

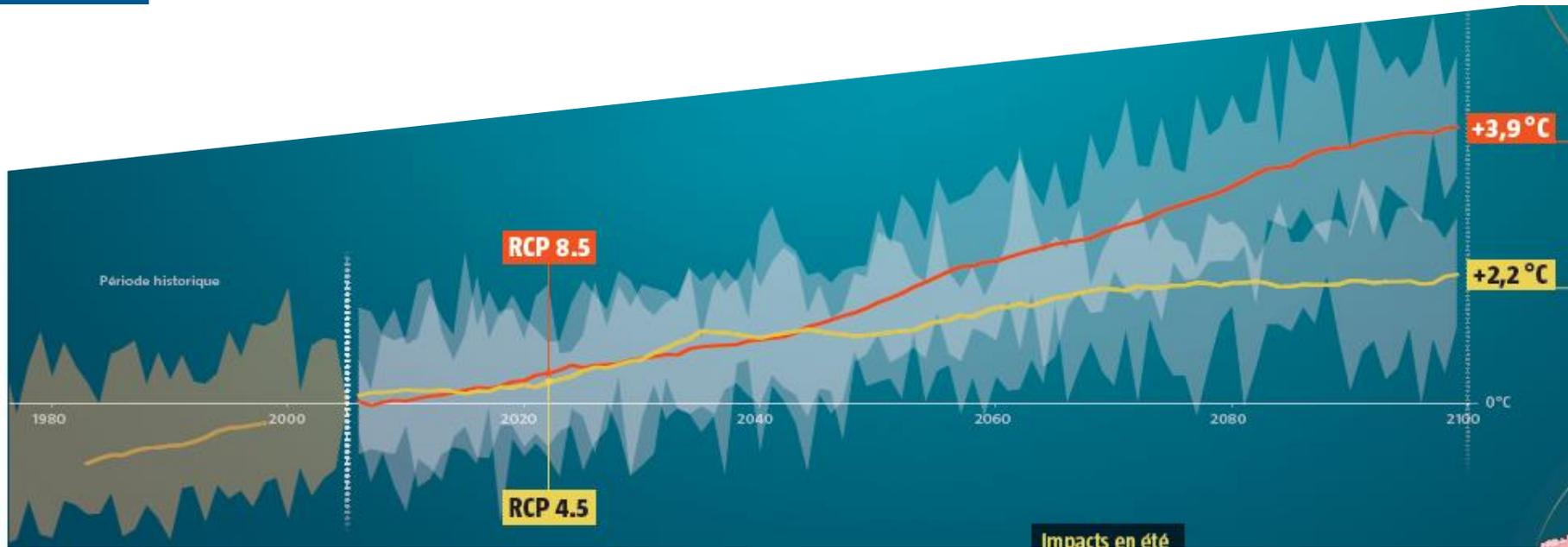
La modélisation de la relation génotype – phénotype à l'ère du changement global

Renaud Rincant

INRAE

UMR Génétique Quantitative et Evolution – Le Moulon

Changement climatique – Projection Météo France RCP 4.5, RCP 8.5



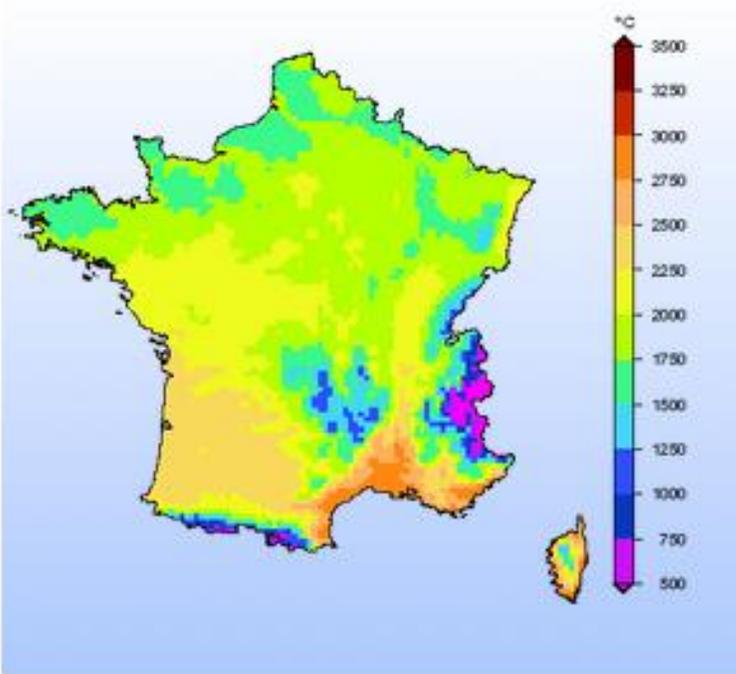
- Réchauffement de 2,2 degrés en 2100 par rapport à 2005 pour le scénario « optimiste ».
- Le réchauffement n'est pas uniforme sur le territoire.
- Plus de canicule, des sécheresses plus longues, beaucoup moins de gel.
- **Conséquences importantes pour l'agriculture.**

Somme des températures en base 6 d'avril à octobre

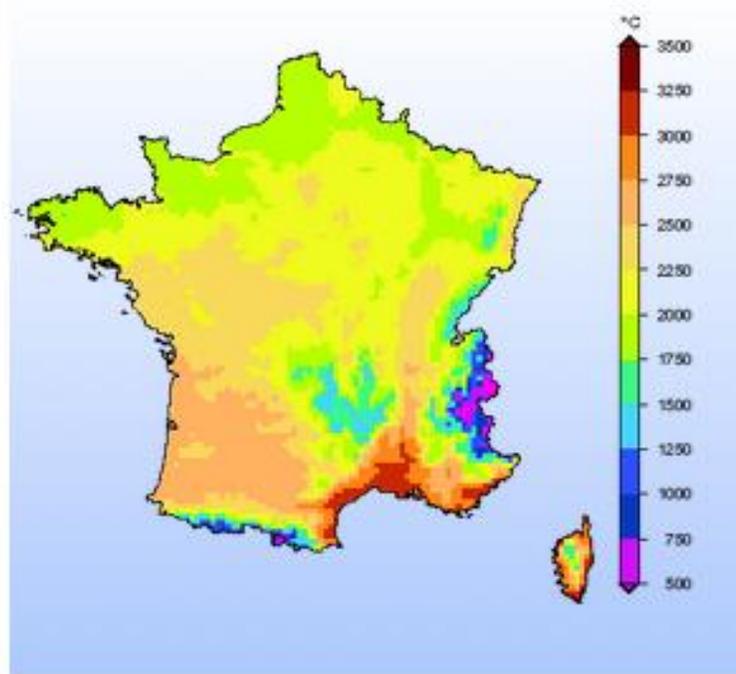


Multi-modèles de DRIAS-2020 : médiane de l'ensemble. Scenarior RCP4.5

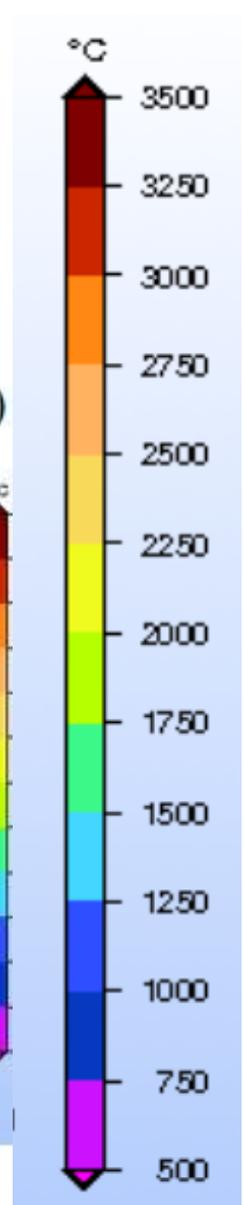
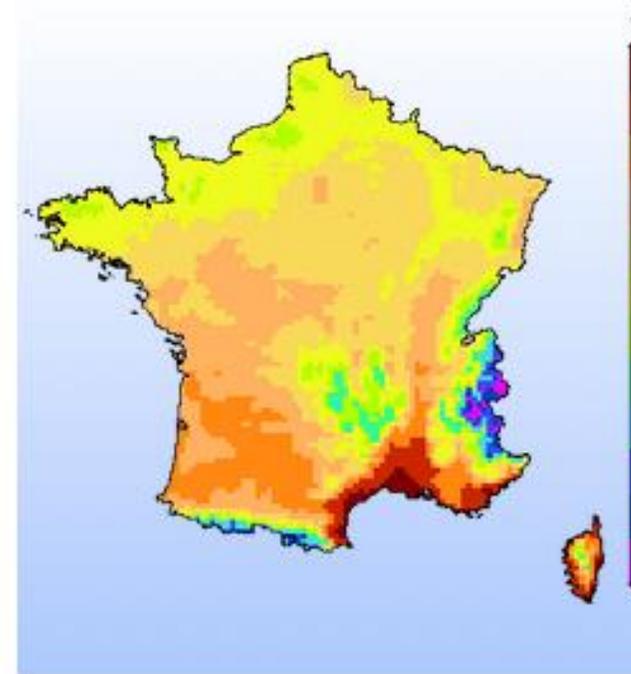
Référence (1976-2005)



Horizon proche (2021-2050)

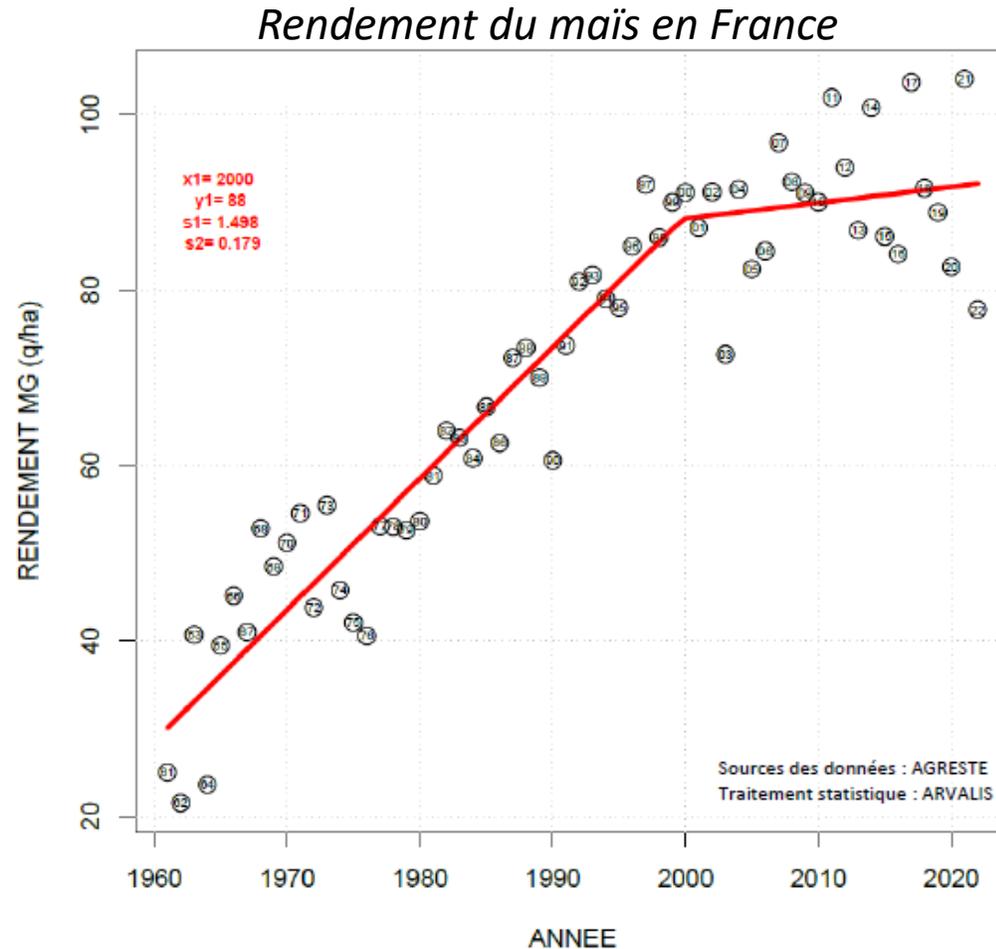


Horizon lointain (2071-2100)



- Les contrastes entre région augmentent.
- En Ile de France, on passe de 2000°C-jour à 2500 °C-jour
- Fort effet sur l'agriculture (phénologie, stress)
- Cache les températures extrêmes.

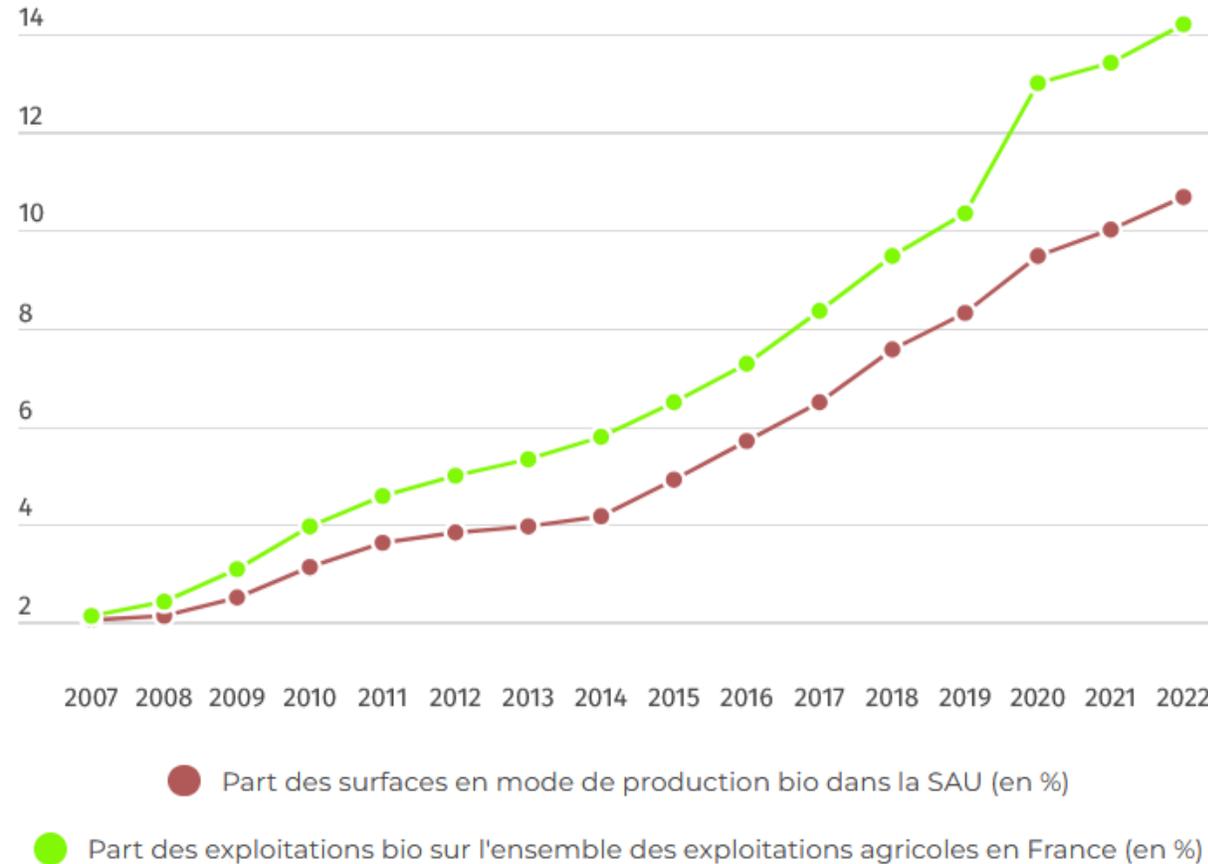
Effet du changement climatique sur la productivité



- Stagnation des rendements (effet tendanciel du changement climatique).
- Augmentation de la variabilité inter-annuelle (dérèglement climatique).
- Effet du changement des pratiques ?
- **Quelle sélection dans ce nouveau contexte ?**

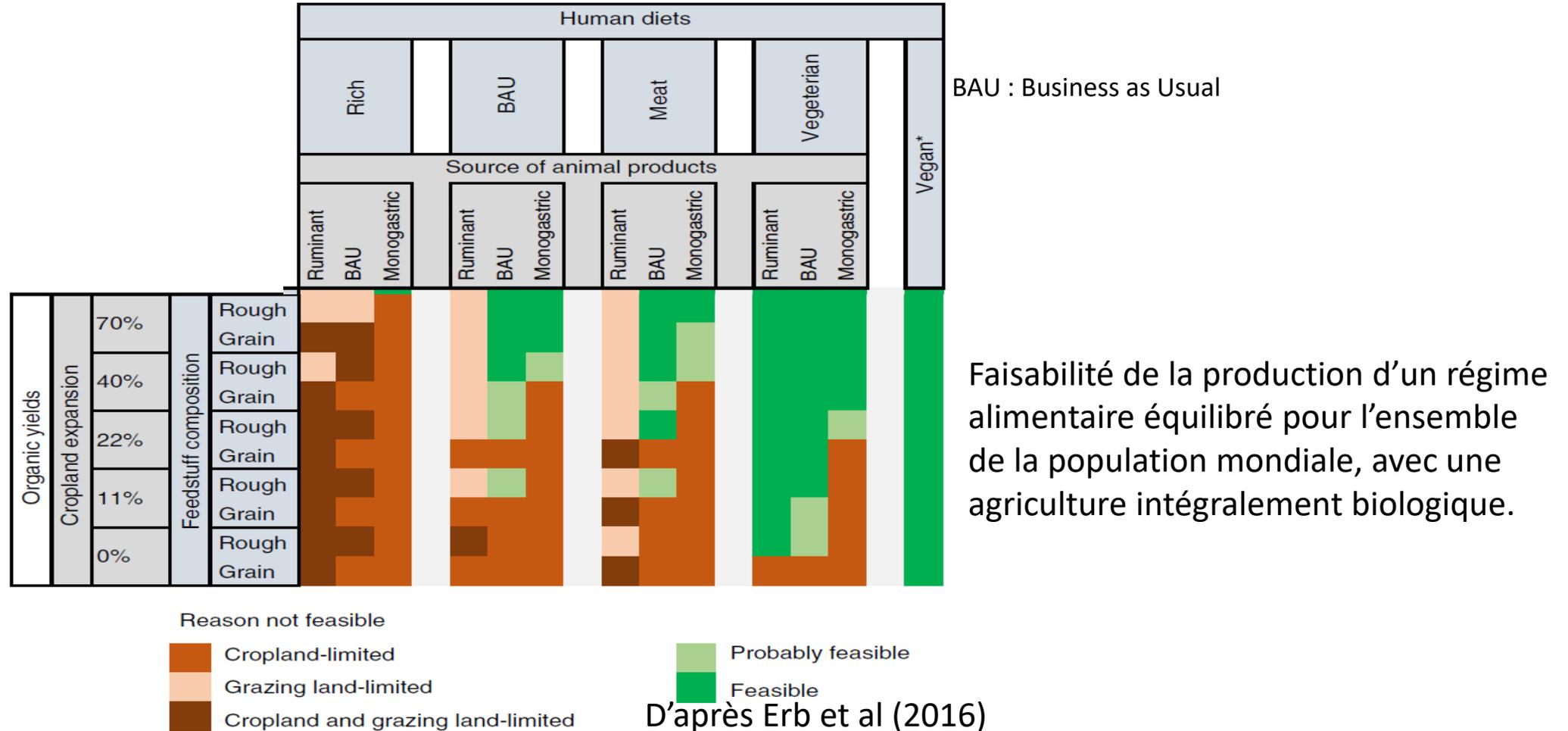
Vers des agricultures moins intensives

Evolution de l'agriculture biologique en France



- Des agricultures moins intensives, plus respectueuses de l'environnement se développent.
- Ceci entraîne une baisse de la productivité.
- **Importance de sélectionner pour la résilience.**

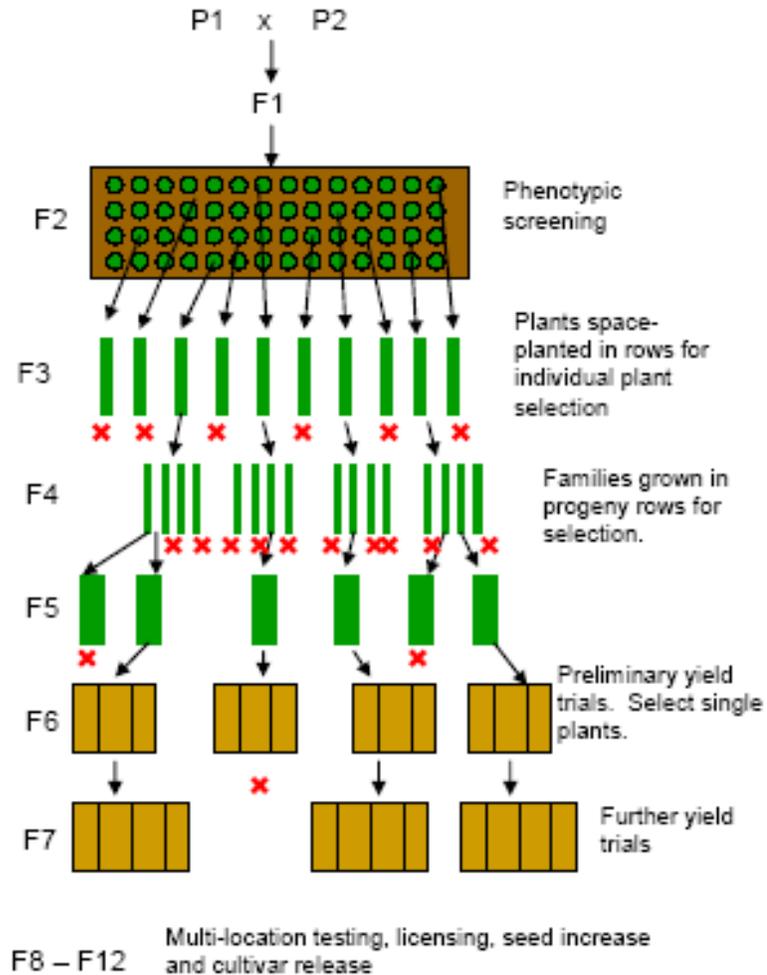
Effet du régime alimentaire sur l'utilisation des surfaces agricoles



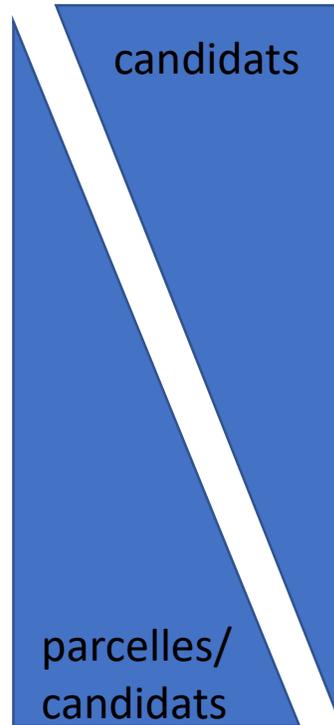
- Un régime alimentaire végétarien permettrait de passer l'ensemble de l'agriculture en biologique.
- Marges de manœuvre importantes dans nos systèmes agricoles pour une transition vers l'agroécologie.
- **Nécessité d'adapter la sélection à ces nouvelles cibles.**

Organisation des programmes de sélection phénotypique

PEDIGREE METHOD

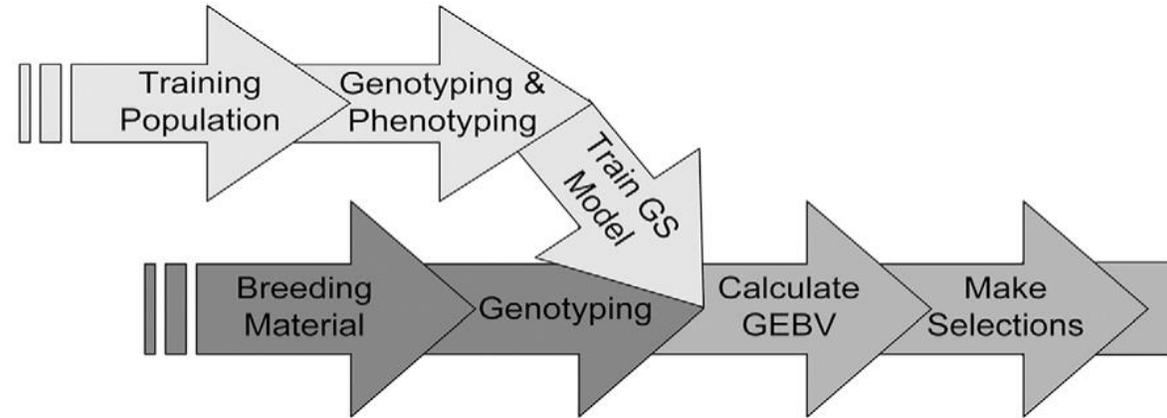


D'après Ribaut & Betran (1999)



- Variabilité générée en croisant des parents complémentaires.
- Sélection phénotypique longue et couteuse.
- L'essentiel des candidats sont éliminés sans connaissance de leur productivité.

La sélection génomique



Heffner et al. 2009

Modèle G-BLUP:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

Phenotype du jeu d'entraînement

Effet aléatoire polygénique

Kinship (apparemment génomique)

$$\text{var}(\mathbf{u}) = \mathbf{K} \sigma_g^2$$

- Prédiction de candidats non phénotypés (mais génotypés).
- Diminution de la durée des cycles, augmentation de l'intensité de sélection, optimisation des plans de croisement.
- Comment la sélection génomique contribue à réinventer les programmes de sélection ? Quelle utilisation optimale en amélioration des plantes ?

Effet de la précision de la prédiction sur le progrès génétique

Réponse à la sélection

Intensité de sélection

Précision de la prédiction

Ecart-type génétique additif

Durée des cycles

$$\Delta G = i \cdot R \cdot \sigma_A / T$$

- Tout facteur ayant un effet sur cette précision affecte le progrès génétique.

1. Optimisation de la composition du jeu d'entraînement
2. Prédiction des interactions génotypes x environnements
3. Explorer la relation génotype – phénotype par la génétique des systèmes

1. Optimisation de la composition du jeu d'entraînement

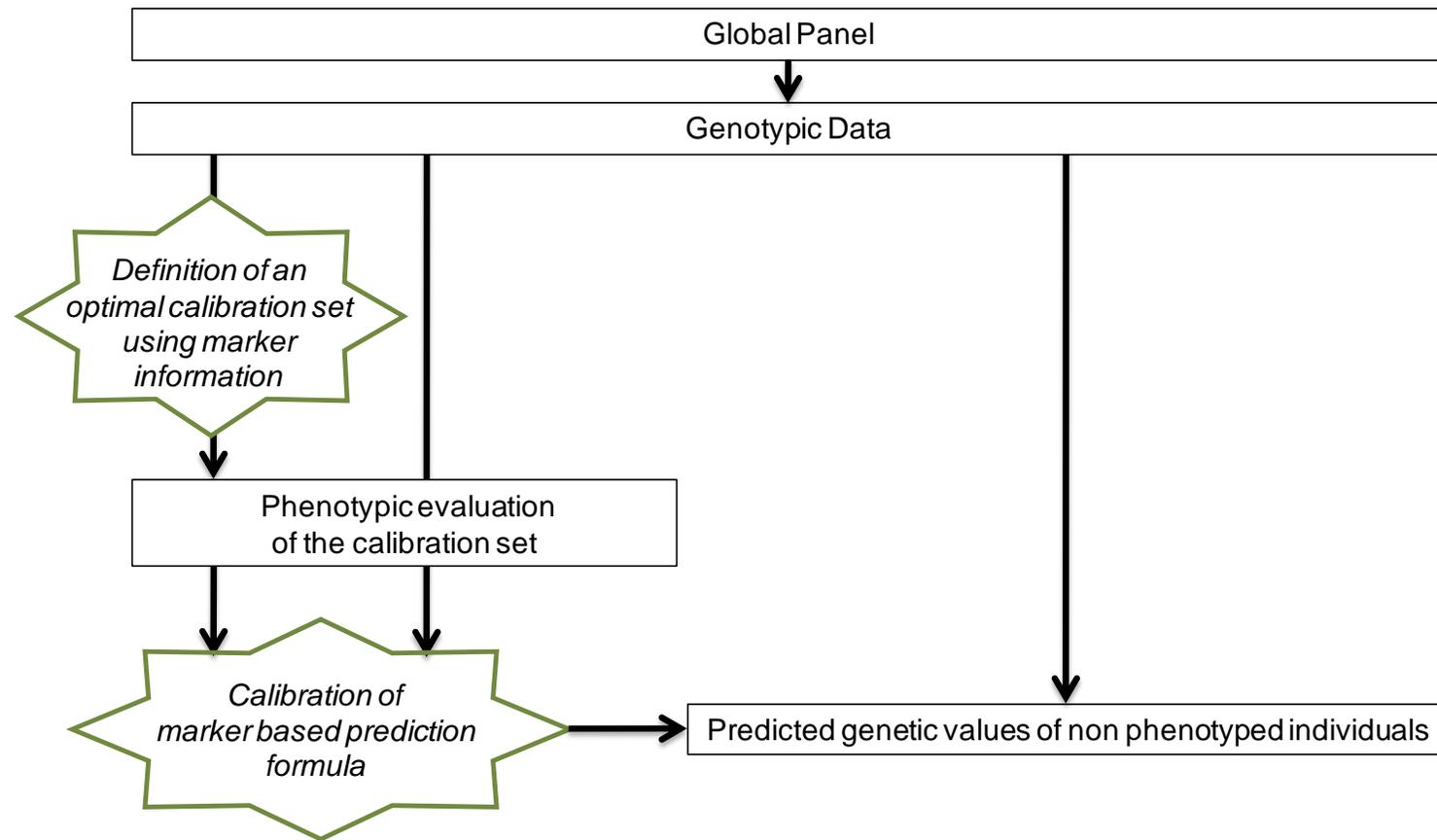
Optimisation de la composition du jeu d'entraînement

Diagram illustrating the equation $\Delta G = i \cdot R \cdot \sigma_A / T$ with labels and arrows indicating the relationship between variables:

- Réponse à la sélection (points to i)
- Intensité de sélection (points to i)
- Précision de la prédiction (points to R)
- Ecart-type génétique additif (points to σ_A)
- Durée des cycles (points to T)

- La composition du jeu d'entraînement a un effet important sur la qualité des prédictions.
- Tout gain en précision de prédiction se traduit en progrès génétique.

Optimisation de la composition du jeu d'entraînement



A. Charcosset



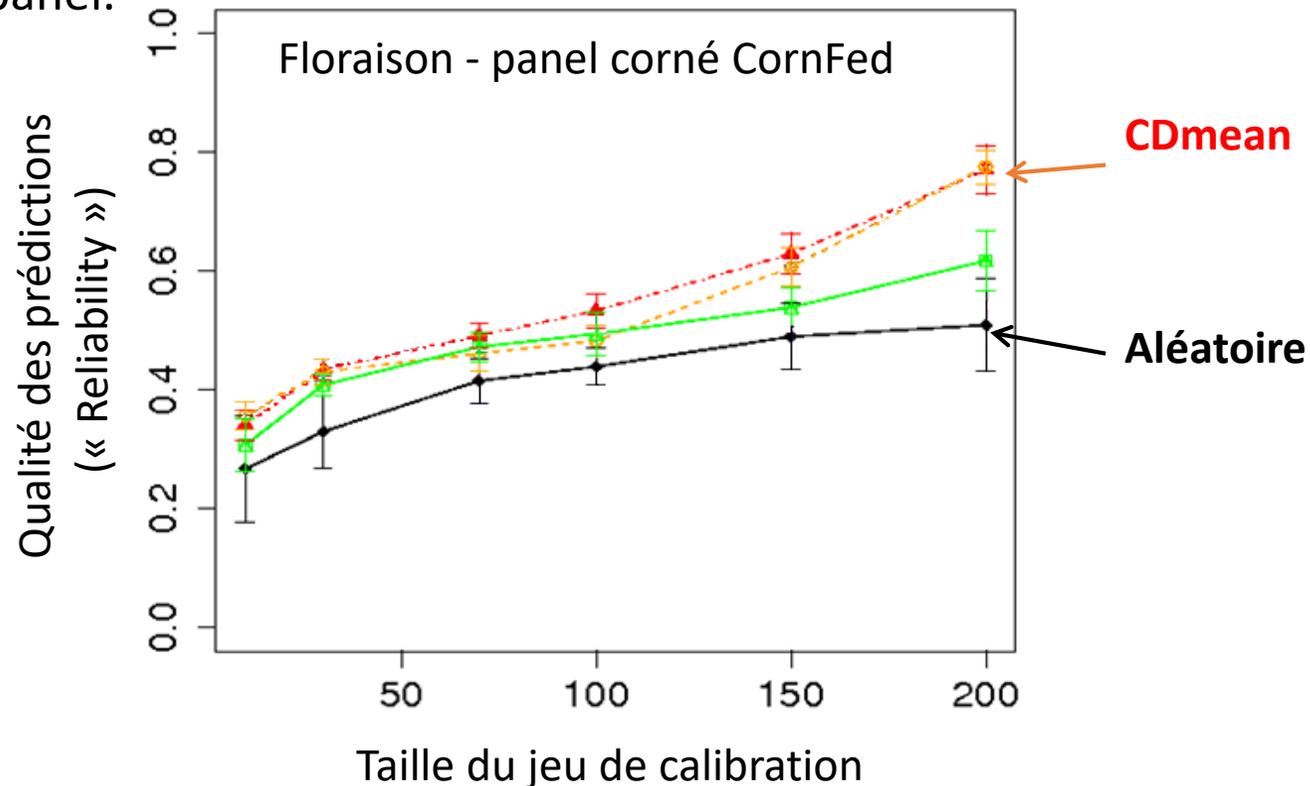
L. Moreau

- Le phénotypage est limitant, mais le génotypage est accessible pour tous les génotypes.
- Quels génotypes phénotyper pour aboutir aux prédictions les plus fiables possible ?

Optimisation de la composition du jeu d'entraînement

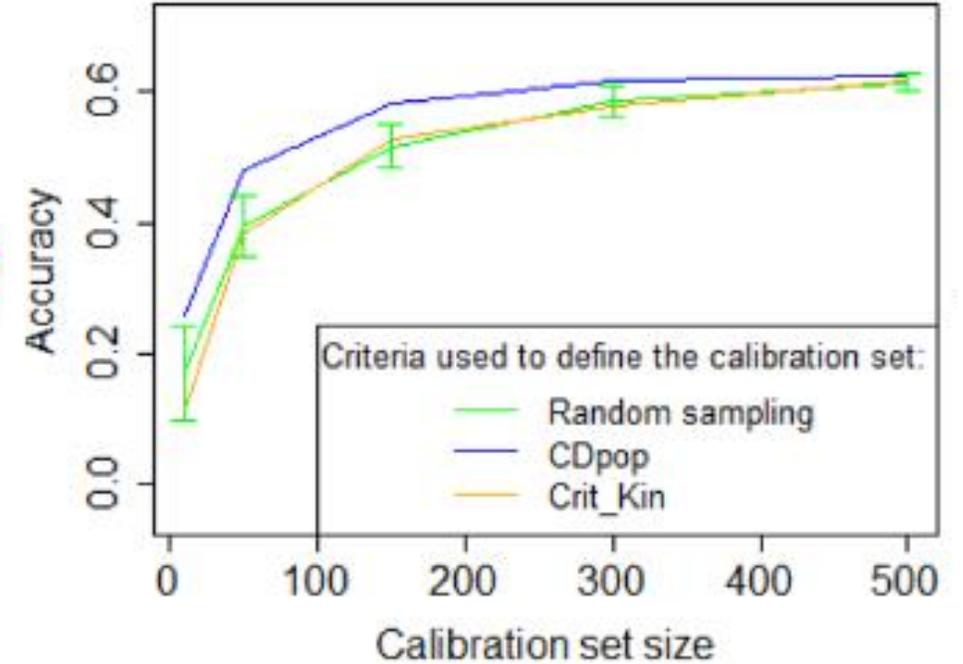
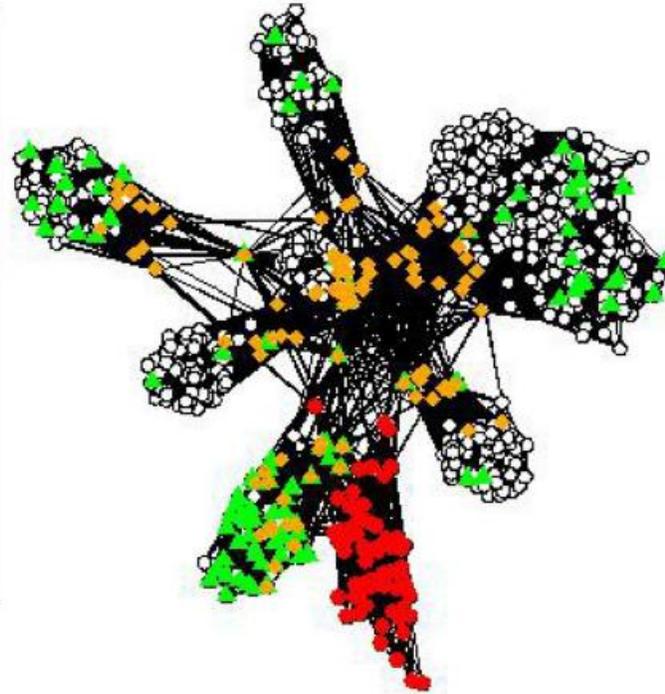
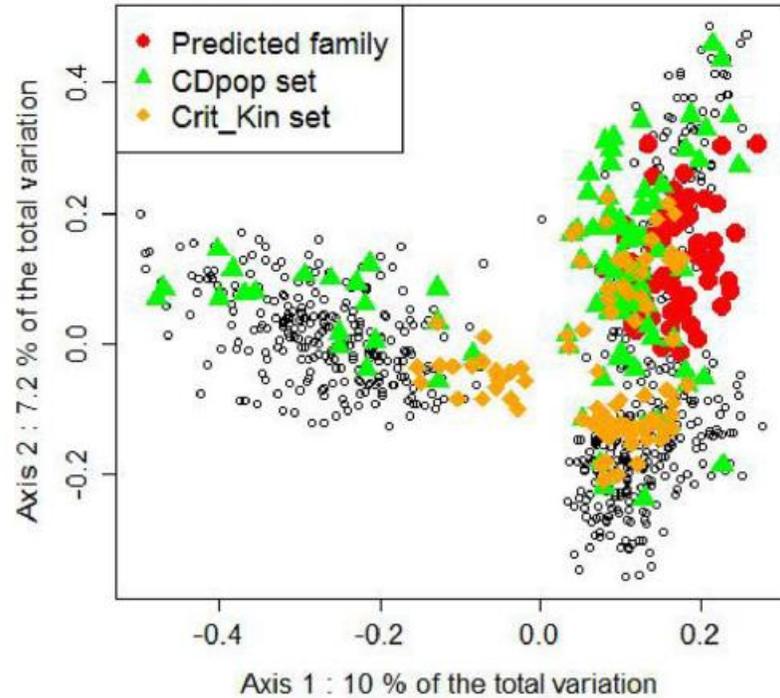
Le **CD généralisé** (Laloë *et al.* 1996): $CD(\mathbf{c}) = \text{diag} \left[\frac{\mathbf{c}'(\mathbf{K} - \lambda(\mathbf{Z}'\mathbf{M}\mathbf{Z} + \lambda\mathbf{K}^{-1})^{-1})\mathbf{c}}{\mathbf{c}'\mathbf{K}\mathbf{c}} \right]$

- On peut estimer *a priori* (avant de phénotyper) la fiabilité des prédictions.
- Définition du critère CDmean : moyenne des CD des contrastes entre chaque génotype à prédire et la moyenne du panel.



Optimisation de la composition du jeu d'entraînement

Adaptation du CD à des populations structurées



- En jouant sur les contrastes on peut adapter le critère à des situations très diverses.
- Gains importants en qualité prédictive.

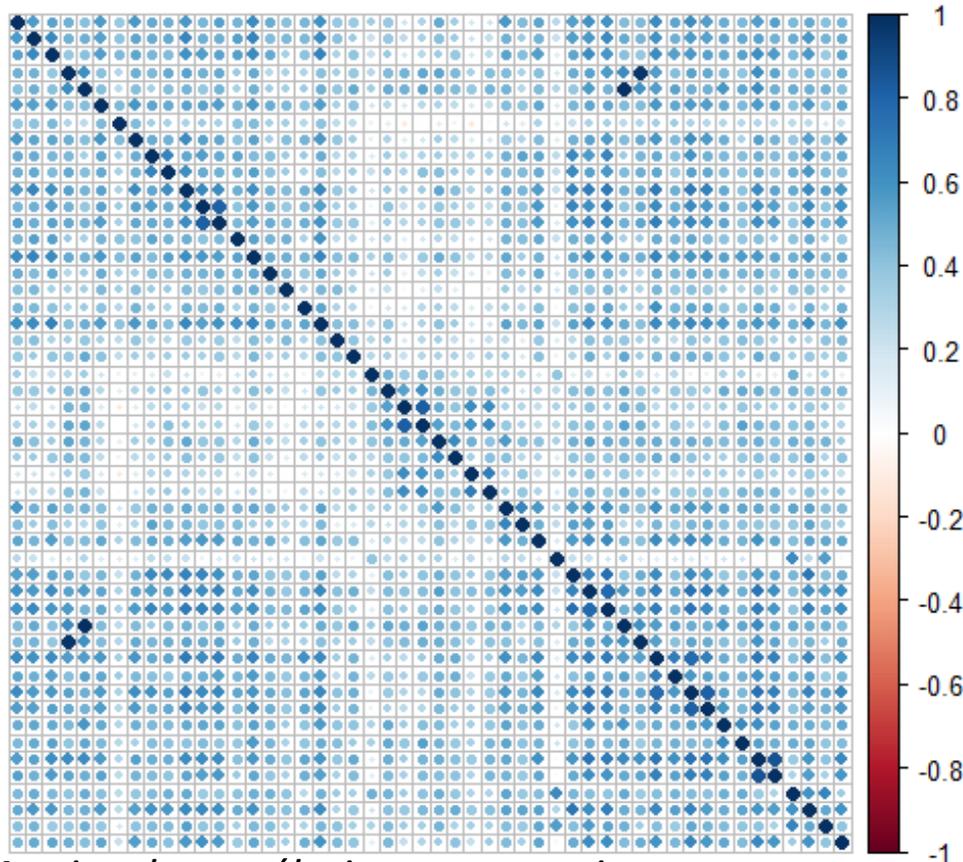
Optimisation de la composition du jeu d'entraînement

- Particulièrement utile pour l'évaluation de caractères difficilement phénotypables.
- Permet de mieux explorer les ressources génétiques (CRB), phénotypage d'un jeu d'entraînement pour des caractères clés pour prédire l'ensemble de la collection.

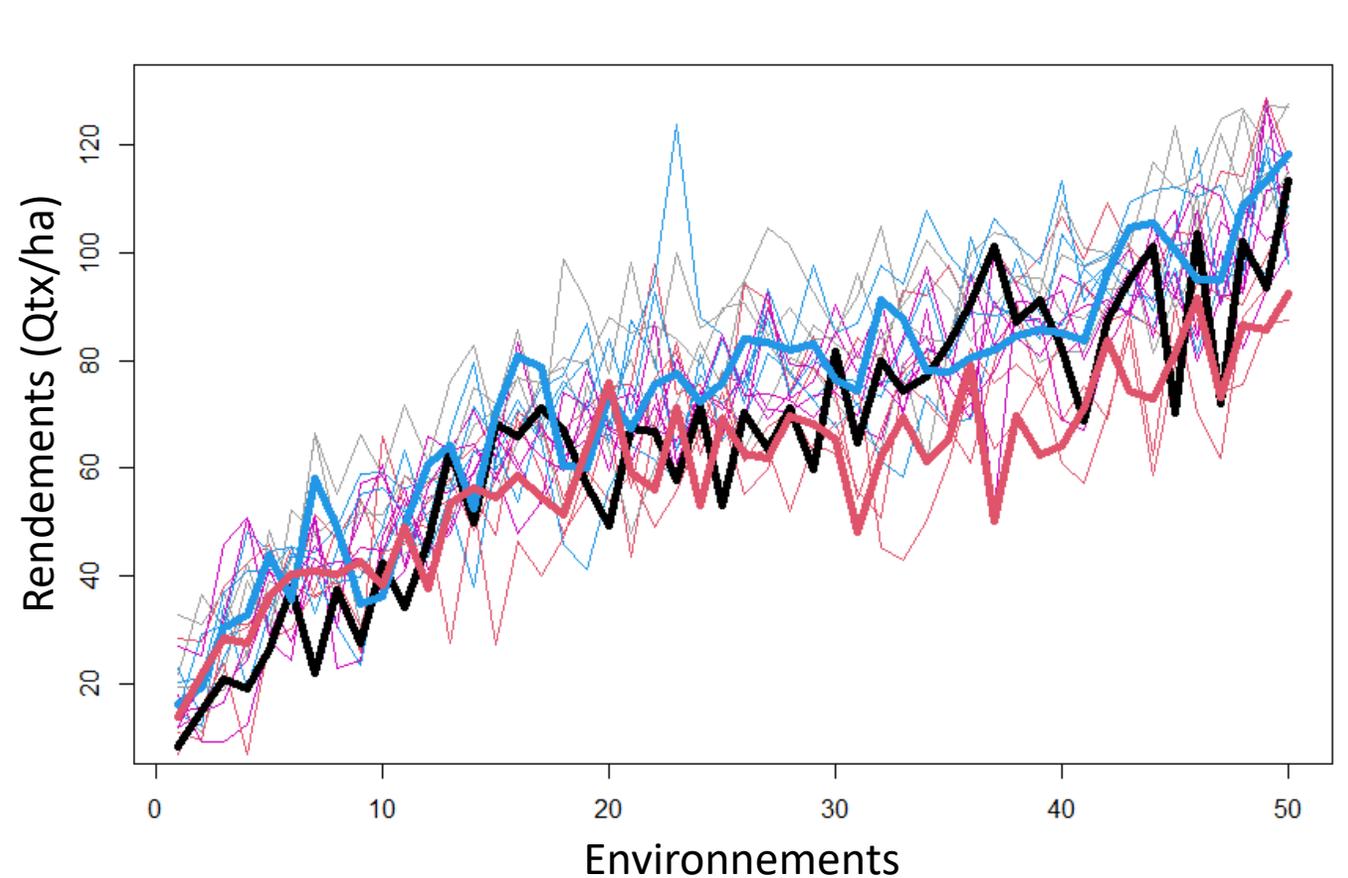
2. Prédiction des interactions génotypes x environnements

Prédiction des interactions génotypes x environnements

- Les génotypes répondent différemment à un changement d'environnement.
- L'ordre des génotypes change.



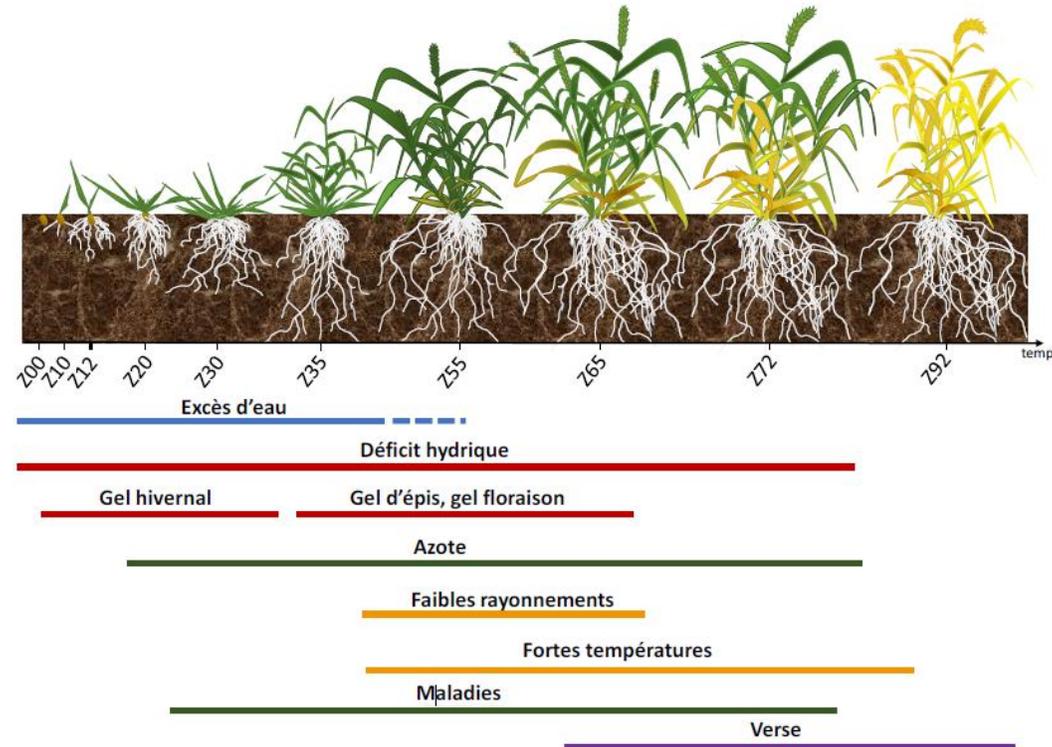
Matrice de corrélation entre environnements
(réseau d'essais DROPS-AMAIZING)



Productivité des variétés dans le réseau d'essais DROPS-AMAIZING

Prédiction des interactions génotypes x environnements

- Impossible d'évaluer tous les candidats dans tous les environnements.
- Pose la question du dispositif expérimental pour représenter au mieux les environnements cibles.
- Reste difficile du fait d'une forte variabilité interannuelle.
- Cela peut être complété par des approches prédictives qui permettent de prédire dans des environnements observés ou dans de nouveaux environnements grâce à des covariables environnementales.



Exemple de covariable: bilan hydrique à l'épiaison

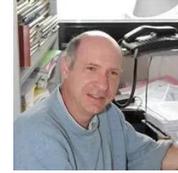
Touzy et al. 2019 (blé tendre)

Prédiction des interactions génotypes x environnements

Intégration de covariables environnementales dans le modèle de prédiction

$$Y_{ij} = \mu_j + G_i + GE_{ij} + \varepsilon_{ij} \quad \text{avec} \quad GE_{ij} \sim N(0, K \otimes W \cdot \sigma_{ge}^2)$$

J. Le Gouis



P. Martre



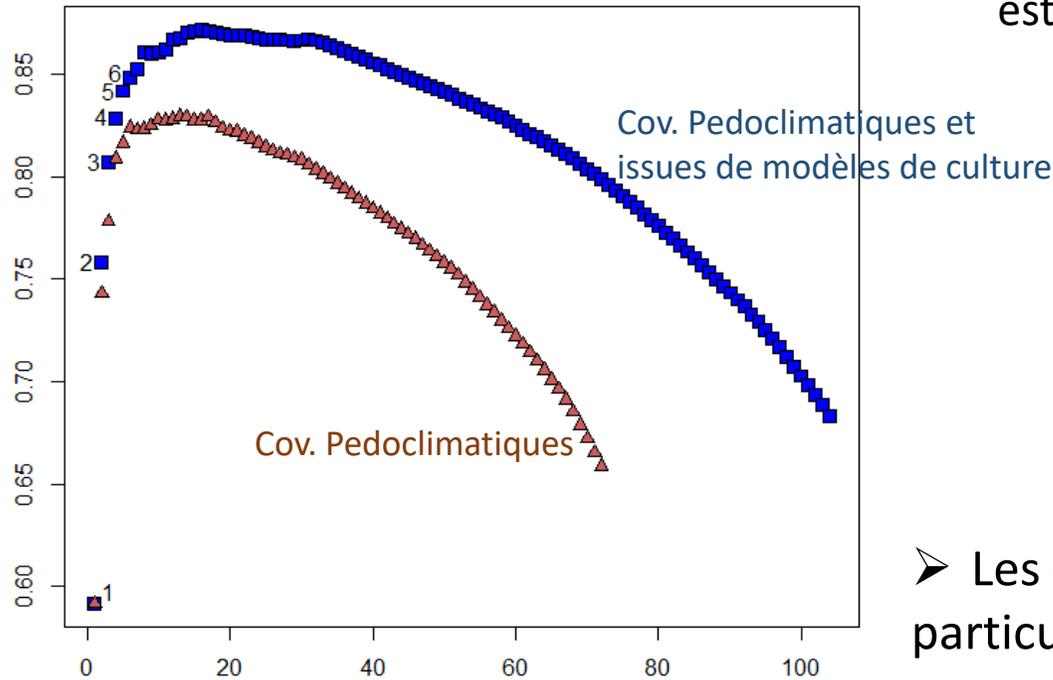
F. Van Eeuwijk



M. Malosetti

Matrice de similarité environnementale
estimée avec les covariables environnementales

Corrélation avec les interactions observées



$$DMSI(d) = \frac{B_{act}(d) - B_{act}(d - 1)}{B_{unl}(d) - B_{act}(d - 1)}$$

➤ Les covariables issues de modèles de culture semblent particulièrement pertinentes.

Réseau d'essais BreedWheat

Prédiction des interactions génotypes x environnements

Intégration de covariables environnementales dans le modèle de prédiction

Qualités prédictives dans le réseau d'essais de blé tendre BreedWheat

Models	Cross-validation scenarios			
	random	newG	newE	newGE
EG	0,55	0,40	0,56	0,39
EG_GxE	0,63	0,52	-	-
EG_GxW	0,74	0,52	0,69	0,45

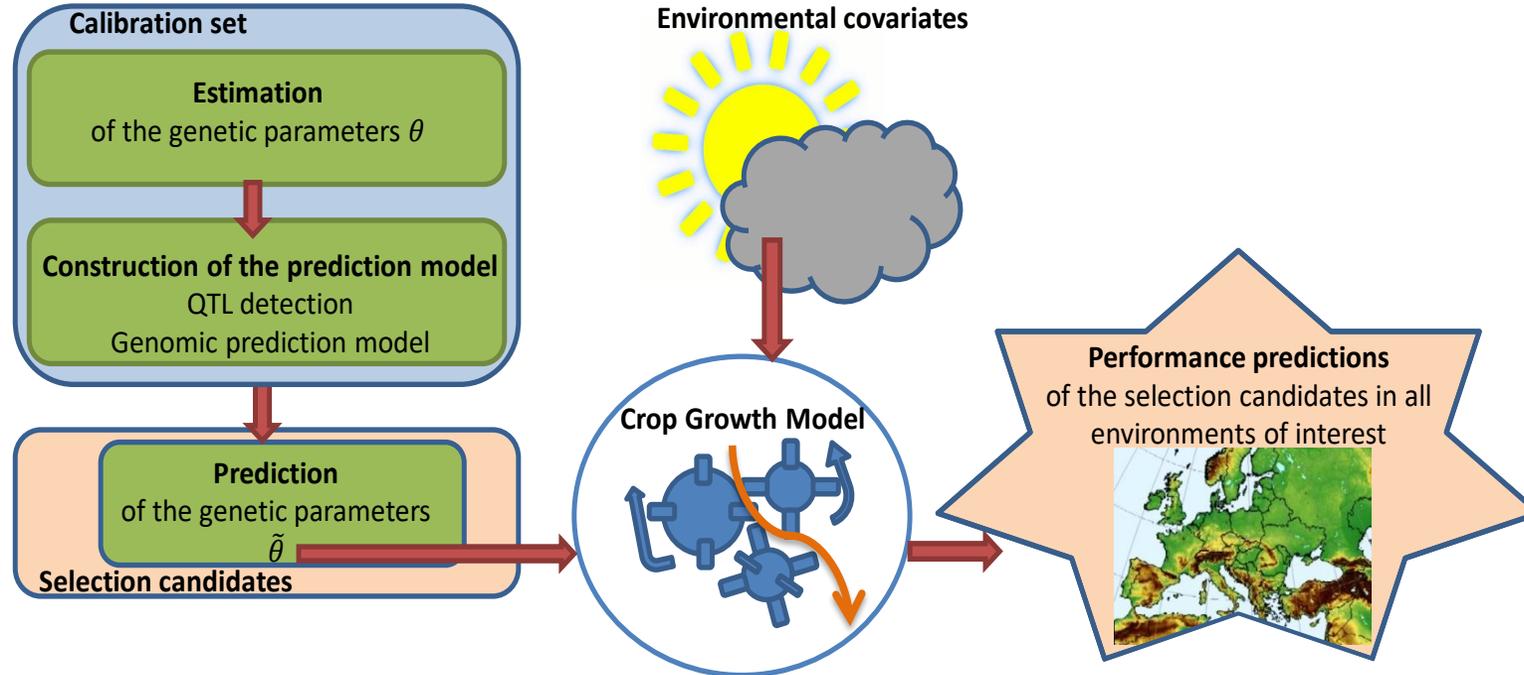
Sans cov. environnementale {

Avec cov. environnementale →

- Meilleures qualités prédictives avec les covariables environnementales.
- Possibilité de prédire de nouveaux environnements.

Prédiction des interactions génotypes x environnements

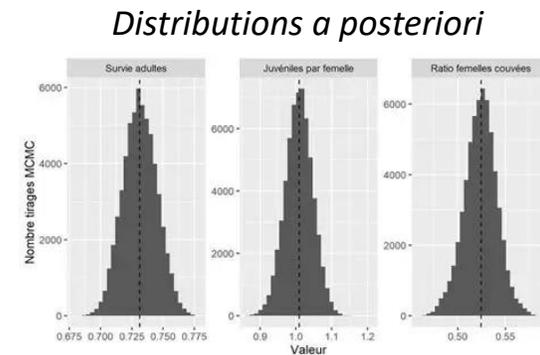
Le « Gene-based modelling » - mettre le modèle de culture au cœur du processus



H. Monod



E. Kuhn

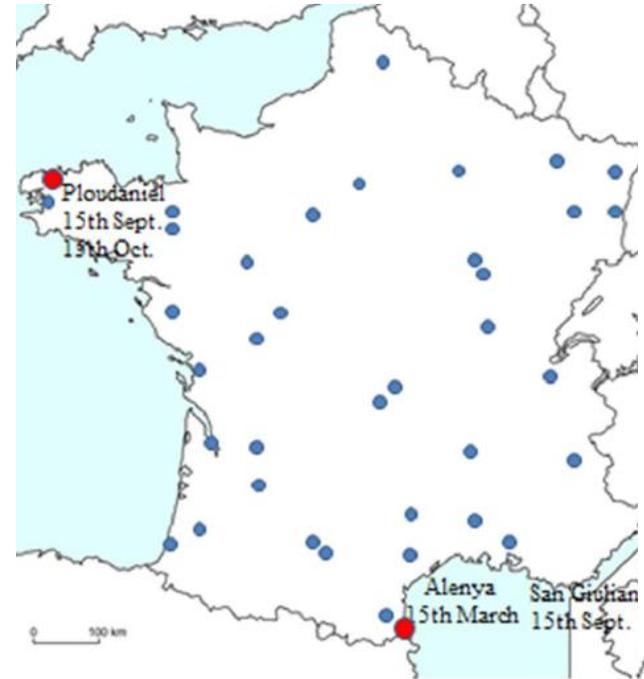


- Modèle fortement non linéaire. Comment estimer les paramètres génétiques ?
- Développement d'un échantillonneur pour réaliser de l'inférence bayésienne sur les paramètres.

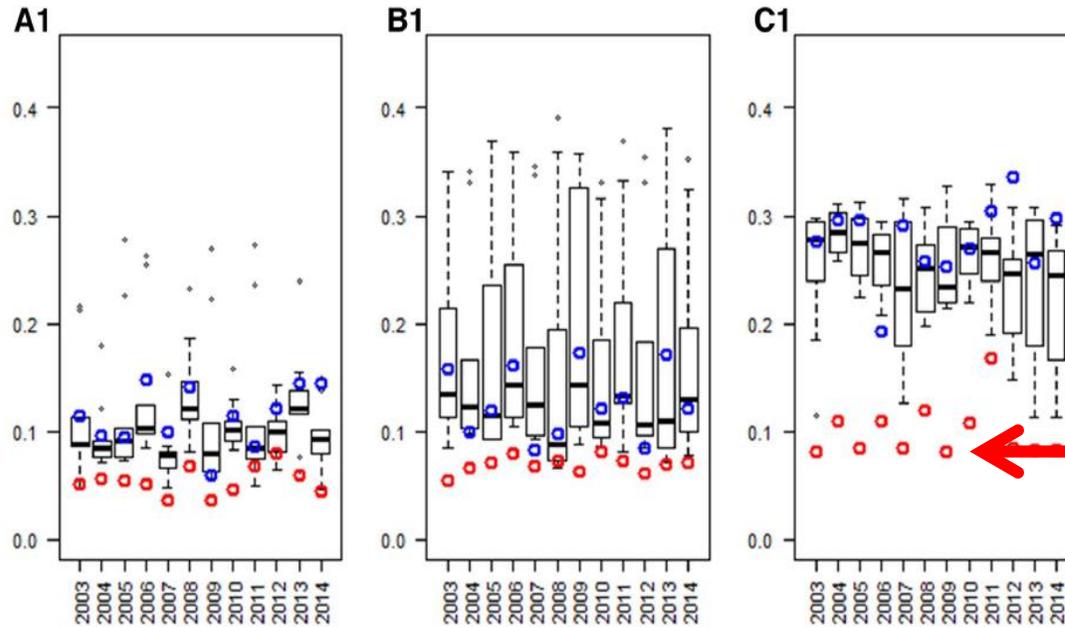
Prédiction des interactions génotypes x environnements

Optimisation du réseau d'essais pour l'estimation des paramètres génétiques

Exemple de la date de floraison chez le blé tendre



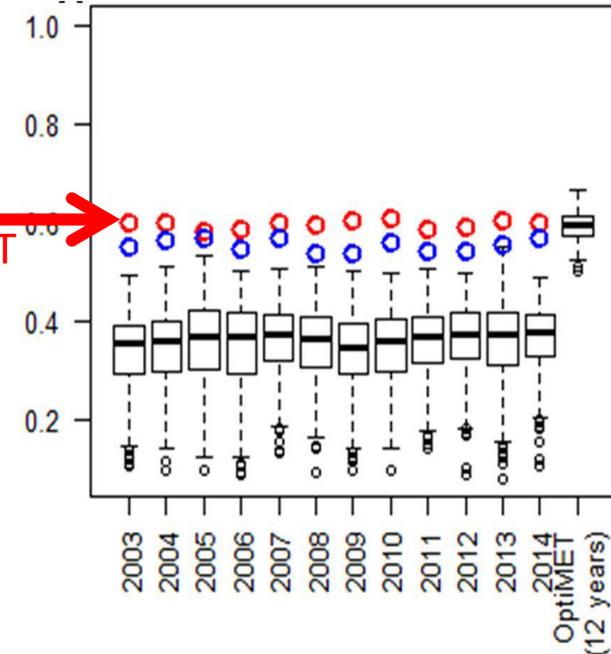
Réseau d'essais optimal (OptiMET)



NRMSE de trois paramètres génétiques avec différents réseaux d'essais (simulations sur blé tendre)

Réseau
OptiMET

Qualités prédictives de
la date de l'épiaison
par l'approche GBM



- Estimation plus fiable des paramètres génétiques.
- Meilleure qualité de prédiction du phénotype cible.

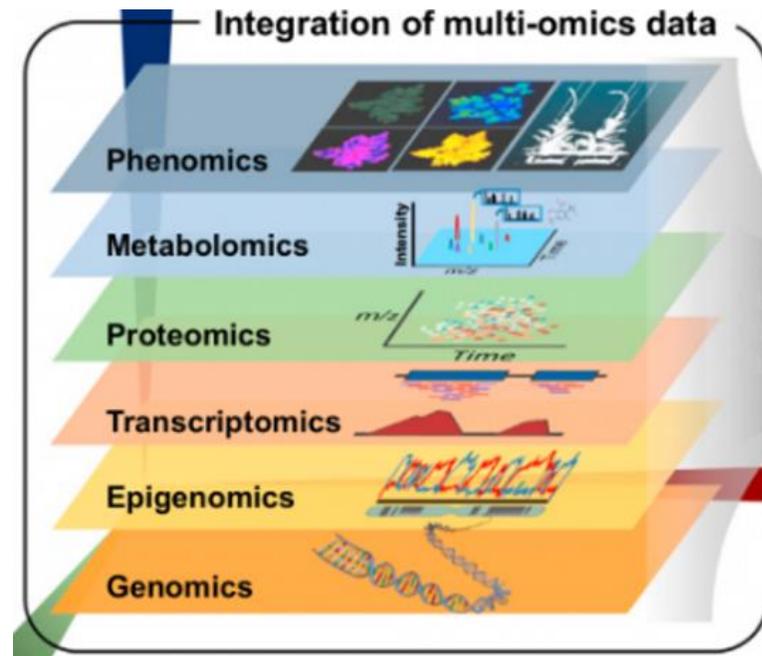
Prédiction des interactions génotypes x environnements

- Développement de variétés adaptées à des conditions pédoclimatiques particulières (choisies : itinéraires techniques faibles intrants, ou subies : stress biotiques et climatiques).
- Prédiction des performances dans des climats futurs.

3. Explorer la relation génotype – phénotype par la génétique des systèmes

La génétique des systèmes...

« La génétique des systèmes est une approche permettant de comprendre les flux d'information qui sous-tendent les caractères complexes. » Civelek et Lusi (2014)



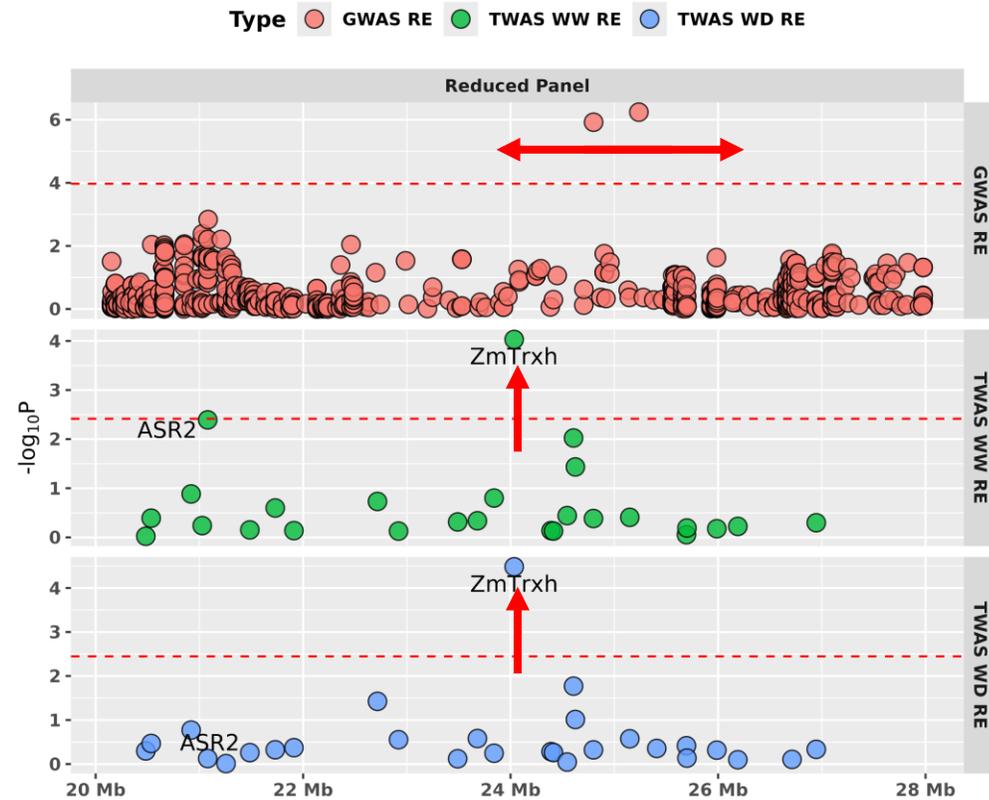
D'après Kim et al. (2016)

- Il est désormais possible de caractériser des panels en multi-omiques (ex. BreedWheat, Amaizing).
- **Objectif** : mieux modéliser la relation génotype – phénotype pour **mieux la comprendre et mieux la prédire**.
- Utiliser des approches intégratives beaucoup plus proches de la biologie que les modèles de génétique quantitative.

La génétique des systèmes... .. pour des études d'association plus puissantes et plus résolutes



B. Ali



Associations SNP - caractère

Associations transcrits (irrigué) - caractère

Associations transcrits (stress hydrique) - caractère

- Les associations avec des transcrits permettent d'identifier directement les gènes.
- Meilleure résolution qu'en génétique d'association.

La génétique des systèmes... ... pour une meilleure prédiction des interactions génotypes x environnements



B. Ali

Model type	Model	Platform traits				Field traits (GY)	
		WD		WW		CV-NH	CV-POH
		CV-NH	CV-POH	CV-NH	CV-POH		
Genomic	GBLUP	0.26 ± 0.11		0.24 ± 0.10		0.57 ± 0.09	
	GxE-BLUP	–	–	–	–	0.58 ± 0.09	0.72 ± 0.06
	GxW-BLUP	–	–	–	–	0.58 ± 0.05	0.72 ± 0.08
	MegaLMM	0.25 ± 0.11	0.88 ± 0.03	0.25 ± 0.10	0.89 ± 0.03	0.59 ± 0.09	0.77 ± 0.05
Standard Multi-omics	OBLUP	0.44 ± 0.09		0.42 ± 0.09		0.55 ± 0.10	
	G + OBLUP	0.43 ± 0.09		0.42 ± 0.09		0.54 ± 0.10	
	GOxE-BLUP	–	–	–	–	0.57 ± 0.10	0.72 ± 0.06
	GOxW-BLUP	–	–	–	–	0.59 ± 0.10	0.72 ± 0.06
	MegaGO	0.15 ± 0.14	0.58 ± 0.12	0.13 ± 0.15	0.59 ± 0.09	0.58 ± 0.09	0.61 ± 0.10

Qualités prédictives de modèles génomiques et omics (Amazing)

- Qualité prédictive des omics bien meilleure qu'avec la génomique en plateforme.
- Illustre la capacité des omics à prédire les interactions génotypes x environnements.
- Vers une caractérisation multi-omics de plein champs (ANR GENESYS) ?

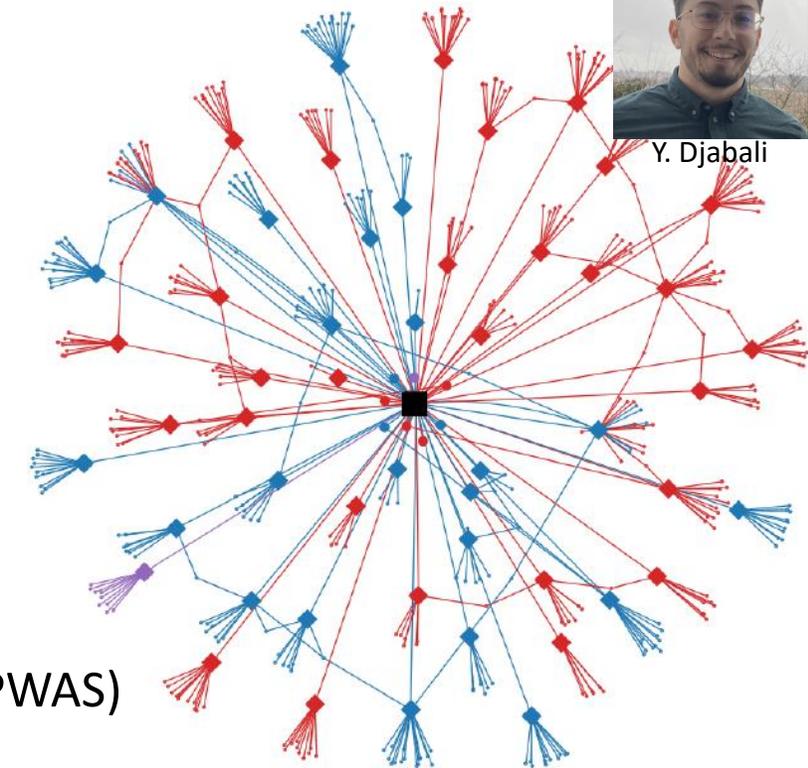
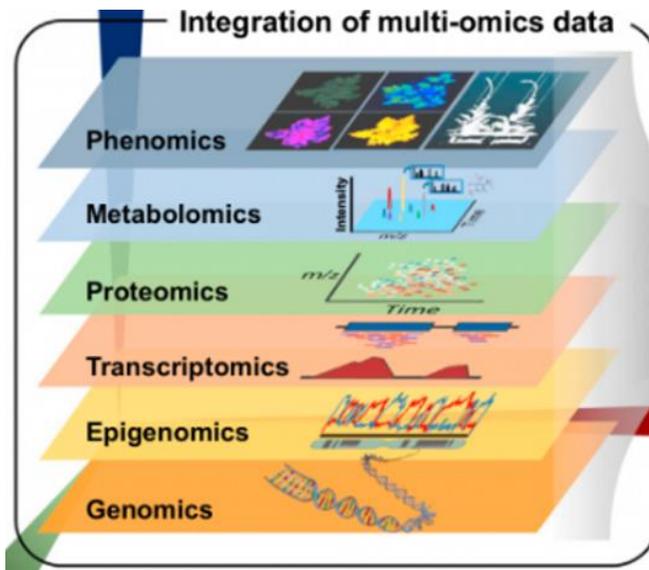
La génétique des systèmes... ... en plein champs

(ANR GENESYS, M. Blein-Nicolas, M. Martin, S. Coursol, R. Chapuis)

M. Blein-Nicolas M. Martin



Y. Djabali



- Associations entre variabilité moléculaires et phénotypes (GWAS, TWAS, PWAS)
- Analyses eQTL et pQTL des molécules significatives.
- Identification de relations directes par modèle graphique gaussien.
- Reconstruire les flux d'information du génotype au phénotype.

Conclusions

- Quels objectifs de sélection dans un contexte de changement global ?
- Les approches prédictives sont des outils essentiels pour aborder les questions d'adaptation, de résilience des variétés, des systèmes d'exploitation face au changement climatique.
- Les nouveaux outils (omics) devraient permettre de modéliser plus finement la relation génotype – phénotype. Meilleures résolutions des détections, meilleures compréhensions des processus...
- Pose beaucoup de questions en terme de modélisation statistique (grande dimension, faibles effectifs, non linéaire, structure de dépendances...).

Merci pour votre attention

