Hologenomic methods to account for host microbiota in genetic evaluations

Advisors:

Dr. Andrea Rau (andrea.rau@inrae.fr; INRAE GABI, Jouy-en-Josas)
Dr. Mahendra Mariadassou (mahendra.mariadassou@inrae.fr; INRAE MaIAGE, Jouy-en-Josas)
Dr. Ingrid David (ingrid.david@inrae.fr; INRAE GenPhySE, Toulouse)

- **Context**

Animals and their microbiota form a composite organism, called holobiont, which can be considered as the ultimate unit on which evolution and selection act (Theis et al., 2016). Host genes and the environment influence the colonization, development and functioning of various microbiota, which in turn help shape host phenotypes (Estellé et al., 2019). Holobiont phenotypes therefore result from the combined action of the host's genes and those of its microbiota, and their determinism can be explored by implementing hologenetic approaches capable of jointly considering host genomes and metagenomes, and possibly their interactions with each other and with the environment. Within the framework of farming systems respecting the principles of agroecology, the objectives are to reduce environmental footprints (emission of greenhouse gases, water and energy supply, reduction of inputs, etc.), the animals must thus cope with changing and complex environments (climate change, various non-competitive nutritional resources, less controlled and less protective environment against pathogens, etc.). In this context, it is important to determine the part of the host/microbiota components in the variability of the determinism of the relevant phenotypes from an agroecological perspective and to understand how the host genetics controls the symbiotic microbiota.

This thesis project is fully funded by the PEPR Agroecology and Digital project HOLOBIONTS, which aims to develop integrative hologenomic approaches for animal breeding using the most innovative technologies to generate, process and analyze sets of genetic and genomic data of the host and its microbiota as well as the phenotypes and environmental parameters in which the holobionts evolve.

- **Scientific state-of-the-art**

Although both the genome and the microbiome are known to contribute traits of agricultural interest, the latter could play a particularly important role in explaining phenotypic similarities between parents, as it is physically transmitted from mother to offspring and is under partial genetic control of the host. Therefore, incorporating metagenomic variability into genomic (and/or transmissible) prediction models holds promise for improving predictions of phenotypes and breeding values (or transmissible potential). Various genomic prediction models have been proposed in recent years to incorporate functional annotations (Mollandin et al., 2022) or intermediate omic features (Christensen et al., 2021; David et al., 2020). A key step in the latter approaches is the construction of a similarity

matrix between individuals based on microbiota information or, equivalently, the inclusion of operational taxonomic units (OTUs) as covariates with effects drawn from previous distributions. However, there are still several questions that need to be investigated to evaluate and improve these methods for use with metagenomic features. Furthermore, thorough evaluation under a wide variety of scenarios is needed to identify the most effective approaches to accurately predict genetic values (or heritable potential) useful for breeding.

- **PhD objectives and methodology**

In this thesis, we will seek to produce innovative quantitative hologenetic methods for the integration of genomic data with knowledge about the microbiota in genomic selection. In particular, we will seek to develop appropriate tools to guide answers to the following three questions:

(1) How should realistic hologenomic data be simulated?
(2) What strategy should be used for the construction of similarity matrices from the microbiota?
(3) What is the optimal combination of genomic, microbiota, and/or multi-omics data for genomic prediction or the transmissibility model?

All of the methods implemented in this thesis will be applied to the datasets generated in the PEPR HOLOBIONTS project in order to define approaches for managing diversity at the holobiont population level for sustainable agriculture systems. The methodological developments mentioned above will be carried out in close collaboration and interactions between the methodologists and biologists involved in this thesis project and the PEPR HOLOBIONTS project, in order to validate and improve the proposed methods.

In order to develop and validate our hologenomic selection approaches, we will rely on existing animal data, as well as simulated data. In particular, we will benefit from the use of data from a multigenerational Japanese quail model returned in the H2020 GERoNIMO project which is particularly well suited to guide methodological developments. An appropriate hologenomic simulation framework will be developed by the doctoral student, and will constitute one of the contributions of the thesis. In the long term, other data produced within the PEPR HOLOBIONTS will allow further validations of the approaches developed, as well as their dissemination to the scientific community

- **Working environment**

This thesis will provide doctoral training focused on the integration of complex biological information into genomic selection methods. Training modules more specific to the analysis of microbiota or genomic data can also be followed, as well as programming (in particular in R and C++) if necessary.

The thesis will be carried out on the INRAE research center at Jouy-en-Josas both in the GABI Unit (Animal Genetics and Integrative Biology) with Andrea Rau and in the MaIAGE Unit (Applied Mathematics and Computer Science from Genome to the Environment) with Mahendra Mariadassou,

as well as with remote co-supervision in the GenPhySE Unit (Genetics Physiology and Livestock Systems) with Ingrid David. We will rely in particular on regular meetings by videoconference as well as short stays in Toulouse to facilitate interactions between the doctoral student and the management team.

The doctoral student will have access to the necessary computational resources and the scientific environment of their research team in Jouy en Josas. The young scientist will be integrated into different working groups on omics data analysis and genomic prediction.

- **References**

Christensen *et al.* (2021) Genetic evaluation including intermediate omics features. Genetics 219, iyab130. https://doi.org/10.1093/genetics/iyab130

Christensen and Lund (2010) Genomic prediction when some animals are not genotyped. Genet. Sel. Evol. 42, 2. https://doi.org/10.1186/1297-9686-42-2

David *et al.* (2020) Inclusive inheritance for residual feed intake in pigs and rabbits. J. Anim. Breed. Genet. 137, 535–544. https://doi.org/10.1111/jbg.12494

David and Ricard (2019) A Unified Model for Inclusive Inheritance in Livestock Species. *Genetics,* 212 (4), 1075-1099.

Estellé (2019) Benefits from the joint analysis of host genomes and metagenomes: Select the holobiont. J. Anim. Breed. Genet. Z. Tierzuchtung Zuchtungsbiologie 136, 75–76. https://doi.org/10.1111/jbg.12383

Kraimi *et al.* (2022) Microbiota and stress: a loop that impacts memory. Psychoneuroendocrinology 136, 105594. https://doi.org/10.1016/j.psyneuen.2021.105594

Li *et al.* 2019. Integrating Gene Expression Data Into Genomic Prediction. Front. Genet. 10.

Mollandin *et al.* (2022) Accounting for overlapping annotations in genomic prediction models of complex traits. BMC Bioinformatics, https://doi.org/10.21203/rs.3.rs-1366477/v1

Pérez-Enciso *et al.* (2021) Opportunities and limits of combining microbiome and genome data for complex trait prediction. Genet. Sel. Evol. 53, 65. https://doi.org/10.1186/s12711-021-00658-7

Theis *et al.* (2016) Getting the Hologenome Concept Right: an Eco-Evolutionary Framework for Hosts and Their Microbiomes. mSystems 1, e00028-16. https://doi.org/10.1128/mSystems.00028-16

Zhang *et al.* (2020) Bayesian modeling reveals host genetics associated with rumen microbiota jointly influence methane emission in dairy cows. ISME J. 14, 2019–2033. https://doi.org/10.1038/s41396-020-0663-x

UMR1313 GABI

Animal Genetics and Integrative Biology unit

Domaine de Vilvert

78350 Jouy-en-Josas - France

+33 1 (0) 1 34 65 22 01