

Cours 1 – Introduction à la génétique quantitative

Jemay Salomon

UMR GQE Le Moulon

Université Paris-Saclay, INRAE, CNRS, AgroParisTech

Outline

Contexte et motivation

Concept théorique

Caractères quantitatifs versus caractères qualitatifs

Concept statistique

Modèle de base

Application à l'amélioration des plantes

Activité - 45 min

Case of varietal mixtures

Case of crop mixtures

Material and methods

Results

Conclusions and perspectives

Outline

Contexte et motivation

Concept théorique

Caractères quantitatifs versus caractères qualitatifs

Concept statistique

Modèle de base

Application à l'amélioration des plantes

Activité - 45 min

Case of varietal mixtures

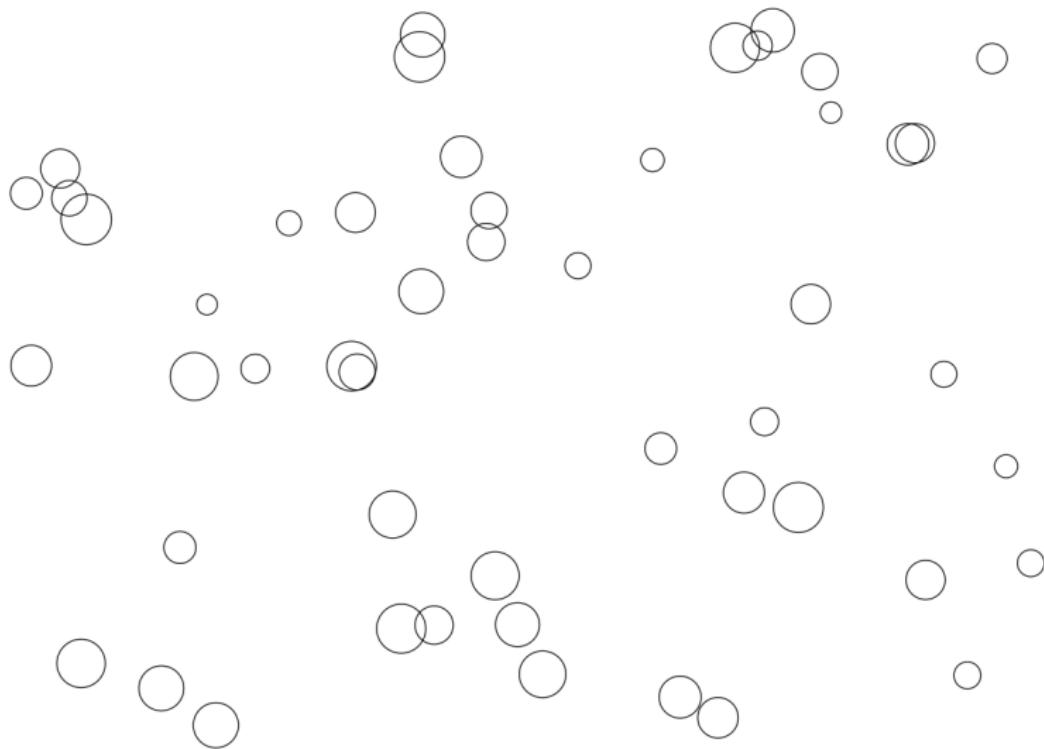
Case of crop mixtures

Material and methods

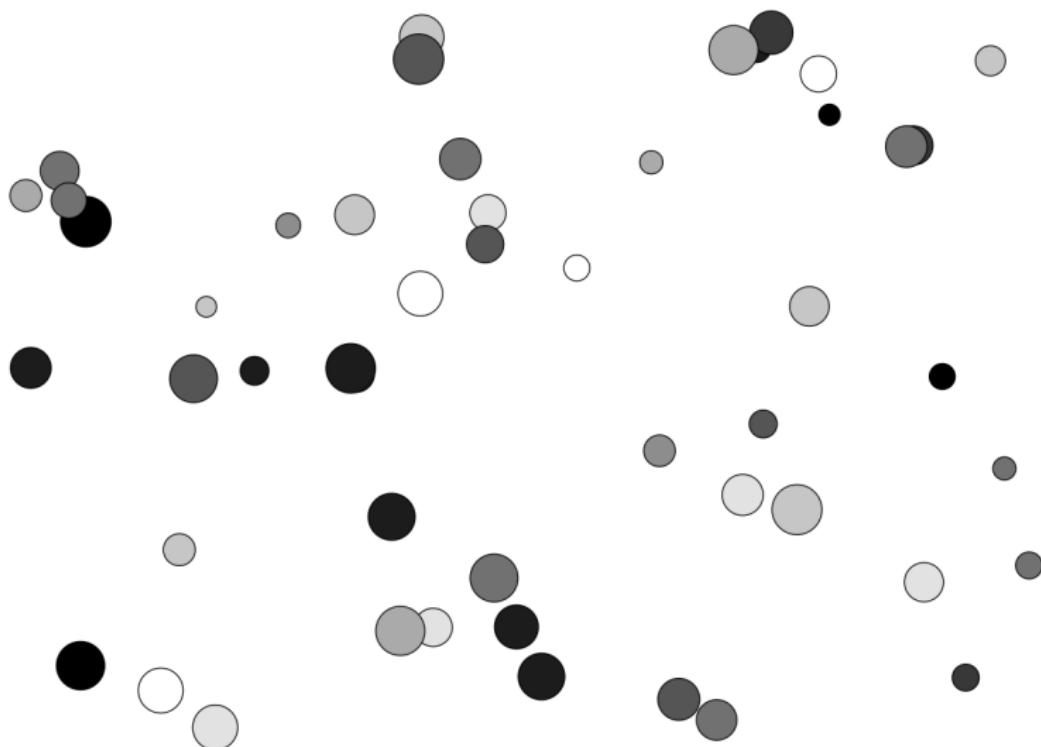
Results

Conclusions and perspectives

Population



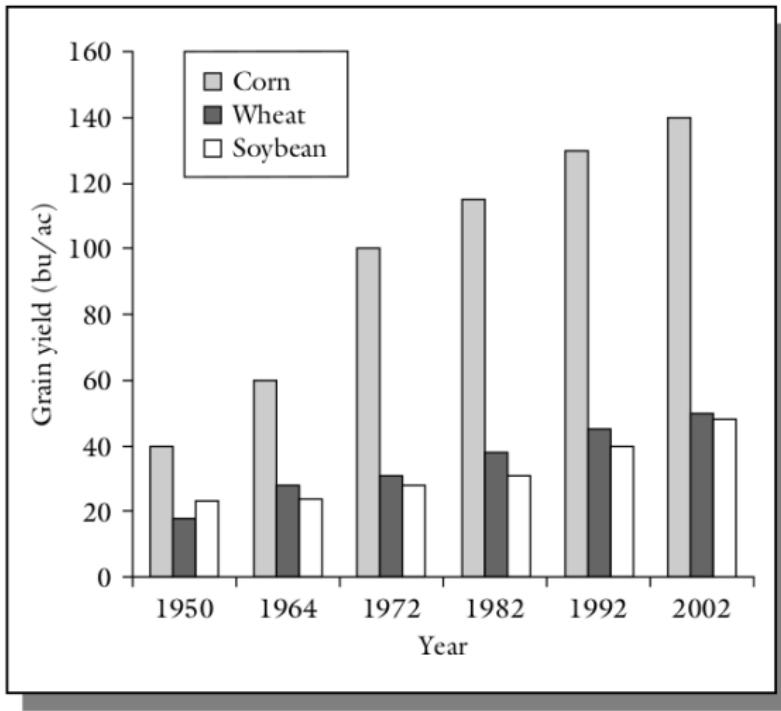
Variation phénotypique



Variation phénotypique



Consequences



cf-Principles of Plant Genetics and Breeding (2012)

Outline

Contexte et motivation

Concept théorique

Caractères quantitatifs versus caractères qualitatifs

Concept statistique

Modèle de base

Application à l'amélioration des plantes

Activité - 45 min

Case of varietal mixtures

Case of crop mixtures

Material and methods

Results

Conclusions and perspectives

Déf.

La génétique quantitative vise à relier la variation des traits complexes à leur base génétique, pour mieux comprendre et prédire l'architecture génétique et l'évolution des populations sur le long terme.

Déf.

- ▶ Génétique quantitative classique

Déf.

► Génétique quantitative classique

- S'intéresse à la variation phénotypique globale et à sa base génétique, en considérant l'ensemble des gènes.

Déf.

► Génétique quantitative classique

- ▶ S'intéresse à la variation phénotypique globale et à sa base génétique, en considérant l'ensemble des gènes.
- ▶ Analyse holistique de tous les gènes, sans distinguer contributions majeures ou mineures.

Déf.

► Génétique quantitative classique

- ▶ S'intéresse à la variation phénotypique globale et à sa base génétique, en considérant l'ensemble des gènes.
- ▶ Analyse holistique de tous les gènes, sans distinguer contributions majeures ou mineures.
- ▶ Fournit une vision globale de la génétique des traits complexes.

Déf.

► Génétique quantitative classique

- ▶ S'intéresse à la variation phénotypique globale et à sa base génétique, en considérant l'ensemble des gènes.
- ▶ Analyse holistique de tous les gènes, sans distinguer contributions majeures ou mineures.
- ▶ Fournit une vision globale de la génétique des traits complexes.

► Génétique quantitative moléculaire

Déf.

► Génétique quantitative classique

- ▶ S'intéresse à la variation phénotypique globale et à sa base génétique, en considérant l'ensemble des gènes.
- ▶ Analyse holistique de tous les gènes, sans distinguer contributions majeures ou mineures.
- ▶ Fournit une vision globale de la génétique des traits complexes.

► Génétique quantitative moléculaire

- ▶ Se concentre sur l'association entre les sites d'ADN polymorphes et les variations phénotypiques.

Déf.

► Génétique quantitative classique

- ▶ S'intéresse à la variation phénotypique globale et à sa base génétique, en considérant l'ensemble des gènes.
- ▶ Analyse holistique de tous les gènes, sans distinguer contributions majeures ou mineures.
- ▶ Fournit une vision globale de la génétique des traits complexes.

► Génétique quantitative moléculaire

- ▶ Se concentre sur l'association entre les sites d'ADN polymorphes et les variations phénotypiques.
- ▶ Analyse détaillée de l'architecture génétique : gènes majeurs (analyse ciblée) et gènes mineurs (vue globale).

Déf.

► Génétique quantitative classique

- ▶ S'intéresse à la variation phénotypique globale et à sa base génétique, en considérant l'ensemble des gènes.
- ▶ Analyse holistique de tous les gènes, sans distinguer contributions majeures ou mineures.
- ▶ Fournit une vision globale de la génétique des traits complexes.

► Génétique quantitative moléculaire

- ▶ Se concentre sur l'association entre les sites d'ADN polymorphes et les variations phénotypiques.
- ▶ Analyse détaillée de l'architecture génétique : gènes majeurs (analyse ciblée) et gènes mineurs (vue globale).
- ▶ Permet de comprendre comment des loci spécifiques contribuent aux variations, utile pour la sélection ou la prédiction génétique.

Outline

Contexte et motivation

Concept théorique

Caractères quantitatifs versus caractères qualitatifs

Concept statistique

Modèle de base

Application à l'amélioration des plantes

Activité - 45 min

Case of varietal mixtures

Case of crop mixtures

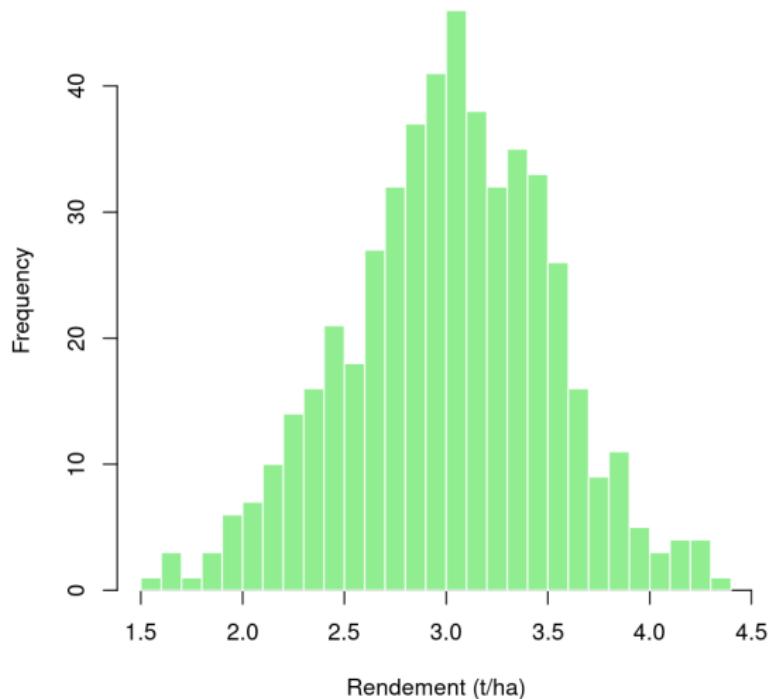
Material and methods

Results

Conclusions and perspectives

Distribution des caractères quantitatifs et qualitatifs

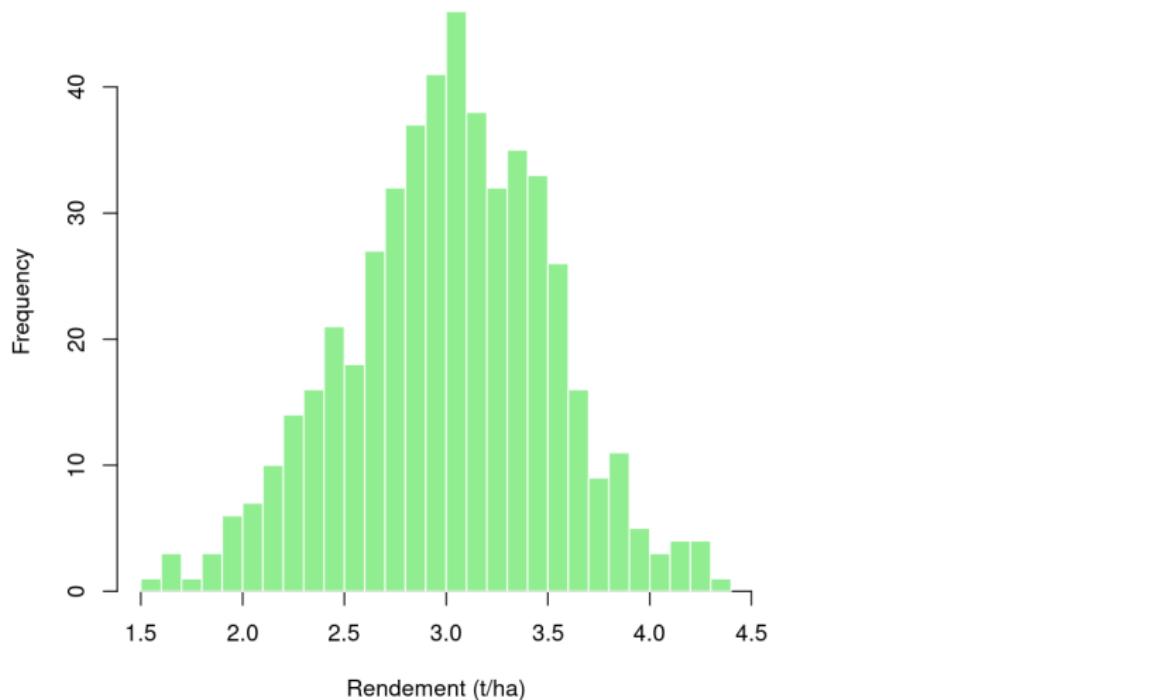
ex. rendement



Distribution des caractères quantitatifs et qualitatifs

ex. rendement

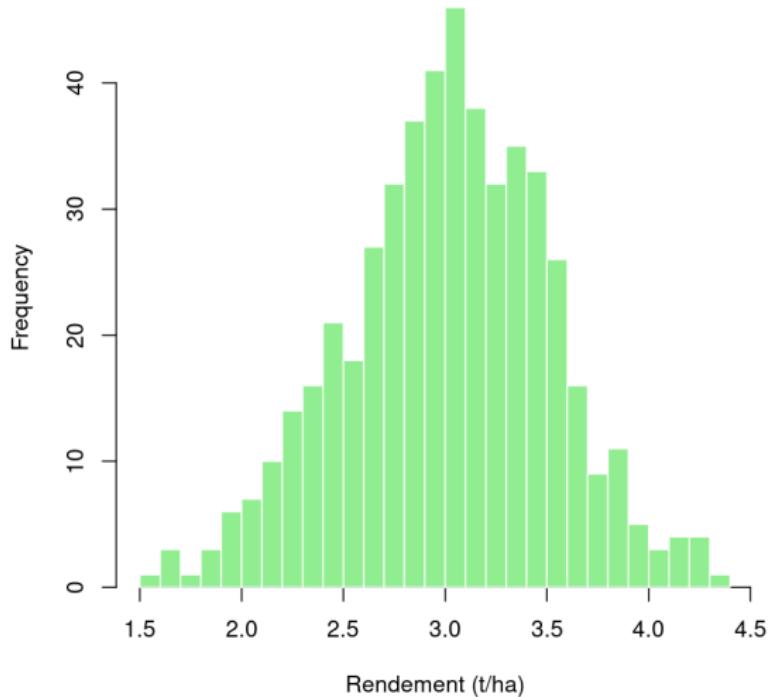
● Variation continue



Distribution des caractères quantitatifs et qualitatifs

ex. rendement

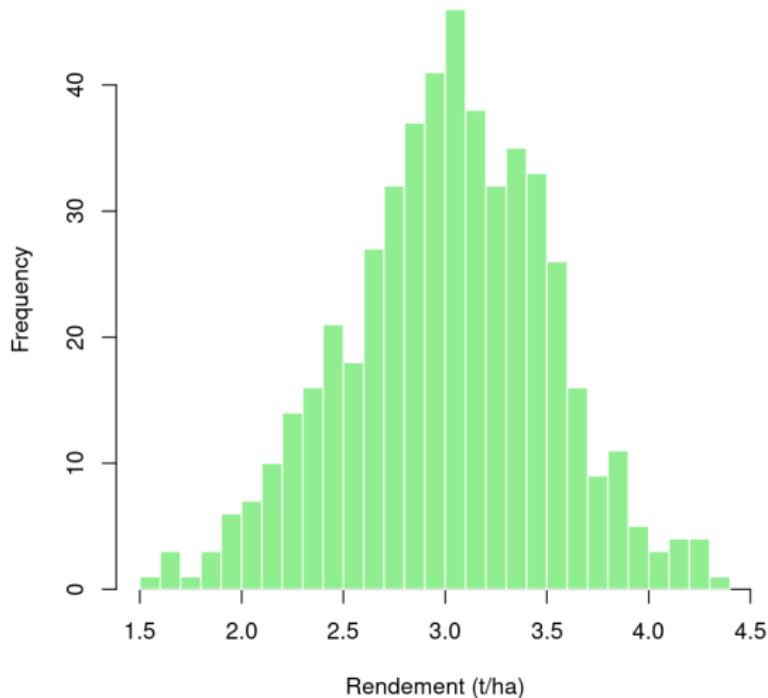
- Variation continue
- Sous contrôle de nombreux gènes



Distribution des caractères quantitatifs et qualitatifs

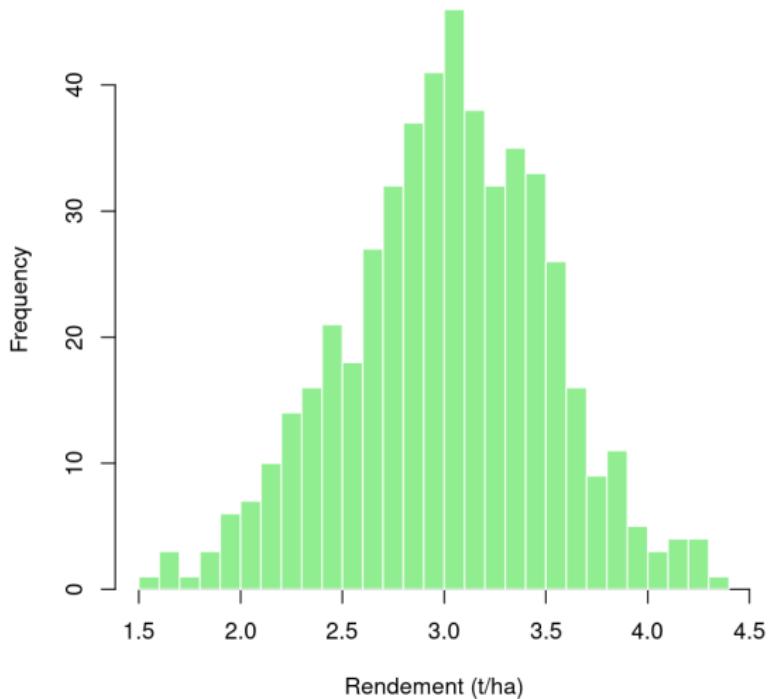
ex. rendement

- Variation continue
- Sous contrôle de nombreux gènes
- Influencer largement par l'environnement
- ...



Distribution des caractères quantitatifs et qualitatifs

ex. rendement

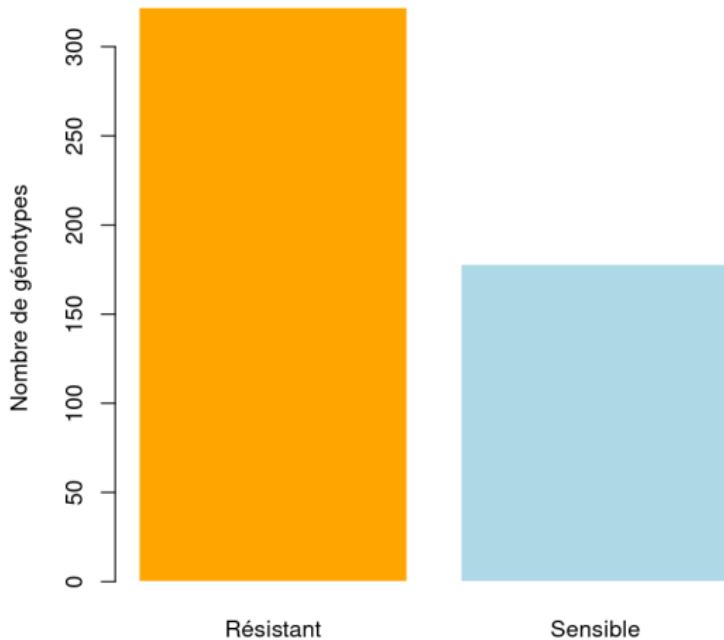


- Variation continue
- Sous contrôle de nombreux gènes
- Influencer largement par l'environnement
- ...
- Ex. Hauteur des plantes, diamètre tige, etc...

Distribution des caractères quantitatifs et qualitatifs

ex. résistance/sensibilité

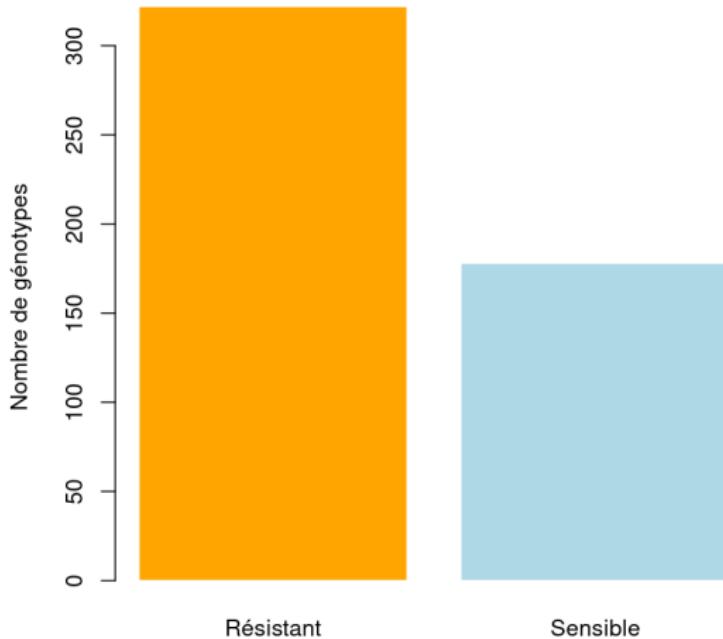
- Variation discontinue



Distribution des caractères quantitatifs et qualitatifs

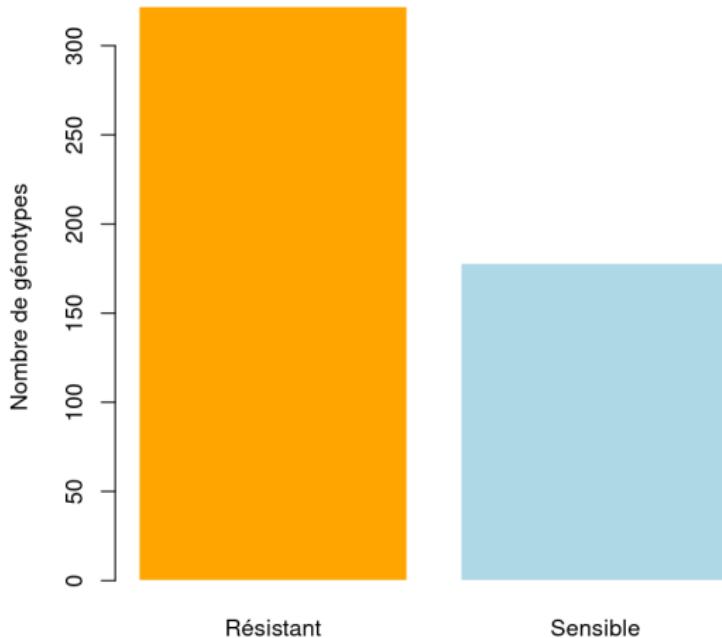
ex. résistance/sensibilité

- Variation discontinue
- Sous contrôle d'un ou peu de gènes



Distribution des caractères quantitatifs et qualitatifs

ex. résistance/sensibilité

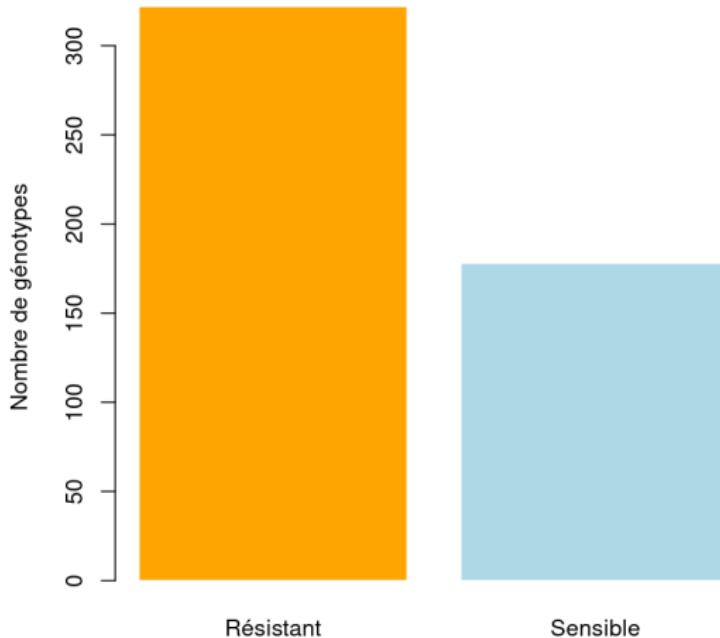


- Variation discontinue
- Sous contrôle d'un ou peu de gènes
- Peu ou pas d'influence de l'environnement

...

Distribution des caractères quantitatifs et qualitatifs

ex. résistance/sensibilité



- Variation discontinue
- Sous contrôle d'un ou peu de gènes
- Peu ou pas d'influence de l'environnement
- ...
- Ex. Couleur des fleurs, résistance/sensibilité, etc..

Outline

Contexte et motivation

Concept théorique

Caractères quantitatifs versus caractères qualitatifs

Concept statistique

Modèle de base

Application à l'amélioration des plantes

Activité - 45 min

Case of varietal mixtures

Case of crop mixtures

Material and methods

Results

Conclusions and perspectives

Outline

Contexte et motivation

Concept théorique

Caractères quantitatifs versus caractères qualitatifs

Concept statistique

Modèle de base

Application à l'amélioration des plantes

Activité - 45 min

Case of varietal mixtures

Case of crop mixtures

Material and methods

Results

Conclusions and perspectives

Outline

Contexte et motivation

Concept théorique

Caractères quantitatifs versus caractères qualitatifs

Concept statistique

Modèle de base

Application à l'amélioration des plantes

Activité - 45 min

Case of varietal mixtures

Case of crop mixtures

Material and methods

Results

Conclusions and perspectives

Outline

Contexte et motivation

Concept théorique

Caractères quantitatifs versus caractères qualitatifs

Concept statistique

Modèle de base

Application à l'amélioration des plantes

Activité - 45 min

Case of varietal mixtures

Case of crop mixtures

Material and methods

Results

Conclusions and perspectives

- Choisir un article
- Donner le contexte/problématique
- Décrire le modèle et les différents termes du modèle utilisé
- Décrire 2 parmi les résultats obtenus (tableau ou graphique)
- Donner votre opinion, force, faiblesse, des changements vous aurez faits si vous devez refaire cette étude
- A envoyer: jemay.salomon@inrae.fr (45 min)
- Max: 3 pages/en pdf

Intra-plot diversification: major agroecological lever

- ▶ caveat: not used in breeding

Intra-plot diversification: major agroecological lever

- ▶ caveat: not used in breeding

Challenge: combinatorial explosion

- ▶ strategy (here): quantitative genetics model to predict unobserved mixtures
- ▶ *joint* analysis of monovarietal and mixed stands to *gradually* introduce mixing ability in breeding programs

Intra-plot diversification: major agroecological lever

- ▶ caveat: not used in breeding

Challenge: combinatorial explosion

- ▶ strategy (here): quantitative genetics model to predict unobserved mixtures
- ▶ *joint* analysis of monovarietal and mixed stands to *gradually* introduce mixing ability in breeding programs

Gaps of knowledge:

- ▶ Which experimental designs?
- ▶ Which magnitude of genetic (co)variances?
- ▶ Which genetic architectures?

Outline

Contexte et motivation

Concept théorique

Caractères quantitatifs versus caractères qualitatifs

Concept statistique

Modèle de base

Application à l'amélioration des plantes

Activité - 45 min

Case of varietal mixtures

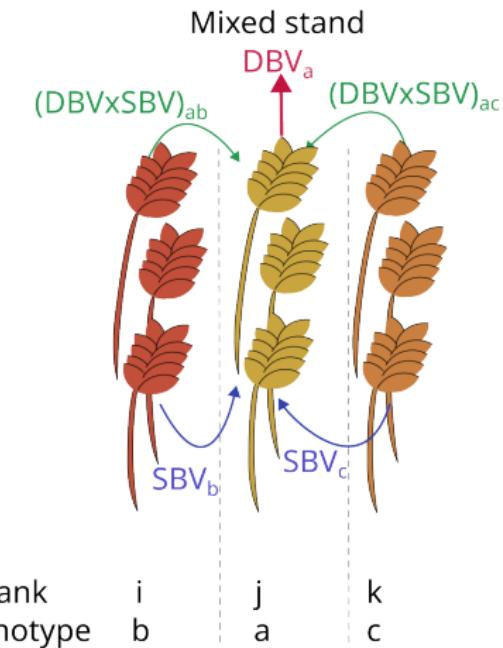
Case of crop mixtures

Material and methods

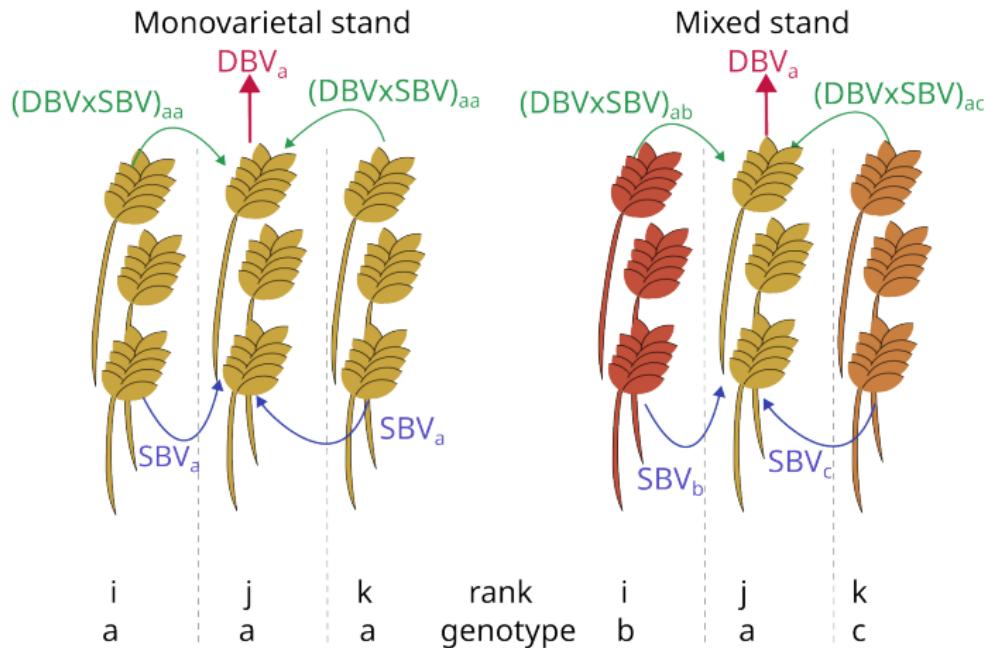
Results

Conclusions and perspectives

Design and model



Design and model



$(DBVxSBV)_{aa}$ corresponds to SMA_{aa} from Forst et al (2019)

Design and model



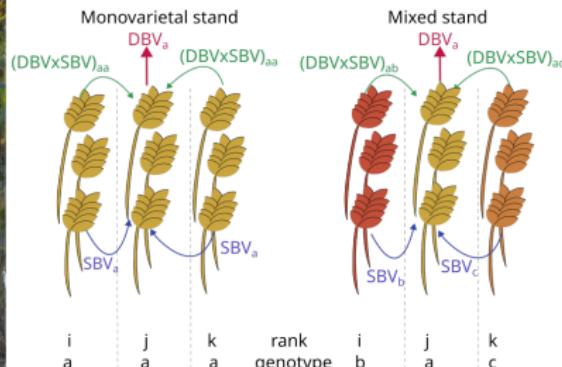
$$y_j = \mu + \alpha_{fix}$$

$$+ DBV_{g_j} + \frac{1}{2}(SBV_{g_i} + SBV_{g_k})$$

$$+ \frac{1}{2}(DBV \times (1/2)SBV)_{g_j g_i}$$

$$+ \frac{1}{2}(DBV \times (1/2)SBV)_{g_j g_k}$$

$$+ \epsilon_j$$



Design and model



- ▶ Genotyping : TaBW420k chip (200k polymorphic SNPs) of 100 lines from a MAGIC population
- ▶ Phenotyping : Yield component for each row (central for monovarietal stands)
- ▶ Use of alternate rank in early selection trials: **can we accurately estimate DBV, SBV and interaction variances?**

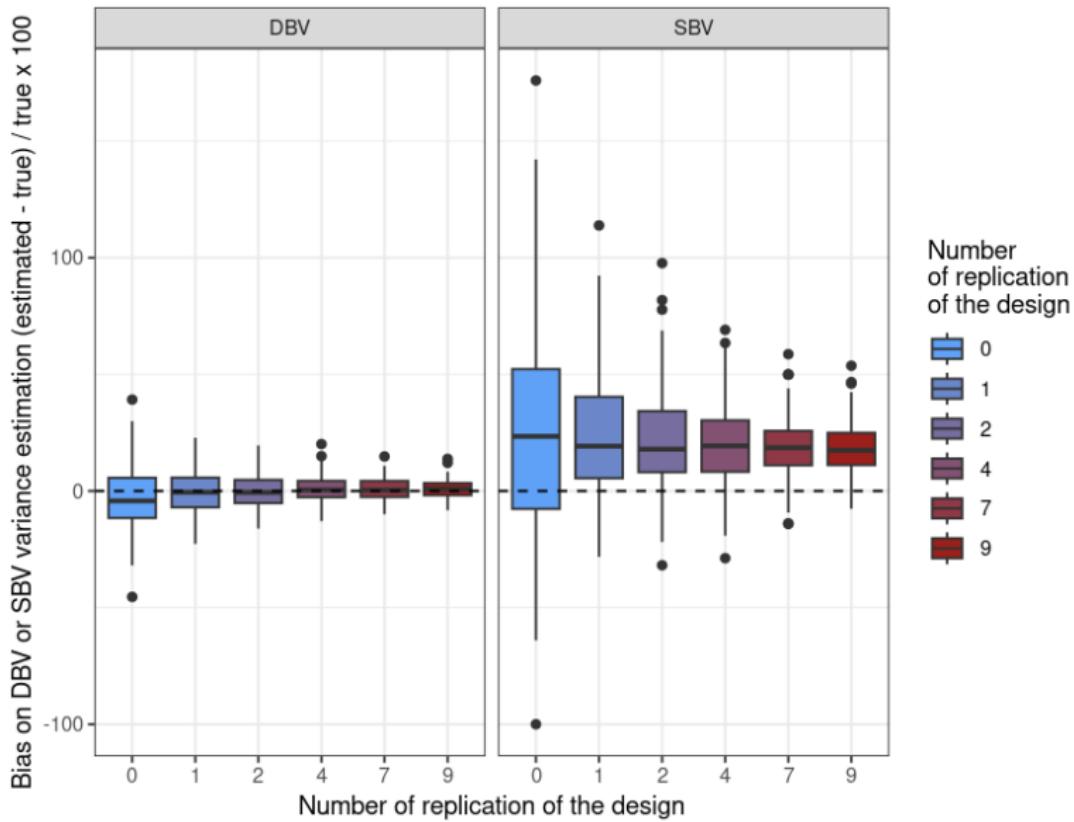
Simulations

- ▶ Simulations based on experimental design
 - ▶ 3-way component mixtures (alternate rows)
 - ▶ per line: 2 monovarietal stands + 4 mixed stands
 - ▶ Total of around 1000 rows (no replicates)
 - ▶ Parameter values: Montazeaud et al. 2023 + our data
- ▶ Test of design accuracy: addition of replicates (0 to 9)

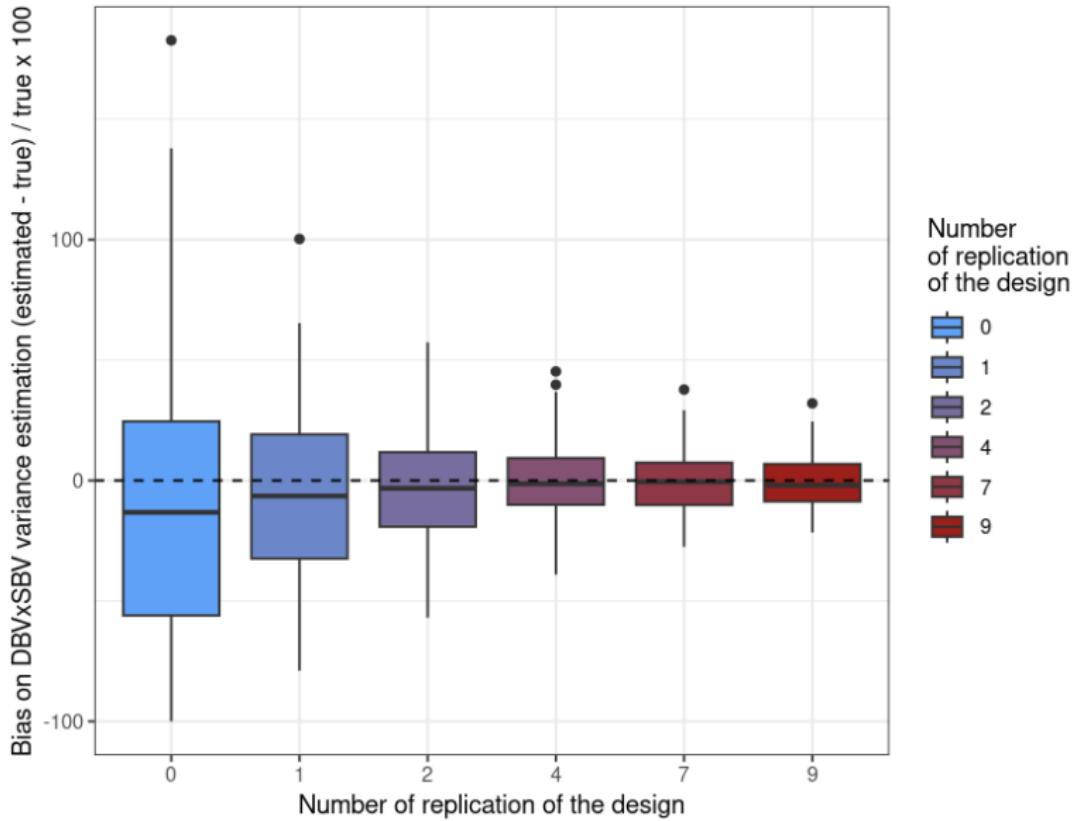
$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_{DS} \mathbf{BV}_{DS} + \mathbf{Z}_{DxS} \mathbf{DBVxSBV} + \mathbf{Z}_e \boldsymbol{\epsilon}$$

- ▶ $\boldsymbol{\beta}$: fixed effects
- ▶ \mathbf{BV}_{DS} : correlated DBV and SBV effects
- ▶ $\mathbf{DBVxSBV}$: uncorrelated interactions between DBV and SBV

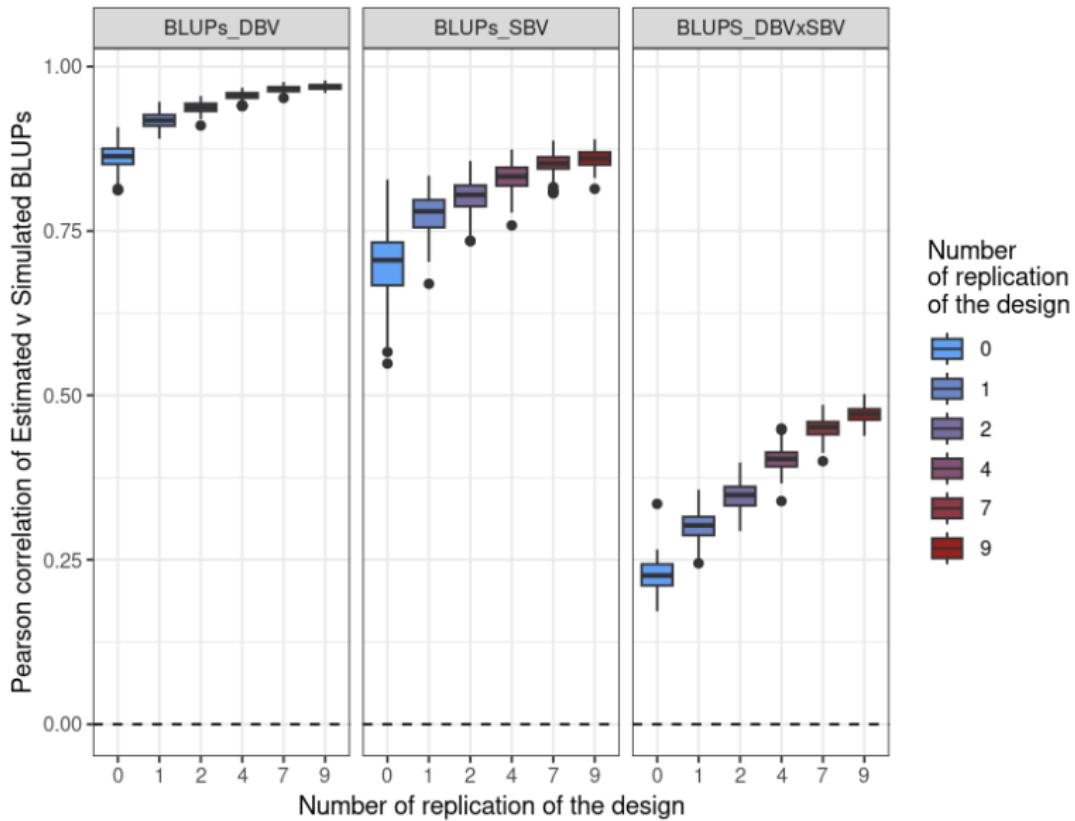
Results



Results



Results



Outline

Contexte et motivation

Concept théorique

Caractères quantitatifs versus caractères qualitatifs

Concept statistique

Modèle de base

Application à l'amélioration des plantes

Activité - 45 min

Case of varietal mixtures

Case of crop mixtures

Material and methods

Results

Conclusions and perspectives

Outline

Contexte et motivation

Concept théorique

Caractères quantitatifs versus caractères qualitatifs

Concept statistique

Modèle de base

Application à l'amélioration des plantes

Activité - 45 min

Case of varietal mixtures

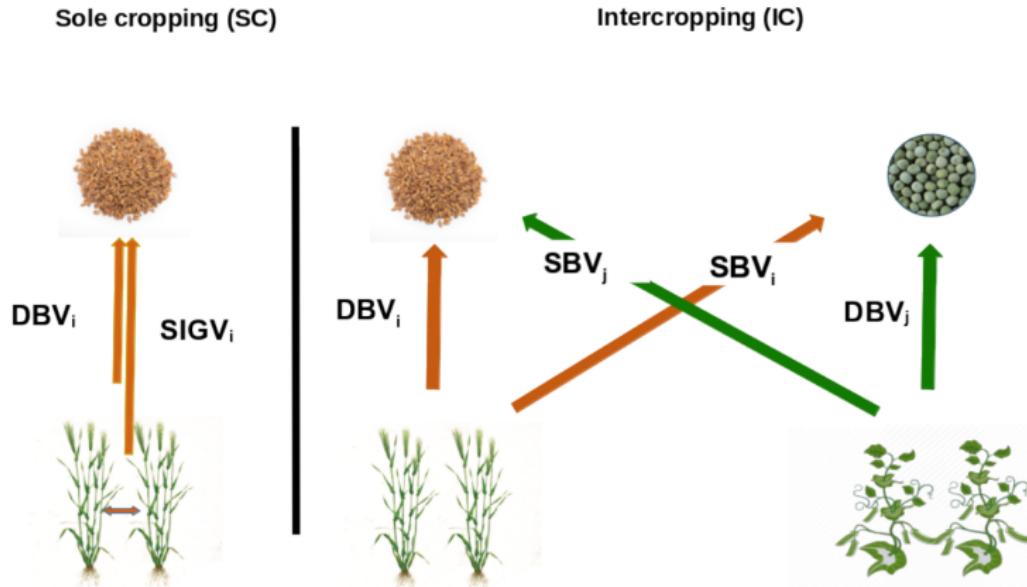
Case of crop mixtures

Material and methods

Results

Conclusions and perspectives

Genetic model



Social intra-genotypic value: $SIGV := SBV^{SC} + (DBV \times SBV)^{SC}$

Genetic model

Intercrop of i (species s_1) and j (species s_2):

$$Y_{i(j)} = \mu^{s_1, IC} + DBV_i + SBV_j^{IC} + (DBV \times SBV)_{ij}^{IC}$$
$$Y_{j(i)} = \mu^{s_2, IC} + DBV_j + SBV_i^{IC} + (DBV \times SBV)_{ji}^{IC}$$

Sole crop of i and sole crop of j :

- ▶ $Y_i = \mu^{s_1, SC} + DBV_i + SIGV_i$
- ▶ $Y_j = \mu^{s_2, SC} + DBV_j + SIGV_j$

Simulation and inference

- ▶ 200 wheat genotypes + 2 pea genotypes (testers)

Simulation and inference

- ▶ 200 wheat genotypes + 2 pea genotypes (testers)
- ▶ 3 experimental designs of 400 plots each:
 - Sole_only (complete): sole cropping system (SC)
 - Inter_only (sparse): intercropping system (IC)
 - Sole_inter_50 (sparse): combination of SC and IC

Simulation and inference

- ▶ 200 wheat genotypes + 2 pea genotypes (testers)
- ▶ 3 experimental designs of 400 plots each:
 - Sole_only (complete): sole cropping system (SC)
 - Inter_only (sparse): intercropping system (IC)
 - Sole_inter_50 (sparse): combination of SC and IC
- ▶ Parameter values: Moutier et al. (2022); Haug et al. (2023)

Simulation and inference

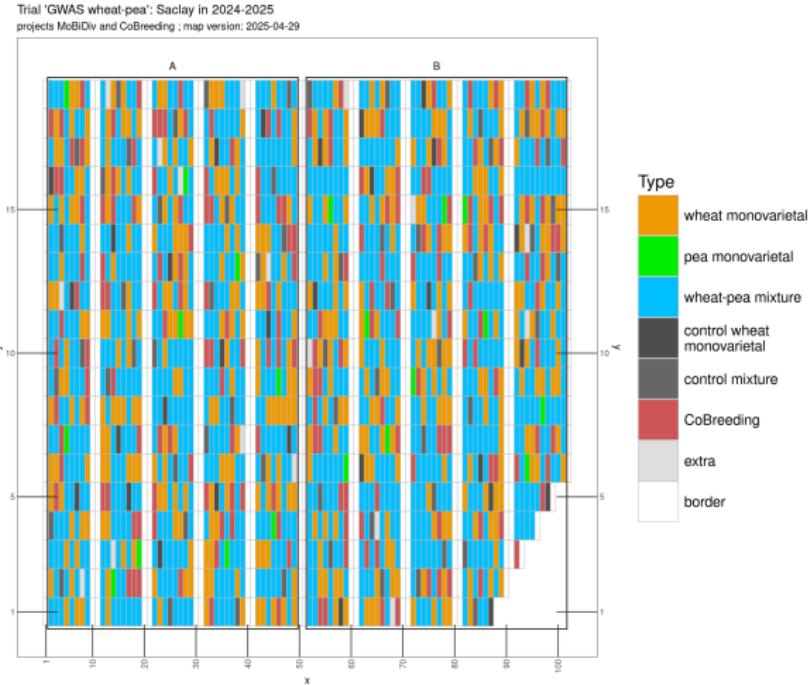
- ▶ 200 wheat genotypes + 2 pea genotypes (testers)
- ▶ 3 experimental designs of 400 plots each:
 - Sole_only (complete): sole cropping system (SC)
 - Inter_only (sparse): intercropping system (IC)
 - Sole_inter_50 (sparse): combination of SC and IC
- ▶ Parameter values: Moutier et al. (2022); Haug et al. (2023)
- ▶ Varying proportions of $\text{var}(\text{SIGV})$ relative to $\text{var}(\text{DBV})$

Simulation and inference

- ▶ 200 wheat genotypes + 2 pea genotypes (testers)
- ▶ 3 experimental designs of 400 plots each:
 - Sole_only (complete): sole cropping system (SC)
 - Inter_only (sparse): intercropping system (IC)
 - Sole_inter_50 (sparse): combination of SC and IC
- ▶ Parameter values: Moutier et al. (2022); Haug et al. (2023)
- ▶ Varying proportions of $\text{var}(\text{SIGV})$ relative to $\text{var}(\text{DBV})$
- ▶ Software implementation: plantmix (R package)

Panel and field trial

- 200 wheat varieties
 - ▶ 395k SNPs
- 2 pea varieties (testers)
- All sole and intercrops in 2 complete blocks
- Traits per species: grain yield, TKW, final height, SLA, soil cover



Panel and field trial



Le Moulon, Saclay, 2025

Outline

Contexte et motivation

Concept théorique

Caractères quantitatifs versus caractères qualitatifs

Concept statistique

Modèle de base

Application à l'amélioration des plantes

Activité - 45 min

Case of varietal mixtures

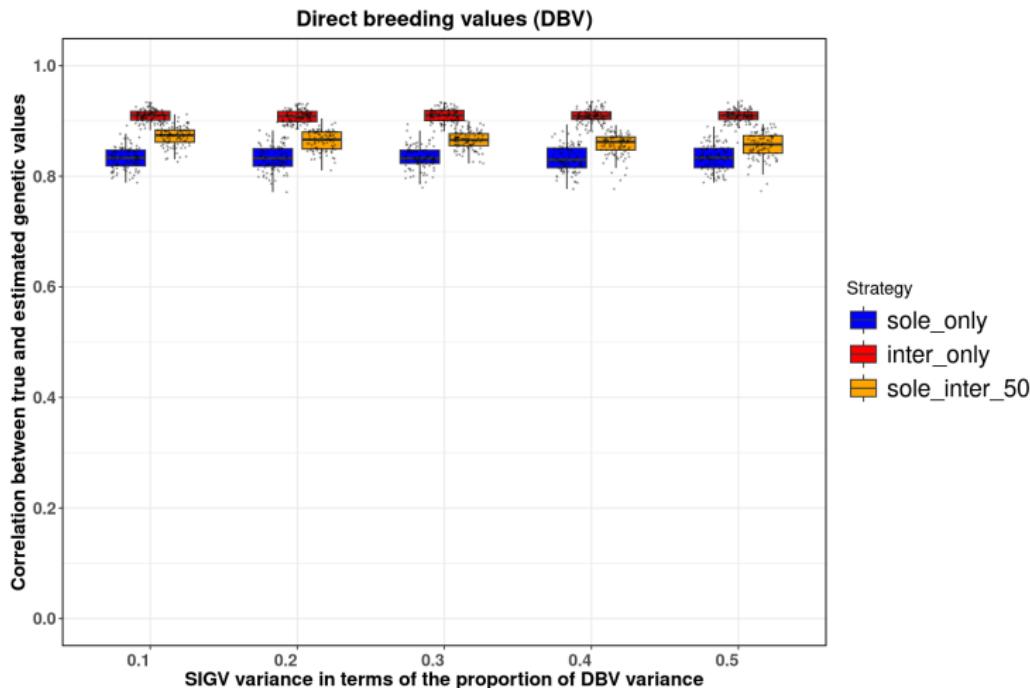
Case of crop mixtures

Material and methods

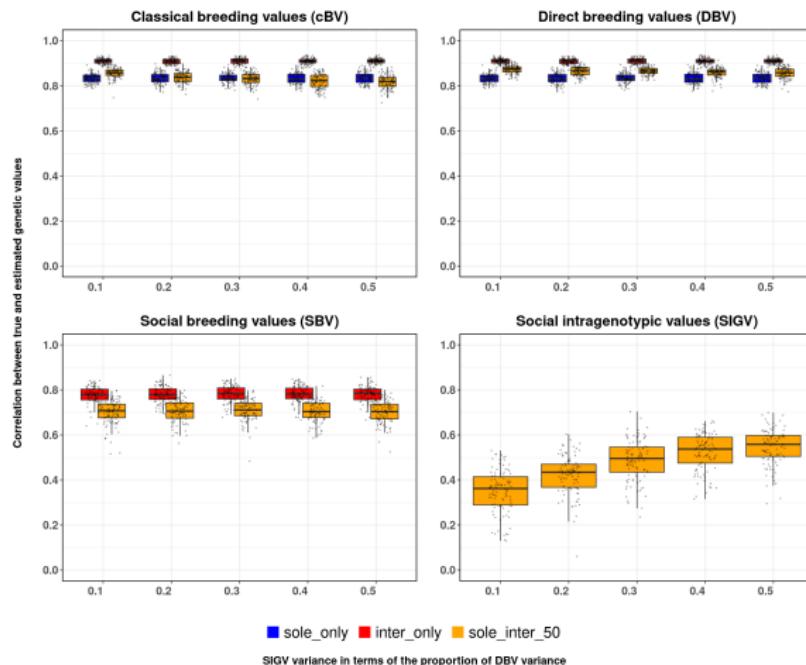
Results

Conclusions and perspectives

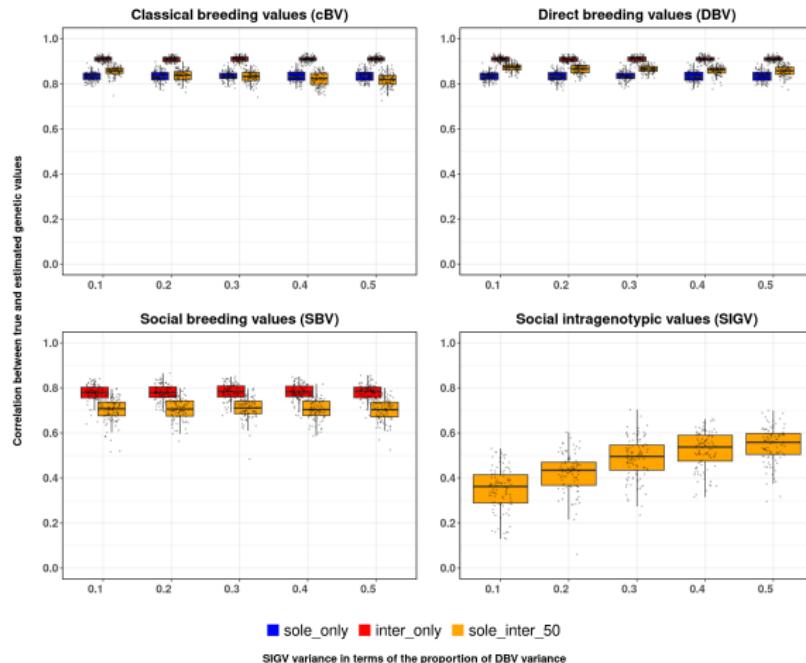
Accuracy of parameter estimation



Accuracy of parameter estimation

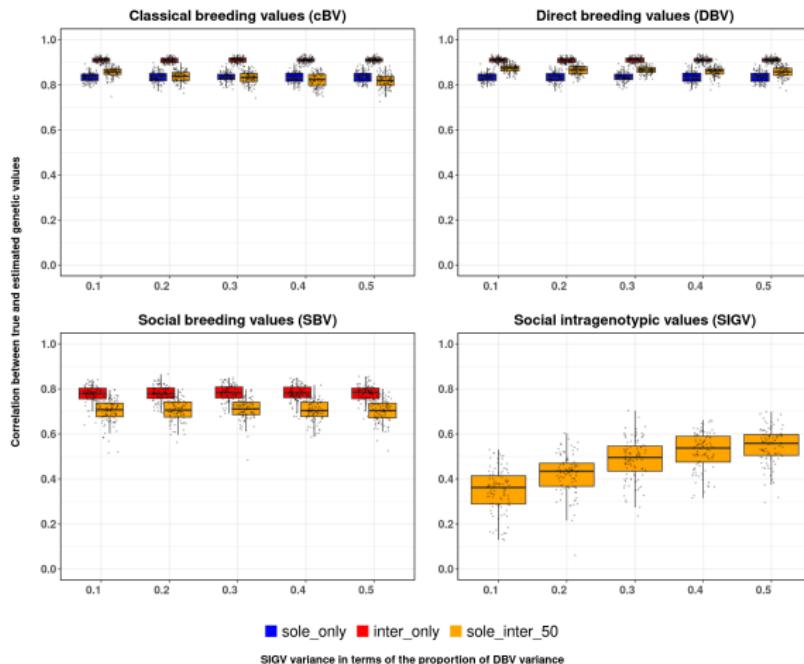


Accuracy of parameter estimation



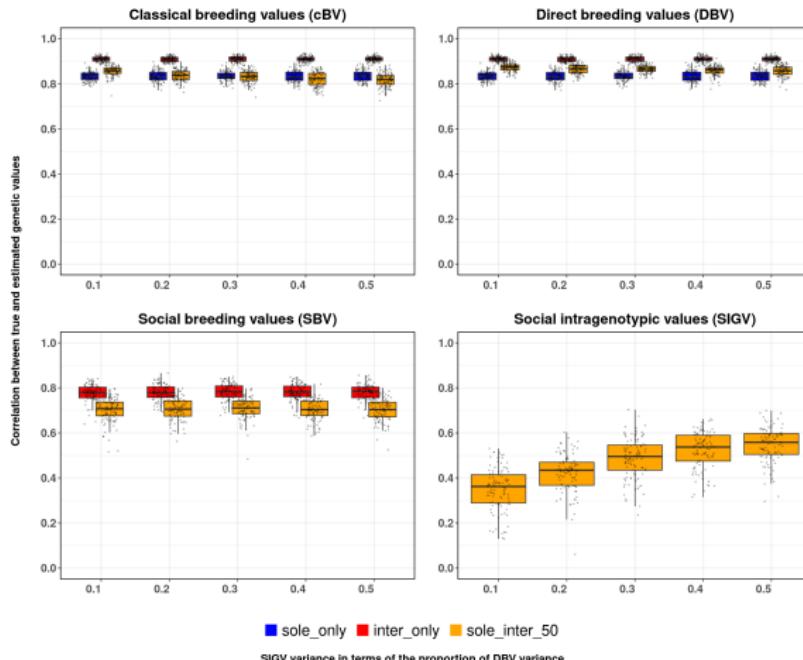
- **Experimental design, genetic effects, and the ratio $\sigma_{SIGV}^2/\sigma_{DBV}^2$ influence accuracy.**

Accuracy of parameter estimation



- **Experimental design, genetic effects, and the ratio $\sigma_{SIGV}^2/\sigma_{DBV}^2$ influence accuracy.**
- **Inter_only provides the most accurate estimates.**

Accuracy of parameter estimation



- **Experimental design, genetic effects, and the ratio $\sigma_{SIGV}^2/\sigma_{DBV}^2$ influence accuracy.**
- **Inter_only provides the most accurate estimates.**
- **Sole_inter_50 helps decoupling SIGV and DBV, improving their estimation.**

Variance estimates for grain yield

σ^2_{DBV}	σ^2_{SBV}	σ^2_{SIGV}	$\sigma^2_{DBV \times SBV}$	$\sigma^2_{e-pur(w)}$	$\sum^2_{e-mix(w)}$	$\sum^2_{e-mix(p)}$
10.1 [8.01, 11.39]	1.25 [0.17, 247]	2.49 [0, 5.23]	2.19 [0.14, 5.76]	16.6 [12.00, 20.00]	12.78 [10.61, 13.98]	18.53 [15.56, 21.30]

Values are shown as the estimate (first line), followed by the 95% confidence interval in brackets (second line), calculated across 500 replicates using bootstrapping.

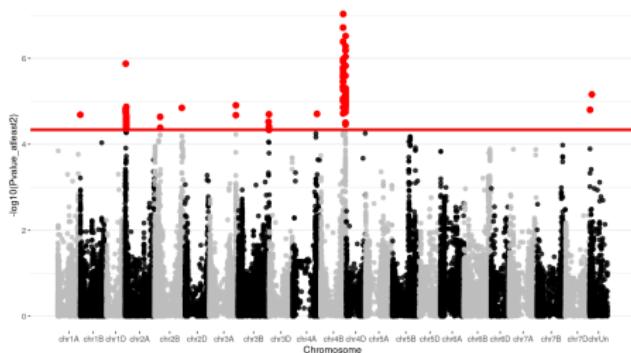
- Social genetic variances < variance of direct breeding values

Correlation estimates for grain yield

$\rho_{DBV-SBV}$	$\rho_{\epsilon\text{-mix}}$
0.24 [0.01, 1]	-0.19 [-0.29, -0.07]

- Unexpected positive correlation observed between DBV and SBV
⇒ favorable wheat-pea interactions

Meta-GWAS on grain yield [DBV, SBV, SIGV]



H0 rejected when a SNP is associated with at least two traits (FDR controlled at 5%)

► 9 pleiotropic regions detected

- Not detected in single-trait GWAS.
- Further analyses underway for validation.

Outline

Contexte et motivation

Concept théorique

Caractères quantitatifs versus caractères qualitatifs

Concept statistique

Modèle de base

Application à l'amélioration des plantes

Activité - 45 min

Case of varietal mixtures

Case of crop mixtures

Material and methods

Results

Conclusions and perspectives

Take-Home Messages

- Joint modeling of sole and intercrops + incomplete design + genomic relationship matrix: accurate estimations of breeding values.

Take-Home Messages

- Joint modeling of sole and intercrops + incomplete design + genomic relationship matrix: accurate estimations of breeding values.
- Field data analyses indicated that social genetic variances were small, but still contribute significantly

Take-Home Messages

- Joint modeling of sole and intercrops + incomplete design + genomic relationship matrix: accurate estimations of breeding values.
- Field data analyses indicated that social genetic variances were small, but still contribute significantly
- Methodology allowing breeders to fine-tune their program depending on the proportion of resources they want to allocate to sole vs intercrops.

Take-Home Messages

- Joint modeling of sole and intercrops + incomplete design + genomic relationship matrix: accurate estimations of breeding values.
- Field data analyses indicated that social genetic variances were small, but still contribute significantly
- Methodology allowing breeders to fine-tune their program depending on the proportion of resources they want to allocate to sole vs intercrops.
- Preliminary GWAS identified several regions associated with at least two types of breeding values.

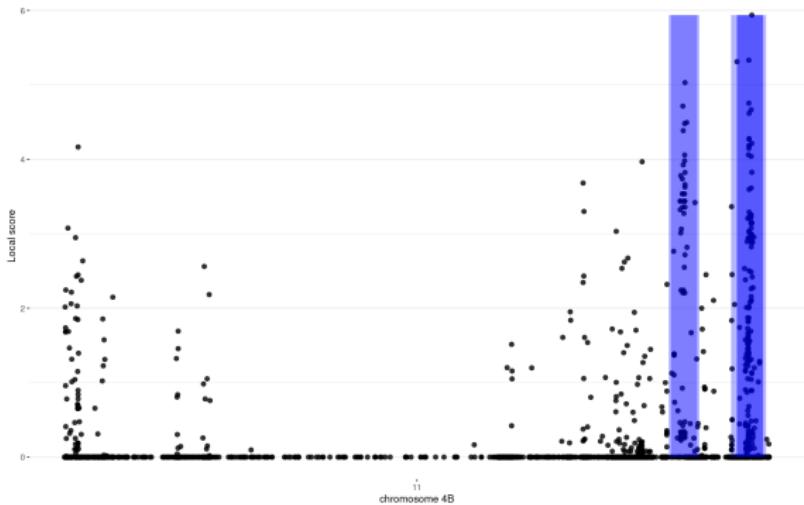
Take-Home Messages

- Joint modeling of sole and intercrops + incomplete design + genomic relationship matrix: accurate estimations of breeding values.
- Field data analyses indicated that social genetic variances were small, but still contribute significantly
- Methodology allowing breeders to fine-tune their program depending on the proportion of resources they want to allocate to sole vs intercrops.
- Preliminary GWAS identified several regions associated with at least two types of breeding values.
- Analyze the other traits, notably those from drone imaging (PhD V. Freitas).

Acknowledgments

- ▶ **GQE:** J. Enjalbert, T. Flutre,
C. Bourhis-Lézier, V. Freitas
 - ▶ M. Lu, P. Briens, S. Roty,
T. Linares, F. Legendre,
J. Hélie, J. Le Gall, L. Salze
 - ▶ E. Akaffou, L. Wang,
M. Mesnil, D. Wang,
F. Petit, R. Diallo, U. Louis,
C. Coste, F. Alléhaut
- ▶ **UEVS:** C. Bédard, F. Barriuso,
D. Sowamber, J. Cannesson,
O. Minguy, A. Belkian
- ▶ **ECOSYS:** J.-M. Gilliot
- ▶ **BIOGER:** T. Vidal,
M. Delrieu
- ▶ **IGEPP:** N. Moutier,
R. Perronne
- ▶ **AGAP:** A. Baranger

Supp 1: meta-GWAS on grain yield [DBV, SBV, SIGV]



Local scores along chromosome 4B. The boxes represent the significant zones identified. The threshold was set to $\xi = 2$, and the nominal FDR level is fixed at 0.05.