# Methodological Advances for Genome-Wide Association Studies (GWAS): multi-locus and multi-trait mixed-models

Vincent Segura

## ▶ To cite this version:

Vincent Segura. Methodological Advances for Genome-Wide Association Studies (GWAS): multi-locus and multi-trait mixed-models. Séminaires de Biologie des Plantes, Nov 2014, Montpellier, France. hal-02793220

**HAL Id: hal-02793220**
**https://hal.inrae.fr/hal-02793220v1**

Submitted on 5 Jun 2020

## *Seminar on association genetics in plants*

**Mercredi 26 novembre 2014**

*Amphithéatre Philippe Lamour*

*Campus INRA SupAgro*

*2 place Viala, 34060 Montpellier*

*9:30 - 10:45 :* ***Methodological Advances for Genome-Wide Association Studies (GWAS): multi-locus and multi-trait mixed-models***

**Vincent Segura**
*Amélioration, Génétique et Physiologie Forestières*
*INRA Val-de-Loire, Orléans, France*

Genome-Wide Association Studies (GWAS) are among the most popular methods to study the genetic architecture of complex traits. They generally consist of testing associations between polymorphisms and one or several phenotypes in populations of individuals deemed to be independent. But in practice, the individuals under study are not independent and the resulting population structure causes long range linkage disequilibrium *i.e.* between regions (loci) not physically linked, leading to the detection of false positives in association studies. Inspired by Fisher's infinitesimal model and used for decades in animal genetics, mixed linear model (MLM) was proposed almost 10 years ago in the context of GWAS and is now considered as the default method to overcome the bias attributed to population structure confounding. However in its current use, the performance of the MLM is quite limited because (i) it does not test more than one polymorphism at a time which may not be suitable for complex traits controlled by several large effect loci, and (ii) it does not consider more than one phenotype at a time and therefore cannot take advantage of potential correlations between phenotypes collected on the same individuals to detect pleiotropic loci. To address these limitations, we have recently proposed two extensions of MLM: MLMM and MTMM which respectively deal with multiple loci and phenotypes. Through dedicated simulations, we have demonstrated that both approaches outperform the existing methods both in terms of power and false discovery rate. Consequently, their application to real data in humans and *Arabidopsis thaliana* has identified new associations and in particular cases of allelic heterogeneity or pleiotropy. Furthermore, by applying MTMM to flowering time data collected in *A. thaliana* grown under several environmental conditions, we have identified a locus whose effect depends on the environment, emphasizing the relevance of this approach for detecting genes involved in phenotypic plasticity. Both approaches have been efficiently implemented, enabling the analysis of large datasets (several thousands of individuals genotyped for hundreds of thousands of markers).

*Organized by the Biochimie et Physiologie Moléculaire (B$_{\&}$PMP) Department*

*Contacts: C. Maurel (maurel@supagro.inra.fr), H. Rouached (hatem.rouached@supagro.inra.fr) and S. Mari (stephane.mari@supagro.inra.fr)*