Jisca Huisman

Gene Flow and Natural Selection in Atlantic Salmon

Thesis for the degree of Philosophiae Doctor

Trondheim, April 2012

"Our comforting conviction that the world makes sense rests on a secure foundation: our almost unlimited ability to ignore our ignorance."

Daniel Kahneman,
Thinking, fast and slow

Acknowledgements

There are many people without whom this PhD thesis would not have been possible, and even more who made the time I spent generally enjoyable. Firstly I would like to thank my parents, for supporting me to move abroad, and especially my father, for stimulating both my skepticism and interest in nature.

I would like to thank my supervisors Jarle and Kjetil for offering me the opportunity to come study in Trondheim. I will never forget the day of my job interview: first the long, mostly silent car ride with Jarle from the airport, confirming my prejudices about Norwegians, and the spontaneous tour off the city by Kjetil afterwards, making me realise that perhaps it would not be all that bad. They have given me lots of freedom, resulting in a PhD that was intended to be purely theoretical to consist half of empirical studies. And while that freedom sometimes made me lose my sense of direction, it eventually resulted in a thesis of which I am proud.

My time in Trondheim would not have been as great without the lively group of PhD students at the biology department. The conversation topics during lunch and coffee breaks have been extremely diverse, from hilariously bizarre anecdotes to wise life lessons passed through the PhD-generations. I especially wish to thank Yngvild and Eirin, my first office mates, who did their best to lure a somewhat shy Dutch girl out of her shell. Grethe, for succeeding in that, and for dragging me along to pubs and mountain trips. Our shared sense of humour has been entertaining for us as well as for others. My current office mates, Lise and Tonje, who made me look forward to come to the office every day. Your daily bugging of me with R related questions was a price I was willing to pay for the good laughs we had, and your willingness to tolerate my quirks. And of course the rest, Camilla, Irja, Line, Aline, Ivar, Kari, Sebastian, Geir, Ingerid, Anna and everyone else. Not to forget Ida, for reintroducing me to knitting, and many enjoyable "strikk drikk" evenings - the pile of socks and mittens is yours to blame. And Marius, for being a good friend through all those years, and for showing me an inside peak into Norwegian culture.

As I expected, doing this PhD was challenging in different ways than anticipated, and I have learned a lot. Whatever I will do next, I will look back on this period with gladness.

Jisca Huisman
Trondheim, 20 April 2012

Contents

List of papers

I de Eyto E, McGinnity P, **Huisman J**, Coughlan J, Consuegra S, Farrell K, OToole C, Tufto J, Megens HJ, Jordan W, Cross T and Stet RJM (2011). Varying disease-mediated selection at different life-history stages of Atlantic salmon in fresh water. Evolutionary Applications 4: 749-762.

II Sundt-Hansen L, **Huisman J**, Skoglund H and Hindar K. Farmed salmon fry may reduce early survival of wild salmon. (Manuscript)

III **Huisman J** and Tufto J. Comparison of non-Gaussian quantitative genetic models for migration and stabilizing selection. (Submitted manuscript)

IV **Huisman J** and Tufto J. Modelling Wild-Domestic Interbreeding: How selection on a quantitative trait affects gene flow at a neutral locus. (Submitted manuscript)

Declaration of contributions

I The project was initiated by PMG, WJ, RS and TC. The field work was carried out by EE, PG and KF, and genotyping by JC, SC, COT, WJ , HJM and TC . JH conducted statistical analysis, overseen by JT, who was involved in the initial design of the analysis. The manuscript was written by EE and PMG with contributions from JC, SC, WJ, TC and RS, and from JH for the statistical parts.

II The experiment was planned by LSH and KH, and was executed by LSH, HS, JH and KH. JH performed the genotyping, statistical analysis and wrote the methods and results sections, LSH wrote the introduction and discussion, both with comments from the other co-authors.

III JT initiated the project. JH developed the models with help of JT, and JH wrote the manuscript with comments from JT.

IV JT and JH initiated the project, JH extended with help of JT a model earlier used by JT . JH wrote the manuscript with comments from JT.

Introduction

Natural processes can be studied from many different perspectives, utilizing many different methods. The ecological, short-time perspective and the evolutionary, long-term perspective have often been studied separately, but more and more the conjunction between those two is studied. It is recognized that evolutionary processes may play an important role on ecological time scales (contemporary evolution), and that ecological processes will affect the evolutionary future of populations (Stockwell et al., 2003).

Selection and migration between populations are two examples of such processes, which are important both ecologically and evolutionary. The strength and shape of natural selection is determined by ecological factors, both physicchemical (temperature, habitat quality) and biological (competition, predation, pathogens). In the long run, only the genes and gene combinations resulting in individuals most adapted to these conditions will survive. Migration from an ecological perspective implies increased competition for the receiving population, and possibly increased or changed disease pressure when the migrants carry pathogens. From an evolutionary viewpoint, it means the influx of genes and new variation on which natural selection may act.

Studying eco-evolutionary processes When studying wild vertebrates, we can generally only observe ecological processes. For most species the time scales on which evolutionary processes take place are longer than the duration of any studies, at most we see only a small part of a much larger process. Different methods than direct observation need to be utilized to make inferences about processes on (eco-)evolutionary time scales; one may use analogies with other, more short-lived organisms, or use computer simulations.

Indirect approaches will never be able to give us the complete picture, but may still give us interesting insights. Either for specific cases, when using detailed information obtained empirically as input, or about general patterns and underlying processes. Nonetheless, natural processes are often complex and may interact with each other in ways we do not fully understand. Any discrepancies between model results and observations from nature may therefore be due to an enormous range of factors, mostly at the nature side of the equation. For that reason it is useful to (first) compare different types of models, and assess how and under what conditions their results differ.

Natural selection and maintenance of variation

Darwin defined natural selection as the "preservation of favourable variations and the rejection of injurious variations" (Darwin, 1859, ch. 4). Many other definitions have been introduced since then, but they all share the common principle that "individuals having any advantage, however slight, over others, have the best chance of surviving and procreating their kind" (Darwin, 1859). In other words, the individuals with the highest fitness contribute most to the gene pool of the next generation, and so the average relative fitness is expected to increase over generations. The traits which result in the highest fitness will differ between environments, so that the increase in fitness results not only in a better adap-

5

tation to local conditions, but also creates differences between populations in different environments (Hereford, 2009; McKay and Latta, 2002).

Generally, fitness is highest for individuals with an intermediate phenotype, rather than a minimum or maximum phenotype (Frankham et al., 2002). Stabilizing selection around an optimum explains both prolonged periods of stasis, in which the mean phenotype does not change, and periods with (strong) directional selection and rapid adaptation, when the mean is some distance away from the optimum (Estes and Arnold, 2007). The biological reason for stabilizing selection may be trade-offs between different traits affecting fitness. Each trait may be under directional selection, but an individual has only a finite amount of resources to allocate to these traits (Cole, 1954).

How well a population is able to respond to selection pressure depends not only on the strength of selection, but also on the amount of genetic variation in the population. A constant strength of selection will tend to deplete variation (Bulmer, 1971; Barton and Turelli, 1989), but when selection is weak, new mutations will replenish variation at an equal rate (Lande, 1976a; Burger et al., 1989; Burger and Lande, 1994). Under less restrictive assumptions, other processes must be considered to explain the maintenance of variation in natural populations. Several hypotheses have been brought forward over the years. One of the earliest was the notion that the environment changes over time, and that the population is evolving towards a constantly shifting optimum (Fisher, 1958; Lande, 2007). Other theories for the maintenance of genetic variation include fluc-

tuating selection as a result of fluctuating population density, frequency dependent selection (Endler and Greenwood, 1988; Mani et al., 1990), heterozygote advantage, population subdivision (Felsenstein, 1976; Slatkin, 1978), pleiotrophy (Keightley and Hill, 1990) and selection on trait combinations (Hunt et al., 2007). All of these have empirical support from some studies, but are refuted by others. It seems possible that each of those factors, and probably other factors as well, are involved in maintaining variation, but that their relative importance differs between systems.

Several of these theories have been tested experimentally with respect to the major histocompatibility complex (MHC), which is both known to be under strong selection as well as the most polymorphic gene in vertebrates. As such, it provides an excellent study model, as we may pick up signals that are undetectable in genes with less variation and/or under weaker selection. The proteins it encodes play a crucial role in the vertebrate immune response (Klein, 1986), and specific alleles are associated with vulnerability or resistance to specific pathogens (Hedrick, 2002). Under natural conditions, animals will be exposed to a varying assemblage of pathogens, whose effects may interact with other variables such as physiological condition, resource availability and abiotic conditions (McGinnity et al., 2009), which in turn may vary over time and space. This, combined with increasing knowledge on the relationship between MHC alleles and resistance to specific diseases, makes it a very good system to test and compare the different hypothesis on the maintenance of genetic variation.

Migration and gene flow

Migration may indicate a wide range of movement patterns, and can be defined on both the individual and the population level (Fox et al., 2001). From an individuals perspective, it can be seasonal, such as in Africa's large grazers, or twice in a lifetime, such as in many salmonids where there is a separation between breeding/early juvenile habitat and feeding/growing habitat. These types of migration are characteristics of the life history of a species. From an evolutionary (population) perspective, they are very different from migration between subpopulations, where an individual leaves its ancestral group and joins another. The tendency for this might be seen as a life history characteristic, but additionally it has evolutionary implications. In this thesis discrete populations are assumed, but in nature the boundaries between populations may be a matter of grey tones and scaling.

When migrants reproduce successfully, this result in the exchange of genetic material between (sub)populations, and an increase of the variation within populations. This may include slight advantages, that allow the population to adapt even better to local conditions, but may also include deleterious variations (Storfer, 1999; Garant et al., 2007). Natural selection will tend to remove any deleterious variations, and prevent or slow down their spread through the population. When migration introduces deleterious variations (much) faster into the population than natural selection can remove them, the reduction in average fitness in the population can potentially be considerable. Under ongoing migration, the number of individuals with this deleterious variation will increase until a balance between migration and selection is reached. A proportion of the population carrying a deleterious variation has negative consequences. As its individuals will on average have a worse chance of surviving and procreating, population viability will decrease (migration load).

Interbreeding with domestic conspecifics A special case of gene flow occurs when domesticated individuals interbreed with wild populations. The occurrence of this is wide spread, as many plant and animal species have been domesticated by humans, some in the ancient past, others just a few decades ago. For most of these species, this has resulted in much higher numbers of individuals under human control then in wild populations. Even when just a very small proportion of the domestic individuals becomes feral, their numbers may be high relative to the size of the wild population. This is the case for example in the European wildcat *Felis silvestris silvestris* with house cats *Felis catus* (Beaumont et al., 2001) and the Eurasian wolf *Canis lupus lupus* with dogs *Canis lupus familiaris* (Randi et al., 2000).

A range of studies, both theoretical and empirical, have shown negative effects of hybridization and introgression (Hall and Ayres, 2009), ranging from the spread of specific (favourable) alleles of cultivated origin (Ellstrand et al., 1999; Fitzpatrick et al., 2010) to genetic assimilation and demographic swamping (Haygood et al., 2003; Wolf et al., 2001) and even extinction (Rhymer and Simberloff, 1996). Many studies have been done in the context of the "escape" of transgenes into the wild, showing that under certain conditions this is likely because transgenes may increase fitness (most commonly via herbicide resistance) (Gepts et al., 2003).

A related scenario occurs in the case of supportive breeding programs (Ryman and Laikre, 1991), when individuals that are bred in captivity are intentionally released into the wild. While they are not as different from their wild conspecifics as the examples mentioned above, genetic adaptation to captivity creates differences due to different selection pressures in the captive and wild environments (Frankham, 2008; Snyder et al., 1996). This may have negative consequences for the wild populations which they are intended to supplement. In Pacific salmon (*Oncorhynchus* spp.) for example, it was shown that fitness can be substantially reduced by artificial propagation, suggesting that this commonly used tool for restoring populations may in fact be detrimental (Reisenbichler and Rubin, 1999; Waples and Drake, 2004).

Feral domestic or captive bred individuals may affect the evolutionary course of populations not only via interbreeding, but also by changing selection pressures. Intrusion of domestic individuals may change the selective landscape by increasing population density if their numbers are large, or by changing competition by their different behaviour. But they may also decrease population density, by introducing pathogens or, in the longer run, if interbreeding decreases average fitness in the population. The combination of these processes may lead to a shift of one stable state to another, comparable to ecosystem shifts like the shift of a clear lake to a murky one. Those shifts tend to be sudden, and hard to reverse (Scheffer et al., 2001).

Genetical Models

As mentioned before, the genetic consequences of introgression, like most genetic processes, are difficult to study in nature. This is due to the time spans involved, but also because we cannot observe genes directly; we have to deduce their effects and frequencies indirectly. To approximate the sometimes complex genetical processes, different mathematical approaches have been developed. Each of these approaches has different strengths and limitations, as a result of trade-offs between simplicity and tractability on the one hand, and generality or biological realism on the other hand.

Genetic models can broadly be grouped based on the number of loci they consider: one or two, a large but finite number, or an infinite number. Models with one or two loci were the first to be extensively studied (e.g. Hardy, 1908; Wright, 1931; Haldane, 1957; Kimura and Crow, 1969). They are inspired by the famous experiments by Mendel in the 19th century, before the field of genetics existed. One- and two-locus models are still useful tools today due to their simplicity, making analytical tractability possible in complex situations, such as under fluctuating selection (Lande, 2007, 2008) or under both migration and strong selection (Zhivotovsky and Christiansen, 1995; Bolnick et al., 2007).

Most traits under selection, however, are quantitative traits, such as body size or offspring number, affected not by one or two loci but by a number ranging from 'a couple' to 'a very large number'. In contrast to one- or two-locus systems, allele frequencies and the effect of each allele can in this case not be estimated from data. Analysis of this type of systems is impossible without some rigorous simplifications and approximations, even with the aid of computers. A widely used as-
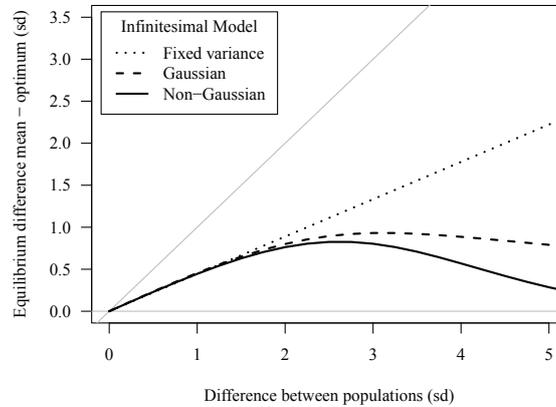
Figure 1: Example of difference in predicted divergence of the population mean from the optimum at migration-selection balance from three approximations to the infinitesimal model: Assuming constant variance (dotted line), assuming a Gaussian distribution (dashed) and assuming neither (solid), for a migration rate of 20% and moderate strength of selection. Difference is expressed in genetic standard deviations (sd) in the original population, grey diagonal shows deviation from optimum in absence of selection.

sumption is that loci have additive effects when the trait is measured on (or transformed to) an appropriate scale (Wright, 1968; Falconer and Mackay, 1996). This ignores that interactions between alleles at the same or different loci (dominance resp. epistasis), or even gene transcription, may play an important role in certain situations (Farrall, 2004; Whitlock et al., 1995). However, our understanding of these non-additive processes is limited (Barton and Turelli, 1989), and what is known may only be true in a few model species under laboratory conditions. While additive models thus have their limitations, they have shown their usefulness in both animal breeding (Falconer and Mackay, 1996) and natural populations (Lande, 1976b).

When one in addition to additivity assumes that the number of loci underlying the trait is very large (i.e. the infinitesimal model, Bulmer (1971)), it follows from the central limit theorem that the distribution of genotypes for this trait is approximately normally distributed (Fisher, 1918; Falconer and Mackay, 1996). Under most conditions it then suffices to keep track of the mean and variance of this distribution, instead of the frequency and properties of all underlying variables, greatly simplifying analysis. When forces such as selection and migration are weak or absent, one can additionally assume that the variance is approximately constant. This is supported by empirical findings that in artificial selection experiments, the mean phenotype often undergoes substantial changes in response to selection while the additive genetic variance remains relatively stable (Falconer and Mackay, 1996; Lande, 1976b). Being able to describe the entire genetic distribution by a single parameter (the mean) allows for analytical analysis of rather complex situations, comparable to the one- and two-locus models.

The variance of the distribution will not remain constant when selection is rel-

atively strong, be it artificial (truncation) or natural (stabilizing) selection. In that case selection will create negative covariance among loci (linkage disequilibrium, LD), which reduces the total additive genetic variance (Bulmer, 1971). Similarly, migration between populations creates positive linkage disequilibrium, increasing the variance. Since the change in mean as a result of selection (response R) is a function of the (additive) genetic variance (V_A) according to the breeders equation $R = h^2 S$, with $h^2 = V_A/V_P$ (where V_P is the phenotypic variance), ignoring this change in variance will affect predicted changes in the mean (figure 1, "Fixed variance" versus "Gaussian").

In certain cases one cannot make the assumption of a genotypic distribution which remains approximately normal. One of those cases is when migration rates are relatively high, and the fitness of immigrants is much lower than that of local individuals. This may be either because of strong selection, or because they are very different (or both). This is for example the case when these immigrants are feral domestic individuals. The more different immigrants are, the more the genotype distribution of the admixed population will get skewed or even bimodal, and the more the assumption of a Gaussian distribution will affect the results (figure 1, "Gaussian" versus "Non-Gaussian").

Atlantic salmon

Atlantic salmon provide a good system to study competition with and introgression of domestic conspecifics for several reasons. Firstly, they occur in large numbers, which translates in high sample sizes and comparatively solid results of field studies. In contrast to for example the wild cat and wolf, experimental studies under semi-natural conditions are possible. For such experiments during the juvenile stage only a limited number of parental individuals is needed, due to the enormous number of offspring fish have relative to mammals (around 10.000 eggs for Atlantic salmon, Hendry and Stearns (2004)). These parentals may be wild caught, and since artificial spawning has been mastered for 150 years (pers. comm. K. Hindar), full control over mating schemes is possible. The high fecundity rate also implies considerable mortality during juvenile stages, part of which will be selective. Because of these reasons, and because of the wide range of life histories between and within species, salmonids have been widely used for studies of life history evolution. As a result, their ecological niche and the evolutionary forces acting on them are relatively well understood (Hendry and Stearns, 2004).

As all study species, the Atlantic salmon has some disadvantages too. The long generation time (2-10 years, Hendry and Stearns (2004)), makes experiments spanning multiple generations very time (and money) consuming. In addition, and not unimportantly when studying wild-domestic interactions, they have been domesticated only about forty years ago. This is very recent compared to most species, many of which have been domesticated thousands of years ago, so that the genetic difference is comparatively small. Norwegian farmed salmon were founded from wild Norwegian populations in the 1970s, and has since undergone about ten generations of artificial selection for growth, meat quality and disease resistance (Gjøen and Bentsen, 1997; Gross, 1998), as well as likely genetic adaptation to captivity (Araki

et al., 2007). This has resulted in a significant difference in traits relevant to survival and breeding success in the wild. For example, seventh generation farmed salmon had more robust bodies, were more risk-prone and showed higher levels of aggression compared to wild salmon (Einum and Fleming, 1997).

Competition experiments Various studies have been performed on the relative performance of and interaction between wild and farmed salmons, and their first and second generation hybrids. These have been conducted throughout the natural range of the Atlantic salmon, in Canada, Ireland and Norway, largely under natural or semi-natural conditions, and are likely to span a wide range of the possible outcomes. Some span (nearly) the whole life cycle, while others have focussed on particular life stages to pinpoint when and how selection and competition may act.

In several studies the reproductive success (number of surviving offspring) of wild salmon has been lower in presence of farmed salmon compared to the expectation in absence of those (Fleming et al., 2000). This implies that intrusion of farmed salmon may not only have an effect on population size on the long run, by decreasing the average fitness in the population, but also more directly, by decreasing the number of surviving wild offspring. Both studies in which this effect was found spanned a large part of the life cycle, and it was not known exactly when this increased mortality occurred (Fleming et al., 2000; McGinnity et al., 2003). It was expected to occur at least partly shortly after hatching when territories are established, as mortality is naturally large in this phase (up to more than 90%, Einum and Nislow (2005)). It

has been speculated that the more aggressive behaviour of farmed juveniles is one of the reasons for this phenomenon.

The outcome of competition between farmed and wild salmon may depend on several factors, including the origin of the wild salmon and the ratio between the two. In a Canadian study using a fourth generation local farm strain, it was found that mortality of wild juveniles increased with an increasing proportion of wild-farmed hybrid juveniles in one population, but was largest at an intermediate proportion (3:7, compared to 5:5 and 15:85) for a second population (Houde et al., 2010). Their results suggested also that for low intrusion rates (15% hybrids), effects on survival and growth of wild juveniles is limited (Houde et al., 2010). The limited effect at low rates may not be valid for populations in Ireland and Norway, where farmed escapes are morphologically and behaviourally more different, as the farm strains used here have undergone more generations of artificial selection.

The enclosure and set-up of the experiment is another factor which may have a considerable effect on the outcome. It is imaginable that farmed fish have adapted to life in tanks, and may outperform wild fish under this type of conditions but not under more natural conditions. With a gravel bed for cover and a diet of insects, different traits are likely to be beneficial than when fed feed pellets in a tank without hiding places. This may explain for example why Houde et al. (2010) found that pairwise tests of competitive ability in tanks could not predict the magnitude of change in fitness-related traits of wild fish exposed to different proportions of hybrids in semi-natural stream environments.

Aims

I Does the direction of selection change between life stages? It has been shown that specific MHC alleles are related with increased or decreased chances of survival under natural conditions between egg and early juvenile stage. The aim of this study was to assess whether the same alleles are associated with survival probability, and in the same direction, during the remainder of the fresh water stage.

II Do farmed juveniles affect the survival and growth of wild juveniles during the earliest life stage? Studies have shown that presence of farmed spawners reduces the number of wild offspring at the end of the freshwater stage. The aim of this study was to find out if these competition effects occur during the earliest weeks of life, when mortality is naturally high, and if those effects are density dependent.

III Is the infinitesimal model appropriate to model migration-selection balance, or are more complex multilocus models needed? The non-Gaussian infinitesimal model is a parsimonious model for quantitative traits, but has some underlying assumptions which are violated under high migration rates and/or strong selection. Whether this affects results compared to more explicit multilocus models, has not been studied previously.

IV How does the relationship between migration rate and gene flow depend on the strength of selection? It is often assumed that selection is negligible and so that the rate of gene flow, inferred from neutral markers, can be translated to a fraction of immigrants. However, low relative fitness of immigrants and their descendants changes the relationship between migration and gene flow in a manner that is not well studied for quantitative traits.

Methods

Study species (Papers I-II)

In this thesis, the Atlantic salmon *Salmo salar* is used as study species. The life history of salmonids is comparatively well understood, as the large variation in life history traits both within and between species has received much research interest (Hendry and Stearns, 2004). Adults commonly spawn in the streams they were born, resulting in discrete, potentially locally adapted populations (Garcia de Leaniz et al., 2007). After hatching juveniles spent up to several years in this stream before the majority migrates to sea for one to several years, where the majority of growth occurs.

During the stream dwelling juvenile stage, mortality is naturally high (Einum and Nislow, 2005), making strong selection pressures likely. They often inhabit small manageable streams, where environmental conditions may vary spatially and temporally. Large scale, ecologically realistic experiments are possible at the experimental facilities at Ims (Norway) and Burrishoole (Ireland).

Results of this type of experiments, together with knowledge from aquaculture on for example heritabilities, are important for seeding eco-evolutionary models, such as the ones developed in papers III and IV. These theoretical papers are inspired by Atlantic salmon, but are of such a general nature that they can be applied to many different species.

Disease-mediated selection (Paper I)

Experimental design The experiment was conducted in the Burrishoole catchment in western Ireland. Wild broodstock from a neighbouring river system was used to increase the change to detect disease-mediated selection, as those are probably not adapted to some of the pathogens at the release site. To ensure a large degree of immunogenetic diversity, a full reciprocal mating system was used, resulting in 63 families and in total 56 000 eggs. Half a year after hatching 746 juveniles sampled, and a further 430 were sampled while migrating towards sea after about 2 years. Those juveniles, as well as their parents, were genotyped for 3 microsatellites embedded in the MHI and MHII alpha loci and 8 putatively neutral microsatellites.

Statistical analysis Allele frequencies in the eggs were compared to the sample frequencies at 6 months and 2 years of age, and changes in frequency were interpreted as differential survival between genotypes. As genetic drift and non-disease related mortality should be detectable at both control and immunogenetic loci, larger frequency change at one or several MH loci than at the control loci must be caused by disease-mediated selection. To distinguish between heterozygote advantage and advantage of specific alleles, alternative models were run for each locus, and the best model selected based on AIC values. These generalized mixed models contained family as random effect to correct for other shared genes among siblings. Where the model with an advantage of specific alleles proved the best fit for a locus, standard errors for the selection coefficients of the alleles were obtained using bootstrap.
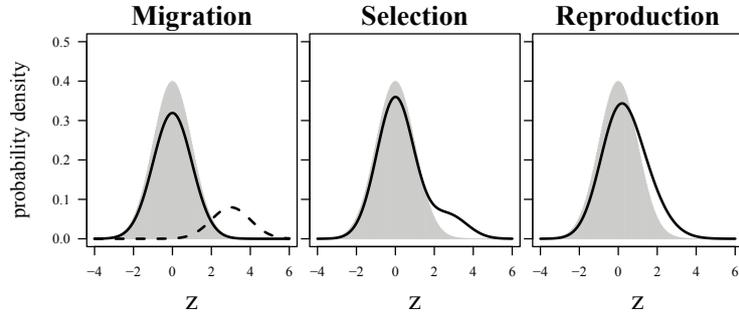
Figure 2: Distribution of the breeding values z in the recipient population (solid line) after 20% is replaced by immigrants from the donor population with $z_1 = 3$ (dashed line) just after migration (left panel), in the selected parents (middle, $s = 0.1$) and in the offspring after random mating (right). The grey shaded area indicates the distribution in the original population, with the mean at the optimum $z = 0$. The post-reproduction distribution is the pre-migration distribution in the next time step.

Wild − farmed juvenile competition (Paper II)

Experimental design Eggs of wild Atlantic salmon from the river Ims and farmed salmon from the Aqua Gen strain were kept in the same hatchery until the yolk-sack fry stage. About three weeks before they would naturally emerge from the gravel and start feeding, they were released in two types of semi-natural environments, one confined and mimicking small streams, the other mimicking a river bend and with opportunity to migrate out (and enter a fish trap). In both enclosure types two treatments were used, either only wild juveniles or an equal ratio of wild and farmed, and both treatments were conducted at high and low density ($n = 48$ resp. $n = 12$ for the small enclosures ($1.25m^2$), $n = 50$ resp. $n = 600$ for the large ones ($21.6 \ m^2$). The experiments were terminated after 58 and 66 days for the small and large enclosures respectively, when remaining fish were caught and weighed. Fish were classified to their origin either based on dye marking of otoliths, or based on the genotype at 11 microsatellite markers.

Statistical analysis The survival rate and body mass at termination was compared between wild juveniles in sympatry with farmed juveniles, wild juveniles in allopatry, and farmed juveniles using a quasi-binomial generalized linear model and a linear mixed model, respectively, with density as an additional factor.

Models for a quantitative trait under selection (Papers III & IV)

Assumptions A non-specified trait is considered which affects fitness, be it via fecundity, survival or both. We model only the additive part of the genotype (breeding value) for this trait and assume dominance effects are negligible and the distribution of genotypes in the population is approximately normal. For simplicity, non-overlapping generations are

assumed with random mating. The relationship between genotype and fitness is presumed constant over time, with stabilizing selection acting around a single, fixed optimum. As fitness is necessarily positive, and gradually declining away from the optimum, we assume fitness is a Gaussian function of the breeding value. The population size is assumed to be constant and either large (Paper III) or infinite (Paper IV). When population size is finite, stochastic processes will cause genetic drift. Stochastic variation in output variables is dealt with by taking averages over hundred replicate runs.

Simulations Each generation, genetically distinct immigrants enter the population, at a rate which is constant over time. Those immigrants originate from a farm population which has undergone several generations of directional selection, so that the average genotype for the trait considered is some distance away from the optimum in the wild (see figure 2). Immigrants breed at random with local individuals. The relative contribution of each individual to the next generation (fitness) is a Gaussian function of its genotypic value (breeding value) and the strength of stabilizing selection.

The output consists of the mean, variance and skew of the breeding value distribution, as well as the total amount of linkage disequilibrium, after fifty generations (Paper III) or at migration-selection equilibrium (paper IV). In paper III, differences between models are assessed based on these parameters, both graphically and using t-tests. For paper IV additionally the frequency of a neutral marker allele is given as output, which at initiation is absent in the receiving population and fixed in the immigrants.

Infinitesimal model As described in the introduction, the infinitesimal model assumes an infinite number of loci underlying the trait, each with an infinitesimal effect. The genotype of an individual can then be described by a single number, its breeding value, and offspring gets the average breeding value of the parents, plus a random term to account for recombination and Mendelian segregation. When modelling a population of finite size, this process is implemented as such in the simulation. When population size is assumed to be infinite, not the properties of individuals are traced but the distribution as a whole, for which various methods have been developed (Turelli and Barton, 1994). Here we divided the range of possible breeding values into 256 bins (like a histogram), and tracked the relative frequency in each of those bins. Reproduction is implemented using Fourier transformations, giving this method its name.

Multilocus models (Paper III only) In a multilocus model, the contribution of each allele to the breeding value and the allele frequencies are modelled explicitly. Optionally, the contribution to the breeding value (allelic effect) differs between loci. Several different models are included in the comparison, with different numbers of alleles per locus and different number of loci. For evolutionary processes, the number of loci is not as important as the relative contributions of loci to the genetic variance (Barton and Turelli, 1989; Chevalet, 1994). This number is related to the number of "freely segregating factors" which can be estimated from crosses of unrelated populations (Lande, 1981). The relation between the actual and "variance effective" number of loci differs between replicate

15

simulations, due to the random sampling of allele frequencies and allelic effects. Therefore, replicates were run using the same effective rather than actual number of loci, and also comparisons between models with different numbers of alleles per locus were done based on this measure.

Effective migration rate (Paper IV only)   The rate of gene flow, or effective migration rate m_e, was calculated from the rate of change of a neutral marker allele. It was defined as that migration rate, that in absence of selection would result in the same rate of allele frequency change as observed. It was calculated as $m_e = (q_B - q_{A,t})/(q_B - q_{A,t-1})$, where q_B is the frequency in the population from which the immigrants originate, and $q_{A,t}$ the frequency in the admixed population at time t.
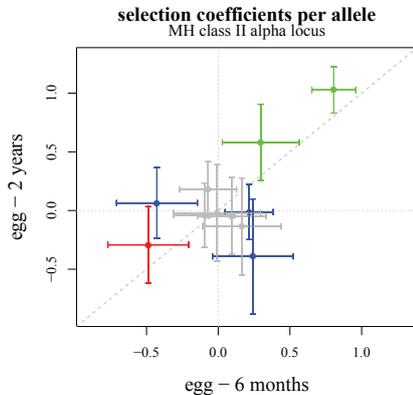

Figure 3: Estimated selection coefficients for each of the alleles on the MH IIα locus between egg and 6 months of age (horizontal) and egg and 2 years of age (vertical). Error bars indicate 95% confidence intervals. Green - alleles associated with increased survival during entire fresh water phase; Red - decreased survival; Blue - first increased, then decreased survival, or vice versa; grey - no significant effect on survival.

Key results and Discussion

Disease-mediated selection (Paper I)

For the MHIIα marker an advantage of specific alleles explained the data significantly better than a general heterozygote advantage, while a combination of the two could not be excluded. No clear indication of frequency-dependent selection was found, but any signal of this might be hard to disentangle from the other two processes.

Evidence was found that certain MHIIα alleles associated with increased survival during the first six months of life, were associated with decreased survival during the subsequent 1.5 year, or vice versa. This resulted in no net survival advantage over the whole period (two alleles), or even an opposite effect over the whole period compared to the first six months (one allele) (blue crosses in figure 3). This means that selective mortality in one time period may be a poor predictor of selection coefficients during other time periods, as found by Kekäläinen et al. (2009). Of the remaining MHIIα alleles, two were during the whole period associated with significantly increased survival, one with decreased survival, and five were not under selection (figure 3).

Allele frequencies at the MHI marker did not change significantly, nor did those at the majority of control microsatellites. Two of the control loci showed signs of selection, but selection coefficients were smaller than for the MHIIα markers. These may be false positives, or the loci were linked to genes under selection.

The results highlight the importance of variation at MH genes, as pathogen pres-

sure and the importance of certain alleles is likely to vary between years and rivers, in addition to the variation shown between life stages.

Wild – farmed juvenile competition (Paper II)

The presence of farmed juvenile Atlantic salmon negatively affected the survival of wild juveniles during the earliest life stage, but not their growth. This effect was independent of density. Increasing the density of wild juveniles in absence of farmed juveniles affected growth but not survival, suggesting the difference in densities was large enough to observe density dependent effects. Interestingly, it also suggests that farmed juveniles not simply increase competitive intensity disproportionally on a one-dimensional scale, but affect wild juveniles in a different manner.

As expected, farmed juveniles grew faster than wild ones. They also had a higher survival than wild juveniles in the small confined enclosures. Under natural conditions, in the presence of predators, they are expected to have a lower survival due to their reduced anti-predator behaviour (Einum and Fleming, 1997; Houde et al., 2010). In the large enclosures with opportunity for out migration, juveniles in the mixed groups entered the trap earlier than those in enclosures with only wild juveniles, which may indicate displacement or avoidance of competition.

Comparison of non-Gaussian models (Paper III)

The main difference between the infinitesimal model and the multilocus models was in the genetic variance. Under the infinitesimal model, the variance changed less relative to the situation at initiation, generally resulting in a larger estimated variance after fifty generations. This was expected, as the infinitesimal model makes some restrictive assumptions about the variance, most importantly a constant variance under linkage equilibrium and a constant, equal within family variance. Despite the difference in variance and skew between the models, the response to selection, measured as the deviation of the mean from the optimum, was still well approximated by the infinitesimal model, and differed never more than about 10% from the multilocus predictions.

The number of alleles per loci had limited effect on the results for most of the parameter space, even when the effective number of loci was limited. Results were more alike between a model with five loci with an infinite number of unique alleles and one with dozens of diallelic loci, than between these models and the infinitesimal model.

Concluding, it seems reasonable to use the more parsimonious (and computationally faster) non-Gaussian infinitesimal model to simulate migration and selection. When one considers the many simplifying assumptions underlying both types of models, the difference between the non-Gaussian infinitesimal model and explicit multilocus models seems of minor consequence. It is possible however that when some of those assumptions are relaxed, the difference between infinitesimal and multilocus models becomes more relevant.

Effective migration rate (Paper IV)

The ratio between gene flow (effective migration rate) and migration rate of individuals m_e/m decreases with increasing strength of selection and with increasing maladaptation of immigrants. Under these conditions, immigrants and their descendants will have lower fitness and contribute relatively less to the next generation. These relative contributions determine the change in allele frequency at the neutral locus, and thus effective migration rate. The effect of (new) immigrants on the effective migration rate can be expressed as a function of the migration rate m and the mean fitness of immigrants W_B relative to the mean fitness in the admixed population W_A according to

$$\widehat{m_e} = m \frac{W_B}{mW_B + (1-m)W_A}.$$

This ignores changes in neutral allele frequency due to indirect selection, via the association between allele frequency and genotype (and thus fitness) in the admixed population. In the limiting case of low migration rate, the ratio m_e/m is independent of m. This is in agreement with studies which investigated the gene flow of a marker allele linked to, or in LD with, one to several loci under selection (Bengtsson, 1985; Barton and Bengtsson, 1986). A new finding is that when the marker allele is in LD with a trait coded for by a large number of loci, the ratio m_e/m rises with increase in m.

Application of the relationships found to real world data, to estimate migration rates of individuals from genetic data or vice versa, is hampered by the simplifying assumptions necessarily made in this model, as well as by the imprecision in measuring/estimating the required input parameters. Nevertheless, the rise in ratio with migration rate highlights that introgression may rise disproportionally when migration rates increases. This includes parameter combinations which are believed to occur in nature, for example in the farmed-wild Atlantic salmon context.

General discussion and perspectives

The papers in this thesis may aid in finding ways to minimize the impact of domesticated individuals on wild conspecifics in general, and the effect of escaped farmed salmon on wild populations in particular. One question in this context is whether one should aim to make farmed salmons as similar or as different as possible from wild salmon. If they are very similar, any interbreeding will lead to minimal genetic changes to the wild population. However, it has been shown that even without artificial selection, a few generations of hatchery breeding may affect fitness performance in the wild (Araki et al., 2007). Massive releases of hatchery-bred (non-)local juveniles are commonplace in salmonids, aiming for example to reduce acute extinction risk in small populations, or simply to increase harvest opportunities (Waples and Drake, 2004). The effect of such releases on the long term viability of populations remains unclear (Waples et al., 2007), but it seems that in at least some cases it may have done more harm than good (Reisenbichler and Rubin, 1999). When considering intrusion with farmed salmon, genetic similarity is most likely not in the economic interest of aquaculture, where considerable genetic gain in commercially important traits has been

achieved in a few generations (Gjøen and Bentsen, 1997). If farmed salmon are very different, the effects of any interbreeding will be large, but the likelihood of this occurring will be very small, and there will be relatively strong selection acting against any descendants of such interbreeding. How different salmon need to be to ensure that the likelihood of interbreeding is very small indeed, depends on the strength of selection acting on relevant traits – which is likely to differ between rivers and years. Both paper III and Baskett and Waples (prep) suggest that the worst case scenario is when farmed salmon are intermediate, so that interbreeding is fairly likely, and each instance has an intermediate impact. Worryingly, for the range of selection intensity believed to be relevant in this scenario (Tufto, 2010), estimates from aquaculture (Gjøen and Bentsen, 1997) and the model used (Paper III) suggest that the current difference between Norwegian farmed and wild salmon is around this worst-case intermediate.

However, this quantitative genetic model is a very simplified version of a complex reality. Salmon populations are believed to be locally adapted (Garcia de Leaniz et al., 2007), and part of this adaptation may be due to a limited number of alleles with a large effect on fitness. Interbreeding with escaped farmed salmons may lead to loss of such alleles, with larger effects on mean fitness in the population than predicted by quantitative genetic models. An example of such alleles are the MH loci, which can have a considerable impact on fitness in a large range of vertebrates (Bernatchez and Landry, 2003), including Atlantic salmon (paper I).

Studies on MH loci have shown that it is possible, with our current methods and technology, to find the specific alleles important for fitness in a certain environment, or under a certain pathogen pressure. Disease resistance is however only one of the many aspects of fitness, and finding the relevant alleles for each aspect would be very time consuming. More importantly, as the environment fluctuates over time, so will the important alleles, and even the relative contribution of traits to overall fitness. Therefore, it is important to conserve the dynamic process of adaptation, by conserving genetic diversity, rather than conserving the adaptation to a given environment.

From this perspective, the threat of intrusion by farmed salmon, or domesticated individuals, is two fold. Firstly, intrusion may lower genetic variation in the population, as domesticated individuals typically have lower genetic diversity due to a limited number of founders (Frankham et al., 2002). Additionally, the decrease in local adaptation will coincide with a decrease in genetic variance in the admixed population, either directly (paper III) or indirectly via a reduction in population size. Decreased variation reduces the population's ability to respond to the second challenge, a changing selection landscape (Lande and Shannon, 1996). Domesticated immigrants will change selection pressures by, amongst others, increasing or changing competitive interactions (paper II) or introducing novel pathogens (McVicar, 1997). The two processes are likely to reinforce each other, similar to the processes in a migrational meltdown (Ronce and Kirkpatrick, 2001) or an extinction vortex (Frankham et al., 2002).

The consequences of introgression may be further enhanced, or at least complicated, by non-additive genetic effects.

The fitness of first generation hybrids is often higher than the mean of both parental groups (hybrid vigour), due to dominance effects and possibly epistasis (Falconer and Mackay, 1996; Lynch, 1997). This may speed up the rate of introgression relative to the expectation under a purely additive model. In later hybrid generations, breakdown of coadapted gene complexes may lead to lower than expected fitness (outbreeding depression) (Lynch, 1997; Edmands, 2007). This may slow down introgression, but may also lead to a sudden decrease in population viability several generations after the start of introgression, when later generation hybrids make up a measurable proportion of the population. The few studies performed in Atlantic salmon in this context suggest these processes are of such magnitude to be biologically relevant (McGinnity et al., 2003; Houde et al., 2010), and that the magnitude may differ between populations (Einum and Fleming, 1997; Houde et al., 2010).

A likely future direction in the empirical study of gene flow and local adaptations lies in the field of "conservation genomics" (Allendorf et al., 2010) such as the use of Single Nucleotide Polymorphisms (SNP). Individuals can be typed for tens of thousands of these genetic markers at a time, and the possibilities seem endless - from deducing pedigrees and "wildlife forensics" to finding the position on the chromosome of genes under selection. The methods to analyse this kind of data are still being developed, and Paper IV may provide one of the building blocks in this process. It shows how the (expected) allele frequency of a neutral marker allele, which most SNPs are believed to be, responds to migration and selection. Paper IV provides only an expectation in the absence of

genetic drift; methods on how to incorporate drift and how to combine multiple alleles in this context will need to be developed. While SNP sets may be a long way ahead for most wild species, a set has been developed recently for Atlantic Salmon (Karlsson et al., 2011). It makes it possible to distinguish escaped farmed salmon and hybrids from pure wild salmon, although this is complicated by differences in SNP signature between rivers, between farm strains and between year classes of these farm strains. Several SNPs have been found which differ consistently in frequency between all farm strains and wild populations, probably because they are closely linked to traits favoured under farm conditions (Karlsson et al., 2011). In species with a longer history of domestication, it seems likely that certain SNPs are fixed for alternate alleles in the domestic and wild populations, enabling identification of hybrids which may be difficult to distinguish morphologically.

Summarizing, this thesis increases our understanding of gene flow by showing how it depends on migration rate, the degree of maladaptation of immigrants and the strength of selection (paper IV), and that migration and selection can be adequately modelled using the non-Gaussian infinitesimal model, as opposed to explicit multilocus models (paper III). These models assume stabilizing selection, which may be the result of fluctuating selection, as was shown to occur between life stages on the MHIIα locus (paper I). Another assumption implicitly made in the modelling papers was that the presence farmed salmon does not affect the fitness of wild salmon, which was falsified by paper II, where the presence farmed juveniles decreased the survival of wild juveniles. Further work may

broaden the scope of the quantitative genetic models used by relaxing some of the assumptions made, allowing simulation of more ecologically realistic scenarios. This, together with future experimental studies, will advance the general knowledge on gene flow and selection, and may aid management decisions to minimize unwanted effects of domesticated individuals in the wild.

References

Allendorf, F., Hohenlohe, P., and Luikart, G. (2010). Genomics and the future of conservation genetics. *Nat Rev Genet*, 11(10):697–709.

Araki, H., Cooper, B., and Blouin, M. S. (2007). Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. *Science*, 318(5847):100–103.

Barton, N. H. and Bengtsson, B. (1986). The barrier to genetic exchange between hybridising populations. *J Hered*, 56:357–376.

Barton, N. H. and Turelli, M. (1989). Evolutionary quantitative genetics: How little do we know? *Annu Rev Genet*, 23(1):337–370.

Baskett, M. and Waples, R. (in prep.). Minimizing unintended fitness consequences of cultured individuals on wild populations: keep them similar or make them different? in prep.

Beaumont, M., Barratt, E. M., Gottelli, D., Kitchener, A. C., Daniels, M. J., Pritchard, J. K., and Bruford, M. W. (2001). Genetic diversity and introgression in the scottish wildcat. *Mol Ecol*, 10(2):319–336.

Bengtsson, B. (1985). The flow of genes through a genetic barrier. In Greenwood, J., Harvey, P., and Slatkin, M., editors, *Evolution Essays in honor of John Maynard Smith*, pages 31–42. Cambridge Univ. Press, Cambridge.

Bernatchez, L. and Landry, C. (2003). Mhc studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *J Evolution Biol*, 16(3):363–377.

Bolnick, D., Nosil, P., and Servedio, M. (2007). Natural selection in populations subject to a migration load. *Evolution*, 61(9):2229–2243.

Bulmer, M. G. (1971). The effect of selection on genetic variability. *Am Nat*, 105(943):201–211.

Burger, R. and Lande, R. (1994). On the distribution of the mean and variance of a quantitative trait under mutation-selection-drift balance. *Genetics*, 138(3):901–912.

Burger, R., Wagner, G., and Stettinger, F. (1989). How much heritable variation can be maintained in finite populations by mutation-selection balance? *Evolution*, 43(8):1748–1766.

Chevalet, C. (1994). An approximate theory of selection assuming of quantitative trait loci. *Genet Sel Evol*, 26:379–400.

Cole, L. (1954). The population consequences of life history phenomena. *Q Rev Biol*, 29(2):103–137.

Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life.* John Murray, UK.

Edmands, S. (2007). Between a rock and a hard place: evaluating the relative risks of inbreeding and outbreeding for conservation and management. *Mol Ecol*, 16:463–475.

Einum, S. and Fleming, I. (1997). Genetic divergence and interactions in the wild among native, farmed and hybrid atlantic salmon. *J Fish Biol*, 50(3):634–651.

Einum, S. and Nislow, K. H. (2005). Local-scale density-dependent survival of mobile organisms in continuous habitats: an experimental test using Atlantic salmon. *Oecologia*, 143(2):203–210.

Ellstrand, N., Prentice, H., and Hancock, J. (1999). Gene flow and introgression from domesticated plants into their wild relatives. *Annu Rev Ecol Syst*, 30:539–563.

Endler, J. and Greenwood, J. (1988). Frequency-dependent predation, crypsis and aposematic coloration. *Proc R Soc B*, 319(1196):505–523.

Estes, S. and Arnold, S. (2007). Resolving the paradox of stasis: Models with stabilizing selection explain evolutionary divergence on all timescales. *Am Nat*, 169(2):227–244.

Falconer, D. and Mackay, T. F. (1996). *Introduction to quantitative genetics*. Longmans Green, Harlow, Essex, UK, 4 edition.

Farrall, M. (2004). Quantitative genetic variation: a post-modern view. *Hum mol genet*, 13(suppl 1):R1.

Felsenstein, J. (1976). The theoretical population genetics of variable selection and migration. *Annu Rev Genet*, 10(1):253–280.

Fisher, R. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *T Roy Soc Edin*, 52:399–433.

Fisher, R. (1958). *The genetic theory of natural selection*. Dover, New York, 2 edition.

Fitzpatrick, B., Johnson, J., Kump, D., Smith, J., Voss, S., and Shaffer, H.

(2010). Rapid spread of invasive genes into a threatened native species. *P Natl Acac Sci USA*, 107(8):3606–3610.

Fleming, I. A., Hindar, K., Mjølnerød, I. B., Jonsson, B., Balstad, T., and Lamberg, A. (2000). Lifetime success and interactions of farm salmon invading a native population. *Proc R Soc B*, 267(1452):1517–1523.

Fox, C., Roff, D., and Fairbairn, D., editors (2001). *Evolutionary ecology: concepts and case studies*. Oxford University Press New York.

Frankham, R. (2008). Genetic adaptation to captivity in species conservation programs. *Mol Ecol*, 17(1):325–333.

Frankham, R., Briscoe, D., and Ballou, J. (2002). *Introduction to conservation genetics*. Cambridge Univ Press.

Garant, D., FORDE, S., and HENDRY, A. (2007). The multifarious effects of dispersal and gene flow on contemporary adaptation. *Funct Ecol*, 21(3):434–443.

Garcia de Leaniz, C., Fleming, I., Einum, S., Verspoor, E., Jordan, W., Consuegra, S., Aubin-Horth, N., Lajus, D., Letcher, B., Youngson, A., et al. (2007). A critical review of adaptive genetic variation in atlantic salmon: implications for conservation. *Biol Rev*, 82(2):173–211.

Gepts, P., Papa, R., et al. (2003). Possible effects of (trans) gene flow from crops on the genetic diversity from landraces and wild relatives. *Environ Biosafety Res*, 2(2):89–103.

Gjøen, H. and Bentsen, H. (1997). Past, present, and future of genetic improvement in salmon aquaculture. *ICES J Mar Sci*, 54(6):1009–1014.

Gross, M. (1998). One species with two biologies: Atlantic salmon (*Salmo salar*) in the wild and in aquaculture. *Can J Fish Aquat Sci*, 55:131–144.

Haldane, J. (1957). The cost of natural selection. *J Genet*, 55(3):511–524.

Hall, R. and Ayres, D. (2009). What can mathematical modeling tell us about hybrid invasions? *Biol Invasions*, 11(5):1217–1224.

Hardy, G. (1908). Mendelian proportions in a mixed population. *Science*, 28:49–50.

Haygood, R., Ives, A. R., and Andow, D. A. (2003). Consequences of recurrent gene flow from crops to wild relatives. *Proc R Soc B*, 270(1527):1879–1886.

Hedrick, P. (2002). Pathogen resistance and genetic variation at mhc loci. *Evolution*, 56(10):1902–1908.

Hendry, A. and Stearns, S., editors (2004). *Evolution illuminated: salmon and their relatives.* Oxford University Press, USA.

Hereford, J. (2009). A quantitative survey of local adaptation and fitness trade-offs. *Am Nat*, 173(5):579–588.

Houde, A. L. S., Fraser, D. J., and Hutchings, J. A. (2010). Fitness-related consequences of competitive interactions between farmed and wild Atlantic salmon at different proportional representations of wild–farmed hybrids. *ICES J Mar Sci*, 67(4):657–667.

Hunt, J., Blows, M., Zajitschek, F., Jennions, M., and Brooks, R. (2007). Reconciling strong stabilizing selection with the maintenance of genetic variation in a natural population of black field crickets (teleogryllus commodus). *Genetics*, 177(2):875.

Karlsson, S., Moen, T., Lien, S., Glover, K., and Hindar, K. (2011). Generic genetic differences between farmed and wild Atlantic salmon identified from a 7K SNP chip. *Mol Ecol Resour*, 11:247–253.

Keightley, P. and Hill, W. (1990). Variation maintained in quantitative traits with mutation-selection balance: pleiotropic side-effects on fitness traits. *Proc R Soc B*, pages 95–100.

Kekäläinen, J., Vallunen, J., Primmer, C., Rättyä, J., and Taskinen, J. (2009). Signals of major histocompatibility complex overdominance in a wild salmonid population. *Proc R Soc B*, 276(1670):3133.

Kimura, M. and Crow, J. (1969). Natural selection and gene substitution. *Genet Res*, 13:127–141.

Klein, J. W. J. (1986). *Natural history of the major histocompatibility complex.* Wiley, New York.

Lande, R. (1976a). The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genet Res*, 26(3):221–235.

Lande, R. (1976b). Natural selection and random genetic drift in phenotypic evolution. *Evolution*, 30(2):314–334.

Lande, R. (1981). The minimum number of genes contributing to quantitative variation between and within populations. *Genetics*, 99(3-4):541–553.

Lande, R. (2007). Expected relative fitness and the adaptive topography of fluctuating selection. *Evolution*, 61(8):1835–1846.

Lande, R. (2008). Adaptive topography of fluctuating selection in a mendelian population. *J Evolution Biol*, 21(4):1096–1105.

Lande, R. and Shannon, S. (1996). The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution*, 50(1):434–437.

Lynch, M. (1997). Inbreeding depression and outbreeding depression. In Grant, W., editor, *Genetic effects of straying of non-native hatchery fish into natural*

populations: proceedings of the workshop, pages 59–70, Seattle, Washington. National Oceanic and Atmospheric Administration. Tech Memo 30.

Mani, G., Clarke, B., and Shelton, P. (1990). A model of quantitative traits under frequency-dependent balancing selection. *Proc R Soc B*, 240(1297):15.

McGinnity, P., Jennings, E., DeEyto, E., Allott, N., Samuelsson, P., Rogan, G., Whelan, K., Cross, T., et al. (2009). Impact of naturally spawning captive-bred atlantic salmon on wild populations: depressed recruitment and increased risk of climate-mediated extinction. *Proc R Soc B*, 276(1673):3601.

McGinnity, P., Prodöhl, P., Ferguson, A., Hynes, R., Ó Maoiléidigh, N., Baker, N., Cotter, D., O'Hea, B., Cooke, D., Rogan, G., Taggart, J., and Cross, T. (2003). Fitness reduction and potential extinction of wild populations of Atlantic salmon, *Salmo salar*, as a result of interactions with escaped farm salmon. *Proc R Soc B*, 270(1532):2443–2450.

McKay, J. and Latta, R. (2002). Adaptive population divergence: markers, QTL and traits. *Trends Ecol Evol*, 17(6):285–291.

McVicar, A. (1997). Disease and parasite implications of the coexistence of wild and cultured atlantic salmon populations. *ICES J Mar Sci*, 54(6):1093–1103.

Randi, E., Lucchini, V., Christensen, M. F., Mucci, N., Funk, S. M., Dolf, G., and Loeschcke, V. (2000). Mitochondrial DNA variability in italian and east european wolves: detecting the consequences of small population size and hybridization. *Conserv Biol*, 14(2):464–473.

Reisenbichler, R. and Rubin, S. (1999). Genetic changes from artificial propagation of pacific salmon affect the productivity and viability of supplemented populations. *ICES J Mar Sci*, 56(4):459–466.

Rhymer, J. M. and Simberloff, D. (1996). Extinction by hybridization and introgression. *Annu Rev Ecol Syst*, 27(1):83–109.

Ronce, O. and Kirkpatrick, M. (2001). When sources become sinks: migrational meltdown in heterogeneous habitats. *Evolution*, 55(8):1520–1531.

Ryman, N. and Laikre, L. (1991). Effects of supportive breeding on the genetically effective population size. *Conserv Biol*, 5(3):325–329.

Scheffer, M., Carpenter, S., Foley, J., Folke, C., Walker, B., et al. (2001). Catastrophic shifts in ecosystems. *Nature*, 413(6856):591–596.

Slatkin, M. (1978). Spatial patterns in the distributions of polygenic characters. *J Theor Biol*, 70(2):213–228.

Snyder, N., Derrickson, S., Beissinger, S., Wiley, J., Smith, T., Toone, W., and Miller, B. (1996). Limitations of captive breeding in endangered species recovery. *Conserv Biol*, 10(2):338–348.

Stockwell, C. A., Hendry, A. P., and Kinnison, M. T. (2003). Contemporary evolution meets conservation biology. *Trends Ecol Evol*, 18(2):94 – 101.

Storfer, A. (1999). Gene flow and endangered species translocations: a topic revisited. *Biol Conserv*, 87(2):173–180.

Tufto, J. (2010). Gene flow from domesticated species to wild relatives: migration load in a model of multivariate selection. *Evolution*, 64(1):180–192.

Turelli, M. and Barton, N. H. (1994). Genetic and statistical analyses of strong selection on polygenic traits: What, me normal? *Genetics*, 138(3):913–941.

Waples, R. and Drake, J. (2004). Risk-benefit considerations for marine stock enhancement: a pacific salmon perspective. In Leber, K., Kitada, S., Blankenship, H., and Svåsand, T., editors, *Enhancement and sea ranching: developments, pittfalls and opportunities*, pages 260–306. Blackwell, Oxford.

Waples, R., Ford, M., and Schmitt, D. (2007). Empirical results of salmon supplementation in the northeast pacific: A preliminary assessment. In *Ecological and genetic implications of aquaculture activities*, pages 383–403. Kluwer Academic Publishers.

Whitlock, M., Phillips, P., Francisco, B., and Tonsor, S. (1995). Multiple fitness peaks and epistasis. *Annu Rev Ecol Syst*, 26:601–629.

Wolf, D. E., Takebayashi, N., and Rieseberg, L. H. (2001). Predicting the risk of extinction through hybridization. *Conserv Biol*, 15(4):1039–1053.

Wright, S. (1931). Evolution in mendelian populations. *Genetics*, 16(2):97–159.

Wright, S. (1968). *Genetic and Biometric Foundations*, volume 1 of *Evolution and the Genetics of Populations*. University of Chicago Press, USA.

Zhivotovsky, L. A. and Christiansen, F. B. (1995). The selection barrier between populations subject to stabilizing selection. *Evolution*, 49(3):pp. 490–501.

Paper I

ORIGINAL ARTICLE

Varying disease-mediated selection at different life-history stages of Atlantic salmon in fresh water

Elvira de Eyto,[1,]*Philip McGinnity,[2,]*Jisca Huisman,[3] Jamie Coughlan,[2] Sofia Consuegra,[4] Killian Farrell,[1] Ciar O'Toole,[2] Jarle Tufto,[5] Hendrik-Jan Megens,[6] William Jordan,[7,†] Tom Cross[2] and Rene' J. M. Stet[8,†]

1 Marine Institute, Newport, Co Mayo, Ireland
2 Department of Zoology, Ecology and Plant Science/Aquaculture and Fisheries Development Centre, University College Cork, Distillery Fields, North Mall, Cork, Ireland
3 Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway
4 Aberystwyth University, Institute of Biological, Environmental & Rural Sciences, Aberystwyth, UK
5 Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway
6 Department of Animal Sciences, Wageningen University, Wageningen, the Netherlands
7 Institute of Zoology, Zoological Society of London, Regent's Park, London, UK
8 Scottish Fish Immunology Research Centre, University of Aberdeen, Zoology Building, Tillydrone Avenue, Aberdeen, UK

Abstract

Laboratory studies on associations between disease resistance and susceptibility and major histocompatibility (MH) genes in Atlantic salmon *Salmo salar* have shown the importance of immunogenetics in understanding the capacity of populations to fight specific diseases. However, the occurrence and virulence of pathogens may vary spatially and temporally in the wild, making it more complicated to predict the overall effect that MH genes exert on fitness of natural populations and over several life-history stages. Here we show that MH variability is a significant determinant of salmon survival in fresh water, by comparing observed and expected genotype frequencies at MH and control microsatellite loci at parr and migrant stages in the wild. We found that additive allelic effects at immunogenetic loci were more likely to determine survival than dominance deviation, and that selection on certain MH alleles varied with life stage, possibly owing to varying pathogen prevalence and/or virulence over time. Our results highlight the importance of preserving genetic diversity (particularly at MH loci) in wild populations, so that they have the best chance of adapting to new and increased disease challenges as a result of projected climate warming and increasing aquaculture.

Introduction

Understanding the genetic basis of the immune response in fish is critical for the conservation of wild stocks that are under threat from many sources. Disease-mediated extinction of local populations is increasingly likely as a consequence of global warming, and the resulting increase in water temperatures that is probably to cause an increase in the diversity and prevalence and/or virulence of pathogens (Harvell et al. 2002; Crozier et al. 2008; Dionne et al. 2009). In addition, commercially important species such as salmonids face additional threats. The increase in salmonid aquaculture (both fish farming and stocking programmes) (ICES 2009) poses significant disease risks to wild populations; wild fish migrating past cages may be exposed to high levels of pathogens (e.g. sea lice) (Krkošek et al. 2007; Ford and Myers 2008; Costello 2009), accidental or deliberate release of farm fish may

bring novel pathogens with them into the wild (Johnsen and Jensen 1994; Bergersen and Anderson 1997) and introgression between farmed escapes and wild populations may lead to changes in the variability of immunogenetic loci of wild populations (Coughlan et al. 2006). While direct genetic effects of introgression between wild and hatchery-reared salmon have been demonstrated (McGinnity et al. 2003; Araki et al. 2007), the impact of diseases originating from aquaculture (Håstein and Lindstad 1991; Johnsen and Jensen 1994; McVicar 1997) on the genetic integrity of wild fish populations has not been sufficiently addressed. A better understanding of how disease-mediated selection impacts on wild populations at all life stages is therefore crucial.

The genes of the major histocompatibility (MH) complex (MHC) encode proteins that play a crucial role in the vertebrate immune response (Klein 1986), and possibly as a result of pathogen-driven balancing selection, MHC genes are the most polymorphic coding regions known in vertebrates (Grimholt et al. 2003; Wegner 2008). Pathogen-driven balancing selection may be the result of heterozygote advantage, negative frequency-dependent selection or varying pathogen resistance over space and time. (Hedrick et al. 1999, 2002; Fraser and Neff 2009; Kekäläinen et al. 2009). The high level of polymorphism in MH genes allows populations to mount an immune response to a wide range of pathogens, but this is only possible if populations have enough variability at MH loci and hence are 'adequately armed to face the challenge of changing environments' (Miller and Vincent 2008).

Atlantic salmon *Salmo salar* L. express single classical MH class I, class II alpha and class II beta loci (Grimholt et al. 2002; Stet et al. 2002). As class I and class II MHC genes in teleosts are unlinked and do not form a single complex, they are therefore known simply as MH genes in this taxon (Stet et al. 2002).Associations between MH genes and resistance or susceptibility to several major salmonid diseases, such as amoebic gill disease (Wynne et al. 2007), furunculosis (Langefors et al. 2001; Lohm et al. 2002), sea lice (Glover et al. 2007), bacterial kidney disease (Turner et al. 2007) and infectious salmon anaemia (Grimholt et al. 2003), have been found in farmed populations, and recently, it has been shown that MH genes are linked with increased susceptibility or resistance to myxozoa in the wild (Dionne et al. 2009). There is, therefore, strong evidence that MH variability can have important implications for the ability of salmon populations to fight disease, but much of this evidence comes from laboratory challenges on adults. However, associations of alleles in single challenge experiments to specific infections cannot explain how the extreme diversity of MH genes is maintained

(Wegner 2008), and indeed, it is highly unlikely that animals in their natural environment are only exposed to one pathogen at a time. Empirical evidence linking MH variability to survival and fitness in wild natural conditions with probable varying pathogen assemblages is rare, but is needed o ascertain and predict the impact of disease-mediated selection on locally adapted wild populations of salmon.

The Atlantic salmon (*Salmo salar* L.) is an anadromous fish that spends the first part of its life in fresh water (typically 1–4 years) and a further one to 3 years feeding in the ocean, before returning to natal rivers to spawn. In the first year after hatching, juvenile salmon (parr) are designated as 0+. After the spring of subsequent years parr are designated as 1+, 2+, etc. until they migrate to sea as smolts. We have previously shown that natural selection on MH genes has fitness consequences for salmon in the first 6 months of their life in fresh water (de Eyto et al. 2007). As salmon generally exhibit high mortality (>90%) in the first few months of freshwater life, this life stage would be most likely to experience disease-mediated selection. However, because of their anadromous life cycle, Atlantic salmon are exposed to pathogens in both the marine and freshwater environments, and as a consequence, salmon undergo a number of other potentially large mortality events both prior to smolting and in the sea, and it is likely that MH-determined survival may also be important at several other life stages. The lack of 'wild immunogenetic' studies has recently been highlighted (Pedersen and Babayan 2011), particularly the need to understand how immunogenetics interact with all the other variables that affects wild animals over time, such as physiological condition, resource availability and abiotic conditions (McGinnity et al. 2009). As salmon grow from parr to smolts, they experience a wide range of these variables, e.g. two subsequent winters of potential resource limitation and cold weather, possible summer drought as 1+ fish, changing physiology as they prepare to leave fresh water as smolts and varying pathogen virulence and prevalence. To fully ascertain the extent of disease-mediated selection, it is crucial to run an experiment for long enough to trigger the full potential of the adaptive immune response (Eizaguirre and Lenz 2010), which in the case of salmon in fresh water is at least 2 years. It may be that the immunogenetic advantage conferred during the salmon's first life stage (de Eyto et al. 2007) ceases to be important 18 months later, as other factors affecting survival increase. Conversely, it may be that the advantage continues to play an important role throughout the 2 years in fresh water. The aim of this study, therefore, is to assess the relationship between MH alleles and survival of salmon throughout the freshwater phase. We wanted to ascertain whether disease-mediated selection is a

determinant for survival over different life stages of salmon, and if so, whether the same alleles associated with survival in parr are also important at later stages of the salmon's development in fresh water.

Materials and methods

Hatchery and field methods

Details of the experiment location and initial set up of family crosses and hatchery procedure can be found in de Eyto et al. (2007). In summary, the experiment was carried out in a contained section of the Srahrevagh River, located in the Burrishoole catchment in western Ireland. As locally adapted fish may not show any signs of disease-mediated selection in their natural environment, we selected wild broodstock from the neighbouring Owenmore system in Co Mayo (Fig. 1). Owenmore fish are not native to the experimental river, and thus are probably not adapted to some of the pathogens endemic to the Burrishoole system. In addition, there has been no history of aquaculture in the Owenmore catchment or the immediate estuary, and thus, fish from this system should have minimal exposure to any aquaculture-associ-

ated diseases. To ensure a large degree of immunogenetic diversity, we used a full reciprocal mating system, which produced 63 families with an average of 889 (±155 SE) eggs per family. We excluded natural spawners from the experiment river over the winter of 2001–2002, ensuring that the only salmon hatching in the river in 2002 were from the experiment. This exclusion was facilitated by the presence of a complete upstream and downstream fish trap at the bottom of the Srahrevagh River. In early February 2002, we counted live eggs, mixed the families and 56 031 eggs were planted in the upper reaches of the river in artificial redds (Donaghy and Verspoor 2000; Fig. 1). We randomly sampled approximately 1500 0+ salmon parr from the river in the summer of 2002. A subsample of 746 of these parr was typed for this study. In addition, any juvenile salmon migrating through the trap at the bottom of the river were retained until after the smolt run in 2004 (the vast majority of salmon in this geographic region migrate to sea as 2+ smolts). Tail clips from each individual were preserved in 99% ethanol. Of the fish migrating through the trap, 110 presmolts (migrating in late 2003) and 320 smolts (migrating in February–April 2004) were randomly selected for
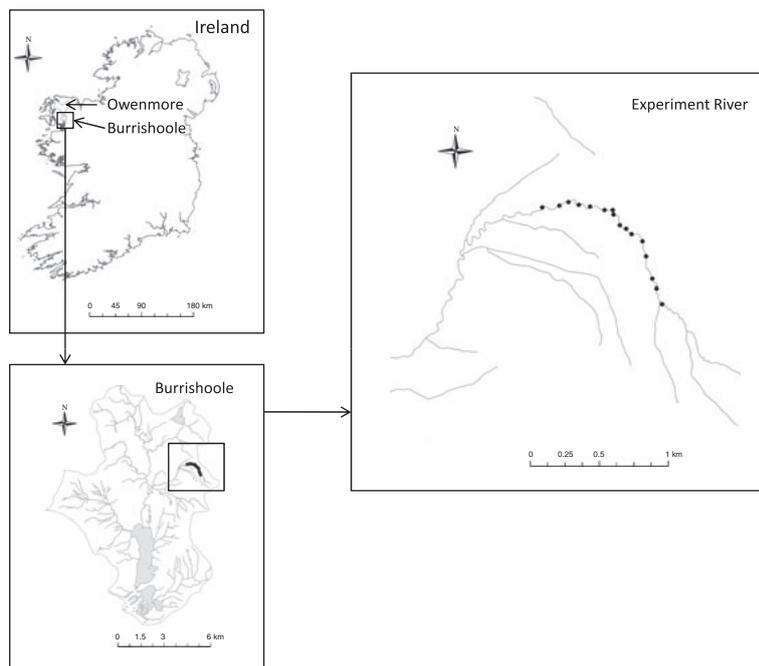
Figure 1 Location of the Burrishoole and Owenmore catchments in Ireland (top left), the experiment river in the Burrishoole catchment (bottom left) and the position of the fish trap and artificial redds in the experiment river (right).

genetic typing. The total number of migrants (smolts and presmolts combined), therefore, typed for this study was 430 individuals. 0+ and 1+ densities of surviving fish in the river were calculated in August 2002 and 2003 using removal sampling [three pass electrofishing, (Zippin 1958)]. The number of surviving fish was divided by the number of eggs planted in the river to calculate mortality at each life stage.

Genetic analysis

Natural selection resulting from disease should only be detectable at immunogenetic loci such as MH, while other forces such as genetic drift, migration and mutation should be detectable at control and immunogenetic loci (Garrigan and Hedrick 2003). To distinguish between disease-mediated selection and other forces that may impact on genetic variation, we included eight putatively neutral microsatellite loci as controls. The immunogenetic loci included in this study were as follows: (i) *Sasa-UBA-3UTR*, a microsatellite marker embedded in the 3′ untranslated region of the MH class I locus; and (ii) *Sasa-DAA-3UTR*, a minisatellite marker embedded in the 3′ untranslated region of the MH class II alpha locus (Stet et al. 2002; Grimholt et al. 2003). Class II alpha loci are highly polymorphic (Stet et al. 2002) in salmonids, and class II alpha and class II beta alleles form unique haplotypes (Stet et al. 2002); each class II alpha allele is associated with a unique class II beta allele. Therefore, characterization of either the alpha allele or the beta allele is sufficient to describe the polymorphism of class II genes. As previous work on the parr from this study indicated a strong signal of selection on *Sasa-DAA* alleles (de Eyto et al. 2007), we also unambiguously determined the *Sasa-DAA* genotype of all parents and progeny. As the relationship between *Sasa-DAA* alleles or genotypes and *Sasa-DAA-3UTR* markers is not one-to-one, in that, some of the markers are associated with more than one allele and *vice versa*, the assignment of *Sasa-DAA* genotype involved the additional typing of an intron length polymorphism in the (linked) MH class II beta (*Sasa-DAB*) locus using two primers (DBIn4ctF: ATAGAACAG AATATGGGATGG; DBIn5ctR: TTCATCAGAACAGGAC TCTCA). *Sasa-DAA/Sasa-DAB* haplotypes have, for the most part, a unique combination of embedded minisatellite and microsatellite markers, respectively (H.-J. Megens, unpublished data). In seven of 63 families, this additional typing did not resolve the *Sasa-DAA* allele, and so these families were excluded from this analysis. This reduced the number of eggs included in the analysis to 51931. Results from both the MH class II embedded marker (*Sasa-DAA-3UTR*), and also the actual MH class II allele (*Sasa-DAA*) are presented and discussed in the

following sections, and are labelled *Sasa-DAA-3UTR* (minisatellite) and *Sasa-DAA* (allele) for clarity.

We typed 746 parr and 430 migrants for the *Sasa-UBA-3UTR* microsatellite, the *Sasa-DAA-3UTR* minisatellite, the *Sasa-DAA* allele and eight control microsatellites [*One*107 (Olsen et al. 2000); *Ssa*171, *Ssa*202, *Ssa*197 (O'Reilly et al. 1996); *Ssp*2215, *Ssa*G7SP (Paterson et al. 2004); SSOSL85 (Slettan et al. 1995); *Ssa*D144 (King et al. 2005)] using fluorescently labelled primers. DNA was extracted using the Wizard SVF Genomic DNA Purification System (Promega, Madison, WI, USA) or using a Chelex method (Estoup et al. 1996). DNA samples were amplified for all the markers in three multiplex reactions using the Qiagen Multiplex PCR kit (Qiagen Ltd., Crawley, West Sussex, UK) in a final volume of 4 μL with 30 cycles of the PCR profile recommended by manufacturers at 58°C annealing temperature or in ten independent PCRs [PCR profile consisted of 3 min at 95°C, followed by 30 cycles of 30 s at 95°C, 30 s at 56°C (or 50°C for MH Class II) and 30 s at 72°]. Alleles were resolved in a ABI 377 automated sequencer (Applied Biosystems, Life Technologies Ltd, Paisley, UK) and allele sizes were evaluated against a TAMRA 350/500 size standards, or on 18 or 25 cm 6% polyacrylamide gels using a LiCor 4200 DNA sequencer with allele sizes evaluated against a 50–350 bp size standard and a cocktail of common alleles to account for any potential differences in scoring between machines.

Statistical analysis

The null hypothesis was that under neutrality (i.e. no selection), genotype frequencies in surviving fish would be equal to those expected from parental crosses. The alternative hypotheses would therefore be that either a) frequencies of heterozygotes would be higher than expected in survivors if selection took place as a result of heterozygote advantage or b) that frequencies of certain alleles would be higher than expected in survivors if additive allelic effects conferred fitness advantages. In theory, if disease-mediated selection events had occurred in the river, we would expect no evidence of selection at the eight neutral markers and conversely evidence of selection as a result of either a) or b) (or a combination of both) at the MH loci. A generalized linear model (McCullagh and Nelder 1989) was used for the analysis of the data, which was based on the comparison of observed genotype frequencies in fish surviving after 6 months and those surviving after 2 years, with expected genotype frequencies calculated from parental crosses. This analysis was conducted for each of the eight neutral microsatellite markers, *Sasa-UBA-3UTR* microsatellite, *Sasa-DAA-3UTR* minisatellite and the *Sasa-DAA* allele. Such models are very similar to standard regression models but are more

general in that the response variable can be non-normal (e.g. Poisson or binomially distributed). We assumed that the observed number of individuals Y_{ij} of each genotype ij at a particular locus (the response variable in the model) followed a Poisson distribution, with an expectation of λ_{ij}. Based on the known number of crossings made between parental genotypes, a measure of the expected number of recaptures under neutrality x_{ij}, was calculated. In the simplest case of neutrality, the expected number of recaptures λ_{ij} should be proportional to expectations based on the crossings made, x_{ij}. Thus, under neutrality, we have $\lambda_{ij} = a \prime x_{ij}$, or, taking logarithmic values,

$$\ln \lambda_{ij} = a + \ln x_{ij} \qquad (H00)$$

This represents our neutral model H00, which is a generalized linear model (McCullagh and Nelder 1989) with a log-link function and a Poisson response variable. This choice of link function ensures that the response variable takes valid values (e.g. only positive values) for all values of the right-hand side of equation (the linear predictor of the model). The term $\ln x_{ij}$ plays the role of an offset in the model, that is, a covariate for which the regression coefficient is not estimated but instead is known *a priori*.

Different extensions of H00 were considered by including terms representing mechanisms of selection. Firstly, terms representing additive allelic effects s_i and s_j of different alleles i and j were added to H00 to form model equation H01.

$$\ln \lambda_{ij} = a + \ln x_{ij} + s_i + s_j \qquad (H01)$$

If this model fitted the data well, it would indicate that fish with certain alleles had higher survival than fish with different alleles, and that this effect was additive on the log scale. This means that the survival of a particular heterozygote, say ij, lies on the arithmetic mean of the survival of the homozygotes ii and jj on the log scale. Secondly, terms representing a dominance deviation d_h where $h = 1$ for $i \neq j$ (heterozygotes) and $h = 0$ for $i = j$ (homozygotes) were added to form model H10.

$$\ln \lambda_{ij} = a + \ln x_{ij} + d_h \qquad (H10)$$

With the constraint that $d_0 = 0$, the parameter d_1 can thus be interpreted as a common increase (or decrease) in survival of heterozygotes relative to the expectation at the arithmetic mean of the respective homozygotes (the expectation under the model with only additive allelic effects, model H01). The advantage of using a common parameter representing dominance deviations for all heterozygote types is increased statistical power and a more parsimonious (simpler) model. We also fitted models where all dominance deviations d_{ij} were free parameters

(H20). Under this model, the survival of at least one heterozygote differs from the expectation under the additive allelic effects model (H01).

$$\ln \lambda_{ij} = a + \ln x_{ij} + d_{ij} \qquad (H20)$$

Finally, Model H11 included both allelic effects and a common dominance deviation:

$$\ln \lambda_{ij} = a + \ln x_{ij} + s_i + s_j + d_h \qquad (H11)$$

The family of fish that the individual belonged to (each family having a different combination of mother and father) was included as a random effect in each model. Theoretically, we could also have tried to model allelic effects and separate dominance deviations together, but estimating all these parameters together along with the family effect proved to be impossible. Thus, five models (H00, H01, H10, H20 and H11) were fitted to observed genotype frequencies of parr and migrant samples and compared separately to the expected genotype frequencies calculated from parental crosses. The parr component of this analysis has been previously published in de Eyto et al. (2007). However, owing to the different number of families included and the inclusion of 'family' as a random effect, we include a reanalysis of the parr data here, so that we can make direct comparison with the migrant data.

The constraint that all allelic effects sum to one was introduced to avoid over-parameterization of the models. The intercept and offset term was included in all models and was fitted using the GLM-function of the software package R (R Development Core Team 2004). Models were assessed based on Akaike Information Criterion (AIC) values (Burnham and Anderson 1998), which were calculated for all model alternatives. Nested model alternatives were also tested against each other using standard tests based on the change in deviance (McCullagh and Nelder 1989). In count data like these, the variance of the response variable is typically larger than expected from the model assumptions, a phenomenon known as overdispersion (McCullagh and Nelder 1989). To assess this, we also computed estimates (McCullagh and Nelder 1989) of the amount of over-dispersion for each selected model, that is, the factor c by which the variance of the response variable exceeds the theoretical variance (the variance is equal to the expectation in case of the Poisson distribution). Thus, an over-dispersion value close to $c = 1$ indicates that no over-dispersion is present.

We utilised a different statistical approach to test whether allele frequencies differed significantly between parr and migrant samples, as treating the genotype counts at the parr stage as true relative frequencies rather than estimates would be problematic as some of the counts were zero, while they were nonzero in the migrant

sample. This makes the inclusion of the ln x_{ij} offset term unfeasible and also leads to bias in the estimates of tests of significance of selection between parr and migrant stage. Instead, we used a chi-square test on a $n \times 2$ contingency table, where n is the number of alleles for the locus under consideration. This test was also carried out on egg and parr comparisons and egg and migrant comparisons to elucidate consistent patterns between the GLM models and the chi-square tests.

Results

Mortality was estimated at 89% in the first 6 months (August 2002) after introduction of the eggs into the experiment river (February 2002). Based on density estimates calculated from electrofishing, 36% of fish occurring in the 0+ age class in August 2002 survived to August 2003 (1+ age class), i.e. 64% mortality was recorded between 0+ and 1+ age classes. An estimated 41% of 1+ fish subsequently migrated through the Srahrevagh River trap, either in late summer of 2003 or as presmolts or 2+ smolts. We presume that the 59% of 1+ fish that did not migrate through the trap died, although there may have been a very small proportion surviving to migrate as 3+ smolts, or staying in the river for a third winter as sexually mature male parr. This number, however, is likely to be very small in the Burrishoole catchment. Cumulative estimated mortality from egg to smolt for the study population was 98.7%. It should be noted that survival to migrate through the Srahrevagh River trap does not give a total estimate of freshwater mortality, as the base of the Srahrevagh River is 10 km upstream of the top of the tide in the Burrishoole catchment, so additional mortality may occur during the migration between the trap and the ocean.

Akaike Information Criterion values of the five model alternatives indicated that the observed parr genotypes of five control microsatellites (*One*107, *Ssa*171, *Ssa*202, *Ssa*D144b and *Ssp*2215) and the *Sasa-UBA-3UTR* marker were close to expectations based on neutrality and the genotypes of parental crosses. Thus, for these loci, the neutral model (H00) was most appropriate as indicated by lower AIC values than for the alternative models (Table 1). AIC values indicated that observed genotypes of two control microsatellites, *Ssa*197 and *Ssa*G7SP, were better fitted by models H10 (common heterozygote advantage) and H01 (additive allelic effects), respectively. However, an explicit test of H10 vs H00 for *Ssa*197 was not significant ($P = 0.13$), while an explicit test of H01 vs H00 for *Ssa*G7SP test was marginally significant ($P = 0.016$), making it unlikely that selection had acted at these loci. For SSOSL85, the H10 model had a lower AIC value than the neutral model (H00), and the difference was highly significant (H10 vs H00, $P = 0.007$). Observed genotypes of the *Sasa-DAA-3UTR* marker and *Sasa-DAA* allele also deviated significantly from neutral expectation, and the model with the lowest AIC values was H01, which included additive allelic effects. In summary, the results from the model selection indicate that at parr stage, the observed genotypes of seven of eight control microsatellites conformed to neutral expectations, and that selection occurred at one control microsatellite (heterozygotes at this locus had higher survival) and at both the Sasa_DAA-3UTR marker and allele (fish with certain alleles had higher or lower survival than expected).

Observed genotypes of migrant fish showed a very similar signal. Observed genotypes of six control microsatellites (*Ssa*171, *Ssa*202, *Ssa*197, *Ssa*D144b, SSOSL85 and *SSp*2215) and the *Sasa-UBA-3UTR* marker were not significantly different from expectations based on

Table 1. Akaike Information Criterion values and of five model alternatives (see text for more detail) for each locus typed in Atlantic salmon surviving in the wild in a section of the Burrishoole river system 6 months (parr) and 2 years (migrants) after introduction. The lowest AIC values (in bold) are considered to be the best fit for that locus. Values for c are the over-dispersion estimates for the best fit model, and n is the number of fish successfully genotyped at that locus.

	Parr							Migrants						
	H00	H01	H10	H20	H11	n	c	H00	H01	H10	H20	H11	n	c
*One*107	**225.7**	236.0	227.1	349.4	237.7	746	1.83	272.0	**271.5**	272.3	371.3	273.1	428	1.31
*Ssa*171	**223.5**	234.1	225.1	271.6	235.5	746	1.85	**258.8**	259.3	260.5	280.5	261.0	430	1.49
*Ssa*202	**178.1**	184.4	180.1	237.1	186.3	746	2.00	**209.8**	217.0	211.7	264.1	218.9	430	1.46
*Ssa*197	293.1	298.4	**292.9**	337.5	299.3	729	1.99	**268.2**	270.7	270.2	310.1	272.0	367	1.28
*Ssa*D144b	**198.1**	208.9	199.3	293.1	210.8	727	1.77	**266.0**	277.7	268.0	331.9	279.3	369	1.37
*Ssa*G7SP	234.2	**231.0**	234.0	247.6	231.9	720	1.62	320.2	**313.3**	317.9	335.5	315.2	363	1.29
SSOSL85	227.9	237.6	**222.7**	261.4	233.8	730	1.92	**240.0**	247.3	241.7	271.8	249.2	365	1.29
*Ssp*2215	**198.7**	204.2	200.6	247.4	206.1	746	1.88	**211.5**	222.0	213.4	270.9	224.0	430	1.39
Sasa-DAA allele	452.5	**361.8**	454.4	414.9	363.7	746	2.02	374.7	**316.5**	376.7	380.0	318.4	430	1.53
Sasa-DAA-3UTR	438.5	**375.1**	440.3	404.6	376.1	746	2.35	367.7	**338.0**	365.9	367.2	339.2	430	1.74
Sasa-UBA-3UTR	**247.4**	258.7	**247.4**	309.6	259.0	746	1.83	**306.3**	310.0	311.1	364.5	311.8	430	1.39

neutrality (H00) (Table 1, lower part). The *Sasa-DAA-3UTR* marker and *Sasa-DAA* allele, as well as two control microsatellites (*One*107 and *Ssa*G7SP), displayed a signal of selection as the best model for these loci included additive allelic effects (H01) (Table 1). In these four cases, explicit tests of H01 vs H00 were significant ($P < 0.010$).

Changes in allele frequencies between egg, parr and migrant were generally very small for most of the loci, with a maximum in the order of 3%. The change in frequency at the *Sasa-DAA* locus was much more substantial, with the maximum change in allele frequency being 7.6% at *DAA*0302* between egg and migrants (Fig. 2). This is further illustrated in Fig. 3 where the pattern and magnitude change in allele frequency is similar for the egg–parr and egg–migrant comparison, with the control microsatellites having a smaller range in frequency change than the two *Sasa-DAA* loci. The change in allele frequency between parr and migrant is fairly equal across all loci (Fig. 3). However, Chi-squared tests for the changes in allele frequencies indicated that significant changes were observed between parr and migrant for two control microsatellites (*One*107 and *Ssa*G7SP) and the two *Sasa-DAA* loci (Table 2), which is consistent with the patterns identified by the GLM analysis.

Estimates of the allelic effects (parameters s_i and s_j) were calculated for each *Sasa-DAA* allele by fitting observed and expected genotype frequencies to model H01. Selection coefficients for each *Sasa-DAA* allele were then calculated from the estimated allelic effect; for example, an estimated allelic effect of 0.25 means that an individual carrying one copy of this particular allele on average experiences an increase in survival by a factor of $e^{0.25} = 1.28$; that is, a 28% increase in survival or a selection coefficient of 0.28, relative to the first allele in the data set. Selection coefficients for the *DAA* alleles ranged from −0.48 to 0.81 for parr and −0.39 and 1.03 for migrants indicating that within this population, the *DAA*-specific allele could affect survival negatively by up to 48% or positively by up to 103% (Fig. 4). The *DAA* alleles *0301* and *0302* were associated with increased survival in both parr and migrant fish, while *DAA*0304*, *1202* and *0601* were associated with decreased survival in both parr and migrant. In six of 12 alleles, selection was either positive or negative depending on the life stage in question – i.e. the direction of selection varied between life stages; for example, *DAA*0901* had a selection coefficient of 0.24 in parr but −0.39 in migrants (Fig. 4).

To assess whether large allelic effects (and hence selection coefficients) and sign changes were important, confidence intervals based on standard errors of the allelic estimates (estimate $\pm z$ 0.95 *SE) were calculated (Fig. 5). For three alleles, which displayed sign changes in their selection coefficients (*DAA*0901*, *DAA*1001* and

*DAA*0201*), the confidence intervals indicate minimal overlap between confidence intervals of egg–parr and egg–migrant estimates, indicating that the sign change in selection coefficient was important. For the three other alleles, which displayed sign changes (*DAA*0501*, *DAA*0101* and *DAA*0401*), the confidence intervals overlapped to a large extent between life stages and were not that different from 0. The two alleles that showed the highest degree of positive selection were *DAA*0301* and *DAA*0302*, and the confidence intervals of these alleles did not include zero. Similarly, the confidence intervals of *DAA*0304*, which was associated with negative selection coefficients at both life stages, were, on the whole, negative. In summary, the confidence intervals around the allelic estimates support the view that three of the alleles displayed a strong selection signal consistent between life stages (positive: *0301* and *0302*, negative: *0304*), and three displayed a selection signal that varied between life stages (*0901*, *1001* and *0201*). For the other five alleles (*0501*, *0101*, *0401*, *1202* and *0602*), the confidence intervals for allelic estimates were generally centred around zero, which indicates that in this experiment these alleles were probably not associated with survival to any large degree.

Discussion

The results presented here indicate the importance of immune genes in determining survival of salmon throughout their life stages in fresh water. At both parr and migrant stages, allele frequencies for the *Sasa-DAA* allele and marker deviated significantly from neutrality, while allele frequencies for the majority of control microsatellites and the *Sasa-UBA-3UTR* marker did not. As natural selection resulting from disease should only be detectable at immunogenetic loci, while other forces such as genetic drift, migration and mutation should be detectable at both control and immunogenetic loci, we conclude that disease-mediated selection during the 2 years of freshwater life was the cause of the deviation of the *Sasa-DAA* locus from neutrality. Although two of the control microsatellites, *One*107 and *Ssa*G7SP, are putatively neutral, our results indicate that they might not necessarily be so, and may in fact be linked to or are in linkage disequilibrium with another locus or loci, which were under selection. The additive allelic effects in migrant fish (as indicated by the selection of model H01) were much stronger for the Sasa-*DAA* immunogenetic locus than for *One*107 and *Ssa*G7SP, as indicated by the difference between observed and expected allele frequencies, which were substantially larger for the *DAA* locus than for the control microsatellites (Fig. 3). This indicates that even though there was some evidence of selection on
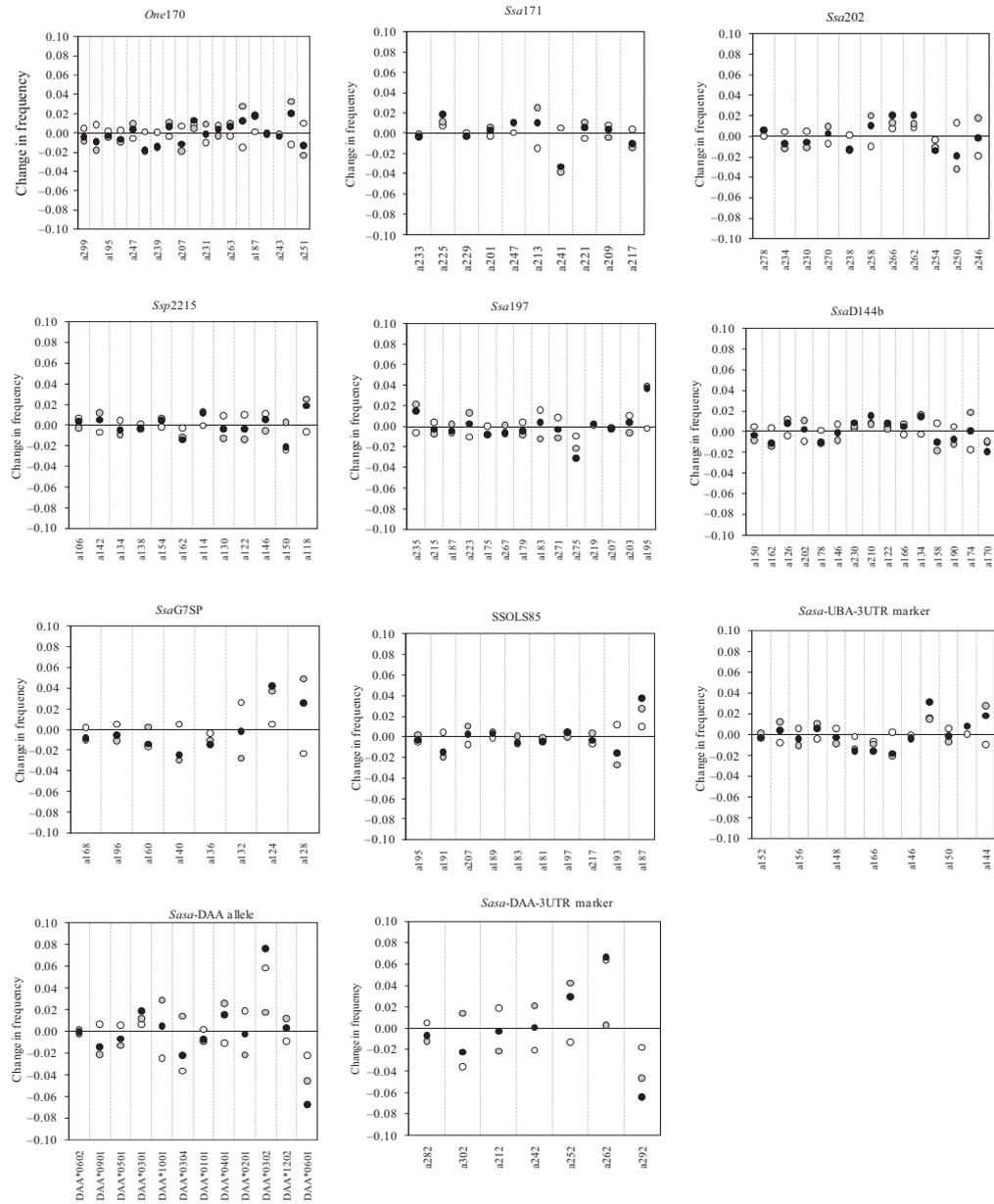
Figure 2 Changes in allele frequencies between egg and parr (white circles), egg and migrant (black circles) and parr and migrant (grey circles). Expected egg allele frequencies were calculated from parental crosses and observed allele frequencies in salmon parr and salmon migrants stage were observed after 6 months and 2 years, respectively, in a wild environment. Alleles are ordered left to right with increasing frequency in eggs.
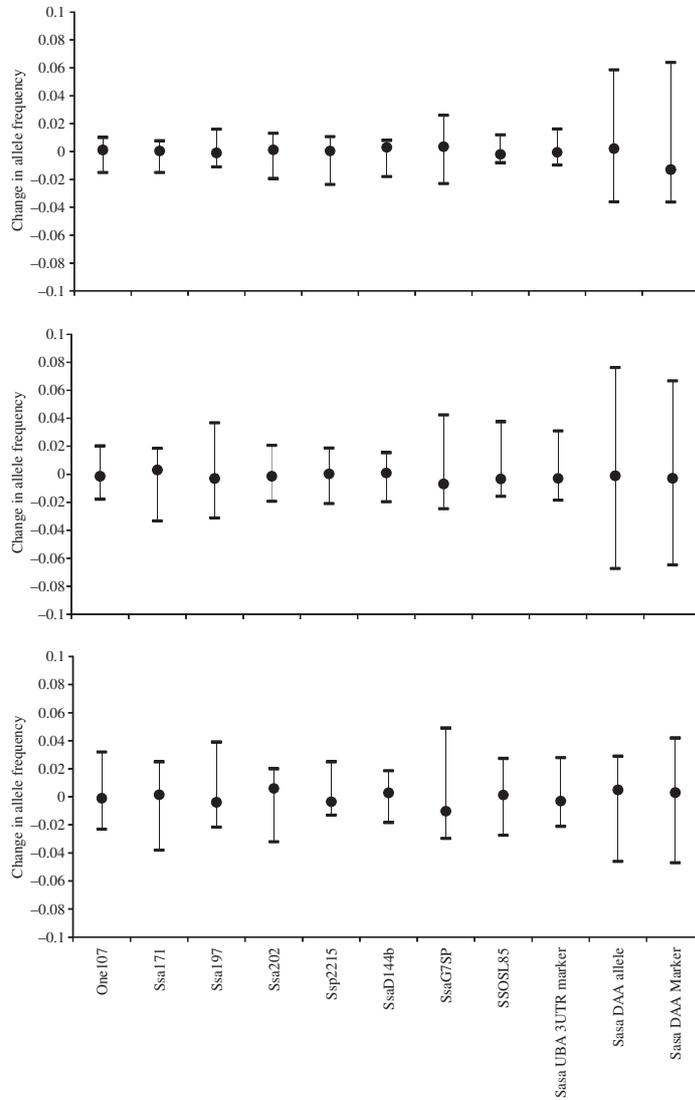
Figure 3 Changes in allele frequencies between egg and parr (top), and egg and migrant (middle) and parr–migrant (bottom) for eight control microsatellites and three immunogenetic loci. Whiskers indicate minimum and maximum changes in allele frequencies. Allele frequencies at egg stage were calculated from parental crosses, and allele frequencies in salmon parr and salmon migrants stage were observed after 6 months ($n = 746$) and 2 years ($n = 430$), respectively, in a wild environment.

these two control microsatellites, it was not as strong as the selection acting on the DAA loci.

There is still much debate over the mechanism by which MH polymorphism is maintained, and how fitness differences are conferred by MH genes. Heterozygote

advantage has been shown to be an important evolutionary mechanism in Arctic charr populations (Kekäläinen et al. 2009), water voles (Oliver et al. 2009), mice (Penn et al. 2002) and salmon (Turner et al. 2007), but has been discounted in other animals (Tollenaere et al. 2008).

Table 2. Per marker *P*-values of chi-square tests on $n \times 2$ contingency tables comparing allele frequencies between two life stages indicated above the column; number of alleles n = d.f. + 1. Significant values ($P < 0.01$) are in bold.

| Locus	d.f.	Egg–parr	Egg–migrant	Parr–migrant
One_107	○	0.132	**0.005**	**0.000**
Ssa171	9	0.638	0.014	0.072
Ssa197	13	0.107	0.035	0.069
Ssa202	10	0.257	0.095	0.068
SsaD144b	14	0.270	0.172	0.196
SsaG7SP	7	0.021	**0.004**	**0.009**
SSOSL85	9	0.533	0.280	0.292
Ssp2215	11	0.129	0.206	0.210
Sasa-DAA-3UTR	6	**0.000**	**0.000**	**0.003**
Sasa-DAA allele	11	**0.000**	**0.000**	**0.000**
Sasa-UBA-3UTR	12	0.419	0.030	0.178
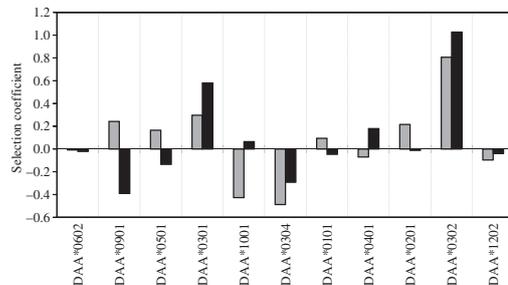
Figure 4 Selection coefficients for *Sasa*-DAA alleles in salmon surviving at 6 months in their natural environment as parr (grey bars) and at 2 years as migrants (black bars). The selection coefficients are derived from allelic point estimates calculated by fitting observed and expected genotype frequencies to model H01 (see text for details). Alleles are ordered left to right with increasing frequency in eggs. DAA*0601 is not included in the graph as it was used as the baseline in the model, against which all other selection coefficients were measured.

There is also evidence supporting the view that specific alleles are more important for disease resistance than heterozygosity (Paterson et al. 1998; Schad et al. 2005; de Eyto et al. 2007). The results presented here support the latter view, as additive allelic effects were more important than dominance deviations in determining the survival of an individual in response to disease-mediated selection. While we cannot rule out heterozygote advantage or superiority as playing a role in determining survival in this experiment, we must conclude that it is less important than the role of specific alleles, as demonstrated by the best fit of the H01 (additive allelic effects) to the data over model H10 (common dominance deviations) or H20 (separate dominance deviations). It is worth noting that the roles of heterozygote advantage and frequency-dependant

selection (or specific allele advantage) are probably not mutually exclusive (Eizaguirre and Lenz 2010), and as the number of alleles in a population increases, the difficulty in separating out the effects becomes increasingly challenging (Oliver et al. 2009). This is especially true in our case, where we had 56 families and 12 *Sasa*-DAA alleles in total. A similar experiment with a smaller number of alleles may allow the fitting of a model including additive allelic effects and separate dominance deviations without over-parameterization (Oliver et al. 2009), which was the problem that we encountered in this analysis. However, it is worth noting that the allelic variation we encountered in our experimental population (12 Sasa-DAA alleles) is what we would expect to find in any locally adapted population in this part of Ireland [14 *Sasa*-DAA alleles were sampled from four rivers close to and including the Burrishoole catchment (Consuegra et al. 2005)], and indeed, it is what we were aiming for when we designed the experiment. This is in line with the suggestion that to test the extent of disease-mediated selection as a result of local immunogenetic adaptation, the allele diversity in the translocated population should reflect the natural diversity present in the wild (Eizaguirre and Lenz 2010). As a result of this trade-off between statistical power and a true representation of the likely adaptive response in a wild population, it is difficult to be definitive about the mechanism accounting for differential survival.

Additive allelic effects are consistent with the theory of frequency-dependant selection, or variable selection in time and space (Hedrick 2002). This study indicates that the second theory may be more applicable to Atlantic salmon, as several *Sasa*-DAA alleles were significantly associated with positive selection coefficients between egg and parr stage but with negative selection between egg and migrant stage, or *vice versa*. This is in agreement with the view that temporal variation in selection means that selective mortality in one time period is a poor predictor of selection coefficients during other time periods (Kekäläinen et al. 2009). Temporal variation in selection may be because of different life stages being susceptible to different pathogens, or alternatively, that the prevalence and/or virulence of pathogens may vary seasonally and annually. Spatial variation in pathogenicity is also very likely in the case of the salmon life cycles, as 0+, 1+ and 2+ salmon use different habitats (Bardonnet and Baglinière 2000), which may harbour differing pathogens. As either spatial or temporal variation in selection (or a combination of the two) is a likely cause of the sign changes in selection coefficients between salmon life stages, our results support the theory that the high polymorphism observed at immunogenetic loci may be maintained by a combination of spatial and temporal variation in selection pressures.
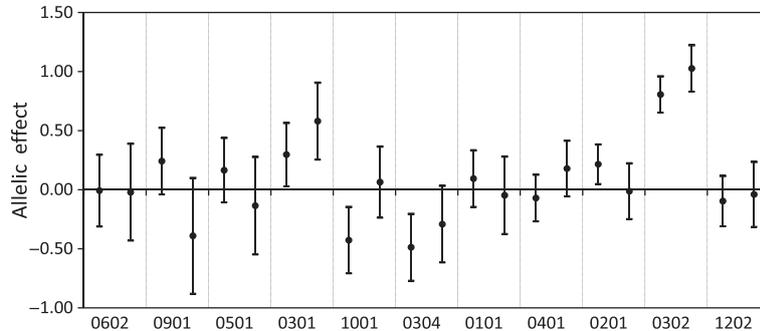
Figure 5 Allelic effects calculated for *Sasa*-DAA alleles in salmon surviving at 6 months in their natural environment as parr (left bars) and at 2 years as migrants (right bars). Allelic effects are calculated by fitting observed and expected genotype frequencies to model H01 (see text for details). Alleles are ordered left to right with increasing frequency in eggs. Whiskers represent the 95% confidence interval of the estimates. DAA*0601 is not included in the graph as it was used as the baseline in the model, against which all other allelic effects were measured.

This study highlights the fact that experiments carried out at particular life stages (e.g. parr) may be insufficient to gain a full understanding of the action of disease-mediated selection; for example, the two *DAA* alleles (*0501* and *0101*) associated with susceptibility to furunculosis in laboratory conditions (Grimholt et al. 2003) were not strongly correlated with survival at parr or migrant stage, even though furunculosis is known to occur in the Burrishoole system. This confirms that resistance or susceptibility in the laboratory does not translate easily to wild conditions, possibly because laboratory studies are generally confined to one age class of fish, most commonly adults (Dionne et al. 2009).

The results presented here show that disease-mediated selection, either carried forward as a selective advantage from the parr mortality event or as a result of subsequent disease episodes, continues to be an important but variable determinant of survival until salmon migrate to sea as smolts. Both of these scenarios are supported by our data; for example, *DAA *0301* and *DAA*0302* were both strongly associated with positive selection at parr stage, and this large effect was carried forward to the migrant stage, when these two alleles still displayed high levels of positive selection. It is also possible that the same pathogen, which these two alleles were conferring some kind of resistance to, continued to be a factor for 1+ and 2+ fish. In contrast, allele *DAA*0901* displayed positive selection at parr stage, but displayed negative selection by the migrant stage. This indicates that fish with this allele, which had survived the initially large mortality event as parr, were disadvantaged during the subsequent 18 months, perhaps as a result of a different pathogen affecting the population. It is interesting to note that the signal of the disease-mediated selection (as indicated by

the difference in the scale of allele frequency changes between neutral and *Sasa*-DAA loci) appears to be strongest at the 0+ life stage of salmon (Fig. 3). This may be simply a reflection of the sheer size of the mortality event, which occurs in the first couple of months after the fry emerge from the redds and the susceptibility of emerged fish to a new habitat. Nevertheless, the fact that three DAA alleles (*DAA*0901*, *DAA*1001* and *DAA*0201*) displayed a sign change in selection coefficient between parr and migrant stages indicates that disease-mediated selection continues to be an important determinant of survival throughout the freshwater stage of the salmon life cycle, and that the selection event at parr stage is not simply carried forward through the freshwater stages.

In the time period of this study, egg-to-smolt survival (to the experiment stream downstream trap) was 1.3%. In other words, for every 1000 eggs spawned, 13 fish survived to migrate towards the sea. However, freshwater mortality is highly variable, as indicated by the egg-to-smolt survival recorded for the entire Burrishoole catchment, which varied up to 400% between 1969 and 2006 (minimum 0.3%, maximum 1.2%, McGinnity et al. 2009). It has been shown that warmer winters reduce the survival of juvenile salmon in their first 2 years in fresh water in the Burrishoole catchment (McGinnity et al. 2009), and it has been hypothesised that this may be the result of a mismatch between photoperiod-determined emergence of fry and temperature-determined energetics of hibernating salmon. It is also possible that it is the result of disease-mediated selection, which may be greater in warmer winters, as pathogen virulence and diversity has been shown to increase with temperature (Hari et al. 2006; Okamura et al. in press). The results presented here indicate that *Sasa*-DAA variability within a population

can play a significant role in determining which fish survive the freshwater stage of the salmon life cycle, and hence how many will migrate to sea. Disease-mediated selection, therefore, may be an important factor determining the annual variation in freshwater survival in salmon. It is highly likely that disease-mediated selection on salmonids in fresh water as reported in this paper also occurs during the marine phase over very short time periods, particularly, as some of the pathogens associated with MH variability occur at sea, such as sea lice (Glover et al. 2007) and infectious salmon anaemia (Grimholt et al. 2003). The importance of MH genes in determining marine survival may be a fruitful avenue of exploration in trying to understand the continuing decline in marine survival of Atlantic salmon (ICES 2010).

The polymorphic nature of immunogenetic loci such as *Sasa-DAA* and other MH loci indicate that these loci are capable of adapting to new pathogen challenges, and it has been shown that increasing temperature may be one of the drivers of immunogenetic diversity in populations (Tonteri et al. 2010). Dionne et al. (2007) showed that large-scale genetic variability at a MH class IIβ gene in Atlantic salmon increased with increasing temperature and bacterial diversity in rivers contrary to patterns with neutral microsatellite markers. It is possible, therefore, that climate change may increase selection on MH genes, and hence increase polymorphism within wild populations in the long term, but only if they are not demographically compromised in the short term by a pathogen, which they do not have the ability to fight at a population level. Locally adapted wild populations may be most at risk of extinction in this case. The increasing production of captive-bred fish for aquaculture and stocking and the related disease risks further increase the risk of disease-mediated extinction. The combination of warmer climates and potential increases in novel pathogens and their virulence will undoubtedly put huge pressure on wild populations (Jonsson and Jonsson 2009; McGinnity et al. 2009). While it might be possible to breed disease resistance into cultured strains using information gathered about susceptibility and resistance conferred by specific MH alleles, it would be impractical to attempt this with wild populations. A more practical application of our results would be to avoid interactions between wild and farmed fish by not locating farms where projected changes in temperature are expected to be the most extreme i.e. northern latitudes (McGinnity et al. 2009) In the face of the unpredictable nature of climate change, our results highlight the importance of conserving genetic diversity (in this case immunogenetic diversity), so that wild populations have the greatest chance to adapt to emerging pathogen challenges.

Literature cited

Araki, H., B. Cooper, and M. S. Blouin. 2007. Genetic effects of captive breeding cause a rapid cumulative fitness decline in the wild. Science **318**:100.

Bardonnet, A., and J.-L. Baglinière. 2000. Freshwater habitat of Atlantic salmon (*Salmo salar*). Canadian Journal of Fisheries and Aquatic Sciences **57**:497–506.

Bergersen, E. P., and D. E. Anderson. 1997. The distribution and spread of *Myxobolus cerebralis* in the United States. Fisheries **22**:6–7.

Burnham, K. P., and D. R. Anderson. 1998. Model Selection and Multimodal Inference. A Practical Information-Theoretic Approach. Springer, New York.

Consuegra, S., H. Megens, K. Leon, R. J. M. Stet, and W. C. Jordan. 2005. Patterns of variability at the major histocompatibility class II alpha locus in Atlantic salmon contrast with those at the class I locus. Immunogenetics **57**:16–24.

Costello, M. J. 2009. How sea lice from salmon farms may cause wild salmonid declines in Europe and North America and be a threat to fishes elsewhere. Proceedings of the Royal Society B: Biological Sciences **276**:3385–3394.

Coughlan, J., P. McGinnity, B. O'Farrell, E. Dillane, O. Diserud, E. de Eyto, K. Farrell et al. 2006. Temporal variation in an immune response gene (MHC I) in anadromous *Salmo trutta* in an Irish river before and during aquaculture activities. ICES Journal of Marine Science **63**:1248–1255.

Crozier, L. G., A. P. Hendry, P. W. Lawson, T. P. Quinn, N. J. Mantua, J. Battin, R. G. Shaw et al. 2008. Potential responses to climate change in organisms with complex life histories: evolution and plasticity in Pacific salmon. Evolutionary Applications **1**:252–270.

Dionne, M., K. M. Miller, J. J. Dodson, F. Caron, and L. Bernatchez. 2007. Clinal variation in MHC diversity with temperature: evidence for the role of host-pathogen interaction on local adaptation in Atlantic salmon. Evolution **61**:2154–2164.

Dionne, M., K. M. Miller, J. J. Dodson, and L. Bernatchez. 2009. MHC standing genetic variation and pathogen resistance in wild Atlantic salmon. Philosophical Transactions of the Royal Society of London B Biological Sciences **364**:1555–1565.

Donaghy, M. J., and E. Verspoor. 2000. A new design of instream incubator for planting out and monitoring Atlantic salmon eggs. North American Journal of Fisheries Management **20**:521–527.

Eizaguirre, C., and T. L. Lenz. 2010. Major histocompatability complex polymorphism: dynamics and consequences of parasite-mediated local adaptation in fishes. Journal of Fish Biology **77**:2023–2047.

Estoup, A., C. R. Largiader, E. Perrot, and D. Chourrout. 1996. Rapid one-tube extraction for reliable PCR detection of fish polymorphic markers and transgenes. Molecular Marine Biology and Biotechnology 5:295–298.

de Eyto, E., P. McGinnity, S. Consuegra, J. Coughlan, J. Tufto, K. Farrell, H. Megens et al. 2007. Natural selection acts on Atlantic salmon major histocompatibility (MH) variability in the wild. Proceedings of the Royal Society B: Biological Sciences 274:861–869.

Ford, J. S., and R. A. Myers. 2008. A global assessment of salmon aquaculture impacts on wild salmonids. PLoS Biology 6:e33.

Fraser, B. A., and B. D. Neff. 2009. MHC class IIB additive and nonadditive effects on fitness measures in the guppy Poecilia reticulata. Journal of Fish Biology 75:2299–2312.

Garrigan, D., and P. W. Hedrick. 2003. Perspective: detecting adaptive molecular polymorphism: lessons from the MHC. Evolution 57:1707–1722.

Glover, K. A., U. Grimholt, H. G. Bakke, F. Nilsen, A. Storset, and Å. Skaala. 2007. Major histocompatibility complex (MHC) variation and susceptibility to the sea louse Lepeophtheirus salmonis in Atlantic salmon Salmo salar. Diseases of Aquatic Organisms 76:57–65.

Grimholt, U., F. Drabløs, S. M. Jørgensen, B. Høyheim, and R. J. M. Stet. 2002. The major histocompatability class I locus in Atlantic salmon (Salmo salar L.): polymorphism, linkage analysis and protein modelling. Immunogenetics 54:570–581.

Grimholt, U., S. Larsen, R. Nordmo, P. Midtlyng, S. Kjoeglum, A. Storset, S. Saebø et al. 2003. MHC polymorphism and disease resistance in Atlantic salmon (Salmo salar); facing pathogens with single expressed major histocompatibility class I and class II loci. Immunogenetics 55:210–219.

Hari, R. E., D. M. Livingstone, R. Siber, P. Burkhardt-Holm, and H. Güttinger. 2006. Consequences of climatic change for water temperature and brown trout populations in Alpine rivers and streams. Global Change Biology 12:10–26.

Harvell, C. D., C. E. Mitchell, J. R. Ward, S. Altizer, A. P. Dobson, R. S. Ostfeld, and M. D. Samuel. 2002. Climate warming and disease risks for terrestrial and marine biota. Science 296:2158–2162.

Håstein, T., and T. Lindstad. 1991. Diseases in wild and cultured salmon: possible interactions. Aquaculture 98:277–288.

Hedrick, P. W. 2002. Pathogen resistance and genetic variation at MHC loci. Evolution 56:1902–1908.

Hedrick, P. W., K. M. Parker, E. L. Miller, and P. S. Miller. 1999. Major histocompatibility complex variation in the endangered Przewalski's horse. Genetics 152:1701–1710.

Hedrick, P. W., R. N. Lee, and D. Garrigan. 2002. Major histocompatibility complex variation in red wolves: evidence for common ancestry with coyotes and balancing selection. Molecular Ecology 11:1905–1913.

ICES. 2009. Report of the Working Group on North Atlantic Salmon (WGNAS). International Council for the Exploration of the Sea CM 2009/ACOM:06.

ICES. 2010. Report of the Working Group on North Atlantic Salmon (WGNAS). International Council for the Exploration of the Sea CM 2010/ACOM:09.

Johnsen, B. O., and A. J. Jensen. 1994. The spread of furunculosis in salmonids in Norwegian rivers. Journal of Fish Biology 45:47–55.

Jonsson, B., and N. Jonsson. 2009. A review of the likely effects of climate change on anadromous Atlantic salmon Salmo salar and brown trout Salmo trutta, with particular reference to water temperature and flow. Journal of Fish Biology 75:2381–2447.

Kekäläinen, J., J. A. Vallunen, C. R. Primmer, J. Rättyä, and J. Taskinen. 2009. Signals of major histocompatibility complex overdominance in a wild salmonid population. Proceedings of the Royal Society B: Biological Sciences 276:3133–3140.

King, T. L., S. Eackles, and B. H. Letcher. 2005. Microsatellite DNA markers for the study of Atlantic salmon (Salmo salar) kinship, population structure, and mixed-fishery analyses. Molecular Ecology Notes 5:130.

Klein, J. 1986. The Natural History of the Major Histocompatibility Complex. Wiley & Sons, New York.

Krkošek, M., J. S. Ford, A. Morton, S. Lele, R. A. Myers, and M. A. Lewis. 2007. Declining wild salmon populations in relation to parasites from farm salmon. Science 318:1772–1775.

Langefors, Å., J. Lohm, M. Grahn, Ø. Andersen, and T. von Schantz. 2001. Association between major histocompatibility complex class IIB alleles and resistance to Aeromonas salmonicida in Atlantic salmon. Proceedings of the Royal Society B: Biological Sciences 268:479–485.

Lohm, J., M. Grahn, Å. Langefors, Ø. Andersen, A. Storset, and T. von Schantz. 2002. Experimental evidence for major histocompatibility complex-allele-specific resistance to a bacterial infection. Proceedings of the Royal Society B: Biological Sciences 269:2029–2033.

McCullagh, P., and J. A. Nelder. 1989. Generalized Linear Models. Chapman & Hall, London.

McGinnity, P., P. Prodöhl, A. Ferguson, R. Hynes, N. Ò Maoileidigh, N. Baker, D. Cotter et al. 2003. Fitness reduction and potential extinction of wild populations of Atlantic salmon, Salmo salar, as a result of interactions with escaped farm salmon. Proceedings of the Royal Society of London. Series B: Biological Sciences 270:2443–2450.

McGinnity, P., E. Jennings, E. de Eyto, N. Allott, P. Samuelsson, G. Rogan, K. Whelan et al. 2009. Impact of naturally spawning captive-bred Atlantic salmon on wild populations: depressed recruitment and increased risk of climate-mediated extinction. Proceedings of the Royal Society B: Biological Sciences 276:3601–3610.

McVicar, A. H. 1997. Disease and parasite implications of the coexistence of wild and cultured Atlantic salmon populations. ICES Journal of Marine Science 54:1093–1103.

Miller, M. P., and E. R. Vincent. 2008. Rapid natural selection for resistance to an introduced parasite of rainbow trout. Evolutionary Applications 1:336–341.

Okamura, B., H. Hartikainen, H. Schmidt-Posthaus, and T. Wahli. 2011. Life cycle complexity, environmental change and the emerging status of salmonid proliferative kidney disease. Freshwater Biology 56:735–753.

Oliver, M. K., S. Telfer, and S. B. Piertney. 2009. Major histocompatibility complex (MHC) heterozygote superiority to natural multi-parasite infections in the water vole (Arvicola terrestris). Proceedings of the Royal Society B: Biological Sciences 276:1119–1128.

Olsen, J. B., E. J. Kretschmer, S. L. Wilson, and J. E. Seeb. 2000. Characterization of 14 tetranucleotide microsatellite loci derived from sockeye salmon. Molecular Ecology 9:2185–2187.

O'Reilly, P. T., L. C. Hamilton, S. K. McConnell, and J. M. Wright. 1996. Rapid analysis of genetic variation in Atlantic salmon (Salmo salar) by PCR multiplexing of dinucleotide and tetranucleotide microsatellites. Canadian Journal of Fisheries and Aquatic Sciences 53:2292–2298.

Paterson, S., K. Wilson, and J. M. Pemberton. 1998. Major histocompatibility complex variation associated with juvenile survival and

parasite resistance in a large unmanaged ungulate population (*Ovis aries* L.). Proceedings of the National Academy of Science of the United States of America **95**:3714–3719.

Paterson, S., S. B. Piertney, D. Knox, J. Gilbey, and E. Verspoor. 2004. Characterization and PCR multiplexing of novel highly variable tetranucleotide Atlantic salmon (*Salmo salar* L.) microsatellites. Molecular Ecology Notes **4**:160–162.

Pedersen, A. B., and S. A. Babayan. 2011. Wild immunology. Molecular Ecology **20**:872–880.

Penn, D. J., K. Damjanovich, and W. K. Potts. 2002. MHC heterozygosity confers a selective advantage against multiple-strain infections. Proceedings of the National Academy of Science of the United States of America **99**:11260–11264.

R Development Core Team. 2004. R: A Language and Environment for Statistical Computing. R Development Core Team, Vienna.

Schad, J., J. U. Ganzhorn, and S. Sommer. 2005. MHC constitution and parasite burden in the Malagasy mouse lemur, *Microcebus murinus*. Evolution **59**:439–450.

Slettan, A., I. Olsaker, and Ø. Lie. 1995. Atlantic salmon, *Salmo salar*, microsatellites at the SSOSL25, SSOSL85, SSOSL311, SSOSL417 loci. Animal Genetics **26**:281–282.

Stet, R. J. M., B. de Vries, K. Mudde, T. Hermsen, J. van Heerwaarden, B. P. Shum, and U. Grimholt. 2002. Unique haploytpes of co-segregating major histocompatibility class II A and class II B alleles in Atlantic salmon (*Salmo salar*) give rise to diverse class II genotypes. Immunogenetics **54**:320–331.

Tollenaere, C., J. Bryja, M. Galan, P. Cadet, J. Deter, Y. Chaval, K. Berthier *et al.* 2008. Multiple parasites mediate balancing selection at two MHC class II genes in the fossorial water vole: insights from multivariate analyses and population genetics. Journal of Evolutionary Biology **21**:1307–1320.

Tonteri, A., A. Vasemagi, J. Lumme, and C. R. Primmer. 2010. Beyond MHC: signals of elevated selection pressure on Atlantic salmon (*Salmo salar*) immune-relevant loci. Molecular Ecology **19**:1273–1282.

Turner, S. M., M. Faisal, and J. A. DeWoody. 2007. Zygosity at the major histocompatibility class IIB locus predicts susceptibility to *Renibacterium salmoninarum* in Atlantic salmon (*Salmo salar* L.). Animal Genetics **38**:517–519.

Wegner, K. M. 2008. Historical and contemporary selection of teleost MHC genes: did we leave the past behind? Journal of Fish Biology **73**:2110–2132.

Wynne, J. W., M. T. Cook, B. E. Nowak, and N. G. Elliot. 2007. Major histocompatibility polymorphism associated with resistance towards amoebic gill diesease in Atlantic salmon (*Salmo salar* L.). Fish and Shellfish Immunology **22**:707–717.

Zippin, C. 1958. The removal method of population estimation. Journal of Wildlife Management **22**:82–90.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. The number of occurrences of alleles of 11 genetic loci in eggs of Atlantic salmon (predicted from parental crosses) and in the same fish surviving in a natural environment after 6 months (parr), and after 2 years (migrants).

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Paper II

Paper III

Comparison of non-Gaussian quantitative genetic models for migration and stabilizing selection

Jisca Huisman[1] and Jarle Tufto[2]

[1]Centre for Conservation Biology, Department of Biology, Norwegian University of
Science and Technology, N-7491 Trondheim, Norway
[2]Centre for Conservation Biology, Department of Mathematical Sciences, Norwegian
University of Science and Technology, N-7491 Trondheim, Norway

Abstract

The balance between stabilizing selection and migration of maladapted individuals has been modelled using a variety of quantitative genetic models of increasing complexity, including models based on a constant expressed genetic variance and models based on normality. The infinitesimal model can accommodate non-normality and a non-constant genetic variance as a result of linkage disequilibrium. It can be seen as a parsimonious one-parameter model which approximates the underlying genetic details well when a large number of loci are involved. Here, the performance of this model is compared to several more realistic explicit multilocus models, with either two, several or a large number of alleles per locus with unequal effect sizes. Predictions for the deviation of the population mean from the optimum are highly similar across the different models, so that the non-Gaussian infinitesimal model forms a good approximation. The infinitesimal model does however generally estimate a larger genetic variance than the multilocus models, with the difference decreasing with an increasing number of loci. The details of the multilocus model seem of secondary importance, as the difference between models depends more strongly on the effective number of loci than on the number of alleles per locus and distribution of effect sizes.

Keywords infinitesimal model, multilocus, admixture, linkage disequilibrium, migration-selection balance

Introduction

Migration of individuals from one population to another is a common phenomenon, and so is local adaptation. This combination results in migrants often having lower fitness in their new environment than resident individuals, so that migration and selection work as opposing forces on the genetic make-up of a population. Migration-selection balance has been thoroughly explored in population genetic single-locus models (Haldane, 1930; Wright, 1931), and to some extend for quantitative traits, coded for

by a large number of loci. To allow analytical tractability, this has generally been done using the simplifying assumtions of normally distributed genotypes and a constant genetic variance (Via and Lande, 1985; Hendry et al., 2001; Tufto, 2001), or a normal distribution with a dynamically changing variance (Barton, 1999; Tufto, 2000; Barton, 2001; Bulmer, 1980, p. 181). It has been less extensively studied in a situation with both high migration rates and strong stabilizing selection, when the approximation of the genotype distribution with a Gaussian breaks down.

This situation potentially occurs when a wild species interbreeds with its domesticated relative. The domestic population is usually much larger than the wild one, and selection acting against escaped or released individuals is generally strong. Such interbreeding occurs in various species of both plants and animals, ranging from sunflowers (Whitton et al., 1997) to wolves (Randi et al., 2000) and wildcats (Beaumont et al., 2001), and may pose a threat to the genetic integrity of a wild population, or even an entire species (Rhymer and Simberloff, 1996). While such interbreeding is often unintentional, similar processes occur during supportive breeding and reintroduction programs, which promote breeding between wild individuals and individuals which have adapted to captivity. In both cases, the domestic or captive bred individuals are expected to have lower fitness than native wild individuals in the local natural environment, due to (genetic) differences created by adaptation to captivity (Frankham, 2008) and/or artificial selection. As most fitness traits have a heritable component, the average fitness in an admixed population resulting from such interbreeding may be lower

than in the original population, depending on the number of domestic/captive bred individuals and the relative fitness of them and their descendants. Lower average fitness in a population increases risk of extinction (Hutchings and Fraser, 2008), by increasing the rate of selective mortality in density-dependent populations, or by decreasing population size in density-independent populations (Lande, 1976b).

An example of interbreeding between a wild species and its domesticated relative can be found in the Atlantic salmon (*Salmo salar*) (McGinnity et al., 1997; Gross, 1998; Naylor et al., 2005). In Norway, the number of farm salmon escaping yearly is several times the estimated size of the wild breeding population (Jonsson, 1997; Gross, 1998; Fleming et al., 2000). Their mortality before they reach the breeding grounds is high, but nevertheless the proportion of farm escapees amongst wild breeders is around 20%, and up to 80% in some rivers (Fiske et al., 2001). Farm salmon differ from wild salmon due to adaptation to the captive environment, which in salmonids has been shown to decrease breeding success in the wild already after one or two generations (Hansen, 2002; Araki et al., 2007). Additionally, they have been selected for increased growth and several other traits for around ten generations (Gjøen and Bentsen, 1997; Gross, 1998). This is much shorter than most other domesticated species, but farm salmon differ already considerably from their wild counterparts (Gjøen and Bentsen, 1997; Fleming and Einum, 1997; Gross, 1998). As in many fish species, growth and body size are correlated with fitness traits like survival and fecundity (Sundt-Hansen et al., 2007; Walsh et al., 2006). Therefore, the change in growth performance is likely

to have moved the multi-trait phenotype away from the fitness optimum in the wild, via indirect response to selection in other traits. Studies on the performance of farm and wild Atlantic salmon and their hybrids under natural conditions have been performed, but do not extend beyond the second generation (Fleming et al., 2000; McGinnity et al., 2003; Araki et al., 2008; Fraser et al., 2010), partly due to the long generation time (3-10 years) of Atlantic salmon.

To make long term predictions about for example the amount of introgression or the decrease in population fitness, modelling studies are useful. Historically, models needed to be analytically tractable in order to solve them without the aid of computers. Examples are the one- and two-locus models, which still prove useful under certain conditions (see e.g. Bolnick et al., 2007). Some fitness traits may be predominantly determined by a single gene; for example in salmon 83% of genetic variation in resistance against a certain viral disease is explained by a single gene (Moen et al., 2009).

Most fitness traits however are continuous (quantitative) rather than discrete. The number of genes typically coding for a quantitative trait is disputed, partly because it is not well understood what a "gene" is in this context, but is generally believed to be large (Lande, 1981; Merilä and Sheldon, 1999). Genetic details such as interactions between alleles at the same or different loci are commonly unknown. Generally it is however assumed a measuring scale can be found on which all genetic variance is close to additive (Lande, 1981). This notion is supported by animal breeding practice and many (but not all) experiments, including experiments in which hybrid per-

formance in salmon was explainable by additive effects only (McGinnity et al., 2003; Fraser et al., 2010).

Infinitesimal models

To make inferences about quantitative genetic traits different models have been developed, with different sets of assumptions to allow analytical tractability. A commonly used model is the infinitesimal model, which assumes an infinite number of unlinked and non-epistatic loci underlying the trait, each with an infinitesimal effect (Fisher, 1918; Falconer and Mackay, 1996; Lynch and Walsh, 2008). From the central limit theorem this leads to normal within-family distributions of breeding values (additive genotypes). Within each family, the means are equal to the midparental value and all variances are equal to half the variance under linkage equilibrium ($V_{G,LE}$, genic variance) (Turelli and Barton, 1994; Dawson, 1997). Under random mating and in absence of migration or selection (i.e., under Hardy-Weinberg and linkage equilibrium), the variance of the population breeding value distribution equals $V_{G,LE}$ (Turelli and Barton, 1994). Also in the limiting case of weak migration and/or weak selection the population variance is reasonably well approximated by $V_{G,LE}$, which stays constant over time. In those cases, one needs to focus only on changes in the mean breeding value to describe the entire distribution (Tufto, 2000). Needing to track only a single variable gives rather straight forward equations, used by for example Via and Lande (1985), Hendry et al. (2001) and Yeaman and Guillaume (2009).

The infinitesimal model assumes that $V_{G,LE}$ stays constant, as even under

strong selection on the phenotype, the strength of selection acting on each of the (infinitely) many loci is very small, so that the change in allele frequency at each locus is very small too and can be ignored (Lynch and Walsh, 2008). Migration will however create positive linkage disequilibrium (LD), as alleles within immigrant gametes tend to occur in different combinations than within local gametes. This increases the population variance relative to the variance expected in a random mating population with the same allele frequencies (and allelic effects), the variance under linkage equilibrium $V_{G,LE}$. We define the total amount of LD as the difference between the observed variance V_z and $V_{G,LE}$; the sum of all covariance terms between and within loci. Similarly, both truncation and stabilizing selection tend to create some negative linkage disequilibrium. When accounting for this changing variance, the breeding value distribution of the admixed population can still be adequately approximated by a Gaussian, even for surprisingly high migration rates and when immigrants are highly maladapted, or when selection is strong (Turelli and Barton, 1994; Barton, 1999; Tufto, 2000).

Under some conditions however, the mean and the variance in the next generation cannot be predicted accurately by the mean and variance in the current generation alone, because the distribution is bimodal or considerably skewed. In those cases it is still possible to use the infinitesimal model, assuming constant $V_{G,LE}$ and equal within-family variances, but one needs to keep track of the whole population distribution (Turelli and Barton, 1994) rather than assume a Gaussian distribution.

In reality however, the number of loci coding for a trait will be finite, so that allele frequencies and thereby $V_{G,LE}$ will change under influence of selection and migration. Additionally, the assumption of equal within-family variances may be violated when LD is strong (Dawson, 1997); one could imagine more variation among offspring of local-by-immigrant matings than of local-by-local matings for example. Multilocus models can incorporate this, but they also involve assumptions that enable analytical solutions or decrease computational costs.

Multilocus models

Two commonly used multilocus models are the continuum-of-alleles model and the exchangeable (equivalent) loci model. The first is (again) inspired by the central limit theorem, and assumes that each locus has an infinite number of possible alleles. Allelic effects are normally distributed, with the same variance for each locus (Kimura and Crow, 1964; Lande, 1976a; Yeaman and Guillaume, 2009). It is more restrictive than the assumption that breeding values are approximately Gaussian, and it has been argued that it is motivated by questionable assumptions about mutation-selection balance (Barton and Turelli, 1989).

The exchangeable loci model assumes all loci are diallelic and have the same allelic effect, such that only the total number of '+' alleles needs to be traced (Barton, 1992; Turelli and Barton, 1994). A special case of this is the symmetric model, which in addition assumes that all allele frequencies are equal; a questionable assumption which leads to coupling of the dynamics of the variance to the dynamics of the population mean (Tufto, 2000).

A more natural assumption may be ex-

ponentially distributed effect sizes across loci (Orr, 1998). Allowing for different allelic effects among loci makes a model analytically intractable and (much) more computationally intensive than the other two types of multilocus models. Now one needs to track the frequency of all possible haplotypes (2^L for L diallelic loci), or all individual genotypes.

Here, the infinitesimal model for will be compared to these three types of multilocus models, as well as a model where the allelic effects are normally distributed as in the continuum-of-alleles model, but the number of possible alleles is limited (≥ 2). For the infinitesimal model, we keep track of the entire distribution of breeding values, thus allowing for linkage disequilibrium and non-normality. Both are expected under a situation of strong stabilizing selection and high rates of immigration of maladapted individuals. The main differences between the infinitesimal model and the multilocus models lie in two assumptions of the infinitesimal model: negligible changes in allele frequencies (constant $V_{G,LE}$) and equal, constant within-family variances.

Whether infinitesimal models or multilocus models more closely resemble the genetic architecture underlying fitness traits is debatable. There are arguments in favour of both models; sustained response to selection after 100 or more generations favours the infinitesimal model, and despite its apparent lack of biological realism it has a proven record in animal breeding practice. On the other hand, quantitative trait loci (QTL) mapping studies frequently identify loci of large effect. Their effect size may however be overestimated (Barton and Keightley, 2002), as one such "locus" may be a cluster of several tightly linked loci. For the current simulation study it is irrelevant

whether a locus consists of such a cluster or a single point mutation; we assume the locus stays intact during the relatively limited time span of the simulations.

The lack of knowledge on the genetic details underlying traits of interest, especially in wild populations, also means it is unknown which multilocus model forms the most suitable approximation. Therefore, the focus of this paper is on how much, and under what conditions, the simpler infinitesimal model differs from the various multilocus models, with minor attention to the difference between the multilocus models.

The infinitesimal model has been compared previously in the same setting to models assuming a normal distribution of breeding values, with and without taking LD into account, as well as to a symmetric multilocus model (Tufto, 2000). It has also been compared to a wide range of different models in a similar setting by Turelli and Barton (1994). It has, to our knowledge, not been compared previously to the "exponential effects" diallelic multilocus model or the model with a limited number of alleles per locus, for which no analytical approximation exist.

The trait considered is a non-specified heritable trait affecting fitness, which is under stabilizing selection in the recipient population. This population receives migrants from a donor population, in which the average trait value is some distance away from the optimum as a result of repeated truncation selection in that population. The models used here deal with a single trait rather than a set of genetically correlated fitness traits, an approximation that has been shown to hold well (Tufto, 2010). Often only the limiting case of weak selection and low migration rates is considered (see e.g Barton and Turelli, 1991; Lynch and Walsh,

2008), but the comparison of the two models considered here is especially relevant under strong selection and high rates of migration, when the breeding value distribution is expected to deviate from a Gaussian.

We shall see that the difference between the models is very small for the deviation of the population mean from the optimum. The difference in variance between the infinitesimal model and the multilocus models is mainly due to changes in $V_{G,LE}$, which are up to 20% when the number of loci is limited. The difference between the models in total genetic variance is generally smaller than the difference in $V_{G,LE}$, as the amount of linkage disequilibrium is smaller for the multilocus models than for the infinitesimal model, especially when selection is strong. Despite these differences between the two model types in the predicted variance, as well as in the skew, they predict a highly similar response to selection and resulting population mean.

Models

We consider the breeding value, the additive genotypic value, for a non-specified trait affecting fitness. The fitness effect includes any effect of this trait on fecundity (breeding success) and/or on survival. The breeding values in the population are assumed to be initially normally distributed, but not necessarily so after selection and migration. Each generation, a fraction of the population is replaced by immigrants with a different mean breeding value, followed by stabilizing selection around a fixed optimum and random mating (figure 1).

A population of fixed size N is simulated, with no age or other structure and no sexes, and discrete generations.

A population of finite rather than infinite size is used, as this allows us to trace individual genotypes rather than a potentially large number of haplotype frequencies(2^L; e.g. about 10^{30} for $L = 100$ loci). To ensure any differences between the models are not due to effects of a finite versus infinite population size, a finite population size model is used for the infinitesimal model as well. To minimize stochastic noise, a large population size is used ($N = 1000$) and averages are taken over 100 replicate runs. Results of the finite population size infinitesimal model are nearly identical to an infinite population size version of this model, using the exact Fourier transform method as described in Tufto (2000), for the range of parameter values considered (unpublished results).

No true deterministic migration-selection equilibrium occurs in a finite population, as fluctuations around this level will keep occurring. Therefore comparisons of the models were done after 50 generations; for most parameter value combinations considered changes in all statistics leveled off after about 10 to 15 generations.

Simulations are written in R, and partly in C for faster computation time.

Simulation details

A population is simulated with breeding values for a fitness trait initially following a standard normal distribution, i.e. mean breeding value $\bar{z} = 0$ and variance under gametic phase equilibrium (genic variance) $V_{G,LE} = 1$, so that all genetic differences are expressed in genetic standard deviations. In the multilocus models, an individual's breeding value is the sum of the allelic effects over all its loci; for details see Appendix. During the sim-

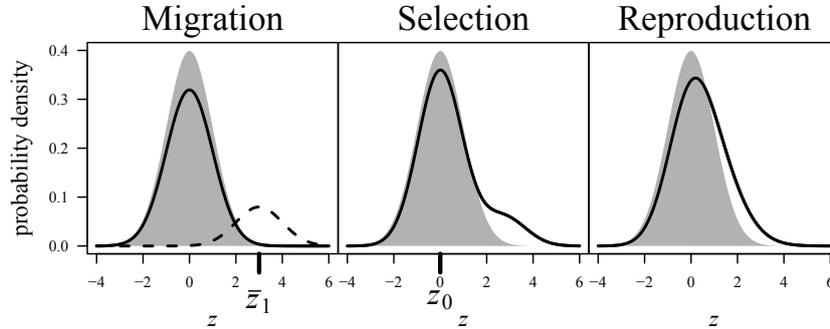## Migration   Selection   Reproduction

Figure 1: Distribution of the breeding values z in the recipient population (solid line) after 20% is replaced by immigrants from the donor population with $z_1 = 3$ (dashed line) just after migration (left panel), in the selected parents (middle, $s = 0.1$) and after reproduction (right). The grey shaded area indicates the distribution in the original population. The post-reproduction distribution is the pre-migration distribution in the next time step.

ulations, the total additive genetic variance V_z is measured as the variance of the breeding value distribution. Under the multilocus model, $V_{G,LE}$ is calculated from the allele frequencies and allelic effects, under the infinitesimal model it is fixed to 1.

The population experiences stabilizing selection around the optimum z_0 which, without loss of generality, is set to 0. The population mean is thus initially at the optimum. Stabilizing selection is implemented by letting the relative probability to be sampled as parent be $w(z) = \exp\left\{-\frac{1}{2}s(z - z_0)^2\right\}$, with parameter s representing the strength of selection acting on the genotypes. The parameter s relates to the strength of selection acting on the phenotype s_P as $1/s = 1/s_P + V_E$, where V_E is the environmental variance. The selection coefficient s_P is the inverse of the variance of the fitness function, $1/s_P = \omega^2$, as used by Lande (e.g. Lande, 1976b). Selection may act on viability, fecundity or both, provided the fecundity of a pair is the product of their individual fecundities (Bodmer, 1965).

Each time step N pairs of individuals are sampled (with replacement) to produce N offspring whose genotypes are either sampled from a normal distribution around the mean of the parents, with a fixed variance equal to half $V_{G,LE}$ (Fisher, 1918; Turelli and Barton, 1994) (infinitesimal model), or a sample of parental alleles following Mendelian inheritance (multilocus model). There is no separate selection on offspring survival; the resulting offspring distribution can be thought of as the distribution of offspring which survive until reproductive age.

Migration takes place before the reproduction and selection steps. A constant fraction m of the population is replaced by Nm individuals from a donor population; individuals from both populations mate at random.

Donor (farm) population The donor population has mean breeding value z_1, and is assumed to have the same alleles as the recipient population, but with different frequencies. Therefore the donor population is created via directional se-
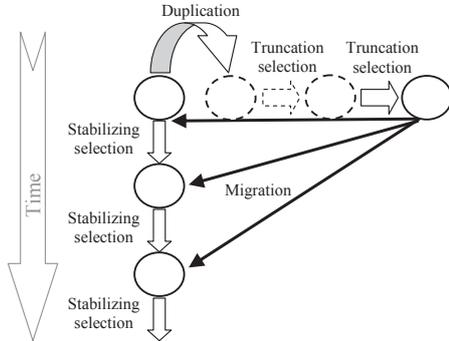
Figure 2: Schematic overview of the model used. The donor population is created by duplicating the original population at $t = 0$ and applying truncation selection until the average breeding value z_1 is obtained. Each time step, migration takes place from this donor (farm) to the recipient (wild, local) population, followed by random breeding and stabilizing selection.

lection from the recipient population (figure 2), rather than by creating a second, independent population (also for infinitesimal model). The donor population is generated by taking a random sample of size N from the first population at $t = 0$ and applying truncation selection on the phenotype P of the trait for several generations, followed by random mating, until the desired genetic difference $z_1 \pm 0.1$ is obtained. Phenotypes are created by adding random terms E from a normal distribution with zero mean and variance V_E to the breeding values z ($P = z + E$). The phenotypes are used solely for the simulation of directional selection in the donor (farm) population.

Due to this set up, variance in the donor population is generally lower than in the base population, and some negative LD among immigrants is created (table 2 in Appendix). In this way it differs from analytical models, where sometimes optionally a lower or higher variance in immigrants can be incorporated (Tufto,

2000), but generally linkage equilibrium among them is assumed.

Parameter ranges The set up described allows three parameters to vary: The migration rate m, strength of stabilizing selection s and genetic difference between the populations $z_1 - z_0$ ($= z_1$ since $z_0 = 0$), the latter two of which are scaled relative to the genetic variance $V_{G,LE}$. The ranges used for those parameters are chosen so as to correspond to values observed in nature. For the migration rate, defined here as the proportion of non-native individuals in the breeding population, the entire range of 0% to 100% is considered, to include potentially low rates in some species and the very high rates in some salmon rivers (up to 80%, Fiske et al. (2001)), and to allow indirect comparison of studies with abruptly changing optima ($m = 1$). Two levels of selection intensity are considered, mild ($s = 0.1$) and relatively strong ($s = 0.5$) (Kingsolver et al., 2001), corresponding for example to $\omega^2 = 9$ respectively $\omega^2 = 1$ if $h^2 = 0.5$. The genetic difference between the recipient and donor population ranges from 0.5 to 5 genetic standard deviations, so that the average relative fitness of immigrants is 94% to 31% for the weakest selection ($s = 0.1$), and 78% to 1.3% for the strongest selection ($s = 0.5$).

Multilocus models

Three different multilocus models are compared, which have as main difference the number of possible alleles per locus: $n = 2$, a random number ($n \sim \text{Pois}(5) + 2$), or $n = 2N$ (infinite) (table 1). For the latter two, the allelic effects are assumed to be normally distributed around 0 with a constant variance across loci. For the diallelic model, exponentially distributed effects across loci are compared to a situation where all loci have the same effect size (exchangable loci). Allelic effects were scaled at initiation to ensure $V_{G,LE} = 1$ (see Appendix). Allelic ef-

Table 1: Characteristics of the multilocus models.

Model	no. alleles per locus	allelic effects	no. loci $L_e = 5$	no. loci $L_e = 20$
diallelic (eq)	2	equal	6–9	26–34
diallelic (exp)	2	exponential	10–53	92–236
multi-allelic (Pois)	Poisson(5) + 2	Normal	6–10	29–40
multi-allelic (2N)	2N	Normal	5	20

fects are constant during a run, no mutations are incorporated.

Initial allele frequencies are distributed either uniform or U-shaped, and are sampled from a Beta distribution for the diallelic models, and from a symmetric Dirichlet distribution for the multi-allelic distributions. For brevity, only the results for the U-shaped distribution will be shown in the Results, which is the distribution expected under mutation-selection equilibrium (Bulmer, 1989). Only for the diallelic model did the allele frequency distribution affect the results, partly via the different scaling of allelic effects at initiation (example in Supplementary Information figure S3).

Effective number of loci   Since replicate runs will have different allele frequencies and allelic effects, the distribution of relative contributions of loci to the genetic variance will differ in shape between runs. To account for this, the effective number of loci L_e is used rather than the actual number L. When all loci contribute equally to the genetic variance (as under the infinite alleles model), the effective number equals the actual number ($L_e = L$), but it is lower otherwise ($L_e < L$) (table 1). The actual number of loci L therefore differs between replicate runs, between boundaries that are based on the empirical relationship between L and L_e, specific to each model and allele frequency distribution (see Appendix for details).

This "variance effective number" is not necessarily the same as the effective (minimum) number of loci contributing to the difference between two lines or populations, es-

timated using crossing experiments (Lande, 1981; Barton and Turelli, 1989). They are probably in the same order of magnitude, therefore (variance) effective number of loci of 5 and 20 are used, as the minimum number of genes contributing to differences between population has been estimated to be about 5 to 10, up to 20 (Lande, 1981). Using fewer than 5 effective loci would allow for more in depth study of the effect of the number of loci underlying the trait, but limits the maximum breeding value that can be created, and therefore hampers the creation of donor populations (figure 9 in Appendix).

Statistical tests

Statistical analysis on the difference between the infinitesimal model and the various multilocus models is performed using two-sided two-sample t-tests assuming unequal variances (Welch's t-test), with sample sizes $n_1 = n_2 = 100$ replicate runs. This is done at regular intervals of m for $z_1 = 3$, and at regular intervals of z_1 for $m = 0.2$, both for $s = 0.1$ and $s = 0.5$. Those parameter combinations were initially chosen to reflect the migration rate and wild-farm genetic difference in Norwegian Atlantic salmon. It was found that the difference between the models is largest around $m = 0.2$ and $z_1 = 3$, and as such they provide a "worst case scenario" for the comparison of the models. P-values of those tests are binned into the standard categories $p < 0.001$ (***), $0.001 < p < 0.01$ (**), $0.01 < p < 0.05$ (*), $0.05 < p < 0.1$ (.) and $p > 0.1$ (n.s.).
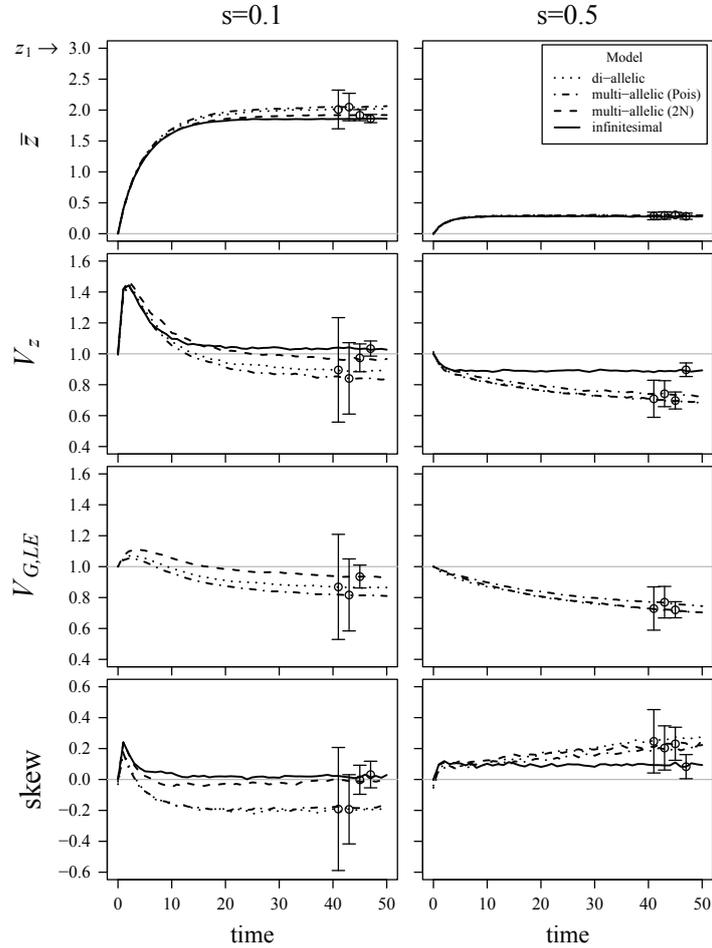
Figure 3: Example of change in mean, variance (observed and under linkage equilibrium) and skewness over time, and amount of variation between models and between replicate runs. For each model, the average over 100 replicate runs is shown (black lines, see legend) and error bars indicate average $\pm$ 1 SE, with order as in legend. (figure with 10% – 90% quantiles over time included in Supplementary Information)

Results

The differences between the various models will be shown as differences in the mean, variance and skewness of the breeding value distribution, as well as in linkage disequilibrium, quantified by the difference between the observed and genic variance. First an example will be given how those variables change over time during admixture, followed by comparisons after 50 generations of admixture between the infinitesimal model and the multilocus models, and among the various multilocus models.

Two examples of how the mean and variance of the breeding value distribution

change during 50 generations of admixture are given in figure 3. The deviation of the mean from the optimum, $\bar{z}$, approaches an asymptotic value within a dozen or so generations. The predictions for $\bar{z}$ of the infinitesimal model and multilocus models differ only slightly for relatively weak selection ($s = 0.1$, left), and are nearly indistinguishable for stronger selection ($s = 0.5$, right). The difference between the models is smaller than the variation between replicate runs of the same model, indicated by the error bars.

In those examples, as under most conditions (see later), the observed variance of the multilocus models is lower than that of the infinitesimal model (figure 3, 2nd row). For weak selection the observed variance V_z is slightly higher than $V_{G,LE}$, and much higher in the first generations of admixture, as migration generates positive LD. Under stronger selection this effect of migration is more than balanced out by the negative LD created by selection, resulting in V_z being somewhat lower than $V_{G,LE}$.

Skew stays very close to zero for the infinitesimal model, while it is more extreme and variable for the multilocus models. When intensity of selection is low, the peak of the distribution is shifted away from the optimum in the direction of the immigrants, to the right, leaving a long tail behind at the left side of the peak — a left-skewed (negatively skewed) distribution. Under stronger selection, the peak stays closer to the optimum, but there is a long tail in the direction of the immigrants — the distribution becomes positively or right-skewed.

Infinitesimal vs. Multilocus models

The small difference between the models in $\bar{z}$ occurs not only in the two earlier described examples, but is a general trend when the values at $t = 50$ are plotted against migration rate (upper row in figure 4) or the immigrant mean (upper row in figure 5). Where the models differ, the multilocus models predict a higher $\bar{z}$ than the infinitesimal model for intermediate values of the immigrant breeding value z_1 (figure 4, $z_1 = 3$), and lower for more extreme values of z_1 (figure 5). The differences are larger for $s = 0.1$ than for $s = 0.5$, and although the difference between the infinitesimal model and the least different multilocus model are sometimes highly significant ($p < 0.001$, indicated with *** in the figures), they are never more than 11% for $z_1 = 3$ to 18% for larger z_1 (difference between averages over replicate runs).

The major difference between the model types is in the estimate of the variance, as expected from theory. For the infinitesimal model all changes in genetic variance V_z are due to changes in LD, as the genic variance $V_{G,LE}$ is fixed at 1 (compare 2nd and 3rd rows in figures 4 and 5, note different scales on the y-axes).

For a given strength of selection and immigrant mean, the negative LD created by stabilizing selection dominates when migration rate is low, so that $V_z < V_{G,LE}$. For intermediate migration rates, positive LD created by migration dominates, causing $V_z > V_{G,LE}$. For still higher migration rates, the negative LD among immigrants dominates, so that again $V_z < V_{G,LE}$. What migration rates are 'intermediate' or 'high' in this context depends on both the strength of selection (figure 4) and z_1 (not shown).

For the multilocus models, any change in V_z is generally mainly due to changes in $V_{G,LE}$, with LD being smaller (closer to zero) than for the infinitesimal model with the same parameter values (figures 4 and 5, 3rd row), while the variance V_z is further away from one (figures 4 and 5, 2nd row). Exception is when selection is strong and migration rate rather high ($m = 0.4$–0.6 for $z_1 = 3$), where the infinitesimal model predicts a balance between those forces such that LD=0, while the multilocus models predict a slightly positive LD (figure 4).
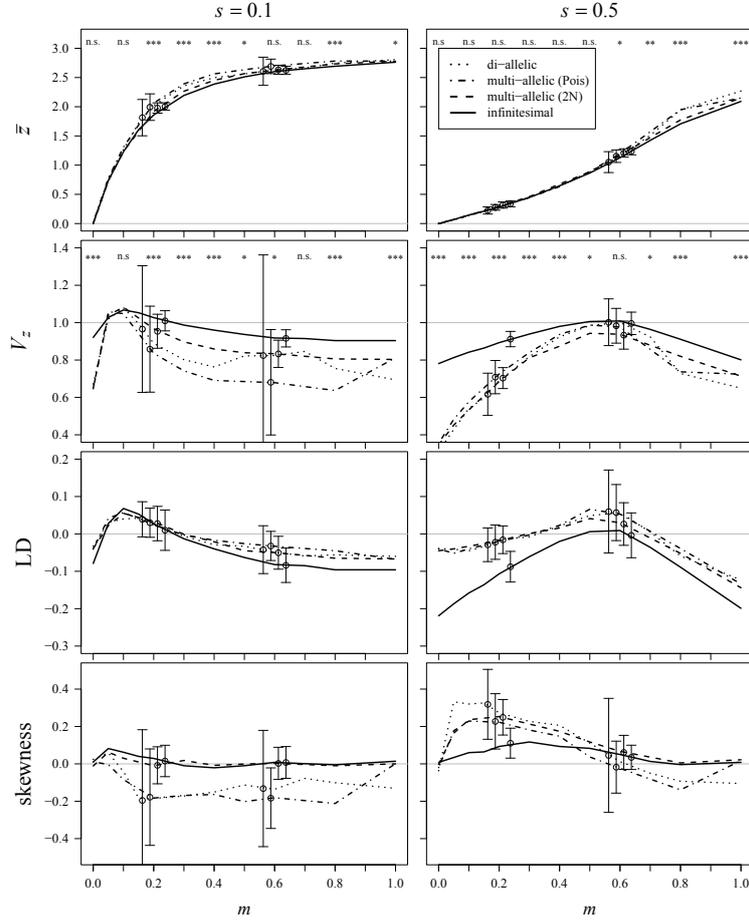
It must be noted that the variance is not

Figure 4: Model differences as function of migration rate m: Mean breeding value $\bar{z}$, genetic variance V_z, linkage disequilibrium LD and skewness for the different models (see legend, multilocus models $L_e = 5$) after 50 generations of recurrent immigration with rate m of immigrants with mean breeding value $z_1 = 3$ and relatively weak ($s = 0.1$, left) or strong ($s = 0.5$, right) stabilizing selection. Error bars indicate average $\pm$ 1 SE (with order as in legend), asterisks indicate significance levels of Welch's t-tests between the infinitesimal model and the least different multilocus model (based on 100 replicate runs).

only determined by the amount of LD generated by migration and selection, but also by the reduced variance in the donor population (table 2 in Appendix), especially when m and z_1 are large. This reduced variance in the donor population is due to both changed allele frequencies (in the multilocus models) and negative LD created by truncation se-

lection. The reduction in both $V_{G,LE}$ and V_z in the donor population is larger for lower L_e and higher z_1. Its effect on the admixed population is generally largest for parameter values resulting in the highest levels of introgression, i.e. for weak to moderate intensity of selection and high migration rate.

The deviation of the breeding value dis-
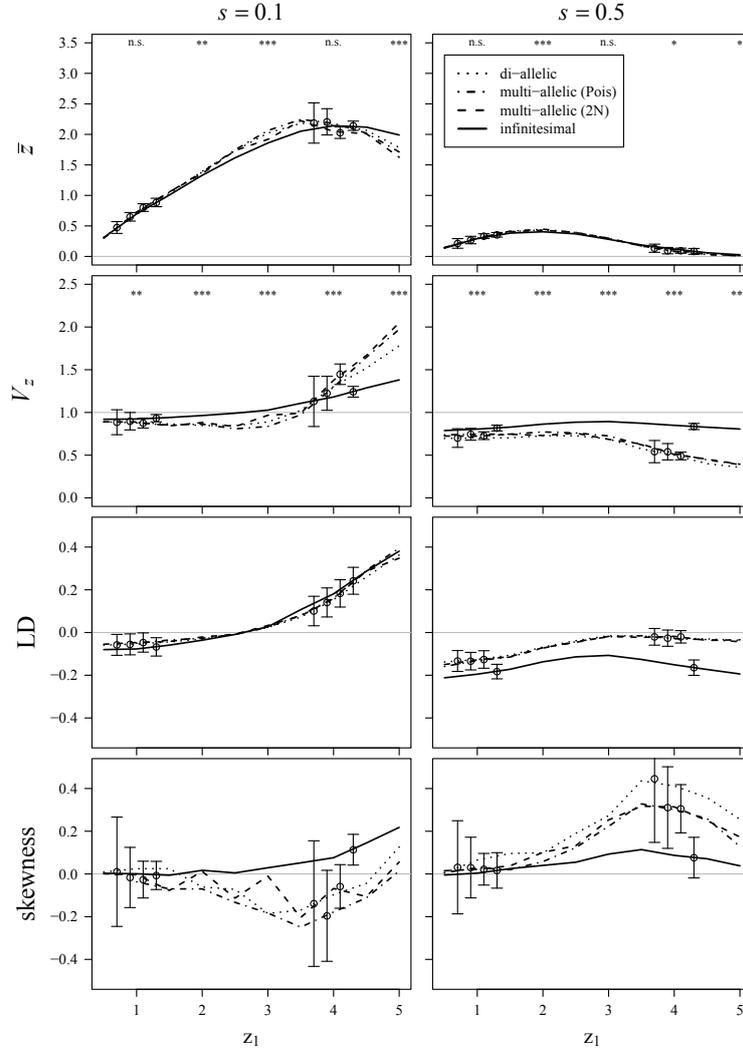
Figure 5: Model differences as function of differences between the populations z_1: Mean breeding value $\bar{z}$, genetic variance V_z and linkage disequilibrium LD for the three different models (see text, for both multilocus models $L_e = 5$) after 50 generations of recurrent immigration with rate $m = 0.2$ of immigrants with mean breeding value $z_1 = 3$ and relatively weak ($s = 0.1$, left) or strong ($s = 0.5$, right) stabilizing selection. Error bars indicate average $\pm$ 1 SE, asterisks indicate significance levels of Welch's t-tests between the infinitesimal model and the least different multilocus model (based on 100 replicate runs). Note different ranges on Y-axes for V_z and LD than in previous figure.

Figure 6: Difference in results between models with $L_e = 5$ (black) and $L_e = 20$ (grey): Mean $\bar{z}$, variance V_z of the breeding value distribution and LD plotted against time for the various models (see legend) un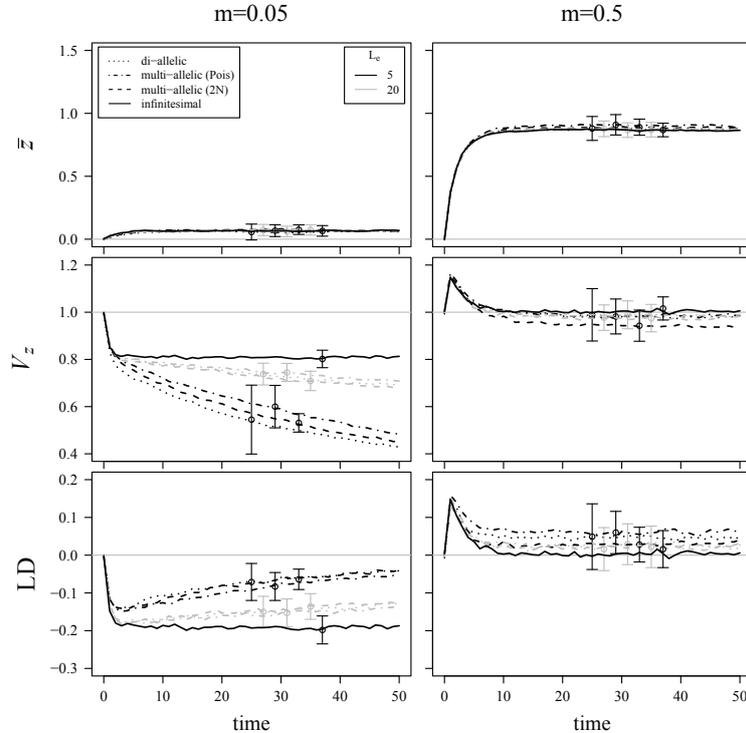der relatively strong stabilizing selection ($s = 0.5$) and a difference between populations of $z_1 = 3$. Error bars indicate average $\pm$ 1 SE based on 100 replicate runs.

tribution from a Gaussian is here depicted by the skew. Under none of the parameter combinations explored was the distribution at $t = 50$, averaged over 100 replicate runs, markedly bimodal (not shown). Individual runs of the diallelic model resulted often in ragged distributions (see Supplementary Information, figure S2), of which modality was hard to judge. Skew is rather small (< 0.2) and positive (distribution 'leaning' to the left) for the infinitesimal model, and does not consistently increase or decrease with any of the parameters (figures 4 and 5, 4th row). For the multilocus models, skew is generally larger but never more extreme than (-)0.4.

Multilocus models: Number of alleles

The number of possible alleles at each locus (2, a Poisson random number or 2N) has only a minor effect on the mean breeding value and LD, but in some cases a considerable effect on the genetic variance and the skewness. These differences in genetic variance V_z are thus due mainly to differences in $V_{G,LE}$. Those are partly caused by the method used to scale the allelic effects, so that the multi-allelic models have on average alleles with smaller effect sizes, and partly due to the larger genetic drift when the total number of alleles is smaller.

The difference between the various multi-
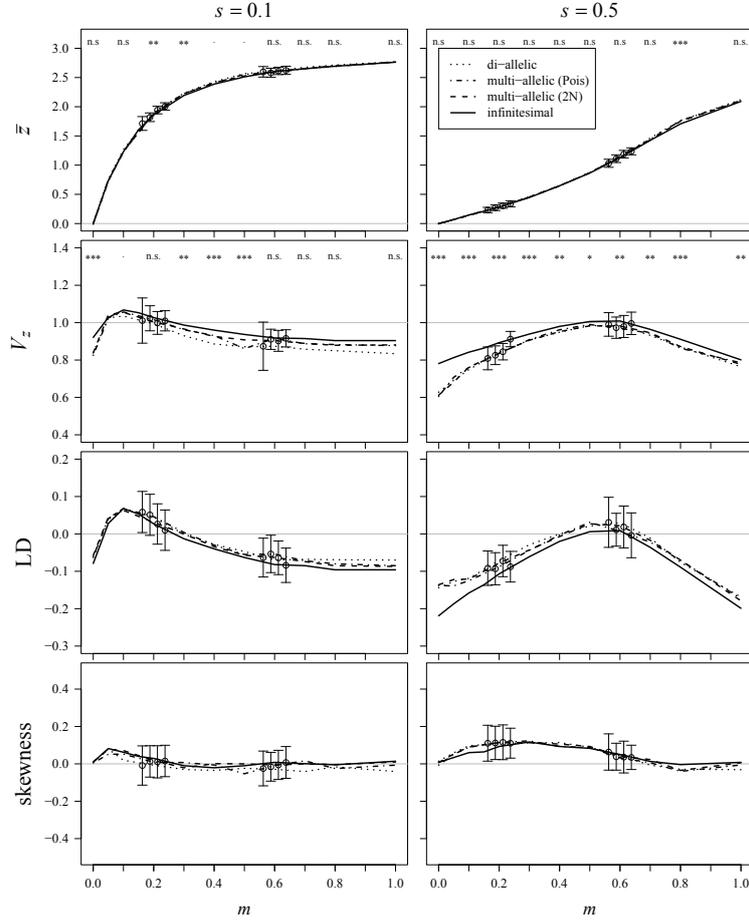
Figure 7: As figure 4, with $L_e = 20$ for the multilocus models instead of $L_e = 5$.

locus models does not increase or decrease with any of the parameters in an obvious, consistent manner. In general, the difference is small when m or z_1 is low, but not necessarily at its smallest. Where the models differ, the model with $2N$ alleles per locus is generally more similar to the infinitesimal model than the diallelic model. The model with a limited, Poisson distributed number of alleles is not always intermediate between the other two models. Considering genetic variance V_z, this model is the most divergent of the three when selection is weak

$(s = 0.1)$ and $z_1 \leq 3$, and the least divergent when selection is stronger $(s = 0.5)$.

For the diallelic models the allelic effects (equal or exponentially distributed) have a small effect on the various characteristics of the breeding value distribution, which is only significant for weak selection and high migration rates ($m \geq 0.5$, Supplementary Information figure S4). Where they differ, the model with equal effects differs more from the infinitesimal model than the model with exponentially distributed effects.

The variation between replicate runs (er-

ror bars in figures 4 and 5) is due to both genetic drift and variation at initiation. It is smallest for the infinitesimal model and model with $2N$ alleles, as for the latter model allele frequencies at initiation are always $1/2N$ and the distribution of allelic effects approaches a continuous distribution.

Number of loci When the effective number of loci is increased from $L_e = 5$ to $L_e = 20$, the results of all multilocus models become more similar to the infinitesimal model. The effective number of loci has a considerable effect on the genetic variance V_z, especially when selection is strong ($s = 0.5$) and migration is limited ($m = 0.05$) (figure 6, left column); the difference in V_z between the same model with $L_e = 5$ versus $L_e = 20$ is larger than the difference between the different models with the same L_e.

The two components that make up V_z, $V_{G,LE}$ and LD, act in opposite directions on this difference. The change in $V_{G,LE}$ is larger for the models with $L_e = 5$ than for those with $L_e = 20$, since allelic effects and therefore strength of selection acting on each allele decrease with an increasing number of loci (due to the scaling method used, ensuring $V_{G,LE} = 1$). The amount of LD, on the other hand, rises with an increasing L_e, not only as a proportion of the total genetic variance, but also in absolute terms (figure 6, bottom left panel; compare right columns figures 4 and 7). This acts to make the difference due to L_e in V_z smaller than the difference in $V_{G,LE}$.

The amount of LD is highly similar for the various models for a given L_e, despite their considerable variation in the actual number of loci; it does not matter whether the trait is coded for by 10-50 diallelic loci or by 6-10 loci with on average 7 alleles each (table 1). This suggests that for LD, as for the variance V_z, L_e is a more relevant parameter than L.

Discussion

We compared the non-Gaussian infinitesimal model to different types of multilocus models, differing in the number of alleles per locus, under a range of migration rates and strong stabilizing selection. We found that all models predicted highly similar deviations of the population mean breeding value from the optimum, suggesting the more parsimonous infinitesimal model is a good approximation for predicting the mean.

When the interest is in the genetic variance however, the infinitesimal model always predicts less change compared to the original situation ($V_z = 1$), which generally implies it predicts a somewhat higher variance than the multilocus models. In the latter models, variance changes not only due to LD, as in the infinitesimal model, but also due to changes in allele frequencies. In addition, in multilocus models the within-family variance may differ between families and across time, while in the infinitesimal model it is constant ($\frac{1}{2}V_{G,LE}$).

As expected, the relative importance of LD and $V_{G,LE}$ in the multilocus models to the total additive genetic variance changed with the effective number of loci (Bulmer, 1971). When the number of loci is large or infinite, changes in variance are mainly due to the generation of LD, and $V_{G,LE}$ stays (nearly) constant. When the number is reduced, allele frequency changes cause $V_{G,LE}$ to change more, and the predicted amount of LD is generally (but not always) smaller. Overall, when the number of loci was decreased, predicted deviation from the optimum slightly increased, observed variance and LD decreased, and skew increased.

The difference between models with different number of alleles per locus, but with the same effective number of loci, was generally smaller than between models of the same type with five versus twenty effective loci. This is in agreement with results by for example Chevalet (1994), who found that "simulation runs involving either several un-

linked loci with many alleles taken from a normal distribution, or several clusters of tightly linked loci with only 2 alleles, lead to very similar responses to directional selection". The multilocus model with an infinite (2N) number of alleles was often more similar to the infinitesimal model than the other two multilocus models, but still most similar to the other multilocus models.

Linkage disequilibrium

In both the infinitesimal and the multilocus models considered here, linkage disequilibrium was incorporated. In many genetic models of selection and migration, it is assumed negligible so that the genetic variance is constant, in order to allow or simplify analytical tractability. While generally only an estimate of the variance at a single point in time will be available (if available at all), this ignores the fact that the variance is a dynamic product of selection and migration. Before migration-selection equilibrium is reached, genetic variance may have been higher than its current level (see figure 3), coinciding with a stronger response to selection. Neglecting this may underestimate the total response to selection, and hence might not result in a good estimate of the equilibrium difference in trait values.

In many systems the difference between local optima will be small and/or the migration rate very limited, so that LD may be ignored and these simpler approximations can be used. Systems in which gene flow is studied, however, will typically be chosen specifically because they have widely different optima, strong selection, and/or high rates of migration (e.g. Moore et al., 2007; Bolnick et al., 2007), such that changing variance should be incorporated in the equations.

Confusingly, models which (implicitly) assume a fixed variance are sometimes referred to as "Gaussian assumption" or "Gaussian approximation" (e.g. Hendry et al., 2001; Lopez et al., 2008; Yeaman and Guillaume, 2009), while "Gaussian" does

not imply this assumption, merely a bell-shaped distribution. This also explains the apparent discrepancy between authors in their judgement of the Gaussian approximation. On the one hand, it was found to work surprisingly well under truncation selection (Turelli and Barton, 1994) and stabilizing selection with one-way migration (Tufto, 2000), even when the distribution departs substantially from normality (Turelli and Barton, 1994). On the other hand, it was found to be "inadequate ... over a large region of the parameter space" under two-way migration (Yeaman and Guillaume, 2009). In the latter case an additional assumption was imposed, known to be violated in this situation (Bulmer, 1971; Tufto, 2000). Actually, if one compares Yeaman and Guillaumes continuum-of-alleles simulation model to the Gaussian approximation that does include changes in variance (Bulmer, 1980, p. 181; see also Tufto, 2000, equations 3 to 6), one can show that results are approximately similar (unpublished data, compared to results in figure 2 in Yeaman and Guillaume (2009)).

When comparing the results presented here to results from models which ignore LD, the most striking difference is in predictions of the deviation of the population mean from its optimum for a range of immigrant means. When ignoring LD, the relationship is linear, with a slope depending on the strength of selection (figure 8, upper panels). When incorporating changes in variance, the deviation of the mean levels off above a certain level of maladaptation of the immigrants, and when additionally allowing for non-normality the deviation even decreases after this point. This agrees with a recent simulation study on supportive breeding of salmon, which included additional parameters such as density dependence, who found the largest drop in population fitness at intermediate difference between wild and hatchery reared salmon (Baskett and Waples, in prep.). A similar result was obtained from simulations
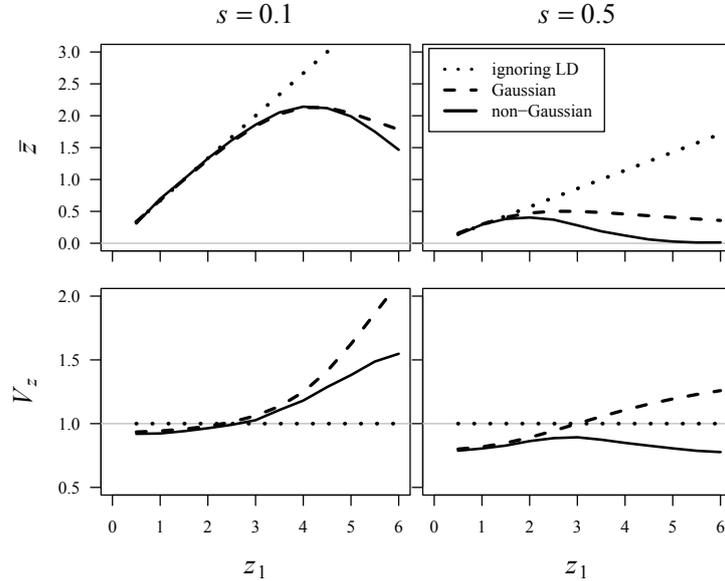
17

Figure 8: Difference between infinitesimal models for infinite population size: Deviation of mean breeding value from optimum ($\bar{z} - z[0]$), divergence between populations at equilibrium ($z[1] - \bar{z}$) and variance V_z at equilibrium when using the non-Gaussian infinitesimal model (solid line), the Gaussian approximation (dashed) and the approximation ignoring linkage disequilibrium (dotted), for a wide range of initial difference between the populations z_1 when $m = 0.2$ and two strengths of stabilizing selection s. (3D plots of $\bar{z}$ vs. m and z_1 in Supplementary Information S6) (equations for those models and detailed comparison in Tufto (2000)).

of a fragmented metapopulation by Lopez et al. (2008), who found that genetic load can increase and then decrease with increasing difference between optima.

Biological interpretation The biological interpretation of this is that as maladaptation of immigrants increases, a smaller fraction of them (and their descendants) successfully reproduces. When their maladaptation exceeds a certain threshold, this reduces their combined effect, despite their larger individual effects on the population mean breeding value. The location of this threshold is dependent on the strength of stabilizing selection. This implies for example that the expected decrease of fitness in the wild of farm salmon across generations, due to ongoing artificial selection, might

make them less harmful to wild salmon populations. Might, as individuals which will not successfully reproduce may still compete for scarce resources with native individuals, lowering their chances of survival (Fleming and Einum, 1997; Houde et al., 2010). In addition, immigrants and especially hybrids may outperform native individuals during some life stages or years, something not considered in many models, including this one.

Non-normality

The importance of the degree of non-normality of the distribution on the relative performance of the models was not investigated in detail. The skew after 50 generations of admixture is given (figures 4 and 5) as an indicator, but the deviation from a

Gaussian is likely to be largest, and most relevant, in the first generations of admixture (example in figure 3), and is not only characterized by skew but also potentially by bimodality and kurtosis. As such, skew after 50 generations explained little of the difference between the infinitesimal and multilocus model.

In contrast, other authors have found skew at mutation-migration-selection equilibrium to correlate with the discrepancy between two models in a situation with symmetric migration and stabilizing selection around local optima (Yeaman and Guillaume, 2009). However, the analytical model to which they compared their continuum-of-alleles simulation did not only assume normality, but also a constant genetic variance (Yeaman and Guillaume, 2009, equation 1). Ignoring the increase in variance due to generation of positive LD by migration underestimates the response to selection (see figure 8). Their analytical model thus overestimates the deviation of the equilibrium means from the optima, or in other words, underestimates the equilibrium divergence between populations. In the results shown here, there is generally a positive correlation between the amount of LD and skew, further strengthening the suspicion that LD caused the observed discrepancy, and skew is merely a confounding factor.

Interestingly, Turelli and Barton (1994) found that greatest steady-state skew is produced with relatively weak truncation selection (but that it was generally small and of negligible biological importance), while Yeaman and Guillaume (2009) found that skew increased with increasing strength of stabilizing selection in their system with two-way migration. Here, with one-way migration, there is no clear correlation between skew and the strength of selection (figures 4 and 5). There does seem to be a trend of increasing skew with increasing deviation of immigrants from the optimum, as was found by Yeaman and Guillaume (2009).

Implications

The difference between results from the non-Gaussian infinitesimal model and multilocus models shown here are (much) smaller than the error margins on parameter estimates in most empirical studies. The same is true for the difference between the results of the Fourier transform and Gaussian approximation (*with* LD) for biologically relevant parameter ranges (Tufto, 2000). It seems that in many practical cases, use of the analytical Gaussian approximation is adequate, and elaborate simulations can be restricted to theoretical studies. Disagreement on this exists however, as some researchers have chosen to ignore LD, rather than to ignore changes in genic variance ($V_{G,LE}$, as all infinitesimal models do), as they found the latter to make up a large part of the total genotypic variance and respond strongly to habitat heterogeneity and dispersal rate (Lopez et al., 2008). In any case, ignoring changes in both $V_{G,LE}$ and LD seems to give unreliable results under a wide range of conditions (see e.g. Tufto, 2000).

Increased complexity of the underlying methods does not necessarily require more input parameters or imply more complicated use. The non-Gaussian and fixed-variance Gaussian infinitesimal models require the same parameters (although some reparametrization may be needed, e.g. $\omega^2 + V_P$ in Hendry et al. (2001) equals $1/s + V_G$ here). Computation time is not a problem either, the infinite population size version of the non-Gaussian infinitesimal model (using discrete Fourier transforms) is very fast, taking less than a second per parameter combination on an ordinary computer. No user friendly interface exists, but "plug and play" R code is available from the authors upon request.

Acknowledgements

References

Araki, H., B. A. Berejikian, M. J. Ford, and M. S. Blouin, 2008. Fitness of hatchery-reared salmonids in the wild. Evol Appl 1:342–355.

Araki, H., B. Cooper, and M. S. Blouin, 2007. Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. Science 318:100–103.

Barton, N. H., 1992. On the spread of new gene combinations in the third phase of Wright's shifting-balance. Evolution 46:551–557.

———, 1999. Clines in polygenic traits. Genet Res 74:223–236.

———, 2001. The role of hybridization in evolution. Mol Ecol 10:551–568.

Barton, N. H. and P. D. Keightley, 2002. Understanding quantitative genetic variation. Nat Rev Genet 3:11–21.

Barton, N. H. and M. Turelli, 1989. Evolutionary quantitative genetics: How little do we know? Annu Rev Genet 23:337–370.

———, 1991. Natural and sexual selection on many loci. Genetics 127:229–255.

Baskett, M. and R. Waples, In prep. Minimizing unintended fitness consequences of cultured individuals on wild populations: keep them similar or make them different?

Beaumont, M., E. M. Barratt, D. Gottelli, A. C. Kitchener, M. J. Daniels, J. K. Pritchard, and M. W. Bruford, 2001. Genetic diversity and introgression in the scottish wildcat. Mol Ecol 10:319–336.

Bodmer, W., 1965. Differential fertility in population genetics models. Genetics 51:411–424.

Bolnick, D., P. Nosil, and M. Servedio, 2007. Natural selection in populations subject to a migration load. Evolution 61:2229–2243.

Bulmer, M. G., 1971. The effect of selection on genetic variability. Am Nat 105:201–211.

———, 1980. The Mathematical Theory of Quantitative Genetics. Clarendon Press, Oxford.

———, 1989. Maintenance of genetic variability by mutation-selection balance: a child's guide through the jungle. Genome 31:761–767.

Chevalet, C., 1994. An approximate theory of selection assuming of quantitative trait loci. Genet Sel Evol 26:379–400.

Dawson, K. K. J., 1997. Linkage disequilibrium and the infinitesimal limit. Theor Popul Biol 52:137–54.

Falconer, D. and T. F. Mackay, 1996. Introduction to quantitative genetics. 4 ed. Longmans Green, Harlow, Essex, UK.

Fisher, R., 1918. The correlation between relatives on the supposition of Mendelian inheritance. T Roy Soc Edin 52:399–433.

Fiske, P., R. Lund, G. Østborg, and L. Fløystad, 2001. Escapees of reared salmon in coastal and riverine fisheries in the period 1989-2000. Tech. Rep. 704, NINA.

Fleming, I. A. and S. Einum, 1997. Experimental tests of genetic divergence of farmed from wild Atlantic salmon due to domestication. ICES J Mar Sci 54:1051–1063.

Fleming, I. A., K. Hindar, I. B. Mjølnerød, B. Jonsson, T. Balstad, and A. Lamberg, 2000. Lifetime success and interactions of farm salmon invading a native population. Proc R Soc B 267:1517–1523.

Frankham, R., 2008. Genetic adaptation to captivity in species conservation programs. Molecular Ecology 17:325–333.

Fraser, D. J., A. L. S. Houde, P. V. Debes, P. O'Reilly, J. D. Eddington, and J. A. Hutchings, 2010. Consequences of farmed-wild hybridization across divergent wild populations and multiple traits in salmon. Ecol Appl 20:935–953.

Gjøen, H. and H. Bentsen, 1997. Past, present, and future of genetic improvement in salmon aquaculture. ICES J Mar Sci 54:1009–1014.

Gross, M., 1998. One species with two biologies: Atlantic salmon (*Salmo salar*) in the wild and in aquaculture. Can J Fish Aquat Sci 55:131–144.

Haldane, J. B. S., 1930. A mathematical theory of natural and artificial selection. part vi, isolation. P Camb Philos Soc 26:220–230.

Hansen, M. M., 2002. Estimating the long-term effects of stocking domesticated trout into wild brown trout (*Salmo trutta*) populations: an approach using microsatellite DNA analysis of historical and contemporary samples. Mol Ecol 11:1003–1015.

Hendry, A., T. Day, and E. Taylor, 2001. Population mixing and the adaptive divergence of quantitative traits in discrete populations: a theoretical framework for empirical tests. Evolution 55:459–466.

Houde, A. L. S., D. J. Fraser, and J. A. Hutchings, 2010. Fitness-related consequences of competitive interactions between farmed and wild Atlantic salmon at different proportional representations of wild–farmed hybrids. ICES J Mar Sci 67:657–667.

Hutchings, J. A. and D. J. Fraser, 2008. The nature of fisheries- and farming-induced evolution. Mol Ecol 17:294–313.

Jonsson, B., 1997. A review of ecological and behavioural interactions between cultured and wild Atlantic salmon. ICES J Mar Sci 54:1031–1039.

Kimura, M. and J. Crow, 1964. The number of alleles that can be maintained in a finite population. Genetics 49:725–738.

Kingsolver, J. G., H. E. Hoekstra, J. M. Hoekstra, D. Berrigan, S. N. Vignieri, C. E. Hill, A. Hoang, P. Gibert, and P. Beerli, 2001. The strength of phenotypic selection in natural populations. Am Nat 157:245–261.

Lande, R., 1976a. The maintenance of genetic variability by mutation in a polygenic character with linked loci. Genetical Research 26:221–235.

———, 1976b. Natural selection and random genetic drift in phenotypic evolution. Evolution 30:314–334.

———, 1981. The minimum number of genes contributing to quantitative variation between and within populations. Genetics 99:541–553.

Lopez, S., F. Rousset, F. H. Shaw, R. G. Shaw, and O. Ronce, 2008. Migration load in plants: role of pollen and seed dispersal in heterogeneous landscapes. J Evolution Biol 21:294–309.

Lynch, M. and B. Walsh, 2008. Evolution and selection of quantitative traits. Available online at http://nitro.biosci.arizona.edu/zbook/NewVolume_2/newvol2.html.

McGinnity, P., P. Prodöhl, A. Ferguson, R. Hynes, N. Ó Maoiléidigh, N. Baker,

D. Cotter, B. O'Hea, D. Cooke, G. Rogan, J. Taggart, and T. Cross, 2003. Fitness reduction and potential extinction of wild populations of Atlantic salmon, *Salmo salar*, as a result of interactions with escaped farm salmon. Proc R Soc B 270:2443–2450.

McGinnity, P., C. Stone, J. B. Taggart, D. Cooke, D. Cotter, R. Hynes, C. Mc-Camley, T. Cross, and A. Ferguson, 1997. Genetic impact of escaped farmed Atlantic salmon (*Salmo salar* L.) on native populations: use of DNA profiling to assess freshwater performance of wild, farmed, and hybrid progeny in a natural river environment. ICES J Mar Sci 54:998–1008.

Merilä, J. and B. C. Sheldon, 1999. Genetic architecture of fitness and nonfitness traits: empirical patterns and development of ideas. Heredity 83:103–109.

Moen, T., M. Baranski, A. Sonesson, and S. Kjøglum, 2009. Confirmation and fine-mapping of a major QTL for resistance to infectious pancreatic necrosis in Atlantic salmon (*Salmo salar*): population-level associations between markers and trait. BMC Genomics 10:368.

Moore, J., J. Gow, E. Taylor, A. Hendry, and M. Peterson, 2007. Quantifying the constraining influence of gene flow on adaptive divergence in the lake-stream threespine stickleback system. Evolution 61:2015–2026.

Naylor, R., K. Hindar, I. A. Fleming, R. Goldburg, S. Williams, J. Volpe, F. Whoriskey, J. Eagle, D. Kelso, and M. Mangel, 2005. Fugitive salmon: Assessing the risks of escaped fish from net-pen aquaculture. Bioscience 55:427–437.

Orr, H., 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. Evolution 52:935–949.

Randi, E., V. Lucchini, M. F. Christensen, N. Mucci, S. M. Funk, G. Dolf, and V. Loeschcke, 2000. Mitochondrial DNA variability in italian and east european wolves: detecting the consequences of small population size and hybridization. Conserv Biol 14:464–473.

Rhymer, J. M. and D. Simberloff, 1996. Extinction by hybridization and introgression. Annu Rev Ecol Syst 27:83–109.

Sundt-Hansen, L., L. Sundström, S. Einum, K. Hindar, I. Fleming, and R. Devlin, 2007. Genetically enhanced growth causes increased mortality in hypoxic environments. Biology Lett 3:165–168.

Tufto, J., 2000. Quantitative genetic models for the balance between migration and stabilizing selection. Genet Res 76:285–293.

———, 2001. Effects of releasing maladapted individuals: A demographic-evolutionary model. Am Nat 158:331–340.

———, 2010. Gene flow from domesticated species to wild relatives: migration load in a model of multivariate selection. Evolution 64:180–192.

Turelli, M. and N. H. Barton, 1994. Genetic and statistical analyses of strong selection on polygenic traits: What, me normal? Genetics 138:913–941.

Via, S. and R. Lande, 1985. Genotype-environment interaction and the evolution of phenotypic plasticity. Evolution 39:505–522.

Walsh, M. R., S. B. Munch, S. Chiba, and D. O. Conover, 2006. Maladaptive changes in multiple traits caused by fishing: impediments to population recovery. Ecol Lett 9:142–148.

Whitton, J., D. E. Wolf, D. M. Arias, A. A. Snow, and L. H. Rieseberg, 1997. The

persistence of cultivar alleles in wild populations of sunflowers five generations after hybridization. Theor Appl Genet 95:33–40.

Wright, S., 1931. Evolution in mendelian populations. Genetics 16:97–159.

Yeaman, S. and F. Guillaume, 2009. Predicting adaptation under migration load: the role of genetic skew. Evolution 63:2926–2938.

Appendix: Initiation of multilocus models

Calculating breeding values

For the diallelic model, individual breeding values z_i are calculated as

$$z_i = \mu + \sum_{l=1}^{L} X_{i,l} a_l$$

where matrix X contains the number of copies of the allele minus one (-1, 0 or 1), with $(X_{.,l} + 1) \sim \mathrm{binom}(2, p_l)$. The allelic effects are scaled such that $V_z = 1$ (see further down), and we define $\mu = -\mathbf{E}\left[\sum_{l=1}^{L} X_{i,l} a_l\right] = -\sum_{l=1}^{L} a_l(2p_l - 1)$ so that $\mathbf{E}[z_i] = z_0 = 0$. The breeding value then becomes

$$z_i = \sum_{l=1}^{L} (X_{i,l} - 2p_l) a_l,$$

For the multi-allelic models, the breeding value is calculated as

$$z_i = \mu + \sum_{l=1}^{L} \sum_{k=1}^{2} A_{i,l,k}, \tag{1}$$

where $A_{i,l,k}$ is the effect size of the allele at locus l on gamete k ($k \in \{1, 2\}$), and μ is again the negative of the expectation,

$$\mu = \sum_{l=1}^{L} \left[\sum_{k=1}^{K_l} a_{l,k} p_{l,k} \right],$$

where K_l is the number of different alleles at locus l, and $a_{l,k}$ is the allelic effect of the kth locus.

For all models, the μ and a in the donor population are equal to those in the recipient population.

Scaling a and calculating L_e

To account for the fact that different allele frequencies and different allelic effects will result in different distribution of relative contributions of each locus to the genetic variance, the effective number of loci L_e is used rather than the actual number L. The two are equal when all allele frequencies p_l and allelic effects a_l are equal, but L_e is lower otherwise.

For the diallelic model, the contribution of each locus to the variance under linkage equilibrium is given by

$$v_l = 2a_l^2 p_l(1 - p_l). \tag{2}$$

To ensure that at initiation the variance of the breeding values equals one, the vector of allelic effects a is scaled,

$$a_l^* = \frac{a_l}{\sqrt{\sum_{l=1}^{L} v_l}}. \tag{3}$$

The effective number of loci is then defined as

$$L_e = \Big(\sum_{l=1}^{L} v_l^{*2} \Big)^{-1}, \tag{4}$$

where v^* are the variance contributions based on the scaled allelic effects.

For the multi-allelic models, the procedure is analogous, with the contribution of a locus l to the variance being

$$v_l = 2 \left[\sum_{k=1}^{K_l} a_{l,k}^2 p_{l,k} - \left(\sum_{k=1}^{K_l} a_{l,k} p_{l,k} \right) \right]. \tag{5}$$

Table 2: Genetic variance V_z and variance under linkage equilibrium $V_{G,LE}$ (Medians over 250 replicates) in the donor population for the various models with 5 or 20 effective loci L_e and for three levels of genetic difference between donor and recipient population z_1 (the total response to selection in the donor population).

Model	L_e	$z_1 = 1$		$z_1 = 3$		$z_1 = 5$	
		V_z	$V_{G,LE}$	V_z	$V_{G,LE}$	V_z	$V_{G,LE}$
diallelic (eq)	5	0.856	0.937	0.613	0.676	0.518	0.568
	20	0.868	0.978	0.796	0.895	0.659	0.736
diallelic (exp)	5	0.832	0.923	0.578	0.642	0.265	0.279
	20	0.866	0.977	0.792	0.901	0.696	0.769
multi-allelic (Pois)	5	0.835	0.921	0.578	0.628	0.251	0.272
	20	0.878	0.979	0.799	0.898	0.676	0.752
multi-allelic (2N)	5	0.854	0.956	0.792	0.884	0.691	0.767
	20	0.872	0.988	0.852	0.968	0.817	0.924
Infinitesimal	N/A	0.882	1	0.876	1	0.878	1

Donor population

Creation of a donor population with a given mean breeding value z_1 is limited by the maximum breeding value, which is obtained when all loci are fixed at the (most) positive allele. This maximum depends on both the number of alleles per locus and the effective number of loci (figure 9). To ensure that replicate runs in which a donor population with a given z_1 could be obtained are not an a-typical subset, and that variance within the donor population was possible, $L_e = 5$ was chosen as the minimum included in the simulations.

Due to directional selection for several generations, variance decreases in the donor population, both due to fixation of alleles (decrease in $V_{G,LE}$) and due to negative LD (Bulmer effect) (table 2).
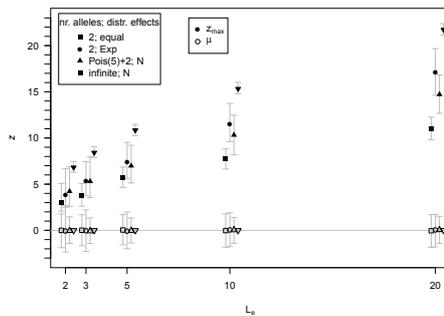
Figure 9: Initial mean breeding value and maximum possible breeding value z_{max} (when fixated for the (most) positive allele at all loci) for the diallelic model with exponential distributed allelic effects (black circles) and the multi-allelic model with normally distributed effects (grey triangles), both with initially U-shaped distribution of allele frequencies.

Supplementary Information
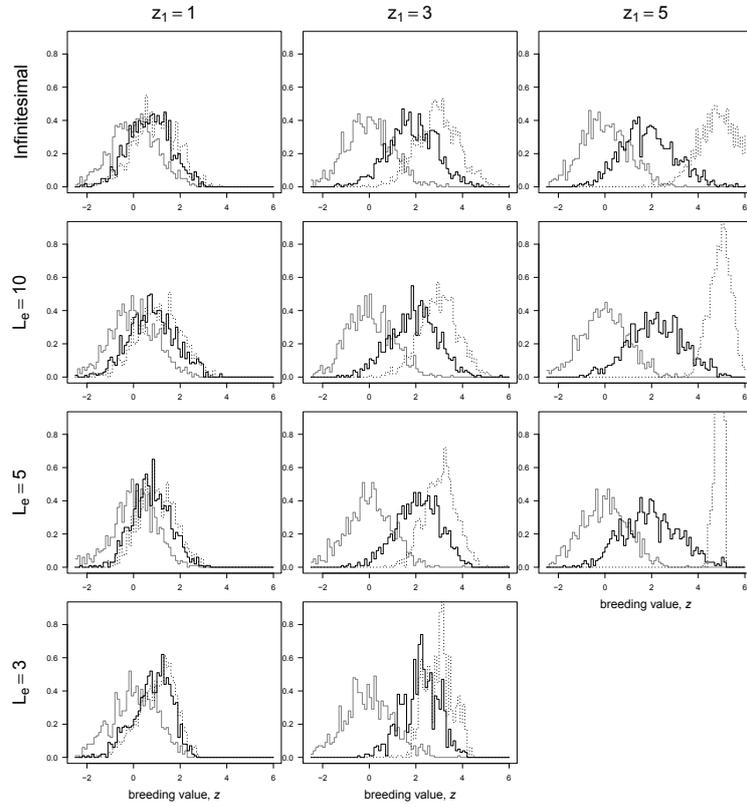
Figure 10: Histogram of breeding values ($N = 1000$), during a single replicate run, for three levels of z_1 (columns) under the infinitesimal model (upper row) and the diallelic multilocus model with three different values for the effective number of loci L_e (rows 2-4). Grey lines represent the distribution in the base population ($t = 0$), dotted in the donor population and black in the admixed population at $t = 25$. For all panels, $m = 0.2$ and $s = 0.1$.
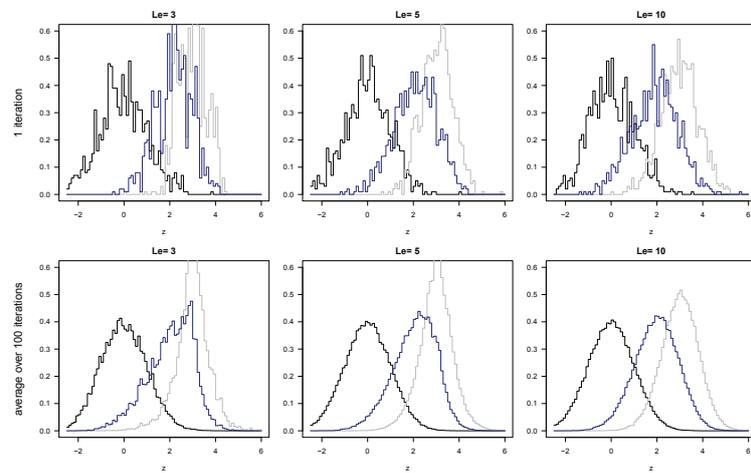
Figure 11: Histogram of breeding values in a single run (upper row) and averaged over 100 replicate runs (lower row), in the base population (grey lines), donor population (dotted) and in the admixed population at $t = 25$ (solid black lines) for $L_e = 3, 5$ and 10. $z_1 = 3$, $s = 0.1$, $m = 0.2$.
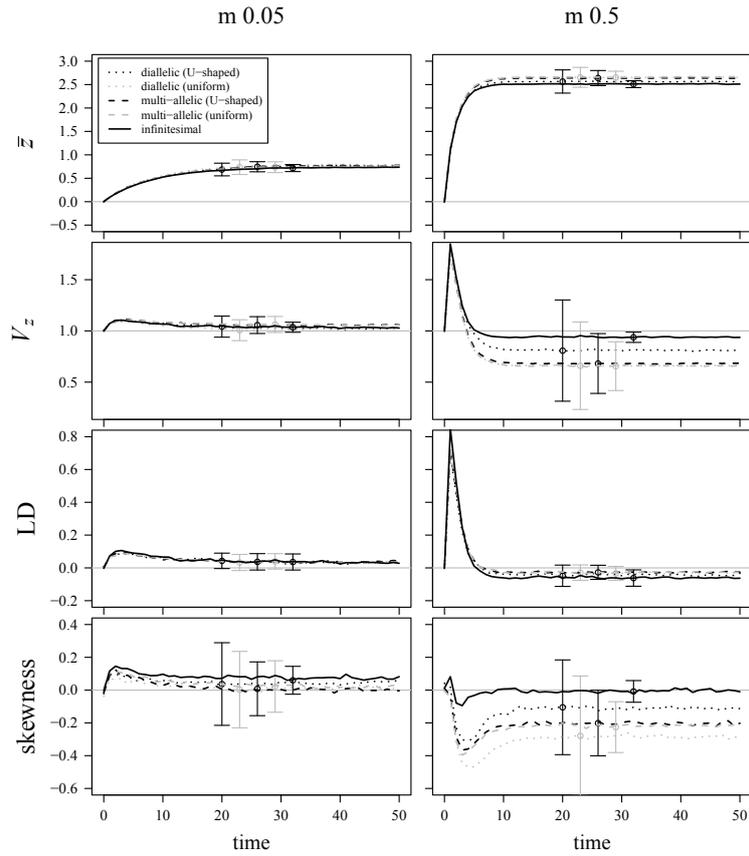
Figure 12: Effect of initial allele frequency distribution: Mean, variance, LD and skewness of the breeding value distribution plotted against time for the diallelic model and the multiallelic model with limited number of alleles, both with $L_e = 5$, when the allele frequencies are initially uniformly distributed (grey) or U-shaped (black). Intensity of selection is $s = 0.1$, difference between the populations $z_1 = 3$. Error bars indicate average $\pm$ 1 SE based on 100 replicate runs.
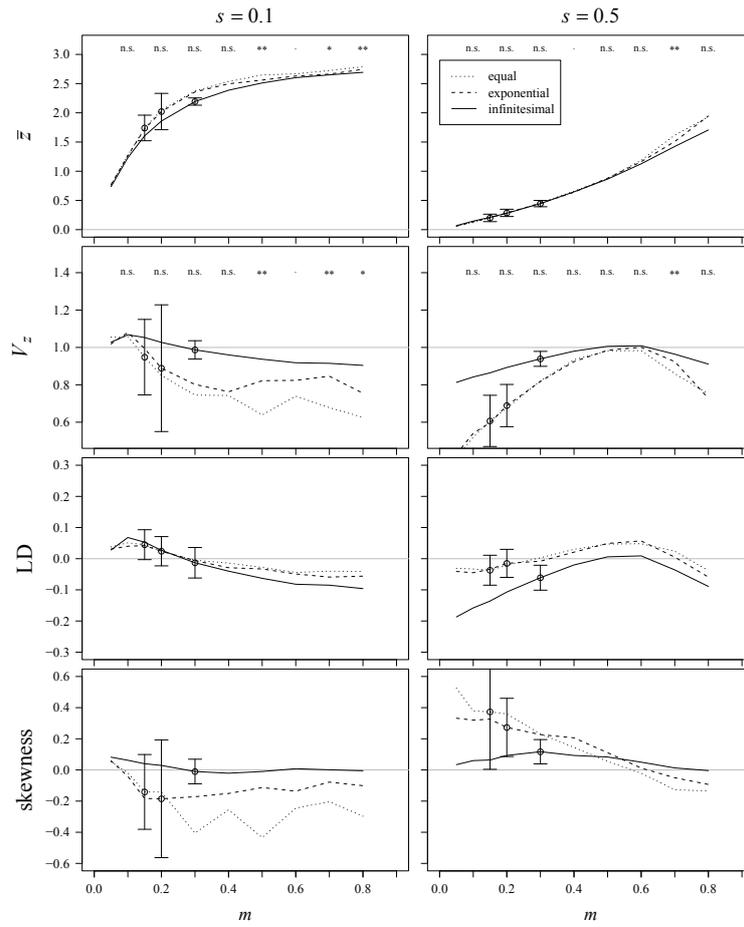
Figure 13: Diallelic models with equal vs. exponentially distributed allelic effect sizes compared for a range of migration rates, $L_e = 5$, $z_1 = 3$, at quasi-equilibrium ($t = 50$). Error bars indicate average $\pm$ 1 SE, asterisks indicate minimum significance levels of Welch's t-tests between the two diallelic models (based on 100 replicate runs).
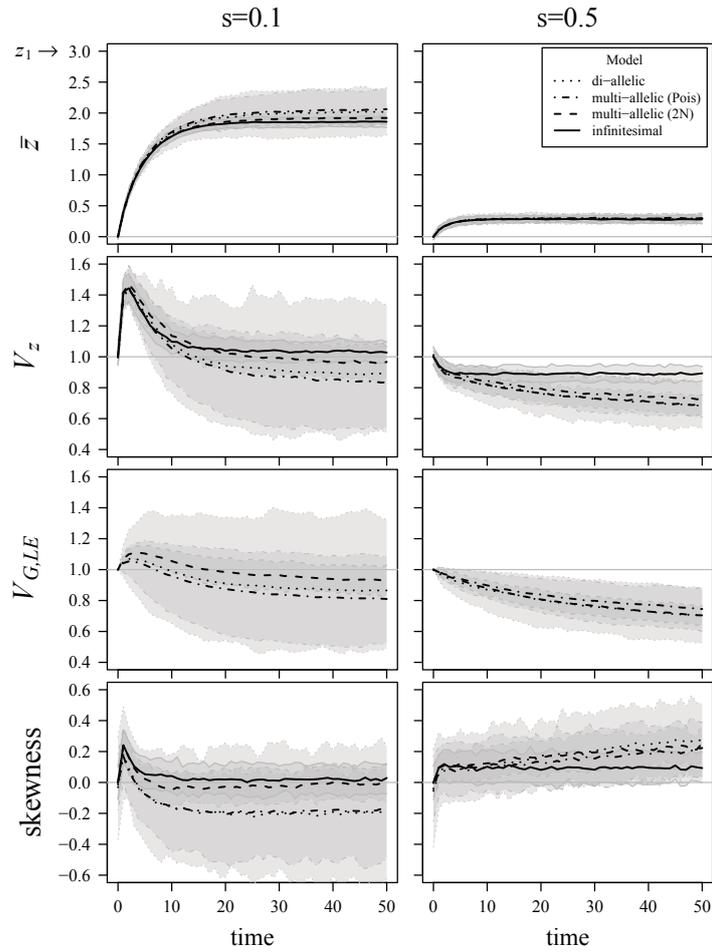
Figure 14: (As figure 3, but with quantiles) Example of change in mean, variance (observed and under linkage equilibrium) and skewness over time, and amount of variation between models and between replicate runs. For each model, the average over 100 replicate runs is shown (black lines, see legend) together with the 10% – 90% quantiles (grey areas, darker grey means more interquantile ranges are overlapping). The shaded areas are delimited by the same line type as the lines representing the averages over all replicates.
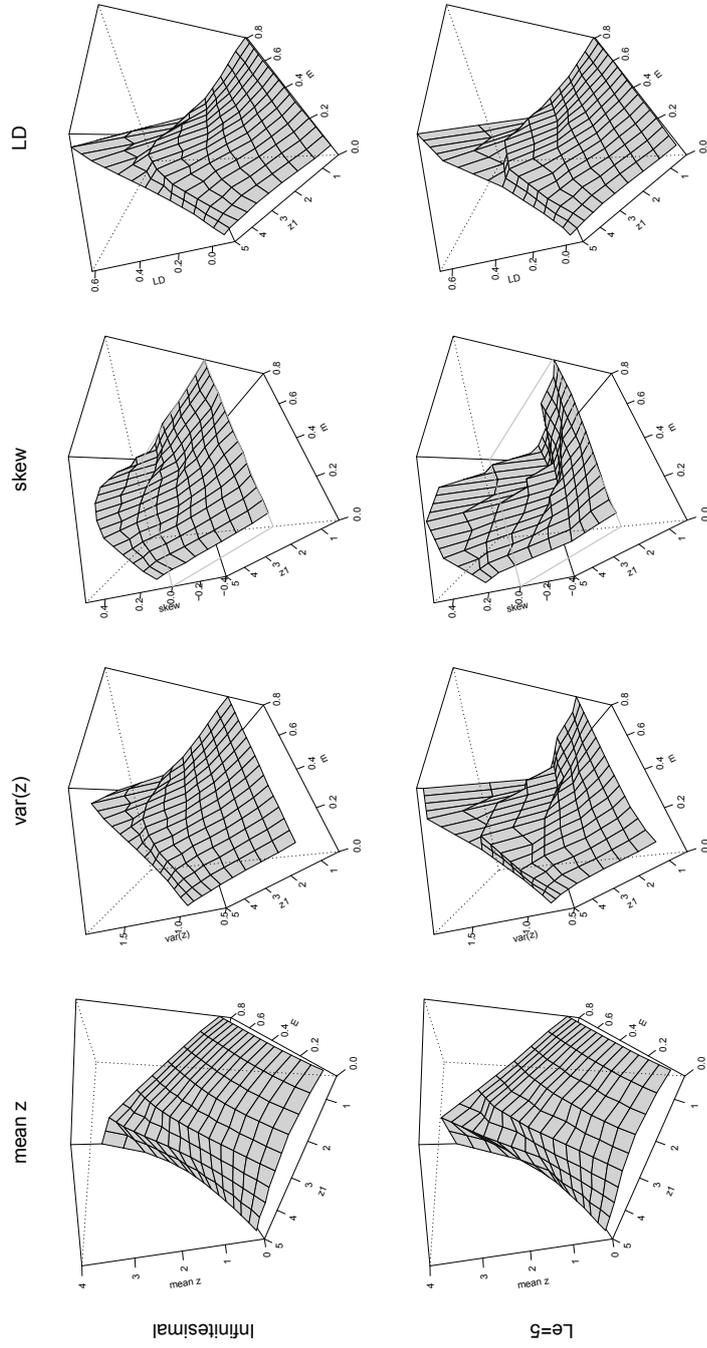
Figure 15: Several properties of the breeding value distribution (average over 100 replicates) after 25 generations of migration with rate m of individuals with average deviation of the optimum $z1$, strength of selection $s = 0.25$, using the infinitesimal model (upper row) or the multilocus model with $L_e = 5$ (lower row).
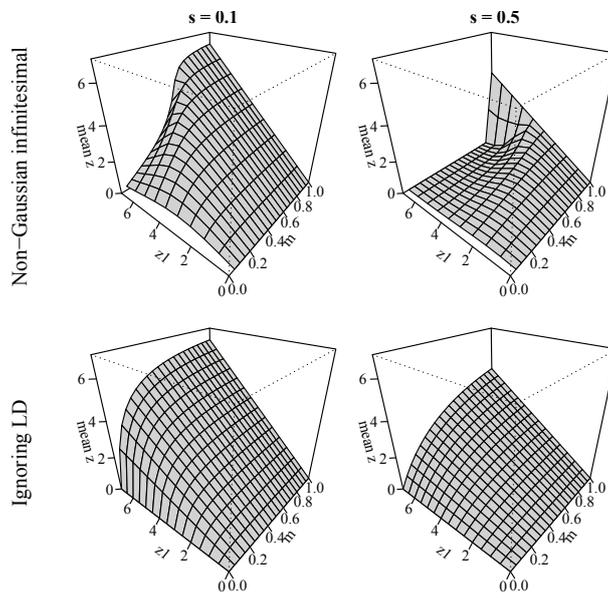
Figure 16: Difference between infinitesimal models: Mean breeding value (mean z) at quasi-equilibrium ($t = 50$) for a wide range of initial differences between the populations ($z1$) and migration rates (m) under weak ($s = 0.1$) and relatively strong ($s = 0.5$) stabilizing selection when using the non-Gaussian infinitesimal model (upper row) and when using the approximation ignoring linkage disequilibrium (lower row).

Paper IV

Doctoral theses in Biology
Norwegian University of Science and Technology
Department of Biology

Year	Name	Degree	Title
1974	Tor-Henning Iversen	Dr. philos Botany	The roles of statholiths, auxin transport, and auxin metabolism in root gravitropism
1978	Tore Slagsvold	Dr. philos Zoology	Breeding events of birds in relation to spring temperature and environmental phenology
1978	Egil Sakshaug	Dr.philos Botany	"The influence of environmental factors on the chemical composition of cultivated and natural populations of marine phytoplankton"
1980	Arnfinn Langeland	Dr. philos Zoology	Interaction between fish and zooplankton populations and their effects on the material utilization in a freshwater lake
1980	Helge Reinertsen	Dr. philos Botany	The effect of lake fertilization on the dynamics and stability of a limnetic ecosystem with special reference to the phytoplankton
1982	Gunn Mari Olsen	Dr. scient Botany	Gravitropism in roots of *Pisum sativum* and *Arabidopsis thaliana*
1982	Dag Dolmen	Dr. philos Zoology	Life aspects of two sympartic species of newts (*Triturus, Amphibia*) in Norway, with special emphasis on their ecological niche segregation
1984	Eivin Røskaft	Dr. philos Zoology	Sociobiological studies of the rook *Corvus frugilegus*
1984	Anne Margrethe Cameron	Dr. scient Botany	Effects of alcohol inhalation on levels of circulating testosterone, follicle stimulating hormone and luteinzing hormone in male mature rats
1984	Asbjørn Magne Nilsen	Dr. scient Botany	Alveolar macrophages from expectorates – Biological monitoring of workers exosed to occupational air pollution. An evaluation of the AM-test
1985	Jarle Mork	Dr. philos Zoology	Biochemical genetic studies in fish
1985	John Solem	Dr. philos Zoology	Taxonomy, distribution and ecology of caddisflies (*Trichoptera*) in the Dovrefjell mountains
1985	Randi E. Reinertsen	Dr. philos Zoology	Energy strategies in the cold: Metabolic and thermoregulatory adaptations in small northern birds
1986	Bernt-Erik Sæther	Dr. philos Zoology	Ecological and evolutionary basis for variation in reproductive traits of some vertebrates: A comparative approach
1986	Torleif Holthe	Dr. philos Zoology	Evolution, systematics, nomenclature, and zoogeography in the polychaete orders *Oweniimorpha* and *Terebellomorpha*, with special reference to the Arctic and Scandinavian fauna
1987	Helene Lampe	Dr. scient Zoology	The function of bird song in mate attraction and territorial defence, and the importance of song repertoires
1987	Olav Hogstad	Dr. philos Zoology	Winter survival strategies of the Willow tit *Parus montanus*
1987	Jarle Inge Holten	Dr. philos Botany	Autecological investigations along a coust-inland transect at Nord-Møre, Central Norway
1987	Rita Kumar	Dr. scient Botany	Somaclonal variation in plants regenerated from cell cultures of *Nicotiana sanderae* and *Chrysanthemum morifolium*

1987 Bjørn Åge Tømmerås	Dr. scient. Zoolog	Olfaction in bark beetle communities: Interspecific interactions in regulation of colonization density, predator - prey relationship and host attraction
1988 Hans Christian Pedersen	Dr. philos Zoology	Reproductive behaviour in willow ptarmigan with special emphasis on territoriality and parental care
1988 Tor G. Heggberget	Dr. philos Zoology	Reproduction in Atlantic Salmon (*Salmo salar*): Aspects of spawning, incubation, early life history and population structure
1988 Marianne V. Nielsen	Dr. scient Zoology	The effects of selected environmental factors on carbon allocation/growth of larval and juvenile mussels (*Mytilus edulis*)
1988 Ole Kristian Berg	Dr. scient Zoology	The formation of landlocked Atlantic salmon (*Salmo salar* L.)
1989 John W. Jensen	Dr. philos Zoology	Crustacean plankton and fish during the first decade of the manmade Nesjø reservoir, with special emphasis on the effects of gill nets and salmonid growth
1989 Helga J. Vivås	Dr. scient Zoology	Theoretical models of activity pattern and optimal foraging: Predictions for the Moose *Alces alces*
1989 Reidar Andersen	Dr. scient Zoology	Interactions between a generalist herbivore, the moose *Alces alces*, and its winter food resources: a study of behavioural variation
1989 Kurt Ingar Draget	Dr. scient Botany	Alginate gel media for plant tissue culture
1990 Bengt Finstad	Dr. scient Zoology	Osmotic and ionic regulation in Atlantic salmon, rainbow trout and Arctic charr: Effect of temperature, salinity and season
1990 Hege Johannesen	Dr. scient Zoology	Respiration and temperature regulation in birds with special emphasis on the oxygen extraction by the lung
1990 Åse Krøkje	Dr. scient Botany	The mutagenic load from air pollution at two work-places with PAH-exposure measured with Ames Salmonella/microsome test
1990 Arne Johan Jensen	Dr. philos Zoology	Effects of water temperature on early life history, juvenile growth and prespawning migrations of Atlantic salmion (*Salmo salar*) and brown trout (*Salmo trutta*): A summary of studies in Norwegian streams
1990 Tor Jørgen Almaas	Dr. scient Zoology	Pheromone reception in moths: Response characteristics of olfactory receptor neurons to intra- and interspecific chemical cues
1990 Magne Husby	Dr. scient Zoology	Breeding strategies in birds: Experiments with the Magpie *Pica pica*
1991 Tor Kvam	Dr. scient Zoology	Population biology of the European lynx (*Lynx lynx*) in Norway
1991 Jan Henning L'Abêe Lund	Dr. philos Zoology	Reproductive biology in freshwater fish, brown trout *Salmo trutta* and roach *Rutilus rutilus* in particular
1991 Asbjørn Moen	Dr. philos Botany	The plant cover of the boreal uplands of Central Norway. I. Vegetation ecology of Sølendet nature reserve; haymaking fens and birch woodlands
1991 Else Marie Løbersli	Dr. scient Botany	Soil acidification and metal uptake in plants
1991 Trond Nordtug	Dr. scient Zoology	Reflctometric studies of photomechanical adaptation in superposition eyes of arthropods
1991 Thyra Solem	Dr. scient Botany	Age, origin and development of blanket mires in Central Norway
1991 Odd Terje Sandlund	Dr. philos Zoology	The dynamics of habitat use in the salmonid genera *Coregonus* and *Salvelinus*: Ontogenic niche shifts and polymorphism
1991 Nina Jonsson	Dr. philos	Aspects of migration and spawning in salmonids

1991 Atle Bones	Dr. scient Botany	Compartmentation and molecular properties of thioglucoside glucohydrolase (myrosinase)
1992 Torgrim Breiehagen	Dr. scient Zoology	Mating behaviour and evolutionary aspects of the breeding system of two bird species: the Temminck's stint and the Pied flycatcher
1992 Anne Kjersti Bakken	Dr. scient Botany	The influence of photoperiod on nitrate assimilation and nitrogen status in timothy (*Phleum pratense* L.)
1992 Tycho Anker-Nilssen	Dr. scient Zoology	Food supply as a determinant of reproduction and population development in Norwegian Puffins *Fratercula arctica*
1992 Bjørn Munro Jenssen	Dr. philos Zoology	Thermoregulation in aquatic birds in air and water: With special emphasis on the effects of crude oil, chemically treated oil and cleaning on the thermal balance of ducks
1992 Arne Vollan Aarset	Dr. philos Zoology	The ecophysiology of under-ice fauna: Osmotic regulation, low temperature tolerance and metabolism in polar crustaceans.
1993 Geir Slupphaug	Dr. scient Botany	Regulation and expression of uracil-DNA glycosylase and O^6-methylguanine-DNA methyltransferase in mammalian cells
1993 Tor Fredrik Næsje	Dr. scient Zoology	Habitat shifts in coregonids.
1993 Yngvar Asbjørn Olsen	Dr. scient Zoology	Cortisol dynamics in Atlantic salmon, *Salmo salar* L.: Basal and stressor-induced variations in plasma levels ans some secondary effects.
1993 Bård Pedersen	Dr. scient Botany	Theoretical studies of life history evolution in modular and clonal organisms
1993 Ole Petter Thangstad	Dr. scient Botany	Molecular studies of myrosinase in Brassicaceae
1993 Thrine L. M. Heggberget	Dr. scient Zoology	Reproductive strategy and feeding ecology of the Eurasian otter *Lutra lutra*.
1993 Kjetil Bevanger	Dr. scient. Zoology	Avian interactions with utility structures, a biological approach.
1993 Kåre Haugan	Dr. scient Bothany	Mutations in the replication control gene trfA of the broad host-range plasmid RK2
1994 Peder Fiske	Dr. scient. Zoology	Sexual selection in the lekking great snipe (*Gallinago media*): Male mating success and female behaviour at the lek
1994 Kjell Inge Reitan	Dr. scient Botany	Nutritional effects of algae in first-feeding of marine fish larvae
1994 Nils Røv	Dr. scient Zoology	Breeding distribution, population status and regulation of breeding numbers in the northeast-Atlantic Great Cormorant *Phalacrocorax carbo carbo*
1994 Annette-Susanne Hoepfner	Dr. scient Botany	Tissue culture techniques in propagation and breeding of Red Raspberry (*Rubus idaeus* L.)
1994 Inga Elise Bruteig	Dr. scient Bothany	Distribution, ecology and biomonitoring studies of epiphytic lichens on conifers
1994 Geir Johnsen	Dr. scient Botany	Light harvesting and utilization in marine phytoplankton: Species-specific and photoadaptive responses
1994 Morten Bakken	Dr. scient Zoology	Infanticidal behaviour and reproductive performance in relation to competition capacity among farmed silver fox vixens, *Vulpes vulpes*
1994 Arne Moksnes	Dr. philos Zoology	Host adaptations towards brood parasitism by the Cockoo
1994 Solveig Bakken	Dr. scient Bothany	Growth and nitrogen status in the moss *Dicranum majus* Sm. as influenced by nitrogen supply
1994 Torbjørn Forseth	Dr. scient Zoology	Bioenergetics in ecological and life history studies of fishes.

1995 Olav Vadstein	Dr. philos Botany	The role of heterotrophic planktonic bacteria in the cycling of phosphorus in lakes: Phosphorus requirement, competitive ability and food web interactions
1995 Hanne Christensen	Dr. scient Zoology	Determinants of Otter *Lutra lutra* distribution in Norway: Effects of harvest, polychlorinated biphenyls (PCBs), human population density and competition with mink *Mustela vision*
1995 Svein Håkon Lorentsen	Dr. scient Zoology	Reproductive effort in the Antarctic Petrel *Thalassoica antarctica*; the effect of parental body size and condition
1995 Chris Jørgen Jensen	Dr. scient Zoology	The surface electromyographic (EMG) amplitude as an estimate of upper trapezius muscle activity
1995 Martha Kold Bakkevig	Dr. scient Zoology	The impact of clothing textiles and construction in a clothing system on thermoregulatory responses, sweat accumulation and heat transport
1995 Vidar Moen	Dr. scient Zoology	Distribution patterns and adaptations to light in newly introduced populations of *Mysis relicta* and constraints on Cladoceran and Char populations
1995 Hans Haavardsholm Blom	Dr. philos Bothany	A revision of the *Schistidium apocarpum* complex in Norway and Sweden
1996 Jorun Skjærmo	Dr. scient Botany	Microbial ecology of early stages of cultivated marine fish; inpact fish-bacterial interactions on growth and survival of larvae
1996 Ola Ugedal	Dr. scient Zoology	Radiocesium turnover in freshwater fishes
1996 Ingibjørg Einarsdottir	Dr. scient Zoology	Production of Atlantic salmon (*Salmo salar*) and Arctic charr (*Salvelinus alpinus*): A study of some physiological and immunological responses to rearing routines
1996 Christina M. S. Pereira	Dr. scient Zoology	Glucose metabolism in salmonids: Dietary effects and hormonal regulation
1996 Jan Fredrik Børseth	Dr. scient Zoology	The sodium energy gradients in muscle cells of *Mytilus edulis* and the effects of organic xenobiotics
1996 Gunnar Henriksen	Dr. scient Zoology	Status of Grey seal *Halichoerus grypus* and Harbour seal *Phoca vitulina* in the Barents sea region
1997 Gunvor Øie	Dr. scient Bothany	Eevalution of rotifer *Brachionus plicatilis* quality in early first feeding of turbot *Scophtalmus maximus* L. larvae
1997 Håkon Holien	Dr. scient Botany	Studies of lichens in spurce forest of Central Norway. Diversity, old growth species and the relationship to site and stand parameters
1997 Ole Reitan	Dr. scient. Zoology	Responses of birds to habitat disturbance due to damming
1997 Jon Arne Grøttum	Dr. scient. Zoology	Physiological effects of reduced water quality on fish in aquaculture
1997 Per Gustav Thingstad	Dr. scient. Zoology	Birds as indicators for studying natural and human-induced variations in the environment, with special emphasis on the suitability of the Pied Flycatcher
1997 Torgeir Nygård	Dr. scient Zoology	Temporal and spatial trends of pollutants in birds in Norway: Birds of prey and Willow Grouse used as Biomonitors
1997 Signe Nybø	Dr. scient. Zoology	Impacts of long-range transported air pollution on birds with particular reference to the dipper *Cinclus cinclus* in southern Norway
1997 Atle Wibe	Dr. scient. Zoology	Identification of conifer volatiles detected by receptor neurons in the pine weevil (*Hylobius abietis*), analysed by gas chromatography linked to electrophysiology and to mass spectrometry
1997 Rolv Lundheim	Dr. scient Zoology	Adaptive and incidental biological ice nucleators

1997	Arild Magne Landa	Dr. scient Zoology	Wolverines in Scandinavia: ecology, sheep depredation and conservation
1997	Kåre Magne Nielsen	Dr. scient Botany	An evolution of possible horizontal gene transfer from plants to sail bacteria by studies of natural transformation in *Acinetobacter calcoacetius*
1997	Jarle Tufto	Dr. scient Zoology	Gene flow and genetic drift in geographically structured populations: Ecological, population genetic, and statistical models
1997	Trygve Hesthagen	Dr. philos Zoology	Population responces of Arctic charr (*Salvelinus alpinus* (L.)) and brown trout (*Salmo trutta* L.) to acidification in Norwegian inland waters
1997	Trygve Sigholt	Dr. philos Zoology	Control of Parr-smolt transformation and seawater tolerance in farmed Atlantic Salmon (*Salmo salar*) Effects of photoperiod, temperature, gradual seawater acclimation, NaCl and betaine in the diet
1997	Jan Østnes	Dr. scient Zoology	Cold sensation in adult and neonate birds
1998	Seethaledsumy Visvalingam	Dr. scient Botany	Influence of environmental factors on myrosinases and myrosinase-binding proteins
1998	Thor Harald Ringsby	Dr. scient Zoology	Variation in space and time: The biology of a House sparrow metapopulation
1998	Erling Johan Solberg	Dr. scient. Zoology	Variation in population dynamics and life history in a Norwegian moose (*Alces alces*) population: consequences of harvesting in a variable environment
1998	Sigurd Mjøen Saastad	Dr. scient Botany	Species delimitation and phylogenetic relationships between the Sphagnum recurvum complex (Bryophyta): genetic variation and phenotypic plasticity
1998	Bjarte Mortensen	Dr. scient Botany	Metabolism of volatile organic chemicals (VOCs) in a head liver S9 vial equilibration system in vitro
1998	Gunnar Austrheim	Dr. scient Botany	Plant biodiversity and land use in subalpine grasslands. – A conservtaion biological approach
1998	Bente Gunnveig Berg	Dr. scient Zoology	Encoding of pheromone information in two related moth species
1999	Kristian Overskaug	Dr. scient Zoology	Behavioural and morphological characteristics in Northern Tawny Owls *Strix aluco*: An intra- and interspecific comparative approach
1999	Hans Kristen Stenøien	Dr. scient Bothany	Genetic studies of evolutionary processes in various populations of nonvascular plants (mosses, liverworts and hornworts)
1999	Trond Arnesen	Dr. scient Botany	Vegetation dynamics following trampling and burning in the outlying haylands at Sølendet, Central Norway
1999	Ingvar Stenberg	Dr. scient Zoology	Habitat selection, reproduction and survival in the White-backed Woodpecker *Dendrocopos leucotos*
1999	Stein Olle Johansen	Dr. scient Botany	A study of driftwood dispersal to the Nordic Seas by dendrochronology and wood anatomical analysis
1999	Trina Falck Galloway	Dr. scient Zoology	Muscle development and growth in early life stages of the Atlantic cod (*Gadus morhua* L.) and Halibut (*Hippoglossus hippoglossus* L.)
1999	Marianne Giæver	Dr. scient Zoology	Population genetic studies in three gadoid species: blue whiting (*Micromisistius poutassou*), haddock (*Melanogrammus aeglefinus*) and cod (*Gradus morhua*) in the North-East Atlantic
1999	Hans Martin Hanslin	Dr. scient Botany	The impact of environmental conditions of density dependent performance in the boreal forest bryophytes *Dicranum majus*, *Hylocomium splendens*, *Plagiochila asplenigides*, *Ptilium crista-castrensis* and *Rhytidiadelphus lokeus*

1999 Ingrid Bysveen Mjølnerød	Dr. scient Zoology	Aspects of population genetics, behaviour and performance of wild and farmed Atlantic salmon (*Salmo salar*) revealed by molecular genetic techniques
1999 Else Berit Skagen	Dr. scient Botany	The early regeneration process in protoplasts from *Brassica napus* hypocotyls cultivated under various g-forces
1999 Stein-Are Sæther	Dr. philos Zoology	Mate choice, competition for mates, and conflicts of interest in the Lekking Great Snipe
1999 Katrine Wangen Rustad	Dr. scient Zoology	Modulation of glutamatergic neurotransmission related to cognitive dysfunctions and Alzheimer's disease
1999 Per Terje Smiseth	Dr. scient Zoology	Social evolution in monogamous families: mate choice and conflicts over parental care in the Bluethroat (*Luscinia s. svecica*)
1999 Gunnbjørn Bremset	Dr. scient Zoology	Young Atlantic salmon (*Salmo salar* L.) and Brown trout (*Salmo trutta* L.) inhabiting the deep pool habitat, with special reference to their habitat use, habitat preferences and competitive interactions
1999 Frode Ødegaard	Dr. scient Zoology	Host spesificity as parameter in estimates of arhrophod species richness
1999 Sonja Andersen	Dr. scient Bothany	Expressional and functional analyses of human, secretory phospholipase A2
2000 Ingrid Salvesen	Dr. scient Botany	Microbial ecology in early stages of marine fish: Development and evaluation of methods for microbial management in intensive larviculture
2000 Ingar Jostein Øien	Dr. scient Zoology	The Cuckoo (*Cuculus canorus*) and its host: adaptions and counteradaptions in a coevolutionary arms race
2000 Pavlos Makridis	Dr. scient Botany	Methods for the microbial econtrol of live food used for the rearing of marine fish larvae
2000 Sigbjørn Stokke	Dr. scient Zoology	Sexual segregation in the African elephant (*Loxodonta africana*)
2000 Odd A. Gulseth	Dr. philos Zoology	Seawater tolerance, migratory behaviour and growth of Charr, (*Salvelinus alpinus*), with emphasis on the high Arctic Dieset charr on Spitsbergen, Svalbard
2000 Pål A. Olsvik	Dr. scient Zoology	Biochemical impacts of Cd, Cu and Zn on brown trout (*Salmo trutta*) in two mining-contaminated rivers in Central Norway
2000 Sigurd Einum	Dr. scient Zoology	Maternal effects in fish: Implications for the evolution of breeding time and egg size
2001 Jan Ove Evjemo	Dr. scient Zoology	Production and nutritional adaptation of the brine shrimp *Artemia* sp. as live food organism for larvae of marine cold water fish species
2001 Olga Hilmo	Dr. scient Botany	Lichen response to environmental changes in the managed boreal forset systems
2001 Ingebrigt Uglem	Dr. scient Zoology	Male dimorphism and reproductive biology in corkwing wrasse (*Symphodus melops* L.)
2001 Bård Gunnar Stokke	Dr. scient Zoology	Coevolutionary adaptations in avian brood parasites and their hosts

Rangifer tarandus platyrhynchus

Castor fiber

Monocytes During Atherosclerosis Development

2002	Terje Thun	Dr.philos Biology	Dendrochronological constructions of Norwegian conifer chronologies providing dating of historical material
2002	Birgit Hafjeld Borgen	Dr. scient Biology	Functional analysis of plant idioblasts (Myrosin cells) and their role in defense, development and growth
2002	Bård Øyvind Solberg	Dr. scient Biology	Effects of climatic change on the growth of dominating tree species along major environmental gradients
2002	Per Winge	Dr. scient Biology	The evolution of small GTP binding proteins in cellular organisms. Studies of RAC GTPases in *Arabidopsis thaliana* and the Ral GTPase from *Drosophila melanogaster*
2002	Henrik Jensen	Dr. scient Biology	Causes and consequenses of individual variation in fitness-related traits in house sparrows
2003	Jens Rohloff	Dr. philos Biology	Cultivation of herbs and medicinal plants in Norway – Essential oil production and quality control
2003	Åsa Maria O. Espmark Wibe	Dr. scient Biology	Behavioural effects of environmental pollution in threespine stickleback *Gasterosteus aculeatur* L.
2003	Dagmar Hagen	Dr. scient Biology	Assisted recovery of disturbed arctic and alpine vegetation – an integrated approach
2003	Bjørn Dahle	Dr. scient Biology	Reproductive strategies in Scandinavian brown bears
2003	Cyril Lebogang Taolo	Dr. scient Biology	Population ecology, seasonal movement and habitat use of the African buffalo (*Syncerus caffer*) in Chobe National Park, Botswana
2003	Marit Stranden	Dr.scient Biology	Olfactory receptor neurones specified for the same odorants in three related Heliothine species (*Helicoverpa armigera, Helicoverpa assulta* and *Heliothis virescens*)
2003	Kristian Hassel	Dr.scient Biology	Life history characteristics and genetic variation in an expanding species, *Pogonatum dentatum*
2003	David Alexander Rae	Dr.scient Biology	Plant- and invertebrate-community responses to species interaction and microclimatic gradients in alpine and Artic environments
2003	Åsa A Borg	Dr.scient Biology	Sex roles and reproductive behaviour in gobies and guppies: a female perspective
2003	Eldar Åsgard Bendiksen	Dr.scient Biology	Environmental effects on lipid nutrition of farmed Atlantic salmon (*Salmo Salar* L.) parr and smolt
2004	Torkild Bakken	Dr.scient Biology	A revision of Nereidinae (Polychaeta, Nereididae)
2004	Ingar Pareliussen	Dr.scient Biology	Natural and Experimental Tree Establishment in a Fragmented Forest, Ambohitantely Forest Reserve, Madagascar
2004	Tore Brembu	Dr.scient Biology	Genetic, molecular and functional studies of RAC GTPases and the WAVE-like regulatory protein complex in *Arabidopsis thaliana*
2004	Liv S. Nilsen	Dr.scient Biology	Coastal heath vegetation on central Norway; recent past, present state and future possibilities
2004	Hanne T. Skiri	Dr.scient Biology	Olfactory coding and olfactory learning of plant odours in heliothine moths. An anatomical, physiological and behavioural study of three related species (*Heliothis virescens, Helicoverpa armigera* and *Helicoverpa assulta*)
2004	Lene Østby	Dr.scient Biology	Cytochrome P4501A (CYP1A) induction and DNA adducts as biomarkers for organic pollution in the natural environment
2004	Emmanuel J. Gerreta	Dr. philos Biology	The Importance of Water Quality and Quantity in the Tropical Ecosystems, Tanzania

2004 Linda Dalen	Dr.scient Biology	Dynamics of Mountain Birch Treelines in the Scandes Mountain Chain, and Effects of Climate Warming
2004 Lisbeth Mehli	Dr.scient Biology	Polygalacturonase-inhibiting protein (PGIP) in cultivated strawberry (*Fragaria* x *ananassa*): characterisation and induction of the gene following fruit infection by *Botrytis cinerea*
2004 Børge Moe	Dr.scient Biology	Energy-Allocation in Avian Nestlings Facing Short-Term Food Shortage
2005 Matilde Skogen Chauton	Dr.scient Biology	Metabolic profiling and species discrimination from High-Resolution Magic Angle Spinning NMR analysis of whole-cell samples
2005 Sten Karlsson	Dr.scient Biology	Dynamics of Genetic Polymorphisms
2005 Terje Bongard	Dr.scient Biology	Life History strategies, mate choice, and parental investment among Norwegians over a 300-year period
2005 Tonette Røstelien	ph.d Biology	Functional characterisation of olfactory receptor neurone types in heliothine moths
2005 Erlend Kristiansen	Dr.scient Biology	Studies on antifreeze proteins
2005 Eugen G. Sørmo	Dr.scient Biology	Organochlorine pollutants in grey seal (*Halichoerus grypus*) pups and their impact on plasma thyrid hormone and vitamin A concentrations
2005 Christian Westad	Dr.scient Biology	Motor control of the upper trapezius
2005 Lasse Mork Olsen	ph.d Biology	Interactions between marine osmo- and phagotrophs in different physicochemical environments
2005 Åslaug Viken	ph.d Biology	Implications of mate choice for the management of small populations
2005 Ariaya Hymete Sahle Dingle	ph.d Biology	Investigation of the biological activities and chemical constituents of selected *Echinops* spp. growing in Ethiopia
2005 Anders Gravbrøt Finstad	ph.d Biology	Salmonid fishes in a changing climate: The winter challenge
2005 Shimane Washington Makabu	ph.d Biology	Interactions between woody plants, elephants and other browsers in the Chobe Riverfront, Botswana
2005 Kjartan Østbye	Dr.scient Biology	The European whitefish *Coregonus lavaretus* (L.) species complex: historical contingency and adaptive radiation
2006 Kari Mette Murvoll	ph.d Biology	Levels and effects of persistent organic pollutans (POPs) in seabirds Retinoids and α-tocopherol – potential biomakers of POPs in birds?
2006 Ivar Herfindal	Dr.scient Biology	Life history consequences of environmental variation along ecological gradients in northern ungulates
2006 Nils Egil Tokle	ph.d Biology	Are the ubiquitous marine copepods limited by food or predation? Experimental and field-based studies with main focus on *Calanus finmarchicus*
2006 Jan Ove Gjershaug	Dr.philos Biology	Taxonomy and conservation status of some booted eagles in south-east Asia
2006 Jon Kristian Skei	Dr.scient Biology	Conservation biology and acidification problems in the breeding habitat of amphibians in Norway
2006 Johanna Järnegren	ph.d Biology	Acesta Oophaga and Acesta Excavata – a study of hidden biodiversity
2006 Bjørn Henrik Hansen	ph.d Biology	Metal-mediated oxidative stress responses in brown trout (*Salmo trutta*) from mining contaminated rivers in Central Norway

2006 Vidar Grøtan	ph.d Biology	Temporal and spatial effects of climate fluctuations on population dynamics of vertebrates
2006 Jafari R Kideghesho	ph.d Biology	Wildlife conservation and local land use conflicts in western Serengeti, Corridor Tanzania
2006 Anna Maria Billing	ph.d Biology	Reproductive decisions in the sex role reversed pipefish *Syngnathus typhle*: when and how to invest in reproduction
2006 Henrik Pärn	ph.d Biology	Female ornaments and reproductive biology in the bluethroat
2006 Anders J. Fjellheim	ph.d Biology	Selection and administration of probiotic bacteria to marine fish larvae
2006 P. Andreas Svensson	ph.d Biology	Female coloration, egg carotenoids and reproductive success: gobies as a model system
2007 Sindre A. Pedersen	ph.d Biology	Metal binding proteins and antifreeze proteins in the beetle *Tenebrio molitor* - a study on possible competition for the semi-essential amino acid cysteine
2007 Kasper Hancke	ph.d Biology	Photosynthetic responses as a function of light and temperature: Field and laboratory studies on marine microalgae
2007 Tomas Holmern	ph.d Biology	Bushmeat hunting in the western Serengeti: Implications for community-based conservation
2007 Kari Jørgensen	ph.d Biology	Functional tracing of gustatory receptor neurons in the CNS and chemosensory learning in the moth *Heliothis virescens*
2007 Stig Ulland	ph.d Biology	Functional Characterisation of Olfactory Receptor Neurons in the Cabbage Moth, (*Mamestra brassicae* L.) (Lepidoptera, Noctuidae). Gas Chromatography Linked to Single Cell Recordings and Mass Spectrometry
2007 Snorre Henriksen	ph.d Biology	Spatial and temporal variation in herbivore resources at northern latitudes
2007 Roelof Frans May	ph.d Biology	Spatial Ecology of Wolverines in Scandinavia
2007 Vedasto Gabriel Ndibalema	ph.d Biology	Demographic variation, distribution and habitat use between wildebeest sub-populations in the Serengeti National Park, Tanzania
2007 Julius William Nyahongo	ph.d Biology	Depredation of Livestock by wild Carnivores and Illegal Utilization of Natural Resources by Humans in the Western Serengeti, Tanzania
2007 Shombe Ntaraluka Hassan	ph.d Biology	Effects of fire on large herbivores and their forage resources in Serengeti, Tanzania
2007 Per-Arvid Wold	ph.d Biology	Functional development and response to dietary treatment in larval Atlantic cod *(Gadus morhua* L.) Focus on formulated diets and early weaning
2007 Anne Skjetne Mortensen	ph.d Biology	Toxicogenomics of Aryl Hydrocarbon- and Estrogen Receptor Interactions in Fish: Mechanisms and Profiling of Gene Expression Patterns in Chemical Mixture Exposure Scenarios
2008 Brage Bremset Hansen	ph.d Biology	The Svalbard reindeer (*Rangifer tarandus platyrhynchus*) and its food base: plant-herbivore interactions in a high-arctic ecosystem
2008 Jiska van Dijk	ph.d Biology	Wolverine foraging strategies in a multiple-use landscape
2008 Flora John Magige	ph.d Biology	The ecology and behaviour of the Masai Ostrich (Struthio camelus massaicus) in the Serengeti Ecosystem, Tanzania

2008	Bernt Rønning	ph.d Biology	Sources of inter- and intra-individual variation in basal metabolic rate in the zebra finch, (*Taeniopygia guttata*)
2008	Sølvi Wehn	ph.d Biology	Biodiversity dynamics in semi-natural mountain landscapes. - A study of consequences of changed agricultural practices in Eastern Jotunheimen
2008	Trond Moxness Kortner	ph.d Biology	"The Role of Androgens on previtellogenic oocyte growth in Atlantic cod (*Gadus morhua*): Identification and patterns of differentially expressed genes in relation to Stereological Evaluations"
2008	Katarina Mariann Jørgensen	Dr.Scient Biology	The role of platelet activating factor in activation of growth arrested keratinocytes and re-epithelialisation
2008	Tommy Jørstad	ph.d Biology	Statistical Modelling of Gene Expression Data
2008	Anna Kusnierczyk	ph.d Bilogy	*Arabidopsis thaliana* Responses to Aphid Infestation
2008	Jussi Evertsen	ph.d Biology	Herbivore sacoglossans with photosynthetic chloroplasts
2008	John Eilif Hermansen	ph.d Biology	Mediating ecological interests between locals and globals by means of indicators. A study attributed to the asymmetry between stakeholders of tropical forest at Mt. Kilimanjaro, Tanzania
2008	Ragnhild Lyngved	ph.d Biology	Somatic embryogenesis in *Cyclamen persicum.* Biological investigations and educational aspects of cloning
2008	Line Elisabeth Sundt-Hansen	ph.d Biology	Cost of rapid growth in salmonid fishes
2008	Line Johansen	ph.d Biology	Exploring factors underlying fluctuations in white clover populations – clonal growth, population structure and spatial distribution
2009	Astrid Jullumstrø Feuerherm	ph.d Biology	Elucidation of molecular mechanisms for pro-inflammatory phospholipase A2 in chronic disease
2009	Pål Kvello	ph.d Biology	Neurons forming the network involved in gustatory coding and learning in the moth *Heliothis virescens:* Physiological and morphological characterisation, and integration into a standard brain atlas
2009	Trygve Devold Kjellsen	ph.d Biology	Extreme Frost Tolerance in Boreal Conifers
2009	Johan Reinert Vikan	ph.d Biology	Coevolutionary interactions between common cuckoos *Cuculus canorus* and *Fringilla* finches
2009	Zsolt Volent	ph.d Biology	Remote sensing of marine environment: Applied surveillance with focus on optical properties of phytoplankton, coloured organic matter and suspended matter
2009	Lester Rocha	ph.d Biology	Functional responses of perennial grasses to simulated grazing and resource availability
2009	Dennis Ikanda	ph.d Biology	Dimensions of a Human-lion conflict: Ecology of human predation and persecution of African lions (*Panthera leo*) in Tanzania
2010	Huy Quang Nguyen	ph.d Biology	Egg characteristics and development of larval digestive function of cobia (*Rachycentron canadum*) in response to dietary treatments -Focus on formulated diets

2010 Eli Kvingedal	ph.d Biology	Intraspecific competition in stream salmonids: the impact of environment and phenotype
2010 Sverre Lundemo	ph.d Biology	Molecular studies of genetic structuring and demography in *Arabidopsis* from Northern Europe
2010 Iddi Mihijai Mfunda	ph.d Biology	Wildlife Conservation and People's livelihoods: Lessons Learnt and Considerations for Improvements. Tha Case of Serengeti Ecosystem, Tanzania
2010 Anton Tinchov Antonov	ph.d Biology	Why do cuckoos lay strong-shelled eggs? Tests of the puncture resistance hypothesis
2010 Anders Lyngstad	ph.d Biology	Population Ecology of *Eriophorum latifolium*, a Clonal Species in Rich Fen Vegetation
2010 Hilde Færevik	ph.d Biology	Impact of protective clothing on thermal and cognitive responses
2010 Ingerid Brænne Arbo	ph.d Medical technology	Nutritional lifestyle changes – effects of dietary carbohydrate restriction in healthy obese and overweight humans
2010 Yngvild Vindenes	ph.d Biology	Stochastic modeling of finite populations with individual heterogeneity in vital parameters
2010 Hans-Richard Brattbakk	ph.d Medical technology	The effect of macronutrient composition, insulin stimulation, and genetic variation on leukocyte gene expression and possible health benefits
2011 Geir Hysing Bolstad	ph.d Biology	Evolution of Signals: Genetic Architecture, Natural Selection and Adaptive Accuracy
2011 Karen de Jong	ph.d Biology	Operational sex ratio and reproductive behaviour in the two-spotted goby (*Gobiusculus flavescens*)
2011 Ann-Iren Kittang	ph.d Biology	*Arabidopsis thaliana* L. adaptation mechanisms to microgravity through the EMCS MULTIGEN-2 experiment on the ISS:– The science of space experiment integration and adaptation to simulated microgravity
2011 Aline Magdalena Lee	ph.d Biology	Stochastic modeling of mating systems and their effect on population dynamics and genetics
2011 Christopher Gravningen Sørmo	ph.d Biology	Rho GTPases in Plants: Structural analysis of ROP GTPases; genetic and functional studies of MIRO GTPases in *Arabidopsis thaliana*
2011 Grethe Robertsen	ph.d Biology	Relative performance of salmonid phenotypes across environments and competitive intensities
2011 Line-Kristin Larsen	ph.d Biology	Life-history trait dynamics in experimental populations of guppy (*Poecilia reticulata*): the role of breeding regime and captive environment
2011 Maxim A. K. Teichert	ph.d Biology	Regulation in Atlantic salmon (*Salmo salar*): The interaction between habitat and density
2011 Torunn Beate Hancke	ph.d Biology	Use of Pulse Amplitude Modulated (PAM) Fluorescence and Bio-optics for Assessing Microalgal Photosynthesis and Physiology
2011 Sajeda Begum	ph.d Biology	Brood Parasitism in Asian Cuckoos: Different Aspects of Interactions between Cuckoos and their Hosts in Bangladesh
2011 Kari J. K. Attramadal	ph.d Biology	Water treatment as an approach to increase microbial control in the culture of cold water marine larvae
2011 Camilla Kalvatn Egset	ph.d Biology	The Evolvability of Static Allometry: A Case Study
2011 AHM Raihan Sarker	ph.d Biology	Conflict over the conservation of the Asian elephant (*Elephas maximus*) in Bangladesh

2011	Gro Dehli Villanger	ph.d Biology	Effects of complex organohalogen contaminant mixtures on thyroid hormone homeostasis in selected arctic marine mammals
2011	Kari Bjørneraas	ph.d Biology	Spatiotemporal variation in resource utilisation by a large herbivore, the moose
2011	John Odden	ph.d Biology	The ecology of a conflict: Eurasian lynx depredation on domestic sheep
2011	Simen Pedersen	ph.d Biology	Effects of native and introduced cervids on small mammals and birds
2011	Mohsen Falahati-Anbaran	ph.d Biology	Evolutionary consequences of seed banks and seed dispersal in *Arabidopsis*
2012	Jakob Hønborg Hansen	ph.d Biology	Shift work in the offshore vessel fleet: circadian rhythms and cognitive performance
2012	Aleksander Handå	ph.d Biology	Cultivation of mussels (*Mytilus edulis*):Feed requirements, storage and integration with salmon (*Salmo salar*) farming