

Atelier de prédiction et sélection génomique

J. David and Co :voir liste ci-dessous

02/02/2024

Abstract

Ce document présente les objectifs, les pré-requis et le programme de l'atelier de prédiction et sélection génomique. Cet atelier est destiné aux étudiants en spécialisation APIMET et SEPMET de l'institut Agro Montpellier, PIST d'AgroParisTech et l'option Amélioration des plantes de l'institut Agro Rennes-Angers.

L'équipe pédagogique de 2024 : Renaud Rincent, Marianne Laurençon, Marion Gauthier, Mathieu Tiret, Sarah Bensadoun, Nicolas Salas, Christopher Sauvage, Anne Laperche et Jacques David.

Un peu d'histoire : En 2024, ce sera la 9^{ème} édition de cet atelier construit au départ par Timothée Flûtre, Jacques David, Friedrich Longin et Philippe Brabant, fortement amélioré par Julien Diot pour le jeu sérieux. Cette année encore, nous aurons les étudiants des 3 écoles, AgroParisTech, l'institut Agro de Montpellier et de Rennes.

Contents

1	Date, lieu et contacts	1
2	Objectifs de l'atelier	2
3	Pré-requis	2
4	Equipement et configuration	3
5	Ressources en ligne	3
6	Programme de la semaine (les horaires sont susceptible de varier)	4

1 Date, lieu et contacts

- Du lundi 12 au vendredi 16 février 2024.
- Ca se passe là : Institut Agro Rennes-Angers, bat 8 Oryza, (<https://www.institut-agro-rennes-angers.fr/campus/environnement-et-acces/plans-dacces-et-des-campus>)
- Contact [Jacques David](#) ([Institut Agro Montpellier](#), équipe [GE2POP](#)),

Equipe pédagogique 2024 :

Renaud Rincent : renaud.rincent@inrae.fr (INRAE, [UMR GQE](#)), Julie Fiévet (INRAE, [UMR GQE](#)),
Marianne Larençon : marianne.laurencon@inrae.fr (INRAE, [UMR IGEPP](#)), Marion Gauthier :

marion.gauthier@agrocampus-ouest.fr (Institut Agro Rennes-Angers, Mathieu Tiret : mathieu.tiret@inrae.fr (INRAE, UMR IGEPP), Nicolas SALAS : nicolas.salas@supagro.fr (Institut Agro Montpellier), Sarah Ben-sadoun : sarah.ben-sadoun@agroparistech.fr (AgroParisTech, UMR GQE), Anne Laperche : anne.laperche@agrocampus-ouest.fr (Institut Agro Rennes-Angers, UMR IGEPP), Christopher Sauvage : Christopher.Sauvage@syngenta.com (Syngenta)

Crédits

Timothée Flutre ([INRA](#), UMR [GQE](#)), Philippe Brabant, Jacques David, Julien Diot

2 Objectifs de l'atelier

Il vise à familiariser les étudiants avec la prédiction génomique de caractères quantitatifs et son utilisation dans les programmes de sélection. Il se déroule en trois étapes.

1. Les bases de la prédiction génomique, par Renaud Rincet (INRAE, GQE).
 - Simulation de données phénotypiques selon différentes architectures génétiques
 - Présentation des principales méthodes d'inférence des effets génétiques
 - Exploration de la qualité des prédictions en fonction de paramètres tels que la structure du matériel végétal, la densité de génotypage, etc.
2. "Jouez au sélectionneur !"
 - Constitution d'équipes de sélection
 - Elaboration d'une stratégie de sélection mêlant essais au champ, génotypage et prédiction
 - Réalisation du schéma de sélection sur une succession de plusieurs années simulant la temporalité et le budget d'un véritable programme de sélection
 - Mise en commun, examen des résultats et comparaison des stratégies
3. "Utilisation de la sélection génomique par les sélectionneurs", par Christopher Sauvage (Syngenta)
 - Présentation des méthodes d'utilisation de la sélection génomique par les sélectionneurs
 - Les avantages et les limites en pratique de la sélection génomique
4. "Optimization of breeding schemes with genomic selection in R", par Jacques David (Institut Agro Montpellier)
 - Présentation de la théorie
 - Mise en pratique via des simulations de schémas de sélection

3 Pré-requis

Les étudiants doivent avoir révisé assez sérieusement les notions suivantes.

1. Les bases de la génétique quantitative
 - 1.1. Décomposition de la variance et notion d'héritabilité
 - 1.2. Effets additifs d'un allèle, modèle bi-allélique de Fisher

- 1.3. Prédiction de l'effet additif par la régression du nombre d'allèles favorables
- 1.4. Covariance entre apparentés et construction d'une matrice de simple apparentement

2. Les outils statistiques suivants

- 2.1 Modèle linéaire mixte (à effets fixes et aléatoires)
- 2.2 Estimation des composantes de la variance par le modèle mixte
- 2.3 Utilisation des matrices d'apparentement dans le modèle mixte

3. Avoir revu les TPs

- 3.1 Estimation des composantes de la variance
- 3.2 Génétique d'association
- 3.3 Prédiction génomique

4. Logiciel R

- 4.1 Bases de la programmation en R
- 4.2 Boucles et stockage de valeurs dans un objet
- 4.3 Ecriture d'une fonction
- 4.4 Ecriture de rapports au format Rmd

4 Équipement et configuration

Les étudiants doivent venir chacun avec leur ordinateur portable avec la configuration suivante:

1. système d'exploitation Windows, Mac ou GNU/Linux;
2. logiciels [R](#) (version ≥ 3) et [RStudio](#) (version ≥ 1), et au moins les paquets suivants:
 - [selection](#): pour la prédiction du gain génétique;
 - [knitr](#), [rmarkdown](#): pour les rapports dynamiques;
 - [MASS](#): pour les simulations multivariées;
 - [lme4](#): pour l'analyse de données phénotypiques via des modèles mixtes
 - [rrBLUP](#), [BGLR](#): pour la prédiction génomique.
 - [cvTools](#): pour la validation croisée
 - [gtools](#): pour trier de manière alphanumérique
 - [MM4LMM](#): algorithme MM pour ajuster une régression par ML ou ReML

5 Ressources en ligne

Ce document, ainsi que les autres nécessaires à la réalisation de l'atelier, se trouve à l'adresse suivante: <https://github.com/timflutre/atelier-prediction-genomique>. Le copyright appartient à l'Institut Agro Montpellier et à l'Institut national de recherche pour l'agriculture, l'alimentation et l'environnement. Le contenu du répertoire est sous licence [Creative Commons Attribution-ShareAlike 4.0 International](#). Veuillez en prendre connaissance et vous y conformer (contactez les auteurs en cas de doute).

6 Programme de la semaine (les horaires sont susceptible de varier)

6.1 Lundi

- 9h30 - 12h: introduction à la modélisation statistique via la simulation en prenant la régression linéaire simple comme exemple, par R. Rincenc (document “premiers-pas.pdf”)
- 14h - 17h: extension au modèle linéaire mixte, puis exploration par simulation de la prédiction génomique, par R. Rincenc (document “prediction-genomique.pdf”) Cette année, pour la première fois, introduction à la prédiction phénotypique.

6.2 Mardi

- 9h - 12h: investigation par simulation, en groupes, des aspects influençant la précision de prédiction
- 14h - 15h30: présentation des résultats et discussion (Sarah BenSadoun & Mathieu Turet)
- 15h30 - 17h: présentation du jeu sérieux sur la sélection, constitution des équipes et réflexions stratégiques

6.3 Mercredi

- 9h - 12h: lancement du jeu (1 heure = 1 année)
- 14h - 17h: suite du jeu (1 heure = 1 année)

6.4 Jeudi

- 9h - 11h40: suite du jeu (40 minutes = 1 année)
- 11h40 - 12h: procédure d’inscription comparant les génotypes proposés des équipes avec les témoins
- 13h30 - 15h: discussion sur les différentes stratégies suivies

6.5 Vendredi

- 9h - 12h: « De la théorie à la routine, comment la sélection génomique est devenue essentielle en sélection variétale – exemples de succès et de limites ». Christopher Sauvage (PhD) – Scientifique en Génétique Appliquée - Syngenta SA France
- 14h - 17h:
 - Cours “Optimization of breeding schemes with GS in R”, par J. David (document “Selgain_Montpellier_2022.pdf”)
 - Travaux pratiques via des simulations de schémas de sélection (document “Example_selgain_students_Supagro.R”)
 - Discussion sur les acquis réalisés durant l’atelier et les améliorations à apporter