individuals are abundant, both in wild and domestic populations. With social interactions, genes of an individual may affect the trait values of other individuals, a phenomenon known as Indirect Genetic Effects (IGEs). IGEs can be estimated using linear mixed models. Current IGE-models assume that individuals interact equally to all group mates irrespective of relatedness. Kin selection theory, however, predicts that an individual will interact differently with family members vs. non-family members. Here we investigate kin and sex–specific social interactions in group-housed mink. Furthermore, we investigated whether IGEs depend on relatedness between interacting individuals or on their sex. In conclusion, our results indicate that male mink show different non-genetic interactions than female mink. Moreover, we have shown how estimates from a family-based model can be translated to the ordinary direct-indirect model, and vice versa. We find no evidence for genetic differences in interactions among related versus unrelated mink.

Key words: Indirect genetic effect, kin recognition, social interaction, mink

4.1 Introduction

Social interactions are common in plants and animals, and are caused by mechanisms such as limited resources, social dominance, or maternal effects (Frank, 2007). Because of social interactions an individual's genes may affect trait values of its social partners, a phenomenon known as Indirect Genetic Effects (IGE, (Griffing, 1967; Moore *et al.*, 1997).

IGEs can be estimated using linear mixed models, by fitting a direct genetic effect for the individual producing the phenotype record, and an IGE for each of its social partners (Arango *et al.*, 2005; Muir, 2005; Bijma *et al.*, 2007a). However, the current models assume that individuals express the same IGE to each of their social partners. This assumption may not be correct when there are systematic differences between group mates. Individuals may, for example, interact differently with family members *vs.* strangers, or with males *vs.* females (Alemu *et al.*, 2014a).

An individual may interact differently with family vs. non-family members because of kin recognition. Kin recognition is a preferential behavior to family members or familiar individuals compared to unrelated or unfamiliar individuals. Kin recognition occurs for instance, in social insects (Hepper, 1986), in Blanding's ground squirrels and Richardson's ground squirrels (Sheppard and Yoshida, 1971; Holmes and Sherman, 1982), in fish (OLSEN 1989; BROWN and BROWN 1996; OLSEN *et al.* 1998), in large mammals, such as pigs (Stookey and Gonyou, 1998; Li and Johnston, 2009), and in plants (Biedrzycki and Bais, 2010). Thus kin recognition is common in both animals and plants (Holmes and Sherman, 1982; Hepper, 1986; Olsen, 1989; Biedrzycki and Bais, 2010).

4 Indirect effects differ due to sex or kin in mink

Differential interactions with different sexes may result from sexual selection. Sexual selection is natural selection arising through preference of one sex for certain characteristics in individuals of the other sex (Darwin, 1874; Thom *et al.*, 2004). In mink, there is a clear difference in size between males and females, which may arise from sexual selection (Darwin, 1874; Thom *et al.*, 2004; Zuffi *et al.*, 2011). Thus, interactions between members of the same sex may be more important than between members of different sexes. Therefore, when estimating indirect genetic effects it is important to take in to account differential interactions due to kin or sex.

Alemu *et al.* (2014b) found substantial IGEs for bite mark traits in group-housed mink. They found a strongly positive genetic correlation between direct and indirect effects, but a near zero environmental correlation between direct and indirect effects on full-sib group mates of the same sex. These results suggest that biting behaviour in mink may depend on relatedness or sex of the individuals. Alemu *et al.* (2014b) however, did not investigate the genetics of kin or sex specific interactions. They did consider sex-specific non-genetic interactions by fitting a covariance between group mates of the same sex. However, this covariance was assumed to be the same for both sexes, suggesting that interactions between males are similar to those between females.

In a theoretical study, Alemu *et al.* (2014a) proposed a model to investigate kin-specific genetic interactions, which distinguished between genetic effects on kin (including self) and strangers (referred to as the "reduced model" in the following). They showed that total breeding values and indirect genetic effects on strangers can be estimated using either the reduced model or the traditional direct-indirect model. When interactions depend on kin, however, the traditional model gives biased estimates for the direct genetic variance and

the direct-indirect genetic correlation. Alemu *et al.* (2014a) did not clarify the relationship between estimates from both models, nor how estimates from the traditional model can be converted to the reduced model.

Here we investigate kin and sex–specific social interactions in group-housed mink. We apply both the traditional direct-indirect model, and the reduced model of Alemu *et al.* (2014a), and show how estimates of both models are related. We also investigate the presence of sex or kin-specific non-genetic interactions in mink, by allowing for a sex-dependent non-genetic covariance between group mates. We discuss the interpretation of such effects, and the need to fit them to avoid bias in the genetic estimates.

4.2 Materials and methods

4.2.1 Materials

Aggressive behavior in mink results in bite marks (EC 2001; Moller *et al.*, 2003; Hansen and Houbak, 2005). Those bite marks can be recorded by visual observation of injuries or scars on the skin on live animals, or on the dead bodies at pelting, or as the number of bite marks on the flesh side of the skin, just after fleshing during the pelting process. We used bite marks recorded at pelting as an indirect measure of the aggressive behavior an individual received over its life time.

Bite marks were collected on mink that were part of a selection experiment started in 2009 at the mink farm at Research Centre Foulum in Denmark, with the objective to reduce the number of bite marks. We analysed data from the first three generations of this experiment. A total of 1969 mink descending from 136 sires and 349 dams were used in our analysis. Two male siblings and two female siblings were placed in 2 storey cages in

years 2009, 2010, and 2011. The female siblings were unrelated to the male siblings within the same cage, but most individuals had siblings present in another cage. The number of records on some of the cages was reduced to three or two, mainly because of loss of id, and partly due to injury or death. Overall we had data from only 2 of the 4 mink from 212 cages, from 3 mink from 85 cages, and from all 4 mink from 325 cages.

Individuals were pelted in November 2010, December 2011, and December 2012. At pelting, the number of bite marks was recorded at the skin side of the pelt. The number of bite marks was subjectively scored on the scale described in Table 1, and expressed as a bite mark score (BMS). Bite marks were scored in the **Neck** (from nose tip to shoulder/front leg), **Body** (from shoulder to 10 cm above the base of the tail) and **Tail** region (from 10 cm above the base of the tail, incl. back legs). Total BMS was computed as the sum of these three scores. We log-transformed the data after adding 100 to each observation, which improved the normality slightly (Alemu *et al.*, 2014b).

Table 1 Bite mark score (BMS) used for subjectively measuring the number of bite marks at pelting

BMS	Number of bite marks
0	0
1	1-5
2	6-10
3	11-15
4	16-20
5	21-25
6	26-30
7	31-35
8	36-45
9	More than 45

Table 2 Mean (standard deviation) of BMSs per sex[1]

Trait	Males	Females
Neck BMS	1.34 (1.60)	2.72 (2.32)
Body BMS	2.16 (2.32)	2.26 (2.54)
Tail BMS	1.53 (1.95)	2.92 (2.97)
Total BMS	5.02 (5.10)	7.88 (6.66)

[1]The number of records on males was 991, and the number of records on females was 978.

4.2.2 Statistical models

First the data were analysed using the GLM procedure in R, to decide which fixed effects should be included. The fixed effects of year, sex, number of individuals in a cage (group size), and the linear regression on the proportion of male cage mates, referred to as the social sex effect, were included in the model. Next, genetic parameters were estimated with residual maximum likelihood and animal models (Henderson, 1975), using ASReml (Gilmour *et al.*, 2002). For all models, the matrix of additive genetic relationships, **A**, was calculated using information on five generations of pedigree, including a total of 2,806 animals.

Following the aim of this work, we compared the traditional direct-indirect genetic model to the reduced model proposed by Alemu *et al.* 2014a, and investigated the need for sex-specific non-genetic covariances between cage members of the same sex. We compared four models in total.

Model 1 was a traditional animal model with IGE and a random group effect, as proposed by (Muir, 2005; Bijma *et al.*, 2007b),

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_D\mathbf{a}_D + \mathbf{Z}_S\mathbf{a}_S + \mathbf{Wg} + \mathbf{e}, \qquad \text{(Model 1)}$$

,

where $\mathbf{Z}_D$ and $\mathbf{Z}_S$ are known incidence matrices for direct and indirect genetic effects, respectively, and $\mathbf{a}_D$ and $\mathbf{a}_S$ are vectors of random direct and indirect genetic effects, with

$$\begin{bmatrix} \mathbf{a}_D \\ \mathbf{a}_S \end{bmatrix} \sim MVN\left(\mathbf{0}, \mathbf{G} \otimes \mathbf{A}\right),$$

in which

$$\mathbf{G} = \begin{bmatrix} \sigma^2_{A_D} & \sigma_{A_{DS}} \\ \sigma_{A_{DS}} & \sigma^2_{A_S} \end{bmatrix},$$

$\otimes$ indicates the Kronecker product, and **A** is the numerator relationship matrix (Falconer, 1960; Lynch and Walsh, 1998). This model assumes that an individual expresses the same IGE on all its cage mates, irrespective of their sex and family. The $\mathbf{g}$ is a vector of random cage effects, with $\mathbf{g} \sim N\left(\mathbf{0}, \mathbf{I}_g\sigma^2_g\right)$, where $\mathbf{I}_g$ is an identity matrix of the appropriate dimension and σ^2_g is the cage effect variance, **W** is an incidence matrix linking records to cages, and $\mathbf{e}$ is a vector of residuals. We fitted different residual variances for male and female individuals,

$$\mathbf{e} = \begin{bmatrix} \mathbf{e}_m \\ \mathbf{e}_f \end{bmatrix} \sim MVN\left(\mathbf{0}, \mathbf{E}\right),$$

$$\mathbf{E} = \begin{bmatrix} \mathbf{I}_m\,\sigma^2_{e_m} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_f\,\sigma^2_{e_f} \end{bmatrix},$$

where $\mathbf{e}_m$ is the vector of residuals for males, $\mathbf{e}_f$ the vector of residuals for females, $\sigma^2_{e_m}$ and $\sigma^2_{e_f}$ the corresponding variances, and $\mathbf{I}_m$ and $\mathbf{I}_f$ are identity matrices of the appropriate dimensions.

Model 2 contained a non-genetic covariance between cage mates of the same sex, rather than a single covariance between all cage mates,

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_D\mathbf{a}_D + \mathbf{Z}_S\mathbf{a}_S + \mathbf{Vk} + \mathbf{e} \quad \text{(Model 2)}$$

,

where **k** is a vector of random cage*sex effects and **V** an incidence matrix for sex*cage, with

$\mathbf{k} \sim N(\mathbf{0}, \mathbf{I}_k \sigma_k^2)$, where $\mathbf{I}_k$ is an identity matrix of the appropriate dimension and σ_k^2 is the variance of the cage*sex effect. The model term **Vk** accounts for non-genetic covariances between cage mates of the same sex, *i.e.*, between the siblings in a cage. (In the data, siblings in the same cage were always either both male, or both female, see above). In Model 2, this covariance has the same magnitude for both sexes. All other elements are the same as in Model 1. Model 2 is identical to the best model of Alemu *et al.* (2014b; Model 5 with results in their Table 6).

Model 3 partitioned the non-genetic covariance between cage mates into a cage-effect common to all cage mates, and a sex-specific covariance of different magnitude for each sex,

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_D\mathbf{a}_D + \mathbf{Z}_S\mathbf{a}_S + \mathbf{Wg} + \mathbf{Vk} + \mathbf{e} \quad \text{(Model 3)}$$

.

The genetic terms in Model 3 are the same as in Model 1. The **Vk** term represents a non-genetic random effect for cage members belonging to the same sex (and thus the same family), with a separate variance for males and for females. Thus **V** is the incidence matrix for sex*cage, as in Model 2, and **k** is a vector of random effects common to the two family

members in the same cage, with $\mathbf{k} \sim N(\mathbf{0}, \mathbf{K} \otimes \mathbf{I}_g)$, where $\mathbf{I}_g$ is an identity matrix with dimensions equal to the number of cages. When data are ordered by sex within cage,

$$\mathbf{K} = \begin{bmatrix} \sigma^2_{k_m} & 0 \\ 0 & \sigma^2_{k_f} \end{bmatrix},$$

subscript m denoting males and f denoting females. Other elements were the same as in Model 1.

The $\sigma^2_{k_m}$ and $\sigma^2_{k_f}$ essentially represent the non-genetic covariance between cage mates of the same sex, and can therefore take negative values. To facilitate the interpretation of these estimates, we expressed them as non-genetic correlations between cage mates of the same sex,

$$\hat{\rho}_m = \frac{\hat{\sigma}^2_{k_m}}{\hat{\sigma}^2_g + \hat{\sigma}^2_{k_m} + \hat{\sigma}^2_{e_m}}$$

$$\hat{\rho}_f = \frac{\hat{\sigma}^2_{k_f}}{\hat{\sigma}^2_g + \hat{\sigma}^2_{k_f} + \hat{\sigma}^2_{e_f}}$$

Hence, $\hat{\rho}_m$ and $\hat{\rho}_f$ measure the non-genetic similarity of male cage mates and female cage mates, respectively, on top of an overall similarity of cage mates due to the random cage effect. The variance of the cage effect, $\hat{\sigma}^2_g$, represents the non-genetic covariance among cage mates of different sex, and will also be expressed as a correlation in the results,

$$\hat{\rho} = \frac{\hat{\sigma}^2_g}{\hat{\sigma}^2_g + \frac{1}{2}(\hat{\sigma}^2_{k_m} + \hat{\sigma}^2_{k_f}) + \frac{1}{2}(\hat{\sigma}^2_{e_m} + \hat{\sigma}^2_{e_f})}$$

The denominator of this expression is the average of the non-genetic variance in males and females.

Model 4 was the reduced model of Alemu *et al.* (2014a) with respect to genetic terms, which partitions genetic effects into a component due to the family (including the focal individual) and a component due to strangers, and includes the same non-genetic effects as Model 3,

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z}_D\mathbf{a}_F + \mathbf{Z}_S^u\mathbf{a}_{S_u} + \mathbf{Wg} + \mathbf{Vk} + \mathbf{e} \qquad \text{(Model 4)}$$

where $\mathbf{Z}_D$ is the incidence matrix for direct genetic effects, identical to the $\mathbf{Z}_D$ in Model 1, and $\mathbf{Z}_S^u$ is a known incidence matrix for indirect genetic effects of cage mates belonging to the other family (hence, subscript *u* indicates "unrelated"), and $\mathbf{a}_F$ and $\mathbf{a}_{S_u}$ are vectors of random family breeding values and indirect genetic effects on unrelated individuals, respectively, with

$$\begin{bmatrix} \mathbf{a}_F \\ \mathbf{a}_{S_u} \end{bmatrix} \sim MVN\left(\mathbf{0}, \mathbf{G}_r \otimes \mathbf{A}\right),$$

in which

$$\mathbf{G}_r = \begin{bmatrix} \sigma^2_{A_F} & \sigma_{A_{FS_u}} \\ \sigma_{A_{FS_u}} & \sigma^2_{A_{S_u}} \end{bmatrix}.$$

Thus Model 4 does not explicitly include the IGE of the sibling of the focal individual; this effect is captured by the family breeding value $\mathbf{a}_F$. The family breeding value captures the sum of the direct genetic effect of the focal individual itself, and the indirect genetic effect of its sibling (Alemu *et al.*, 2014a).

Model 4 is equivalent to Model 3 in terms of likelihood, but partitions the genetic effects differently. With kin or sex specific interactions, genetic estimates from Models 1 through 3

are a mix of genetic effects on the same family and on strangers, whereas Model 1 groups genetic effects by family (Alemu *et al.*, 2014a). With respect to the non-genetic model terms, a comparison of Model 2 to either Model 3 or 4 will test whether individuals interact systematically different depending on their sex or family relationship to cage mates.

Table 3 Model comparison using likelihood and AIC[1]

		Neck BMS		Body BMS		Tail BMS		Total BMS	
Model	# Param.	Log L	AIC	Log L	AIC	Log L	AIC	Log L	AIC
1	12	-17.99	31.98	-5.77	7.54	-10.14	16.28	-5.43	6.86
2	12	-16.16	28.32	-5.61	7.22	-6.87	9.66	-4.98	5.96
3	14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4[2]	14	-0.93	1.86	-5.65	11.30	-2.14	4.28	-3.40	6.80

[1]AIC and likelihood value of best model according to AIC was set to zero as reference. AIC = 2× # parameters – 2 × log-likelihood; thus smaller values indicate a better model. [2]Model 3 and 4 are statistically equivalent in theory. Differences in likelihood originate probably from deviations of normality.

Heritable variation

For Models 1, 2 and 3, total heritable variation available for response to selection is given by (Bijma *et al.* 2007a)

$$\sigma^2_{A_T} = \sigma^2_{A_D} + 2(n-1)\sigma_{A_{DS}} + (n-1)^2\sigma^2_{A_S} .$$

For Model 4, total heritable variation is given by Alemu *et al.* (2014a)

$$\sigma^2_{A_T} = \sigma^2_{A_F} + n\sigma_{A_{FS_u}} + \tfrac{1}{4}n^2\sigma^2_{A_{S_u}} .$$

For all models, total heritable variation was expressed relative to phenotypic variance, using

$$T^2 = \frac{\sigma^2_{A_T}}{\sigma^2_P} ,$$

to facilitate easy comparison with ordinary heritability (Bergsma *et al.* 2008).

Relationship between the reduced and the traditional IGE model

Traditional IGE-models (Models 1 through 3) yield estimates of the direct and indirect genetic (co)variances, $\sigma^2_{A_D}$, $\sigma_{A_{DS}}$ and $\sigma^2_{A_S}$. The reduced IGE-model (Model 4), in contrast, yields estimates of the family-variance, $\sigma^2_{A_F}$, the variance of IGE on unrelated individuals, $\sigma^2_{A_{S_u}}$, and their covariance, $\sigma_{A_{FS_u}}$. The relationship between those estimates can be derived using Equations 12 and 16 of Alemu *et al.* (2014a). The result shows that estimates for the reduced model can be calculated from those of the traditional model, using

$$\sigma^2_{A_F} = \sigma^2_{A_D} + (n-2)\sigma_{A_{DS}} + (\tfrac{1}{2}n-1)^2\sigma^2_{A_S}$$

$$\sigma_{A_{FS_u}} = \sigma_{A_{DS}} + (\tfrac{1}{2}n-1)\sigma^2_{A_S}$$

$$\sigma^2_{A_{S_u}} = \sigma^2_{A_S} ,$$

and vice versa,

$$\sigma^2_{A_D} = \sigma^2_{A_F} - (n-2)\sigma_{A_{FS_u}} + (\tfrac{1}{2}n-1)^2\sigma^2_{A_{S_u}}$$

$$\sigma_{A_{DS}} = \sigma_{A_{FS_u}} - (\tfrac{1}{2}n-1)\sigma^2_{A_{S_u}}$$

$$\sigma^2_{A_S} = \sigma^2_{A_{S_u}} .$$

Therefore, when the non-genetic effects in the model are the same, all genetic parameters of the reduced model are linear combination of the traditional IGE model, and vice versa. Furthermore, the likelihoods of both models are identical (Alemu *et al.*, 2014). Thus, the reduced model and traditional IGE models are statistically equivalent. We validated the equivalency of both models using simulated data, and found exact agreement (Results not shown).

4.3 Results and Discussion

Mean BMS were higher in females than in males, for all body regions (Table 2). The fixed year effect was significant for all traits. Effects of sex and group size were significant for bite mark traits, except for body bite marks. Effects of social sex were significant for tail and total bite mark traits.

Models 3 & 4 were statistically superior over Models 1 & 2 when compared based on likelihood and AIC (Table 3). Thus the model term for sex-specific non-genetic interactions was statistically significant. This result indicates that non-genetic interactions among male siblings differ from non-genetic interactions among female siblings in mink. Hence, estimation of genetic parameters for group-housed mink should take such systematic interaction into account.

Though Models 3 & 4 are theoretically equivalent in terms of likelihood, they produced somewhat different likelihoods. When data were simulated under multivariate normality, however, they produced identical likelihoods (Alemu *et al.*, 2014b). Hence, the difference in likelihood is probably due to deviations of the data from normality (See Alemu *et al.* 2014b for histograms of residuals).

Table 4 shows estimated parameters from Model 4. Estimated total heritable variance from Model 4, $\hat{\sigma}^2_{A_T}$, was ~22% smaller than the corresponding estimate from Model 2 (15.00 vs. 19.13; estimates from Model 2 are shown in Table 6 of Alemu *et al.* 2014b). This suggests that a model ignoring the sex-specific interaction among cage mates may result in overestimated genetic variance. Previous studies also showed that estimation of genetic parameters for indirect effect is sensitive to non-genetic terms in the model. Van Vleck and Cassady (2005) observed this in simulated data, and Bijma *et al.* (2007b) in data on mortality in laying hens. When we omitted both the cage and the cage*sex random effects, the estimated total heritable variation for total BMS was 83 % higher than with Model 4.

Table 4 Estimated parameters from Model 4

Parameter	Neck BMS	Body BMS	Tail BMS	Total BMS
$\hat{\sigma}^2_{A_F}$	0.52±0.14	0.93±0.22	0.75±0.18	5.20±1.12
$\hat{\sigma}_{A_{F,S_u}}$	0.24±0.06	0.40±0.1	0.25±0.06	2.17±0.47
$\hat{\sigma}^2_{A_{S_u}}$	0.15±0.04	0.27±0.07	0.11±0.04	1.13±0.31
$\hat{r}_{A_{F,S_u}}$	0.86±0.13	0.80±0.10	0.86±0.15	0.89±0.10
[1] $\hat{\sigma}^2_{A_T}$	1.67±0.38	2.91±0.0.59	1.81±0.40	15.00±2.95
$\hat{\rho}$	0.05±0.04	0.03±0.06	0.06±0.04	0.08±0.05
$\hat{\rho}_m$	0.00±0.04	-0.10±0.07	-0.13±0.05	-0.07±0.06
$\hat{\rho}_f$	0.40±0.05	0.08±0.08	0.01±0.08	0.20±0.06
$\hat{\sigma}^2_{e_m}$	1.30±0.12	3.12±0.26	2.66±0.21	11.21±1.1
$\hat{\sigma}^2_{e_f}$	2.30±0.19	3.57±0.29	5.68±0.41	16.6±1.43
[2] $\hat{\sigma}^2_P$	3.50±0.11	4.95±0.16	5.29±0.17	28.00±1.00
$\hat{T}$ [2]	0.48±0.14	0.59±0.14	0.33±0.80	0.54±0.11

[1]We used n = 3.18 to calculate the TBV. [2]Phenotypic variance was calculated in a separate analysis using the model $\mathbf{y} = \mathbf{Xb} + \mathbf{e}$. The reason is that we wanted a single number for phenotypic variance and heritability, covering both sexes, because also a single genetic variance was fitted covering both sexes. As we didn't want to estimate the other model terms assuming same residual variance for both sexes, we fitted a separate model for phenotypic variance.

The estimated genetic correlation between the family breeding value and the IGE on unrelated individuals, r_{A_F,S_u}, was strongly positive. For total BMS, the estimate was very similar to the estimated direct-indirect genetic correlation from Model 2, which does not distinguish between kin and non-kin (0.89 in Table 4 *vs.* 0.90 in Table 6 of Alemu *et al.* 2014b). Hence, this result suggests that there is little difference in genetic interactions among kin *vs.* non-kin in mink. Thus, though such differences may be expected based on kin selection and/or selection theory, we find no indications for them.

The random group-effect, as measured by the non-genetic correlation between cage mates, $\hat{\rho}$, was not significantly different from zero. For neck BMS, there was a clear difference between the non-genetic correlation between male *vs.* female cage mates; $\hat{\rho}_m = 0.00 \pm 0.04$ whereas $\hat{\rho}_f = 0.40 \pm 0.04$. This result suggests that females fight in a reciprocal way, resulting in a similar number of neck bite marks in both female cage members. As total BMS is the sum of BMS of the three body regions, this difference for the neck region also resulted in different correlations for total BMS. Overall, this result shows that female cage mates tend to be similar for non-genetic reasons, whereas there was no such similarity for male cage mates.

It is an integral part of mating behaviour of males to bite (and thus hold) females in the neck. Non-genetic differences in mating behaviour among males would result in non-genetic correlations among females, whereas genetic differences in mating behaviour would have resulted in genetic differences in interactions (Robert 1952).

In conclusion, our results indicate that male mink show different non-genetic interactions than female mink, and that ignoring this difference may inflate estimated genetic variance. Moreover, we have shown how estimates from a family-based model can be translated to

the ordinary direct-indirect model, and vice versa. We find no indications for genetic differences in interactions among related versus unrelated mink.

References

Alemu SW, Berg P, Janss L, Bijma P (2014a). Indirect genetic effects and kin recognition: estimating IGEs when interactions differ between kin and strangers. *Heredity* **112:** 197-206.

Alemu SW, Bijma P, Moller SH, Janss L, Berg P (2014b). Indirect genetic effects contribute substantially to heritable variation in aggression-related traits in group-housed mink (Neovison vison). *Genetics, selection, evolution : GSE* **46:** 30.

Arango J, Misztal I, Tsuruta S, Culbertson M, Herring W (2005). Estimation of variance components including competitive effects of Large White growing gilts. *J Anim Sci* **83:** 1241-1246.

Biedrzycki ML, Bais HP (2010). Kin recognition in plants: a mysterious behaviour unsolved. *J Exp Bot* **61:** 4123-4128.

Bijma P, Muir WA, Van Arendonk JAM (2007a). Multilevel selection 1: Quantitative genetics of inheritance and response to selection. *Genetics* **175:** 277-288.

Bijma P, Muir WM, Ellen ED, Wolf JB, Van Arendonk JAM (2007b). Multilevel selection 2: estimating the genetic parameters determining inheritance and response to selection. *Genetics* **175:** 289-299.

Darwin C (1874). *The descent of man, and selection in relation to sex*. A. L. Burt: New York,.

Scientific Committee on Animal Health and Animal Welfare: *The welfare of animals kept for fur production*. European Commission; 2001.

http://ec.europa.eu/food/animal/welfare/international/out67_en.pdf

Falconer DS, Mackay TFC (1996). *Introduction to Quantitative Genetics*, Ed 4. Longmans Green, Harlow, Essex, UK.

Frank SA (2007). All of life is social. *Current Biology* **17:** R648-R650.

Gilmour AR, Gogel BJ, Cullis BR, Thomson R (2006). *ASReml user guide release 2.0*. VSN International Ltd. Hemel Hempstead,UK.

Griffing B (1967). Selection in reference to biological groups .I. Individual and group selection applied to populations of unordered groups. *Australian Journal of Biological Sciences* **20:** 127-139.

Hansen, S. W., B. T. Houbak (2005) To skridt frem og tre tilbage – gruppeind-husning af mink. Faglig Årsberetning 2004, Pelsdyr-erhvervets Forsøgs- og Forsknings-Center. In. pp. 29-47.

Henderson CR (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics* **31:** 423-447.

Hepper PG (1986). Kin recognition: functions and mechanisms. A review. *Biol Rev Camb Philos Soc* **61:** 63-93.

Holmes WG, Sherman PW (1982). The ontogeny of kin recognition in two species of ground-squirrels. *Am Zool* **22:** 491-517.

Li YZ, Johnston LJ (2009). Behavior and performance of pigs previously housed in large groups. *Journal of Animal Science* **87:** 1472-1478.

Lynch M, Walsh B (1998). *Genetics and analysis of quantitative traits*. Sinauer: Sunderland, Mass.

Moller SH, Hansen SW, Sorensen JT (2003). Assessing animal welfare in a strictly synchronous production system: The mink case. *Animal Welfare* **12:** 699-703.

Moore AJ, Brodie ED, Wolf JB (1997). Interacting phenotypes and the evolutionary process .1. Direct and indirect genetic effects of social interactions. *Evolution* **51:** 1352-1362.

Muir WM (2005). Incorporation of competitive effects in forest tree or animal breeding programs. *Genetics* **170:** 1247-1259.

Olsen KH (1989). Sibling Recognition in Juvenile Arctic Charr, Salvelinus-Alpinus (L). *J Fish Biol* **34:** 571-581.

Sheppard DH, Yoshida SM (1971). Social behavior in captive Richardson´s ground squirrels. *J Mammal* **52:** 793-799.

Stookey JM, Gonyou HW (1998). Recognition in swine: recognition through familiarity or genetic relatedness? *Appl Anim Behav Sci* **55:** 291-305.

Thom MD, Harrington LA, Macdonald DW (2004). Why are American mink sexually dimorphic? A role for niche separation. *Oikos* **105:** 525-535.

Zuffi, M. A. L., Sacchi, R., Pupin, F., Cencetti, T. (2011) Sexual size and shape dimorphism in the Moorish gecko (Tarentola mauritanica, Gekkota, Phyllodactylidae). *NorthWestern Journal of Zool* 7, 189

5

Genomic prediction of survival time in a population of brown laying hens showing cannibalistic behaviour

S.W. Alemu[1,2],MPL.Calus[1], William M.Muir[3], Katrijn peeters[1,4], Addie Vereijken[4], and P.Bijma[1]

[1]Animal Breeding and Genomics Centre, Wageningen University, Wageningen, Netherlands, [2]Department of Molecular Biology and Genetics, Aarhus University, Tjele DK-8830, Denmark, [3]Department of Animal Science, Purdue University, West Lafayette, IN, USA, [4]Hendrix Genetics, Research and Technology Centre, Boxmeer, Netherlands

To be submitted

Abstract

Mortality due to cannibalism is an economic and welfare problem in laying hens. Beak trimming and genetic selection are two strategies to reduce mortality and to increase survival time. Genetic selection becomes more efficient by taking into account heritable variation originating from social interaction. Social interactions may lead to so-called Indirect Genetic Effects (IGE), which are heritable effects of an individual on trait values of others. Though there is a considerable heritable variation in survival time when the contribution of IGE is included, genetic improvement of survival time in laying hens is still challenging for the following reasons. The heritability of the trait even with IGE is still limited, ranging from 0.06 to 0.26, and the individuals that are still alive at the end of the recording period are censored. Furthermore, survival time records are available late in life and only on females. Thus, we need new genetic tool such as genomic selection to cope with these challenges. Here we tested the hypothesis that genomic selection increases the accuracy of estimating the breeding value compared with parental average and the response to selection for the survival time compared with a traditional breeding scheme in two lines of brown layers showing cannibalism. We also tested the hypothesis that the rate of inbreeding per generation for genomic selection is lower compared with the traditional breeding scheme. Genetic parameters and breeding values were estimated using residual maximum likelihood with an animal model, using the programme BLUPF90. Response to selection and rate of inbreeding were predicted using the programme SelAction. The total genetic standard deviations in genomic EBVs were around 20 days for both lines, indicating good prospects for selection against mortality due to cannibalism in these brown layer lines. The accuracy of the EBV was increased by 35% when we included the marker information, when compared to the parent average EBV in both lines. The response to selection using genomic information was substantially higher than with the traditional breeding scheme for line B1 and slightly higher for line BD. The higher response in B1 line

for genomic selection is due to higher linkage disequilibrium. The predicted rate of inbreeding per generation was substantially lower for genomic selection than for the traditional breeding scheme for both lines. Our results show that genomic selection is a promising tool for the improvement *of* socially affected traits.

Key words: socially affected trait, survival time, ssGBLUP, genomic selection, response to selection

5.1 Introduction

Mortality due to cannibalism is an economic and welfare problem in laying hens. As a consequence, survival time is reduced (Blokhuis and Wiepkema, 1998). Beak trimming and genetic selection are two strategies to reduce mortality and to increase survival time. Genetic selection has been implemented to increase survival time; however, responses to selection have been limited, partly because the heritability of the trait is low (around 0.02 to 0.10) which leads to low accuracy (Ellen *et al.*, 2008). Moreover, survival in laying hens showing cannibalistic interactions depend on social interactions among cage mates, and this interaction may have a heritable component (Griffing, 1967; Muir, 1996; Muir, 2005; Bijma *et al.*, 2007a). Ignoring the heritable components due to social interactions contributes to the low accuracy and low response to selection, and may even cause a negative response to selection (Griffing 1967).

Recently, genetic selection methods have become more efficient by taking into account the additional heritable variation created by social interactions among cage mates (Craig and Muir, 1996; Muir, 1996; Arango *et al.*, 2005; Ellen *et al.*, 2008). In laying hens showing cannibalism, social interactions increase the heritable variation two to five times compared to the classical direct additive genetic variance (Ellen *et al.*, 2008; Peeters *et al.*, 2012). Though there is a considerable heritable variation in survival time when the contribution of IGE is included, genetic improvement of survival time in laying hens is still challenging for two reasons. First, heritable variation of the trait even with IGE is still low (proportion of total heritable variation to phenotypic variation 0.06 to 0.26), and some of the individuals are censored (Ellen *et al.*, 2008; Peeters *et al.*, 2012). Second, and more important, survival time records are available later in life and available only in females. Furthermore, breeding females are kept in single bird cage which makes that for survival time own performance for females is not available. Thus, selection of females is based on

pedigree and sib information which leads to limited accuracy. We have to rely on information of relatives for males, mainly on progeny information which leads to long generation interval for males. Consequently, the response to selection for survival time will be low. Thus, we need a new genetic tool such as genomic selection to increase response to selection.

Currently, breeding programmes for laying hens are also changing from progeny testing to genomic selection and this is mainly for response in egg number. The response to selection per year for egg number using genomic selection is expected to be higher than progeny testing. A relevant question is whether this will also work for survival time in cannibalistic laying hens is when using genomic selection compared with the traditional breeding scheme. In the following "traditional breeding scheme" refers sires are selected based on progeny testing and dams are selected based on sib information and pedigree information.

Genomic selection is a genetic selection method in which genetic markers covering the whole genome are used so that all quantitative trait loci are in linkage disequilibrium at least with one single nucleotides polymorphism (SNP) (Meuwissen *et al.*, 2001). Genomic selection increases the response to selection compared with traditional breeding for the following reasons. Genomic selection may increase the accuracy of estimating the breeding values, particularly when compared with the parent-average EBV (Gonzalez-Recio *et al.*, 2009; Hayes *et al.*, 2009b; Daetwyler *et al.*, 2010a) and genomic selection can reduce the generation interval compared to e.g. schemes based on progeny testing. Thus, genomic selection schemes may result in higher response to selection per year compared with traditional breeding programs. For instance, in dairy cattle the genetic gain is increased by a factor of ~2 (Schaeffer, 2006).

Genomic selection can be implemented using GBLUP that uses a relationship matrix derived from a genome-wide markers (Zhang *et al.*, 2007; VanRaden, 2008). In this method, we estimate the genomic breeding values using individuals that are both genotyped and phenotyped. However, we often do not have genotypes for all phenotyped individuals. Exploiting these non-genotyped but phenotyped individuals is one of the challenges of genomic selection. Recently, single-step genomic BLUP (SSGBLUP) has been developed. This procedure combines the relationship matrix derived from the pedigree (**A**) and from genome wide markers (G) into a single relationship matrix (**H**) that allows for genomic selection in a single step (ssGBLUP; (Legarra *et al.*, 2009; Aguilar *et al.*, 2010; Christensen and Lund, 2010). The accuracy of predicting breeding values with correct blending for ssGBLUP is higher than with GBLUP and BLUP (Christensen *et al.*, 2012). In conclusion, when the population consists of a substantial number of non-genotyped but phenotyped individuals, ssGBLUP is a promising method to use as it utilizes all available information.

Genetic selection with IGE coupled with genomic information may increase the survival time for layers compared with pedigree-BLUP. This can be tested by comparing the accuracy of estimated breeding values for survival time using genomic selection vs. pedigree-BLUP. More importantly, it can be tested by comparing the responses to selection and rates of inbreeding using genomic selection vs. a traditional breeding scheme. The aim of this paper is therefore to investigate whether genomic selection increases the accuracy of estimating breeding values compared with pedigree-BLUP for survival time in two crossbred brown layers, with the data currently available. Furthermore, we investigate whether genomic selection increases response to selection per year compared with a traditional breeding scheme. Finally, we investigate whether genomic selection reduces the rate of inbreeding per generation compared with a traditional breeding scheme.

5.2 Materials and Methods

5.2.1 Population and pedigree

Data were provided by the Institute de Sélection Animale B.V., the layer breeding division of Hendrix Genetics. Sires were mated to approximately 8 dams and each dam produced ~five males and ~five females offspring. 19,755 crossbred laying hens had B1 as sire line, and 10,910 had BD as sire line (the crosses and their number are given in Table 1). Of each crossbred individual, only the sire ID was recorded; dam IDs were unknown.

Table1: Number of individuals per different combination of crosses

cross(♂x♀)	Number of individuals	Line
B1 × BA	3,570	B1
B1× BB	1,270	B1
B 1× BD	5,735	B1
B1× BE	1,365	B1
B1× BF	4,715	B1
B1× BH	3,100	B1
BD× B1	790	BD
BD× B5	5,415	BD
BD× B6	4,705	BD

Post-hatching, the chicks were wing-banded, sexed, and vaccinated for infectious bronchitis and Marek's disease. Their beaks were kept intact. At approximately 17 weeks of

age, each batch was placed in a different laying house. Five paternal half sibs were placed in a cage. Nine commercial crosses were produced.

The trait of interest, “survival time”, was defined as “the number of days from the start of the laying period till either death or the end of the experiment”. We have different censoring moments for different batches and it is therefore necessary to set a limit (Table 2). We avoided to set either the youngest age group to the maximum number of survival time or taking the oldest age group as the maximum number of survival time. In both cases, many individuals will be censored or removed. For example for BD line in Table 2 if we take 351 as cut of point then all individuals with survival time >= 351 will be censored which is more than 80% of individuals are censored. On the other hand if we take 413 as cut of point, we need to remove the other two batches because those individuals will have unknown survival time which means we lose88 % of the data. Therefore, for BD line we took 372 as cutting point so that 57% of the individual are censored, and we lose 45% of the data which is optimal compared with the other cutting points. We did the same for B1 line. Finally in BD line we removed cages which do not have five individuals. For BD line we have about 70 cages which had four individuals and 11 cages which had 3 individuals, and we removed these cages. For line B1 all cages had five individuals.

Table 2: Number of individuals (No ind) censored at different censoring points for line BD

Batch	No ind	No indi>=351	No indi>=372	No ind >= 413
201042	5122	2461	0	0
201182	4786	2859	2774	0
2009191	6437	3805	3705	3567

5.2.2 Genomic data

Quality control

Genotypes were available on part of the sires. Out of 509 B1 sires, 207 were genotyped, and out of 284 BD sires 234 were genotyped, both with 60k SNP chips. The following quality controls were undertaken in the programme. Markers were excluded with call rate ≤0.90 or, with a minor allele frequency ≤2%. (Based on cross validation, this gave the highest accuracy of EBVs compared with values of 0.01, 0.03, 0.04, and 0.05). SNPS with a deviation of $\chi^2 \geq 600$ from Hardy-Weinberg equilibrium were excluded. Overall, we had a total of about 35k SNP satisfying all the criteria in both lines.

Data analysis

To decide which fixed effects should be included in the model to estimate the genetic parameters, data on survival time were analysed using GLM procedure in R. The fixed effects of batch, cross and the interaction term for each laying house*row*level were fitted.

We fitted a traditional sire model, with a "direct" sire effect only. Data on both sire lines were analysed separately. Because cages consist of full sibs, the sire effect captures the total sire effect, rather than the classical (direct) sire effect. In other words, when cages are composed of relatives, the EBV from an ordinary sire model is an estimate of the total sire effect, including both direct and indirect genetic effects, and the estimated additive genetic variance is an estimate of the total genetic variance, rather than of the classical direct additive genetic variance. This is explained in detail in (Peeters *et al.*, 2013). The model was

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Wg} + \mathbf{Zu} + \mathbf{e},$$

where $\boldsymbol{y}$ is a vector of survival times, $\mathbf{X}$ is the incidence matrix for fixed effects, $\mathbf{b}$ is the vector of fixed effects including cross, batch, and an interaction term for each laying house*row*level, which was included to correct for infrastructural effects (e.g., differences in light intensity). The $\mathbf{g}$ is a vector of random group effects, with $\mathbf{g} \sim N\left(0, \mathbf{I}_g \sigma_g^2\right)$, $\mathbf{I}_g$ is an identity matrix of the appropriate dimension, $\mathbf{W}$ an incidence matrix for cage effect , σ_g^2 is the group variance, and $\mathbf{Z}$ is an incidence matrix for additive sire genetic effect, $\mathbf{u}$ is vector of sire effects and $\mathbf{e}$ is a vector of residuals.

Model 1 was pedigree-BLUP, where the breeding value was assumed normally distributed as: $\mathbf{u} \sim N\left(0, \mathbf{A}\sigma_u^2\right)$, where $\mathbf{A}$ is the genetic relationship matrix derived using the pedigree information and five generations of pedigree was included to calculate **A.** where σ_u^2 is the additive genetic sire variance

Model 2 was SSGBLUP, where the breeding value was assumed normally distributed as $\mathbf{u} \sim N(0, \mathbf{H}\sigma_u^2)$, **H** is the relationship matrix that combines both pedigree and marker based genomic relationships,

$$\mathbf{H} = \begin{bmatrix} \mathbf{G}_\omega & \mathbf{G}_\omega\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \\ \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{G}_\omega & \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{G}_\omega\mathbf{A}_{11}^{-1}\mathbf{A}_{12} + \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \end{bmatrix},$$

where $\mathbf{A}_{11}$ is the sub-matrix of the pedigree based relationship matrix ($\mathbf{A}$) for genotyped animals only, $\mathbf{A}_{22}$ is the sub-matrix of $\mathbf{A}$ for non-genotyped animals, $\mathbf{A}_{12}$ and $\mathbf{A}_{21}$ are the sub matrices of $\mathbf{A}$ for the relationship between genotyped and non-genotyped animals, $\mathbf{G}_\omega = (1-\omega)\mathbf{G} + \omega\mathbf{A}_{11}$ where ω is the weight which is the default value of the software PreGSf90 (0.95)(Aguilar *et al.*, 2011)

The inverse of the $\mathbf{H}$ is (Aguilar *et al.*, 2010; Christensen and Lund, 2010)

$$\mathbf{H}^{-1} = \begin{bmatrix} \mathbf{G}_\omega^{-1} - \mathbf{A}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \mathbf{A}^{-1}$$

G and **H** are calculated using preGSf90 (Aguilar *et al.*, 2011). The allele frequencies of the current population were used for the calculation of **G**. **G** and **A** are constructed to have the same base population. It is constructed using, mean diagonal of (**G**) equals mean diagonal of **($\mathbf{A_{11}}$)** and mean off diagonal of (**G**) equals mean off diagonal of **($\mathbf{A_{11}}$)** (Aguilar *et al.*, 2011). Genetic parameters and breeding values were estimated using residual maximum likelihood with an animal using the programme BLUPF90 (Misztal, 1997; Misztal, 2013).

5.2.3 Cross validation

We considered two scenarios for breeding value estimation. In the first scenario both genotyped and non-genotyped sires were used as reference population. In the second scenario only genotyped sires were used as reference population. We compared the accuracy of predicted breeding values for the two models, pedigree-BLUP and ssGBLUP using cross validation. We randomly sampled approximately twenty percent of the genotyped sires (n = 207 for B1 and n = 234 for BD) without replacement, to produce five mutually exclusive validation data sets. For each validation data set, the remaining 80% of the data set served as a training data set. The breeding values for all individuals were estimated using the training data set.

The use of cross validation requires observed phenotypes on the individuals in the validation data set. However, part of the individuals were censored, and records were available on offspring, whereas EBV were predicted on sire. Therefore, the following steps were undertaken to estimate the Spearman correlation between true breeding value and estimated breeding value.

First the observed phenotypes were adjusted for fixed effects, using a linear model $(\mathbf{y} = \mathbf{Xb} + \mathbf{e})$. The residuals from this model served as the corrected phenotypes. Second the corrected phenotypes were sorted, so that the rank of censored individual followed by the rank of uncensored individuals. For example, assume we have 100 individuals with 30% censoring. Then the phenotype is known for 70 individuals, but the phenotype is unknown for the 30 individuals that were still alive at the end of the experiment. The individuals with known phenotype have known rank of 1 to 70, while the individuals with unknown phenotype will have rank of 71 to 100 in unknown order. We assumed that the rank of censored individuals was in random order. Then the expected rank correlation can be calculated by substituting the unknown ranks by the mean rank of censored individuals, in this case 85.5 (Ellen *et al.*, 2010). Third, the rank of the sire was calculated as the mean

rank of its daughters. Forth, the correlation between the estimated breeding values of sires and the mean corrected rank of the offspring of the sires was calculated ($\rho_{\hat{A}_s,\bar{P}_{off}}$). This procedure was repeated 5 times, once for each validation set. We calculated the standard error of the correlation for each validation using the following equation, $SE(\hat{r}) = \frac{1-\hat{r}^2}{sqrt(n)}$, where n refers to the number of individuals in the validation set and $\hat{r}^2$ refers to the correlation between the estimated breeding values of sires and the mean corrected rank of the offspring (Stuart and Ord, 1994). The standard error over the total of the five sets was calculated using bivariate analysis of mean rank of corrected phenotypes for the sires vs. rank of estimated breeding value for the sires, with a fixed effect for validation set, using the ASREML software (Gilmour *et al.*, 2002)**.**

Finally, accuracy of estimated breeding value ($\rho_{A_s,\hat{A}_s}$) is calculated by taking the ratio of correlation calculated from cross validation ($\rho_{\hat{A}_s,\bar{P}_{off}}$) with the accuracy of progeny testing ($\rho_{A_s,\bar{P}_{off}}$). Thus, $\rho_{A_s,\hat{A}_s} = \frac{\rho_{\hat{A}_s,\bar{P}_{off}}}{\rho_{A_s,\bar{P}_{off}}}$.

The accuracy of progeny testing was calculated using the following equation

$$\rho_{A_s,\bar{P}_{off}} = \sqrt{\frac{0.25\sigma_A^2}{0.25\sigma_A^2 + \frac{0.25\sigma_A^2}{n_d} + \frac{0.5\sigma_A^2 + \sigma_E^2}{n_o n_d}}} \quad (1)$$

Where n_d is number of dams mated to a sire and n_o the number of offspring per dam. On average a sire mated 8 dams and each dam gave five offspring in both lines.

5.2.4 Response to selection

Breeding programmes in laying hens are shifting from traditional progeny testing to genomic selection. The reason is it is expected and found that the egg number is improved using genomic information. Here, we want to investigate the potential of genomic selection for survival time in brown layer compared with a traditional breeding scheme.

To investigate the benefit of genomic selection with the current reference population, response to selection and rate of inbreeding were compared between a traditional breeding scheme and a genomic selection scheme. Response to selection was predicted using deterministic simulation based on selection index theory, using the SelAction software (RUTTEN et al. 2002). Response to genomic selection can be predicted by treating the genomic EBV as a correlated trait with full heritability (Rutten *et al.*, 2002; Schrooten *et al.*, 2005; Dekkers, 2007). SelAction predicts the response to selection and accuracy of selection for breeding programmes by accounting for the reduction in variance due to selection, known as the "Bulmer effect". This is essential when comparing genomic selection and traditional breeding programs, particularly when accuracies differ substantially between the sexes (Dekkers, 1992; Bijma, 2012).

The following inputs were used in the SelAction programme to predict response to selection and rate of inbreeding and these inputs were provided by Hendrix genetics (Table 3). We used 8% of male and female for breeding for the traditional breeding scheme and 2% of males and 8% of females for genomic selection. The generation interval for traditional breeding scheme for males is 1.9 and for females is 1.06 years. The generation interval for genomic selection for males is 0.63 and for females is 1.06 years. We used 20 breeding males and 400 breeding females for both genomic selection and traditional breeding scheme.

Table 3: Inputs used to estimate response to selection and rate of inbreeding using SelAction

input	Progeny testing	Genomic selection
Selection intensity for males	8 %	2 %
Selection intensity for females	8 %	8%
Generation interval for males	99 weeks	33 weeks
Generation interval for females	55 weeks	55 weeks
Information used for males	Pedigree, progeny (40)	own
Information used for females	Pedigree	own
Number of sire (dam)	20 (400)	20 (400)

5.3 Results

A significant effect on survival time was found for cross, batch, and laying house*row*level for both lines. The average survival time for line B1was higher than for line BD (see Figure1).

Figure 1 : Proportion of survival individuals

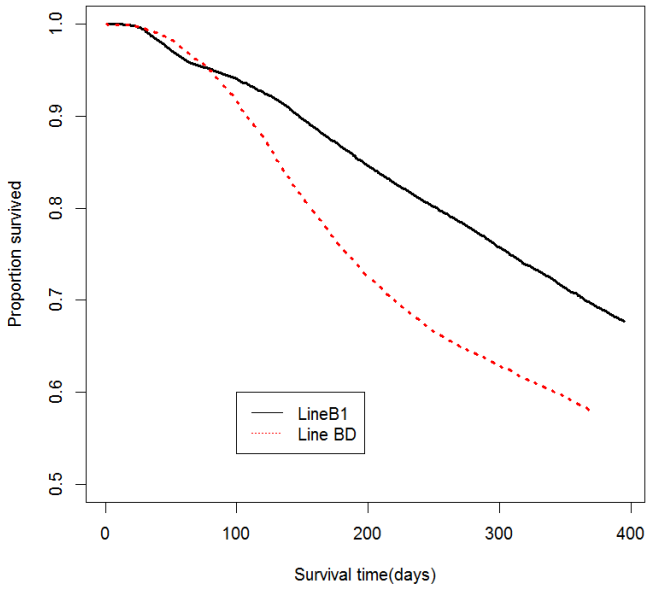

Table 4 shows estimated genetic and non-genetic parameters. The proportion of total heritable variation relative to phenotypic variance, T^2, was 0.18 for line B1 and 0.22 for line BD (Table 3). The estimated genetic variance was very significantly different from zero ($P < 0.001$). Estimated total genetic standard deviations were ~44 days for line B1 and ~ 53 days for line BD, indicating good prospects for genetic improvement. Note that this estimate includes both the contribution of the direct and of the indirect genetic effects to total genetic variance.

Table 4: Estimated parameters for survival time for lines B1 and BD

Variance components	B1 line	BD line
$\hat{\sigma}_e^2$	8,885.00±99.00	10,270.00±154.66
$\hat{\sigma}_g^2$	1,084.00 ±72.00	1,431.00 ±116.50
$\hat{\sigma}_u^2$	1,912.00± 244.19	2,739.00± 421.25
$\hat{\sigma}_P^2$	10,446.00±115.61	12,386.00±187.35
$\hat{T}^2$	0.18±0.02	0.22±0.03

$\hat{\sigma}_u^2 = \sigma_{A_D}^2 + 2(n-1)\sigma_{A_{DS}} + (n-1)^2\sigma_{A_S}^2$. Thus, $\hat{\sigma}_u^2$ is an estimate of total heritable variation

Table 5 a &b shows the correlation between the estimated breeding values of sires and the mean corrected rank of the offspring of the sires, calculated from the cross validation. The correlation between estimated breeding values of sire vs. mean phenotype of the offspring was higher for ssGBLUP than for parental average, for both line B1 and BD and for both scenarios: for both genotyped and non-genotyped sires as reference population and for genotyped sires only as reference population. Correlations were only slightly higher when including non-genotyped sires in the reference population (0.37 vs 0.35 for line B1, and 0.28 vs 0.27 for line BD).

Table 5a: Correlation of estimated breeding value of sire with average phenotype of the offspring sire ($\rho_{\hat{A},\bar{P}_{off}}$) for ssGBLUP and pedigree-BLUP for line B1 line for two scenarios: genotyped + non genotyped sires and genotyped sires only as reference population($\rho_{\hat{A},\bar{P}_{off}}$)

	Genotyped+ sire	non-genotyped	Genotyped sires	
method	ssGBLUP	pedigree-BLUP	ssGBLUP	pedigree-BLUP
1st 20%	0.29±0.13	0.15±0.15	0.16±0.15	0.08±0.15
2nd 20%	0.30±0.13	0.08±0.14	0.38±0.12	0.17±0.14
3rd 20%	0.37±0.15	0.26±0.16	0.30±0.15	0.21±0.16
4th 20%	0.48±0.12	0.42±0.13	0.44±0.12	0.43±0.13
5th20%	0.42±0.13	0.43±0.13	0.43±0.13	0.42±013
average	0.37±0.06	0.27±0.07	0.35±0.06	0.26±0.07
Accuracy[1]	0.48±0.08	0.35±0.09	0.46±0.08	0.34±0.09

[1]The accuracy of estimating the true breeding value

Table 5b: Correlation of estimated breeding value of sire with average phenotype of the offspring sire ssGBLUP vs BLUP for BD line for two scenarios for BD line: genotyped +non genotyped and genotyped only as reference population($\rho_{\hat{A},\bar{P}_{off}}$)

	Genotyped+ sire	non-genotyped	Genotyped sires	
method	ssGBLUP	p-BLUP	ssGBLUP	p-BLUP
1st 20%	0.36±0.12	0.24±0.14	0.36±012	0.23±0.14
2nd 20%	0.33±0.12	0.19±0.14	0.30±0.13	0.19±0.14
3rd 20%	0.35±0.12	0.18±0.14	0.29±0.13	0.11±0.14
4th 20%	0.14±0.12	0.18±0.14	0.16±0.14	0.18±0.14
5th20%	0.19±0.14	0.19±0.14	0.25±0.14	0.27±0.13
average	0.28±0.06	0.20±0.060	0.27±0.06	0.20±0.07
Accuracy[1]	0.35±0.06	0.25±0.08	0.34±0.08	0.25±0.09

[1]The accuracy of estimating the true breeding value

Table 6 shows the predicted response to selection per year and rate of inbreeding per generation. The predicted response using genomic selection is significantly higher than the traditional breeding scheme response to selection for line B1, and slightly higher than the traditional breeding scheme for line BD. This table shows that the rate of inbreeding for genomic information is lower than for the traditional breeding scheme for lines B1 and BD.

Table 6: Predicted response to selection(R) and rate of inbreeding for both line B1 and BD .

parameter	B1	BD
[1] $r_{\hat{A}A}$	0.48	0.35
[2] $r_{\hat{A}A}$	0.35	0.25
$\hat{T}^2$	0.18	0.22
r_p	0.21	0.17
R traditional per year	24.6	30.5
R genomic per year	39.4	34.6
Rate of inbreeding per generation traditional	2.76	2.74
Rate of inbreeding per generation genomic selection	0.76	0.76

[1]refers genomic accuracy ,[2] refers parental average accuracy and r_p is the phenotypic correlation and is calculated as $hr_{\hat{A}A}$

5.4 Discussion

We provide evidence that genomic selection increases the accuracy of estimating the breeding value for survival time in brown layers compared with the parental average (Table 5). More importantly, despite the currently small reference population, genomic

selection results in substantially higher response to selection per year for survival time in line B1 compared with the traditional breeding scheme. Thus, genetic selection with IGE coupled with marker information increases survival time for line B1, mainly by reducing the generation interval in males and by improving the accuracy of predicting the breeding values for females. The standard deviations in genomic EBVs are around 20 days for both lines, indicating good prospects for selection against mortality due to cannibalism in these brown layer lines.

Genetic parameters: The data structure with paternal half-sibs in a cage allows us to estimate the linear combination of the direct genetic effect and the indirect genetic effect, which is the total breeding value (Peeters *et al.*, 2013). The proportion of the total heritable variation to phenotypic variance for survival time was 0.18 for B1 line and 0.22 for BD line. Similar results for survival time were found in white layer by Ellen *et al.* (2008) and by Peeters et al. (2013) in crossbred white laying hens, who found values around 0.1 to 0.2. With cages composed of sibs, the contribution of direct and indirect genetic effects to the total genetic variance cannot be estimated (Peeters et al. 2013). Hence, these contributions are unknown for brown layers. In white layer lines, the indirect genetic variance contributed the majority of the total genetic variance (Ellen et al. 2008; Peeters et al. 2013).

Accuracy using cross validation: We found that the accuracy of predicting the breeding value is higher when we include the marker information, when compared to the parent average EBV. We compared two scenarios, scenario one with genotyped and non-genotyped sires as the reference population, and scenario two with genotyped sires only as reference population. In both scenarios the accuracy of predicting breeding values increased by 35% by including the genomic information for both lines.

Response to selection: the response to selection per year for traditional breeding scheme was about 24.6 days for B1 and 30.5 days for BD increase in survival time . For genomic selection it was about 39.4 days for B1 and 34.6 days for BD increase in survival time. Thus, using genomic selection response to selection per year increased by 60 % for B1 and by 13 % for BD compared with traditional breeding scheme.

The large increase in response to selection for genomic selection compared with traditional breeding scheme for B1 line is due to the following reason. The correlation between the estimated breeding value and true breeding value is higher for B1 than for BD (0.48 vs. 0.36). The reason is that B1 shows higher linkage disequilibrium than BD.

We estimated the genome-wide linkage disequilibrium using the following relationship, $M_e = \frac{1}{\bar{r}_{LD}^2} = \frac{1}{\text{var}(\mathbf{G})}$ where M_e is the effective number of segments in the genome, which is an indirect measure of linkage disequilibrium between each SNP across the genome (Daetwyler *et al.*, 2008; Goddard, 2009; Daetwyler *et al.*, 2010b). $\bar{r}_{LD}^2$ refers the genome-wide average linkage disequilibrium. Instead of using ***G***, we used $\mathbf{D} = \mathbf{G} - \mathbf{A}$, where the expected value for all elements of $\mathbf{D}$ is zero since ***G*** and ***A*** were constructed so that they had the same base. Thus, we used $M_e = \frac{1}{\bar{r}_{LD}^2} = \frac{1}{\text{var}(\mathbf{D})}$ because our populations consisted of related individuals (Wientjes *et al.*, 2013). We found M_e = 799 for B1 and M_e = 1020 for BD. Thus, the variation in family relationship for B1 line is higher than BD and consequently the linkage disequilibrium in B1 is higher than BD that causes to have higher accuracy of predicting the breeding value even if the reference population for B1 is slightly lower than for BD.

Finally, we compared the accuracy of estimating the true breeding value from empirical estimates from cross validation with the theoretical value based on the formula of Daetwyler *et al.* (2008):

$$r^2_{\hat{A}A} = \frac{N_p h^2}{N_p h^2 + N_g}$$

h^2 refers the reliability of the trait and it equals the heritability if own information is used. Our data is not own performance rather it is progeny information. Thus, we calculated the reliability which equals the square of the accuracy of progeny testing which is given in equation 1. N_P refers to the number of individual in the reference population and N_g refers the the number of loci underling the trait. For our case we used M_e as measure of N_g.

We found out for B1 line $r_{\hat{A}A}$ 0.46 and for BD line $r_{\hat{A}A}$ was 0.38. We previously showed from cross validation $r_{\hat{A}A}$ for B1 line was 0.48 and for BD line was 0.36. Thus, there is remarkable agreement between the empirical cross validation and theoretical expected value based on Daetwyler *et al.* (2008) both for B1 and BD.

Rate of inbreeding: We also compared the predicted rates of inbreeding for the genomic selection vs. traditional breeding scheme. We found that the rate of inbreeding using the traditional breeding scheme is 2.75 % increase per generation for both lines, but using genomic selection it is 0.77% per generation for both lines. Thus, genomic selection reduces the rate of inbreeding compared with the traditional breeding scheme due to the following reason. There is a strong Bulmer effect for genomic selection since the genomic estimated breeding value has a heritability of one. This reduces the correlation between EBV of sires and more distance relatives. Thus, the strong Bulmer effect reduces the rate of inbreeding for genomic selection (Bijma *et al.*, 2000). The rate of inbreeding increases for

traditional breeding scheme because in traditional breeding scheme we used parental average estimated breeding values for females. Thus, females which originate from the same family have the same EBV, which leads to family selection. This family selection is the main reason to have high rate of inbreeding for the traditional breeding scheme.

Overall, despite the small reference population (207 genotyped sires for B1 and 242 for BD) genomic selection gives a reasonable good accuracy of predicting breeding true breeding values compared with pedigree-BLUP. More importantly, it gives a substantially higher response to selection and lower rate of inbreeding compared with the traditional breeding scheme for line B1 and slightly higher response to selection and substantial lower rate of inbreeding for line BD. Thus, if the reference population increases in the future, the accuracy of estimating the breeding value and the response to selection will increase further (Hayes *et al.*, 2009a; VanRaden *et al.*, 2009).

References

Aguilar I, Misztal I, Johnson DL, Legarra A, Tsuruta S, Lawlor TJ (2010). Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *Journal of Dairy Science* **93:** 743-752.

Aguilar I, Misztal I, Legarra A, Tsuruta S (2011). Efficient computation of the genomic relationship matrix and other matrices used in single-step evaluation. *J Anim Breed Genet* **128:** 422-428.

Arango J, Misztal I, Tsuruta S, Culbertson M, Herring W (2005). Estimation of variance components including competitive effects of Large White growing gilts. *J Anim Sci* **83:** 1241-1246.

Bijma P (2012). Accuracies of estimated breeding values from ordinary genetic evaluations do not reflect the correlation between true and estimated breeding values in selected populations. *J Anim Breed Genet* **129:** 345-358.

Bijma P, Muir WM, Ellen ED, Wolf JB, Van Arendonk JAM (2007a). Multilevel selection 2: Estimating the genetic parameters determining inheritance and response to selection. *Genetics* **175:** 289-299.

Bijma P, Van Arendonk JAM, Woolliams JA (2000). A general procedure for predicting rates of inbreeding in populations undergoing mass selection. *Genetics* **154:** 1865-1877.

Blokhuis HJ, Wiepkema PR (1998). Studies of feather pecking in poultry. *Vet Quart* **20:** 6-9.

Christensen OF, Lund MS (2010). Genomic prediction when some animals are not genotyped. *Genet Sel Evol* **42:2**.

Christensen OF, Madsen P, Nielsen B, Ostersen T, Su G (2012). Single-step methods for genomic evaluation in pigs. *Animal* **6:** 1565-1571.

Craig JV, Muir WM (1996). Group selection for adaptation to multiple-hen cages: Beak-related mortality, feathering, and body weight responses. *Poultry Science* **75:** 294-302.

Daetwyler HD, Hickey JM, Henshall JM, Dominik S, Gredler B, van der Werf JHJ *et al.* (2010a). Accuracy of estimated genomic breeding values for wool and meat traits in a multi-breed sheep population. *Anim Prod Sci* **50:** 1004-1010.

Daetwyler HD, Pong-Wong R, Villanueva B, Woolliams JA (2010b). The Impact of Genetic Architecture on Genome-Wide Evaluation Methods. *Genetics* **185:** 1021-1031.

Daetwyler HD, Villanueva B, Woolliams JA (2008). Accuracy of Predicting the Genetic Risk of Disease Using a Genome-Wide Approach. *Plos One* **3(10)**.

Dekkers JCM (1992). Asymptotic Response to Selection on Best Linear Unbiased Predictors of Breeding Values. *Anim Prod* **54:** 351-360.

Dekkers JCM (2007). Prediction of response to marker-assisted and genomic selection using selection index theory. *J Anim Breed Genet* **124:** 331-341.

Ellen ED, Ducrocq V, Ducro BJ, Veerkamp RF, Bijma P (2010). Genetic parameters for social effects on survival in cannibalistic layers: Combining survival analysis and a linear animal model. *Genet Sel Evol* **42:27**.

Ellen ED, Visscher J, van Arendonk JAM, Bijma P (2008). Survival of laying hens: Genetic parameters for direct and associative effects in three purebred layer lines. *Poultry Science* **87:** 233-239.

Gilmour, A., B. Gogel, B. Cullis, S. Welham and R. Thompson, 2002 *ASReml User Guide Release 2.0*, pp. Hemel Hempstead: VSN International; 2002. http://www.vsn-intl.com/

Goddard M (2009). Genomic selection: prediction of accuracy and maximisation of long term response. *Genetica* **136:** 245-257.

Gonzalez-Recio O, Gianola D, Rosa GJM, Weigel KA, Kranis A (2009). Genome-assisted prediction of a quantitative trait measured in parents and progeny: application to food conversion rate in chickens. *Genet Sel Evol* **41:3**.

Griffing B (1967). Selection in reference to biological groups .I. Individual and group selection applied to populations of unordered groups. *Australian Journal of Biological Sciences* **20:** 127-139.

Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME (2009a). Invited review: Genomic selection in dairy cattle: Progress and challenges. *J Dairy Sci* **92:** 433-443.

Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME (2009b). Invited review: Genomic selection in dairy cattle: progress and challenges (vol 92, pg 433, 2009). *J Dairy Sci* **92:** 1313-1313.

Legarra A, Aguilar I, Misztal I (2009). A relationship matrix including full pedigree and genomic information. *J Dairy Sci* **92:** 4656-4663.

Meuwissen THE, Hayes BJ, Goddard ME (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157:** 1819-1829.

Misztal, I., 1997 BLUPF90 - a flexible mixed model program in Fortran 90. University of Georgia, pp. 1-24. http://nce.ads.uga.edu/html/projects/blupf90.pdf,.

Misztal, I., 2013 BLUPF90 family of programs.http://nce.ads.uga.edu/wiki/doku.php.

Muir WM (1996). Group selection for adaptation to multiple-hen cages: selection program and direct responses. *Poultry Science* **75:** 447-458.

Muir WM (2005). Incorporation of competitive effects in forest tree or animal breeding programs. *Genetics* **170:** 1247-1259.

Peeters K, Ellen ED, Bijma P (2013). Using pooled data to estimate variance components and breeding values for traits affected by social interactions. *Genet Sel Evol* **45:27**.

Peeters K, Eppink TT, Ellen ED, Visscher J, Bijma P (2012). Indirect Genetic Effects for Survival in Domestic Chickens (Gallus gallus) Are Magnified in Crossbred Genotypes and Show a Parent-of-Origin Effect (vol 192, pg 705, 2012). *Genetics* **192:** 1559-1559.

Rutten MJM, Bijma P, Woolliams JA, van Arendonk JAM (2002). SelAction: Software to predict selection response and rate of inbreeding in livestock breeding programs. *J Hered* **93:** 456-458.

Schaeffer LR (2006). Strategy for applying genome-wide selection in dairy cattle. *J Anim Breed Genet* **123:** 218-223.

Schrooten C, Bovenhuis H, van Arendonk JAM, Bijma P (2005). Genetic progress in multistage dairy cattle breeding schemes using genetic markers. *J Dairy Sci* **88:** 1569-1581.

Stuart A, Ord J (1994). *Kendall's Advanced Theory of Statistics, Distribution theory Vol. 1*: London: Hodder Education.

VanRaden PM (2008). Efficient Methods to Compute Genomic Predictions. *J Dairy Sci* **91:** 4414-4423.

VanRaden PM, Van Tassell CP, Wiggans GR, Sonstegard TS, Schnabel RD, Taylor JF *et al.* (2009). Invited review: Reliability of genomic predictions for North American Holstein bulls. *J Dairy Sci* **92:** 16-24.

Wientjes YCJ, Veerkamp RF, Calus MPL (2013). The Effect of Linkage Disequilibrium and Family Relationships on the Reliability of Genomic Prediction. *Genetics* **193:** 621-31.

Zhang Z, Todhunter RJ, Buckler ES, Van Vleck LD (2007). Technical note: Use of marker-based relationships with multiple-trait derivative-free restricted maximal likelihood. *J Anim Sci* **85:** 881-885.

6

General discussion

6.1 Introduction

Social interactions such as competition or cooperation are common both in plants and animals (Hamilton, 1964; Moore *et al.*, 1997; Muir, 2005; Frank, 2007). When there are social interactions, the trait value of an individual may be influenced by genes of its interacting partners, a phenomenon known as Indirect Genetic Effects (IGE; (Griffing, 1967; Muir, 2005). An IGE is a heritable effect of an individual on the trait value of its interacting partner (Griffing, 1967; Muir, 2005). A large body of literature confirmed that social interactions can create addition heritable variation in both plants and animals, for both behavioural and production traits (Moore *et al.*, 1997; Muir, 2005; Bijma *et al.*, 2007; Wilson *et al.*, 2009; Wilson *et al.*, 2011; Alemu *et al.*, 2014b).

In this thesis, we developed statistical methods to estimate IGE when interactions differ between kin and non-kin (Chapter two). We also investigated whether social interactions create additional heritable variation for bite mark traits in group-housed mink. More than 85% of the heritable variation for total bite mark score originated from social interactions (Chapter three). Furthermore, we showed that IGE estimation should take into account systematic interactions between individuals due to their sex or kin for bite mark traits in group-housed mink (Chapter four). Finally, we studied the use of genomic information to increase survival time in poultry showing cannibalistic social interactions. We found that the predicted response to selection for line B1 is higher using genomic information than with progeny testing, and the accuracy of estimating the breeding value for lines B1 and BD is higher for genomic selection compared with the parent-average EBV.

In this general discussion, I want to discuss the following five topics:

1. Kin recognition mechanisms and kin recognition consequences for animal breeding

2. The use of genome wide markers to estimate IGEs when IGE differ between kin and non-kin
3. IGE estimation using a trait based model when IGE differ between kin and non-kin
4. Accuracy of estimating the total breeding value when IGE differ between kin and non-kin
5. Prospects of genetic selection for reducing bite mark in group-housed mink

6.2 Kin recognition mechanisms

Kin recognition is the ability of an individual to distinguish kin from non-kin and interact differently with kin vs. with non-kin (Holmes and Sherman, 1982; Holmes and Sherman, 1983; Hepper, 1986; Gamboa *et al.*, 1991). It is a widespread phenomenon in various species of animals. For instance, there is kin recognition in bees (Greenberg, 1979), frogs (BLAUSTEIN AND OHARA, 1981), spiny mice (Porter *et al.*, 1981), pigs (Horrell and Hodgson, 1992b; Horrell and Hodgson, 1992a), lambs (Porter *et al.*, 2001; Ligout *et al.*, 2002; Ligout and Porter, 2003), and cloned heifers in cattle (Coulon *et al.*, 2010). There are four possible kin recognition mechanisms mentioned in the scientific literature. These are recognition based on spatial distribution, prior association, phenotypic matching and recognition alleles.

6.2.1 Recognition based on spatial distribution: With this mechanism, an individual considers as kin any other individuals encountered within a given distance (Holmes and Sherman, 1983; Hepper, 1986). Recognition based on spatial distribution may evolve when individuals found in specific locations are genetically related to one another. For instance, individuals located in a nest may be treated as kin (Waldman *et al.*, 1988). I agree with Tang-Martinez (2001) that kin recognition based on spatial distribution should not be considered as a kin recognition mechanism, because the individual treats all individuals encountered in a particular area equally irrespective of their true genetic relatedness (Tang-Martinez (2001).

6.2.2 Recognition based on prior association: This is an important kin recognition mechanism that enables an individual to recognize familiar kin (Holmes and Sherman, 1983; Tang-Martinez, 2001). The kin recognition process is as follow: during the rearing period an individual learns the phenotype of its relatives (sibs, parents). Later they use this learned template to distinguish these familiar relatives from newly encountered individuals (Holmes and Sherman, 1983; Tang-Martinez, 2001). Sibling recognition in the prairie vole, (*Microtus ochrogaster) (Gavish et al., 1984),* in the white-footed mouse (*Peromyscus leucopus) (Halpin and Hoffman, 1987)* and in the spiny mouse (*Acomys cilicicus)* (Porter *et al.*, 1981) depends on association prior to weaning.

Prior association may be the most important kin recognition mechanism for livestock and fish. In pigs, for example, sibs weaned together from birth to weaning learn the phenotype of their relatives (in this case sibs). Later when families are mixed at the start of the fattening phase, individuals fight with members of other families but not with their family members (Erhard and Mendl, 1997; Giersing and Andersson, 1998; Stookey and Gonyou, 1998; D'Eath, 2004). Lambs interact differently with familiar individuals (kin) than with non-familiar individuals, and mainly they use prior association to recognize familiar kin (Porter *et al.*, 2001; Ligout *et al.*, 2002; Ligout and Porter, 2003). Cattle can discriminate the familiar herd members in learning experiments (Hagen and Broom, 2003). Laying hens prefer to associate with the familiar group rather than with a group of strangers (D'Eath and Keeling, 2003). Both cattle and laying hens most probably use prior association to recognize familiar individuals (D'Eath and Keeling, 2003; Hagen and Broom, 2003).

A series of studies has shown kin recognition using prior association in different species of fish, such as sticklebacks (*Gasterosteus aculeatus*) (Frommen *et al.*, 2007) and bluegill

sunfish (*Lepomis macrochirus)* (Hain and Neff, 2006). Thus, kin recognition using prior association is a common phenomenon in livestock and fish.

6.2.3 Recognition based on phenotypic matching: This is a useful kin recognition mechanism that enables an individual to recognize unfamiliar kin. The kin recognition process is as follow: animals learn the template of their own phenotype (Mateo and Johnston, 2000) and/or those of familiar kin (Sherman *et al.*, 1997) and subsequently match the phenotypes of newly encountered individuals to this template (Hepper, 1986). The template may include visual (Cooke *et al.*, 1972), chemical (Hepper, 1986), and auditory cues (Beecher, 1982). Using phenotype matching, an individual can recognise kin without previous experience. It is observed in Belding's ground squirrels (HOLMES, 1986), frogs (BLAUSTEIN AND OHARA, 1981), fish (Olsen, 1989; Olsen *et al.*, 1998; Hesse *et al.*, 2012), western bluebirds (Akcay *et al.*, 2013) and monkeys (Wu *et al.*, 1980). Phenotype matching relies on correlation between phenotypic and genotypic similarity, so that recognizable traits are more similar among relatives than non-relatives (Holmes and Sherman, 1983; Gerlach *et al.*, 2008).

There is almost no literature on kin recognition due to phenotypic matching in livestock. I found one study that stated that cloned heifers might use phenotypic matching to recognize unfamiliar kin (Coulon *et al.*, 2010). Cloned heifers from a specific genotype appeared to be more associated and interact with each other than with others (Coulon *et al.*, 2010). This association could also be based on morphological and behavioural affinity (Coulon *et al.*, 2010). Thus, there should be further investigation on the importance of kin recognition due to phenotypic matching in livestock species.

With respect to fish, there is some literature that shows that fish species such as Arctic charr (*Salvelinus alpinus*; (Olsen, 1989) and coho salmon (*Oncorhynchuskisutch*; (Quinn and Busack, 1985) use phenotypic matching to recognize unfamiliar kin.

6.2.4 Kin recognition based on recognizing alleles: With this mechanism, a single gene or a group of genes produces a phenotypic cue such as an odor. Using these cues the individual will recognize similar cues in other individuals. The individual will show preferential treatment towards individuals that carry similar cues (Hamilton, 1964). Phenotypic matching and recognizing allele are more or less similar because recognition using recognizing allele is expressed in form of phenotype matching (Blaustein, 1983; Waldman, 1987).

6.2.5 Consequence of kin recognition for animal breeding

A genetic model that estimates indirect genetic effect was developed by (Griffing, 1967; Muir, 2005; Bijma *et al.*, 2007a). The model assumes that an individual interacts equally with all group mates. This assumption may not be true in a mixed group of kin and non-kin. In a mixed group of kin and non-kin, an individual may interact differently with kin vs. non-kin. In this thesis (Chapter two), we developed a statistical method that estimates IGEs when interactions differ between kin and non-kin. The method takes into account both heritable and non-heritable systematic interactions in a mixed group of kin and non-kin. For example, the heritable systematic interactions are accounted for by modelling two types of IGE: IGE to kin and IGE to non-kin. The non-heritable systematic interactions are accounted for by modelling two types of indirect environmental effects: indirect environment to kin and indirect environment to non-kin. One important question I want to address here is the consequences of ignoring systematic interaction due to kin recognition or other systematic interactions factors such as sex when estimating IGEs.

In animal breeding, the aim is mainly improving the response to selection. The response to selection for socially affected trait depends on total heritable variation (Bijma, 2011). In chapter two, we showed that the total heritable variation can be estimated either using the reduced model or the usual indirect genetic effects model, regardless of whether IGE to kin vs. IGE to non-kin are the same or not. Thus, if an individual shows different IGE to kin and to non-kin due to kin recognition or due to sex, the total heritable variation can still be estimated. However, if an individual shows different non-genetic indirect effects to kin and to non-kin due to kin recognition or sex, it has a consequence for the non-genetic parameters of the model. Thus, ignoring non-heritable kin recognitions has an impact on the estimation of genetic parameters.

In chapter two, we showed that the model accounts for non-genetic systematic interaction by fitting residual correlations for family members (kin) in a group in addition to random group effects. If this correlation is positive, it can be fitted easily by adding an additional non-genetic random effect (cage*kin) in addition to the cage effect and the residual (Bijma *et al.*, 2007a). If this correlation is negative, it can be fitted by fitting residual correlations for family members (kin) in a group (BIJMA *et al.* 2007a, Chapter two). The cage*kin correlation represents non-heritable systematic interaction. Ignoring this parameter will result in biased estimates of the genetic parameters. For example, ignoring the non-genetic systematic interaction resulted in biased estimates of the genetic parameters for average daily gain in pigs (DUIJVESTEIJN 2014), for harvest weight in the gift strain of Nile tilapia (*Oreochromis niloticus*) (Khaw *et al.*, submitted), and for bite mark traits in group-housed mink (Chapter three).

One last point I want to add is that the IGE model takes in to account systematic interactions due to kin recognition irrespective of the mechanism of the kin recognition.

Thus, from animal breeding point of view it is not needed to know the mechanism of kin recognition. It is sufficient to know whether there is kin recognition or not.

In summary, kin recognition is a widespread phenomenon in various species of animals. Non-heritable kin recognition can have a substantial effect on estimates of genetic parameters. Thus, genetic parameter estimation for socially affected traits should take into account non-heritable systematic interaction either due to kin recognition or other non-genetic systematic interaction, for example, systematic interaction due to sex. Ignoring the non-heritable systematic interactions may result in biased estimates of the genetic parameters.

6.3 The use of genome wide markers to estimate IGE when interaction differ kin vs. non-kin

In chapter two, we showed that not all genetic parameters are statistically identifiable when interactions differ between kin and non-kin. We showed that a meaningful linear combination of genetic parameter is estimable and suggested a possible method to estimate all six genetic parameters. We suggested that all six genetic parameters are statistically identifiable in this situation when groups consist of a mix of full-sibs, half-sibs, and unrelated individuals. A statistically more powerful approach may come from cross-fostering designs, where full siblings that grow up in different litters may interact as if they were unrelated. When cross-fostering is impossible and a mix of full and half siblings is unavailable, variation in relatedness among pairs of full siblings, estimated using genome-wide genetic markers may provide a solution. Here, I discuss the estimation of all genetic parameters using the variation in relatedness among pair of full siblings estimated from genetic markers.

6.3.2 Estimation of IGE when interactions differ between kin vs. strangers using genome wide markers

Estimation of genetic parameters for direct and indirect genetic effects depends on the covariance between phenotypes of relatives and their social partners (Lynch and Walsh, 1998). The degree of additive genetic covariance between relatives is expressed by the additive genetic relationship and is estimated from pedigree data (Odegard and Meuwissen, 2012). The pedigree additive genetic relationship has a limitation, as it is the expected relationships rather than the actual relationship (Hill, 1993; Odegard and Meuwissen, 2012). Thus, the same types of relatives, for example, two pairs of full-sibs, have the same additive genetic relationship which makes the pedigree based analysis based

on between family covariance. Thus, the residuals and the Mendelian sampling deviations of non-parents are fully confounded on the same individuals (Odegard *et al.*, 2010).

Recently, Visscher *et al.* (2006) have estimated actual relationships among human-full sibs, which range from 0.37 to 0.62. Using these actual relationships they quantified the additive genetic variation in human height using the within-family segregation only. In this way, they use the Mendelian sampling deviation. The reason behind this is the actual relationship from genetic marker utilises the variation in relatedness among pairs of twins.

As stated in chapter two, not all the genetic parameters are estimable when interactions differ kin vs. non-kin. Direct genetic effects and IGE to kin are fully confounded and their linear combinations known as family breeding value is estimable in the reduced model. Thus, the main reason for not estimating all the genetic parameters is that there are not enough informative covariances, i.e., six genetic parameters need to be estimated using five informative covariances. The expected covariance between the phenotypes of group members of the same family members (full-sib) within a group is the same (Chapter 2 Appendix A1). Thus, the actual relationship derived from marker enables us to use the within full-sib variation to estimate the components of the family breeding values, because the actual relationship for full-sibs in each group is different. In statistical term, we have enough informative covariances to estimate the components of the family breeding value variance. The covariance between each pair of individuals is different, though they may vary little. Thus, we may need large sample size because the power is limited. This method is more efficient for fecund species that have large family size (Odegard and Meuwissen, 2012).

Kin recognition experiments have been undertaken mainly in fecund species such as amphibians (Blaustein and Ohara, 1981; Waldman, 1981), rodents (Holmes and Sherman,

1982), and fish (Quinn and Busack, 1985; Olsen, 1989). As the family size increases, the actual variation within a family for pair of sib will be larger which makes the method more efficient. Thus, in fecund species the sample size required to test kin recognition may not be as large as in human data used by Visscher *et al.* (2006).

6.3.3 Genomic model when IGE differ between kin and strangers

In this section I discuss how the statistical model would look like. I propose a model that includes all SNP (single nucleotide polymorphisms) simultaneously (Habier *et al.*, 2007). This model can be referred to as SNPBLUP. This model is equivalent with GBLUP (Goddard, 2009). In this model it is assumed that all SNPs contribute equally to the genetic variance. In addition to SNPGBLUP, Bayesian methods may be needed to allow different contributions to genetic variance across SNPs by differential shrinkage (Meuwissen *et al.*, 2001b; Habier *et al.*, 2007; Calus *et al.*, 2008). SNPGBLUP or Bayesian methods help to estimate SNP associations with respect to direct genetic effects and indirect genetic effects to kin and indirect genetic effects to non-kin (Duijvesteijn, 2014). The proposed model is

$$\mathbf{y} = \mathbf{Xb} + \mathbf{M_D}\boldsymbol{\alpha}_\mathbf{D} + \mathbf{M_{Ik}}\boldsymbol{\alpha}_\mathbf{Ik} + \mathbf{M_{Ink}}\boldsymbol{\alpha}_\mathbf{Ink} + \mathbf{Wg} + \mathbf{e} \qquad 6.1$$

where $\boldsymbol{y}$ is a vector of phenotypic observations; $\mathbf{X}$ is the design matrix for the fixed effects, $\mathbf{b}$ is a vector of unknown fixed effects; $\boldsymbol{\alpha}_\mathbf{D}$ is a vector of direct SNP effects for each marker from the focal individual and $\mathbf{M_D}$ is a design matrix of which the entries are SNP genotypes of the focal individual coded as the count of a given allele. Vector $\boldsymbol{\alpha}_\mathbf{Ik}$ contains the indirect genetic effects to kin for each marker from the members of the same family found in the group, and $\mathbf{M_{Ik}}$ is a design matrix of which the entries are the marker genotypes of *j* family members in the cage, coded as the count of a given allele. Vector $\boldsymbol{\alpha}_\mathbf{INk}$

contains the indirect genetic effects to non-kin for each marker from the non-family members found in the same group and $\mathbf{M}_{\mathbf{Ink}}$ is a design matrix of which the elements are the marker genotypes of ***Ink*** non-family members in the cage coded as the count of a given allele. The SNP effects in $\boldsymbol{\alpha}_{\mathbf{D}}, \boldsymbol{\alpha}_{\mathbf{Ik}}$ and $\boldsymbol{\alpha}_{\mathbf{Ink}}$ can be derived from a normal distribution, e.g. in SNPBLUP, where $\alpha \sim N(0, \mathbf{I}\sigma^2_{SNP})$, where $\mathbf{I}$ is an identity matrix and σ^2_{SNP} the variance due to a single SNP (Meuwissen *et al.*, 2001a). The $\mathbf{g}$ is a vector of random group effects, with $\mathbf{g} \sim N(\mathbf{0}, \mathbf{I}_g\sigma^2_g)$ and incidence matrix $\mathbf{W}$ linking records to groups. The covariance structure for the residual term is $\text{var}(\mathbf{e}) = \mathbf{R}\sigma^2_e$, where $R_{ii} = 1$, $R_{ij} = \rho$ when i and j are group mates from the same family, and $R_{ij} = 0$ otherwise (Chapter two).

With large sample size and dense genomic information it is possible to test whether IGE differ between kin and non-kin.

6.4 Indirect genetic effect estimation using a trait based model

Traits affected by indirect genetic effect can be modelled in two ways. First using the variance component model, where the trait value of the focal individual is partitioned into a direct genetic component originating from focal individual and an indirect genetic effect originating from the interacting partners and the residuals (Griffing, 1967; Muir, 2005; Bijma *et al.*, 2007a).The variance component method divides the total phenotypic variance of the trait into the direct genetic variance which originates from an individual's own genotype and the indirect genetic variance which originates from interacting partners' genotypes and the residual variances. With this method direct and indirect genetic (co)variances are estimated without knowledge of the social traits that cause the IGEs (Willham, 1963; Griffing, 1967; Muir, 2005; Bijma *et al.*, 2007a; Bijma *et al.*, 2007b). After finding (co)variance of the genetic terms, DGEs, and IGEs are predicted as random effects, in the same way as prediction of ordinary breeding values (Henderson, 1975).

The second method uses a trait-based model. Trait-based models define IGEs as variation of focal individual traits caused by one or more heritable traits in interacting individuals (McGlothlin and Brodie, 2009). Thus, these models specify the phenotype of the focal individual as a function of the phenotypic trait values of interacting partners, and are referred to as "interacting phenotypes" (Moore *et al.*, 1997; Wolf *et al.*, 1998). The strength of IGEs in the trait-based model is determined by ψ (psi), a matrix of "regression coefficients" of trait values of the focal individual on trait values of its interacting partner(s) (Moore *et al.*, 1997). In this topic, first I will review the compatibility of the variance component model and the trait-based model when IGE is the same between kin and non-kin. Second, I will show that the variance component model and trait-based model are compatible with each other when IGE differ between kin and non-kin.

6.4.2 Trait-based model when IGE is same for kin vs. non-kin

Here, I am going to show the simplest situation, in which the trait of interest and the trait causing the IGE is the same trait, and the interaction is between two individuals. For this case, the trait based model equals (Moore *et al.*, 1997)

$$P_i = A_i + e_i + \psi P_j, \qquad 6.2$$

where *i* represents the focal individual, *j* its interacting partner, A_i the additive genetic component of P_i originating fully from the focal individual, ψ (psi), the "regression coefficient" explaining the strength and direction of the effect of the trait value of *j* on the trait value of *i*. With symmetric interaction (reciprocal interaction), the same model applies to the interacting partner, $P_j = A_j + e_j + \psi P_i$, which creates a feedback loop. We overcome this problem by substituting P_j into the model for P_i. This gives (Moore *et al.*, 1997)

$$P_i = \frac{A_i + e_i + \psi(A_j + e_j)}{1-\psi^2}. \qquad 6.3$$

Thus , the trait value is undefined for $|\psi| = 1$.

From Equation 6.3 it follows that the magnitude of DGEs and IGEs using the trait-based model is given by (McGlothlin and Brodie, 2009)

$$A_{D,i} = \frac{A_i}{1-\psi^2} \qquad 6.4a$$

$$A_{S,i} = \frac{\psi A_i}{1-\psi^2} \qquad 6.4b$$

$$corr(A_D, A_S) = \begin{cases} -1 \text{ for } \psi < 0 \\ 0 \text{ for } \psi = 0 \\ 1 \text{ for } \psi > 1 \end{cases} \quad 6.4c$$

Therefore, in the trait-based model the variance of *A* together with value of ψ determines the (co) variances of DGEs and IGEs. If the traits causing IGE are known, the trait-based model quantifies the strength of the interaction for each trait using the interaction coefficients (Bijma, 2014). Thus, when the traits causing IGE are known, trait-based model can help us to understand the biological mechanism of social interactions.

6.4.3 Trait based model when interactions differ between kin vs. non-kin

In this section, I show the trait-based model when interactions differ between kin and non-kin and clarify how the trait-based model relates to the variance component model. The trait-based model when interactions differ between kin and non-kin is

$$P_i = A_i + e_i + \psi \sum_j P_j + \gamma \sum_k P_k \quad 6.5$$

where ψ explains the strength and direction of the effect of related individuals *j* on *i* (kin effect) and γ explain the strength and direction of the effect of unrelated individuals *k* on *i* (non-kin effect). Here both ψ and γ are population parameters.

After defining the model, now we are going to show how this model is related to the variance component model. For simplicity assume we have a population divided into groups and each group consists of four individuals originating from two families. For example, we may have a group with members 1, 2, 3 and 4, where 1 and 2 are family members and 3 and 4 are members of another family. The phenotype of each individual will be:

$$P_1 = A_1 + e_1 + \psi P_2 + \gamma(P_3 + P_4)$$

$$P_2 = A_2 + e_2 + \psi P_1 + \gamma(P_3 + P_4)$$

$$P_3 = A_3 + e_3 + \psi P_4 + \gamma(P_1 + P_2)$$

$$P_4 = A_4 + e_4 + \psi P_3 + \gamma(P_1 + P_2)$$

Putting these equations in a matrix gives the following expression

$$\begin{bmatrix} 1 & -\psi & -\gamma & -\gamma \\ -\psi & 1 & -\gamma & -\gamma \\ -\gamma & -\gamma & 1 & -\psi \\ -\gamma & -\gamma & -\psi & 1 \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \end{bmatrix} = \begin{bmatrix} A_1 + e_1 \\ A_2 + e_2 \\ A_3 + e_3 \\ A_4 + e_4 \end{bmatrix}.$$

Thus, we can express the phenotypes in terms of the underlying breeding values by solving for the vector of phenotypes,

$$\begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ P_4 \end{bmatrix} = \begin{bmatrix} 1 & -\psi & -\gamma & -\gamma \\ -\psi & 1 & -\gamma & -\gamma \\ -\gamma & -\gamma & 1 & -\psi \\ -\gamma & -\gamma & -\psi & 1 \end{bmatrix}^{-1} \begin{bmatrix} A_1 + e_1 \\ A_2 + e_2 \\ A_3 + e_3 \\ A_4 + e_4 \end{bmatrix}.$$

Solving the matrix inverse using the mathematical software Mathematica (Wolfram Research, 2010) , we can write the phenotype in terms of the trait-based model. For instance, for individual one:

$$P_1 = \left(\frac{1-\psi^2-2\gamma^2-2\psi\gamma^2}{(1-\psi)^2-4(\gamma+\psi\gamma)^2} \right)(A_1 + e_1) + \left(\frac{\psi-\psi^3+2\gamma^2+2\psi\gamma^2}{(1-\psi)^2-4(\gamma+\psi\gamma)^2} \right)(A_2 + e_2) + \frac{\gamma+2\psi\gamma+\psi^2\gamma}{(1-\psi)^2-4(\gamma+\psi\gamma)^2}(A_3 + e_3) + \left(\frac{\gamma+2\psi\gamma+\psi^2\gamma}{(1-\psi)^2-4(\gamma+\psi\gamma)^2} \right)(A_4 + e_4) \quad 6.6$$

Thus, the trait value is undefined when the denominator is zero, which occurs for $\gamma = \left| \frac{1-\psi}{2(1+\psi)} \right|$.

The corresponding phenotypic value using variance component model for individual one is:

$$P_1 = A_{D_1} + e_{D_1} + A_{S_{f2}} + e_{S_{f2}} + A_{S_{nf3}} + e_{S_{nf3}} + A_{S_{nf4}} + e_{S_{nf4}} \quad 6.7$$

Thus, the relationship between the trait-based model and the variance component model is given by:

$$A_{D_i} = \left(\frac{1-\psi^2-2\gamma^2-2\psi\gamma^2}{(1-\psi)^2-4(\gamma+\psi\gamma)^2} \right) A_i \quad 6.7a$$

$$A_{S_{fi}} = \left(\frac{\psi-\psi^3+2\gamma^2+2\psi\gamma^2}{(1-\psi)^2-4(\gamma+\psi\gamma)^2} \right) A_i \quad 6.7b$$

$$A_{S_{nfi}} = \left(\frac{\gamma+2\psi\gamma+\psi^2\gamma}{(1-\psi)^2-4(\gamma+\psi\gamma)^2} \right) A_i \quad 6.7c$$

Therefore, in the trait-based model the variance of A together with value of ψ and γ determines the (co) variances of DGEs, IGEs to kin and IGE to non-kin.

One way of estimating ψ and γ may seem to be as regression coefficients, for example using a model $Y = A_i + b_k P_j + b_{nk}(P_k + P_l) + e$, where b_k and b_{nk} are expected estimates of ψ and γ. However, the relationship between regression coefficients and ψ and γ is complex, even when IGE are the same between kin and non-kin and fitting the simplest possible case in which the trait of interest and the trait causing IGE is the same and the interaction is between two individuals (Bijma, 2014). In this simple case, it might be expected that b is an estimator of ψ. However, the relationship between ψ and b is complex, b is a quadratic function of ψ (Bijma, 2014). Therefore, when IGE differ between kin and non-kin, estimating ψ and γ as functions of b_k and b_{nk} will be even more difficult. Both ψ and γ will be complex functions of b_k and b_{nk}.

The second solution is fitting the variance component model and solving ψ and γ as functions of the estimated variance components, using Equations 6.7a-6.7c. However, as can be seen from Equations 6.7a-c, the relationship is very complex, so that it is difficult to solve for ψ and γ when fitting the trait- based model, even for the simplest possible case of a single trait. Therefore, when we fit variance components we could never back solve ψ and γ.

In practical situations, estimation of ψ requires that all the traits that cause IGEs to kin should be known and recorded and estimation of γ requires that all the traits that cause IGE to non-kin should be known and recorded. However, it is difficult to know all the traits of the interacting partner that affect the focal individual.

Fitting trait-based model to estimate IGE when interactions differ between kin and non-kin is complex. The theoretical relationship between the variance component model and the trait-based model can be developed, but estimating ψ and γ is difficult. Thus, it is very difficult to use empirically the trait-based model when IGE differ between kin and non-kin.

6.5 Accuracy of estimated breeding value when IGE differ between kin vs. non-kin

Accuracy of estimating the breeding value is the correlation between true breeding value and estimated breeding value (Falconer, 1960). Accuracy of estimating the true breeding value is the most important genetic parameter (Falconer, 1960). It is directly proportional to the response to selection (Falconer, 1960). We have different types of accuracy for socially affected traits depending on the types of selection, such as mass selection, multilevel selection and group selection. Here I want to discuss the accuracy of mass selection when IGE differ between kin and non-kin.

The accuracy of mass selection when IGE is the same for kin and non-kin was given as (WADE et al. 2010)

$$\rho = \frac{r\sigma^2_{TBV} + (1-r)\left[\sigma^2_{A_D} + (n-1)\sigma_{A_{DS}}\right]}{\sigma_{TBV}\sigma_P}$$

Here, I want to derive the accuracy of mass selection when IGE differ between kin and non-kin. The data structure is the same as in Chapter two, which is two full-sib families in each group, the relationship between both families in a group is zero, and the relationship within families in a group is 0.5.

From Chapter two we have the phenotypic value and total breeding value as follow:

$$P_i = A_{D,i} + E_{D,i} + \sum_{j=1}^{\frac{n}{2}-1} A_{S_f,j} + \sum_{j=1}^{\frac{n}{2}-1} E_{S_f,j} + \sum_{k=1}^{n/2} A_{S_u,k} + \sum_{k=1}^{n/2} E_{S_u,k} \qquad 6.8a$$

$$A_{T,i} = A_{D,i} + (½n - 1)A_{S_f,i} + ½n A_{S_u,i} \qquad 6.8b$$

The accuracy of mass selection will be

$$\rho = cor(A_{T,i}, P_i)$$

$$\rho = \frac{\text{cov}(A_{T,i}, P_i)}{\sigma_p \sigma_T}$$

Using Equations 6.8a and 6.8b, the covariance between total breeding value and phenotypic value of an individual will be

$$\text{cov}(A_{T,i}, P_i) = \sigma^2_{A_D} + \left(\frac{n}{2} - 1\right)(1+r)\sigma_{A_{Dsf}} + \left(\frac{n}{2} - 1\right)r\left(\left(\frac{n}{2} - 1\right)\sigma^2_{A_{Sf}} + \frac{n\sigma_{A_{sfsu}}}{2}\right) + \frac{n\sigma_{A_{DSu}}}{2} \qquad 6.8c$$

where r is the relatedness within families within a group. The phenotypic variance will be

$$\sigma^2_P = \sigma^2_{A_D} + (n-2)r\sigma_{A_{DS\,f}} + \left(\frac{n}{2} - 1\right)\sigma^2_{A_{S_f}}\left(1 + \left(\frac{n}{2} - 2\right)r\right) + \frac{n\sigma^2_{A_{Su}}\left(1 + \left(\frac{n}{2} - 1\right)r\right)}{2} \qquad 6.8d$$

Thus, the true accuracy of mass selection when IGE differ between kin and non-kin will be:

$$\rho = \frac{\sigma^2_{A_D} + \left(\frac{n}{2}-1\right)(1+r)\sigma_{A_{Dsf}} + \left(\frac{n}{2}-1\right)r\left(\left(\frac{n}{2}-1\right)\sigma^2_{A_{Sf}} + \frac{n\sigma_{A_{SfSu}}}{2}\right) + \frac{n\sigma_{A_{DSu}}}{2}}{\sigma_{A_T}\sigma_P} \quad 6.8e$$

As can be seen from Equation 6.8e, the accuracy is directly proportional to direct genetic variance and IGE variance to kin and the three types of covariances ($\sigma_{A_{Dsf,}}\sigma_{A_{Dsu,}}\sigma_{A_{SfSu}}$). The accuracy is independent of IGE variance to non-kin. This shows there is no direct selection for IGE to non-kin. The response in IGE to non-kin, therefore, depends on $\sigma_{A_{SfSu}}$. If $\sigma_{A_{SfSu}}$ is negative, the response in IGE to non-kin will be negative. Furthermore relatedness has smaller impact to the accuracy compared with the traditional IGE model (when IGE is the same kin and non-kin) (Figure 1).

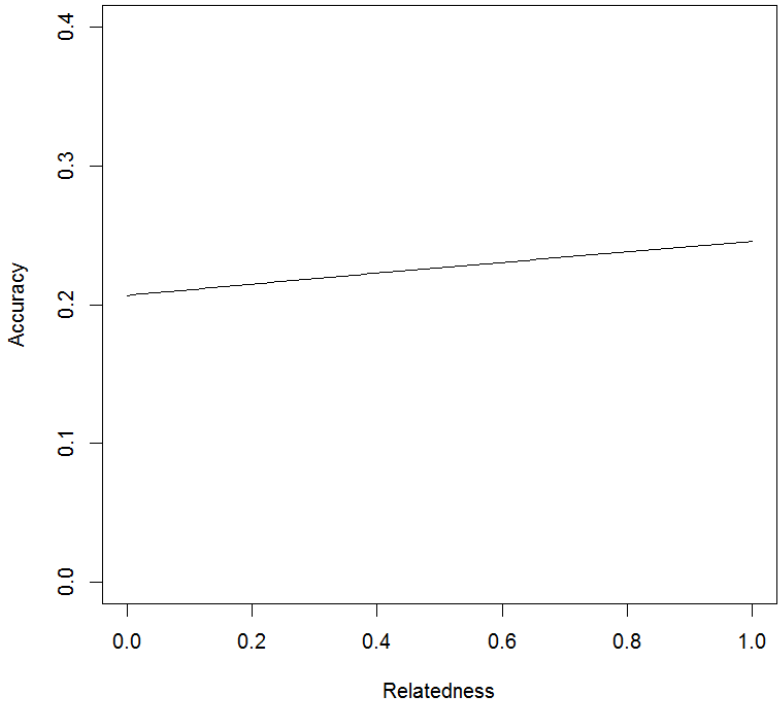

Figure1. Accuracy of mass selection when IGE differ between kin and non-kin. The following inputs were used to plot this figure $r_{A_{DS_f}} = 0.5, r_{A_{DS_u}} = r_{A_{S_f S_u}} = -0.5$ and $h_D^2 = 0.5,\ h_{S_f}^2 = h_{S_u}^2 = 0.1$

As shown in Figure 1 as relatedness between the same families within a group increases, accuracy of mass selection increases slightly.

6.5.2 Perceived accuracy of mass selection when IGE differ between kin and non-kin using traditional-IGE model

In this topic, I want to derive the perceived accuracy of mass selection when IGE differ between kin and non-kin, but when this is not taken into account. In other words, IGE may depend on kin, but breeders may not be aware of this, and thus may calculate an accuracy based on the traditional model. This will result in a perceived accuracy that differs from the true accuracy. Here, for simplicity, I assumed that the non-genetic part of the model is correct (error structure from the full model is used). The perceived trait value and total breeding value are given by

$$P_i = A_{D,i} + \sum_{j=1}^{n-1} A_{S,j} + E_{D,i} + \sum_{j=1}^{\frac{n}{2}-1} E_{S_f,j} + \sum_{k=1}^{n/2} E_{S_u,k} \qquad 6.8f$$

$$A_{T,i} = A_{D,i} + (n-1)A_{S,i} \qquad 6.8g$$

Equation 6.8f and 6.8g are the models that are used to derive the accuracy if we assumed that IGE are the same for kin and non-kin.

Using Equations 6.8f and 6.8g the perceived covariance between total breeding value and phenotypic value of an individual will be:

$$\text{cov}(A_{T,i}, P_i) = \sigma^2_{A_D} + (n-1)\sigma_{A_{DS}} + \left(\frac{n}{2}-1\right) r\sigma_{A_{DS}} + \left(\frac{n}{2}-1\right)(n-1) r\sigma^2_{A_S}$$

Thus, calculating the perceived accuracy requires values for $\sigma^2_{A_D}$, $\sigma_{A_{DS}}$ and $\sigma^2_{A_S}$. When IGE differ between kin and non-kin, the estimates of $\sigma^2_{A_D}$ and $\sigma_{A_{DS}}$ using the

traditional IGE model are biased, but the estimates of $\sigma^2_{A_S}$ is unbiased (see Chapter two). In Chapter two, we derived, using the effective record approach, the expectations of the estimates of $\sigma^2_{A_D}$, $\sigma_{A_{DS}}$ and $\sigma^2_{A_S}$ from the traditional model expressed as a function of the parameters of the full model.

The expected value for these are (Chapter 2)

$$E(\hat{\sigma}^2_{A_D}) = \sigma^2_{A_D} + (½n-1)^2\left(\sigma^2_{A_{S_f}} - 2\sigma_{A_{S_f}A_{S_u}} + \sigma^2_{A_{S_u}}\right) + (n-2)(\sigma_{A_DA_{S_f}} - \sigma_{A_DA_{S_u}})$$

$$E(\hat{\sigma}_{A_{DS}}) = \sigma_{A_DA_{S_u}} + (½n-1)(\sigma_{A_{S_f}A_{S_u}} - \sigma^2_{A_{S_u}})$$

$$E(\hat{\sigma}^2_{A_S}) = \sigma^2_{A_{S_u}}$$

Substituting $E(\hat{\sigma}^2_{A_D})$ for $\sigma^2_{A_D}$, $E(\hat{\sigma}_{A_{DS}})$ for $\sigma_{A_{DS}}$ and $E(\hat{\sigma}^2_{A_S})$ for $\sigma^2_{A_S}$ from above equation yields

$$\mathrm{cov}(A_{T,i}, P_i) = \sigma^2_{A_F} + \left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)(r-1)\sigma^2_{A_{Su}} + \left(\left(\frac{n}{2}-1\right)r+1\right)\left(\sigma_{A_{DSu}} + \left(\frac{n}{2}-1\right)\sigma_{A_{SfSu}}\right) \quad 6.8h$$

This provides the denominator of the perceived accuracy.

The next step is to derive the phenotypic variance. The phenotypic variance is

$$\sigma^2_P = \hat{\sigma}^2_{A_D} + (n-1)\hat{\sigma}^2_{A_S} + 2r\left(\frac{n}{2}-1\right)\hat{\sigma}_{A_{DS}} + \left(\frac{n^2}{4}-\frac{n}{2}\right)r\hat{\sigma}^2_{A_S} + \sigma^2_{E_D} + \left(\frac{n}{2}-1\right)\sigma^2_{E_{Sf}} + \frac{n}{2}\sigma^2_{E_{Su}}$$

Substituting $E(\hat{\sigma}^2_{A_D})$ for $\sigma^2_{A_D}$, $E(\hat{\sigma}_{A_{DS}})$ for $\sigma_{A_{DS}}$ and $E(\hat{\sigma}^2_{A_S})$ for $\sigma^2_{A_S}$ into the expression for phenotypic variance yields

$$\sigma_P^2 = \sigma_{A_F}^2 + \frac{n^2\sigma_{A_{Su}}^2}{4} + 2(r-1)\left(\frac{n}{2}-1\right)\left[2\left(\frac{n}{2}-1\right)\sigma_{A_{S_f S_u}} + \sigma_{A_{DS_u}}\right] + \sigma_{E_D}^2 + \left(\frac{n}{2}-1\right)\sigma_{E_{Sf}}^2 + \frac{n}{2}\sigma_{E_{Su}}^2 \quad 6.8i$$

Thus, the perceived accuracy of mass selection will be

$$cor(A_{T,i}, P_i) = \frac{\sigma_{A_F}^2 + \left(\frac{n}{2}-1\right)\left(\frac{n}{2}\right)(r-r)\sigma_{A_{Su}}^2 + \left(\left(\frac{n}{2}-1\right)r+1\right)\left(\sigma_{A_{DSu}} + \left(\frac{n}{2}-1\right)\sigma_{A_{SfSu}}\right)}{\sigma_P \sigma_{A_T}} \quad 6.8j$$

Hence, Equation 6.8j shows the perceived accuracy of mass selection when IGE depend on kin but this is ignored in the calculation of the accuracy. In the following, I investigate whether ignoring the dependency of IGE on kin leads to over- or under estimation of the accuracy.

6.5.3 Comparing accuracy of mass selection using traditional IGE model vs. full model

In this section, I want to compare the true accuracy of mass selection when IGE differ between kin and non-kin with the perceived accuracy when this dependency is ignored. As can be seen from Equation 6.8e and 6.8j, the two correlations are different. For instance, the covariance of the first correlation is independent of the variance of IGE to non-kin, but the covariance of the second correlation is dependent on the variance of IGE to non-kin. Furthermore, the phenotypic variances for the two correlations are also different. Thus, the accuracy of mass selection when IGE differ between kin and non-kin will be biased if we obtain it from the traditional-IGE model.

In the next section, I want to compare both accuracies (full model accuracy vs traditional-IGE model accuracy) for a range of correlations between IGE to kin and

IGE to non-kin for three scenarios. In the first scenario, the variance of IGE to kin is the same as the variance of IGE to non-kin (Figure 2). In the second scenario, the variance of IGE to kin is four times the variance of IGE to non-kin (Figure 3). In the third scenario, the variance of IGE to kin is a quarter of the variance of IGE to non-kin (Figure 4).

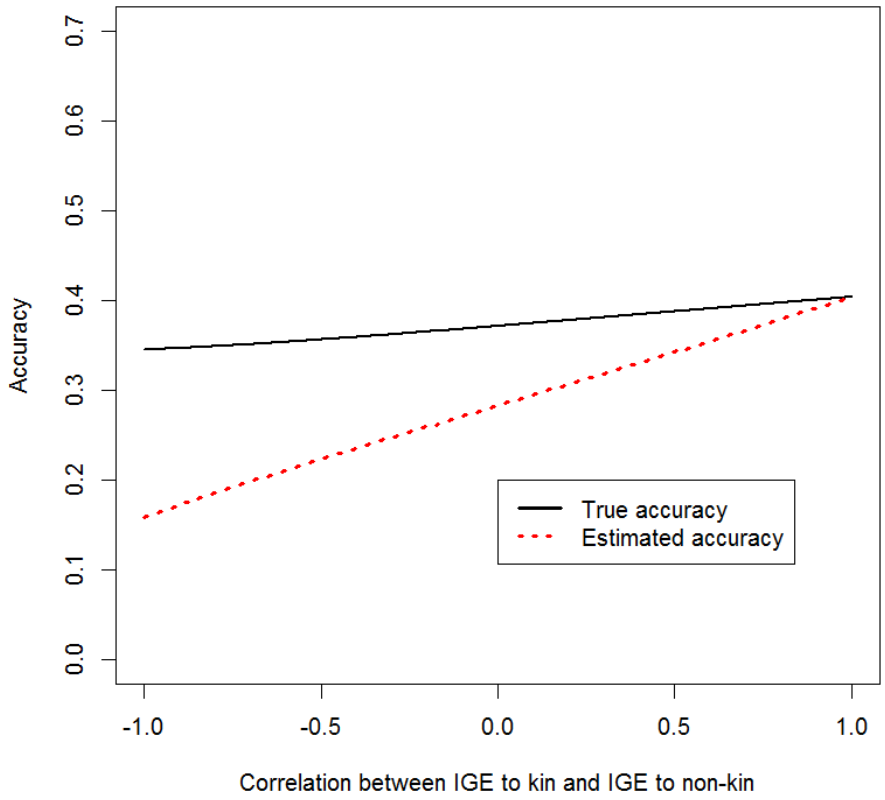

Figure 2 True accuracy from full model vs. perceived accuracy from traditional model when $\sigma^2_{A_{Sf}} = \sigma^2_{A_{Su}}$. The following inputs were used to plot this figure $r_{A_{DS_f}} = r_{A_{DS_u}} = h^2{}_D = h^2_{S_f} = 0.5, h^2_{S_u} = 0.2$

For this scenario, the true accuracy using the full model is higher than the perceived accuracy using the traditional-IGE model. Also, the difference in accuracy between the full model and the traditional IGE model is lower as the correlation between IGE to kin and IGE to non-kin is higher.

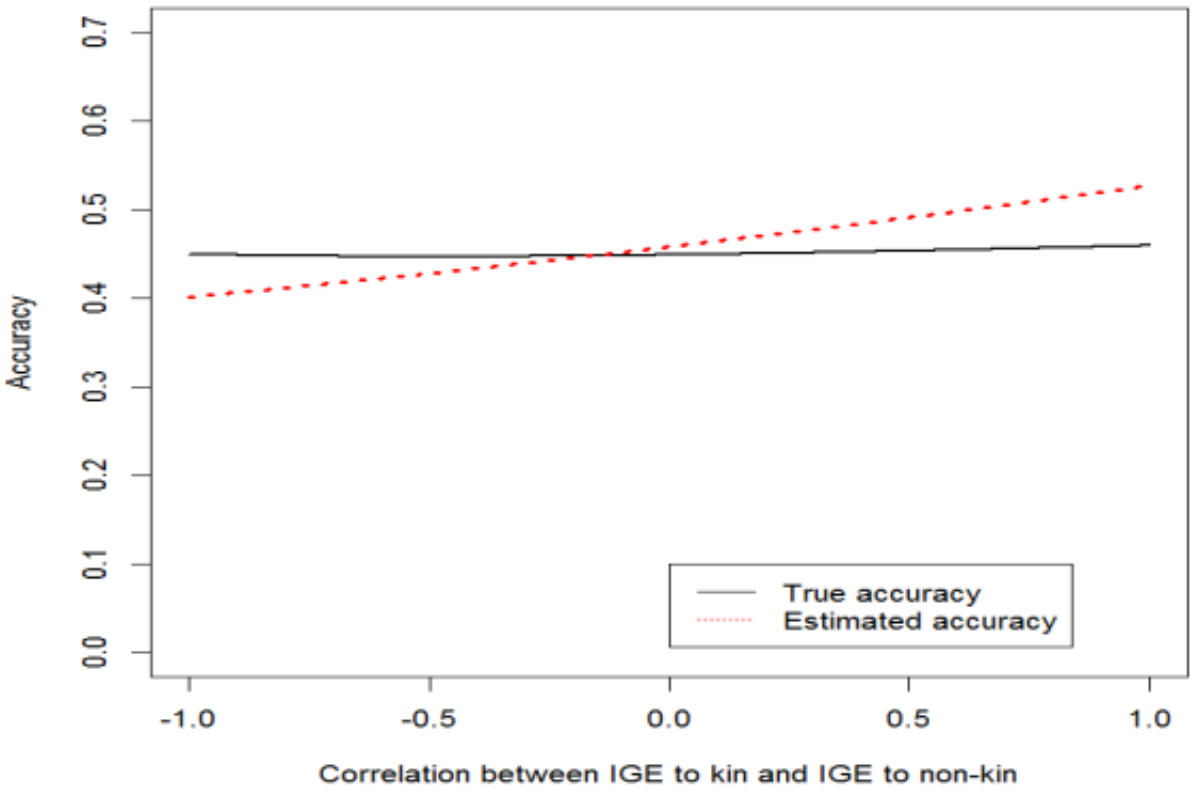

Figure 3 True accuracy from full model vs. perceived accuracy from traditional model when $\sigma^2_{A_{Sf}} = 4\sigma^2_{A_{Su}}$. The following inputs were used to plot this figure $r_{A_{DS_f}} = r_{A_{DS_u}} = h^2{}_D = h^2_{S_f} = 0.5, h^2_{S_u} = 0.2$

For this scenario, the true accuracy using the full model is higher than the perceived accuracy using the traditional-IGE model when the correlation between IGE to kin and IGE to non-kin is higher than zero. The true accuracy using the full model is lower than the perceived accuracy using the traditional-IGE model when the correlation between IGE to kin and IGE to non-kin is lower than zero. Also, the difference between in true accuracy and perceived accuracy is higher as the absolute value of the correlation between IGE to kin and IGE to non-kin is higher.

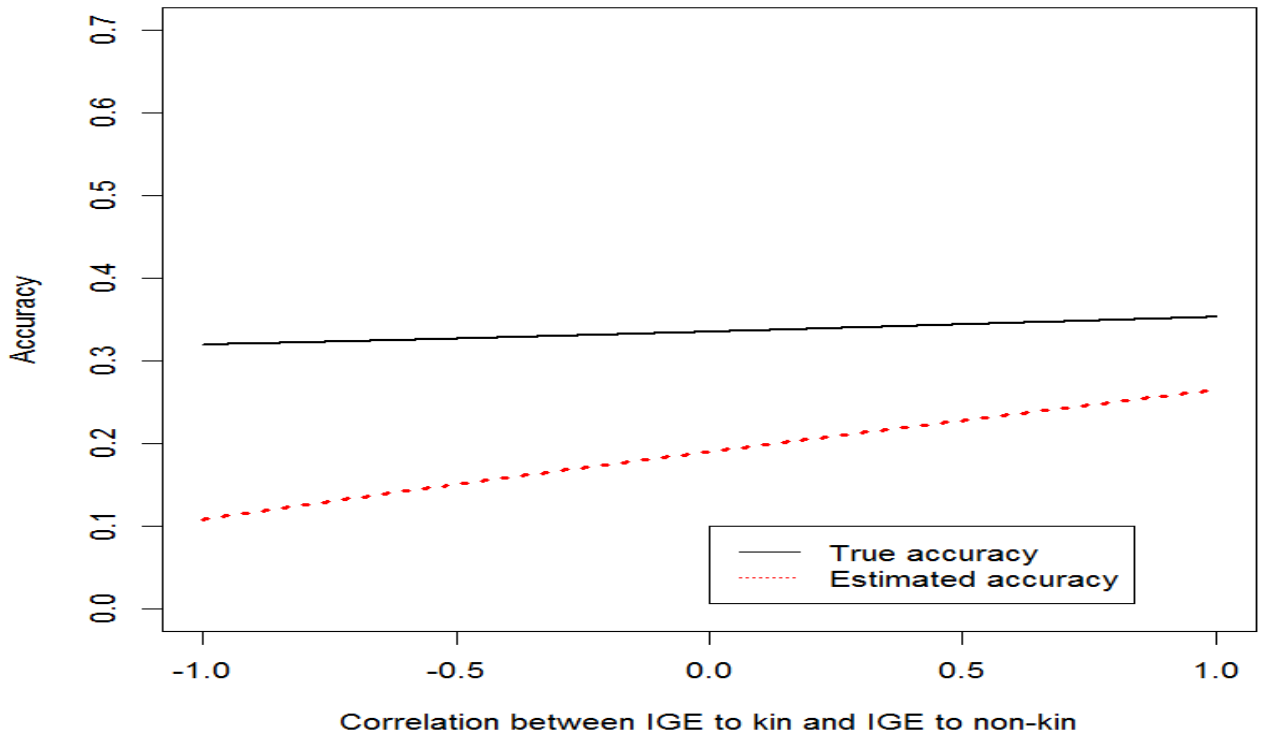

Figure 4 True accuracy from full model vs. perceived accuracy from traditional model when $\sigma^2_{A_{Sf}} = 0.25\sigma^2_{A_{Su}}$.

The following inputs were used to plot this figure $r_{A_{DS_f}} = r_{A_{DS_u}} = h^2{}_D = h^2_{S_u} = 0.5, h^2_{S_f} = 0.2$

For this scenario, the true accuracy using the full model is higher than the perceived accuracy using the traditional-IGE model. Also, the difference in accuracy between the full model and the traditional IGE model is lower as the correlation between IGE to kin and IGE to non-kin is higher. This scenario is similar with the first scenario with the following difference. The difference in accuracy in scenario one become smaller than the difference in accuracy in this scenario at higher correlation. At correlation one there is no difference in accuracy in scenario one but there is still difference in this scenario. There should be further investigation on the accuracy difference for different scenarios.

6. 6 Group housing of mink

Naturally, mink is a solitary and territorial species. Mink defend their territory by aggression towards mink of the same sex. The territory of a male can overlap with several females but not with other males (Dunstone, 1993). Thus, the overlap between mink of the same sex is not reported (Dunstone, 1993). The young mink leave the territory occupied by their mother to have their own territory. This dispersion happens when they are 12-16 weeks old (Dunstone, 1993). This characteristic is the main reason for keeping a pair of male and female sibs in a cage in Denmark. Through this type of housing, mink maintain their natural behaviour (e.g. solitary), but it has some limitations such as limited space for the individuals which makes the mink stressful. Group housing of mink (more than two mink) has been suggested as a potential way to improve the welfare.

Group housing of mink is recommended by the Council of Europe (European Commission 1999). This is because it may improve welfare from 'social enrichment' as outlined in (European Commission 2001). It also increases the stocking density in the cages and thereby decreases housing investments. Though group housing offers some advantages, it has still some limitations such as increased food competition and aggression behaviour (Pedersen and Jeppesen, 2001; Moller *et al.*, 2003).

The aggressive behaviour in group-housed mink is higher than in pair-wise housing and it is reflected by increased bite mark in grouped house mink than pair housing (Hansen and Damgaard, 1991; Pedersen and Jeppesen, 2001; Moller *et al.*, 2003). Thus, for continuity of group housing the welfare of mink needs to be improved. For example, mink from group housing should have a low level of biting. One solution to improve the welfare in group-housed mink is improving the

management by the use of environmental enrichment such as plastic tubes (Hansen, 2012). However, the use of plastic tubes is not sufficient (Hansen, 2012).

The other promising solution is genetic selection. Producing mink that have a lower level of aggression using genetic selection is a good solution. One possible way of measuring aggression behaviour in mink is by its consequences, for example, by using bite marks. Bite marks are an excellent indicator of aggression behaviour (Hansen *et al.*, 2014). Thus, genetic selection using bite mark traits can be a solution to reduce aggression behaviours in group-housed mink.

6.6.2 Prospects of genetic selection for reducing bite marks in group-housed mink

In 2009, a selection experiment was initiated to select for a reduced number of bite marks at pelting, at the mink farm at the Research Centre Foulum in Denmark. The experiment lasted three generations from 2009 till 2011. In chapter three, we analysed data from the first three generations of that selection experiment. The genetic parameters for the bite mark traits in the neck, body and tail region as well as total bite mark score were estimated. We found a substantial amount of heritable variation for bite mark traits in group-housed mink for all parts of the body. For example, for total bite mark score the total heritable variation expressed as a proportion of phenotypic variance was about 0.6. Thus, there are good prospects for genetic improvement of the trait. We predicted the response to selection. Using mass selection, the accuracy of selection based on estimates equals ~0.4 and if we used 10% of the population used for breeding the predicted response to selection will be equal to ~3.07 and the total BMS is predicted to reduce from ~6.47 to ~3.4, which is a very substantial reduction in a single generation of selection. When using

group selection for groups of four sibs, two males and two females that all belong to the same family, it is possible to reach an even higher accuracy of ~0.65, and thus the predicted response to selection will be ~5 and the total BMS is predicted to reduce from ~6.47 to ~1.47 in a single generation of selection (see chapter three for more details). However, mass selection and group selection are difficult to use, since currently bite mark scores are recorded on the pelts of the dead animals. Hence, recording the trait would require sacrificing the selection candidates. Thus sib selection is more appropriate. The predicted accuracy of sib selection for groups of four sibs, two males and two females that all belong to the same family will be equal to ~ 0.54 and the predicted response to selection to ~4.14. Thus, total BMS will be reduced from ~6.47 to ~ 3.33, again a very substantial reduction in a single generation of selection (Chapter four). I concluded that within a few generations, for instance three, it is possible to produce mink that have a much lower level of biting.

This result is supported by a previous experiment at a research farm in Ederveen in the Netherlands, where no systematic difference in the number of bite wounds and bite mark between group housing and pair wise housing was found (de Rong J. and van Willigen, 2012). The reason might be that the mink in the Netherlands are adapted for group housing because group housing has been practised since the 1990's. Thus, there are good prospect for producing mink that have a lower level of biting using genetic selection.

We found that bite mark score is a highly heritable trait and that there are good prospects for genetic improvement. Based on the results so far I hypothesize that bite mark for group-housed mink will be as low as for the pair-housed mink in later generations, for example about generation three. To test this hypothesis, there

should be a selection experiment for number of generations. The selection experiment should consist of control line which has pair-housed mink and a selection line which has group-housed mink. The genetic selection will take place only in selection line. This experiment should last at least three generations.

High responses to selection have been found in a selection experiment against mortality in laying hens. Genetic selection using group selection reduced the mortality of the layer line from 68% in generation two to 9 % in generation six (Muir, 1996). In the sixth generation, the mortality of the group-housed layers was similar to mortality of layers placed in single bird cages. Thus, it is essential to compare bite marks in group-housed vs. bite marks in pair-housed mink to have good conclusions on the prospects of genetic selection to reduce bite marks. I do preliminary comparisons based on Hansen and Møller (2012) findings on total bite mark comparisons between pair-housed mink and group-housed mink. They compared bite mark between pair-housed mink and group-housed mink in four farms. In one of the farm the pair-housed mink (one male and one female) had total bite mark score of 6.9 and the group-housed mink (two male and two female) had total bite mark score 14.9. (Hansen and Møller, 2012). Thus, the difference in bite mark between group-housed mink and pair-house mink is about 8. If we apply sib selection scheme for this farm, the bite mark in group-housed mink will be the same as bite mark in pair-housed mink within three generations.

There are good prospects of producing mink that have a low level of biting using genetic selection.

References

Akcay C, Swift RJ, Reed VA, Dickinson JL (2013). Vocal kin recognition in kin neighborhoods of western bluebirds. *Behav Ecol* **24:** 898-905.

Alemu SW, Berg P, Janss L, Bijma P (2014a). Indirect genetic effects and kin recognition: estimating IGEs when interactions differ between kin and strangers. *Heredity* **112:** 197-206.

Alemu SW, Bijma P, Moller SH, Janss L, Berg P (2014b). Indirect genetic effects contribute substantially to heritable variation in aggression-related traits in group-housed mink (Neovison vison). *Genet Sel Evol* **46**.

Beecher MD (1982). Signature Systems and Kin Recognition. *Am Zool* **22:** 477-490.

Bijma P (2011). A general definition of the heritable variation that determines the potential of a population to respond to selection. *Genetics* **189:** 1347-1359.

Bijma P (2014). The quantitative genetics of indirect genetic effects: a selective review of modelling issues. *Heredity* **112:** 61-69.

Bijma P, Muir WA, Van Arendonk JAM (2007b). Multilevel selection 1: Quantitative genetics of inheritance and response to selection. *Genetics* **175:** 277-288.

Bijma P, Muir WM, Ellen ED, Wolf JB, Van Arendonk JAM (2007a). Multilevel selection 2: Estimating the genetic parameters determining inheritance and response to selection. *Genetics* **175:** 289-299.

Bijma P, Muir WM, Van Arendonk JAM (2007). Multilevel Selection 1: Quantitative Genetics of Inheritance and Response to Selection. *Genetics* **175:** 277-288.

Blaustein AR (1983). Kin Recognition Mechanisms - Phenotypic Matching or Recognition Alleles. *Am Nat* **121:** 749-754.

Blaustein AR, Ohara RK (1981). Genetic-Control for Sibling Recognition. *Nature* **290:** 246-248.

Calus MPL, Meuwissen THE, de Roos APW, Veerkamp RF (2008). Accuracy of genomic selection using different methods to define haplotypes. *Genetics* **178:** 553-561.

Cooke F, Seiger MB, Mirsky PJ (1972). Color Preferences in Lesser Snow Goose and Their Possible Role in Mate Selection. *Can J Zoolog* **50:** 529-&.

Coulon M, Baudoin C, Abdi H, Heyman Y, Deputte BL (2010). Social behavior and kin discrimination in a mixed group of cloned and non cloned heifers (Bos taurus). *Theriogenology* **74:** 1596-1603.

D'Eath RB (2004). Consistency of aggressive temperament in domestic pigs: The effects of social experience and social disruption. *Aggressive Behav* **30:** 435-448.

D'Eath RB, Keeling LJ (2003). Social discrimination and aggression by laying hens in large groups: from peck orders to social tolerance. *Appl Anim Behav Sci* **84:** 197-212.

De Rong J. , van Willigen GCK (2012). Mortality in pair- and group-housed mink after weaning. Proceedings of the Xth International Scientific Congress in Fur Animal production , Scientifur 36, 103-105.

Duijvesteijn N (2014). Sociable Swine: Prospects of indirect genetic effects for the improvement of productivity, welfare and quality, PhD thesis, Wageningen university, Wageningen.

Dunstone N (1993). *The mink. T &A.D. Poyster, London.*

Erhard HW, Mendl M (1997). Measuring aggressiveness in growing pigs in a resident-intruder situation. *Appl Anim Behav Sci* **54:** 123-136.

European Commission (1999). Standing committee of the European convention for the protection of animals kept for farming purposes: Recommendation concerning fur animals. European Commission; 1999. http://.www.coe.int/.

European Commission (2001). Scientific Committee on Animal Health and Animal Welfare: The welfare of animals kept for fur production. European Commission;. http://ec.europa.eu/food/animal/welfare/international/out67_en.pdf.

Falconer DS (1960). *Introduction to quantitative genetics*. Oliver etc.: Edinburgh etc.

Frank SA (2007). All of life is social. *Current Biology* **17:** R648-R650.

Frommen JG, Luz C, Bakker TCM (2007). Kin discrimination in sticklebacks is mediated by social learning rather than innate recognition. *Ethology* **113:** 276-282.

Gamboa GJ, Berven KA, Schemidt RA, Fishwild TG, Jankens KM (1991). Kin Recognition by Larval Wood Frogs (Rana-Sylvatica) - Effects of Diet and Prior Exposure to Conspecifics. *Oecologia* **86:** 319-324.

Gavish L, Hofmann JE, Getz LL (1984). Sibling Recognition in the Prairie Vole, Microtus-Ochrogaster. *Anim Behav* **32:** 362-366.

Gerlach G, Hodgins-Davis A, Avolio C, Schunter C (2008). Kin recognition in zebrafish: a 24-hour window for olfactory imprinting. *P Roy Soc B-Biol Sci* **275:** 2165-2170.

Giersing M, Andersson A (1998). How does former acquaintance affect aggressive behaviour in repeatedly mixed male and female pigs? *Appl Anim Behav Sci* **59:** 297-306.

Goddard M (2009). Genomic selection: prediction of accuracy and maximisation of long term response. *Genetica* **136:** 245-257.

Greenberg L (1979). Genetic Component of Bee Odor in Kin Recognition. *Science* **206:** 1095-1097.

Griffing B (1967). Selection in reference to biological groups .I. Individual and group selection applied to populations of unordered groups. *Australian Journal of Biological Sciences* **20:** 127-139.

Habier D, Fernando RL, Dekkers JCM (2007). The impact of genetic relationship information on genome-assisted breeding values. *Genetics* **177:** 2389-2397.

Hagen K, Broom DM (2003). Cattle discriminate between individual familiar herd members in a learning experiment. *Appl Anim Behav Sci* **82:** 13-28.

Hain TJA, Neff BD (2006). Promiscuity drives self-referent kin recognition. *Curr Biol* **16:** 1807-1811.

Halpin ZT, Hoffman MD (1987). Sibling Recognition in the White-Footed Mouse, Peromyscus-Leucopus - Association or Phenotype Matching. *Anim Behav* **35:** 563-570.

Hamilton WD (1964). Genetical Evolution of Social Behaviour *J Theor Biol* **7:** 1-16.

Hansen S (2012). Plastrør og halmbriketter reducerer pelsgnav. (Ed Steffen W.hansen og birthe M Damgaard) Temadag om aktuel Minkforsking - DCA Rapport nr. 010, spetember 2012. Aarhus university.p. 13-19.

Hansen SW, Damgaard BM (1991). Stress Physiological, Hematological and Clinical-Chemical Status of Farm Mink Placed in Groups or Singly. *Acta Agr Scand* **41:** 355-366.

Hansen SW, Moller SH, Damgaard BM (2014). Bite marks in mink-Induced experimentally and as reflection of aggressive encounters between mink. *Appl Anim Behav Sci* **158:** 76-85.

Hansen SW, Møller SH. (2012). *Proceeding of the Xth International Congress in fur animal produktion, Scientifur volume 36(3/4), 350-359.*

Henderson CR (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics* **31:** 423-447.

Hepper PG (1986). Kin Recognition - Functions and Mechanisms a Review. *Biol Rev* **61:** 63-93.

Hesse S, Bakker TCM, Baldauf SA, Thunken T (2012). Kin recognition by phenotype matching is family- rather than self-referential in juvenile cichlid fish. *Anim Behav* **84:** 451-457.

Hill WG (1993). Variation in Genetic Composition in Backcrossing Programs. *J Hered* **84:** 212-213.

Holmes WG (1986). Kin Recognition by Phenotype Matching in Female Belding Ground-Squirrels. *Anim Behav* **34:** 38-47.

Holmes WG, Sherman PW (1982). The Ontogeny of Kin Recognition in 2 Species of Ground-Squirrels. *Am Zool* **22:** 491-517.

Holmes WG, Sherman PW (1983). Kin Recognition in Animals. *Am Sci* **71:** 46-55.

Horrell I, Hodgson J (1992a). The Bases of Sow-Piglet Identification .1. The Identification by Sows of Their Own Piglets and the Presence of Intruders. *Appl Anim Behav Sci* **33:** 319-327.

Horrell I, Hodgson J (1992b). The Bases of Sow-Piglet Identification .2. Cues Used by Piglets to Identify Their Dam and Home Pen. *Applied Animal Behaviour Science* **33:** 329-343.

Khaw HL, Ponzoni RW, Yee HY, Aznan bin Aziz M, Bijma P (submitted). Genetic and non-genetic indirect effects for harvest weight in the GIFT strain of Nile tilapia (*Oreochromis niloticus*).

Ligout S, Porter RH (2003). Social discrimination in lambs: the role of indirect familiarization and methods of assessment. *Animal Behaviour* **65:** 1109-1115.

Ligout S, Porter RH, Bon R (2002). Social discrimination in lambs: persistence and scope. *Appl Anim Behav Sci* **76:** 239-248.

Lynch M, Walsh B (1998). *Genetics and analysis of quantitative traits*. Sinauer Associates, Sunderland, MA.

Mateo JM, Johnston RE (2000). Kin recognition and the 'armpit effect': evidence of self-referent phenotype matching. *Proceedings of the Royal Society B-Biological Sciences* **267:** 695-700.

McGlothlin JW, Brodie ED (2009). How to Measure Indirect Genetic Effects: The Congruence of Trait-Based and Variance-Partitioning Approaches. *Evolution* **63:** 1785-1795.

Meuwissen T, Hayes B, Goddard M (2001a). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157:** 1819-1829.

Meuwissen THE, Hayes BJ, Goddard ME (2001b). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157:** 1819-1829.

Moller SH, Hansen SW, Sorensen JT (2003). Assessing animal welfare in a strictly synchronous production system: The mink case. *Animal Welfare* **12:** 699-703.

Moore AJ, Brodie ED, Wolf JB (1997). Interacting phenotypes and the evolutionary process .1. Direct and indirect genetic effects of social interactions. *Evolution* **51:** 1352-1362.

Muir WM (1996). Group selection for adaptation to multiple-hen cages: selection program and direct responses. *Poultry Science* **75:** 447-458.

Muir WM (2005). Incorporation of competitive effects in forest tree or animal breeding programs. *Genetics* **170:** 1247-1259.

Odegard J, Meuwissen THE (2012). Estimation of heritability from limited family data using genome-wide identity-by-descent sharing. *Genet Sel Evol* **44**.

Odegard J, Meuwissen THE, Heringstad B, Madsen P (2010). A simple algorithm to estimate genetic variance in an animal threshold model using Bayesian inference. *Genet Sel Evol* **42:** 1-7.

Olsen KH (1989). Sibling Recognition in Juvenile Arctic Charr, Salvelinus-Alpinus (L). *J Fish Biol* **34:** 571-581.

Olsen KH, Grahn M, Lohm J, Langefors A (1998). MHC and kin discrimination in juvenile Arctic charr, Salvelinus alpinus (L.). *Anim Behav* **56:** 319-327.

Pedersen V, Jeppesen LL (2001). Effects of family housing on behaviour, plasma cortisol and performance in adult female mink (Mustela vison). *Acta Agriculturae Scandinavica Section a-Animal Science* **51:** 77-88.

Porter RH, Desire L, Bon R, Orgeur P (2001). The role of familiarity in the development of social recognition by lambs. *Behaviour* **138:** 207-219.

Porter RH, Tepper VJ, White DM (1981). Experiential Influences on the Development of Huddling Preferences and Sibling Recognition in Spiny Mice. *Dev Psychobiol* **14:** 375-382.

Quinn TP, Busack CA (1985). Chemosensory Recognition of Siblings in Juvenile Coho Salmon (Oncorhynchus-Kisutch). *Anim Behav* **33:** 51-56.

Sherman P, Reeve H, Pfennig D (1997). Recognition Systems. In: Krebs J and Davies N (eds) *Behavioural Ecology, 4th Edition*. Oxford: Blackwell Science, pp 69-96.

Stookey JM, Gonyou HW (1998). Recognition in swine: recognition through familiarity or genetic relatedness? *Appl Anim Behav Sci* **55:** 291-305.

Tang-Martinez Z (2001). The mechanisms of kin discrimination and the evolution of kin recognition in vertebrates: a critical re-evaluation. *Behav Process* **53:** 21-40.

Visscher PM, Medland SE, Ferreira MAR, Morley KI, Zhu G, Cornes BK *et al.* (2006). Assumption-free estimation of heritability from genome-wide identity-by-descent sharing between full siblings. *Plos Genet* **2:** 316-325.

Waldman B (1981). Sibling Recognition in Toad Tadpoles - the Role of Experience. *Z Tierpsychol* **56:** 341-358.

Waldman B (1987). Mechanisms of Kin Recognition. *J Theor Biol* **128:** 159-185.

Waldman B, Frumhoff PC, Sherman PW (1988). Problems of Kin Recognition. *Trends Ecol Evol* **3:** 8-13.

Willham RL (1963). Covariance between Relatives for Characters Composed of Components Contributed by Related Individuals. *Biometrics* **19:** 18-&.

Wilson AJ, Gelin U, Perron MC, Reale D (2009). Indirect genetic effects and the evolution of aggression in a vertebrate system. *P Roy Soc B-Biol Sci* **276:** 533-541.

Wilson AJ, Morrissey MB, Adams MJ, Walling CA, Guinness FE, Pemberton JM *et al.* (2011). Indirect genetics effects and evolutionary constraint: an analysis of social dominance in red deer, *Cervus elaphus*. *J Evol Biol* **24:** 772-783.

Wolf JB, Brodie ED, Cheverud JM, Moore AJ, Wade MJ (1998). Evolutionary consequences of indirect genetic effects. *Trends Ecol Evol* **13:** 64-69.

Wolfram Research I. (2010). *Vol. Version 8.0*. Wolfram Research, Inc.: Champaign, Illinois.

Wu HMH, Holmes WG, Medina SR, Sackett GP (1980). Kin Preference in Infant Macaca-Nemestrina. *Nature* **285:** 225-227.

Summary

Social interactions among individuals are common both in plants and animals. Social interactions can be cooperative or competitive. In both types of social interactions the trait value of an individual may be influenced by genes of its interacting partners, a phenomenon known as indirect genetic effects (IGE). An IGE is heritable effect of an individual on the trait value of interacting partner. A large body of literature confirmed that social interactions may create addition heritable variation in both plants and animals, for both behavioural and production traits.

When IGE are estimated, it is usually assumed that an individual interacts equally to all its social partners, irrespective of genetic relatedness. This assumption may not be true in mixed groups of kin and non-kin, as suggested by kin selection theory. Kin selection theory states that individuals interact systematically different to kin vs. non kin. One obvious reason for systematic interaction is kin recognition. Kin recognition is the ability of an individual to recognize kin and interact differently with kin vs. non kin. Thus, in mixed groups of kin and non-kin, an individual may distribute different IGEs to kin vs. non-kin. We, therefore, need a method that can estimate IGEs when interactions differ between kin and non-kin. In **Chapter two**, we developed a statistical method to estimate IGEs when interactions differ between kin and non-kin. The results show that not all genetic parameters are statistically identifiable. However, a genetic parameters of a family breeding value and of IGE expressed on non-kin can be estimated.

Social interactions are important are mink kept for the production of fur. Recently, group housing of mink has become common. Group housing of mink increases aggression behaviours, which is reflected by an increase in the number of bite marks on the pelts, and reduces the welfare of the animals. Thus, for continuity of group housing, the biting in mink should be reduced. One solution to improve the welfare in group-housed mink is to

improve the management, for example by the use of environmental enrichment, such as plastic tubes. However, such measures do not reduce the level of biting sufficiently. Another promising solution is genetic selection.

To judge the prospects for genetic improvement, we estimated the genetic parameter for bite mark trait in group housed mink in **Chapter three**. For total bite mark score, we found a total genetic variance equal to 61% of phenotypic variance, indicating very good prospects for genetic improvement. Results showed that a substantial amount of heritable variation originated from IGE. We concluded there are good prospects to produce mink that have a low level of biting. In the analysis in Chapter three we did not fully investigate systematic interactions due to sex or kin.

In **Chapter four** we further investigated estimation of IGE for bite mark traits in group-housed mink, by taking in to account systematic interactions due to the sex of individuals or the family relationship between individuals ("kin"). We found that IGE estimation needs to take in to account systematic interactions due to sex or kin in group-housed mink. Ignoring such systematic interactions biased estimates of all genetic parameters.

Bite mark traits are recorded after life, after the mink are culled and pelleted. Thus recording of bite mark traits requires sacrificing the individual. Thus, we will not have own performance records for bite mark traits on the selection candidates. We can use sib information or progeny testing. With sib information the accuracy of the EBV is limited, and progeny testing increases the generation interval. Thus, using sib information or progeny testing may reduce the response to selection. We, therefore, need a new genetic tool such as genomic selection to increase the response to selection.

Due to unavailability of genomic data on mink, we investigated genomic selection for socially affected traits by considering survival time in two lines of brown egg layers. Thus,

in **Chapter five** we investigated whether genomic selection can increase the accuracy of EBVs and the response to selection compared with a traditional breeding programme for survival time in layers. Despite the limited reference population of ~234 progeny tested sires, the accuracy of estimating the breeding value was ~35% higher for genomic selection compared with parental average-EBV. We found that the response to genomic selection per year for line B1 was substantially higher than for the traditional breeding scheme, whereas for line BD response was slightly higher than for the traditional breeding scheme.

In the **general discussion** in chapter six, I discuss five topics that are related to the thesis. The first topic is kin recognition mechanisms, and I describe four kin recognition mechanisms found in the literature. The second topic is the use of genome wide marker to estimate all the genetic parameters when IGE differ between kin and non-kin. I argue that it is possible to estimate all genetic parameters when using genome wide markers. As the third topic focusses, I present a trait-based model for the case when IGE differ between kin and non-kin. The theoretical relationship between the variance component model and the trait-based model can be developed, but estimating of the parameters of the trait-based model is difficult. Thus, it is very difficult to use empirically the trait-based model when IGE differ between kin and non-kin. The forth topic is the accuracy of the estimated breeding value when IGE differ between kin and non-kin. I found that, when IGE differ between kin and non-kin, the accuracy of estimating the total breeding value using mass selection most likely will be biased if we use a statistical analysis that ignores the systematic difference in interactions between kin and non-kin. As final topic I discuss the prospects of genetic selection for reducing bite marks in group-housed mink. I conclude that there are good prospects of producing mink that have a low level of biting using genetic selection, within a few generations.

Sammendrag

Sociale interaktioner mellem individer er almindelige både i planter og dyr. Sociale interaktioner kan være kooperative eller kompetitive. I begge typer af sociale interaktioner kan egenskaber hos èt individ være påvirket af gener hos sine interagerende partnere, et fænomen kendt som indirekte genetiske effekter (IGE). IGE er en arvelig effekt af et individ på egenskaben hos en interagerende partner. En stor mængde litteratur bekræftede, at sociale interaktioner kan skabe yderligere arvelig variation i både planter og dyr, for både adfærdsmæssige og produktionsegenskaber.

Når IGE estimeres, antages det normalt, at et individ interagerer ligeligt med alle sine sociale partnere, uanset genetisk slægtskab. Denne antagelse er ikke nødvendigvis rigtig i blandede grupper af beslægtede og ubeslægtede, som foreslået af "kin selection theory". "Kin selection theory" forudsiger, at individer interagerer systematisk anderledes med beslægtede vs. ikke beslægtede individer. En oplagt årsag til systematisk forskellig interaktion er genkendelse af slægtninge. Genkeldelse af slægtninge er et individs evne til at genkende beslægtede individer og interagere forskelligt med beslægtede vs. ikke beslægtede I blandede grupper af beslægtede og ubeslægtede, kan et individ bidrage med en forskellig IGE til beslægtede vs. ubeslægtede. Vi har derfor brug for en metode, der kan estimere IGE's når interaktioner varierer mellem beslægtede og ubeslægtede. I **kapitel to**, udviklede vi en statistisk metode til at estimere IGE'er når interaktioner er forskellige mellem beslægtede og ubeslægtede. Resultaterne viser, at ikke alle genetiske parametre er statistisk identificerbare. Men genetiske parametre for en familie avlsværdi og IGE udtrykt på ubeslægtede individer kan estimeres.

Sociale interaktioner er vigtige for mink der holdes til pelsproduktion. For nylig er gruppeindhusning af mink blevet almindelig. Gruppeindhusning af mink øger aggressiv

adfærd, hvilket afspejles i en stigning i antallet af bidmærker på skindene, og reducerer dyrenes velfærd. For fortsat brug af gruppeindhusning er der behov for at bid og aggressiv adfærd reduceres. En løsning til at forbedre velfærden i gruppeindhuste mink er at forbedre management, for eksempel ved brug af miljøberigelse, såsom plastrør. Men sådanne foranstaltninger reducerer ikke niveauet af bid tilstrækkeligt. En anden lovende løsning er genetisk selektion.

For at vurdere perspektivet for genetisk forbedring, estimerede vi de genetiske parametre for bidmærker i gruppeindhuste mink i **kapitel tre**. For total bidmærke score, fandt vi en total genetisk varians på 61% af fænotypisk varians, hvilket indikerer meget gode muligheder for genetisk forbedring. Resultaterne viste, at en væsentlig del af den arvelige variation stammede fra IGE. Vi konkluderede, at der er gode muligheder for at producere mink, der har et lavt niveau af bid. I analysen i kapitel tre undersøgte vi ikke fuldt ud systematiske interaktioner på grund af køn eller pårørende.

I **kapitel fire** undersøgte vi yderligere estimering af IGE for bidmærke egenskaber i gruppeindhuste mink, ved at tage hensyn til systematiske interaktioner på grund af køn eller slægtskab mellem individer. Vi fandt, at estimering af IGE skal tage hensyn til systematiske interaktioner på grund af køn eller slægtskab i gruppeindhuste mink. Ignoreres disse systematiske interaktioner opstår bias i estimater for alle genetiske parametre.

Bidmærker registreres efter at mink er aflivet , og forudsætter derfor at dyret er aflivet. Vi vil derfor ikke have observationer for bidmærker på de dyr der er kandiater til avl. Vi kan alternativt anvende sæskende information eller afkomsprøver. Med søskende information er sikkerheden på EBV begrænset, og afkomsprøver øger generations intervallet. Således

kan anvendelse af søskende information eller afkomsprøver reducere avlsfremgangen. Vi har derfor brug for et nyt genetisk værktøj som genomisk selektion for at at øge avlsfremgangen.

På grund af manglende adgang til genomiske data på mink, undersøgte vi genomisk selektion for socialt påvirkede egenskaber ved at analysere overlevelsestid i to linier af brune æglæggere. I **kapitel fem**, undersøgte vi derfor om genomisk selektion kan øge sikkerheden på EBVs og avlsfremgangen sammenlignet med et traditionelt avlsprogram for overlevelsestid i æglæggere. På trods af den begrænsede reference population på ~ 234 afkomstestede haner, var sikkerheden på estimerede avlsværdier ~ 35% højere for genomisk selektion i forhold til forældrenes gennemsnit-EBV. Vi fandt, at avlsfremganegn med genomisk selektion for linje B1 var væsentligt højere end for den traditionelle avlsplan, mens der for linje BD var lidt højere avlsfremgang end for den traditionelle avlsplan.

I den **generelle diskussion** i kapitel seks, diskuteres fem emner, der er relateret til specialet. Det første emne er mekanismer for genkendelse af slægtninger, og jeg beskriver fire mekanismer for genkendelse af slægtninge fundet i litteraturen. Det andet emne er brugen af markører på tværs af genomet for at estimere alle de genetiske parametre, når IGE varierer mellem slægtninge og ikke-slægtninge. Det fremhæves at det er muligt at estimere alle genetiske parametre ved brug af markører på tværs af genomet. Som det tredje emne præsenterer jeg en egenskabs-model for det tilfælde hvor IGE er forskellig mellem slægtninge og ikke-slægtninge. Det teoretiske forhold mellem varians komponent modellen og egenskabs modellen kan udvikles, men estimering af parametrene for egenskabs modellen er vanskelig. Det er derfor meget vanskeligt at empirisk anvende

egenskabs modellen, når IGE varierer mellem slægtninge og ikke-slægtninge. Det fjerde emne er sikkerheden på den estimerede avlsværdi når IGE varierer mellem slægtninge og ikke-slægtninge. Jeg fandt, at når IGE er forskellig mellem slægtninge og ikke-slægtninge, vil sikkerheden på estimering af den samlede avlsværdi sandsynligvis være biased hvis vi bruger en statistisk analyse, der ignorerer den systematiske forskel i interaktioner mellem slægtninge og ikke-slægtninge. Som sidste emne, diskuteres mulighederne for genetisk selektion for at reducere bidemærker i gruppeindhuste mink. Jeg konkluderer, at der er gode udsigter til at producere mink, der har et lavt niveau af bidmærker, ved hjælp af genetisk udvælgelse i nogle få generationer.

samenvatting

Sociale interacties tussen individuen zijn wijdverspreid in zowel planten als dieren. Dergelijke sociale interacties kunnen coöperatief of competitief van karakter zijn. In beide gevallen kunnen eigenschappen van dieren beïnvloed worden door genen in hun sociale partners. Dergelijke effecten staan bekend als indirect genetische effecten (IGE). Een indirect genetisch effect is dus een erfelijk effect van een individu op de kenmerken van zijn sociale partners. Uit de wetenschappelijke literatuur blijkt dat indirect genetische effecten additionele genetische variatie creëren in populaties van dieren en planten, voor zowel gedragskenmerken als productiviteit.

Bij het in kaart brengen van indirect genetische effecten wordt meestal aangenomen dat individuen dezelfde interactie vertonen met al hun sociale partners, ongeacht de genetische verwantschap met die partners. Dat is waarschijnlijk niet het geval als groepen uit een mix bestaan van verwante en onverwante individuen. Kin selectie theorie suggereert dat individuen systematisch anders interacteren met verwante individuen dan met onverwanten. Als individuen verwanten kunnen herkennen, dan kan dit leiden tot systematisch verschillende interacties tussen verwanten en vreemden. In gemende groepen zouden individuen dus systematisch anders kunnen interacteren met verwanten en vreemden. Om dit fenomeen in kaart te brengen zijn methodieken nodig die een onderscheid maken tussen indirect genetische effecten op verwante en onverwante sociale partners. In **hoofdstuk 2** wordt een statistische methodiek ontwikkeld om indirect genetische effecten te schatten wanneer interacties systematisch verschillen tussen verwante en onverwante individuen. Uit de resultaten blijkt dat niet alle genetische parameters statistisch identificeerbaar zijn, maar dat de fokwaarde voor effecten op de familie en voor effecten op onverwante dieren wel gescheiden kunnen worden.

Sociale interacties spelen een belangrijke rol in nertsen die gehouden worden voor de productie van bont. In nertsen is recentelijk groepshuisvesting ingevoerd. In groepshuisvesting komt meer agressief bedrag voor, wat leidt tot bijtplekken in de pelzen en verminderd welzijn van de dieren. Om nertsen in groepen te kunnen houden moet het bijtgedrag dus worden verminderd. Verrijking van de kooien, bijvoorbeeld met plastic buisjes, is een manier om het bijtgedrag te verminderen maar heeft onvoldoende resultaat. Genetische selectie is een andere mogelijkheid.

Om de mogelijkheden voor fokkerij tegen bijtgedrag te onderzoeken zijn in **hoofdstuk 3** genetische parameters geschat voor bijtplekken in de pelzen van nertsen gehouden in groepen. De geschatte erfelijke variatie voor het totaal aantal bijtplekken in een pels bedroeg 61% van de fenotypische variatie. Uit de resultaten blijkt ook dat een groot deel van de erfelijke variatie door indirect genetische effecten wordt verklaard. Samenvattend betekent dit dat er zeer goede mogelijkheden zijn voor fokkerij tegen bijtplekken. In hoofdstuk 3 is niet gekeken of sociale interacties in nertsen systematisch verschillen tussen verwante en onverwante dieren, of tussen de beide seksen.

In **hoofdstuk 4** zijn indirect genetische effecten voor bijtplekken bij nertsen verder onderzocht, waarbij er een onderscheid is gemaakt tussen de seksen en tussen verwante en onverwante dieren. Uit de resultaten blijkt dat sociale interacties systematisch verschillen, ofwel tussen de seksen of tussen verwante en onverwante dieren. Hiermee moet in het statistisch model rekening worden gehouden om schattingsfouten in de genetische parameters te voorkomen.

Bijtplekken in pelzen worden gemeten aan de binnenkant van de pels, nadat de nertsen zijn gedood en gestroopt. Om het aantal bijtplekken te tellen moet een dier

dus gedood worden, en kan dan niet meer gebruikt worden voor de fokkerij. Dus fokdieren kunnen niet worden geselecteerd op basis van hun eigen aantal bijtplekken. Dit bemoeilijkt de fokkerij tegen bijtgedrag. In principe kan gebruik worden gemaakt van gegevens aan bijtplekken die gemeten zijn aan broers en zussen of aan nakomelingen. Dergelijke fokprogramma's hebben echter een lagere nauwkeurigheid of een hoger generatie-interval, wat leidt tot minder genetische verbetering. Voor dit soort situaties is zgn. genomische selectie een veelbelovende fokmethode om de genetische vooruitgang te versnellen. Op dit moment zijn er echter geen DNA gegevens om genomische selectie bij nertsen mogelijk te maken. Daarom is er in **hoofdstuk 5** een vergelijkbare case bij legkippen onderzocht, waarvoor wel DNA gegevens beschikbaar zijn. Legkippen vertonen kannibalisme, waardoor er aanzienlijke sterfte kan optreden als het puntje van de snavel van kippen niet wordt verwijderd. In hoofdstuk 5 is gekeken naar de mogelijkheden van genomische selectie voor levensduur, in twee lijnen van bruine legkippen waarvan de snavels intact waren. De nauwkeurigheid van genomische fokwaardes voor levensduur is vergeleken met die van klassieke fokwaardes, en de mate van genetische verbetering is vergeleken tussen beide methodieken. Ondanks dat de referentie-populatie op dit moment nog maar klein is (ca. 234 hanen met nakomelingen) was de nauwkeurigheid van genomische fokwaardes ca. 35% hoger dan van klassieke fokwaardes op basis van afstamming. Voor beide onderzochte lijnen was de genetische verbetering van een fokprogramma met genomische selectie hoger dan van het traditionele fokprogramma.

De algemene discussie in hoofdstuk 6 bespreekt vijf onderwerpen die gerelateerd zijn aan dit proefschrift. Als eerste onderwerp worden vier mechanismen besproken die het herkennen van verwanten mogelijk maken. Vervolgens wordt beargumenteerd

dat DNA gegevens het mogelijk maken om alle genetische parameters te schatten als indirect genetische effecten verschillen tussen verwanten en onverwanten. Als derde onderwerp wordt een zgn. kenmerk-gebaseerd model gepresenteerd, voor indirect genetische effecten die verschillen tussen verwanten en onverwanten. De theoretische relatie van dit model met het variantie-componenten model dat in de rest van het proefschrift wordt gebruikt wordt uitgewerkt. De resultaten laten zien dat het zeer lastig is om de parameters van het kenmerk-gebaseerde model te schatten. Dit betekent dat het kenmerk-gebaseerde model in de praktijk erg lastig te gebruiken is als indirecte effecten verschillen tussen verwanten en onverwanten. Als vierde onderwerp wordt de nauwkeurigheid van fokwaardes besproken, wanneer indirect genetische effecten verschillen tussen verwanten en onverwanten. De resultaten laten zien dat de nauwkeurigheid van de zgn. totale fokwaarde niet goed wordt geschat als in de statistische analyse genegeerd wordt dat indirect genetische effecten verschillen tussen verwante en onverwante individuen. Het laatste onderwerp dat aan de orde komt is de mogelijkheid om bijtplekken te verminderen door middel van fokkerij in nertsen gehouden in groepshuisvesting. Het blijkt dat er goede mogelijkheden zijn om binnen een paar generaties nertsen te fokken die veel minder bijten.

Curriculum Vitae

About the author

Setegn Worku Alemu was born on 28 November 1981 in Gondar, Ethiopia. When he came of school age, he joined the nearby elementary school, Gendewuha elementary school and later completed his high school in Azezo in 2001. Setegn followed his higher education in Animal science at Hawassa University, Ethiopia and obtained BSc in Animal science. Then after, he worked at Wollo University as Instructor. After three years of service he got a scholarship to pursue his education further in Animal Breeding at Wageningen University, The Netherland.. He did his Master with major thesis entitled deriving standard error for genetic parameters for socially affected traits and minor thesis entitled comparison of linkage disequilibrium computed using haplotype vs. genotype. In 2011 he received the best MSc thesis award for his very good performance at Wageningen University. Following his MSc graduation, he started his PhD research entitled as Indirect genetic effect for group housed animals of which the results are described in this thesis. The findings of his PhD project were presented in international conferences and published in journals. Since March 2015 he has got a post doc position at Aarhus University on a project called optimization of breeding program using genomic selection for broiler.

Peer-Reviewed Papers

Alemu, S. W., Berg, P., Janss, L., and Bijma, P. (2014a) **Indirect genetic effects and kin recognition: estimating IGEs when interactions differ between kin and strangers**.*Heredity* **112**, 197-206.

Alemu, S. W., Bijma, P., Moller, S. H., Janss, L., Berg, P. (2014b) **Indirect geneticeffects contribute substantially to heritable variation in aggression-related traits in group-housed mink (Neovison vison)**. Genet Sel Evol 46:30

Alemu, S. W., Berg, P., Janss, L., and Bijma, P. (2015) **Estimation of indirect genetic effects in group-housed mink (Neovison vison) should account for systematic interactions either due to kin or sex**. J. Anim. Breed. Genet. (2015) 1–8.

Alemu S.W., MPL.Calus, W.M.Muir, Katrijn peeters, Addie Vereijken, and P.Bijma. **Genomic prediction of survival time in a population of brown laying hens showing cannibalistic behaviour**. In preparation

Conference proceedings

Alemu SW, Berg P, Janss L, Møller SH, Bijma P. 2014. **Genetic and non-genetic indirect effects for bite mark traits in group housed mink.** Paper presented at 10th World Congress on Genetics Applied to Livestock Production (WCGALP), Vancouver, Canada.

Alemu SW, Berg P, Bjima P. **Estimating indirect genetic effect when interactions differ between kin and non- kin**. In Programme & Book of Abstracts, ICQG 2012. Edinburgh. 2012. p. 161-161.

Alemu SW, Berg P, Bjima P. **Evidence for genetic variation in bite marks in group-housed mink** proceeding of the Xth International Congress in fur animal produktion, Scientifur volume 36, 240-245

Alemu SW, Janss L, Berg P, Bijma P. **Genetic selection can reduce aggression behaviour in group-housed mink**. In Book of Abstracts of the 64th Annual Meeting of the European Association for Animal Production. 2013 ed. Vol. 19. The Netherland: Wageningen Academic Publishers. 2013. p. 171

Berg P, Alemu SW, Møller SH, Janss L, Bijma P. **Kan vi avle for mere sociale mink?**. In Hansen SW, Damgaard BM, editors, Temadag om aktuel minkforskning: DCA rapport nr. 028. Vol. 28. DCA - Nationalt center for fødevarer og jordbrug. 2013. p. 68-71.

Berg P, Møller SH, Alemu SW. **Selektion imod bidskader i gruppeindhusning – resultater fra første generation**. In Berg P, editor, Temadag om aktuel minkforskning. Aarhus Universitet, Forskningscenter Foulum. 2011. p. 48-52. (Intern rapport; No. 109)

Training and supervision

AARHUS UNIVERSITY
DEPARTMENT OF MOLECULAR BIOLOGY AND GENETICS

The Graduate School
WAGENINGEN INSTITUTE of ANIMAL SCIENCES

Basic packages (4 ECTS)

Introduction to PhD Aarhus university	2010
EGS-ABG Summer Research School	2013

Advanced scientific courses (31 ECTS)

Quantitative genetics with integration of genomic selection	2012
Introduction to Statistical Modeling in Agricultural, Biological and Environmental Sciences	2010
Interpretation of stress response	2012
Quantitative Genetics	2012
Genetics of competition	2012
Mixed model in quantitative genetics	2012
Linkage and association study	2013
Advanced statistical and genetical analysis of complex data using ASReml 4	2014
PhD student's discussion groups	2010 ,2014
Quantitative genetics discussion group	

Professional Skills Support Courses (7.5)

Introduction to R in the biological sciences	2011
Techniques for writing and presenting a scientific course	2012
R/Bioconductor Workshop	2012
Academic English	2013

Didactic skills Training (4 ETs)

Assisting Advanced statistics course	2011, 2012

Scientific exposure (9)

International conferences

Proceeding of the Xth International scientific congress in Fur Animal	2012

4th international conference on quantitative genetics	2012
64th Annual meeting of the European Association for Animal Production	2013
10th world congress on genetics applied to livestock production system	2014

Seminars and workshop

WIAs science day,wageningen	2014

Presentations

Evidence for genetic variation in bite marks in group-housed mink *,IFASA ,Oral presentation*	2012
Estimating indirect genetic effect when interactions differ between kin and non-kin,*ICQG, Poster presentation*	2012
Genetic selection can reduce aggression behaviour in group-housed mink., EAAP,oral presentation	2013
Genetic and non-genetic indirect effects for bite mark traits in group housed mink.WCGALP *Poster*	2014

Education and training totals:55.5 ECTS

Acknowledgments

Acknowledgments

This thesis is submitted in partial fulfilment of the requirement for the Joint Doctor of Philosophy Degree (PhD) at Aarhus and Wageningen University. This PhD thesis would not be possible without the support of many people to whom I take this opportunity to express my sincere gratitude.

My sincere thanks go to my promoter, Professor Johan, who supported my steps in the field, inspire my growth as a student and a researcher. Dear Prof.Dr.ir Johan Van Arendonk I sincerely appreciate your patience and productive comment throughout the life of the project. I remain indebted to Dr.ir. Piter Bijma who were instrumental in getting done this project in time with good quality. Dear Piter at this point, I would like to express my special thank as your continuous follow up and guidance were unforgettable. You allowed me to carry out my studies with much freedom and your advice helped me to shape my writing, organizing and analytical skills as a student who wants to be to be a researcher in the field. I am grateful to be your student and indeed I am indebted more than you know. Many thanks go to Dr.Peer Berg for giving me the opportunity to work freely at my own pace, for the valuable inputs while I was writing scientific papers and for advising and arranging to have joint PhD degree, and I am also grateful to Dr.Luc Janss for the support and back up during writing papers and fruitful discussions about Bayesian statistical concepts. Many thanks go to Dr.Steen for his help during writing paper and constructive discussions in mink behavior and Dr. Bernt for fruitful discussion in quantitative genetics and statistics. Besides, Thanks to Karin who for being very helpful in arranging my accommodation and administration stuffs at Aarhus University.

I want to thank all my colleagues at QGG who made my stay in Denmark enjoyable. Zhou Lei, Huiming, Naveen, Mao were unforgettable, we elapsed nice time together. My Ethiopian friends in Denmarks: Gizachew and his wife Meseret, Daniel and his wife Frie, Tsehaye were supportive for me. The time I had with you were energy for me. Thank you all.

During my stay at Wageningen University I had a wonderful time and want to thank all my colleagues and friends at ABG group. My sincere thanks go to Mario Calus, for his critical reflection towards genomic selection, katrijn for the valuable advice in Asreml and BLUPF 90 programmes. I am also grateful to Alex for linux programe, Bas Rodenberg for the productive animal welfare discussion. To My friends Hadi, Tessa, Hamed, utlwanang, Gabriel, Mahlet,Amabel, Peng thanks for having a nice time with you. My special thanks

goes to Ada and Lisette for being sincere and kind in helping me for all the necessary arrangement during my stay at Wageningen. I want to thank Dr. EJ (Evert Jan) Bakker and his families for helpful discussion of statistical concepts and for the nice dinner.

I am also grateful to Ethiopian friends Asmare, Mezgebu, Alemayehu, Biruhtesfa, Mekuria, Tesfahun, Yalemzewd, Dr.Tesema, Dr.Akalu, Dn Gebreyohanes, Dn Tariku, Dn Hailu, Mequanint, Habtamu and Henok for making my stay in Wageningen pleasant and fruitful.

I want to thank all of the members of De Stuiterd (table tennis association of Wageningen University) for having nice time with you. I enjoyed with you playing table tennis in both training secession and competition in different cities of Netherlands.

I want to thank my friends in Ethiopia Birhan, Alexander, Silesh, and Dagne who are always with me in my bad and good times. Finally I want to thank my Mother Birkie Chanie, and Brothers Birhan Worku Alemu and Fikadu Worku Alemu for their continuous help and inspirations. Last but not least I am thankful to my fiancée Helen Tesfaye for her patience while I was in abroad and in taking care of my family.

.

Colophon

Colophon

This research in this thesis was financially supported by Danish Ministry of Food, Agriculture and Fisheries, Program Innovationsloven , Aarhus University and wageningen University.

The Photo of the cover was provided by Aarhus University

The cover of this thesis was designed by

This thesis was printed by