Animal Breeding and Genomics (ABG) MSc thesis topics

Table of contents

Quantitative Genetics

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

Genomic selection has revolutionized 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 nationallevel 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 populations-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 inflenced 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 Bonifzi

Modelling erosion of genomic breeding values

Genomic selection has been implemented in most important breeding programs in the last decade. With genomic selection, animals are selected based on genomic breeding values relatively accurately in early life, before performance is recorded on the animals themselves or on close relatives. Genomic breeding values are estimated using estimated SNP effects. Current genomic breeding value estimation models assume that SNP effects are constant across the years spanned by the data that is used. This assumption is likely violated in practice, which implies that estimated SNP effects and genomic breeding values erode across generations.

The aim of this thesis is to develop and validate a model that is able to explicitly model and estimate the extent of this erosion. To enable validation of the model, the work will be based on simulated data.

Supervisor(s): Mario Calus

What is the impact of genomic selection on genetic diversity?

Selection in a livestock population 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.

For this thesis, 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): Mirte Bosse, Yvonne Wientjes

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 programs 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, the student will investigate the impact of the different parameters on the long-term 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. The student 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: breeding programs for racing pigeons

Homing pigeons are bred and kept by pigeon enthusiasts for participating in racing contests. In these contests, the pigeons are released from a location far away from their homes (lofts), and points are awarded based on the speed at which they return home. Homing speed seems to have a strong genetic component, because breeders have been able to improve performance over the years by means of selection. However, genetic trends and genetic parameters (such as heritability) for homing speed have never been estimated. One of the problems that pigeon breeders face, is the loss of young birds during training. Birds can, for example, be lost due to predation, accidents, or the inability to find their way back home. In the last 10-20 years, the fraction of young birds that are lost seems to have increased, and a clear explanation for this trend is missing. Inbreeding and unfavorable genetic correlations might play a role here, but this has not been investigated. In this project, you will work on data provided by a pigeon breeder. This data contains a well-document pedigree, and results from racing contests of the last 13 years. From this data, you could study past and current inbreeding levels, or estimate genetic parameters and breeding values for racing performance or 5 survivability. Which of these research questions you will work on depends on your personal interest.

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 5 commercial dairy farms.

In a combined research project of Animal Breeding and Animal Nutrition, 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, Eugenie 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 6 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

Improving genomic predictions by using individualized reference populations

Genomic prediction typically uses a single reference population of genotyped and phenotyped individuals, to estimate breeding values for selection candidates that are only genotyped. Over time, as data accumulates, the reference population becomes increasingly heterogenous. As a result, the differences in allele frequency and linkage disequilibrium patterns between the reference population and selection candidates can limit prediction accuracy. It may therefore be beneficial to use only part of the reference population. But how do we decide which animals to use? One can imagine that the subset of the reference population that maximizes accuracy differs between selection candidates. In theory, for each selection candidate, it would be beneficial to only include animals in the subset that are closely related. Lopez-Cruz and de Los Campos (2021) proposed an approach that identifies, for each selection candidate, an optimal subset from the reference population (i.e. a set of support points) from which predictions are derived. They showed that this approach works well in plants (maize).

The objective of this project is to determine in which scenarios this algorithm could be used. Is it also useful in livestock populations? The secondary objective is to understand how and why the algorithm works. You will learn about genomic prediction, programming, and how to analyse large datasets (data science).

Literature: Optimal breeding-value prediction using a sparse selection index. Lopez-Cruz M., de Los Campos G. Genetics. 2021 May 17; 218(1):iyab030. doi: 10.1093/genetics/iyab030.

Supervisor(s): Henri Heuven, 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 programs 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

Improving decisions for selection based on useful genetic variance

Genomic prediction models allow estimating SNP effects. Using these estimated SNP effects may be beneficial for diversity management that is targeted at increasing the genetic variance in commercial breeding programs. Increasing the genetic variance ultimately allows to realize larger genetic gains in the future. This concept is currently developed in ABG and is different to traditional inbreeding management which aims at maintaining diversity regardless of the underlying effect on the trait. The complexity to make selection decisions when considering estimated effects increases the more animals are considered.

In this project, you will investigate and develop approaches rooted in quantitative genetics theory, to reduce this complexity so that breeders can make better decisions faster. You will improve your programming skills and learn how to use high performance computing and simulate data. You will also improve your understanding of inheritance, quantitative genetics and operational breeding considerations.

Supervisor(s): Tobias Niehoff, Mario Calus

Combining long-term and short-term diversity management

Genomic prediction models allow estimating SNP effects. Using these estimated SNP effects may be beneficial for diversity management that is targeted at increasing the genetic variance in commercial breeding programs. Increasing the genetic variance ultimately allows to realize larger genetic gains in the future. This concept is currently developed in ABG and is different to traditional inbreeding management which aims at maintaining diversity regardless of the underlying effect on the trait. Because breeding goals may change in the future, only focusing on the genetic variance of a trait that is of interest in the present is likely not sufficient to stay competitive.

The aim of this project is to investigate concepts how breeders can maintain more useful diversity for short-term gain and at the same time preserve enough diversity in the long-term. You will learn how to express the value of animals for short-term gain and how to perform optimum contribution selection. You will improve you programming skills, learn how to use high performance computing and simulate data.

Supervisor(s): Tobias Niehoff, Mario Calus

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

Tail biting is a common problem on commercial pig farms. Tai 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 anecdoteds 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, the student 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. The student will also get acquainted with ASReml a program frequently used in animal breeding for variance component estimation.

Supervisor(s): Piter Bijma, Bernadett Hegedus

If you are a victim in early life, will you always stay a victim? Correlation of tail damage at different life stages 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 reduced. 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 tail biting. Some studies have also shown the importance of genetics for the development of tail biting. Tail biting 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.

Data is available on tail damage status for around 2000 animals at three stages of their lives. Here you will investigate whether tail biting is the same trait in different stages of a pig's life. In others words, if you are bitten once, are you a victim for the rest of your life? You will estimate heritabilities for these three stages and the phenotypic and genetic correlations between them. This is valuable knowledge for the breeding company as it helps them decide whether it is enough to select for only one of the traits. If tail damage turns out to be the same trait in all stages then the company only needs to record the trait once. While working on this project you will learn to handle data, improve your R-skills, and learn to work with software for estimating heritability. Therefore, we are looking for a student who is keen on improving their data science skills. The software used will be ASReml, a program which is frequently used in animal breeding.

This project is in close collaboration with the breeding company Topigs Norsvin, so you will have the opportunity to visit the company and also present the final results there. Depending of the results, this thesis project could furthermore result in a scientific publication.

Supervisor(s): Piter Bijma, Bernadett Hegedus

Estimating heritability of pilling 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 initializing 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. The student 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.

Literature: Winter, J., Toscano, M. J., & Stratmann, A. (2021). Piling behaviour in Swiss layer flocks: Description and related factors. Applied Animal Behaviour Science, 236,105272 https://doi.org/10.1016/j.applanim.2021.10527

Supervisor(s): Piter Bijma, Tzayhri Osorio

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 programs, 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

Consequences of enhancing PRRS resistance in pig populations through selection at the GBP5 locus

Porcine Reproductive and Respiratory Syndrome (PRRS) remains a significant challenge in swine production worldwide, leading to substantial economic losses. The identification of genetic markers associated with increased PRRS resistance offers a promising opportunity for improving disease resilience in pig populations. Among these markers, the GBP5 SNP (formerly known as the WURSNP) has emerged as a key determinant, exhibiting a substantial impact on PRRS resistance (Koltes et al 2015, Dunkelberger at al. 2017).

This MSc project aims to investigate the consequences of selection pressure on the GBP5 SNP, its implications for genetic diversity, and the consequence on other traits. The primary objective is to assess how genetic diversity has evolved in the GBP5 SNP region following the implementation of selection for PRRS resistance via the GBP5 SNP, focusing specifically on changes in allele frequencies and haplotype diversity. To achieve this, a comprehensive analysis will be conducted using genomic data (genotypes and whole genome sequence) from four pig populations that were subject to GBP5 selection (2017-2024). More specifically, the student can evaluate population parameters, including nucleotide diversity, heterozygosity, and genetic differentiation over time. Furthermore, the student can explore potential tradeoffs or unintended consequences associated with selection for PRRS resistance. By examining correlations between the GBP5 SNP and other economically important traits, such as growth performance and reproductive efficiency. Overall, this study will provide valuable insights into the long-term effects of selection for single gene variants on genetic diversity and phenotypic traits in commercial pig population.

Supervisor(s): Martijn Derks, Yvonne Wientjes

Exploring the potential applications of randomized 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 programs globally. However, the rapid growth of genomic datasets, which can exceed 1 million genotypes, poses significant computational challenges. Randomized linear algebra has emerged as a promising tool for analyzing 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 randomized 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.

Throughout this thesis, you will become familiar with the theory of randomized linear algebra and gain practical experience with genomic prediction, programming languages (e.g., Julia, Fortran), and high-performance computing.

Supervisor(s): Jeremie Vandenplas

Genomics

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 characterize the CNV regions by looking at overlapping genes. You will 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): Martien Groenen, 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 rumpless 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 rumpless 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 Boolstaart 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 of 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

Studying the transcriptome of turkeys (Meleagris gallopavo) at different developmental stages

The transcriptome is very dynamic and differs extensively between tissues and developmental stages. However, transcriptome studies were mostly performed on adult tissues and there is still a lack of knowledge in the developmental stages including different embryo stages. In this project you will examine the transcriptome of developing turkeys in day 5, day 7, and day 9 of development. Leveraging the Isoform Sequencing (Iso-Seq) technology, the thesis delves into the landscape of alternatively spliced variants and differential gene expression patterns. Iso-Seq is a state-of-the-art long-read sequencing method, to capture full-length transcript isoforms, thus offering insights into the transcriptomic dynamics and alternative splicing.

By examining the embryonic transcriptome at these time points, the study examines dynamic changes in alternative splicing events and identifies key genes associated with developmental transitions. This research contributes to a deeper comprehension of the molecular mechanisms underlying turkey embryogenesis and enhances our understanding of gene regulation and expression diversity in turkey, the second largest contributor to world poultry production.

Supervisor(s): Martijn Derks

Assessing allelic specific expression in porcine muscle and its influence on meat quality

Differences between the expression of the two alleles of a gene is known as allelespecific expression (ASE), an event commonly found in the transcriptomes of mammals. For discovering ASE variants, it is needed to integrate transcriptomics and genotype data. Interestingly, we have a population of pigs that are a result from a cross between two pure lines. That means that perhaps we might encounter ASE in certain variants that are originally from one of the two breeds.

It is expected that the student knows how to work in UNIX. The student will perform bioinformatic analyses in a porcine dataset that includes genotyping data from the Axiom™ Porcine Genotyping Array (targeting 660K SNPs) and 100 RNA-seq datasets of muscle tissue. The student will perform extensive bioinformatic analyses including mapping, SNP calling, filtering and functional analysis. At the end, the student will extract which are the biological functions of those genes that contain ASE and identify if there is a potential link with meat quality characters.

Supervisor(s): Marta Godia Perello, Martijn Derks

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, you 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 a missing link to fertility

The negative energy balance (NEB) characterized by high serum non-esterified fatty acid (NEFA) concentrations, due to increased mobilization 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 induce 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 postfertilization 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 unsaturated fatty acid and appears to protect the oocyte against elevated levels of saturated NEFA (Aardema et al., 2017). Interestingly, in dairy cattle two 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 hypothesize 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, Rayner Gonzalez-Prendes

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 analyze potential detected genes.

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

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 Commission on Zoological Nomenclature suggested a taxonomic classification i.e. *Bos mutus* (Przewalski, 1883) for wild yaks and B*os 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 herders in the high mountains of Central Asia. The traditional habitat area extends through southern part of Russia, north and western Mongolia, Kyrgyzstan, Tajikistan, Afghanistan, north-western 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, Rayner Gonzalez-Prendes 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 recognized, based on morphological and behavioural 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 are going to assemble a new reference genome based on nanopore long read sequences and Ilumina 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, Rayner Gonzalez-Prendes
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, Rayner Gonzalez-Prendes
and Martijn Derks

The role of methylation in determination of fish egg quality

We would welcome a MSc student who is interested in investigating the role of methylation in the determination of egg quality in fishes, by literature review and by sampling analyses. The link between epigenetics and gamete quality has received little attention in fish. However, several recent studies have revealed very specific methylation patterns of several gene promoters. Samples will come from our reproductive model species European eel. Worldwide, eel populations have decreased strongly in numbers since the 1970's. The existing eel farms still depend on 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. But eels cannot be reproduced in captivity yet. Only by hormonal treatment larvae can be produced, but they die before feeding exogenously. The major biological bottleneck is the inferior egg quality and the potential role of methylation is still unclear.

Supervisor(s): Arjan Palstra, Ole Madsen

Molecular fingerprinting of exercise effects in fish muscle

We would welcome a MSc student who is interested in investigating the molecular fingerprint of exercise effects in the muscles of fishes, by literature review and by analyses of our own RNAseq database. In many fishes, exercise stimulates growth by building muscle (hypertrophy) but interspecific variations exist. Over the years, we have executed long-term swim training experiments with the vertebrate model zebrafish and with aquaculture species such as Gilthead seabream and Yellowtail kingfish. In all these cases, swimmers were longer and heavier. We collected white muscle samples and performed RNAseq to determine the genes that are important in exercise-induced muscle growth. We would like you to do a comparative analysis and determine key genes and their function in this process.

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

Full genome analysis of the African buffalo

Various studies on basis of microsatellites show genome-wide positive selection of deleterious alleles. This positive seleciton is related to a sex-ration meiotic drive. Preliminary analysis of buffalo genomies indicates associations between these deleterious alleles and GC content. Earlier microsattelite 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 analsysis. Some 30 full genome sequences are available (from the genebank).

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

Genome-wide characterization 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 uncharacterized 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, biomineralization, 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 characterizing 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 candidate 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, i.e., 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 (ABG30806) or Genetic Improvement of Livestock (ABG31306). 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 Sa

Genetic diversity

Genetic diversity across local Dutch breeds

In the Netherlands we have a number of unique rare cattle breeds. Efforts have been made to preserve these breeds for the future. Their DNA is stored in the Dutch gene bank, and breeding strategies to maintain genetic diversity in the living population are in place. The breeds look very different, but besides those breed characteristics we know little about the genetic diversity across the different breeds. With whole genome sequence data we can investigate which regions in the genome differ significantly between the breeds and which genes are located in those regions that may explain differences in phenotypes. Are you interested in working with whole genome sequence data on genomic diversity of Dutch local breeds? You will learn to handle whole genome sequence data, to estimate genomic diversity across breeds, and interpret the results.

Supervisor(s): Aniek Bouwman

Inbreeding due to breed characteristics

Is inbreeding higher at loci related to breed appearance? Several local Dutch breeds have very unique breed characteristics. For example the white belt of Dutch belted cattle (Lakenvelder), the white head with coloured blisters around the eyes of the Groninger White Headed (Blaarkop), or the double muscling of Dutch Improved Red cattle (Verbeterd Roodbont). The regions on the genome responsible for these breed characteristic are known. With whole genome sequence data we can characterize these areas and investigate the uniqueness and length of haplotypes in those areas. By comparing those regions to the remainder of the genome we can answer the research question. Are you interested in working with whole genome sequence data on genomic diversity of Dutch local breeds? You will learn to handle whole genome sequence data, to estimate genomic diversity, to identify haplotypes and interpret the results.

Supervisor(s): Aniek Bouwman

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 subject 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 need to have affinity with zoo populations and genetic diversity. They 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

Evolutionary genomics of common carp (Cyprinus carpio)

The common carp (Cyprinus carpio) is an iconic fish and important in aquaculture and as a model species. Carp is often preferred for experimental work due to its impressive size over its relative, the diminutive zebrafish which is a model for vertebrate development and physiology. To increase its use as an experimental species, but also to facilitate selective breeding, a better understanding of the functioning if the carp genome is required. The carp genome however has an interesting feature: it is tetraploid. The carp and its relative the goldfish originated from a hybrid between two species, about 12 million years ago. That hybridization led to a genome duplication – the genomes of both species were combined into a single nucleus. This situation more or less continues until today. But this leads to many questions concerning the fate of all those duplicated genes – are all genes still active? Only one of the copies is perhaps enough. Or are genes gradually lost, something you see in other polyploid species. And if both copies of genes are still there, do they perform different functions? We are currently doing a new genome assembly using Nanopore sequencing. This improved genome assembly allows us to tackle these questions. First we try to improve the interpretation of the function of the genome (annotation) by looking at gene expression in the context of the genome assembly. Second, we will compare the two 'subgenomes' to see what is lost, what is still there, and what is still there but doing something different. We are also interested in a broader comparative genomic study that includes zebrafish and many other fish species. To this end we will generate functional genomic data such as gene expression data, epigenetic modifications, and more, for several species including carp.

Supervisor(s): Annemiek Blasweiler (AFI, ABG), Hendrik-Jan Megens

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 analyze pedigree data from fish breeding programs that reproduce fish in breeding groups, and use the results to predict the rate of inbreeding in such a program. In addition, you can apply this knowledge to design a breeding program 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

F*rom sequence to sushi: genetic diversity in a newly domesticated fish species*

The Yellow Tail 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, a highly prized product. 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 Yellow Tail 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 Yellow Tail 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 be working 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. The number of breeding animals however, is limited, especially within each colour variety. No pedigree is kept. However, breed 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 (Blaarkop in Dutch) 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 what 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 from pedigree and DNA analysis, can improve their computer skills and will contribute to better safeguarding different breeds in Europe.

Supervisor(s): Jack Windig, 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

Sustainability and breeding programs

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 programs 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 programs. 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. Students 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 programs.

More reading about the observed behavioral patterns and their association with health and disease can be found at:

https://doi.org/10.1016/j.applanim.2019.05.007
https://doi.org/10.1016/j.cbd.2020.100700
https://doi.org/10.3389/fvets.2020.00587

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

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 saleable milk 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 calculatons.
- 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 student and supervisor(s).

Chair group(s): combined thesis 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 faming conditions the same as the ones that show excellent performance under organic farming conditions? In other words, is there genotype by environment 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 data base 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: Nauta, Wytze J. 2009. Selective Breeding in Organic Dairy Production. PhD thesis Wageningen University.

Supervisor(s): Henk Bovenhuis

Difference between morning and evening goat milk composition in relation to heat stress

In general, we assume that milk composition (fat% and protein%) is similar between morning and evening milk samples. However, we have strong indications that this is not the case for goat milk. Furthermore, we have indications that goat milk composition is (strongly) affected by heat stress. The aim of this study is to quantify differences in goat milk composition between morning and evening milk samples. We hypothesize that this difference is affected by temperature; it is expected that during warm summer days the evening milk sample is stronger affected by heat stress than the morning milk sample.

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 high value fish species: sea bream and sea bass. The company is taking photographs of all the fish in their breeding program 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, color, belly shape, etc.

In this project you will analyze 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 program 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 metabolizing 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 program 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 upto-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-WUR 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, color, 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, color, and/or ear position on goat produciton-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 MSc 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

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 programs 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 program. 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 realized 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 program 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 realized 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 realized 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 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, then this thesis project may result in a scientific publication.

Supervisor(s): Piter Bijma, Egbert Knol

Big data and data sciences

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 labor 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 program. The programming part will be done using Python (a general and very popular programming language) and OpenCV (a specialized 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 thesis 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 organizations 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, the student 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

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 by emailing a description of your project idea (pascal.duenk@wur.nl).

Contact information

Bastiaansen, John	john.bastiaansen@wur.nl
Bijma, Piter	piter.bijma@wur.nl
Blasweiler, Annemiek	annemiek.blasweiler@wur.nl
Bonifazi, Renzo	renzo.bonifazi@wur.nl
Bosse, Mirte	mirte.bosse@wur.nl
Bourke, Peter	peter.bourke@wur.nl
Bouwman, Aniek	aniek.bouwman@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
Fodor, István	istvan.fodor@wur.nl
Godia Perello, Marta	marta.godia@wur.nl
Gonzalez-Prendes, Rayner	rayner.gonzalezprendes@wur.nl
Groenen, Martien	martien.groenen@wur.nl
Guennoc, Eugenie	eugenie.guennoc@wur.nl
Gulzari, Benan	benan.gulzari@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
Niehoff, Tobias	tobias.niehoff@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. taghavirazavizadeh@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