



GENE-SWitCH

Newsletter - Issue 2
February 2021



The GENE-SWitCH project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement N°817998.



Summary

From FAANG to fork: application of highly annotated genomes to improve farmed animal production	2
News.....	4
First Annual General Meeting of GENE-SWitCH ..	4
Joint dissemination platform in EuroFAANG	4
GENE-SWitCH Brochure	6
GENE-SWitCH Fact Sheet	6
BovReg workshop.....	6
Influence of maternal diet on epigenetic programming of offspring.....	6
New datasets from WP1 are available on BioSamples and ENA.....	7
An improved pig reference genome sequence to enable pig genetics and genomics research	8
Young researchers.....	9
Combining multiple layers of data to augment genomic prediction	9
Becoming a scientist (bioinformatics and epigenomics) within the GENE-SWitCH project.	10
GENE-SWitCH on Twitter	11
Upcoming Events.....	11

From FAANG to fork: application of highly annotated genomes to improve farmed animal production

A summary by Emily Clark (The Roslin Institute, University of Edinburgh, Co-PI in BovReg)

Members of GENE-SWitCH, AQUA-FAANG and BovReg together with colleagues from the United States recently published an editorial manuscript in the journal *Genome Biology* describing the research priorities for the next stages of the Functional Annotation of Animal Genomes consortium (FAANG)¹. These research priorities build on

1 Clark EL, Archibald AL, Daetwyler HD, Groenen MAM, Harrison PW, Houston RD, Kühn C, Lien S, Macqueen DJ, Reecy JM, Robledo D, Watson M, Tuggle CK and Giuffra E (2020) From FAANG to fork: application of highly annotated genomes to improve farmed animal production. *Genome Biol* 21, 285 (2020). <https://doi.org/10.1186/s13059-020-02197-8>

the work that the current EU and US funded FAANG projects will accomplish and are framed around meeting the challenges facing food production in the 21st Century.

The FAO predicts the world's human population will reach 9.7 billion by 2050² which will pose considerable challenges to global food production. The food production industry will need to adapt quickly to increase production capacity whilst meeting the goals of the European Green Deal³. These goals include, conserving biodiversity, enhancing animal welfare and improving the efficient use of natural resources. Sustainable improvements in animal production will be required, and high on the list of priorities will be an enhanced ability to use an animals' genomic information to predict its characteristics in specific environments. This is often termed 'genotype to phenotype' or 'G2P'. For example, some pigs may require less food to reach the same body weight than others, this reduces the resources they require and will depend on the genetics of the animals and also on the environmental conditions in which the pigs are raised. The FAANG consortium is working to discover basic functional knowledge of genome function to link G2P in farmed animals. In the manuscript we describe a set of research priorities for the next decade of FAANG research to link G2P in diverse populations of animals and apply innovative new technologies to model G2P at the cell, tissue and whole animal scale.

We outline six main research priorities (Figure 1), that will inform aspects of animal breeding and fundamental biology and ultimately lead to healthier and improved animals. 1) Capturing all available population level genomic information for a species, called a 'pangenome' approach, will facilitate the development of innovative guidelines for conservation of locally adapted breeds. 2) A 'FAANG_{GTE} Project' will build on the expectation that we will be able to collect functional genomic data, across diverse populations of animals that have extended pedigrees, for the application of statistical genomics. This information will allow researchers, in partnership with industry, to i) build models that can predict G2P more accurately, ii) improve our understanding of G2P relationships and iii) incorporate this information in existing technologies to improve the accuracy with which we can use genomic information to predict the characteristics of farmed animals. 3) New technologies are now available in the form of GPS wearables to collect information about the characteristics of farmed animals at large scale. This information is essential for linking G2P in combination with FAANG_{GTE}. It also provides a

2 The Food and Agriculture Organisation of the United Nations (FAO) (2015) *World Population Prospects*. https://population.un.org/wpp/Publications/Files/Key_Finding_s_WPP_2015.pdf

3 The European Commission (2020) *A European Green Deal*. https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal_en



foundation for genome enabled management to meet the needs of individual animals based on knowledge of their genome.

In addition to providing information that could advise breeding decisions directly, we also outline research priorities to improve our understanding of fundamental biology, linking cell, tissue and whole animal scale knowledge. For research priority 4) we will create laboratory collections known as 'biorepositories' of tissues, cells and 'organoids' which are miniature copies of organs that can be produced in the lab. These will provide a hugely valuable resource to understand genome function at a very high resolution through research priority 5) FAANG_{SingleCell}. They will also provide high-throughput systems to test or validate the information from, for example, FAANG_{GTEX} in research priority 6) using genome editing technologies. These research priorities would provide unprecedented understanding of the fundamental biology driving G2P in farmed animals, allowing us to model the effects of changing environmental conditions and supporting immunology, vaccinology, physiology, nutritional and biodiversity conservation studies.

Each of the research priorities outlined in Figure 1 will require a means to store, analyze and distribute data, and as such expanding FAANG's Data Coordination Centre will be crucial. The molecular assays required to deliver this data now cost much less than they did even five years ago. However, the costs for the research priorities outlined above remain substantial, especially considering the rapid increase in number and diversity of target species, particularly to include fish, shell fish and farmed marine invertebrates.

The timely achievement of all of the research priorities we outline in the manuscript for the next decade of FAANG will enhance the capacity of the farmed animal production industry to face the challenges of the future. We underline that a strong commitment to invest in this research is needed. New measures of international cooperation will be essential to coordinate existing projects, secure new funding and harmonize the global FAANG efforts.

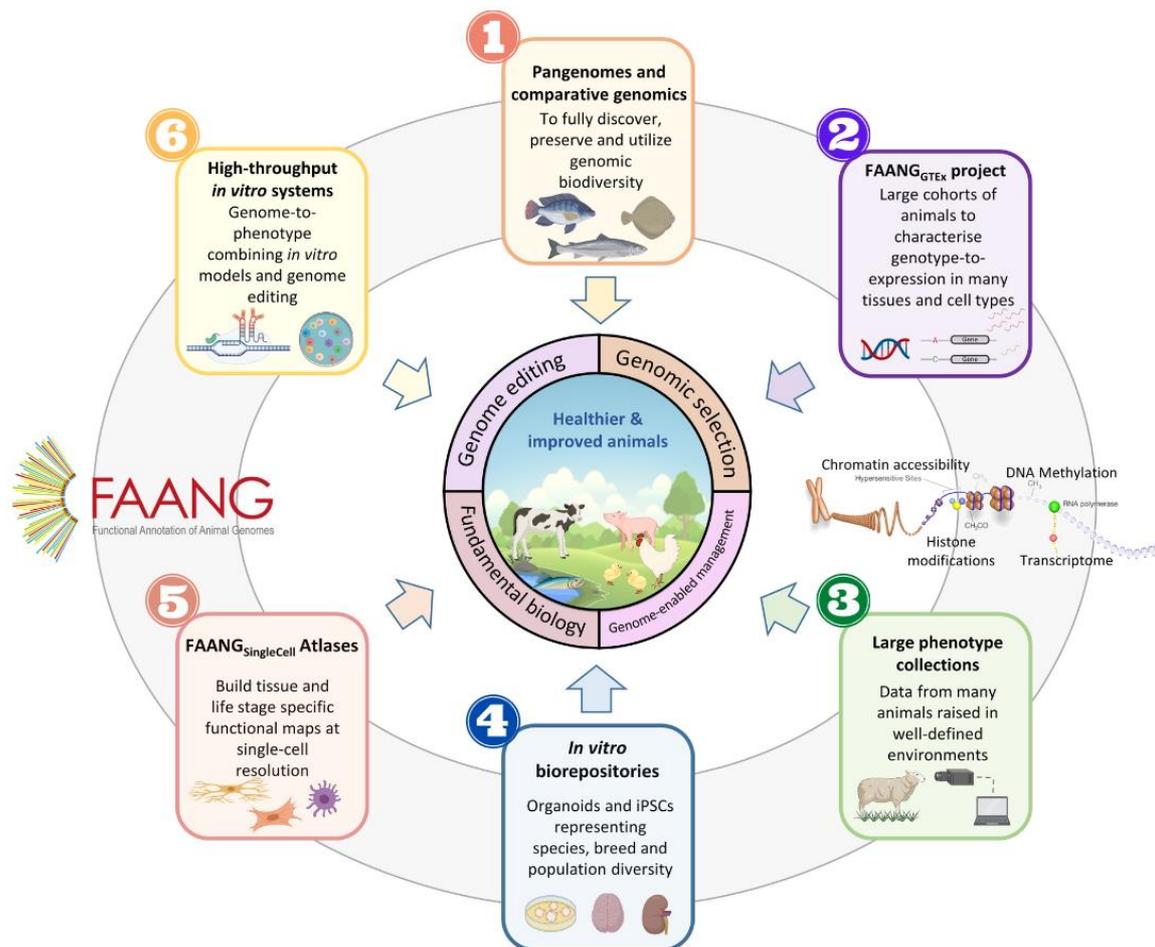


Figure 1: Priorities for the next decade of FAANG research (Clark et al (2020) Genome Biol 21, 285)



News

First Annual General Meeting of GENE-SWitCH

By Elisabetta Giuffra (INRAE)

GENE-SWitCH held its first virtual Annual Meeting on 8-9 July 2020, with general presentations on project progress by Work Package (WP) and Pillar leaders followed by workshops on specific topics. We also had a chance to meet some of our Stakeholder Advisory Board (SAB) members. We received very interesting suggestions and useful remarks from them. They appreciated the strong interconnections between Pillars and WPs in the project, with the recommendation of keeping in mind the time and skills required to exploit it at its full potential.



As the GENE-SWitCH results are inherently relevant to sustainable development and global changes regarding climate and environment, it was also recommended to disseminate them as efficiently as possible. Moreover, the planned actions to contribute to the education of the younger generation of scientists in the public and livestock industry sectors are of particular relevance.

The clustering with the other two European projects funded under the SFS30 call topic (AQUA-FANG and BovReg), with the recent establishment of EuroFAANG, offers also an important change of scale for dissemination and outreach, and as such, a strong opportunity to inspire the public at large about the genomics of FAANG.



Joint dissemination platform in EuroFAANG



By establishing EuroFAANG (the European node of the FAANG Data Coordination Centre), AQUA-FAANG, BovReg and GENE-SWitCH have formed a closer relationship to coordinate their objectives within Europe in link with the international FAANG initiative. EuroFAANG brings together a wide range of genomics, bioinformatics, modelling and open data expertise, as well as multiple platforms for dissemination and outreach.

In EuroFAANG, the projects identify joint strategies to empower research as well as for communication, dissemination and training activities. Examples are the common planning and development of collaborative analyses and the coordinated organization of training courses. EuroFAANG aims as well to maximise the stakeholders' engagement and uptake of results by targeting a wide range of stakeholder groups.

Objectives for EuroFAANG communication are:

1. Support AQUA-FAANG, BovReg and GENE-SWitCH in achieving their objectives by issuing effective, coordinated and timely communications,
2. Raising awareness within the relevant scientific communities of these initiatives
3. Promoting the tools and data resources that EuroFAANG projects make available to the community,
4. Providing clear information and call to actions to the scientific community about how they can get involved in the projects and dissemination,
5. Promote the public acceptance of the tools/techniques developed by the three projects,
6. Cross-promote the three projects whenever appropriate,
7. Encourage synergy, maximize mutual benefits and map a collaboration path between the three projects' communication and dissemination objectives,
8. Target different stakeholder groups jointly, in particular non-academic partners to increase the impact and effectiveness of the dissemination activities.



In order to streamline activities, a Communication and dissemination working group was created:

AQUA-FAANG	BOVREG	GENE-SWITCH
Cagla Kaya (EFFAB)	Riccardo Carelli (EAAP)	Cagla Kaya (EFFAB)
Lise Marie Fjellsbø (NMBU)	Johanna Vilkki (LUKE)	Riccardo Carelli (EAAP)
Peter Harrison (EMBL-EBI)	Peter Harrison (EMBL-EBI)	Camille Benard (IT)
Oana Stroe (EMBL-EBI)	Oana Stroe (EMBL-EBI)	Peter Harrison (EMBL-EBI)
		Oana Stroe (EMBL-EBI)

A training working group was also created:

AQUA-FAANG	BOVREG	GENE-SWITCH
Sigbjørn Lien (NMBU)	Christa Kuhn (FBN)	Elisabetta Giuffra (INRAE)
Dan Macqueen (Roslin)	Johanna Vilkki (LUKE)	Herve Acloque (INRAE)
Peter Harrison (EMBL-EBI)	Peter Harrison (EMBL-EBI)	Peter Harrison (EMBL-EBI)
Elisabeth Tangstad (NMBU)	Cedric Notrdame (CRG)	Sylvain Foissac (INRAE)
		Sarah Djebali (INSERM)

EuroFAANG message house for different audiences are summarized in the following headings:

Scientific audiences

EuroFAANG will create a suite of free resources that link genome to phenome to improve animal production		
Increase efficiency through precision breeding	Reduce disease resistance	Minimise environmental impact

Policy-makers

EuroFAANG leverages genomic knowledge to make animal production more robust and sustainable		
Increase efficiency through precision breeding	Reduce disease resistance	Reduce the environmental impact

General audience

EuroFAANG uses genetics to improve how we breed animals		
Feed a growing population	Reduce antibiotic use	Reduce the environmental impact

For more information, please check our communication channels:





GENE-SWitCH Brochure

The [introductory brochure](#) of the GENE-SWitCH project is now available on our website! The brochure provides compact information on the project aim, structure, partners, expected outputs and impact.

Why?
To date, most of the research has been focused on identifying genomic dynamics in tissues of adult animals. However, little is known on how the fetal genome integrates genetic information and environmental constraints. To unlock our knowledge on this dynamic integration, it is essential to characterize functional genomic elements across cells, tissues, development and a range of environments. The characterization of these elements is a non-trivial task that requires a highly advanced and costly procedure. This underlines the importance of full coordination and cooperation with ongoing projects from the FAANG community including AQUA-FAANG (new sequencing and analysis tools) and BovReg (new breeding and production) to produce a comprehensive data repository for genomic and epigenetic data to produce a readily available, openly accessible, metadata and phenotypic dataset of research's results, covering both developmental stages and adult life.

How?
Identification and characterization of functional genomic elements
Development of innovative genomic predictive models that integrate functional annotations followed by validation in commercial pig and poultry population
Profiling the epigenetic effects of maternal diet on pig herds/flocks and their possible persistence through the rearing stage
Masking the value of results for scientists, breeders and livestock production through the integration of new annotation maps to the FAANG data portal
Collaboration with projects BovReg and AQUA-FAANG to coordinate GENE-SWitCH's efforts within the global FAANG action
Enhance dissemination, communication and training activities to facilitate uptake of the GENE-SWitCH outcomes
Standardization of data and genomic annotation and outreach

A. Functional annotations across tissues and developmental stages
Early Organisms, Life Organisms, Newborn, Adult

B. Using functional annotations for precision animal breeding
Integration of genetic data, Integration of epigenetic data, Epigenetic marks, Implementation of new Genomic Selection Models, Epigenetic effects triggered in vitro by maternal diet

C. Towards more sustainable pig and poultry production systems

Impact
GENE-SWitCH will deliver new knowledge and tools for immediate transition to the pig and poultry sectors:
• Generate comprehensive genome annotation map of high-quality
• Contribute to the international FAANG cooperation on genome annotation
• Use these maps to improve genomic selection and breeding practices in monogenic species
• Pave the way to enhance the sustainability of farmed animal production

GENE-SWitCH Fact Sheet

We are happy to share that [GENE-SWitCH's Fact Sheet](#) is published! The fact sheet includes the projects' objectives and impacts in relation to the EU Green Deal and the Sustainable Development Goals (SDGs).

GENE-SWitCH will support the main objectives of EU Green Deal's Farm to Fork strategy by supporting the animal production sector to preserve affordability of its products, while fostering fair economic returns and reducing the impact on the environment and ecosystems. Understanding the biological blueprint of animals will pave the way for innovative breeding strategies supporting sustainable animal production thereby aiding the SDGs.

GENE-SWitCH
A European Research Project aiming to provide in-depth understanding of the regulation of genes and how these bring about genetic variation in pig and poultry (swine, turkey and ducks). Characterising the most important genomic elements based on key traits for sustainable production will be an important step for the European livestock sustainable farming. GENE-SWitCH will focus on identifying the dynamics of these elements, analyzing how external factors, such as nutrition and environmental surroundings, influence an animal's biological blueprint along its development from embryo to adult age.

Background
Together, poultry and pigs account for the two primary sources of meat production worldwide. With an expected world population of 9 billion by 2050, it is evident that production should be optimized while also improving the sustainability of farmed animal production to make sure we can produce enough food for not only this but also future generations.

Project results
GENE-SWitCH will deliver new knowledge and tools for immediate transition to the pig and poultry sectors:
1. NO POVERTY: Increasing our knowledge on pigs and poultry will help European producers to improve their businesses, providing a framework for improving production in underdeveloped areas.
2. ZERO HUNGER: Promoting cost-efficient and sustainable animal production will pave the way in making sources of high-quality protein available to everyone.
3. GOOD HEALTH AND WELL-BEING: GENE-SWitCH will enhance the production of high-quality animal products, thereby supporting good health and well-being of humans.
4. INDUSTRY, INNOVATION AND INFRASTRUCTURE: Training on how to use the tools developed in this project will ensure full utilization of research outcomes, supporting an innovative and sustainable industry.
5. RESPONSIBLE CONSUMPTION AND PRODUCTION: GENE-SWitCH builds trust and transparency in the food supply by explaining the project to society and helping people to understand the value of genetic information.
6. CLIMATE ACTION: The tools developed in this project will ensure the reduction of GHG emissions from pig and poultry production.
7. LIFE ON LAND: Through the availability of genetic blueprints for pigs and poultry, we can improve the management of genetic resources.
8. PARTNERSHIPS FOR THE GOALS: GENE-SWitCH brings together experts from the animal breeding and biotech industry thereby contributing to international cooperation.

Impact
GENE-SWitCH will support the main objectives of **Green Deal's Farm to Fork strategy** by supporting the animal production sector to preserve affordability of its products while fostering fair economic returns and reducing the impact on the environment and ecosystems.
Understanding the biological blueprint of animals will pave the way for innovative breeding strategies supporting sustainable animal production thereby aiding the Sustainable Development Goals.

BovReg workshop

By Sarah Djebali (INSERM), Sylvain Foissac (INRAE)

The BovReg workshop on “Reproducible genomics workflow using Nextflow and nf-core” was organized by the Centre for Genomic Regulation in Barcelona and took place online from 17 November to 20 November 2020. It was a very [complete bioinformatics workshop](#) including a Nextflow training and hackathon, keynote speaker talks and selected talks from participants. It gathered more than 80 people [Figure 1]. The event focused on computational pipelines for the analysis of sequencing data, featuring in particular Nextflow and nf-core. The new Nextflow language DSL2 was introduced by Paolo di Tommaso, creator and project leader of Nextflow. The general talks were live streamed on [BovReg Youtube channel](#) and a Slack discussion forum was available for technical questions. During this workshop, Cyril Kurylo from INRAE Toulouse, presented the [GENE-SWitCH Nextflow RNA-seq pipeline](#). This pipeline requires minimal input data, which makes it very easy to use. It is portable on multiple platforms and able to both model new transcripts and genes present in a set of samples and quantify their expression. A total of 28 members of GENE-SWitCH and other EU FAANG projects attended the event.

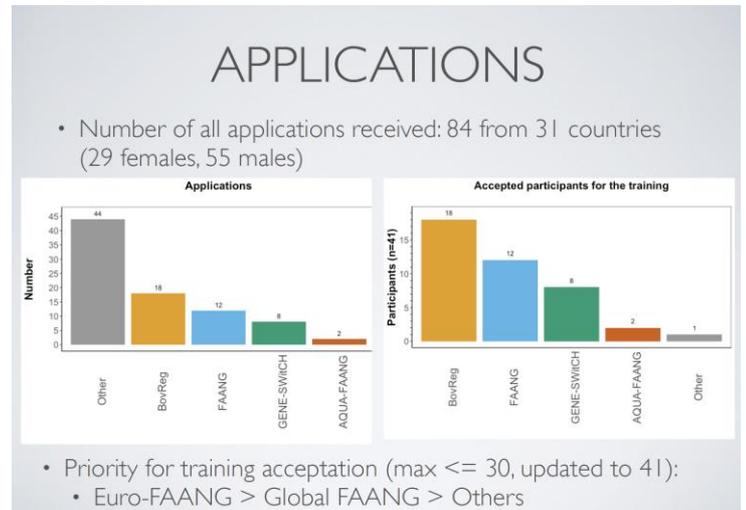


Figure 1: Participants

Influence of maternal diet on epigenetic programming of offspring

By Linda Loonen (WU)

Within WP5 of the GENE-SWitCH project, we focus on the effect of the maternal diet on the epigenome of the offspring. The epigenome describes a set of heritable “marks” (chemical modifications) that alter gene expression and are influenced by environmental variables. The diet was chosen as an environmental variable which is of relevance to sustainable animal production. The aim of this work package is to examine

the impact of the diet on the epigenome and associated gene expression in two tissues. The tissues chosen for this are the skeletal muscle and the liver. Muscle has been shown to be epigenetically modulated by diet in sheep fetuses, and the liver is a key metabolic organ. In the diets used in this WP, the levels of dietary fiber vary, leading to different amounts of short chain fatty acids (SCFA) in the mother.

The short-chain fatty acids (SCFA) acetate, propionate and butyrate, are produced by the microbiota. The amount of SCFA production by the microbiota can be influenced by the diet of the host (in this case the sow). SCFA play a role in intestinal homeostasis, helping to explain why changes in the microbiota (due to amongst others diet) can contribute to the pathophysiology of human diseases. SCFA also have effects on tissues and organs beyond the gut, through their circulation in the blood. SCFA not only signal through binding to cognate G-protein-coupled receptors on endocrine and immune cells in the body, but also induce epigenetic changes in the genome through effects on the activity of histone acetylase and histone deacetylase enzymes. Furthermore, epigenetic imprinting, the process in which DNA and/or chromatin structure is modified or altered according to the parent's epigenetic marks, likely occurs *in utero*, highlighting the importance of the maternal diet in early life. Therefore, the effect of the maternal diet will be assessed in this work package on both epigenetic and transcriptional differences in fetuses and in piglets after weaning, which will be indicative of the persistence of effects occurring *in utero*.

The animal experiment performed in this work package was done in collaboration with Schothorst feed research. In total 21 sows were impregnated twice, the first pregnancy led to pigs that were sampled at 10 weeks of age (the post-weaning group), and the second round of pregnancy led to the harvest of day 70 post fertilization fetuses (the *in utero* group). Two male and two female pigs or fetuses were used per sow. The two main tissues (liver and skeletal muscle) were collected to be used for mRNA-seq and ATAC-seq, and a biorepository was made with other tissues from these animals (see Figure 1). The biorepository is currently stored at WU, and can be used for future studies identifying links between genotype and phenotypes. Metadata was collected and will be entered into the FAANG data portal, together with the protocols.

The animal sampling is done now for this work package (many thanks to the people from Schothorst, who did an amazing amount of work, and to my colleagues from the HMI group that helped during the dissections!), what will be done in the coming months is the sequencing (ATAC-seq (Diagenode) and RNA-seq (tender pending)) and analysis of that, measurement and analysis of SCFA levels in serum of the animals, and analysis of the microbiota of the sows.

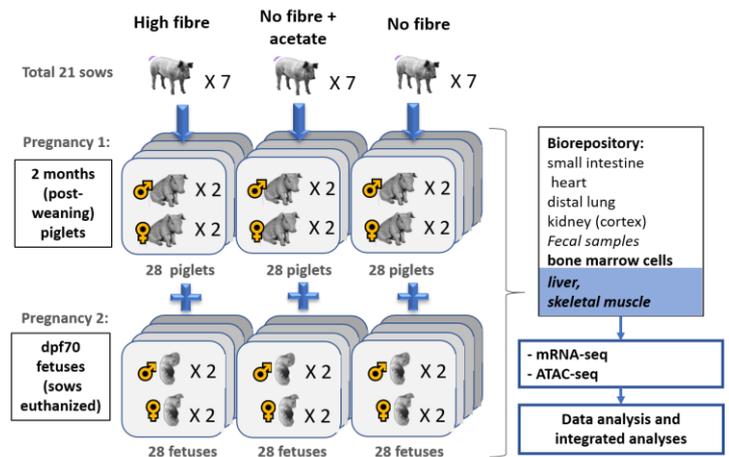


Figure 1: Overview of sampling strategy in WP5

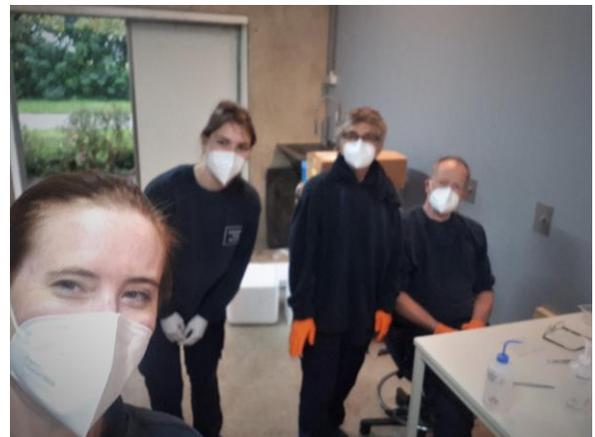


Figure 2: Sampling team ready for action

New datasets from WP1 are available on BioSamples and ENA

By *Hervé Acloque (INRAE) and Wassim Lakhel (Diagenode)*

The finalization of the sampling of biological material took place in the summer of 2020. The information relating to the biological samples has been deposited on the BioSamples database and is also accessible on the web page of the GENE-SWitCH project of the [FAANG Data portal](#). The production of sequencing data could therefore be initiated during the second half of 2020. All the DNA methylation, mRNA-seq and smallRNAseq data have been deposited on the European Nucleotide Archive and are now accessible to the scientific community following the [FAANG data sharing statement](#). ATAC-seq, long-read-RNAseq and promoter Capture Hi-C datasets will be available soon in early 2021. We warmly thank all the contributors to WP1 who did their best to deliver these valuable data, despite the current difficulties.



An improved pig reference genome sequence to enable pig genetics and genomics research

By Prof Alan L. Archibald, The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh

The most accurate map of the genetic code of pigs ever produced paves the way to more precise agriculture and biomedical research.

Scientists have decoded the whole genetic make-up of pigs – known as their genome – in a development that will facilitate the discovery of genetic variations linked to key traits. The findings could help improve selective breeding and enable more accurate use of gene-editing technologies to develop pigs with desired characteristics, such as resistance to infectious diseases. Knowledge of the genome also enables identification of genetic similarities between pigs and humans, which will enhance biomedical research in which pigs are used as models to study human health.

The study was led by scientists at The Roslin Institute and the US Department of Agriculture's, ARS Meat Animal Research Center, and involved 40 scientists from 15 laboratories in the UK and US.

Improved pig reference genome assembly

The new reference genome (Sscrofa11.1), which was built using the latest cutting edge DNA sequencing technologies, provides information on the location of more than 21,000 pig genes and is freely available [online](#). To summarize, whilst the earlier draft reference genome sequence (Sscrofa10.2) was assembled into 243,021 contigs (i.e. sections of continuous DNA sequence information), the new reference has been assembled into 1,118 contigs (i.e. many fewer much longer stretches of continuous DNA sequence) (see Figure 1).

The improved reference genome assembly represents a much better framework for the discovery of the functional elements in the genome including protein coding genes, non-coding genes and regulatory sequences that determine where and when the genes are expressed. Thus, the improved reference will greatly facilitate the parts of the GENE-SWitCH project that are concerned with describing (annotating) these functional elements. The new reference also serves not only as an improved template for discovering genetic variation but also has advantages for genomic selection, including improved accuracy for predicting genomic information

based on a reference panel (imputation). The improved quality of the reference genome also allowed the researchers to identify a further 2,500 pig genes with an evolutionary link to a human gene, increasing the known number of such genes to 15,500.

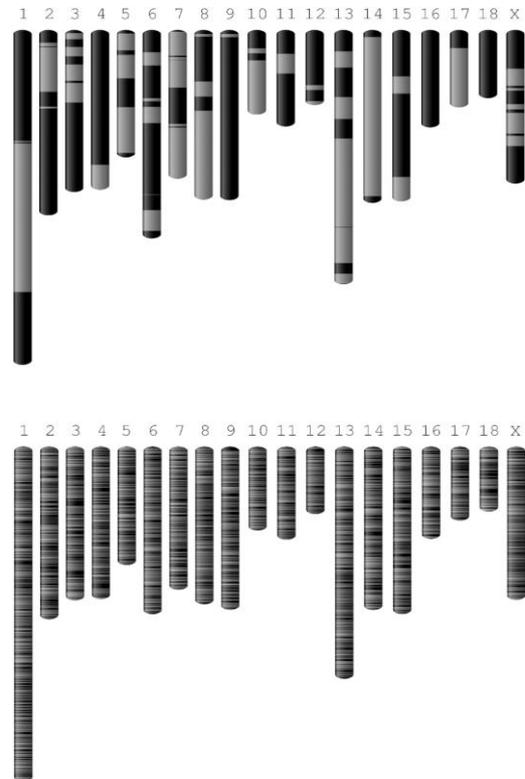


Figure 1. Diagram showing pig chromosomes with alternating grey and black bands representing the assembled contigs or fragments of sequence information for the new Sscrofa11.1 assembly (top) and earlier draft Sscrofa10.2 assembly (bottom). The genetic make-up of chromosomes 16 and 18 can each be read as a single continuous sequence.

Scientists characterized the genomes of two different pigs – a sow from the Duroc breed and a male crossbred white composite boar, and included detailed information of a further 11 pigs from European and Asian breeds that had been sequenced by Chinese researchers.

The results of the study, which was funded by the UK Biotechnology and Biological Sciences Research Council, the US Department of Agriculture, the Wellcome Trust and the Roslin Foundation, are published in the journal *GigaScience*.

Warr et al. [An improved pig reference genome sequence to enable pig genetics and genomics research](#). *Gigascience*. 2020 Jun 1;9(6):giaa051. doi: 10.1093/gigascience/giaa051. PMID: 32543654; PMCID: PMC7448572.



Young researchers

Combining multiple layers of data to augment genomic prediction

By Bruno Perez (Hendrix Genetics)

Current livestock breeding situation

The identification and selection of individuals with superior performance is crucial for the future of sustainable food production worldwide. For a long time, traditional animal breeding methods have been applied with great success by considering the individual's genealogical information to estimate their genetic merit. In the past 10 years, genomic selection has become a reality for most livestock species, making use of information from thousands of genetic markers to leverage more accurate prediction and deliver higher genetic gains after each generation. The recent technical revolution has allowed a drastic decrease in costs from state-of-art genomic technologies, making them fit for large-scale data collection in the farm. The consequence is that there are not only new sources, but also big amounts of data available in livestock breeding. However, the addition of multiple sources of data in huge volumes may also require new methods that are able to combine these in a single model to improve prediction ability.

Machine Learning and its advantages to current methods

Machine learning is an application of artificial intelligence able to build models with the ability to learn and improve from experience without being explicitly programmed. It is increasingly used to deal with big data, such as in image processing, audio recognition and text mining. Recently, machine learning models made their way into animal breeding and current results show that these methods have the potential to match or overcome other approaches commonly used by geneticists. When compared to traditional methods used in animal breeding, machine learning offers the ability to capture more complex relationships existing in data which could at least partially have a genetic basis. It is also particularly fit for combining multiple layers of data to improve performance of a single model, such as environmental features, gene expression levels and functional annotation.

Biological knowledge in genomic prediction

There has been strong interest in bridging the gap between statistical genetic models and the underlying biological mechanisms of quantitative traits. The idea

behind this concept is that multiple layers of data can help unraveling complex relationships hidden within large datasets that were not possible by only considering genetic markers and direct effects. Two strategies are commonly considered for the use of biological knowledge to improve performance of prediction: using prior knowledge on biological properties of genes to prioritize specific genetic markers and the use of other -omics data (ex.: gene expression levels) together with genetic markers. This is a developing field of animal breeding, as the best strategies to incorporate this kind of data in genomic prediction models are still being explored.

In the scope of GENE-SWitCH, we are working on the development of machine learning models that can accommodate different sources of prior biological knowledge to improve performance of genomic selection. In this context, these sources can be annotations from gene functions or their expression levels in different tissues or moments of the animal's life. At a first step we are comparing different machine learning and traditional models using publicly available data of mice. These models will be further applied to datasets generated from WP2 within the current project. Results obtained are expected to provide new insight into the integration of multiple layers of data into genomic selection models and into the applicability of such models to improve genetic gains in animal breeding programs.

For further details:

Gao *et al.* (2017). Incorporating gene annotation into genomic prediction of complex phenotypes. *Genetics*, 107,2. doi:10.1534/genetics.117.300198

Montesinos-López *et al.* (2021) A review of deep learning applications for genomic selection. *BMC Genomics*, 22:19. doi: 10.1186/s12864-020-07319-x



Bruno Perez obtained his Master's degree in Animal Breeding and Genetics at the University of São Paulo (Brazil) in 2016 and his PhD degree in 2019 at the same institution. He has joined Hendrix Genetics (HG) as a geneticist in 2019 and has been supervised by Marco

Bink (HG) and Mario Calus (WU) in activities related to WP4 in GENE-SWitCH.



Becoming a scientist (bioinformatics and epigenomics) within the GENE-SWitCH project

By Jani de Vos (Wageningen University)



After finishing the MSc in Animal Breeding and Genetics at the University of Pretoria in South Africa, Jani de Vos started her PhD at Wageningen University (WU) in The Netherlands in July 2019. The GENE-SWitCH project had captured her attention a few months before when she saw the PhD position available at

WU. From there it was a happy journey to begin this PhD, and her work is under the supervision of Dr. Ole Madsen and Prof. Dr. Martien Groenen.

Branches of this PhD

This PhD is a part of the WP2 of GENE-SWitCH, where development of standardized bioinformatics pipelines is one of the objectives. These pipelines will be used for the analyses of data from the project to predict regulatory elements as well as tissue and species specific patterns of gene regulation and comparisons between tissues and species.

This PhD combines the field of bioinformatics and (epi)genomics. Using a bioinformatics approach, we can develop pipelines for the analysis of methylation data. Such pipelines can be used to analyze raw data and receive useful results (figures, statistics and graphs). Regulatory elements in the genomes of pig and chicken in specific tissues during different phases of fetal development are not fully understood. The activation status of these elements may vary between tissues as well as during the different developmental stages. The identification of the regulatory elements controlling the genome during the different stages of development in both species contributes to the identification of the switches that differentiate the different phases; especially as some of these elements may be undetectable or inactive in adults. Comparable developmental stages in pigs and chickens will be investigated, enabling a functional comparison and enabling studying the evolution of these elements.

Process of developing in bioinformatics and epigenomics

One of my first steps was attending a bioinformatics and a genomics course within the first 6 months of my PhD which was a good kick-start for learning new skills. During the COVID-19 pandemic, together with colleagues in our

group, a python learning group was formed. This provided an opportunity to learn new basic python concepts such as loops, conditional statements and dictionaries. This has now expanded to studying an advanced book on python skills and has greatly helped in my development of coding skills. Further training in using a workflow manager (Nextflow) has been very beneficial for the development of the pipeline and understanding how to 'glue' different processes and sub-processes together. The next and hopefully final challenge is to pack the pipeline into a format that can be transferred between different computer systems. This will be done with so-called containers. Containers are very useful for a pipeline as tools used within the pipeline are stored here. This provides another way of standardizing data analysis (version of tools), as well as ease of use. A diagram of the strategy of developing the pipeline is shown in the figure. Data from the project is already arriving and the prospect of starting the analyses and unravelling the genomes of these species is an exciting prospect!

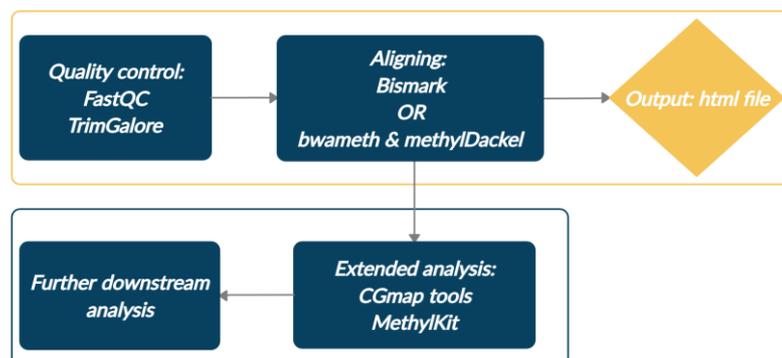


Figure: For the development of this pipeline, I will expand an already existing pipeline (yellow block) to be more robust and produce further results relevant for the GENE-SWitCH project (blue block).

Further reading and listening:

More information about the project and data types: <https://www.gene-switch.eu/project.html>

Klemm SL et al (2019) Chromatin accessibility and the regulatory epigenome. Nat Rev Genet. 20:207-220 (review)

[Nature Podcast Extra: Epigenetics.](#)

GENE-SWitCH on Twitter

Peter Harrison @PeterWHarrison · Nov 26, 2020

Great to see the FAANG to Fork editorial published in @GenomeBiology. Thanks to all for gathering (in particular @emilylouise221 for coordinating) our collective views on the priorities and future challenges for the exciting next phase of @faangomics research.

Emily Clark @emilylouise221 · Nov 24, 2020

Our new FAANG to Fork editorial was published in @GenomeBiology today! We describe a new set of research priorities to meet present and future challenges to farmed animal production that build on progress, successes and resources from the FAANG project genomebiology.biomedcentral.com/articles/10.11...

The diagram illustrates the FAANG research priorities, centered around 'Genomic editing' and 'Genomic selection' leading to 'Healthier & improved animals'. It branches into several key areas: 'Genome to phenotype combining in vitro models and genome editing', 'Animals to characterize genotype-to-expression in many tissues and cell types', 'Chromatin accessibility, DNA Methylation, Histone modifications, Transcription', 'Large phenotype collections (Data from many animals raised in well-defined environments)', 'In vitro biorepositories (Organoids and iPSCs representing species, breed and)', 'FAANG special Atlases (Build tissue and life stage specific functional maps at single-cell resolution)', and 'Comparational biology (Genome-wide association studies, GWAS)'. The FAANG logo is prominently displayed on the left.

You Retweeted

Hervé Acloque @HervePiccolo · Nov 20, 2020

Wonderful workshop and training on Nextflow and nf-core. Thanks to all the organizers, to @BovReg and @CRGenomics. Many people from @GeneSwitch also assist and contribute. See more details here, many presentations available on the BovReg youtube channel: bit.ly/3pNFQdR

The screenshot shows a slide titled 'Reproducible genomics workflows using Nextflow and nf-core' dated 17-20 November 2020. It features logos for BovReg, nextflow, nf-core, and @CRG. Below the logos, there are several social media-style posts and links related to the workshop, including mentions of 'All necessary will take place in Cambridge European Standard Time (GMT+0)', 'Use following BOVREG Project Youtube', and 'Training in @genus'.

GENE-SWitCH @GeneSwitch · Jun 25, 2020

During the @EFFAB Annual General Meeting, @GeneSwitch project's aim and progress were presented to the breeding industry members of @EFFAB and @FABRETP. #genomics #FAANG #H2020 #genome #monogastrics

The screenshot shows a presentation slide titled 'The regulatory GENoM of SWine and CHicken: functional annotation during development'. It lists the goals of the GENE-SWitCH project:

- GENE-SWitCH aims to;
 - deliver **new knowledge and tools** for immediate translation in the pig and poultry sectors
 - deliver **comprehensive genome annotation maps** of high quality, paving the way for subsequent use of these maps to **improve breeding in monogastric species**
 - Contribute to the international cooperation on genome annotation
 - On the long-term, **enhance the sustainability of farmed animal production**

The slide also includes the EFFAB logo and the website www.gene-switch.eu and the Twitter handle @GENE_SWITCH.

Upcoming Events

Due to the COVID-19 pandemic, events in 2021 have been either postponed to a later date or moved online. Physical conferences are continuously following the evolution of the COVID-19 situation and updating information regularly on the relevant websites of the events. Below is the list of the upcoming international conferences and events in 2021. Please check the websites of the events for the most updated information.

EVENT	DATE	LOCATION
World's Poultry Congress	08-12 August 2021	Paris, FRANCE
EAAP 2021	30 August - 03 September 2021	Davos, SWITZERLAND



Disclaimer: the sole responsibility of this publication lies with the authors. The European Commission and the Research Executive Agency are not responsible for any use that may be made of the information contained therein.

Copyright 2021 GENE-SWitCH Project. All rights reserved.