Animal Breeding and Genomics MSc thesis document

02

Quantitative genetics

08

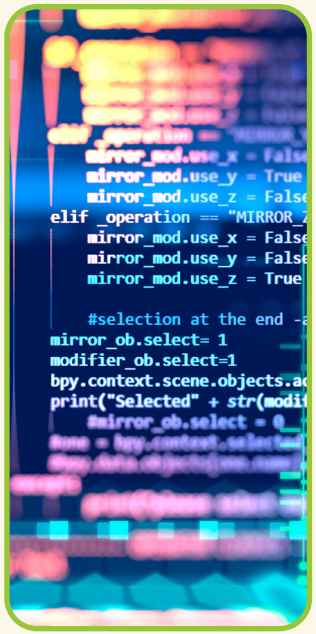
Genomics

19

Genetic diversity

25

Sustainability and breeding programmes



31

Big data and data sciences

34

Alternative options and contact information

Quantitative genetics

Selection history models in animal breeding

With the continuous shift in breeding goals towards more balanced breeding, it may become interesting to add new traits to the breeding goal. Before a trait can be included in the breeding goal, knowledge of the (indirect) selection history of these traits would be insightful. This information on the selection history of new traits can indicate whether there are strong genetic correlations with other breeding goal traits, which is important to consider before the trait is included in the breeding goal. Two models ($\hat{G}$ and s-value estimation) have been developed that are able to assess the selection history of traits. We have evaluated the performance of these models in a simulated pig breeding population that was undergoing phenotypic selection with varying heritabilities (0.05, 0.1, 0.3) and 30 generations of ongoing selection. Both models were able to detect selection, but their performance varied for the different heritabilities. Furthermore, the selection interval influenced performance of the s-value method. This is useful information, however, we were not able to study the impact of different parameters on the detection of selection, using the existing simulated dataset. Options are to investigate the impact of the number of genotyped animals, the number of genotyped variants (i.e. SNPs), or different types of selection. In this MSc thesis, you will develop programming skills, learn high-performing computing and work with large simulated data sets.

Supervisor(s): Yvonne Wientjes, Anne Jansen

Pure or mixed? Investigating signals of hybridisation in Dutch mussel species

In this project you will explore whether Dutch mussel populations show signs of hybridisation between the native blue mussel (Mytilus edulis) and the non-invasive Mediterranean mussel (Mytilus galloprovincialis). Using existing genetic data and a subset of diagnostic single nucleotide polymorphisms (SNPs) you will screen mussel samples for genetic signatures of mixing between the two species. The project will focus on comparing allele frequencies, identifying potential hybrid individuals, and mapping where hybridisation is most likely to occur. You will need basic experience in understanding genomic data. The project will help assess whether non-native mussels are genetically influencing native populations and provide useful information for conservation and aquaculture management in the Netherlands.

Supervisor(s): Chris Barratt, Sophie Valk

Estimating heritability of piling behaviour in laying hens using video-recorded phenotypes

Recurrent piling is the sudden aggregation, with no apparent trigger, of laying hens in densities larger than expected. This type of behaviour can result in injuries or in the death by smothering of the individuals. Hence, piling events impact both economic gains and animal welfare, raising the interest in reducing the frequency of occurrences. Winter et al. (2021), found a correlation between age and piling frequency in brown layers, while piling events where equally frequent in white layers regardless of their age. This study used videos to document piling events and the number of individuals participating in them. However, no individual identification was available. Until this day, no estimations of heritability for recurrent piling in laying hens are available. The main reason is that recording of piling is very challenging, because it occurs frequently and is therefore difficult to record by direct human observation.

In the IMAGEN project, video recordings of Dekalb White hens in 12 pens were obtained, and all piling events were automatically registered with computer vision technology. The individuals initialising each event, the immediate followers, and the day time when the events occurred were documented in a dataset. This dataset is the first of its kind, and has no known equivalent. Moreover, all birds are genotyped, so that a genetic analysis can be done.

The objective of this study is to perform a genetic analysis of piling behaviour, and to obtain estimate of variance components and heritability of piling behaviour. You will be in charge of fitting the best possible model for variance components estimation, while learning how to handle real field data. This thesis project will be executed in collaboration with the breeding company Hendrix Genetics.

Supervisor(s): Piter Bijma

The long-term effects of multi-trait selection on genetic correlations between traits

In commercial breeding programmes, selection takes place on multiple traits using an index of (genomic) breeding values. In this index, the genetic variances of the traits and the genetic correlations between traits are often considered to be constant. However, as a result of selection, those variance components are changing over time. Those changes are likely to be stronger and more unpredictable for the genetic correlation than for the genetic variances. Use of incorrect variance components in a selection index can lead to suboptimal breeding programme designs and failure to achieve maximum selection response. Moreover, incorrect variances can lead to incorrectly predicted selection responses.

Using an existing simulation programme, you will investigate the impact of different selection methods on the genetic correlation between traits and how this is related to the value of the genetic correlation. This will increase our understanding about the changes in the genetic correlation under selection and on how often breeding programmes should update the (co)variance components for their breeding value estimation. You will learn to use different multi-trait selection strategies, such as pedigree and genomic selection, and to investigate the long-term effects of selection.

Supervisor(s): Yvonne Wientjes

Validation of deterministic equations for predicting the accuracy of international genomic beef cattle evaluations

Genomic selection has revolutionised animal breeding, and genomic evaluations are increasingly becoming adopted in beef cattle. Although genotyping is becoming cheaper, accurate genomic predictions often require large and representative reference populations, which can be expensive and time-consuming to build and maintain, especially for small populations and difficult-to-measure traits.

International genomic evaluations are one of the possible ways to improve national-level genomic predictions. These evaluations pool and combine phenotypic, pedigree and genomic data from different countries into a single genomic evaluation allowing to build larger reference populations. However, not all participating countries can contribute the same amount of genomic data to the international evaluations. Thus, countries must evaluate the benefits and trade-offs of their participation in international genomic evaluations upfront. Deterministic equations, i.e. equations that use population specific parameters, can be used to predict the expected gain in the accuracy of genomic predictions when combining data from different sources of information, countries in this case.

In this project you will use a simulated dataset consisting of pedigree, phenotypic and genomic information mimicking two beef cattle populations exchanging different amounts of genomic data and at different levels of genetic connectedness between them. After computing the required parameters for the prediction equations, you will validate them using the true simulated breeding values. You can also compare these results with commonly-used validation methods. In this project you will develop your programming skills and learn how to use genomic data in the context of (inter)national cattle evaluations.

Supervisor(s): Renzo Bonifazi

The impact of data structure and sires exchange on estimating (direct-maternal) genetic correlations

In livestock, recorded traits can be influenced in their expression by the mother. For example, the weaning weight of a calf is affected by the milk

production of the cow. For maternally affected traits, genetic evaluations require modelling of the so-called direct-maternal genetic correlation (rdm), i.e. the correlation between the direct and the maternal genetic effects. Different sources of information are needed to estimate genetic correlations of maternally affected traits compared to conventional traits.

In beef cattle, estimates of rdm are often reported to be negative. However, some studies underline that rdm could be affected by different sources of bias due to, for example, having an incomplete data structure during the estimation process. International beef cattle evaluations aim to assess animals' estimated breeding values (EBVs) in different countries. In the context of international evaluations, genetic correlations within and between countries are crucial to define how the information from one country contributes to the estimated breeding value of an animal in another country.

This project aims to identify the optimal data structure required to accurately estimate genetic correlations within and between countries in beef cattle maternally influenced traits and to identify and validate possible sources of bias due to missing information. You will use a previously simulated dataset with different levels of sires exchanges between two populations and compare estimates obtained with pedigree and genomic data. In this project you will improve your programming skills and learn to estimate genetic correlations in national and international genetic evaluations.

Supervisor(s): Renzo Bonifazi

What is the impact of genomic selection on genetic diversity?

Selection in a livestock aims to improve the future performances of the population. As a result of selection, changes are taking place on the genome, such as changes in allele frequencies and genetic variation. Insights in those changes help to open up the black box of selection and will lead to a better understanding of the effects of selection. One of the outstanding questions is whether those changes are depending on the used selection strategy (phenotypic, pedigree or genomic selection). The aim of this thesis is to compare the genomic changes in a livestock population under the three abovementioned different selection strategies in order to learn whether the impact on genomic variation is different. This is important to know because it has implications for how we should preserve valuable genetic diversity in the future. You will use a simulated dataset of a livestock population that was under 50 generations of selection. You will learn to quantify the genomic changes by performing analyses on runs of homozygosity, selective sweeps and changes in allele frequencies across the genome.

Supervisor(s): Yvonne Wientjes, Mirte Bosse

Social genetic effect of tail damage and correlation with maternal ability in pigs

Tail biting is a common problem on commercial pig farms. Tail biting affects the health and welfare of the animals and therefore it needs to be managed. The underlying motivations for this behaviour are still partly unclear and can potentially range from boredom to social hierarchy. Many environmental factors such as housing, management and diet have been shown to affect this harmful behaviour. Some studies have also shown the importance of genetics for the development of tail biting. This behaviour is hard to study because it often lasts only a few seconds and therefore requires continuous observation at the individual level. An alternative method is to record the tail damage on the animals and thereby look at the victims of tail biting. With the help of social genetic models, we can use tail damage data to estimate genetic parameters of tail biting behaviour. Social genetic effects are genetic effects that are visible in the phenotypes of pen mates. According to anecdotes from pig farmers, animals that are more aggressive towards humans take better care of their piglets. The question is if harmful behaviour towards other pigs also correlates with better mothering ability. Data on tail damage and maternal ability is provided on around 15k animals by a breeding company.

In this thesis you will develop social genetic models to estimate the genetic parameters of tail damage and use multivariate models to estimate the correlation with maternal ability. This thesis offers an opportunity to

work with animal models and learn about social genetic effects. You will also get acquainted with ASReml, a program frequently used in animal breeding for variance component estimation.

Supervisor(s): Piter Bijma, Bernadett Hegedus

The long-term effects of genomic selection

Genomic selection was introduced at the beginning of this century. Since that time, genomic selection has been rapidly implemented in most of the livestock breeding programmes worldwide. However, at the moment not much is known about the long-term effects of genomic selection and whether enough genetic variation is maintained for future selection. Using an existing simulation program you will investigate the impact of the different parameters on the longterm effects of genomic selection. Options are to investigate the impact of the number of loci underlying the trait or the change in breeding goals over time, or to investigate the usefulness of older generations in a reference population for genomic selection. You will learn to use different selection methods such as pedigree and genomic selection and to investigate the long-term effects of selection.

Supervisor(s): Yvonne Wientjes

Breed for speed? Designing sustainable breeding programmes for racing pigeons

Pigeon racing is a popular pastime, especially in the Netherlands, Belgium and China, where many people compete in challenging races. In pigeon breeding, breeding decisions are usually based on intuition rather than on systematic data collection. In addition, inbreeding is commonly applied, which could threaten genetic diversity and the sustainability of pigeon breeding.

This MSc project focuses on the development of a breeding programme for racing pigeons, grounded in data analysis. The project may involve interviews with pigeon breeders to learn how they currently select birds and what the most important challenges are that pigeon breeders face. From there, you will develop a breeding programme for pigeons and compare the efficiency of different breeding strategies using simulations. The goal is to find the best way to improve racing performance and animal health while maintaining genetic diversity. This project is perfect for students interested in learning how to design a breeding programme from scratch.

Supervisor(s): Pascal Duenk

Genetic differences in feed digestibility in dairy cows

In recent years, interest in feed efficiency has increased. This is due to economic reasons, societal issues (feed-food competition) and environmental concerns (for example nitrogen pollution). Recently, the Dutch breeding company CRV included feed efficiency in its breeding objective. However, there is still a lack of knowledge on the biological background of differences between cows in feed efficiency. Some studies showed that differences in feed efficiency might be partly explained by feed digestibility. Digestibility is measured as the difference between the feed ingested and the faeces excreted. There are indications that feed digestibility is heritable and that genetic improvement is possible, but as it is a laborious and expensive trait to record, there are few studies on its genetic background. Currently, the breeding company CRV records routinely feed intake on five commercial dairy farms. In a combined research project of Animal Breeding and Genomics (ABG) and Animal Nutrition (ANU), feed digestibility will be measured on approximately 400 genotyped cows on commercial farms. Feed intake data will be provided by CRV and faecal samples will be collected to estimate feed digestibility. The aim is to better understand differences between cows in feed digestibility through the analysis of the genetic background, including estimation of heritability, genetic variance and a genome wide association study. The project also aims to better understand the relationship between feed efficiency and digestibility, so the correlation between digestibility and other traits, including feed related traits, will be examined.

Supervisor(s): Henk Bovenhuis, Eugénie Guennoc

Estimating heritability when genes interact

Heritability is an important parameter in animal breeding because it partly determines the potential for genetic improvement in a population. When estimating heritability, the focus is on the so-called 'additive effects' of alleles because only these additive effects are inherited. In reality, however, alleles interact with each other, resulting in non-additive genetic effects such as dominance and epistasis. In simulated data, we observed that the presence of dominance can lead to inaccurate estimates of heritability.

In this project you will investigate the impact of non-additive genetic effects on estimates of heritability using simulations. Some examples of questions are: What is the impact of non-additive genetic effects on estimates of heritability? Should we include non-additive genetic effects in our models? Can genotype data improve estimates of heritability in the presence of non-additive effects? This topic is meant for students who are up for a challenge and want to learn more about quantitative genetics, and estimation of heritability with (genomic) models. This project requires a good understanding of genetics and statistics (from courses like GIL and MSLS), and some affinity with or interest in programming.

Supervisor(s): Pascal Duenk

Estimating genetic variance in the current population using genetic data

The genetic variance in the current population determines the response to selection that can be achieved in the next generation. Monitoring the genetic variance in recent generations can therefore inform us whether the current breeding programmes are sustainable and maintaining enough genetic variation to sustain selection in the future. However, by using standard animal models, the genetic variance is estimated in the base generation of the population and not in the current population. Recently, a new approach was developed based on marker (SNP) information that could be used to estimate current genetic variance components.

In this MSc thesis you will investigate how accurate this approach can estimate the genetic variance in different scenarios. You will use a simulated dataset in which the actual genetic variance is known. You will develop your programming skills, learn how to handle large data sets and learn how to estimate genetic variance components using marker data.

Supervisor(s): Yvonne Wientjes

Evaluating the efficiency of a two-step approach for routine single-step genomic evaluations

Single-step genomic evaluations have become a crucial tool in animal breeding programmes, enabling the simultaneous estimation of breeding values for genotyped and non-genotyped individuals. However, as the volume of genomic data grows, routine single-step evaluations are becoming increasingly computationally challenging. This MSc study aims to investigate a novel two-step approach to alleviate this challenge. This novel two-step approach consists as follows. First, single-nucleotide polymorphism (SNP) effects and associated prediction error covariances (PEC) are estimated for historical data (e.g., pre-2010) using a single-step SNPBLUP model. Second, these estimated SNP effects and PEC are integrated into a second single-step SNPBLUP using only recent data (e.g., post-2010). The efficiency of the proposed approach will be compared using the results of a single-step SNPBLUP evaluation using all available data. The results of this thesis will demonstrate how this two-step approach may mitigate the computational constraints associated with routine single-step genomic evaluations, enabling more efficient estimation of breeding values. Throughout this thesis you will become familiar with the theory of mixed model equations and routine genomic evaluations using large real datasets, and gain practical experience with genomic prediction tools (e.g., MiXBLUP), programming languages (e.g., R, Julia), and high-performance computing.

Supervisor(s): Jeremie Vandenplas, Renzo Bonifazi

Developing genomic evaluations across countries for novel sustainability traits

Data collection for novel sustainability traits (e.g. feed efficiency and

methane emissions) is challenging as these traits are difficult to measure and expensive. Consequently, data collected on sustainability traits within a single country is usually limited and small reference population sizes are present at the national level, limiting the accuracy of genomic evaluations for such traits. Combining data collected in multiple countries and potentially on multiple breeds into a joint international genomic evaluation can improve the accuracy of genomic predictions and enhance the genetic progress for such societally-relevant traits. However, to combine national datasets containing data recorded on both purebred and crossbred animals and from different (small) populations into a joint international evaluation is challenging.

In this MSc project you will use a large beef cattle dataset with feed efficiency or longevity data collected from three different European countries and investigate one (or more) of the following research areas (depending on your progress and interest): 1) Imputation of genotypes across different panels for pure and crossbred animals, 2) Estimating genetic parameters such as heritability and genetic correlations across countries and breeds, 3) Implementing and validating single-step genomic evaluations using MiXBLUP software. Your work will provide insights into the genetic basis of international evaluations for novel sustainability traits in beef cattle. In this project you will gain hands-on experience with quantitative genetics, genomic prediction models, and statistical software and tools. Moreover, you will develop advanced data analysis skills for handling large datasets (e.g., bash, awk) and programming experience (e.g., R, Python, or similar).

Supervisor(s): Renzo Bonifazi, Jeremie Vandenplas

Exploring the potential applications of randomised linear algebra in genomic selection

The increasing availability of SNP genotype data has transformed the field of animal breeding, enabling the widespread adoption of genomic selection in livestock breeding programmes globally. However, the rapid growth of genomic datasets, which can exceed 1 million genotypes, poses significant computational challenges. Randomised linear algebra has emerged as a promising tool for analysing high-dimensional data, but its potential applications in animal breeding and genomics have yet to be fully explored. This thesis aims to investigate the potential of randomised linear algebra to improve the computational efficiency of various genomic prediction approaches in animal breeding and genomics, and to evaluate its impact on the accuracy of these methods. In this thesis you will become familiar with the theory of randomised linear algebra and gain practical experience with genomic prediction, programming languages (e.g. Julia, Fortran), and high-performance computing.

Supervisor(s): Jeremie Vandenplas

Back to contents page

Genomics

Characterizing variation in immunocompetence genes in racing pigeons

Racing pigeons are subject to intense selective breeding for performance traits such as homing speed and endurance. Selection is typically carried out within closed breeding lines, which may reduce overall genetic diversity through inbreeding and genetic drift. However, immune genes (e.g. those that are part of the major histocompatibility complex (MHC)) are expected to resist this loss of diversity because of balancing selection driven by pathogens. This thesis will use whole-genome sequence data from 148 racing pigeons belonging to a single breeding loft to characterize variation at MHC class I and class II loci and other immunocompetence genes. The main goals are to describe the level and distribution of allelic diversity at these loci, to test whether MHC diversity is maintained, and to identify signals of balancing selection. The results will give insight into the immunogenetic consequences of artificial selection in racing pigeons and may have practical value for breeders interested in maintaining healthy and disease-resistant birds.

Supervisor(s): Pascal Duenk, Richard Crooijmans

Conservation genomics of the short-tail nurse shark

The short-tailed nurse shark (Pseudoginglymostoma brevicaudatum) is a critically endangered shark species. It is found only near the coasts of a few East African countries, notably Kenya, Tanzania, and Madagascar. A breeding programme, coordinated by ARTIS Zoo, aims to maintain the genetic diversity of the species in European aquaria and zoos. However, little information is available on the origin of the fewer than 100 sharks in the breeding programme, and even less is known about the genetic diversity represented by these animals.

Because wild populations are currently extremely vulnerable, ARTIS Zoo is investigating not only how to optimally retain diversity and avoid inbreeding problems in captivity, but also whether the captive population could form the basis for a rewilding project - reintroducing the species to reefs where it once naturally occurred but has now gone extinct.

For this project, we are building genomic resources to inform the breeding programme. Since no reference genome is currently available, this is our first step. It is not an easy one, as the genome is around 1.5 times larger than the human genome. We are also fully re-sequencing individual sharks to assess diversity at the whole-genome scale. This information will inform several critical biodiversity indicators, such as relatedness among animals (even across multiple generations), evidence for inbreeding due to unguided breeding in captivity, and possibly inbreeding in the remaining wild populations. Moreover, genome analysis may reveal geographic differentiation, if present, which could have important implications for captive breeding as well as decisions regarding rewilding.

Sequence data for genome assembly (Oxford Nanopore) and for population analysis (Illumina-based re-sequencing) are available. You will be expected to carry out analyses on the Wageningen computer cluster (Anunna), so affinity with Linux-based systems and scripted analyses (e.g. Python) is desirable. The project is a collaboration between Wageningen University and ARTIS Zoo, Amsterdam. As a student on the project, you will be invited to participate in project meetings in Amsterdam.

Supervisor(s): Hendrik-Jan Megens, Mirte Bosse

The effect of fragmented swampscapes on genetic diversity in the autotetraploid weatherfish

The weatherfish (Misgurnus fossilis) is a characteristic fish species of Dutch freshwater wetlands. Historically, the species could move freely throughout the floodplains of the major rivers during periods of high water. However, since flooding is no longer possible in most areas, the species is now restricted to relic populations. Many of these remaining

populations are so small that local inbreeding may pose a serious threat to their survival.

In this project, we investigate the diversity of Dutch weatherfish populations and specifically aim to link habitat size (and expected population size) to diversity parameters and inbreeding risk. We do this by re-sequencing individual genomes and thereby characterizing the majority of genomic variation in each fish. However, there is a complication: European weatherfish populations are autotetraploid. Interestingly, polyploidy is a common feature in the family of true loaches. Fascinating as this is, it also makes genome characterization more challenging.

Students can work on various aspects of the project, ranging from genome assembly and annotation to variation analysis, estimation of inbreeding risk, and the search for signatures of local selection. Which aspects are available will depend on the timing of the project. During some parts of the year, especially from late spring to early autumn, there may also be opportunities to participate in fieldwork. Although the weatherfish is relatively rare, its historically broad range means that it can still be found in many places in the Netherlands. The project is a collaboration between Wageningen University, RAVON, INBO (Belgium), and several Dutch water boards and provinces.

Supervisor(s): Hendrik-Jan Megens, Arjen Palstra, Reindert Nijland (MAE)

Inbreeding risk in Dutch burbot populations

The burbot (Lota lota) is, in many ways, a remarkable fish. It is the only freshwater representative of the cod family (Gadidae). It is also a species that depends strongly on cold Dutch winters, which are becoming increasingly rare. Burbot spawn in winter in flooded areas when water temperatures are below six degrees, preferably even colder. Due to warmer winters, spawning success has been very low in recent years, and population numbers are declining steeply in some areas. In other areas, however, burbot are doing relatively well, somewhat surprisingly. In some places, burbot populations have been restored through captive breeding programmes, with varying levels of success. There is concern that the limited number of reproducing individuals could accelerate extinction risk through inbreeding.

In this project, we are generating a genome assembly for a local burbot population. In addition, re-sequencing data are available from two populations. Analyses will be carried out on the Wageningen computer cluster (Anunna), which requires some familiarity with the Linux command line and scripted analysis (e.g. Python), or a willingness to learn these skills during the project. Depending on the time of year, there may also be limited opportunities to participate in fieldwork in the Netherlands. The project is a collaboration between Wageningen University, RAVON, and INBO (Belgium).

Supervisor(s): Hendrik-Jan Megens, Reindert Nijland (MAE)

Two species or one? River lamprey and brook lamprey diversity in Dutch streams and rivers

Lampreys are jawless fish. One could even argue that lampreys are not fish at all, but instead represent an ancient split in the vertebrate lineage almost half a billion years ago. In the Netherlands, there are three lamprey species, two of which are common in freshwater: the river lamprey (Lampetra fluviatilis) and the brook lamprey (Lampetra planeri).
River lampreys are migratory: they mature at sea and then migrate upstream through the major rivers. Driven by pheromones excreted by lamprey larvae, they find suitable spawning habitat. After spawning, they die. The larvae remain in small streams, blind and filter-feeding, until they undergo metamorphosis and return to sea. Brook lampreys, in contrast, undergo metamorphosis, mature immediately, spawn, and then die. The two species are quite distinct. River lampreys are much larger and have a different mouth structure.

Each species has its own conservation concerns. River lampreys are hindered in their migration by the many barriers in the highly regulated Dutch freshwater system. Brook lampreys, by contrast, depend on very small streams with sufficient gravel substrate, which are relatively rare in the Netherlands. In addition, these streams are highly vulnerable to

drought, creating significant extinction risks, especially in strongly disconnected brook systems.

Intriguingly, both species often use the same streams for reproduction, facilitated by the fact that river lampreys are attracted to the pheromones released by brook lamprey larvae. Evidence is mounting that these two forms may not be separate species at all, but rather ecotypes of the same species. Brook lampreys may therefore represent formerly migratory lampreys that became landlocked.

In this project, we address two specific conservation questions:

1. Are river lampreys and brook lampreys truly separate species? Is there genetic exchange between them, and if so, how much?
2. What are the inbreeding risks in disconnected brook lamprey populations?

Analyses will be carried out on the Wageningen computer cluster (Anunna), which requires some familiarity with the Linux command line and scripted analysis (e.g. Python), or a willingness to learn during the project. Depending on the time of year, there may also be limited opportunities to participate in fieldwork (Veluwe, Achterhoek, Limburg). The project is a collaboration between Wageningen University, RAVON, and INBO (Belgium).

Supervisor(s): Hendrik-Jan Megens, Stan Coppis, Reindert Nijland (MAE)

Riverscape genomics of perch

The perch (Perca fluviatilis) is a very common species in Dutch freshwaters. It is found in small streams and ponds, but also in large lakes such as Lake IJssel, and in all major rivers in the Netherlands. The species is therefore highly adaptable, which raises the question of how this adaptation is achieved. Phenotypic plasticity is certainly an important factor, but as always, adaptive potential also has a genetic component. There may therefore be genetic differences among populations that have not yet been explored.

One of the most notable differences among populations is body size. In particular, the perch population in the Haringvliet, at the downstream end of the Rhine river system, grows exceptionally large. So large, in fact, that it attracts enthusiastic anglers from abroad. Haringvliet perch are famous. Food abundance is expected to play a role in this gigantism. However, the Haringvliet is also a man-made lake: it used to be an estuary, but today a large dam separates it from the sea. Do Haringvliet perch show adaptation to these unique habitat conditions?

In this project, we want to investigate diversity in perch across the Dutch delta, from small ponds to large lakes, and from small canals to major rivers. The Rhine in particular acts as an "express highway" for fish genetic diversity from Switzerland to the North Sea. Are there differences among perch populations at different locations along the Rhine? How do these differences compare with populations in connected river systems such as the IJssel and Meuse? How does diversity in the Rhine compare with that in formerly connected water bodies now cut off by dikes? And is there evidence of adaptation along this riverscape?

We started resequencing perch throughout the Dutch delta and along the Rhine into Germany. In addition to sequence data, we also intend to collect other types of data to characterize local habitats. The main component of the thesis will be the analysis of re-sequencing data, but depending on the time of year, there will also be ample opportunity to participate in fieldwork. There is also considerable scope to combine genetic and phenotype-based analyses using morphological parameters. The project is a collaboration between Wageningen University and the Dutch Angling Society (Sportvisunie).

Supervisor(s): Hendrik-Jan Megens (ABG), Reindert Nijland (MAE), Leo Nagelkerke (AFI), Niels Breve (Sportvisunie)

An ide with a predatory streak: hybridization between ide (Leuciscus idus) and an exotic predator (Leuciscus aspius) in the Netherlands

The ide is a very common fish in the Netherlands, especially in larger rivers and streams. It does, however, show a distinct preference for spawning in smaller streams, and every year in late winter and early spring ide undertake a rather spectacular migration - one that is sadly often overlooked. We strongly recommend following the Visdeurbeljournaal to witness this phenomenon. While perhaps not on the scale of the Serengeti migration, it is still something quite special.

Ide is not the only fish species to migrate in this way. There is also a newcomer in Dutch ichthyofauna: the asp. Originally from the Danube, it reached the Rhine system through the Main-Danube Canal in the 1980s. While ide is a jack-of-all-trades, readily feeding on a wide variety of macrofauna, the asp is a voracious predator that specializes in fish.

As both species undertake their annual migration, they meet. And mate. And produce offspring that are clearly recognizable as "something in between". Hybridization between closely related cyprinid fish is common, especially in areas where species do not naturally co-occur, suggesting that no strong reinforcement mechanisms have yet evolved to keep them reproductively separate.

For this study, we address several questions:

1. Are the hybrids sterile, or do they contribute to further generations?
2. Is hybridization symmetric, or is it mainly initiated by one sex of one of the species?
3. Is there evidence of reinforcement between the species?
4. Can we develop quick, field-ready genetic tests to identify hybrids from skin swabs?

We have already started re-sequencing ide, asp, and hybrids for this project. The project is a collaboration between Wageningen University and the Dutch Angling Society (Sportvisunie). If you are interested in angling, there will be ample opportunity to participate in fieldwork.

Supervisor(s): Hendrik-Jan Megens (ABG), Reindert Nijland (MAE), Leo Nagelkerke (AFI), Niels Breve (Sportvisunie)

Genomic analysis of flight performance in racing pigeons

Racing pigeons are subject to intense selective breeding for specific performance traits, such as flight speed and endurance. This process often leads to the development of specialized lines with distinct athletic characteristics. This MSc thesis project will use whole-genome sequence data from 148 racing pigeons representing two different lines: one selected for long-distance and the other for short-distance flights.

The primary goal of the project is to identify allele frequency differences between these two lines and to detect genes that show signals of differential selection. By analyzing these genomic regions, the project aims to link specific genetic variations to differences in flight performance. Throughout the study, the student will perform detailed genomic analyses and learn to use essential bioinformatics tools. The results will provide insight into the genetic basis of athletic traits in birds and contribute to our understanding of how artificial selection shapes the pigeon genome.

Supervisor(s): Pascal Duenk, Martijn Derks

Can homozygosity reliably measure inbreeding?

In many current breeding programmes, inbreeding is measured as observed genome-wide homozygosity, delivered from genotypes obtained with medium density SNP (single nucleotide polymorphism) chips. Those SNP chips are often constructed by selecting e.g. ~50,000 SNPs that have a relatively high minor allele frequency in breeds and populations of interest. This procedure to construct SNP chips can lead to so-called ascertainment bias of the included SNPs. The ascertainment bias is typically ignored when changes in homozygosity are measured over time.

In this MSc thesis you will compare observed homozygosity based on a medium density SNP chip versus all SNPs in the entire genome. Homozygosity based on all SNPs in the entire genome will serve as a reference of 'true' homozygosity. This thesis will be based on simulated data. Setting

up the simulations, including a selection history and construction of a medium density SNP chip, will be part of the thesis work.

Supervisor(s): Mario Calus

Genetic diversity, inbreeding, and purifying selection in global chicken populations

Single nucleotide polymorphisms (SNPs) and small insertions and deletions (indels) are the most abundant forms of genetic variation and contribute significantly to phenotypic diversity and evolutionary processes. Studying these variants across diverse populations can reveal patterns of genetic load, inbreeding, and natural selection.

While commercial chicken breeds have been extensively studied, much less is known about the genome-wide patterns of deleterious variation and purifying selection across the full spectrum of global chicken diversity. In this thesis the student will analyse a unique whole genome sequence dataset comprising over 3,000 chickens representing a wide range of breeds and geographic origins from across the world. The goal is to characterise SNP and indel variation, assess genetic load, and quantify inbreeding levels in each population. The student will use state-of-the-art computational tools, including a chicken-specific Combined Annotation Dependent Depletion (CADD) score, to predict the functional impact of variants and identify signatures of purifying selection.

The outcome of this thesis will be a comprehensive overview of genome-wide variation and selective constraint in chickens, providing key insights into how breeding history, domestication, and environment have shaped the genome. The findings will be of broad relevance to evolutionary biology, animal breeding, and conservation of genetic resources in one of the world's most important livestock species.

Supervisor(s): Martijn Derks

Two populations, one breed: how to conserve the Mergelland sheep (together)?

Mergelland sheep is an endangered breed from the South of the Netherlands and North of Belgium, with very small population size in both countries. Conserving local breeds such as the Mergelland is important for maintaining biodiversity, preservation of cultural heritage and possibly ensuring future resilience to changing environmental conditions. This project aims to assess the level of genetic differentiation between the Dutch and Belgian Mergelland sheep, identify any unique diversity present in either population, and explore the potential for optimal exchange of breeding animals between them. Pedigree information for both populations is available, along with genotype data from approximately 100 animals (NL + BE combined). Additionally, the role of the Dutch gene bank could be analysed and evaluated to see whether its stored material can contribute meaningful in the conservation of the Mergelland sheep in NL & BE. This project is a collaboration between KU Leuven (Center of Animal Breeding and Genetics) and the Centre for Genetic Resources, the Netherlands (CGN), of Wageningen University & Research (Animal Breeding and Genomics).

Supervisor(s): Mira Schoon (CGN/WUR-ABG) and Steven Janssens (KU Leuven-CABG)

Cross-species mapping of embryonic pain-neuron development

When do animals first become capable of pain-related processing during development, and how much of the underlying biology is shared across species? The biological focus of this project is on parts of the nervous system that are directly involved in sensing and processing pain. One tissue of interest is the trigeminal ganglion, which contains the cell bodies of sensory neurons that detect touch, temperature, and pain from the face and head. These neurons are among the first 'pain sensing' cells to form during embryonic development. In addition, the project includes parts of the forebrain, such as the thalamus and pallial regions. The thalamus is a central relay station that receives sensory information (such as pain signals) and sends them to the higher-order pallium that contains areas involved in the processing of the information (perception of pain). Together, these tissues could provide how pain-sensing neurons emerge and mature

during the embryonic development.

In this MSc project you will analyse chick embryonic single-nucleus RNA-seq (snRNA-seq) and compare the results to other species. snRNA-seq is a method that measures which genes are on/off in thousands of individual cells, which gives a detailed snapshot of the cell types and developmental states present in a tissue. This high-resolution transcriptomic analysis will be carried out across multiple embryonic stages (planned: days 5, 7, 9, 11, and 13). This allows you to track how relevant neural cell populations emerge and mature over time. The project will start with in-person training in single-cell analysis (quality control, clustering, cell type/state annotation, and general methodology), so you will learn how we go from raw sequencing data to biologically meaningful cell groups. After this, you will perform cross-species mapping by aligning chick cell states to mouse/rat/macaque/human reference atlases using ortholog mapping (matching comparable genes between species) and atlas alignment/label transfer (matching comparable cell types/states). This step is important because it lets you test in a quantitative way, which development programmes related to pain-sensing and sensory processing are conserved across species versus avian-specific. Your results could highlight candidate genes/markers that explain the similarities and differences. The expected outcome is a 'conservation map' where you link chick cell types/states to mammalian counterparts, and together with a list of marker genes relevant to the emergence of pain pathways and the maturation of the involved neural circuits.

For requirements, some basic working level of R and Unix/Linux is expected (e.g. running scripts, handling/parsing tables, and working on a server/HPC). Python is helpful too. Even though the topic is about pain, this thesis will not be painful. It will be a fun, supportive, and a data-driven project focused on learning something new.

Supervisor(s): Ole Madsen, Hamid Fotowatikha

Recalibrating breeding goals including antagonistically correlated traits

Current breeding goals include many traits, with often antagonistic genetic correlations between them. Despite this antagonism, each trait has some genetic variance that is not associated with genetic variance in other traits. In this thesis, you will work on a conceptual recalibration of a breeding goal, by splitting each trait in multiple traits that do or do not covary with the other breeding goal traits. Based on derived index weights for all traits in this recalibrated breeding goal, you will then evaluate the possible advantage by putting more emphasis on selecting variation private to traits of interest. The breeding goal recalibration will involve working with matrix algebra and selection index theory. To enable validation of the recalibrated breeding goal, the work will be based on simulated data.

Supervisor(s): Mario Calus

From microsatellites to SNPs: how to modernise parentage testing in the Dutch Draft Horse?

The Dutch Draft Horse (Trekpaard) is a rare native breed in the Netherlands, where parentage verification has long relied on microsatellite markers. With SNP genotyping now the global standard, a transition is needed to ensure accuracy, cost-efficiency, and long-term sustainability. This project aims to design a roadmap for moving from microsatellites to SNPs, making use of the available pedigree records, extensive microsatellite data, and SNP genotypes from ~150 horses (partly overlapping with microsatellite profiles). Beyond parentage testing, SNP data will also enable valuable studies of genetic diversity and population structure, strengthening conservation efforts. The goal is to establish a validated, affordable protocol that secures the genetic integrity of the Dutch Draft Horse and can serve as a model for other small horse studbooks in the future.

Supervisor(s): Bart Ducro, Mira Schoon

CNV analysis in different pig breeds using 650k SNP genotypes

You will work with a large SNP genotype data set of around 100 pig samples of different breeds. By using PennCNV-software you will identify CNVs. You will characterise the CNV regions by looking at overlapping

genes and look at breed differences and function of the genes involved and compare results with SNP genotypes from whole-genome sequence data and from 60K SNP chip.

Supervisor(s) Richard Crooijmans

Insights of the Lakenvelder phenotype in farm animals

Coat colouring in farm animals is a complex phenotype. The Lakenvelder phenotype is a phenotype occurring in several species. We would like to know the genetics and genomics behind this trait. Furthermore, we would like to have a summary of all genes involved in coat colour over species. We do have SNP genotypes of Lakenvelder chicken and cattle available. A selective sweep analysis in chicken and in cattle Lakenvelder can be performed to find potential new genes associated with this trait. This project is partly a literature study and SNP data analysis. A comparative study between species will be part of the task.

Supervisor(s): Richard Crooijmans

The search for the missing tail

A fowl without a tail might sound odd, but the Drenthe fowl has been around for centuries in the Netherlands. A variant of this breed called Klomphoen is the rumples Drents. The Drenthe Fowl Bolstaart shows a strikingly rounded rump due to the lack of a tail, or – to be more precise – uropygium, from which the tail grows. The trait of caudal truncation is kept by farmers for the reason that the rumples chickens escape better from the attack of foxes, because the predators have no grip due to the absence of the tail. Here we would like to develop a MSc thesis study to investigate the genetic nature underlying the tailless phenotype by using the genetic data of Drenthe Fowl Bolstaart and Drenthe Fowl. It is possible to use the whole-genome sequence data to work on the genetic variants that are associated with the tailless phenotype. In order to tackle the research question, genome-wide association study, detection of signal selection will be possible. We are looking for a student with a basic knowledge of genomics (variations, genome sequencing and functional annotation) and working experience with linux machine (basic bash coding) and R.

Supervisor(s): Richard Crooijmans

Building the European mink genome using long-read data

The mink is a semiaquatic carnivorous mammal that is found in various habitats. The European mink (Mustela lutreola) is considered as critically endangered in the IUCN red list, and is already extinct in large parts of its ancestral range. Until now, the only reference genome available was the American mink (from 2018), a non-endangered mink often used in fur farming. To study the evolution and genomic health of the endangered European mink the ERGA (European Reference Genome Atlas) is working to produce a chromosome level European mink genome. In this MSc project you will work with long-read sequence data (Nanopore) from the European mink that has recently been generated in the department of ABG in collaboration with Copenhagen University. You will be in charge of generating a de novo assembly and, more specifically, will perform comparative analysis between the American and European mink genome (i.e. structural variation).

Supervisor(s): Peter Karlskov-Mortensen (Copenhagen University), Marta Godia Perello and Martijn Derks

SCD1 genotypes in dairy cattle may be missing a link to fertility

The negative energy balance (NEB) characterised by high serum non-esterified fatty acid (NEFA) concentrations, due to increased mobilisation of body fat reserves, is a risk factor for reduced fertility performance in dairy cows. Saturated NEFAs can have lipotoxic effects and may negatively affect the quality of the oocyte and embryo. In contrast, high levels of unsaturated NEFA do not include such effects and can even prevent the lipotoxic effects. Studies at Utrecht University show that the composition and balance of saturated and unsaturated free fatty acids in follicular fluid is crucial for postfertilisation developmental competence of oocytes (Aardema et al., 2013). The enzyme stearoyl-VoA desaturase1 (SCD1) enzyme that is also present in cumulus cells converts saturated into unsatu-

rated fatty acid and appears to protect the oocyte against elevated levels of saturated NEFA (Aardema et al., 2017). Interestingly, in dairy cattle with tho isoforms of SCD1 have been identified. These genetic variants of SCD1 have been associated with carcass fatty acid composition in Japanese Black Cattle (Taniguchi et al., 2004) and with milk fatty acid composition (e.g. Schennink et al., 2008): the A and the V variant of SCD1 differ in their capacity to desaturate specific fatty acids. Therefore we hypothesise that that SCD1 genotypes have an effect on the fertility of cows, especially during the period of negative energy balance that precedes the insemination period.

A dataset of cows from parity 1, 2 and 3 cows forms the basis to unravel a potential link between SCD1 expression and fertility. During this MSc thesis you will have regular discussions with veterinarian and reproduction specialist dr. Hilde Aardema and you will work in close collaboration with the reproduction laboratory of the faculty of Veterinary Medicine at Utrecht University.

Supervisor(s): Hilde Aardema (Utrecht University)

Genomic analysis of milk oligosaccharides and their nutritional benefits

Goat milk oligosaccharides have unique characteristics that make them an attractive choice for improving human nutrition. In this study you will describe and identify the genomic region and genes affecting milk oligosaccharides in livestock species. The aim is to learn more about the known genetic features already studied in the goat populations.

Supervisor(s): Richard Crooijmans

Morphological characterisation of goat population

The production of goat milk is expected to grow in the coming years, due to its beneficial effect on infant health. However, despite the importance of this species for the human population, goats have been significantly less studied than other livestock species. The aim of this thesis is to evaluate the functional effect of goat phenotypes such as beard, wattles color and ear position by performing GWAS and analyse potential detected genes.

Supervisor(s): Richard Crooijmans (multiple thesis topics available)

Building a reference genome of Yak (Bos grunniens)

Wild and domesticated yaks are classified as two subspecies of grunniens species, under the Bos genus and Bovidae family. The international Commision on Zoological Nomenclature suggested a taxonomic classification i.e. Bos mutus (Przewalski, 1883) for wild yaks and Bos grunniens (Linnaeus, 1766) for domesticated yaks. Morphological difference between two subspecies is that wild yak appeared to be larger in size and fiercer in behaviour. The domesticated yak (Bos grunniens) is a main subsitence species for shepherds in the high mountains of Central Asia. The traditional habitat area extends through southern part of Russia, north and western Mongolia, Kyrgyzstan, Tajikistan, Afghanistan, northwestern China (Tibet), Nepal, Bhutan and northern India. These areas are elevated at the approximate altitude between 2000 and 5500 meters. Both wild and domesticated yak are distributed over the highlands of Central Asia. Domesticated yaks produce meat, milk and fiber.

In this MSc project you will work with long and short read sequences of one individual to build a reference genome. Variant calling including structural variation detection within the new reference and compare these with public available WGS data.

Supervisor(s): Richard Crooijmans, Marta Godia Perello and Martijn Derks

Structural variant calling in water buffalo (Bubalis bubalis)

The water buffalo (Bubalis bubalis), also called the domestic water buffalo or Asian water buffalo, is a large bovid originating in the Indian subcontinent and Southeast Asia. Today, it is also found in Europe, Australia, North America, South America and some African countries. Two extra types of water buffalo are recognised, based on morphological and be-

havioural criteria: the river buffalo of the Indian subcontinent and further West to the Balkans, Egypt and Italy and the swamp buffalo found in Asia. Water buffaloes are especially suitable for tilling rice fields, and their milk is richer in fat and protein than that of dairy cattle.

In this project you will assemble a new reference genome based on nanopore long read sequences and Illumina short read sequences. You will compare detected variants (SNPs and SV) and compare these with public available WGS data.

Supervisor(s): Richard Crooijmans, Marta Godia Perello and Martijn Derks

Structural variant calling and admixture analysis of tauros cattle (Bos tauros)

The auroch is the ancestor of all cattle and thereby the most important animal in the history of mankind. The keystone species for many European ecosystems was hunted to its extinction in 1627. However, its DNA is still alive and distributed among a number of the ancient original cattle breeds. The aim of the Tauros Programme is to create a modern-day equivalent of a long-dead animal. The principal technique is 'back-breeding': by combining cattle breeds with desired characteristics, a bovine can be created with physical attributes, behaviour and genetics that closely match those of Europe's original wild aurochs. Today, many old European cattle breeds still retain a genetic similarity to the aurochs. Following the selection of six or seven such breeds from an original list of over thirty. In this project you will assemble a reference genome of a tauros animal (7th generation) and make a variant map of this animal. You will compare this with the WGS data of the seven breeds used to create this animal and determine the admixed genome structure. Also, you will compare the newly assembled genome of the tauros animal with the WGS of public available aurochs animals.

Supervisor(s): Richard Crooijmans, Marta Godia Perello and Martijn Derks

Full genome analysis of the African buffalo

Various studies on basis of microsatellites show genome-wide positive selection of deleterious alleles. This positive selection is related to a sex-ration meiotic drive. Preliminary analysis of buffalo genomics indicates associations between these deleterious alleles and GC content. Earlier microsatellite studies also indicate that these deletrious alleles occur in LD blocks. We are looking for a student who wants to try to support these indications with a more elaborate full genome analysis. Some 30 full genome sequences are available (from the genebank).

Supervisor(s): Pim van Hooft (Wildlife Ecology and Conservation group)

DNA methylation patterns and relation with sex and age in eel

We would welcome a MSc student who is interested in investigating DNA methylation patterns in eel and the relation with sex and age. Sex differentiation in eels occurs rather late in life at a size of 20-25 cm. From 25+ cm eels we can determine sex invasively and only from larger 40+ cm eels we can determine sex by external appearance. Age can only be determined by counting the growth rings on the otolithes for which the eel has to be dissected. Ultimately, we could use methylation markers for sex determination and age prediction, raising an important genetic tool for eel management. You will analyse the individual variations in DNA methylation patterns of eels differing in sex, age and size and investigate their relation.

Supervisor(s): Arjan Palstra, Hendrik-Jan Megens

Fish migration genetics and physiology, and disturbing contamination

Together with our partner Waterschap Aa en Maas, we have identified several research priorities under the title fish migration genetics and physiology for which we would welcome MSc students:

- Impact of fish migration measures on population structure and species distributions (Gobio gobio, Barbatula barbatula, Misgurnus fossilius);
- Negative effects of contaminant cocktails in fish migration (literature

study);
- On-site testing swimming performance of a diversity of fish species in relation to water quality;
- Methylation patterns in finclips of fish in relation to contamination.

Supervisor(s): Arjan Palstra, Hendrik-Jan Megens, Carlo Rutjes (Waterschap Aa en Maas)

Genome-wide characterisation of microRNAs in the European flat oyster (Ostrea edulis) and their conservation across bivalve molluscs

The European flat oyster (Ostrea edulis) is a bivalve mollusc species native to Europe and has been an integral part of the European diet for centuries. This species displays extraordinary genetic diversity and thrives across a range of physiological conditions. However, anthropogenic activities and disease outbreaks have severely affected wild stocks and farming of this iconic species. Large-scale restoration efforts are now focused on restocking O. edulis to develop sustainable populations. However, parasitic disease outbreaks have significantly hampered these efforts. While multiple reference genome assemblies, genomic, and transcriptomic datasets have been developed for the species to better understand traits of interest, such as growth and disease resistance, small noncoding RNAs remain uncharacterised in this species. MicroRNAs are highly conserved, small noncoding RNAs in the genome involved in the regulation of gene expression. These molecules contribute to a variety of biological processes, including development, immune response, biomineralisation, and host-microbe interactions. Evolutionarily, microRNAs are known to be highly conserved across species, with novel families added over time rather than lost.

This project involves characterising microRNAs across the genome of O. edulis using small RNA data generated from a pooled sample of eight tissues. Additionally, the identified microRNAs will be compared against the miRNAomes of other evolutionarily and commercially important bivalve species to identify conserved and novel microRNAs in the flat oyster genome. The outcomes of this project will provide the first comprehensive annotation of microRNAs in the European flat oyster genome, paving the way for an improved understanding of the genetic basis of various biological processes. We expect the student to possess basic bioinformatic skills or should be willing to learn.

Supervisor(s): Manu Kumar Gundappa, Marta Godia Perello

The curious case of the moving sperm: defining heritability of variation in porcine sperm kinematics

Sperm cells transmit the paternal DNA to offspring. These important cells exhibit complex motion characteristics that influence fertility outcome. These kinematic traits, such as velocity and travel distance, display considerable variation, which may be partially explained by genetic factors. Understanding the heritability of this variability is crucial for advancing reproductive performance. This thesis will focus on quantifying the heritability of kinematic traits in porcine sperm, such as velocity and distance. The student will work in a collaborative project with a breeding company (Topigs Norsvin). They will have access to a large dataset on phenotypic records of kinematic traits measured on 250,000 ejaculates collected from over 3,000 boars and additional pedigree information. The student will use advanced statistical tools to partition genetic and environmental sources of variation and estimate genetic contributions to these traits and phenotypic and genetic correlations between them. Familiarity with statistics and quantitative genetics is recommended, as well as completion of courses such as Analysis for Plant and Animal Breeding or Genetic Improvement of Livestock. This work will provide insights into the genetic basis of sperm motility traits, paving the way for novel approaches to improving fertility management in swine breeding.

Supervisor(s): Henk Bovenhuis, Rodrigo Mezencio Godinho (Topigs Norsvin), Pedro Sá

Alternative splicing in the gut transcriptome of Tilapia larvae

We are looking for a student interested in investigating alternative splicing in Oreochromis niloticus (Nile tilapia) using Nanopore sequence data.

Alternative splicing plays a key role in gene regulation, but little is known about its patterns in fish gut tissue, especially in response to dietary interventions. This project focuses on building a transcriptome from first-feeding tilapia larvae fed different types and levels of beta-glucans, a dietary supplement known to influence gut health. Using long-read sequencing, we aim to identify full-length transcripts, compare them to existing reference databases, and find evidence of novel splicing events. If possible, we will also examine whether different dietary treatments influence alternative splicing patterns. This study will provide new insights into transcriptome complexity in tilapia, with potential implications for fish nutrition and health. The project will involve bioinformatics analysis, including transcriptome assembly using Nanopore sequencing data, identification of novel splice variants, and interpretation of the results. Students with an interest in molecular biology, bioinformatics, or aquaculture genetics are encouraged to apply.

Supervisor(s): Manu Kumar Gundappa, Fotini Kokou (AFI)

Back to contents page

Genetic diversity

Population analysis of the Groninger horse: balancing diversity and heritage in an open studbook

The Groninger horse is one of the smallest Dutch horse breeds in the Netherlands, with deep historical roots and strong cultural significance. Like many horse studbooks they have an open studbook, making use of related breeds and stallions from different populations. In this project you will begin with a genetic population analysis to assess the current risk status of the breed. This includes evaluating pedigree completeness, trends in inbreeding, effective population size, and patterns of stallion usage. We will also quantify the genetic contributions of foreign breeds and assess how they have influenced genetic diversity. Next, using simulations of various breeding scenarios, we will explore strategies to support the long-term sustainability of the population. Based on these outcomes, we aim to provide tailored recommendations on measures such as mating restrictions, the strategic use of gene bank material, and finding the optimal balance between preserving heritage and maintaining sufficient genetic openness for the future.

Supervisor(s): Bart Ducro, Mira Schoon

Restoring the dace: biogeography and conservation genomics of the freshwater fish Leuciscus leuciscus in the Netherlands

The dace - serpeling in Dutch - is not a particularly uncommon fish. However, it does have rather specific habitat requirements, such as clean, freely flowing water. Its sensitivity to migration barriers and pollution has made this species an indicator of improving water and habitat quality in the Netherlands. Unfortunately, because dace is not a particularly strong natural disperser, it has difficulty recolonizing waters from which it disappeared during the past century.

This project focuses on informing restoration efforts in the Netherlands using genome data. We know that dace populations differ genetically among drainage basins: streams connected to the Eems, Rhine, Meuse, and Scheldt all contain distinguishable populations. However, little is known at the level of genome-wide variation. Restoration efforts should ideally take these differences into account.

A second objective is to support reintroduction efforts aimed at creating viable populations that harbour sufficient genetic diversity to avoid inbreeding and maintain local adaptive potential. Two real-world cases are addressed:

1. Reintroduction of dace in the north of the Netherlands (eastern Drenthe) by increasing numbers through captive breeding.
2. Reintroduction in Achterhoek and Twente (eastern Netherlands) by direct translocation from local, larger populations.

The project is a collaboration between Wageningen University, the Dutch Angling Society (Sportvisunie), RAVON, INBO (Belgium), and four Dutch water boards (Hunze en Aa's, Noorderzijlvest, Rijn en IJssel, and Vechtstromen).

Supervisor(s): Hendrik-Jan Megens (ABG), Reindert Nijland (MAE)

Biogeography and conservation genomics of the bullhead (Cottus perifretum/Cottus rhenanus complex)

The bullhead - donderpad in Dutch - is a fish with a personality that does not match its appearance. Despite its tough-looking exterior, it is actually a rather docile fish, quietly minding its own business. It is so good at this, in fact, that it tends to isolate itself in small streams, where unique populations emerge through genetic differentiation, sometimes with distinctive external traits.

In Europe, it was once assumed that there was only one species (Cottus gobio), but this has since been split into 14 species. In the Netherlands,

two species are now recognized, each with its own Red List status and management strategy. Unfortunately, bullheads have come under severe pressure, particularly from invasive gobies that outcompete these far more introverted fish. As a result, bullheads are rapidly disappearing from open waters in the Netherlands. Remaining populations are mainly found in smaller waters, streams, and brooks that are either inaccessible to invading gobies or too cold for them. Small brooks fed by springs and wells therefore remain strongholds for the species, but these local, isolated populations may be too small to remain viable in the long term.

In this project, we aim to:

1. Understand genetic differentiation among regions in the Netherlands, especially between open waters and small streams, with a particular focus on local adaptation and signatures of incipient speciation. This biogeographic perspective is highly important for guiding future population restoration efforts.
2. Understand the role of isolated populations in forming distinct genetic management units and assess the possible role of inbreeding risk in the long-term persistence of local populations.

The main component of the thesis will be the analysis of re-sequencing data, but depending on the time of year, there may also be opportunities to participate in fieldwork. The project is a collaboration between Wageningen University, RAVON, INBO (Belgium), and four Dutch water boards (Brabantse Delta, Vallei en Veluwe, Rijn en IJssel, and Vechtstromen).

Supervisor(s): Hendrik-Jan Megens, Reindert Nijland (MAE)

Paws and pedigrees: evaluating population status and genetic diversity of cat breeds

There are about 40 to 70 distinct cat breeds officially recognised worldwide. Some of these breeds are popular and have relatively large population sizes (e.g. British shorthair, Ragdoll and Maine Coon), whereas others are rare (e.g. Sokoke and Ural Rex). In all such (closed) breeds, there is a risk of losing genetic diversity and inbreeding, and genetic management is important to keep the breeds genetically healthy. In this project, you will use publicly available data to study one or multiple cat breeds. You will investigate the population status, population substructure and estimate levels and rates of kinship and inbreeding. Given that the data are coming from publicly available databases that are maintained by volunteers, an important part of the project will involve checking data quality, completeness (and consider representativeness). Last, depending on the specific interest of the student, we may also try to extract some data on simple traits (when available) and investigate the inheritance.

Supervisor(s): Harmen Doekes

Inbreeding and kinship with pedigree and genomics in the Schapendoes

In dog breeding it is now possible to send in a sample of your dog to a commercial company and receive its inbreeding coefficient based on DNA in return. Potentially, genomically estimated inbreeding coefficients have a number of advantages over pedigree based estimates. They are insensitive to pedigree errors, do not depend on pedigree length, and can determine differences in inbreeding within litters. However, there are different ways to estimate inbreeding coefficients from DNA. In practice, pedigree based and different DNA based methods from multiple companies provide different estimates that vary at different scales. Owners and breeding organisations are generally at loss how to deal with these estimates. For the Schapendoes a database is available with dogs tested at different companies, with some of them by more than one company, plus an excellent pedigree. In this thesis you will make an inventory of the different methods and determine the correlation between them and with pedigree estimates. From a scientific viewpoint it will be interesting to unravel the different ways of estimating inbreeding and their relevance for genetic management. Furthermore, we will explore whether the different methods can be integrated into a single measure for dog owners and breeding organisations, and set up an advice for breeding organisations how to deal with the different estimates.

Supervisor(s): Jack Windig

A holistic population genomic comparison of two stickleback species

The three-spined and nine-spined stickleback are considered supermodels in evolutionary biology because of their tendency to adapt quickly to new environments. Both species are studied worldwide and have high-quality annotated reference genomes of around 460 Mb. Both species also occur across the same environmental gradient (brackish water to freshwater), allowing us to compare the evolutionary forces driving or constraining adaptation, and the underlying genomic basis. Our preliminary results based on phenotypic data and SNPs show that populations of the three-spined stickleback show much stronger adaptation than populations of the nine-spined stickleback (www.nature.com/articles/s41467-017-00256-6). These contrasting patterns are striking given that both species are closely related and are rather similar in ecology.

We hypothesise that the difference between the species is rooted in their genomic architecture, making the three-spined stickleback more sensitive for natural selection, while nine-spined stickleback might respond more to environmental pressures through plastic and regulatory mechanisms. This is what we would like to find out in this study, starting from an in-depth comparison of structural variation. In both species, we will (1) evaluate how structural variants (CNVs, indels, duplications and inversions) are distributed in the genome, (2) compare how these structural variants are connected to ecologically relevant genes, and (3) analyse which of these structural variants are the target of natural selection. Our ambition is to present the first comparative analyses of the role of structural variation for adaptation in two species across the same ecological landscape.

Supervisor(s): Joost Raeymaekers (Nord University), Manu Kumar Gundappa (ABG)

Genetic management of zoo populations living in groups

For species living in groups such as fish in aquariums, birds in flocks and deer in herds, no individual pedigrees are available for zoo populations. Consequently, genetic management based on kinship and inbreeding coefficients estimated with pedigrees is not available. However, alternative strategies such as exchange of animals between zoos according to a breeding circle are available. The effectiveness of these strategies can be assessed with computer simulations. In this project we will investigate – in cooperation with Antwerp zoo Research Centre – the best way to set up genetic management for zoo populations of different size, in different numbers of zoos, etc. and the best way to use computer simulations. Students who are interested in this topic need to have affinity with zoo populations and genetic diversity. You will learn how to translate theoretical computer results to practical management of live zoo populations.

Supervisor(s): Jack Windig

Genetic relationship between free ranging Scottish Highlander herds on the national reserve 'De Hoge Veluwe'

By using GPS data and SNP genotype (100K) data we want to investigate inbreeding and migration between herds.

Supervisor(s): Richard Crooijmans

Inheritance analysis in Tauros cattle: a free ranging cattle population in the Netherlands

We have 5 different sets of SNP data varying from 20k till 770K with in total more than 500 semi-wild cattle individuals with unknown/or incomplete pedigree information. By using different methods we want to reconstruct the pedigree.

Supervisor(s): Richard Crooijmans

Inbreeding when matings are out of control

In many fish species that are grown in aquaculture the reproduction is not completely controlled. Often this means that we can select potential parents, but we cannot decide on the individual contributions that parents make to the next generation and/or which parents will mate together.

The theory and methods that are generally applied to control inbreeding assume that you can make specific matings and keep offspring of those.

In this project you will analyse pedigree data from fish breeding programmes that reproduce fish in breeding groups, and use the results to predict the rate of inbreeding in such a programme. In addition, you can apply this knowledge to design a breeding programme to control the rate of inbreeding at an accepted level. The design work can involve simulation or theoretical derivation of inbreeding rates based on the results from the analysis of the breeding groups.

Supervisor(s): John Bastiaansen

From sequence to sushi: genetic diversity in a newly domesticated fish species

The Yellowtail Kingfish (Seriola lalandi) is one of several fish species that has recently made the transition from exclusively wild-caught to being cultured, and is now on its way to becoming a domesticated species. This species has a biology that is similar to tuna (although no direct 'family relationship'): it lives in the oceans and is very mobile. These fish are found all around the world. The domestication process and related breeding are directed to efficient production and high quality meat. Most of the meat is produced for sushi. There is, however, currently hardly any knowledge on the variation in this species in the wild, or in cultured systems. We recently sequenced the genome of the Yellowtail Kingfish, and conducted physiological experiments and gene expression analyses to gain a better insight in the biology of culturing this species. We have also sequenced 20 Yellowtail Kingfish that were caught from the wild near South Africa from cultured populations in Chili and Australia. This whole-genome sequence data holds a very large amount of information that can give us knowledge on the evolution of the species and on the variation that is present in the populations that are currently in the process of becoming domesticated. Questions that you can work on: what is the effective population size? How much variation is there in the genome? What is the pattern of evolution in protein altering-, versus neutral evolution? And is there reason to fear inbreeding depression when a species that may have a very large population size in the wild is transferred to an enclosed and relatively small breeding system?

Supervisor(s): Hendrik-Jan Megens, John Bastiaansen

Managing genetic diversity in the Friesian chicken

The local Friesian chicken (Dutch: Friese Hoen, Friesian: Fryske Hinne) is an ancient breed from the North of the Netherlands. To safeguard the breed, the breed society was founded 100 years ago. The breed consists of 13 different colour varieties. However, the number of breeding animals is limited, especially within each colour variety. No pedigree is kept, but breeding circles might be an efficient way to limit the inbreeding rate within varieties. The breed society is interested to set up breeding circles. In this project we will look at the population structure (i.e. how many animals within each variety), the inheritance of the colour variety (i.e. what variety is produced when you cross them), and determine the effect of breeding circles on inbreeding rates and how to set them up. Breeding circles will be evaluated with computer simulations mimicking the varieties and their population structure. If you are interested in genetic diversity and safeguarding ancient breeds, this is an interesting project.

Supervisor(s): Jack Windig

Genetic diversity in Groningen White Headed cattle and use of genebank material

The Groningen White Headed (Dutch: Blaarkop) is currently a rare cattle breed in the Netherlands. Due to the small population number, its genetic diversity is under threat. However, its genetic management has been relatively successful, with a low inbreeding rate in 2017. In this project we will investigate the current status of the breed and the role the genebank can play. Is the material stored in the genebank still representative of the live population? How often is genebank material used in the live population? And what if more material would have been used? Moreover, which animals should be sampled to supplement current genebank material? We will use a combination of pedigree analysis to evaluate current diversity

and computer simulations to evaluate future diversity. Answering these questions will not only help the Groningen White Headed, but also help to improve Dutch and international genebank policies. Students with an interest in genetic diversity and ancient breeds can increase their quantitative genetic skills and learn more on conservation genetics.

Supervisor(s): Jack Windig

How to quantify the genetic differences of (sub-)populations or breeds?

There is often discussion about the status of different populations and whether they should be conserved as separate breeds and if exchange of animals should be allowed. Genetic relatedness is a good measure to quantify the difference between two groups of animals. It can be measured using pedigree data or DNA data. However, these two methods tend to vary on a different scale, and there are several ways in which DNA data can be quantified. Moreover, results may depend on what DNA markers are used and their allele frequencies. We are working on a fact sheet that explains the differences between the various methods, what thresholds to use to distinguish between different situations (e.g. one single population/ two related but distinct populations/two separate populations). We will use data from cattle populations stored in the genebank in the Netherlands and computer simulations to evaluate the different measurements. Students will learn more about pedigree and DNA analysis, can improve their computer skills, and will contribute to better safeguarding different breeds in Europe.

Supervisor(s): Hendrik-Jan Megens, Jan ten Napel

Use of DNA-derived inbreeding coefficients in dog breeding

Nowadays dog owners can send in a sample of their dog to a commercial company and get in return – for a hundred dollars or more – a measure of its inbreeding. These measurements can be quite different from the traditional pedigree-based inbreeding coefficients. In principle, DNA measurements can be more precise that pedigree measurements, but do depend on the marker sets used and the allele frequencies. In this project we will use computer simulations where we know the real levels of inbreeding to compare them with inbreeding coefficients estimated using subsets of DNA markers. Next, we will evaluate the effectiveness of genetic management using pedigree or marker-derived coefficients. We will do so in breeds with low rates of inbreeding and in breeds with high rates, and compare with results found in real dog breeds such as the Saarloos Wolfdog and the Stabyhoun.

Supervisor(s): Jack Windig

Loss of function variation in the Friesian Horse

Loss-of-function (LoF) variants represent a particularly impactful class of genetic variation, as they are predicted to severely disrupt gene function. These include mutations such as stop-gained variants (introducing premature stop codons), frameshift mutations (altering the reading frame of a gene), and splice-site variants (affecting proper RNA splicing). Because LoF variants can abolish or strongly impair protein function, they are often associated with deleterious effects on fitness, health, or development.

In this thesis, you will analyse whole-genome sequencing (WGS) data from 50 influential Friesian sires to identify and characterise candidate LoF variants. The Friesian horse is a historic Dutch breed known for its distinctive black coat and elegant conformation, but it is also characterised by relatively high levels of inbreeding due to historical bottlenecks and the extensive use of a limited number of popular sires. This makes it an ideal population to study the accumulation and impact of potentially deleterious variants.

A key component of this project is the careful filtering and validation of candidate LoF variants, as variant annotation pipelines can produce a substantial number of false positives. You will apply stringent quality control and filtering strategies to derive a high-confidence set of LoF variants. These variants will then be evaluated in the context of gene function and potential phenotypic consequences. In addition to WGS data, you will leverage a large cohort of horses genotyped with 70K SNP arrays and

imputed to sequence level. This dataset will be used to validate the identified LoF variants, assess their population frequencies, and investigate their distribution across the breed. Using the imputed sequence data, you will also test for deviations from Hardy–Weinberg equilibrium as a signal of purifying selection, which may help identify deleterious alleles affecting fitness.

Together, this project will generate a curated catalogue of LoF variants in the Friesian horse and provide insights into how deleterious variation contributes to inbreeding-related health and fitness challenges in the breed

Supervisor(s): Martijn Derks, Marije Steensma

Back to contents page

Sustainability and breeding programmes

Hotwings: dual-purpose chickens for tropical environments

Dual-purpose chickens are bred for both eggs and meat, and are therefore a popular choice among smallholder farmers in developing countries. Many of these smallholder farms are located in countries with a tropical climate, while they buy birds from European breeding companies such as Hendrix Genetics (HG). Genetic improvement of dual purpose breeds (such as Sasso from HG) typically takes place in Europe, and is often based on data collected from farms in more temperate (cooler) climates. This contrast between the environments of the smallholder farm and breeding locations may be problematic, because selection favours individuals that are well adapted to the European climate. The welfare and productivity of birds placed in the tropical climate may therefore be at risk, because they are not well-adapted to high temperatures. In this project, you will use data from the breeding company Hendrix Genetics to study genotype-by-climate interactions between hens producing eggs in France and Burkina Faso. In addition, you may study the benefit of using data from Burkina Faso for improving heat tolerance and overall robustness of dual purpose chickens.

Supervisor(s): Pascal Duenk, Naomi Duijvesteijn (HG)

Selection traits for alternative housing systems in laying hens

Cage-free housing systems have been suggested as a welfare-friendly alternative to conventional cage housing of poultry. Such housing systems allow animals to express a wider range of natural behaviours, such as movement in outdoor areas, sand bathing, etc. Modern poultry, however, have been genetically selected for conventional systems and may not be well-suited for cage-free environments. Furthermore, birds housed in cage-free systems experience various behavioural, physical and disease challenges which may negatively affect them. Given the increasingly global trend towards cage-free housing, genetic selection programmes must be adapted to select birds appropriate for such housing systems. This involves the development of relevant cost-effective phenotyping strategies, in which a large number of birds are measured and ranked on their performance. Those animals with the best performance will be selected to parent the next generation of birds, ultimately contributing to improved health and welfare of poultry.

In collaboration with a large avian genetics company, animal welfare specialists at the University of Bern in Switzerland and geneticists at the University of Guelph in Canada, this project aims to develop a series of novel traits associated with movement, locomotion and health and welfare traits in poultry for use in breeding programmes. Sensor tracking technology applied to commercial laying hens, as well as information on a number of conventional traits, will be collected and analysed within this project. You will gain knowledge and skills in animal behaviour, genetics, data analysis and statistics. This project aims to develop novel phenotypes that can be used for genetic selection within laying hen breeding programmes. More reading about the observed behavioural patterns and their association with health and disease can be found here: DNA methylation variation in the brain of laying hens in relation to differential behavioral patterns

Supervisor(s): Pascal Duenk, Henk Bovenhuis, Prof. Christine Baes (University of Guelph and Bern)

Selective breeding for cow-calf contact (CCC) dairy systems

There is an increasing interest in dairy production systems that allow for prolonged period of contact between cow and calf. In those systems calves can suckle with their mother until a specific age and only part of the milk that is produced is made available for sale. Consequently, milk recording – for selective breeding and management purposes – will only record milk that can be sold and the total amount of milk produced by a cow is unknown. Current selective breeding and management systems cannot handle milk production records from these cow-calf contact (CCC) dairy systems. Furthermore, it is unknown if milk production from cow-

calf dairy systems is genetically the same trait as that of conventional dairy production systems.

There are several relevant questions in relation to this topic, for example:

- What should a breeding scheme for CCC dairy systems look like? This can be investigated by performing model calculations.
- Model lactation curves for CCC dairy systems.
- Genetic parameters for milk production and other traits in CCC dairy systems (estimate the heritability for milk production traits; estimate the genetic correlation between milk production during suckling and after suckling; estimate the genetic correlation between milk production in CCC and conventional dairy production systems).

The aim of this thesis will be determined after discussion between the student and supervisor(s). This is a combined thesis for Animal Production Systems (APS), Animal Breeding and Genomics (ABG), and Adaptation Physiology (ADP).

Supervisor(s): Eddie Bokkers, Henk Bovenhuis, Ariette van Knegsel

Do organic and circular dairy farming require different genetics?

Are cows that show their best performance under conventional farming conditions the same as the ones that show excellent performance under organic farming conditions? In other words, is there genotype by environment (GxE) interaction? This has been the topic of an ongoing debate. Nauta (2009) estimated a relatively high genetic correlation between milk yield under conventional and organic production of 0.80. Standard errors of the estimated genetic correlations were high, which did not allow drawing firm conclusions. Since the work of Nauta (2009) the number of organic farms in the Netherlands has increased, allowing for more accurate estimates of genetic parameters. In addition, quantifying genotype by environment interaction will benefit from an accurate definition of 'environment' which has been problematic in previous studies as there are considerable differences among organic farms as well as among conventional farms, for example in the use of concentrates. The excretion indicator (BEX) is used to calculate specific N and P excretion from dairy cattle on farm level. This database contains detailed information on several farm characteristics which can be used to better characterise farms (i.e. the environment). How differences in the use of concentrate interact with genetics is a relevant topic in view of the switch to circular agriculture. The aim of the current study is to quantify genotype by environment interaction for dairy farms in the Netherlands. The first project will focus on the difference between organic and conventional farming conditions. The second project will focus on differences in the use of concentrates.

Literature: Selective breeding in organic dairy production

Supervisor(s): Henk Bovenhuis

How do cultured fish keep their shape?

Fish for aquaculture are typically selected for growth. Selection for growth can have effects on the shape of fish. In this project we will use image analysis to investigate the genetics of shape and appearance. For this project we collaborate with a fish breeding company in Greece that produces sea bream and sea bass. The company is taking photographs of all the fish in their breeding programme at multiple times during their growth. The shape and appearance of the fish, i.e. the exterior traits, are very important because they are sold as whole fish. In horses and cattle it is common to give scores for the exterior of the animals and to breed for this. Fish are also scored in a similar way, but obviously for different traits as horses, cattle or dogs. The important traits in fish are roundness, colour, belly shape, etc.

In this project you will analyse the digital images of the animals to develop smart and objective measures to describe the exterior of the fish. In a second phase you try to estimate heritabilities for your new measures and investigate how they correlate, genetically, to the current scores given by human judges. The aim is to develop measures that can be included in the breeding programme in Greece to improve the appearance of the fish.

Supervisor(s): John Bastiaansen

Growing fish also healthy?

Fish is generally known to be a healthy food for humans. To become a healthy food, farmed fish should be healthy itself. In this study we will investigate the genetic relationship of growth and health of the fish. For this project health is measured as the condition of several organs, like the gills, the heart and the liver that play an important role in supporting the growth of fish. For instance by supplying oxygen, or by metabolising fatty acids. These organs are expected to be good checkpoints for the overall health of the animal. For this study a large number of fish from a breeding population have been dissected and the health of their organs was scored by a veterinarian. The heritabilities of these scores as well as their correlations with growth will be estimated by you. In addition, you will investigate whether difference in environmental conditions such as the season have an impact on the variation that is observed in organ health. From your results you will make recommendations on what is needed in a breeding programme to make sure the fish stay healthy while selecting for improved growth.

Supervisor(s): John Bastiaansen

Fish reproduction

At ABG, we study the reproductive physiology of fish in nature and aquaculture and develop innovative methodology for solving reproduction bottlenecks of captive fish. Research topics include sexual maturation, fertility, and larval development. Closing the production cycle of species difficult to propagate (e.g. European eel) is a focus area. Within this line of research we always have projects running within which up-to-date thesis topics can be formulated.

Supervisor(s): Arjan Palstra

Swimming of fish and use for farming fit fish

At ABG, we study the swimming physiology of fish in nature (e.g. migration) and aquaculture and develop innovative methodology for farming fit fish. Research topics include the use of swimming for phenotyping, for inducing (muscle) growth, for improving health and welfare, and for studying the energy economy with sensor technology. Within this line of research we always have projects running (e.g. on seabream, yellowtail kingfish, tilapia, zebrafish within which up-to-date thesis topics can be formulated.

Supervisor(s): Arjan Palstra

Social interactions, breeding for diversity and competition between plants

Global biodiversity is declining rapidly, one of the reasons being the widespread use of pesticides. Increasing biodiversity may be a natural way to make agriculture more resilient to pests, but this will also require other breeding strategies. One way to combat pests by increasing biodiversity is the use of mixed cropping, where a field consists of a mixture of different crop species. This system is also known as 'intercropping', and a well-known example is the combination of maize and faba bean. Genetic improvement (breeding) for such systems introduces new challenges, such as heritable competition between the two species. In an ideal system, taking maize and faba bean as example, genetic improvement of the maize should also benefit performance of the faba bean, and vice versa.

This topic is very closely related to the quantitative genetic study of social interactions in animal breeding, a field where ABG has a lot of expertise. In this thesis project we aim to integrate knowledge on social interactions coming from the field of animal breeding with systems for intercropping in plant breeding. An important objective is to design and evaluate experimental setups for the estimation of direct and social breeding values in intercropping systems. This will be done at least by computer simulation, and potentially also in the field. You will develop programming skills (e.g. R or Python) and data analysis skills, and become familiar with the application of quantitative genetics in plant breeding (for which er are currently excellent job opportunities).

Supervisor(s): Piter Bijma (ABG), Peter Bourke (PBR)

Genetic analysis of behavioural traits in poultry

Livestock are social animals and they are frequently engaged in social interactions. Some of the interactions can be harmful, such as aggressive pecking in poultry. Aggressive pecking can lead to 15-20% of mortality in modern group-housed poultry farms. Evidence has shown that aggressive pecking behaviour is heritable, so there is a potential for genetic improvement. Before we can actually breed for the behaviour, a few questions should be addressed. For example: How can we phenotype the behaviour? How can we best analyse the collected data to estimate breeding values? If we are able to answer these questions, we will be one step closer to further improvements in efficiency and animal welfare in poultry. In this project, you will have the chance to do either simulation study or real data analysis on feeding behaviour, and thereby improve our understanding of the genetics of behavioural traits, using turkeys as an example.

Supervisor(s): Zhuoshi Wang, Harmen Doekes, Piter Bijma

Exploring the influence of goat phenotypes on production efficiency

In recent years, the production of goat species has grown significantly, increasing its importance for agricultural development in the twenty-first century. Different morphological characteristics, such as beard, colour, or ear position, can affect the performance of goats in terms of milk production, meat production, and reproductive efficiency. Further research is needed to fully understand the effects of these morphological characteristics on goat production traits. The aim of this thesis is to study the effect of goat phenotypes such as beard, colour, and/or ear position on goat production related traits.

Supervisor(s): Richard Crooijmans, Rayner Gonzalez-Prendes

Investigating the connection between milk protein and oligosaccharides in goats

Milk protein and oligosaccharides are two important components of goat milk. Milk protein content can affect the nutritional value of goat milk, as well as its functionality in food applications. Goat milk oligosaccharides have unique characteristics that make them an attractive choice for improving human nutrition. However, goats have been significantly less studied than other livestock species. The aim of this thesis is to study the relationship between protein content and oligosaccharides composition in goat milk.

Supervisor(s): Richard Crooijmans, Rayner Gonzalez-Prendes

Cryopreservation of eel sperm

We would welcome a student who is interested in investigating by literature review how cryopreservation of eel sperm (and perhaps even eggs, embryos and gonads) can be improved. Worldwide, eel populations have decreased strongly in numbers since the 1970's. The existing eel farms still depend on the catches of glass eels in nature which are then raised to market size. Only a restricted number of glass eels is available for aquaculture and societal concern exists about the lack of sustainability. Successful reproduction in captivity could supply aquaculture with glass eels and close the production cycle. This way, both eel aquaculture as well as management of the natural populations could become sustainable. The current status of European eel reproduction is that some research groups can produce larvae regularly which are then kept alive up to about 20 days post hatching. However, the larvae cannot be fed yet, which means they die around that age. Problematic is still the high individual variation in maturation response of female eels. Therefore it would be useful to have good quality sperm storage available. What we need is a review on cryopreservation methodology in fish, specifically eels. This review would provide a status update and recommendations on how to improve the current protocols. Your improved protocol will then be validated by us and compared with the currently available protocol and the use of fresh sperm.

Supervisor(s): Arjan Palstra, Julie Lamy, Pauline Jéhannet

How does genetic change in pig breeding translate to pig farms in practice?

Animal breeding is an important tool to improve livestock populations. Historically, breeding has resulted in substantially increased productivity. Today, breeding efforts are focused largely on animal health, survival and welfare traits However, we often see that genetic improvement created in breeding programmes is not fully expressed in livestock production practise. In other words, the genetic improvement that we see in production farms is often smaller than the improvement we see in the breeding programme. But why does this happen? Why does actual progress fall short of expected progress? Answering this question is important to better use animal breeding for sustainable genetic change. The realised change on production farms depends on both the genetic change, but also on changes in animal management in production farms, such as in the feeding regime and the husbandry system, and also on potential genotype by environment interaction between the breeding programme and the production herds. At present, we do not understand the contributions of these different factors.

In this thesis project, you will try to understand why realised improvement falls short of expected improvement in pigs. You will use data provided by the Dutch pig breeding company Topigs Norsvin, and will collaborate closely with Topigs Norsvin in this research. Hence, critical analysis of existing data will be your main tool to answer the above research question. This involves both estimating the impacts of genetic change on realised progress, but also investigating impacts of other changes, particularly in the feeding system. For this thesis project, we are looking for a student with a good natural ability for doing calculations, an ability to link theory to practise, and good common sense. In this thesis, you will improve your programming and data handling skills (e.g. in R). Because the project will be executed in close collaboration with Topigs Norsvin, you will develop a professional network in the field of animal breeding, which is useful for your career after finishing your MSc. If you perform an excellent job, this thesis project may result in a scientific publication.

Supervisor(s): Piter Bijma, Egbert Knol

Population management tool for captive hamadryas baboon populations

Hamadryas baboons live in an exceptionally complex multilevel society in which the one-male unit (OMU) forms the core of the social structure, multiple OMUs combining into clans, and several clans grouping together into larger bands. This social organization is strongly influenced by factors such as harem size, age structure, sex-specific reproductive maturity, and kinship between individuals. In captivity, zoological institutions aim to replicate these natural social structures as closely as possible; however, imbalances in group composition can result in social instability, increased aggression, and reduced animal welfare. Although population management advice is generally approached at population levels (thus across zoos) limited attention is currently given to the advice on institution level for these socially complex primates. Developing a population management tool that integrates natural OMU dynamics, age-related changes in harem size, and reproductive parameters could support institutions in creating more stable and socially appropriate captive groups.

Research question: How can a population management tool integrate demographic and social parameters (such as age-specific harem size, sex-specific reproductive maturity and sex-related mortality) improve social stability and reduce aggression in captive hamadryas baboon populations?

Sub-questions:

- Which demographic and social parameters are essential for modeling stable groups in captive hamadryas baboons?
- How does harem size change with male age in hamadryas baboons, and how can this information be incorporated into population management strategies?
- What role do sex-specific differences in reproductive maturity and moratility play in the stability of captive hamadryas baboon groups?
- How can a population management tool assist zoological institutions in predicting and preventing social unrest and excessive aggression?

Supervisor(s): Harmen Doekes, Bart Ducro, Heleen Post-van Engelendorp

Gastelaars (Dierernpark Amersfoort)

Population management plan for captive prairie dog populations

Prairie dogs are highly social rodents that live in complex colony systems characterized by strong social bonds, cooperative behaviors, and structured family groups known as coteries. In the wild, these coteries typically consist of one or several males, multiple females, and their offspring, forming stable social units within larger colonies. In captivity, maintaining socially compatible groups can be challenging, as inappropriate group composition, skewed sex ratios, or disruptions in age structure may lead to increased aggression and reduced welfare. Despite the importance of social organization in prairie dog societies, population management in zoological institutions is often focused primarily on demographic sustainability rather than social stability. Developing a population management plan that incorporates natural social structures, reproductive parameters, dispersal tendencies, and age-specific group dynamics may help institutions establish more stable and welfare-oriented captive populations.

Research question: How can a a population management plan/tool integrate social structure, age composition, dispersion and reproductive parameters to improve social stability and welfare in captive prairie dog populations?

Sub-questions:

- Which social and demographic parameters are essential for modeling stable captive prairie dog groups?
- How do sex ratio and age composition influence aggression and social cohesion in captive prairie dog colonies?
- What role do natural dispersal patterns and reproductive dynamics play in the long-term stability of captive populations?
- How do natural dispersal patterns and reproductive dynamics play in the long-term stability of captive populations?

Supervisor(s): Harmen Doekes, Bart Ducro, Heleen Post-van Engelendorp Gastelaars (Dierenpark Amersfoort)

Back to contents page

Big data and data sciences

Creating interactive visualisations of biodiversity in the Netherlands

This project will focus on developing interactive tools to visualise biodiversity patterns across the Netherlands. You will compile existing datasets, such as genetic diversity metrics, and environmental DNA (eDNA) records and use them to build spatial visualisations using R Shiny or similar interactive frameworks. The goal is to translate complex biodiversity data into clear, user-friendly maps and dashboards that allow exploration and data handling. You will develop practical skills in spatial data processing, interactive visualisation, and simple app development, while applying these skills to real-world biodiversity datasets. The final outcome will be a functional prototype tool that demonstrates how interactive data visualisation can support biodiversity research, conservation planning, and environmental decision-making.

Supervisor(s): Chris Barratt

A fish with a personality: identifying individual weatherfish through image analysis

The weatherfish (Misgurnus fossilis) is a fish species typical of Dutch freshwater wetlands. Historically, the species could move freely throughout the floodplains of the major rivers during periods of high water. However, since flooding is no longer possible, the species is now restricted to relic populations.

Inbreeding is related to the number of breeding individuals in a population. But what is a population? How much do fish move between areas, taking their genes with them? The viability of a population therefore depends on both numbers and migration. Since inbreeding appears to be rather common in weatherfish, local populations may be small and movement between them may be limited.

Weatherfish can become quite old, up to 20 years. The black-and-gold scale patterns of these fish offer an interesting opportunity: identifying individuals over multiple years, even as they grow. Recognizing individual fish from photographs would make it possible to monitor population demography, without invasive procedures such as PIT-tagging, reducing both costs and animal welfare concerns.

In this study, we aim to develop a computer vision-based approach to identify individual weatherfish. Fish will be captured over several days in the nature reserve Het Klooster, near Nieuwegein, Utrecht. This area was recently developed as a nature compensation site (2018). The newly established weatherfish population has since grown substantially and is very well studied. Although the population appears stable at present, concerns remain about its long-term viability due to inbreeding.

By collecting an extensive set of photographs, we will create a training dataset for machine learning-based identification. Part of the work will involve developing standardized procedures for field-ready photo recognition. This will go hand in hand with identifying which features of the fish are most suitable for robust recognition - head, belly, and sides all have distinctive patterns, but which is best suited for reliable recognition and easiest to use in the field? To test the accuracy of identifications, we will also collect genotype data from skin swabs.
The project therefore integrates fieldwork, computer vision, and genetics. Depending on your interests, it can be tailored to fit your profile. The project is a collaboration between Wageningen University and RAVON (Reptielen-Amfibieën-Vissen-Onderzoek Nederland).

Supervisor(s): Hendrik-Jan Megens (ABG), John Bastiaansen (ABG), Jelger Herder (RAVON)

Big data from a school of fish: a computer vision system to extract data from images of fish

In animal breeding, one of the biggest challenges is to measure many phenotypes efficiently and accurately. Measuring animals is time consuming, and measurements taken by humans have a limit on how accurate they are and how many animals can be measured cost-effectively. In fish breeding, regular imaging of fish is becoming the norm, to facilitate semi-automated phenotyping of the animals. Although improvements can be made by this approach, it does not take full advantage of the image data. Rather, what is needed, is a fully automated image analysis approach, that can take many more measurements. In addition, automated systems result in easily scaling up in numbers of animals measured without increasing labour costs that much, and in higher reproducibility of the measurement. From a scientific perspective the exciting prospect is to capture quantitative aspects of the phenotype that can only be measured subjectively by humans. How do you quantify, for instance, 'shape', from looking at a picture or when only taking 2 or 3 manual measurements? A computer vision system can measure hundreds of points in a reproducible way.

In this project you will develop a prototype of a fully automated fish phenotyping programme. The programming part will be done using Python (a general and very popular programming language) and OpenCV (a specialised computer vision library that can be used in Python). However, more importantly, you will investigate and apply new ways to define shape and growth parameters that can be derived from your software. Depending on your interests, this can be approached either from biologically informed phenotype definitions, such as lateral line shape, operculum shape, etc., or from more general 'big data' type approaches, collectively called 'Machine Learning'. Interested students will have the opportunity to learn basic programming in Python.

Supervisor(s): Hendrik-Jan Megens, John Bastiaansen

Monitoring animal behaviour with accelerometers and/or video

Automated monitoring of animal behaviour is highly relevant for both management (e.g. to be able to intervene timely when needed) and breeding purposes (e.g. to improve phenotyping and subsequent genetic improvement of health, resilience and efficiency traits). In this project you will develop methods for behavioural monitoring and/or analyse behavioural data. The focus is on infectious disease studies, in which animals are infected with pathogens (e.g. with avian influenza virus) to investigate, for example, disease transmission or vaccine effectiveness. You can either focus on general activity patterns, or on specific behaviours such as drinking or feeding behaviour. You will analyse accelerometer data and/or video data of, for example, sheep or chickens. The accelerometer data consists of 25 Hz accelerations in x-, y- and z-direction, and can be used to estimate an activity index or possibly more advanced metrics. The video approach uses Aruco markers (simple QR-codes) to determine the position of the animals over time and, based on those positions, you can estimate activity patterns. In addition, based on proximity and movement at a drinker/feeder, you can estimate the time spent drinking feeding. Depending on your interests, you could work on more methodological questions or on interpretation of behavioural patterns based on the accelerometers and/or video data. Thereby, you can contribute to reduction and refinement of animal experiments. This is a collaboration project with Wageningen Bioveterinary Research.

Supervisor(s): Harmen Doekes

The history of Dutch Large White breeding lines, a genomic perspective

The modern commercial Large White breeding lines of the pig breeding company Topigs Norsvin have been created out of the consolidation of a number of lines from a variety of Dutch breeding organisations around the turn of the last century. Hence, the current large white breed is a genetically diverse breed, influenced by a mix of various ancestral Large White/Yorkshire breeds. During the consolidation phase, which resulted in merging of populations and phasing out of other populations, sperm of breeding boars was deposited at the Dutch Centre for Genetic Resources (CGN). Recently, we sequenced the complete genomes (WGS) of 42 boars from eight breeding lines with a Large White background, stored in the CGN genebank. Most sequenced boars were stored approximately 20 years ago (1995-2002), but also more recent samples have been sequenced (2010-

2020). Moreover, within our department we have access to a large dataset of the current elite Large White breeding lines from Topigs Norsvin. This dataset opens new possibilities to investigate the genetic development and diversity of the Large White populations over the last few decades. More precisely, you will investigate changes in allele frequency, the genetic load (inbreeding risk), and inbreeding levels, in time. The goal of the project is to get a genomic perspective of the Large White breeds over the last decades, and assess the consequence of consolidation of breeding lines (e.g. loss of genetic diversity).

Supervisor(s): Martijn Derks

Back to contents page

Alternative options

Didn’t find what you were looking for? There are some alternative options for thesis projects at ABG that are not listed in this document.

These options include:

- Propose your own topic
- Write your thesis in collaboration with a different chair group
- Write your thesis in collaboration with a company
- Write your thesis at another university (even abroad!)

Note that for all the above options, you should take the initiative in contacting collaborators. We can facilitate your search by providing contact persons that could be useful. All projects should be at an academic level, and contain either

1) Analysis of existing data; the student should either provide data (open source, from a previous project, from a breed organisation he/she/they are a member of, etc.), or have a specific and realistic dataset in mind that we have available;

2) Simulation study; the student suggests a research question and possible approach;

3) A meta-analysis of results in literature (on which they also perform some data analysis). A literature review without (quantitative) analysis is not sufficient for a MSc thesis.

If you are interested in any of these options, please contact the MSc thesis coordinator.

Contact information

Bastiaansen, John	john.bastiaansen@wur.nl
Bijma, Piter	piter.bijma@wur.nl
Bonifazi, Renzo	renzo.bonifazi@wur.nl
Bosse, Mirte	mirte.bosse@wur.nl
Bourke, Peter	peter.bourke@wur.nl
Bovenhuis, Henk	henk.bovenhuis@wur.nl
Calus, Mario	mario.calus@wur.nl
Crooijmans, Richard	richard.crooijmans@wur.nl
Derks, Martijn	martijn.derks@wur.nl
Doekes, Harmen	harmen.doekes@wur.nl
Duenk, Pascal	pascal.duenk@wur.nl
Ellen, Esther	esther.ellen@wur.nl
Godia Perello, Marta	marta.godia@wur.nl
Gonzalez-Prendes, Rayner	rayner.gonzalezprendes@wur.nl
Guennoc, Eugénie	eugenie.guennoc@wur.nl
Gundappa, Manu Kumar	manu.gundappa@wur.nl
Hegedus, Bernadett	bernadett.hegedus@wur.nl
Heuven, Henri	henri.heuven@wur.nl
Madsen, Ole	ole.madsen@wur.nl
Megens, Hendrik-Jan	hendrik-jan.megens@wur.nl
Mencarelli, Angelo	angelo.mencarelli@wur.nl
Mulder, Han	han.mulder@wur.nl
Osorio Gallardo, Tzayhri	tzayhri.osoriogallardo@wur.nl
Palstra, Arjan	arjan.palstra@wur.nl
Sluis, Malou van der	malou.vandersluis@wur.nl
Taghavi, Marjaneh	marjaneh.taghavizavizadeh@wur.nl
Vandenplas, Jeremie	jeremie.vandenplas@wur.nl
Wientjes, Yvonne	yvonne.wientjes@wur.nl
Windig, Jack	jack.windig@wur.nl
Yu, Yun	yun.yu@wur.nl