ABG MSc thesis topics: 2021-2022

Version, March 2022

Breeding Programs and genetic diversity:

ABG_MSc-thesis2021_011

Immunogenome diversity and evolution in African chicken

The chicken immune system has been study as a model system to study basic immunology knowledge by establishing major landmarks and discoveries that defined concepts widely used today. Immunity involves many genes and gene families. One such important gene family is the Major Histocompatibility Complex (MHC). The MHC genes are an example of genes that are clustered and often duplicated within a single locus in the genome. The MHC genes code for cell surface proteins essential for the adaptive immune system, playing a vital role in the recognition of pathogens. The chicken MHC is a cluster of 46 genes in a region of about 209 kb on micro-chromosome 16 and plays an important role in immune response. It is highly polymorphic and has been used to study adaptive genetic diversity in chickens.
In this thesis you will study the MHC gene family in chicken. We have a unique whole genome sequence dataset of 26 Ethiopian indigenous chicken sample populations which were collected from four environmental gradients (6-7 populations per gradient) making a total of 513 birds from the whole country. These village chickens are raised under scavenging production systems that are characterized by frequent exposure to diseases pathogens as well as fluctuations in feed supplies. More specifically, you will identify single nucleotide polymorphisms and structural variation at the MHC locus in chicken. You will characterize the frequency and functionality of genetic variation and link that to disease resistance. Therefore, Knowledge on bioinformatics (working with large datasets) is essential. Together, this dataset provides a unique opportunity to identify novel markers associated with disease resistance. Your thesis will provide novel insights into the immune system and disease resistance in African chickens and will facilitate ongoing work in chicken immunology.
Supervisor: Martijn Derks, Richard Crooijmans

ABG_MSc-thesis2022_002

Genetic diversity across local Dutch breeds

In the Netherland 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.
Contact: aniek.bouwman@wur.nl

ABG_MSc-thesis2021_007

Breeding program for the Dalmatian Pelican (Kroeskoppelikaan – Pelecanus crispus)

The Dalmatian pelican is a rare breeding bird in SE-Europe, which is kept in a number of zoos, amongst others in Rotterdam Blijdorp zoo. The species breed in colonies, and in zoos with variable success. Individual matings cannot be set up and consequently breeding programs have to be set up at group level. One kind of breeding programs are breeding circles, whereby animals are exchanged in a systematic way across groups. This research will investigate in which way a European breeding program could be set up with the help of a breeding circle. First an inventory of numbers of animals and breeding success across zoos will be made. This may include interviews with European zoos keeping pelicans. Furthermore exchange schemes including breeding circles will be designed and evaluated with the help of computer simulations to determine potential inbreeding rates and (loss of) genetic diversity. Finally recommendations for further research will be made, including DNA research to better determine population structure. We are looking for a student with an interest in animal breeding and genetic diversity.
Supervisors: Jack Windig and Jeroen Kappelhof

ABG_MSc-thesis2020_013

What is the impact of selection on genomic 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 (namely phenotypic, pedigree and genomic selection). In this thesis, the aim is to compare the genomic changes in a livestock population under these three different selection strategies, 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. The student will learn to quantify the genomic changes by performing analyses on runs of homozygosity, selective sweeps and changes in allele frequencies across the genome.
Contact persons: Mirte Bosse and Yvonne Wientjes.

ABG_MSc-thesis2022_003

Inbreeding due to breed characteristics

Is inbreeding higher at loci related to breed appearance? Several local Dutch breeds have very unique breed characteristics such as the white belt of Dutch belted cattle (Lakenvelder), or 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 characterise 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.
Contact: aniek.bouwman@wur.nl

ABG_MSc-thesis2020_002

Inbreeding and genetic management in dog breeds

In the past 5 years inbreeding and the effect of genetic management have been analyzed and reports written on 15 dog breeds by WUR. These dog breeds differ in size, inbreeding rate, population structure and breeding policy. For this thesis the student will assemble data from the 15 studies to infer trends over dog breeds. Some of the questions are:

- Is there a general trend in inbreeding over dog breeds and if so, is it going in the right direction?
- Are the policies applied sufficient or is more needed?
- What causes a certain policy to be efficient in one breed and not in another, or is there a general policy that can be applied to all breeds.
- Are there situations in which an outcross is the only possibility left to solve inbreeding problems?

Answers to these and other questions will help to compile a general advice for dog breeding organizations to better manage inbreeding and relatedness in their breeds. The student will need to have an interest in population genetics, inbreeding and its consequences and to translate research into practical advice.
Supervisor: Jack Windig

ABG_MSc-thesis2019_035

Collaboration will help to reduce loss of genetic diversity in the Dutch and Belgian Draft horse studbooks?

A recent population analysis on the Dutch Draft horse breed indicated that the inbreeding rate is critically high. The population is closely related to the Belgian Draft horse and collaboration between the studbooks could possibly be an option to reduce loss of genetic diversity in both studbooks. In this study you will perform pedigree analyses of both studbooks separately and jointly to study the genetic relation between them. Subsequently you will use simulations to develop the most optimal strategy to reduce inbreeding in both studbooks.
This study is a collaboration between CGN and the University of Leuven in Belgium.
Supervisors: Mira Schoon, Marjolein Neuteboom and Jack Windig

ABG_MSc-thesis2021_012

Structural variation in local African chicken populations

Structural Variants (SV) are an important form of genetic variance including deletions, duplications, and other types of structural rearrangements. Structural variations (SVs), like copy number variants (CNVs) have roles in chicken evolution and explain phenotypic differences among populations. Although much is known about SVs in commercial populations of chickens and their significant effects on phenotype, little is known about the extent of SVs in indigenous chicken populations.
For this thesis, you will use a unique whole genome sequence dataset of 26 Ethiopian indigenous chicken populations which were collected from four environmental gradients (6-7 populations per gradient) making a total of 513 birds from the whole country. In this thesis you will characterize deleterious and functional SVs in the a diverse range of African chicken populations, using a bioinformatics approach. You will generate a catalog of SVs and perform functional annotation of identified SVs to assess the contribution of structural variation on trait variation. Your thesis will provide an important resource for chicken variation from a diverse set of breeds and regions in Africa. This research will provide novel insights into the evolutionary history of African chickens and will facilitate ongoing genetic and functional studies on the world’s most well-known domestic species.
Supervisor: Martijn Derks

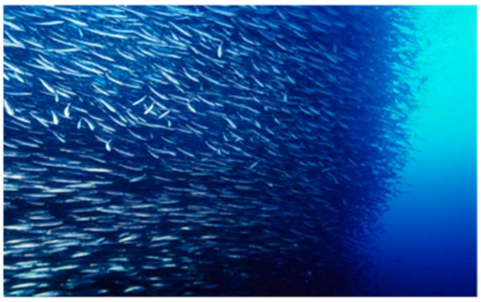
ABG_MSc-thesis2019_001
Inbreeding when matings are out of control

Summary: 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.
Supervision: John Bastiaansen

ABG_MSc-thesis2018_006
How do cultured fish keep their shape?

Summary: 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.
Supervision: John Bastiaansen, Hans Komen

ABG_MSc-thesis2018_007
Big data from a school of fish : A computer vision system to measure shape and size.



Summary: 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.
Supervision: Hendrik-Jan Megens, John Bastiaansen.

ABG_MSc-thesis2018_009
Fast growing fish also healthy?

Summary: 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.
Supervision: John Bastiaansen, Hans Komen

ABG_MSc-thesis2018_010
From sequence to sushi: Genetic diversity in a newly domesticated fish species.

Summary: 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?
Supervision: Hendrik-Jan Megens, John Bastiaansen

ABG_MSc-thesis2018_011
Potential of genomic selection for better health in dogs.

Summary: Health issues are prominent in most dog breeds. Selection may help to decrease diseases with a polygenic background. In the UK and Sweden breeding values are now estimated for hip-dysplasia in some of the larger breeds . Extension to other countries and breeds is not straightforward, especially because pedigree data across breeds and countries is often lacking. One possibility to overcome this would be to use DNA information. With genomic selection DNA profiles are linked to breeding values in a reference population. This then enables the estimation of breeding values for animals with only DNA information. An important aspect determining the accuracy of their breeding values

is their relatedness to the reference population. This subject will investigate the feasibility of genomic selection in dog breeds. It will answer questions such as the required size of the reference population, breed composition of the reference population and the influence of different relatedness among breeds or among populations in different countries, whether or not in the reference population. Computer simulations will be used to provide answers.
Supervision: Jack Windig

ABG_MSc-thesis2018_012

Genetic diversity in Dutch local cattle breeds stored in the gene-bank

Summary: The DNA of all local cattle breeds stored in the Dutch gene-bank has been typed with SNP-chips at the end of 2016. Now a wealth of information on genetic diversity in the gene-bank and in the local Dutch cattle breeds is available, which can help to manage both the gene-bank and the breeds themselves. In this subject this genetic diversity will be explored. Diversity can be quantified in different ways (e.g. allelic diversity, marker estimated kinships, rows of homozygosity etc.) to answer questions such as:

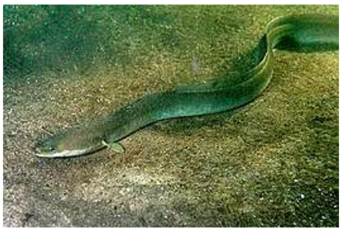
1. Do breeds differ in the amount of diversity stored in the gene-bank?
2. How much does the genetic diversity of the different breeds overlap (and differ)?
3. How does the diversity stored in the gene-bank relate to the diversity in live populations?
4. Has the diversity in the breeds changed over time?

Student interested in the conservation of genetic diversity can apply. Having followed animal breeding courses or wildlife genetics will be helpful but is not required.
Supervision: Jack Windig

ABG_MSc-thesis2018_013

Cryopreservation of eel sperm

Summary: We would welcome a MSc student who is interested to investigate by literature review how cryopreservation of eel sperm can be improved. World-wide, eel populations have decreased strongly in numbers since the 1970s. 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 till about 20 days post hatching. However, larvae cannot be fed yet and therefore 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 a good quality sperm storage available.
What we need from you 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.
Supervision: Arjan Palstra and Henri Woelders



Quantitative genetics:

ABG_MSc-thesis2022_005
Inbreeding and genetic parameters in homing 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 unfavourable 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 survivability. Which of these research questions you will work on depends on your personal interest.
Contact person: Pascal Duenk

ABG_MSc-thesis2022_004
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

Contact person: Henk Bovenhuis, Prof Christine Baes (Univ. Bern and Guelph)

ABG_MSc-thesis2021_004
Selecting/restricting training populations for improved prediction accuracy.

The essence of genomic prediction is to predict the phenotype of genotyped individuals. This is achieved by using the results of a statistical model that determines the association between phenotype and SNP-genotypes in a so-called training population. The accuracy of this prediction is mainly driven by the genetic relationship between training population and the selection population. Underlying assumption is that the relationship based on the SNPs reflects the relationships based on the QTLs underlying the traits of interest. Genomic prediction works well in a closed breeding population.

If, however, the training population consists of a combination of (sub) populations, the relationship (Linkage Disequilibrium) between SNP genotypes and underlying QTL genotypes may not be consistent. For example, SNPs can be segregating in the whole training population, while QTL alleles might be segregating in only part of the training population, which will hamper the prediction accuracy. This has been extensively documented in plants, e.g. corn breeding (Melchinger/Bernardo), and also prediction across breeds in animals results in low accuracies.

A PhD student of De los Campos has, recently, published an algorithm that could be used to determine which part of the training population is most informative for a given group of genotyped individuals. The idea is that selecting the optimal part of the training population would result in higher prediction accuracy. The goal of this MSc project is to better understand this algorithm and to discover when and how to apply it using real and/or simulated data.

Reference

Lopez-Cruz et al. 2021. Optimal breeding-value prediction using a Sparse Selection Index. Genetics. iyab030. doi: 10.1093/genetics/iyab030

Contact: Henri Heuven, Pascal Duenk

ABG_MSc-thesis2021_003

How good is good enough? The inter- and intraobserver reliability of annotations for pose-estimation.

In poultry production, locomotion is an important health and welfare trait. An animal’s locomotion can affect its ability to access feed and water, perform motivated behaviors, and avoid pecks of conspecifics. Locomotion is often assessed using a gait score, a score that depends on the judgment of a human expert based on the visual observation of an animal. Efforts have been made to make gait scoring less time-consuming and more scalable. Computer vision could help, though, machine learning algorithms often require a large amount of annotated data to learn what you are interested in. This data should be of sufficient quality and quantity, as poor data results in poor models. In this thesis, you will work on pose-estimation in broilers with DeepLabCut (open-source pose-estimation software). You will investigate the human variability of annotations and will go through all steps to train your own pose-estimation network to check the performance of your network against the found human variability.

Contact person: Jan Erik Doornweerd

Status: available, some affinity with programming required

ABG_MSc-thesis2020_016

Detecting and characterising de novo structural mutations in pigs

Much of the current published research on de novo mutations (DNM), new mutations arising in the germline, has focused on detecting single nucleotide polymorphism DNM. However, structural variation, including copy number variation (CNV), also plays an important role in individual fitness, making CNV DNM important to include when studying DNM and mutation rate. In humans, CNV DNM have been associated with the presence of a few developmental disorders, however, to date, the effects of CNV DNM have not been estimated for complex traits in any species.

For this project:

I. You will detect CNV DNM using high depth whole genome sequencing data of 46 pig trios from two commercial lines.

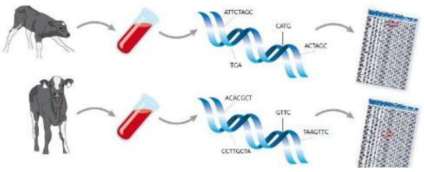
II. You will validate the CNV DNM by detecting these new CNV in the offspring of probands. Each proband has >30 offspring that have been genotyped with a 50K SNP chip and this information can be used to call CNV.

III. You will estimate the effects of the CNV DNM. Each of the >30 offspring also have phenotypic records which can be used in addition to the CNV calls to estimate the effects of DNM CNV on several quantitative traits.

Supervisors: Young Lim Lee and Christina Rochus

ABG_MSc-thesis2020_014

The long-term effects of genomic selection



In the beginning of this century, genomic selection has been introduced. 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.
Contact person: Yvonne Wientjes.

ABG_MSc-thesis2020_006

Updating genomic predictions without updating the reference population

With genomic prediction, breeders typically use the same reference population for a number of generations to make selection decisions, because updating the reference population is costly. As a result, the accuracy of genomic prediction reduces every generation, because marker effects may change considerably due to changes in allele frequencies. It may therefore be possible to improve genomic predictions by accounting for these changes in allele frequency. In this project, you will use simulations to investigate the benefit of updating marker effects each generation without updating the reference population. Depending on your interests and skills, you can investigate the benefit for accuracy, response to selection, and profitability of the breeding program. You will learn about simulation of breeding programs, genomic prediction, and you will obtain valuable programming skills.
Contact person: Pascal Duenk

ABG_MSc-thesis2021_008

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).
Joint thesis project Animal Breeding and Plant Breeding; 36 ECTS.
Contact persons:
Animal Breeding and Genomics: piter.bijma@wur.nl
Plant Breeding: peter.bourke@wur.nl

ABG_MSc-thesis2021_005

Data selection for improved association studies.

In GWAS studies the goal is to identify QTL affecting complex traits. In many cases the assumption is that the underlying QTL is biallelic (Q/q). However, if the population is heterogeneous this assumption may be wrong and more QTL alleles (Q1/Q2/Q3…) are segregating. Especially because SNPs only have 2 alleles, finding an association between SNP and QTL is easier if data is used where only two QTL alleles are segregating, i.e. in the part of the data that is homogeneous with respect to QTL-effects.
The idea is that an algorithm developed by De los Campos for improving prediction accuracy, could be used to select data which is most informative to find QTL in an association study.
The goal of this project is to study the algorithm for identifying QTL in heterogeneous data set using real and/or simulated data.
Reference
Lopez-Cruz et al. 2021. Optimal breeding-value prediction using a Sparse Selection Index. Genetics. iyab030. doi: 10.1093/genetics/iyab030
Contact: Henri Heuven, Pascal Duenk

ABG_MSc-thesis2020_003

To estimate genetic parameters for activity and locomotion traits in broilers

Locomotion is an important trait in broilers. However, measurements for locomotion are subjective and data collection is time-consuming. In this thesis, you will investigate if activity could be used as a proxy for locomotion traits. Data collected at Cobb will be used to estimate genetic parameters for activity and locomotion traits, to investigate the genetic relation between traits.
Contact: Esther Ellen/Malou van der Sluis

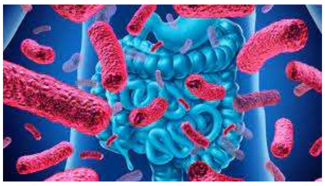
ABG_MSc-thesis2020_001

A computer vision system to extract data from images of fish

Among the most commonly selected traits of the fish for aquaculture are growth and fillet yield. While growth can be measured on the selection candidates themselves, fillet yield require sib information. Thus, measuring fillet yield is more laborious and costly than measuring growth. Semi-automated data collection systems using computer vision may be a solution to predict the fillet yield of live fish, which facilitates the selection for this trait. For this project, we collected 2D and 3D images of commercially produced gilthead seabream and recorded their harvest weight and fillet yield. Also, 2D images of heart and liver of the fish were taken. The MSc student will be responsible for extraction of the information from the images and using this information to create a solution to predict the fillet yield of live fish. Interested students will have the opportunity to learn basic programming in Halcon, R, Python and ASReml according to their preferences.
Supervisors: John Bastiaansen, Angelo Mencarelli, Benan Gulzari

ABG_MSc-thesis2021_002

Combine animal, health, management and PCR data to optimise output and help us unchain the value of the microbiome

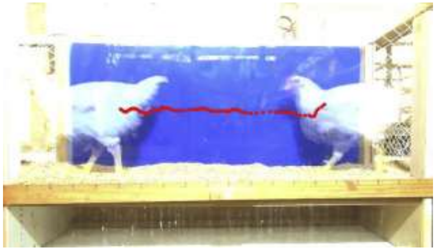


The animal microbiome is widely recognized to have a large impact on day-to-day animal health and welfare. The microbiome sequencing data grow exponentially but the microbiome remains an extremely complex system. Our newly developed molecular assays help to unchain the value of the animal microbiome.

Currently we are running the assays on field samples and the first results of analysis of the PCR data show great potential. However, by adding smart analysis, data processing and animal and management information we believe we can optimize these results in order to further grow the insights of the animal microbiome and thereby support our customers.

You will help us by analysing and interpreting the results of a large field experiment. Questions to be answered will be (among others):
- Do we see differences in microbiome in heathy and sick animals?
- What is the influence of animal and management factors on the micriobiome?
- Does the sampling method influence the results?

Contact person: Kirsten van Hees, Bart Ducro



ABG_MSc-thesis2020_005
Individual differences in activity patterns of broilers

Differences in average activity levels of broilers have been suggested to provide insight into the health and welfare of broilers. However, it is not yet studied in detail how individual broilers may differ in their patterns in activity over time and what the underlying causes of these different patterns are. In this thesis, a dataset on activity over time of a large group of broilers is available and you will investigate whether individual broilers show different patterns in activity over time. You can, for example, think of patterns in the use of the area over time, the most active periods during the day, and the frequency of feeder and drinker visits. Furthermore, data on different health and welfare traits are available for these broilers, which could potentially be linked to different patterns in activity.

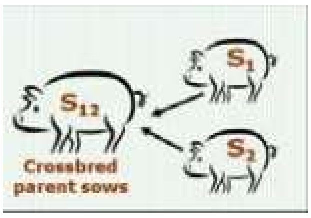
Contact: Malou van der Sluis/Esther Ellen

ABG_MSc-thesis2019_036
Impact of misspecified variances on estimated breeding values

Summary: Breeding value estimation models rely on previously estimated genetic and other variances. In the era of genomic selection, the genetic variances used may still be estimated based on pedigree data, and may therefore be not fully consistent with genomic breeding value estimation models. This thesis will investigate the impact of misspecified variances on estimated breeding values and genetic trends, in terms of accuracy and bias. To enable evaluation of accuracy and bias, the work will be based on simulated data.

Daily supervisor: Mario Calus

ABG_MSc-thesis2020_009
Improving genomic prediction for crossbred performance without crossbred information



Crossbred breeding programs aim to improve performance of crossbred animals. Selection takes place in purebred lines, and is usually based on purebred performance measurements. Breeders can decide to collect data on crossbred animals to improve genomic predictions, but this approach is costly and may be difficult in practice. It may therefore be interesting to investigate whether predictions can be improved without collecting crossbred data, for example by improving the model. In this project, you will learn how to simulate a crossbred breeding program that applies genomic selection. The aim is to simulate and evaluate strategies that aim to optimize genomic prediction for crossbred performance, without collecting crossbred information.

Contact person: Pascal Duenk

ABG_MSc-thesis2019_033
Definition of phantom parent groups in international evaluations of beef cattle.

In national genetic evaluations, recorded animals may miss pedigree information for one or both of their parents. Nevertheless, producers desire to have an animal Estimated Breeding Value (EBV) also for such individuals to make their selection decisions and, therefore, these animals are still included in the genetic evaluations. However, in genetic evaluation models, animals with missing pedigree information are considered as offspring of the same unrelated and unselected animals in the base population, disregarding of their year of birth. This, in turn, may lead to potential bias in the animals' EBV. In a similar way, when animals are evaluated across countries in an international genetic evaluation, they may miss parental information. To overcome such problem, Phantom Parent Groups (PPG) can be defined. However, in current beef cattle

international evaluations PPG are not yet implemented. In this MSc project, after a brief literature review on the topic, you will define PPG for across-country evaluations in Limousine beef cattle European populations. Your goal will be to find the best approach to use for PPG in estimating beef cattle animals' EBVs. You will have the opportunity to work with real data and to improve programming skills.
Please keep in mind that organizing the data before the start of the thesis might take some time. So please inform us ~3 months in advance about your interest in this topic.
Supervisor: Renzo Bonifazi

ABG_MSc-thesis2020_007
Proximity as an indicator of social bonds in broilers

Social behaviour is an important component of an animal's behavioural repertoire. One way to assess whether broilers have preferences for specific individuals in their environment is to study the proximity between individuals over time. In this thesis, a dataset on locations over time of individual broilers within a group is available and you will investigate the proximity between individuals to determine whether broilers show preferences in who they associate with.
Contact: Malou van der Sluis/Esther Ellen

ABG_MSc-thesis2019_026
Machine Learning to predict lifetime resilience scores using at-market sensor technologies

The resilience of a cow indicates how well she responds to disturbances (for example, a heat wave or disease event). Although resilience is a trait of high interest for farmers, it is also a complex trait that is very difficult to measure. Therefore, we have tried to develop proxies for resilience with data from sensor technologies that are commercially available (measuring for example activity, or body weight). We have analysed these traits using ordinal logistic regression analyses, unfortunately with limited success. Modern machine learning algorithms, however, may be more suitable to deal with our large volumes of (noisy) data to find sensor patterns specific for resilient cows. Your thesis will contribute in our search to develop proxies for resilience, and will be part of GenTORE (www.GenTORE.eu). This is an EU-funded project aiming at the development of management tools to optimise resilience and efficiency of cattle across production systems (beef and dairy) and environments.
Contact person: Claudia Kamphuis

ABG_MSc-thesis2019_027
Methane traits for dairy cows

Dairy cattle are a significant contributor to greenhouse gases. There is a possibility that breeders can select cows that produce less methane but are still efficient milk producers. We are exploring, methods of recording methane on farm, defining methane phenotypes, and estimating genetic parameters, all of which would contribute to the breeding of low methane producing cows. Two different types of machines have been installed in barns across the Netherlands. Both measure methane of individual cows, but with different methods.
The aim of this thesis is to help determine if cows would have the same ranking for methane with the two different methods.
Contact person: Michael Aldridge
Status: Available

ABG_MSc-thesis2019_028
Estimation of heritabilities and genetic correlations in production environment of gilthead seabream using genotypic data

Heritabilities and genetic correlations of production traits in fish are commonly calculated with the data collected in the environment where the broodstock is kept, which is generally far away from the commercial production environment. Variance components estimations with the data collected in the commercial production environment are rare. In this project, data of the commercially produced gilthead seabream for human consumption were collected for harvest weight, fillet weight, fillet yield, fat percentage, heart weight and liver weight. Harvest weight, fillet yield and fat percentage of the fish are traits of commercial importance. Heart and liver weights can be used as indicators of overall health. The heritabilities of these traits as well as genetic correlations among them will be estimated using genotypic data by the MSc student. Estimations performed by the MSc student will be an important contribution to fish breeding literature. Interested students will have the opportunity to learn basic programming in R and ASReml.
Supervision: Benan Gulzari, John Bastiaansen

ABG_MSc-thesis2019_022

Effects of the TLR1A polymorphism in layer chickens

Based on a genome wide association study we showed that a single nucleotide polymorphism (SNP) in the toll-like receptor 1 family member A (TLR1A) has a major effect on natural antibody levels in chickens. Subsequently a new group of birds were genotyped for this polymorphism in order to investigate in more detail the effects of this polymorphism. These birds were from generation 6 and 7 of a selection line that has been selected for either high or low levels of natural antibodies. These birds have information on body weight at different ages. In previous studies we also found that natural antibody levels in male and female birds are genetically different traits. Therefore the aims of this study are:

- Estimate the effect of this TLR1A polymorphism on body weight at different ages
- Estimate the effect of the TLR1A polymorphism on natural antibodies in male and female birds

Supervisors: Henk Bovenhuis

ABG_MSc-thesis2019_034

Data requirements for the estimation of across-genetic correlations in international beef cattle genetic evaluations: a simulation study.

International beef cattle evaluations aim to assess animals' Estimated Breeding Value (EBV) in different environments, i.e. countries. To accurately estimate animals' EBV, a crucial aspect to consider is how much the information collected on animals recorded in one country is going to contribute to the genetic merit of its relative in another country. This is achieved by the usage of across-country genetic correlations, which are estimated from animals with recorded relatives in multiple countries, e.g. a popular bull. Such animals ensure the presence of across-country genetic links, called genetic connections. Therefore, to correctly estimate genetic correlations across-country good data structure is required, with many genetic connections in place. However, in beef cattle, compared to dairy cattle, there is a structural lack of genetic connections due to the lower usage of artificial insemination and, in turn, estimating genetic correlations across countries may be difficult and require long time. Moreover, many recorded traits in beef cattle are affected by a maternal genetic effect expressed in the offspring's phenotype and that requires a particular data structure in place to be accurately estimated.
In this project, your goal will be to simulate a beef cattle international dataset with phenotypic and genomic information. You will gain experience with programming, simulation studies and international cattle evaluations. Once the dataset is in place, you will perform a simulation study with the aim of understanding the required amount and structure of the data to have an accurate estimation of genetic correlations across-country.
Supervisor: Renzo Bonifazi

ABG_MSc-thesis2019_023

How to use on-farm sensor technologies to define resilience dairy cows?

The resilience of a cow indicates how well she responds to disturbances (for example, a heat wave or disease event). Although resilience in a trait of high interest for farmers, it is also a complex trait that is very difficult

to measure. Therefore, we have tried to develop proxies for resilience with data from sensor technologies that are commercially available (measuring for example activity, or body weight). A major challenge we faced, however, was the lack of a gold standard definition of resilience. Your thesis will contribute in this search for a gold standard, by focussing on the analyses of sensor data patterns and to test whether these patterns differ between cows that stay on a farm for a long time (assuming these to be resilient) and those leave the farm at an early stage (assuming these to be non-resilient). Your thesis will be part of GenTORE (www.GenTORE.eu), which is an EU-funded project aiming at the development of management tools to optimise resilience and efficiency of cattle across production systems (beef and dairy) and environments.
Contact person: Claudia Kamphuis

ABG_MSc-thesis2019_021

Finding targets for Crispr-Cas using two different methods (Comparing methods (BVS and IBSS) for analysis of genome wide association (GWAS) data).

Crispr-Cas is considered to revolutionize animal breeding, i.e. replacing or adding the alleles that are beneficial. A strong pre-requisite is that the targets (causal mutation are known). Identifying the causal mutations, especially for complex traits, will be challenging. Identifying the regions on the genome and then fine mapping is needed to find the targets. Recently a new method for finemapping is proposed but is not obvious if this method is superior to existing methods for finemapping.
The goal of an association study is to identify regions/genes/causal mutations associated with phenotypes. In most cases there are some phenotypes (n=250-5000) but many genotypes (SNPs; 1k-600k). Many SNPs have a similar segregation pattern and it is therefore hard to determine which SNP is best associated with the phenotype. Currently a Bayesian Variable Selection (BVS) method is used to address this problem. However a new method is recently (dec 2018) proposed called Iterative Bayesian Stepwise Selection, which identifies groups of SNPs (credible sets). Wang et al 2018),
The goal of this project is to compare BVS with IBSS on a real dataset (cattle, pigs, chicken or plants) to determine the pros and cons of either method for identifying genes/regions affecting a phenotype.

Requirements: Interest (background/courses) in statistical genetics.
You will learn programming in R, Bayesian statistics and writing an article.
Contact person: Henri Heuven Henri.heuven@wur.nl

Ref: A simple new approach to variable selection in regression, with application to genetic fine-mapping. Gao Wang and Abhishek Sarkar and Peter Carbonetto and Matthew Stephens.
https://www.biorxiv.org/content/biorxiv/early/2018/12/19/501114.full.pdf

ABG_MSc_thesis2019_030

The genetic background of milk composition

There is remarkable genetic variation in detailed milk composition, e.g. fat-, protein- and mineral composition. The genetic background of milk composition is not only relevant for selective breeding but milk composition is also an important indicator for farm management. Several research topics related to milk composition are available, e.g.

- Estimation of genetic parameters
- Genome wide association studies
- Change in the genetic background of milk production traits across parities
- Milk infrared analyses as indicator for breeding goal traits

Supervisors: Henk Bovenhuis

ABG_MSc_thesis2019_020

Identifying and monitoring individual birds in a group, why is one bird more active than another? (minor thesis)

Broiler chickens are kept in large groups, which makes it difficult to visually identify birds and monitor their activity levels. In this thesis, ultra-wideband tracking data on activity of individual birds of two genetic lines is available, which can be used to study whether there are individual differences in activity. Also, data on the bodyweight and gait score of the individuals is available. During your thesis you will be analysing the data to identify possible correlations between these traits and activity levels.
Contact persons: Esther Ellen (esther.ellen@wur.nl) & Malou van der Sluis (malou.vandersluis@wur.nl)

ABG_MSc-thesis2019_005
Increasing response to genomic selection, by creating selection candidates that give more variable offspring

In animal breeding, we make progress by selecting offspring that are on average better than their parents. Recent research shows that some parents give more genetically variable offspring than others. Genomic selection of parents that give more variable offspring can be used to increase response to selection, because they have a higher chance of producing an extremely good offspring.

In dairy cattle, for example, some bulls have been very successful as fathers of the next generation of bulls (so-called bull-sires) because they occasionally produce an extremely good son, whereas other bulls with equally good EBV mainly produced average sons.

In this thesis project, you will investigate whether we can create individuals that give more genetically variable offspring. For example, by mating the least related parents based on their genotypes, we can create individuals with more heterozygosity. In the next generation, such individuals will give more genetically variable offspring. Thus the focus of this thesis project is the use of mating strategies to create individuals that give more variable offspring.

We are looking for a quantitatively oriented student, who also likes to learn to program (in R and/or Fortran). You will develop programming skills and considerably strengthen your understanding of quantitative genetics and animal breeding. The knowledge and skills developed in this thesis can be used in breeding for a broad range of species.
Contact person: Piter Bijma

ABG_MSc-thesis2019-006
Genetics of longevity in the Friesian horse

Selection for longevity has shown its value in cattle and pigs and it is therefore worthwhile to study the possibilities in horse breeding. This selection trait requires that the day of death is known; in the Friesian horse this is facilitated since a reimbursement is paid when the passport of the horse is returned to the studbook. Alternatively, productive life of a horse can be approximated from recordings during events like competition, inspection or birth dates of progeny.

In this study you will investigate what is the best way to combine recordings on a horse to construct the selection trait longevity and to infer what the possibilities for improvement of longevity are based on heritability estimates.
Contact person: Bart Ducro
Status: not available

ABG_MSc-thesis2019_019
Minor student project: Validation of a novel tracking system for broilers

Broilers are kept in large groups, which makes it difficult to identify and monitor individual animals. However, individual data on broiler behaviour is valuable for breeding programmes. We are currently developing and implementing a novel tracking system, to monitor activity of individual broilers throughout life. This student project focuses on the validation of the tracking system. Data is currently being collected on a broiler farm. To validate the novel tracking system, we want to compare the output of the tracking system to video analysis. In this student project, video recordings will be studied and the location of birds will be compared to the tracking system output. Furthermore, there are options to visit the farm where the novel tracking system is implemented and to help in data collection.
Contact persons: Esther Ellen (esther.ellen@wur.nl) & Malou van der Sluis (malou.vandersluis@wur.nl)

ABG_MSc-thesis2019_008
Genetics of fish growth in Greek vs. Spanish waters

Summary: In the fall of 2019 we have a unique opportunity to do an MSc thesis that includes field work to do measurements on fish that you will subsequently use for quantitative genetic analysis. The question that is asked in this project is whether growing in warmer vs. colder water puts different (genetic) requirements on fish. To measure this genotype by environment interaction (GxE) we let fish from the same families grow in two different farms, one in Spain and one in Greece. The experiment is part of the EU project MedAID (http://www.medaid-h2020.eu/). In this project you will analyze phenotypes collected in the project to estimate heritability and genetic correlations of performance in the two different environments.

Supervision: John Bastiaansen, Benan Gulzari

ABG_MSc-thesis2019_009
Individual activity patterns of seabream in a sea cage by accelerometry

Summary: Like salmon, cultured seabream (dorade) is raised in tanks on land after which they are transferred to sea cages for more advanced growth up to market size. In order to decrease the mortalities that accompany this transfer, there is a trend of keeping fish longer in the tanks before they go to sea. The longer on-land phase allows to raise fish in flowing water that stimulates swimming exercise. Exercise may have beneficial physiological effects that make fish more robust which could lead to improved growth, and better stress coping and immune capacity. Exercise may also impact body shape and activity patterns in the seacages. For this study seabream was raised at three different flow regimes for eight months. In the first two regimes the fish were forced to swim continuously at two different speeds, and in the 3rd regime they were not forced to swim. Accelerometers were implanted in fish and after release in an experimental sea cage, acceleration as measure of activity was monitored for 6 weeks. In this study you will determine the individual activity patterns of seabream in a sea cage by accelerometry. How did individual fish behave over time and in relation to temperature? Were exercised fish also more active in the cages?

Supervision: Arjan Palstra, Hans Komen

ABG_MSc-thesis2019_010
Exercise effects on body shape, curvature and deformities in seabream

Summary: Like salmon, cultured seabream (dorade) is raised in tanks on land after which they are transferred to sea cages for more advanced growth up to market size. In order to decrease the mortalities that accompany this transfer, there is a trend of keeping fish longer in the tanks before they go to sea. The longer on-land phase allows to raise fish in flowing water that stimulates swimming exercise. Exercise may have beneficial physiological effects that make fish more robust which could lead to improved growth, and better stress coping and immune capacity. Exercise may also impact body shape and activity patterns in the seacages. For this study seabream was raised at three different flow regimes for eight months. In the first two regimes the fish were forced to swim continuously at two different speeds, and in the 3rd regime they were not forced to swim. In this study you will design a morphological parameter set and analyze body shape of fish that were kept at the three different flow regimes. How does exercise affects body shape? Does exercise cause/prevent deformities? Does swimming in a round tank lead to curved bodies?

Supervision: Arjan Palstra, Hans Komen

ABG_MSc-thesis2019_011
Exercise effects on the molecular regulation of immune capacity in pikeperch

Summary: Exercise has beneficial effects on immune capacity in mammals but what about fish? Juvenile pikeperch was exercised or not exercised for four weeks by making them swim in a large swim gutter. Exercised and non-exercised fish were then immune challenged mimicking a bacterial infection. Head kidney tissue of the fish was dissected and used for RNA sequencing of the transcriptomes. In this study you will determine the exercise effects on the molecular regulation of immune capacity in pikeperch. How does exercise modulate immune activation? Which genes are important markers?
Supervision: Arjan Palstra, Geert Wiegertjes, Hans Komen

ABG_MSc-thesis2019_012
The use of ultrasound for non-invasive physiotyping

Summary: Ultrasound can be used for non-invasive morphological and physiological measurements in fish. In the past we have used ultrasound for determining the sex of sole; measuring fillet thickness in salmon and seabream; determining cardiac outflow in yellowtail, and determining liver and gonad size in eels. In this study you will assess the applicability of ultrasound for high throughput physiotyping. What can be measured live and high throughput? Which of these parameters are valuable for breeding?
Supervision: Arjan Palstra, Pauline Jehannet, Hans Komen

ABG_MSc-thesis2019_013
Novel treatments for the induced sexual maturation of eels

Summary: World-wide, eel populations have decreased strongly in numbers since the 1970s. 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. But eels cannot be reproduced in captivity yet. Only by hormonal treatment larvae can be produced but they subsequently die before feeding exogenously. Somehow they do not eat. In this study you will analyze the effects of novel hormone treatments on the sexual maturation of the eel. Which treatments are more efficient in increasing reproductive success?
Supervision: Arjan Palstra, Hans Komen

ABG_MSc-thesis2019_014
Effects of the altitude on qualitative phenotypic characteristics of local Ethiopian chicken

Summary: You will investigate the effect of the environment (variation in altitude) on morphological variability of indigenous chicken populations. The phenotypes will be collected in Ethiopia. So far we have been aiming to utilize the quantitative measurements for landscape genomic studies. You will collect and learn to analyze biogeographic information together with data on qualitative traits from indigenous chicken populations. The traits could include plumage colour, comb type, earlobe colour, shank colour, head shape, shank morphology. Based on the results of analysis you will develop hypotheses for why certain phenotypes might be useful in a particular environment.

Field work in Ethiopia to collect data and to work with the researchers at ILRI campus in Addis Ababa could be included. In this case you will collect data on chicken production systems, qualitative morphological data, geographic data (coordinates), and possibly other geospatial and environmental data such as altitude or climate. The correlation of traits with environmental parameters is part of your analyses.
Supervision: Fasil Getachew, John Bastiaansen.

ABG_MSc-thesis2018_015

Relation between litter size and survival in pigs

As litter sizes are further increasing in pigs there is a risk this will be at the expense of correlated traits like vitality of individual pigs. A closer look is required to this relationship to properly define breeding programs to avoid undesirable effects. Large datasets are routinely collected by breeding companies to study such a relationship. You will work with advanced statistical techniques like random regression to tackle this research question and will talk with people from the industry about practical consequences.

Contact: Han Mulder

Status: not available

ABG_MSc-thesis2018_016

Use of automatic oestrus data in breeding values for reproduction in dairy cattle

Oestrus detection is often facilitated using sensor devices like pedometers. These devices record many more data than actually used to alert oestrus Additional traits might be defined which have a relationship with reproduction in dairy cattle. In this study you will work with large data streamings from pedometers and you will develop new traits and investigate whether these are geneticallly determined. You will work together with people from Nedap and CRV

Contact: Bart Ducro and Han Mulder

Status: not available

ABG_MSc-thesis2018_017

Validating the optimal modelling of traits when predicting animal breeding values

Summary: Optimal modelling has a key role in improving predictability of breeding values of traits of interest. Better predictions increase the genetic gain of the breeding program. When optimizing a prediction model it is important to have a validation strategy that works well for assessing accuracy and bias of model predictions to selection candidates. Different approaches have been attempted to validate accuracy of model predictions such as; theoretical accuracy derived from the prediction error variance, correlation between predicted breeding values and corrected phenotypes, and semi-parametric estimates of population accuracy for prediction of future phenotypes. However, it is little known which approach is most appropriate for traits such as indirectly observed traits, traits with scaling effects for different sexes, traits measured in group records, and traits that are affected by maternal effects. You will investigate validation strategies for these traits in different breeding programs with traditional BLUP selection and GBLUP selection for pig, chicken or fish. In this thesis project you will learn to perform stochastic simulations and learn about the design and optimization of breeding programs and genetic evaluation models.

Supervision: Thinh Tuan Chu, John Bastiaansen

ABG_MSc-thesis2018_020

Analysis of Genotype by Environment interaction in laying hens: the effect of beak trimming on egg production

Summary: Beak-trimming is a common procedure in laying hens to reduce the damage due to feather pecking and to avoid feather pecking turning into cannibalism. There is a desire to get rid of beak-trimming, since it may affect welfare of the laying hens. However, beak trimming may affect egg production. In other words, families that are good egg producers when beak-trimmed may be suboptimal without beak-trimming. This may result in the selection of the wrong families if the selection and production system differ with respect to beak trimming. This is a form of GxE-interaction, which can be quantified by estimating the genetic correlation between egg production in beak-trimmed and intact birds. In this thesis project, you will statistically analyse data on egg production in laying hens, so as to quantify this GxE-interaction.

Daily supervisor: Piter Bijma

ABG_MSc-thesis2018_021

Breeding for general disease resistance in chickens

Summary: Livestock production has been, and is still changing dramatically due to societal concerns. Antibiotic resistant bacteria (amongst other reasons) led to a ban on antibiotic use. Group housing of animals is mandatory, and increases welfare. In addition group housing also increases disease pressure on animals. Robust livestock animals are needed for these challenging environments. In chicken, levels of natural antibody were associated to an increased survival during laying period. Therefore natural antibody levels seem to be a promising selection criterion for general disease resistance in layer chicken. Currently we are selecting two chicken lines for either high or low levels of natural antibodies. As expected, the high line showed a 2-fold lower mortality due to an *E. coli*-infection than the low line. In addition, some relevant genomic regions were identified. However many questions remain concerning natural antibodies, disease resistance, and consequences for production parameters. The project offers various opportunities for students interested in the link between animal breeding and genetics and immunology. However, a background in immunology is not needed! For more information on thesis possibilities, please send an e-mail to Marleen Visker.

Possible techniques: estimation of genetic parameters, estimation of maternal effects, genome wide association studies, estimation of correlated responses, (q)PCR, mutations, determination of immune repertoire

Supervisors: Henk Bovenhuis

ABG_MSc-thesis2018_022

Improvement of the efficiency of milking robots

Summary: At this moment, efficiency of milking robots is determined by measuring the number of kg milk per minute robot time, and is only determined based on successful milking events. This efficiency can be estimated better when also unsuccessful milking events are taken into account, and breeding for this trait can result in a cow that makes more efficient use of the milking robot by generating less unsuccessful milking events and/or producing more kg milk per minute robot time. In this way, the output in kg milk per milking robot can be further improved. Unsuccessful milking events can be caused by poor udder conformation, or by agitation of the animal (starting phase of lactation, character, environment (flies), stress-experience). In this project, the student is asked to extract, check and structure the data from the milking robots at farms, define the traits, and to estimate the relevant genetic parameters. In addition, genetic correlations with other traits will be estimated. The project will be executed at CRV, Arnhem, under supervision of CRV and Wageningen University.

Contact person ABG: Han Mulder

Status: not available

ABG_MSc-thesis2018_025

Estimating genetic parameters for body weight in crossbred laying hens

Summary: In this study, you will estimate genetic parameters for body weight in four crossbred layer lines. Body weight has been collected when the birds were approximately 40 weeks of age. To estimate genetic parameters, you will include genomic information in the analysis. During this thesis, you will work with several computer programs, including R and ASReml. You will have the opportunity to visit the breeding company Hendrix Genetics to discuss and present your results.

Daily supervisor: Piter Bijma/ Katrijn Peeters

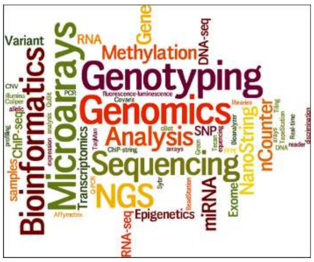
ABG_MSc-thesis2018_026

Finding the solution for estimating significant SNPs in a GWAS analysis using a Bayesian approach.

Summary: Bayesian methods rely only MCMC-methods which is basically using an iterative approach to calculate the SNP-effects in a GWAS analysis. It might require 100,000's of iterations. Finding the solution assumes that the solution exist and it a matter a searching for it in an efficient way.
Evolutionary algorithms have been described to do this efficiently. The idea is to develop an evolutionary algorithm for the Bayesian method such that it will find the answer quicker than by iteration. This might be especially useful if many SNPs become available due to sequencing (instead of genotyping) individuals. You will learn about R, Bayesian analysis, evolutionary algorithms; it is going to be very challenging!
Daily supervisor: Henri Heuven

ABG_MSc-thesis2018_027

Efficient genomic evaluation with a large number of genotyped animals



Summary: Single-step genomic evaluations combining data from genotyped and non-genotyped animals are becoming the method of choice to predict genomic breeding values. However, such genomic evaluations could become infeasible with more than 100,000 genotyped animals, mainly due to expensive inversions of the genomic relationship matrices. An algorithm computing an approximation of the inverted relationship matrices was recently proposed to remove this limitation. The aim of this thesis is to test and to validate this algorithm for genomic evaluations involving purebred and crossbred genotyped animals.
Daily supervisor: Jeremie Vandenplas
Status: not available

ABG_MSc-thesis2018_028

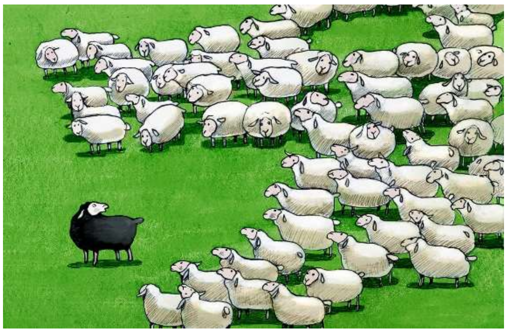
Disentangling the (antagonistic) correlation between production and disease and reproduction using markers.

Summary: Just as genetic variation can be explored using markers, i.e. determine which loci affect a trait, also genetic covariation can be explored. Some genes/loci have a positive effect on two traits while other loci have opposite effects on the two traits. The overall effect of all loci together is the genetic correlation. Once this genetic background is understood it might be easier to select for antagonistic traits. In this project you will earn how to use R to handle data files, ASreml to estimate genetic correlations and Bayz to disentangle genetic correlations using markers.
Daily supervisor: Henri Heuven

ABG_MSc-thesis2018_030

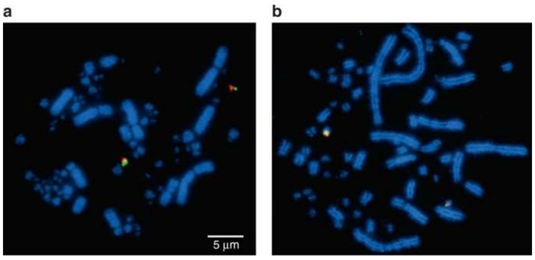
Estimation of breeding values for longevity in sheep?

Summary: Considering sustainability in a breeding program is important in various livestock species and often this is done by including longevity in breeding. Breeding for longevity requires that the date of death or culling is known, which is often not the case in some species. One could consider the last date known to be alive as a proxy of longevity. E.g. in sheep lambing date is often recorded (or birthdate of the lamb and its dam is registered). In thesis you will analyse whether last date of lambing (or weaning of the lamb) is a good proxy for longevity in the Swifter sheep breed, and if this trait can be used for breeding for higher longevity. The dataset consists of ca. 1000 ewes and their lambing dates from the last 15 years. Issues you have to deal with in this thesis are censoring (we don't know the true culling date), and connectedness between herds (although exchange of rams is taking place).
Daily supervisor: Gerben de Jong (CRV) and Bart Ducro
Status: not available



ABG_MSc-thesis2018_031

Painting chromosomes based on local ancestry



Summary: Analysis of genotype data from admixed populations provides important insights in the mapping of disease genes, in the control of population structure for genome-wide association studies, or even in the study of population genetic processes. In livestock, genotype data of animals from multiple breeds, or crosses, can be also beneficial for genomic selection. In this context of admixed populations, "painting" accurately the chromosomes of current individuals (i.e., estimating the ancestry at every locus) may be an important step for accurately performing such analyses, and several methods have been proposed. In this Msc project you will perform a literature study to make an overview of the different available methods and software, and their characteristics. You will potentially compare the efficiency of selected software using simulated datasets for crossbred animals.

Daily supervisor: Jérémie Vandenplas

Status: not available

ABG_MSc-thesis2018_033

Developing an state of the art routine for finding (regions containing) causal mutations.

Summary: Recently Sun et al. (GSE (2016) 48:77) have described how linkage information and co-segregation can be disentangled for genomic prediction. The hypothesis is that this model is also suited to discover causal mutations affecting quantitative traits. Knowledge regarding causal mutations become very important for across population predictions as well as for genome editing.

I'm looking for an enthusiastic person that likes to be trained in the field of R and Bayesian Analysis, i.e. in the field of quantitative genetics combined with statistics and programming.

Daily supervisor: Henri Heuven henri.heuven@wur.nl

ABG_MSc-thesis2018_034

Genomic evaluations that combine large number of genotyped and non-genotyped animals

Summary: Genomic evaluations are now applied in several production systems, such as those for cattle, pigs and chickens. A current challenge is to deal with the (future) large amount of genomic information. Recently, several models were proposed to combine large amount of data from genotyped and non-genotyped data. Comparison of these models is required to develop an optimal genomic evaluation that is able to deal with many genotypes. Therefore, the aim of this thesis is to review the different models proposed in the literature and to test these models using "small" datasets. During this thesis, you will learn more about genomic evaluation, and you will gain programming experience.

Daily supervisor: Jeremie Vandenplas, Mario Calus

Status: not available

ABG_MSc-thesis2018_036

Genomics in honey bees

Summary: The pedigree in honey bees differs considerably from the usual situation studied in animal breeding, like in cattle, pigs or poultry. Observations are not done on one single individual as in dairy cows, growing pigs or laying hens, but on thousands of workers, jointly called a colony. The dam of the workers, of the colony, is the queen that is the only single individual in the pedigree. The sire again is a group of animals, the group of queens producing the multitude of drones that mate with a virgin queen. This pedigree poses considerable challenges in estimation of breeding values of honey bees and in the application of genomics as well. The thesis is about the optimal design of a reference population. What 'animals' to include to reach maximum accuracy with least effort. The thesis also is about the question what avenues there are to utilise genomics. Is it primarily an issue of unravelling pedigree relationships or an issue of shorting of generation intervals, or both.

Supervision: Pim Brascamp

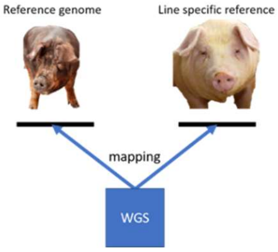
ABG_MSc-thesis2019_016

Linking microbiota profiles of human gut diseases to low performance pigs.

In this project we want to acquire published (faecal) microbiota profiles from diseased human and if available other relevant species. Subsequently, we want to investigate if similar 'disease' microbiota profiles are encountered in pigs with a low performance. In other words, are there (faecal) microbiota profiles that are predictive for a subclinical status in these pigs.

Supervisor: Dirkjan Schokker, Lisanne Verschuren

Genomics, including bioinformatics:



ABG_MSc_Thesis2022_001

Improvement in imputation accuracy with a line specific reference genome

The public reference genome of species is based on an single individual of a specific breed. In pig this is a Duroc animal. However, different breeds also have some differences in the genomes, especially with respect to structural variation. In pigs we have generated breed-specific reference genomes using long-read sequences. The line-specific reference genome should be a more accurate reference genome compared to the public reference genome. Our research question is whether this accuracy in the reference genome also leads to more accurate imputation of SNP chips to WGS level. Whole genome sequencing (WGS) is still too costly to do routinely on large number of animals. Imputation from SNP chips up to WGS is therefore a good alternative to create datasets with potential causal variants of sufficient size for e.g., GWAS studies. However, the success of such studies depends largely on the imputation accuracy. We have WGS data available from a pig breed with a line-specific reference genome. You will learn to align this data to this new line-specific reference genome with bioinformatic tools. Thereafter you will impute SNP chips up to WGS using that data aligned to the line-specific reference genome, as well as using the WGS data aligned to the public Duroc reference genome. You will ultimately compare the imputation accuracy across different genomic regions and MAF categories to identify situations were imputation accuracy is improved with a line-specific reference genome.
Contact person: martijn.derks@wur.nl

ABG_MSc_Thesis2021_006

Sugar Pig

Milk composition has been under natural selection for thousands of generations. Milk composition affects both the fitness of the mother and the offspring and therefore is optimized to transfer essential ingredients as efficient as possible from mother to offspring. Remarkably, milk contains oligosaccharides which are not digestible by the offspring and therefore seem like a waste of maternal energy. These oligosaccharides, however, provide a selective advantage to specific microbes which are specialized in digesting oligosaccharides and might be beneficial to the host. In humans, FUT2 gene variants have been associated with activity and composition of the gut microbiome. More than a hundred different milk oligosaccharides have been identified and there are large differences between and within species in milk oligosaccharide composition. Part of these differences have been attributed to genetic differences e.g. in expression of certain glycosyltransferases. In humans 4 "milk groups" are distinguished based on milk oligosaccharide composition which can be related to mutations in FUT2 and FUT3. These fucosyltransferase genes are also linked to ABO histo-blood groupings, illustrating that the function of these genes is not limited to fucosylation of milk oligosaccharides. In pigs, polymorphisms in the FUT1 gene have been associated with control of the intestinal expression of enterotoxigenic E. coli (ETEC) F18 receptors; a pathogen causing edema disease and post-weaning diarrhea (e.g. Meijerink et al 1997).
The aim of the current study is to use whole genome sequence data to identify variants in genes involved in the production of pig milk oligosaccharides. These genes include among others FUT1 and FUT2.
Supervisors: Martijn Derks & Henk Bovenhuis

ABG_MSc_Thesis2021_001

Heat stress adaptation mechanisms of chicken

Background

Global warming increases (the effect of) heat stress in livestock. Heat stress reduces welfare, health, and productivity of livestock.

Project

We have three datasets from experiments with very different experimental design: two datasets with transcriptomic data from different experimental design studies, and one epigenomic dataset with eggs instead of mature chickens (the epigenome

regulates the transcriptome, also in later life). The challenge is to perform a meta-analysis with these datasets. Expected methods to be used will be scripting (e.g. in R and / or Python) and use of diverse databases (e.g. KEGG and Genecards) for biological interpretation of the data. The results can lead to a publication.
Supervisors: Marinus te Pas, Ole Madsen, Dirkjan Schokker

ABG_MSc_Thesis2020_011

Genetic relationship between free ranging Scottish Highlanders herds on the national reserve "de Hoge Veluwe"

By using GPS data and SNP genotype (100K) data we want to investigate inbreeding but also migration between herds.
Contact person: Richard Crooijmans

ABG_MSc_Thesis2020_008

Exploring the Unmapped DNA and RNA reads in Chicken and Turkey

Summary: A key step in analysing next-generation sequencing projects is the alignment of reads to a reference genome. The mapped reads can subsequently be used for downstream analysis (e.g. variant calling). However, a significant percentage of reads frequently remain unmapped. These unmapped reads are usually ignored and discarded from the analysis process, but interesting biological information can be revealed by these data. In this thesis project you will explore the unmapped reads of whole genome sequence and RNA-sequence data from both Chicken and Turkey samples. The DNA was extracted from blood which could yield important information about viral material (i.e. the virome) within the blood samples of the sequenced animals. The goal of this project is to identify important viral, microbial, parasitic or other types of DNA in the unmapped reads of our WGS samples. More specifically, you will generate de novo assemblies of the unmapped reads, after which the generated sequence contigs can be aligned to identify the closest known matching sequence. The goal of the project is to identify genetic material which does not belong to the host sample (contaminants) and relate this to possible interesting traits (e.g. past infections). You will ultimately write a pipeline to discover sequences that are either absent or misassembled in the reference genome, and sequences that indicate infection or sample contamination.
Contact person: Martijn Derks
Status: Available

ABG_MSc_Thesis2021_009

Predictive capacity of microbiome/metabolome data on complex traits (feed efficiency or health)

Summary: Omics data, including microbiome and metabolome, are increasing within WUR as well as in the public domain. For the livestock domain we are interested in getting easy access samples, that is acquire samples in a (less) invasive way. Examples include saliva or faeces samples. The question remains what their predictive capacity is on more complex traits, such as feed efficiency or health (resilience). Within this project you will translate knowledge from the human field to livestock species. Furthermore, by searching the public domain for microbiome and/or metabolome datasets on a specific trait of interest you can test the predictive capacity. At WUR we also have some datasets that could be suited for the question at hand and can be made available.
Daily supervisor: Dirkjan Schokker

ABG_MSc_Thesis2019_024
Evolution of immune genes in livebearing fish (Poeciliidae).

The vertebrate immune system is vitally important in animal health because it confers defense against viral, bacterial, and parasitic pathogens. Immune involves many genes, and, very often, these genes are part of complex gene families. Complex gene families usually arise through gene duplication and subsequent 'neofunctionalization' or specialization. Complex gene families are quite difficult to study in large, vertebrate genomes because genome assemblies are often of poor quality. This picture of genome duplications is further complicated by a whole-genome duplication early in the evolution of the bony fish (the 'teleost-specific whole genome duplication, or TGD). More than 27,000 species of teleost fish have been described, making it by far the largest of all vertebrate groups. As more and more fish species are sequenced, it is becoming increasingly clear that striking differences can exist in the immunogenomes of fishes.

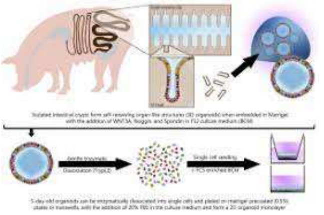
Among the bony fishes, the livebearing fishes of the family Poeciliidae (350 species), that includes the guppy, have emerged as an important model system for evolutionary research. At ABG and EZO, the genomics of adaptation in these species is studied. We recently created new genome assemblies for two species in the genus *Poeciliopsis* (*P. retropinna* and *P. turrubarensis*), a.o., using the Pacific Biosciences long read sequencing technology. These genome assemblies are among the best ever achieved for a fish species. The excellent quality of the assembled genomes enables detailed annotation and evolutionary research in complex gene families. For this project we are looking for a student interested in:

1) Improving the annotation of key immune gene families
2) Studying the evolutionary dynamics of immune-related genes in the livebearing fish

Apart from the two *Poeciliopsis* genomes, genomes of the guppy (*Poecilia reticulata*) and platy (*Xiphophorus* spp.) are available. There is also some expression data available for some of the species. The gene families of highest interest are MHC (Major Histocompatibility Complex), TLR (Toll-Like Receptors), and cytokines such as the Interleukins.

Supervisors: Hendrik-Jan Megens; Henri van Kruistum (This project is a collaboration with AFI and EZO)

ABG_MSc_Thesis2021_010
Transcriptional landscape of pig intestinal organoids



Summary: Organoids are powerful experimental models for studying biological function on the molecular level. In contrast to biomedical research in the human field, the application of organoids in farm animal research is still in its infancy. Farm animal derived organoids are of particular interest for studying important, but complex, traits such as feed efficiency (FE), disease resistance, and host-microbe interactions. In this project you will have several datasets at hand from various pig intestinal organoid experiments. In the intake we will try to match your focus on some fundamental biological phenomena, you could think of generating transcriptional landscapes for the different intestinal segments, or during health and disease. And based on these available data possibly generate data-driven gene expression networks.
Daily supervisor: Dirkjan Schokker

ABG_MSc_Thesis2019_037
Specific selection signatures in dog breeding

Summary: Currently there are nearly 400 distinct dog breeds. Dog breeds are highly divergent on appearance, body size, talent, lifespan, and many other phenotypes. These breed-specific traits were selected artificially and fixed by inbreeding. In a long history, dog breeding programs were isolated from Western countries to Oriental countries. This yielded the diversification in many traits between Western and Oriental dog breeds. Since genotype data of more than 6000 dogs from more than 160 dog breeds across the whole world is available which makes it possible to investigate the

breed-specific selection signatures. Meanwhile, in this thesis, you will collect the general standard characteristics of dog breeds, like standard height and lifespan, as your phenotypes to be investigated. Optional questions to work on: Based on the genotype data and breeds information you could detect the distinct major selection signature(s) for specific dog breeds, for instance, terrier. The measures you will work on are Fst, inbreeding level, ROH (run of homozygosity), and so on. Furthermore you will interpret how the genes underlying the selection signatures contribute to the specific traits or special dog breeds. You could also work on identifying the selection signatures linked to the geographical origin of the breeds. Another one interesting diverse trait is the average lifespans of dog breeds. To investigate the underlying genetic basis of the lifespan in dogs, you will first collect the lifespan parameters of various dog breeds, and then perform GWAS analysis to identify the genes associated with the average lifespan of the dog breeds.
Supervisors: Yun Yu, Richard Crooijmans

ABG_MSc_Thesis2019_023
Immunogenome diversity and evolution in pigs, wild boar, and related species.

The vertebrate immune system is vitally important in animal health because it confers defense against viral, bacterial, and parasitic pathogens. Immunity involves many genes, that often are part of complex gene families. Complex gene families usually arise through gene duplication and subsequent 'neofunctionalization' or specialization. Complex gene families are difficult to study in large, vertebrate genomes because genome assemblies are usually of poor quality, especially in regions that harbour such gene families. The MHC genes are a clear example of genes that are clustered and often duplicated. These duplications can even be specific for certain individuals or populations, conferring potentially unique resistance to some pathogens.

Recent advances in genome sequencing (e.g. Pacific Biosciences) have greatly improved the quality of genome assemblies of some species, such as the pig. For the pig we have large amounts of re-sequence data available that, together with the improved genome builds, could result in new insights in immunogenome diversity in pigs. With new diseases on the rise (African swine fever) and old diseases never far off (e.g. Mexican flue), understanding diversity in disease resistance in pigs and their wild counterparts is important. In this project, we intend to study immunogenome diversity in pigs by:

1) Determine the quality of the assembly and annotation of the latest pig genome build in regions that harbor important immune-related genes, and where necessary improve the annotations.
2) Use the vast collection of resequencing data present at ABG to determine the allelic diversity present at these loci. We have sequence data of many hundreds of pigs and wild boars. For pigs we also have sequence data of many rare breeds that may harbor unique resistance variation.
3) Estimate differences between individuals, populations, and species, and, where possible, infer selection on genes or pathways.

The student will choose one immune-family, e.g. MHC, TLRs, or Interleukins. Depending on the immune-gene family, further research questions could be formulated on gene evolution, untilizing sequence data that we have for related pig species, such as the Pygmy Hog, or African Warthog.
Supervisors: Hendrik-Jan Megens; Martijn Derks

ABG_MSc_Thesis2020_012
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
Contact person: Richard Crooijmans

ABG_MSc_Thesis2018_039

Whole genome bisulfite sequencing (WGBS) is a technique for profiling genome-wide DNA methylation. However, analysis of WGBS data still poses considerable challenges especially in the mapping and SNP –calling phases. Although many computer programs have been developed for mapping and calling, most of the programs have barely succeeded in improving either quantity or quality of the results. In this project we would like to compare extensively different mappers and callers to develop the best combination for WGBS data analysis by using the Parus major datasets. We will also test how this pipeline will work with reduced representation bisulfite sequencing (RRBS) data.
<u>Contact:</u> Veronika Laine (NIOO) and Ole Madsen

ABG_MSc-thesis2018_043

Estimation of de novo mutation rates in pigs using family-based sequencing analysis

<u>Summary:</u> *De novo* mutations generates new genetic variations which potential can have an effect on phenotypes. Our understanding of *de novo* mutation in farm animals is still limit, specific in terms of e.g. 1) where in the genome do *de novo* mutations arise and 2) are there local differences in *de novo* mutation rates in the genome? Whole genome sequencing data from families makes it possible to estimate *de novo* mutations. In this project whole genome sequence data of 1 family trio and 1 family quartet (parent + offspring) of pigs will be use to estimate *de novo* mutation rates. Different tools available for detecting *de novo* mutation from family data will be used and compared.
<u>Daily supervisor</u>: Ole Madsen

ABG_MSc-thesis2018_047

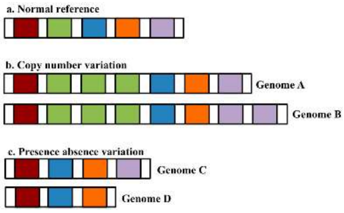
Correlations between tissue specific miRNAs and mRNA expression

FAANG
Functional Annotation of Animal Genomes

<u>Summary:</u> In recent years it has become evident that phenotypic variation is often caused by variation in gene expression. Large scale sequencing projects in humans and model organisms (ENCODE projects) have shown that gene expression is a complex process where both genetic and epigenetic variations play a role. Understanding this complex process of gene regulation is still deficient in farm animals and the need for such an understanding is not only important for gaining fundamental insight in the functional genome of farm animals but will likely also facility future breeding goals. In this study you will use two type of epigenetic data (miRNA and DNA methylation) from three different pig breeds to assess the role of these epigenetics marks in gene expression. Material & Methods: Analysis of miRNA-seq and RNA-seq data from 5 tissues from 3 individual pigs using NGS software like Tophat, Cufflinks and the miRDeep2 pipelines
<u>Daily supervisor</u>: Ole Madsen

ABG_MSc-thesis2018_048

CNV analysis in different pig breeds using 650k SNP genotypes



<u>Summary:</u> 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. Compare results with SNP genotypes from whole genome sequence data and from 60K SNP chip.
<u>Daily supervisor</u>: Martien Groenen/Richard Crooijmans

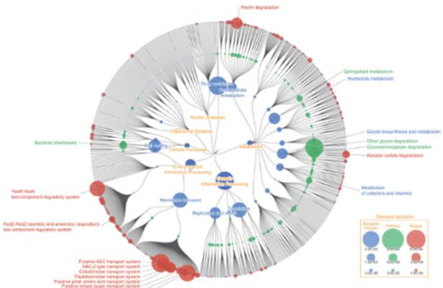
ABG_MSc-thesis2018_049
Comparison of gene copy numbers between placental and non-placental livebearing fish

Summary: Livebearing fish of the family Poeciliidae show a striking difference in the amount of nutrients that are transferred to the embryo after fertilization of the egg. In some species, the embryos start with a large amount of yolk and receive few nutrients from their mother during pregnancy (non-placental or lecithotrophic) while for other species the embryo receives many nutrients from its mother during pregnancy (placental or matrotrophic). Currently, the genomic basis for this difference is unknown. However, it is known that gene duplications can be an important driving force of the evolutionary processes that cause differences between closely related species. Using newly assembled reference genomes, you will assess the differences in gene copy numbers between placental and non-placental poeciliid fish, and test the hypothesis of gene duplications acting as a driving force of the evolution of the placenta in poeciliid fish.
Daily supervisor: Henri van Kruistum

ABG_MSc-thesis2018_052
Conservation genomics of wild boar

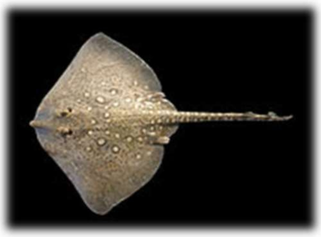
Summary: Population genomics analysis of whole genome sequence data of wild boars from different regions in Europe and Asia.
Daily supervisor: Hendrik-Jan Megens

ABG_MSc-thesis2018_055
Meta-analysis of intestinal transcriptomics data of broilers



Summary: The first aspect is to perform a literature study, focusing on the temporal and spatial development of the gut in broilers, what is already known in broilers mostly regarding the genes involved in the following processes; functional, morphological, immunological, and also the colonization of the gut microbiota. The second aspect is getting hands-on experience in data acquisition from public online repositories for transcriptomics data, as well as data handling and analysis in R. The emphasis of this project is on the functional analyses of the acquired data, such as cluster analysis and generating gene association networks. The last aspect is to combine the meta-analysis of transcriptomics data followed by functional analyses, and (re)view this in light of known literature and describe the added value of such an approach. Deliverable: Concept scientific article.
Daily supervisor: Dirkjan Schokker

ABG_MSc-thesis2019_040
Annotation and comparative genome analysis of North Sea rays (*Raja* spp.)



Summary: Elasmobranchs (sharks and rays) are very vulnerable to overfishing due to their slow reproduction. Ray species of the genus Raja are caught in high numbers as bycatch in commercial flatfish (sole, place etc) fisheries. To understand the impact of bycatch on vulnerable populations, genomics can play an important role – genomics can provide tools for population monitoring, from census size estimation, to functional effects of population decline, including human induced selection and inbreeding risk. At ABG, in collaboration with AFI, we are developing genomics tools for two Raja species (*Raja clavata*, the thornback ray, and *Raja brachyura*, the blonde ray). We have assembled the genomes of these two species, and want to further annotate and analyse these genomes from an evolutionary and functional perspective. If you are interested in bioinformatics or evolutionary genomics of aquatic species, don’t hesitate to contact us for further information.
Daily supervisor: Hendrik-Jan Megens (This project is a collaboration with AFI)

ABG_MSc-thesis2018_056
What is needed to keep a 'old' laying hen vital?

Summary: In the (near) future laying hens will be getting older (up to 100 weeks), the underlying question is how to keep these 'old' laying hens vital. Important aspects are the energy/metabolism homeostasis and immune system, as well as the calcium requirements. Underlying mechanisms can possible be steered via nutrition and/or genetics, so it is important to investigate which contribution both can have. The focus of this project is on the genes and their context, such as signalling pathways or molecular interaction networks. The aim is to generate and analyse molecular interaction networks based on 'known' literature. Subsequently, expand these networks by integrating transcriptomics data. Another aspect of this project is to elaborate on potential indicators for health, for example which genes/metabolites or organs/tissues are most discriminatory for 'health'. Deliverable: Concept scientific (review) article with case-study of network approach.
Daily supervisor: Dirkjan Schokker

ABG_MSc-thesis2018_057
Insights of the Lakenvelder phenotype in farm animals

Summary: Coat colouring in farm animals is a complex phenotype. The" Lakenvelder" phenotype is a phenotype occurring in several species. We would like to know how 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.
Daily supervisor: Richard Crooijmans
Status: available

ABG MSc Thesis2018_059
Inheritance of a chromosomal-wide inversion in great tit

Summary: Large inversions in bird genomes are known to affect different phenotypes. A striking example is the non-canonical reproductive strategy in ruff, which is associated with an inversion that in turn underlies lower testosterone levels. Moreover, inversions have a prominent role in the evolution because they can lead to strong differentiation between the inverted and non-inverted region. Recently, we identified a chromosome-wide inversion on PM1A present in 5-10% of wild great tit (*Parus major*) populations, which indicates an evolutionary role for the inversion haplotype.
This 1A inversion is indeed highly differentiated from the normal version of the chromosome 1A and until this point was demonstrated to occur just in heterozygotic state. From 2017 onwards we will start to breed birds containing the inverted chromosome 1A to allow detailed evaluation of inversion inheritance. The student shall then be able to explore whether the heterozygotic state is mandatory for birds carrying the inversion haplotype. Therefore, the student should illustrate different levels of genomic complexity within inverted and non-inverted 1A chromosomes, discuss associated phenotypes as well as their putative evolutionary roles. This is a collaborative project between Wageningen University & Research (WUR) and Netherlands Institute of Ecology (NIOO-KNAW), supervised by two researchers in each institute and will require knowledge on bioinformatics and genetics.
Supervisors: Vinicius Henrique da Silva (WUR), Mirte Bosse (WUR), Veronika Laine (NIOO-KNAW), Marcel E. Visser (NIOO-KNAW)
Status: available

ABG_MSc-thesis2019_039
Adaptation to salinity stress in Tilapia

Summary: Tilapia is an important fish species in freshwater aquaculture in tropical and subtropical regions. Over the past years fast-growing strains have been developed that have led to substantial increase in Tilapia production. However, Tilapia are grown in very diverse environments, which leads to the genetic growth potential not always being attained. One solution could be to generate environment specific strains that are better adapted to local circumstances. Generating such adapted strains will be more efficient if it is understood how the adaptation works. A highly relevant example is salinity tolerance in Tilapia. In Indonesia, there is interest to develop Tilapia strains that can be used in a polyculture system that includes shrimp. The shrimp, however, require salt-to-brackish water, which Tilapia can’t tolerate very well. It is, however, clear that Tilapia can be selected to become more tolerant to high salt content. In this project we approach this adaptation from the genomics perspective: which genes are involved, and how are these genes selected? The work includes whole genome sequencing, selective sweep analysis, and functional genome analysis, including gene expression studies.
Daily supervisor: Xiaofei Wu, Hendrik-Jan Megens

ABG MSc Thesis2018_060
The search for the missing tail

Summary: A tail-less chicken sounds like an odd. But the Drenthe fowl has been around for centuries in the Netherlands. A variant of this breed called ’Klomphoen’ is the rumpless/ tail-less Drents, as shown in the pictures. The Drenthe Fowl Bolstaart shows striking rounded rump due to the lack of tail, or to be 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, then the predators have no grip because of the absence of the tail. Here we would like to develop a Master thesis study to investigate the genetic nature underlying the tail-less 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 tail-less 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:

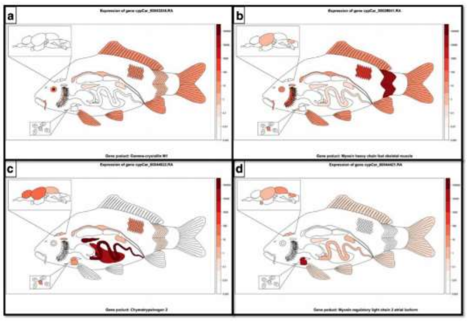
1. Basic knowledge of genomics.(variations; genome sequencing and functional annotation).
2. Working experience with linux machine(basic bash coding) and R.


© Foto: Willem en Martijn Hoeksma


© Foto: Willem en Martijn Hoeksma

Supervisors:: Zhou Wu; Richard Crooijmans
Status: available

ABG MSc Thesis2019_041
Evolutionary genomics of common carp *(Cyprinus carpio)*



Summary: The common carp *(Cyprinus carpio)* is an iconic fish and important in aquaculture and as 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 of 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, around 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 being 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 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.
Daily supervisor: Annemiek Blasweiler (AFI, ABG), Hendrik-Jan Megens (This project in collaboration with AFI)

New subjects that soon will become available in 2022 (supervisors Ole and Marta)

Functional genome analysis during development of chicken and pig

Apply allele specific analysis to detect Genome aneuploidy

Single cell RNA-seq analysis of organoids (data not yet available but suppose it will be there in mid Spring)

"Nanopore sequencing"