Title of the PhD thesis:
Improved knowledge and pan-genomic prediction of complex phenotypes using a multi-omic integrative approach in black poplar

Context:
Plants are fixed organisms, with continuous development, no pre-established germline, and highly varied life cycles ranging up to hundreds of years in trees. Thus, they are relevant biological models for studying the relationship between development and adaptation to the environment (Mladenov et al., 2021). The woody species Populus nigra, the black poplar, constitutes a reference model where so-called "omics" resources such as phenomics, transcriptomics, genomics and epigenomics have been accumulated for more than 20 years in collaboration between BioForA and the LBLGC.
Genome-wide evaluation is an innovation in the way genetic progress is managed and generated in breeding programs. However, like traditional genetic evaluation, it relies on very little consideration of the underlying genetic architecture of traits of interest. However, new resources from omics approaches are bridging the gap between the phenotypes studied and the DNA sequence, paving the way for systems approaches to integrative biology. Recently, systems biology approaches allow, through the integration of different types of data (genomic, transcriptomic), to better understand the interaction of functional factors that condition the expression of phenotypes (Chateigner et al 2020).
Epigenetics is involved in phenotypic plasticity (Maury et al., 2019), stress memory, and even transgenerational transmission and is of interest in plant breeding (Kakoulidou et al., 2021). Epigenomic data thus offer a new and innovative perspective for improving phenotypic predictions through the integration of information from heterogeneous omics data. The challenge will therefore be to decipher the complexity of quantitative traits of interest using explicit predictive models, while maintaining the predictive qualities necessary for operational selection.
This thesis project is thus placed at the interface between 2 teams (LBLGC and BioForA) for which it reinforces the synergies by developing an integrative and ambitious subject.

Objectives:
How can we integrate epigenomic, genomic and transcriptomic data to improve predictive models of complex phenotypes? What benefits would this bring to genomic prediction and to the knowledge of the functional architecture of the phenome? What variants impact the phenotype and potentially the adaptation of this species to its environment?

Work plan:
A set of 241 black poplar genotypes, representative of the natural range of this species in Western Europe, across 11 river basins in 4 countries, was planted at the INRAE experimental site in Orleans in 2008. During their fourth year of growth, a sampling of the developing wood was taken for RNA and methylome sequencing. In addition, the trees were evaluated for several

phenotypic traits (growth, phenology, morphology, ...) and biochemical properties were obtained by prediction from near infrared spectra (SPIR).
The first part of the project will consist of a literature review of methodological and empirical studies on inferring gene networks from omics data and their use for phenotype prediction. There is no single gold standard method to achieve this goal, but rather a set of approaches often combined in a sequential manner. Among these approaches, we can cite holistic approaches without a prior model, such as machine learning, and classical genetic approaches mostly based on a predefined model, such as mixed models.
A second component will consist of analyzing the data obtained from the methylome (in progress 2022). First, with a population genomics approach that will consist in describing the variation of different statistics calculated along the genome, based on the identified SMP (population epigenomics). Subsequently, the objective will be to carry out a study integrating the methodological developments previously identified in the literature review on transcriptomic, genomic and epigenomic data. These omics data will thus be used in an integrative analysis allowing the identification of loci associated with methylation variation (meQTL) (Sow et al., 2018), correlations between methylation and each omics, methylation and phenotypes or generalized canonical correlations can be applied. In addition, we will implement newly developed genomic prediction models (Wade et al., 2021) involving multiple layers of omics data.
Finally, a third component will be dedicated to validate the benefits of the previously developed pipeline in a study on a new environment. The acquisition of new data to perform this task will be done on a similar field trial installed in Toulenne (Gironde, EU INRAE, plantation 2017).

Expected results:
Different results are expected from these analyses, on the role of omic variations on the adaptation of trees to climate change, and more practically on the enrichment of prediction models directly usable in breeding and on the identification of key genetic pools for the conservation of genetic resources. An original methodological contribution on the integration of different layers of omics data is also expected.

List of references:
Amaral J, Ribeyre Z, Vigneaud J, Sow MD, Fichot R, Messier C, Pinto G, Nolet P, Maury S (2020) Advances and Promises of Epigenetics for Forest Trees. Forest. doi:10.3390/f11090976
Chateigner, A., Lesage-Descauses, M.-C., Rogier, O., Jorge, V., Leplé, J.-C., Brunaud, V., Roux, C.P.-L., Soubigou-Taconnat, L., Martin-Magniette, M.-L., Sanchez, L., et al. (2020). Gene expression predictions and networks in natural populations supports the omnigenic theory. BMC Genomics *21*, 1–16. doi:10.1186/s12864-020-06809-2
Kakoulidou, I.; Avramidou, E.V.; Baránek, M.; Brunel-Muguet, S.; Farrona, S.; Johannes, F.; Kaiserli, E.; Lieberman-Lazarovich, M.; Martinelli, F.; Mladenov, V.; Testillano, PS, Vassileva, V., Maury, S. (2021) Epigenetics for Crop Improvement in Times of Global Change. Biology, 10, 766. doi:10.3390/biology10080766
Maury S, Sow MD, Le Gac A-L, Genitoni J, Lafon-Placette C, Mozgová I (2019) Phytohormone and chromatin Crosstalk: The Missing Link For Developmental Plasticity? Frontiers in Plant Science 10, 395. doi:10.3389/fpls.2019.00395
Mladenov, V.; Fotopoulos, V.; Kaiserli, E.; Karalija, E.; Maury, S.; Miroslav, B.; Segal, N.; Testillano, P.S.; Vassileva, V.; Pinto, G.; et al. (2021) Deciphering the Epigenetic Alphabet Involved in Transgenerational Stress Memory in Crops. Int. J. Mol. Sci., 7118. doi:10.3390/ijms22137118
Sow MD, Allona I, Ambroise C, Conde D, Fichot R, Gribkova S, Jorge V, Le-Provost G, Pâques L, Plomion C, Salse J, Sanchez L, Segura V, Tost J, Maury S (2018) Epigenetics in forest trees: state of the art and potential implications for breeding and management in a context of climate change. Advances in Botanical Research 88, 387-453. doi:10.1016/bs.abr.2018.09.003

Sow MD, Le Gac AL, Fichot R, Lanciano S, Delaunay A, Le Jan I, Lesage-Descauses MC, Citerne S, Caius J, Brunaud V, Soubigou-Taconnat L, Cochard H, Segura V, Chaparro C, Grunau C, Daviaud C, Tost J, Brignolas F, Strauss SH, Mirouze M, Maury S (2021) RNAi suppression of DNA methylation affects the drought stress response and genome integrity in transgenic poplar. New Phytologist, 232, 80-97
Wade, A. R., Durufle, H., Sanchez, L., & Segura, V. (2021). eQTLs are key players in the integration of genomic and transcriptomic data for phenotype prediction. bioRxiv. doi:10.1101/2021.09.07.459279

Candidate profile:
A background in statistics, with advanced knowledge of R coding will be desirable. A knowledge of quantitative genetics, genetic and/or genomic improvement will be appreciated.

Contract conditions:
3-year contract starting October 1, 2022
Gross monthly remuneration: ~1750 €.

Administrative information:
ED Santé, Sciences Biologiques et Chimie du Vivant (SSBCV) n°549
PhD supervisor: Stéphane Maury (Professor Université d'Orléans, LBLGC)
PhD co-supervisor: Leopoldo Sanchez (Directeur de Recherches INRAE, BioForA)
Close collaborator: Harold Duruflé (Chargé de Recherches INRAE, BioForA)
Location: Orléans
LBLGC: *Laboratoire de Biologie des Ligneux et des Grandes Cultures* (team ARCHE)
https://www.univ-orleans.fr/fr/lblgc
BioForA: *Laboratoire en biologie intégrée pour la valorisation de la diversité des arbres et de la forêt* (team GA2)
https://www6.val-de-loire.inrae.fr/biofora

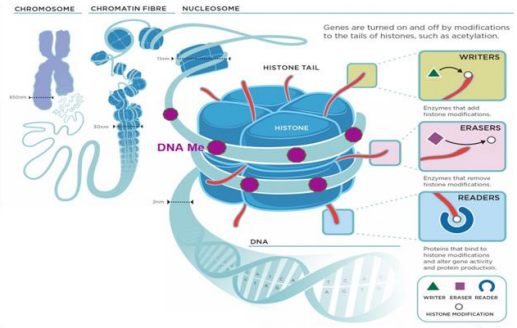
To apply:
The https://www.adum.fr/ portal should be used for the final proposal submission. Please also send your application to:

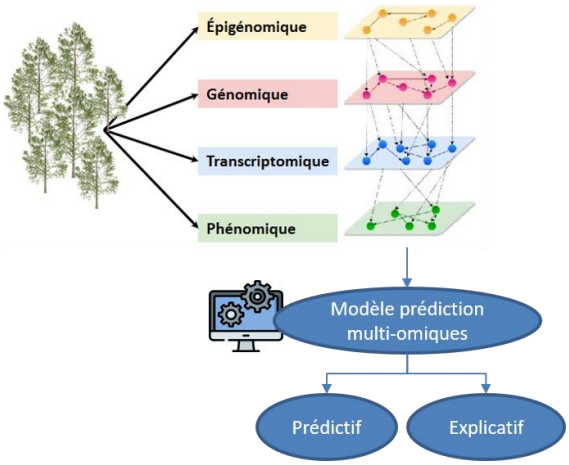
- Stéphane Maury (stephane.maury@univ-orleans.fr)
- Leopoldo Sanchez (leopoldo.sanchez-rodriguez@inrae.fr)
- Harold Duruflé (harold.durufle@inrae.fr)

The items below:
- A letter stating your motivation for this project;
- A Curriculum Vitae;
- The contact information of two references.

Application deadline: 27/04/2022.

Addition on the subject:
https://collegedoctoral-cvl.fr/as/ed/voirproposition.pl?site=CDCVL&matricule_prop=40857&langue=en
https://www6.inrae.fr/epitree-project/


CHROMOSOME
CHROMATIN FIBRE
NUCLEOSOME
Genes are turned on and off by modifications to the tails of histones, such as acetylation.
HISTONE TAIL
HISTONE
DNA Me
DNA
WRITERS
Enzymes that add histone modifications.
ERASERS
Enzymes that remove histone modifications.
READERS
Proteins that bind to histone modifications and alter gene activity and protein production.
WRITER
ERASER
READER
HISTONE MODIFICATION


Épigénomique
Génomique
Transcriptomique
Phénomique
Modèle prédiction multi-omiques
Prédictif
Explicatif