

Exploitation des données multi-omiques pour élucider les bases génétiques et moléculaires de caractères complexes: Une étude de génétique des systèmes de la réponse du maïs à la sécheresse

Yacine DJABALI

Directrice et co-directrice de thèse: Mélisande Blein-Nicolas et Marie-Laure Martin

Table des matières

I. CONTEXTE

- A. Implication socio-économique de la culture du maïs et des challenges associés
- B. Détection de QTLs par les approches de génétique d'association
- C. Les données omiques, un moyen de s'approcher de la complexité de la réponse au stress hydrique

II. MATERIEL & DONNEES

- A. Données de génotypage
- B. Données de phénotypage
- C. Objectifs de thèse

III. G x E & PLASTICITE

- A. Impact de la variabilité inter-essai sur la détection de QTLs impliqués dans la réponse au stress hydrique
- B. Les GWAS multi-trait améliorent la détection de QTLs de réponse au stress hydrique
- C. Les QTLs de plasticité sont spécifiquement impliqué dans la variance GxW

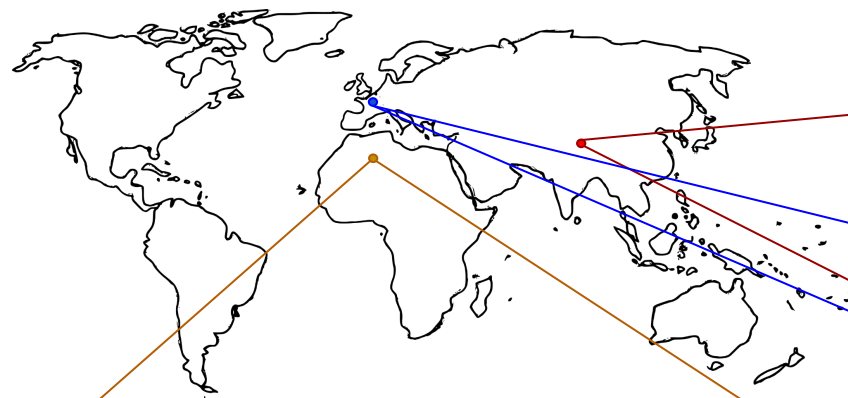
IV. INTÉGRATION DE DONNÉES OMIQUES

- A. Une étude de génétique des systèmes réalisé sur le même jeu de données
- B. Un tier de la variance GxW n'est pas expliquée par les QTLs détectés sur les caractères
- C. Apport fonctionnel des données de protéomique sur le GxW
- D. Apport quantitative des données de protéomique sur le GxW
- E. Inférence du réseau multi-échelle maximisant la part de variance capturé sur le GxW

V. DISCUSSION

- A. Interprétations des résultats obtenues
- B. Perspectives

Avec de fascinantes capacités d'adaptation



Sahara, Algérie



Photo de l'U.S. Grains Council

Gif-Sur-Yvette, France



Photo de Sacha Revillon, Laboratoire GQE (Équipe BASE)

Mustang, Népal

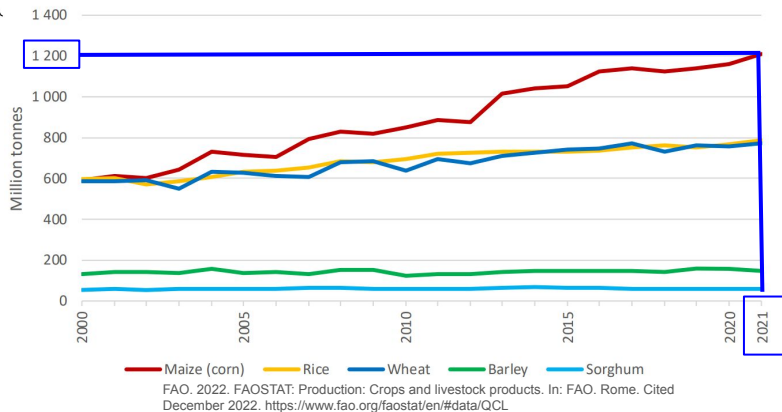
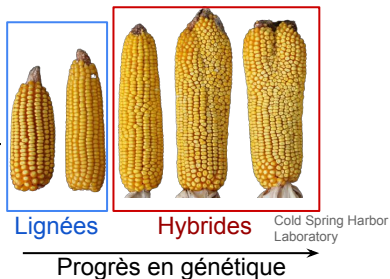


Photo de Chandan Chaurasia sur Unsplash

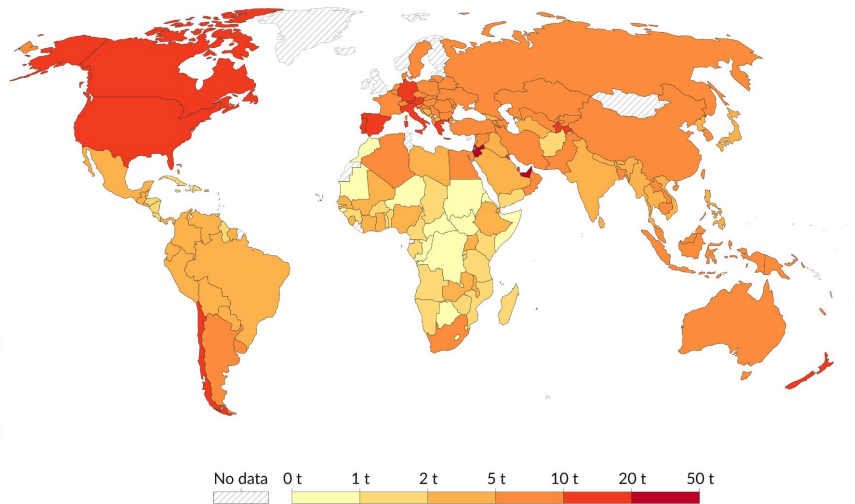
Lui permettant d'être impliqué dans des forts enjeux socio-économiques



Zea mays



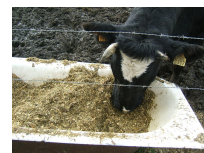
Rendement de maïs 2022 en fonction des pays (FAOSTAT, 2022)



Sécurité alimentaire mondiale



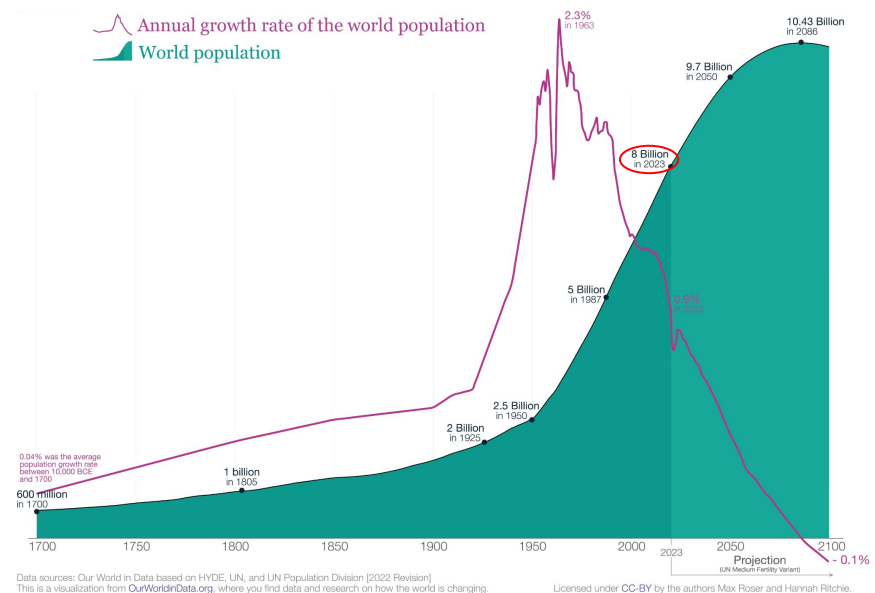
Consommation humaine et animale



Biocarburant

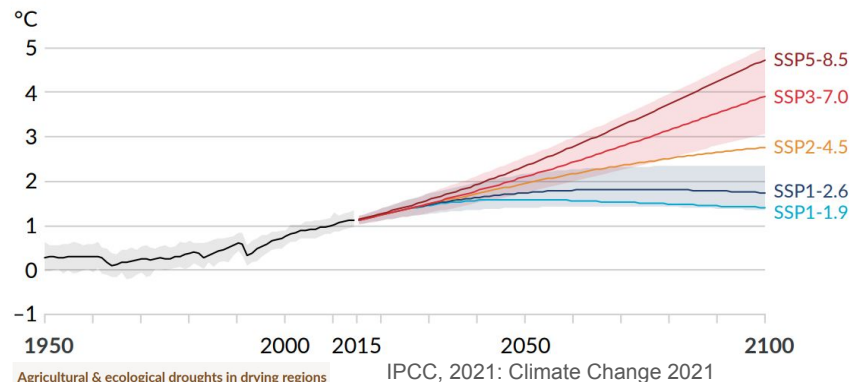
Sécurité alimentaire mondiale certes,... mais pour combien de temps ?

Croissance de la population mondiale



Prédiction de la montée des température à horizon 2100

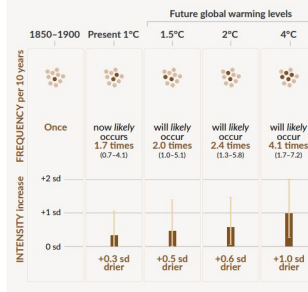
(a) Global surface temperature change relative to 1850–1900



Agricultural & ecological droughts in drying regions

10-year event

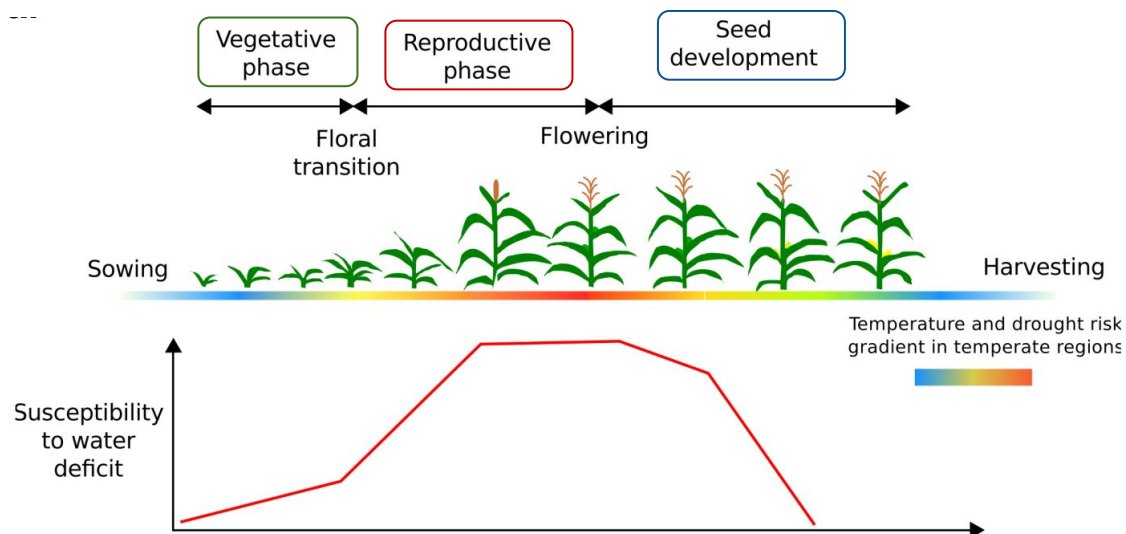
Frequency and increase in intensity of an agricultural and ecological drought event that occurred once in 10 years on average across drying regions in a climate without human influence



Des pertes agricoles s'élevant à 37 milliards de \$ selon la FAO



Un besoin d'eau significatif pour maintenir les rendements de maïs



Floraison: Développement des organes reproducteurs de la plante

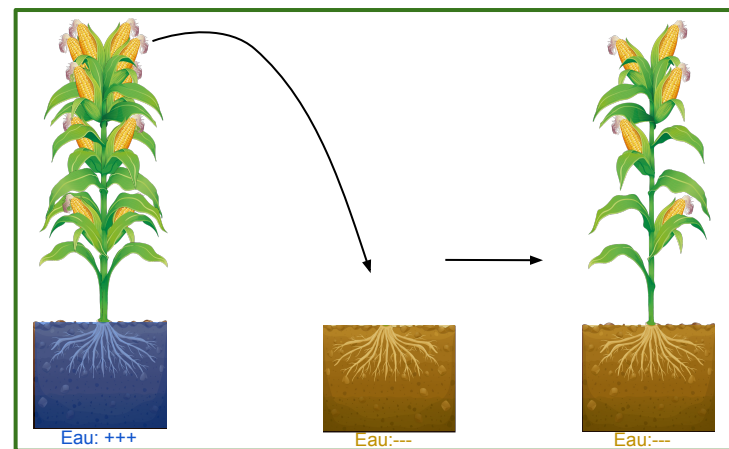
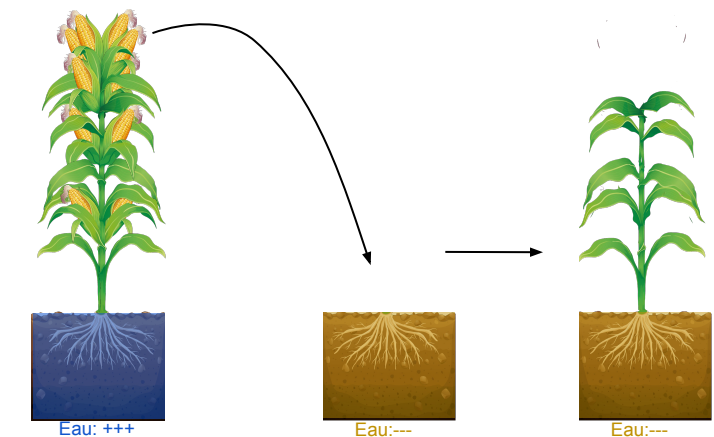
Conception des graines

Un manque d'eau durant:

- La **floraison** peut conduire une perte de rendement d'au moins **50%**
- Le stade **végétative** peut conduire une perte de rendement d'au moins **25%**
- Le **remplissage des graines** peut conduire une perte de rendement d'au moins **21%**

Denmeand and Shaw, 1960

Une solution: désigner des variétés tolérantes au stress hydrique



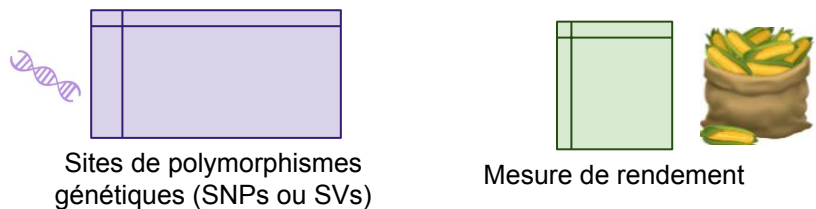
Sélection assistée par marqueurs

Panel de diversité



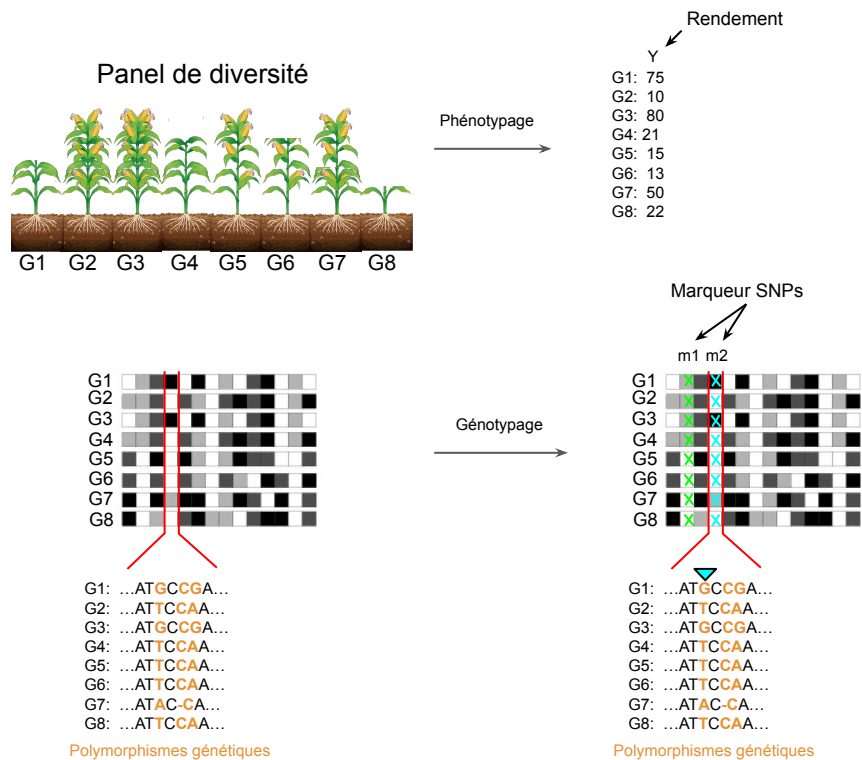
Génotypage

Phénotypage



Identification de gènes de tolérance par la détection de **Locus de caractères quantitatifs (QTLs)**

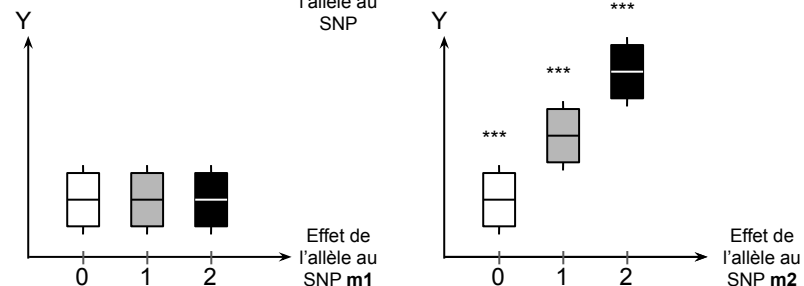
Détection de QTLs par étude d'association pangénomique (GWAS)



Modèle GWAS (Yu *et al.* 2006)

$$Y_g = \mu + \alpha \cdot X_g + G_g + \epsilon_g$$

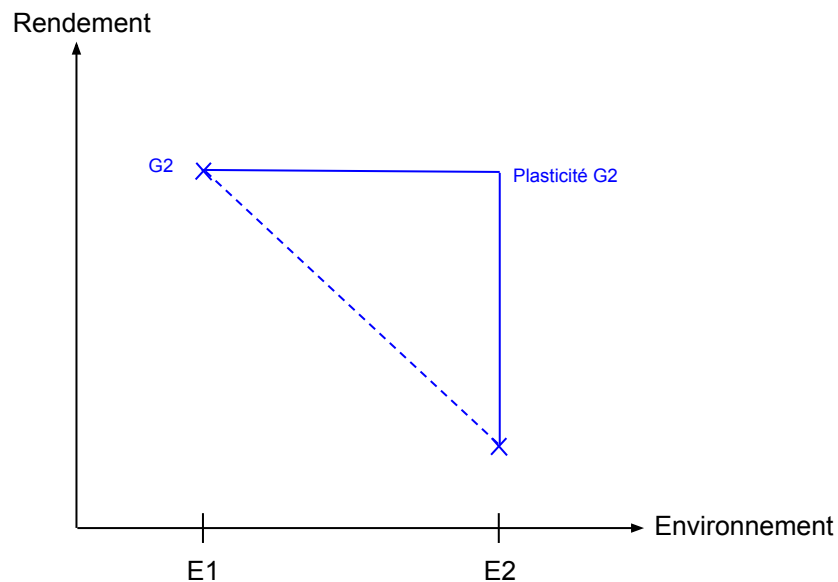
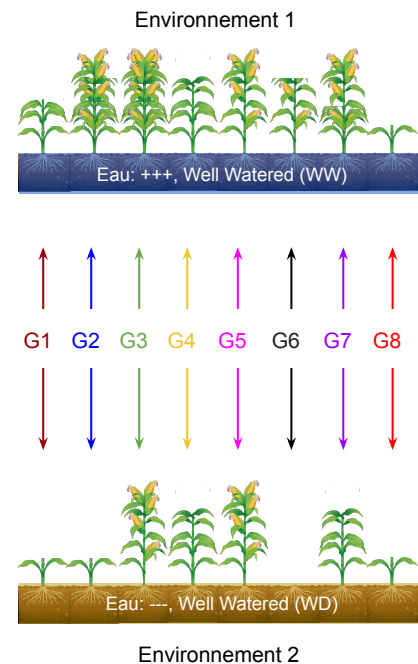
Effet de l'allèle au SNP



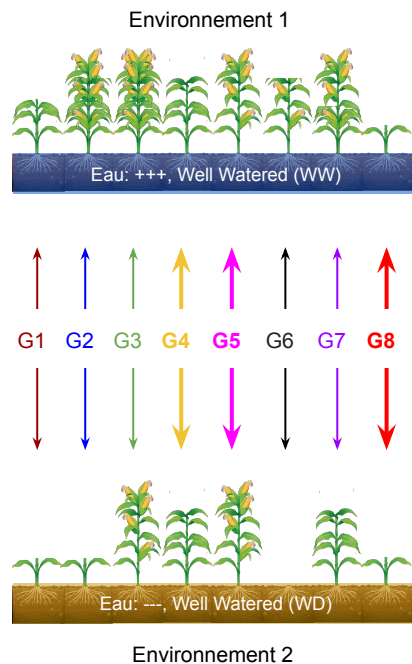
Association QTL



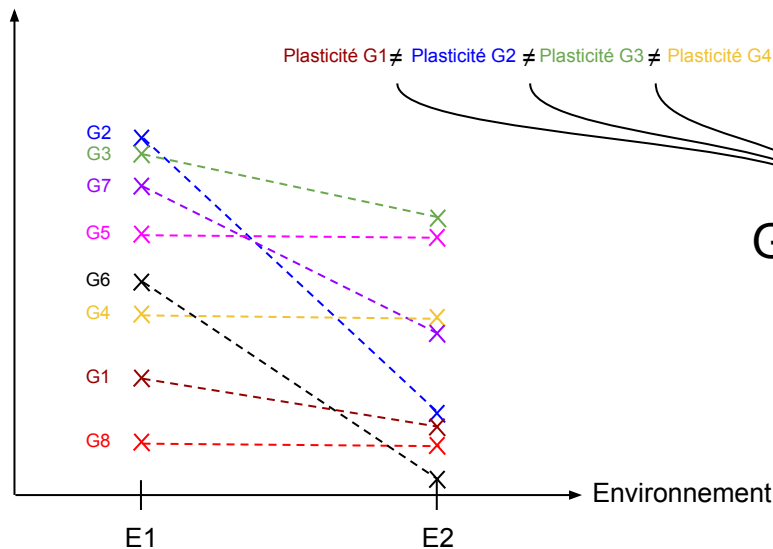
L'interaction Génotype x Environnement, une notion clé de la tolérance au stress



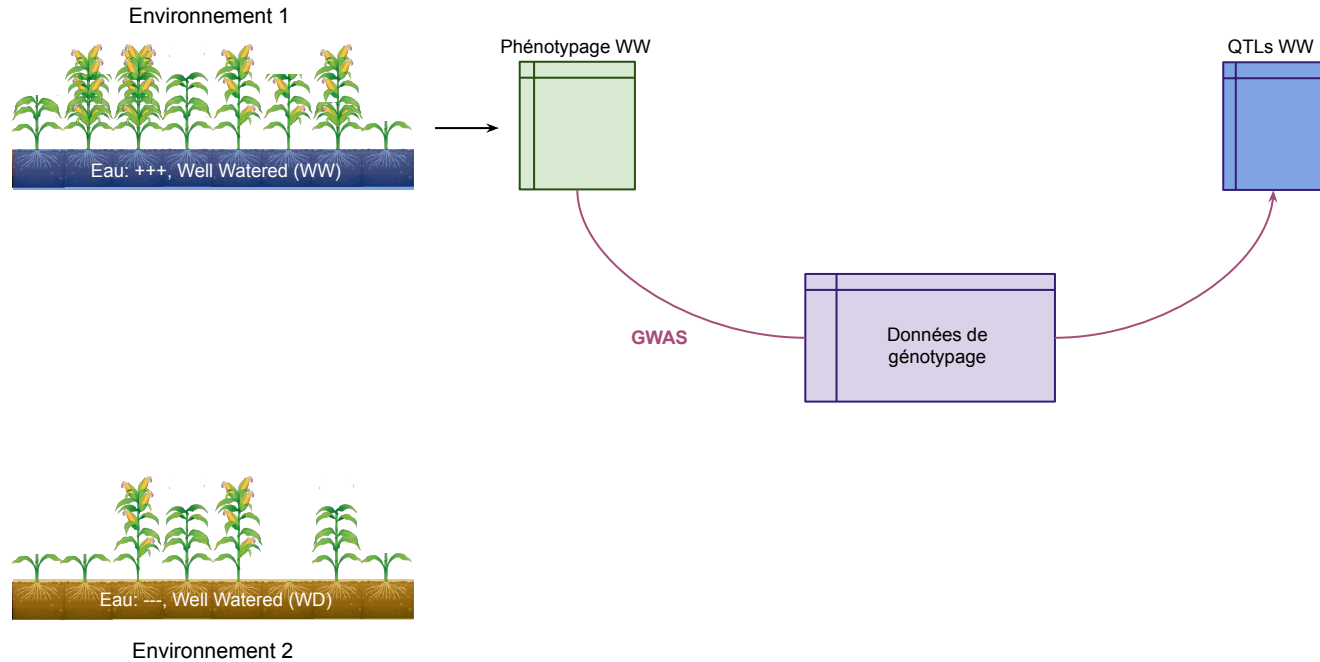
L'interaction Génotype x Environnement, une notion clé de la tolérance au stress



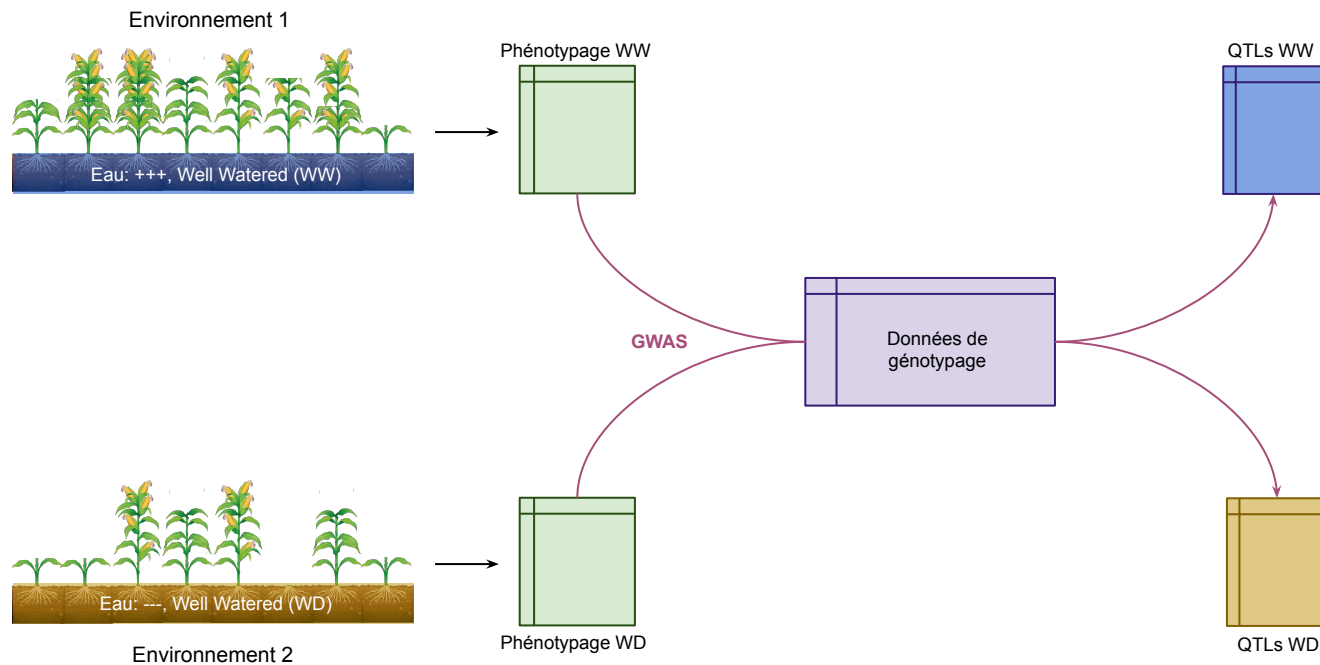
Rendement



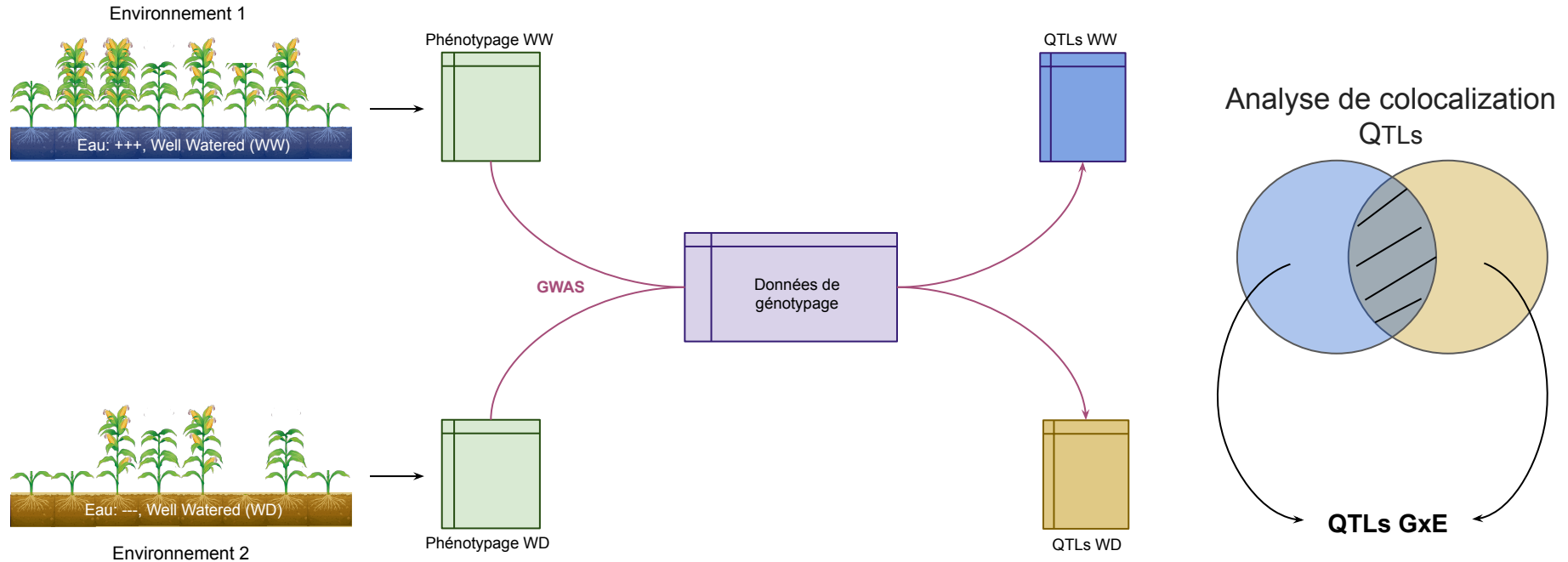
L'interaction Génotype x Environnement, une notion clé de la tolérance au stress



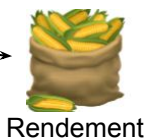
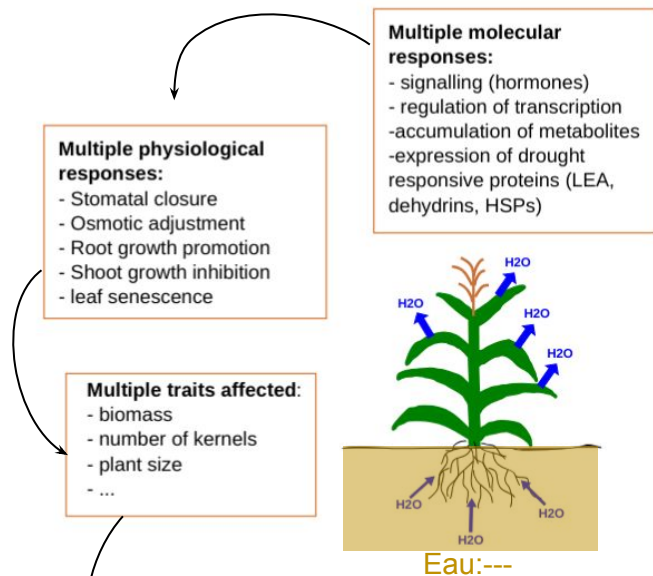
L'interaction Génotype x Environnement, une notion clé de la tolérance au stress



L'interaction Génotype x Environnement, une notion clé de la tolérance au stress

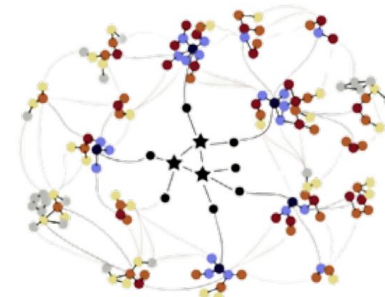
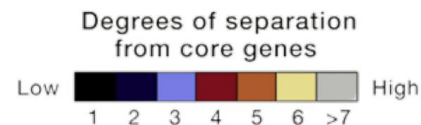


Un manque de puissance indéniable si l'étude est menée sur un seul caractère



Rendement

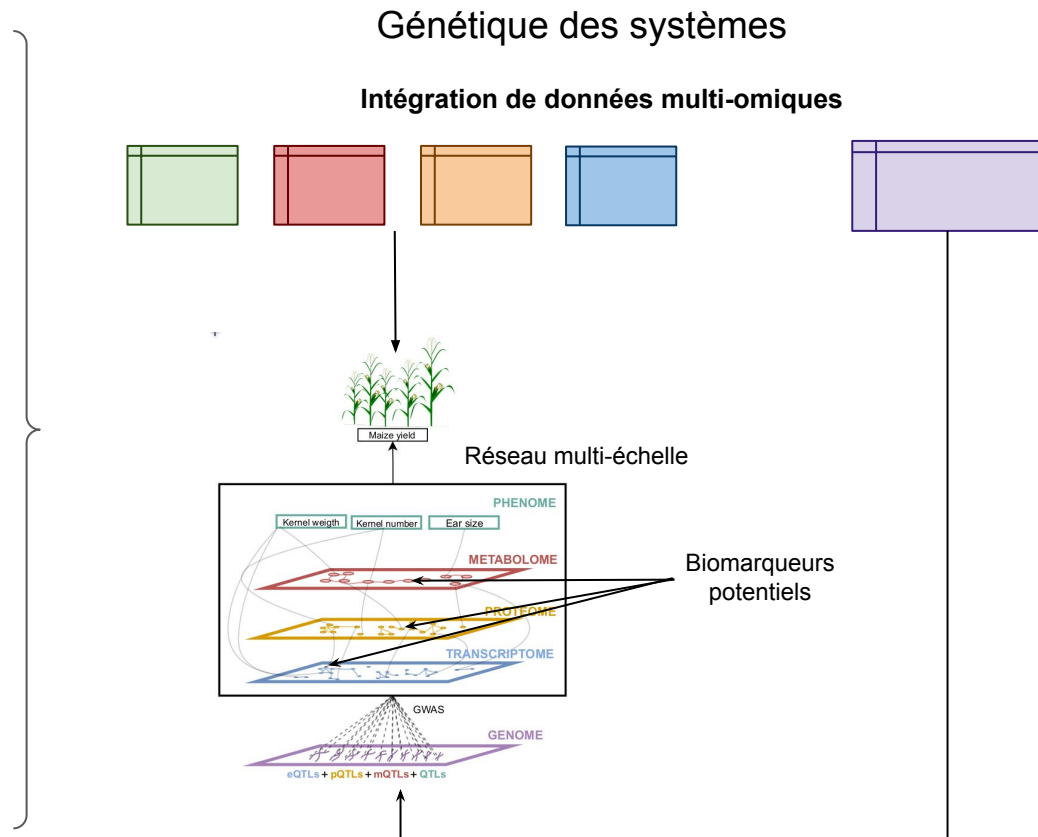
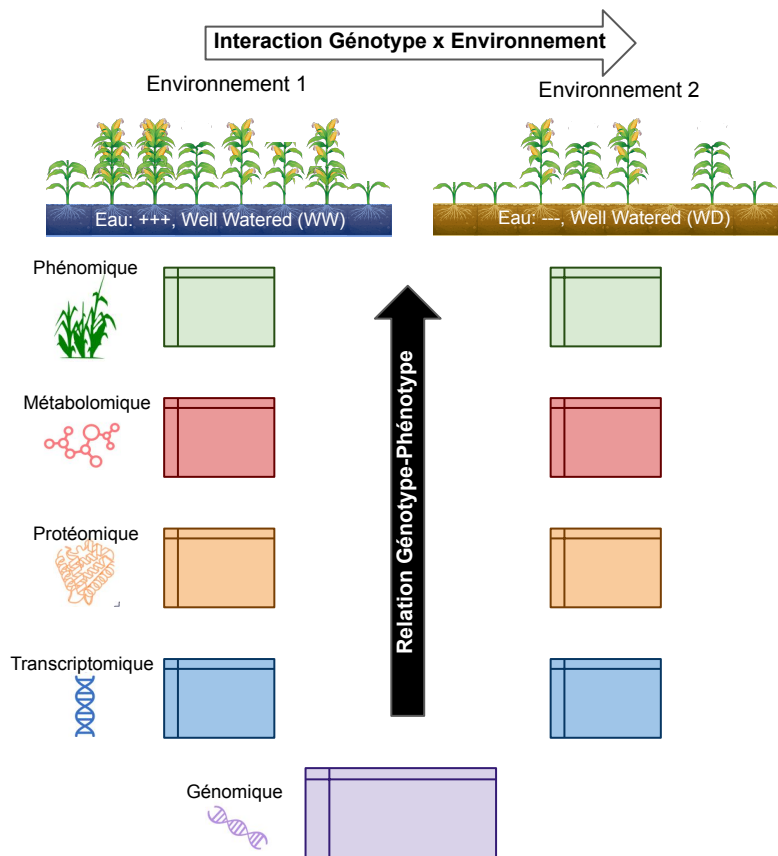
Caractère complexe



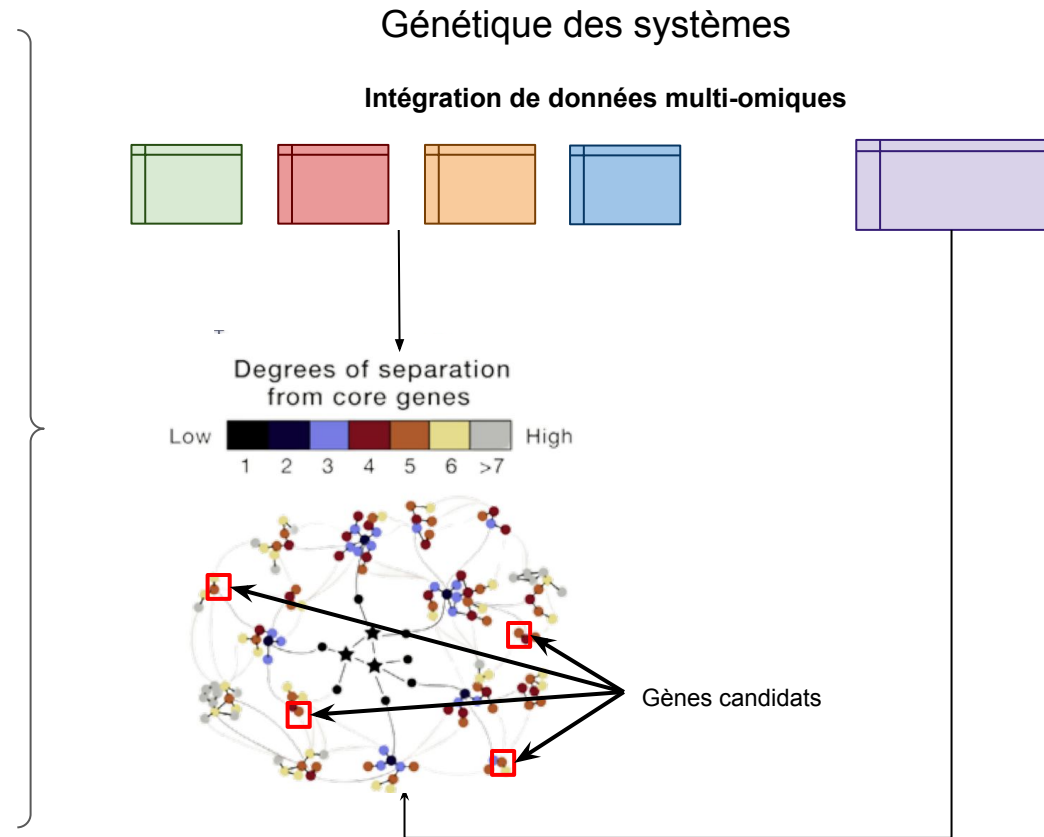
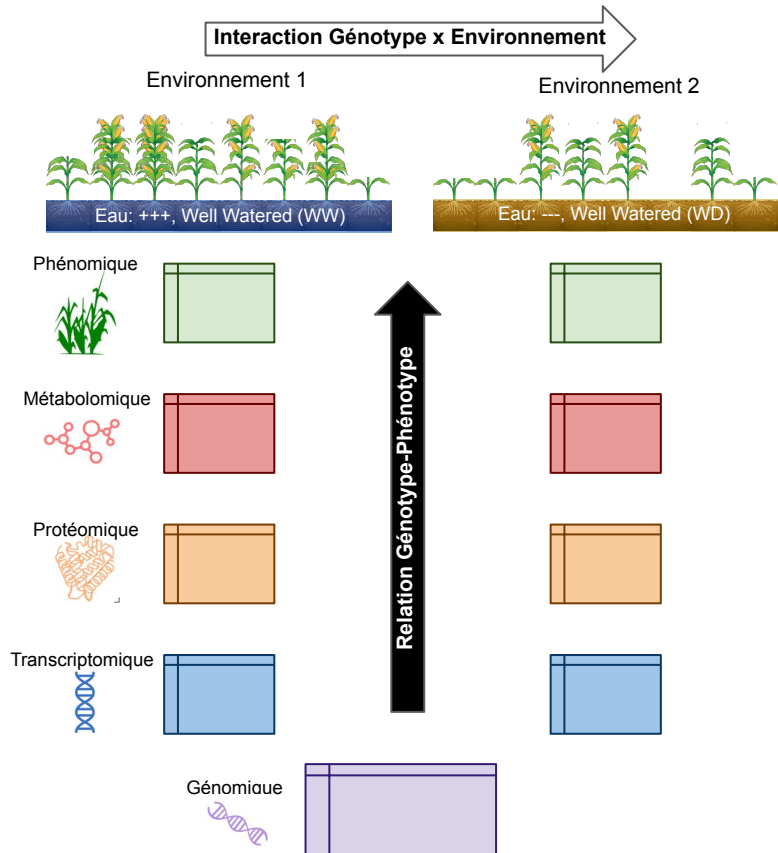
Modèle omni-génique (Boyle *et al.*, 2017)

Les QTLs détecté uniquement à partir du rendement ne sont pas suffisant pour expliquer les différences de réponse au stress hydrique

Ancrer la relation Génotype-Phénotype et le GxE dans des analyses systémiques



Ancrer la relation Génotype-Phénotype et le GxE dans des analyses systémiques



Ancrer la relation Génotype-Phénotype et le GxE dans des analyses systémiques

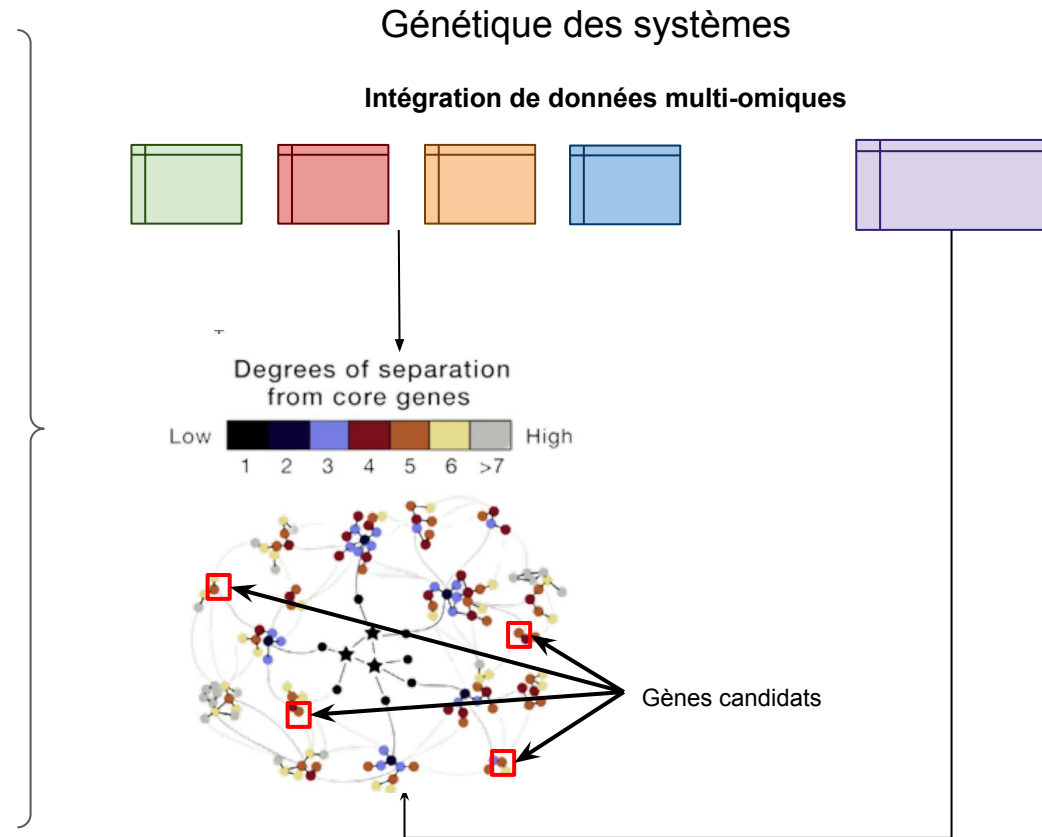
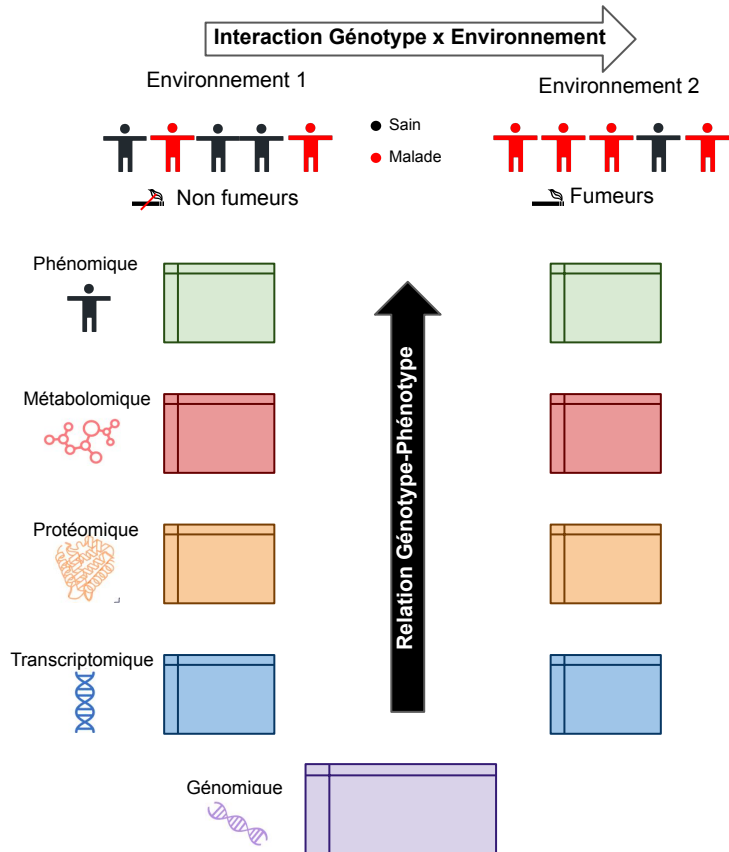
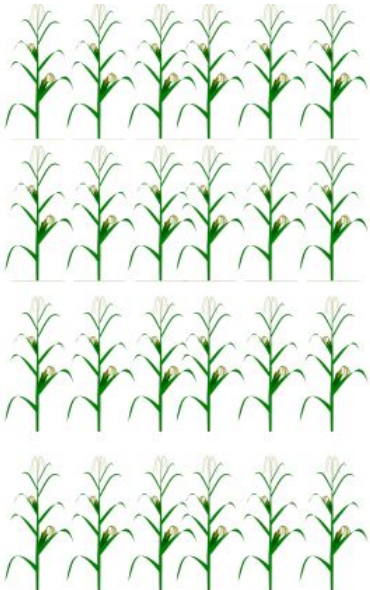


Table des matières

- I. CONTEXTE
 - A. Implication socio-économique de la culture du maïs et des challenges associés
 - B. Détection de QTLs par les approches de génétique d'association
 - C. Les données omiques, un moyen de s'approcher de la complexité de la réponse au stress hydrique
- II. **MATERIEL & DONNEES**
 - A. Données de génotypage
 - B. Données de phénotypage
 - C. Objectifs de thèse
- III. G x E & PLASTICITE
 - A. Impact de la variabilité inter-essai sur la détection de QTLs impliqués dans la réponse au stress hydrique
 - B. Les GWAS multi-trait améliorent la détection de QTLs de réponse au stress hydrique
 - C. Les QTLs de plasticité sont spécifiquement impliqué dans la variance GxW
- IV. INTÉGRATION DE DONNÉES OMIQUES
 - A. Une étude de génétique des systèmes réalisé sur le même jeu de données
 - B. Un tier de la variance GxW n'est pas expliquée par les QTLs détectés sur les caractères
 - C. Apport fonctionnel des données de protéomique sur le GxW
 - D. Apport quantitative des données de protéomique sur le GxW
 - E. Inférence du réseau multi-échelle maximisant la part de variance capturé sur le GxW
- V. DISCUSSION
 - A. Interprétations des résultats obtenues
 - B. Perspectives

Matériel végétale et production des données de génotypage

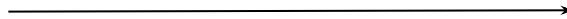


254 variétés dentés

Rincent *et al.*, 2014

Génotypage

- 50K Infinium HD Illumina array
- 600K Axiom affymetrix array
- Genotyping-by-sequencing



Données génomiques



977, 459 SNPs

Ganal *et al.*, 2011

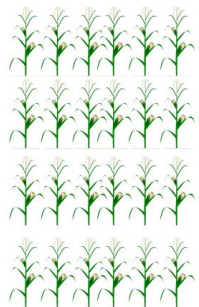
Unterseer *et al.*, 2014

Negro *et al.*, 2019

Production des données de phénotypage



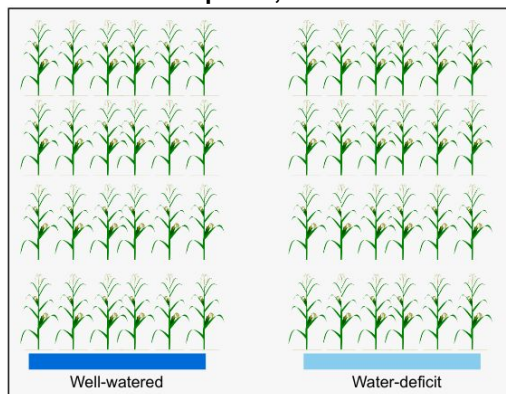
254 variétés dentés



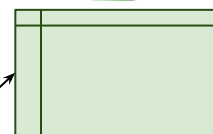
1 variété corné

Croisements

254 hybrides
Plateforme de phénotypage (serre) Phenoarch,
Montpellier, France



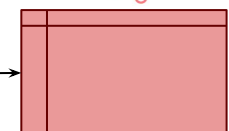
4 essais expérimentaux: Spring 2012, Spring 2013, Winter 2013, and Spring 2016



6 caractères
écophysologiques



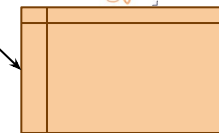
Alvarez Prado
et al., 2017



1,416 abondances de
métabolites



BORDEAUX
METABOLOME



2.055 abondances de
protéines



Blein-Nicolas
et al., 2020

Objectifs de thèse

Comprendre les bases génétiques et moléculaires de la réponse du maïs à la sécheresse à l'aide d'une approche intégrative basée sur les données multi-omiques.

- ❖ Améliorer la détection des déterminants génétiques impliqués dans la réponse au stress hydrique en considérant les variations inter-essai et les indices de plasticités.
 - Prise en compte d'un effet "essai" dans les modèles GWAS
 - Décomposition de la variance GxE en Genotype x Disponibilité en eau (GxW) et Genotype x Essai (GxT)
 - Utilisation des indices de plasticité comme variable réponse dans les modèle GWAS

- ❖ Caractérisation génétique et moléculaire approfondie de la réponse du maïs au déficit hydrique en exploitant les données de phénomique, protéomique et génomique en suivant une approche de génétique des systèmes.
 - Utilisation d'un modèle multi-caractères pour l'identification du sous-ensemble de optimal de QTLs expliquant la variance GxW
 - Exploiter les données de protéomique pour fournir des réseaux d'interaction protéines-protéines (PPI) comme support fonctionnel aux QTLs
 - Exploiter les données de protéomique pour apporter de l'information génétique supplémentaire impliqué dans la variance GxW
 - Inférence d'un réseau multi-échelle contenant la totalité des entités biologique d'intérêt et ainsi mettre en valeur des biomarqueurs potentiels

Table des matières

- I. CONTEXTE
 - A. Implication socio-économique de la culture du maïs et des challenges associés
 - B. Détection de QTLs par les approches de génétique d'association
 - C. Les données omiques, un moyen de s'approcher de la complexité de la réponse au stress hydrique
- II. MATERIEL & DONNEES
 - A. Données de génotypage
 - B. Données de phénotypage
 - C. Objectifs de thèse
- III. G x E & PLASTICITE**
 - A. Impact de la variabilité inter-essai sur la détection de QTLs impliqués dans la réponse au stress hydrique
 - B. Les GWAS multi-trait améliorent la détection de QTLs de réponse au stress hydrique
 - C. Les QTLs de plasticité sont spécifiquement impliqué dans la variance GxW
- IV. INTÉGRATION DE DONNÉES OMIQUES
 - A. Une étude de génétique des systèmes réalisé sur le même jeu de données
 - B. Un tier de la variance GxW n'est pas expliquée par les QTLs détectés sur les caractères
 - C. Apport fonctionnel des données de protéomique sur le GxW
 - D. Apport quantitative des données de protéomique sur le GxW
 - E. Inférence du réseau multi-échelle maximisant la part de variance capturé sur le GxW
- V. DISCUSSION
 - A. Interprétations des résultats obtenues
 - B. Perspectives

Objectifs de thèse

Comprendre les bases génétiques et moléculaires de la réponse du maïs à la sécheresse à l'aide d'une approche intégrative basée sur les données multi-omiques.

- ❖ Améliorer la détection des déterminants génétiques impliqués dans la réponse au stress hydrique en considérant les variations inter-essai et les indices de plasticités.
 - Prise en compte d'un effet "essai" dans les modèles GWAS
 - Décomposition de la variance GxE en Genotype x Disponibilité en eau (GxW) et Genotype x Essai (GxT)
 - Utilisation des indices de plasticité comme variable réponse dans les modèle GWAS

Theoretical and Applied Genetics (2023) 136:228
<https://doi.org/10.1007/s00122-023-04458-z>

ORIGINAL ARTICLE



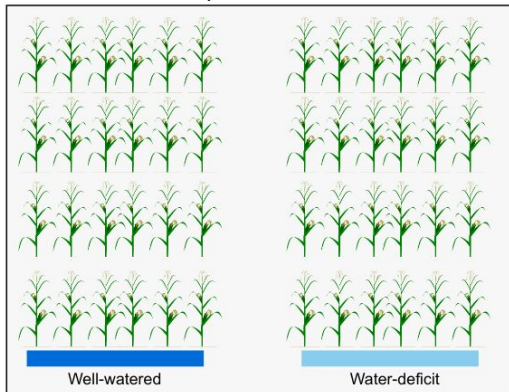
Plasticity QTLs specifically contribute to the genotype × water availability interaction in maize

Yacine Djabali^{1,2,3} · Renaud Rincent³ · Marie-Laure Martin^{1,2,4} · Mélisande Blein-Nicolas³

Approche classique de GWAS menée sur les données phénotiques

254 hybrides

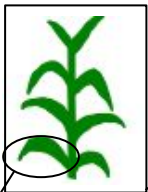
Plateforme de phénotypage (serre) Phenoarch,
Montpellier, France



4 essais expérimentaux: Spring 2012, Spring 2013, Winter 2013, and Spring 2016

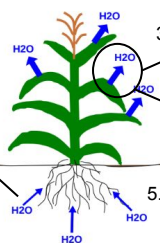
6 caractères écophysiologiques

1. Biomasse (Biom)



2. Surface foliaire (LA)

6. Utilisation efficace de l'eau (WUE)



3. Taux de transpiration (Trate)

4. Conductance stomatique (gs_max)

5. Absorption de l'eau (WU)



X 6

Spring 2012	0.50	0.30
	0.40	0.15
Spring 2013	0.70	0.54
	0.28	0.09
Winter 2013	0.33	0.11
	0.39	0.26
Spring 2016	0.04	0.65
	0.50	0.17

MOYENNES PHÉNOTYPIQUES

ORIGINAL ARTICLE

Phenomics allows identification of genomic regions affecting maize stomatal conductance with conditional effects of water deficit and evaporative demand Alvarez Prado *et al.*, 2017

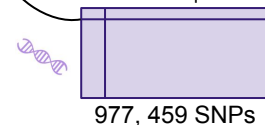
Absorption de l'eau (WU)

Spring 2012	0.50	0.30
	0.40	0.15
Spring 2013	0.70	0.54
	0.28	0.09
Winter 2013	0.33	0.11
	0.39	0.26
Spring 2016	0.04	0.65
	0.50	0.17

Modèle GWAS

$$Y_g = \mu + \alpha.X_g + \frac{G_g}{G_g} + \frac{\varepsilon_g}{\varepsilon_g}$$

Génomique

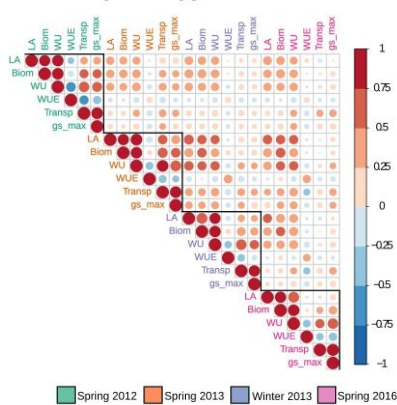


Détection de 531 QTLs par caractères, par condition, **par essai**

Variabilité inter-essai non prise en compte lors de la détection de QTLs

Forte variabilité inter-essai

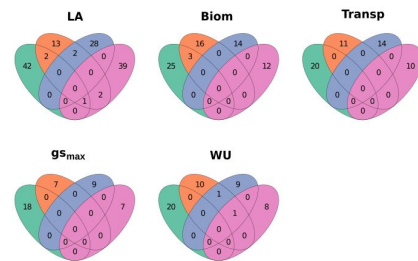
WD phenotypic means



GWAS

QTLs très spécifique de l'essai

Colocalization de QTLs



Problématique

On souhaite comprendre le déterminisme génétique en relation avec la réponse au stress hydrique et **non pas de l'environnement fluctuant présent au sein de la plateforme de phénotypage**

Solution

$$Y_g = \mu + \alpha.X_g + \underline{G}_g + \underline{\varepsilon}_g$$

→

$$Y_{gt} = \mu + \underbrace{T_t}_{\text{Effet essai}} + \alpha.X_g + \underline{G}_g + \underline{\varepsilon}_{gt}$$

Effet essai

Est-ce que le fait de prendre en compte l'effet "essai" dans le modèle de GWAS permet de détecter des sous ensembles de QTLs plus consistant de la variance GxW ?

Détection de QTLs consistant au GxW par la prise en compte de l'effet essai

GWAS multi-essai

$$Y_{gt} = \mu + T_t + \alpha X_g + \underline{G_g} + \underline{\varepsilon_{gt}}$$



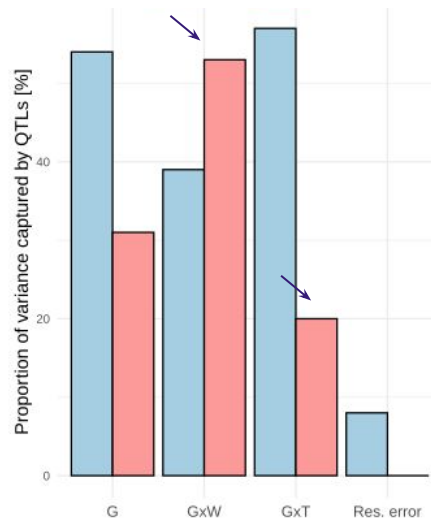
Détection de 102 QTLs par caractères, par condition

Modèle linéaire mixte de référence (sans QTLs)

$$Y_{gwt} = \mu + E_{wt} + PC_g + (PC \times E)_{gwt} + \underline{G_g} + \underline{(G \times W)_{gw}} + \underline{(G \times T)_{gt}} + \underline{\varepsilon_{gwt}}$$

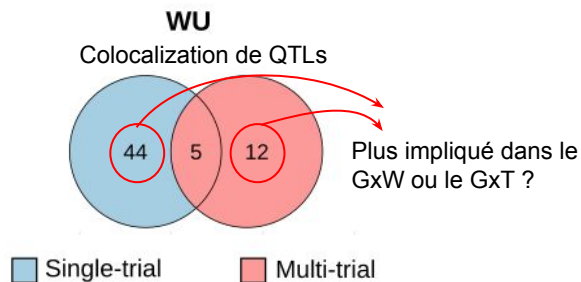
Modèle linéaire mixte multi-locus (avec QTLs)

$$Y_{gwt} = \mu + E_{wt} + PC_g + (PC \times E)_{gwt} + \underline{G_g} + \underline{(G \times W)_{gw}} + \underline{(G \times T)_{gt}} + \underline{\varepsilon_{gwt}}$$

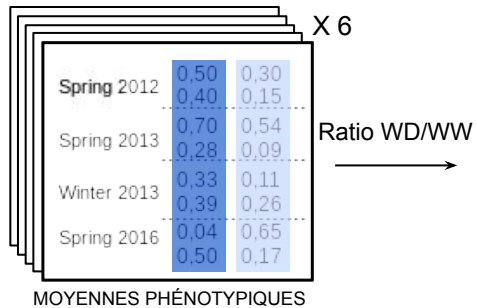
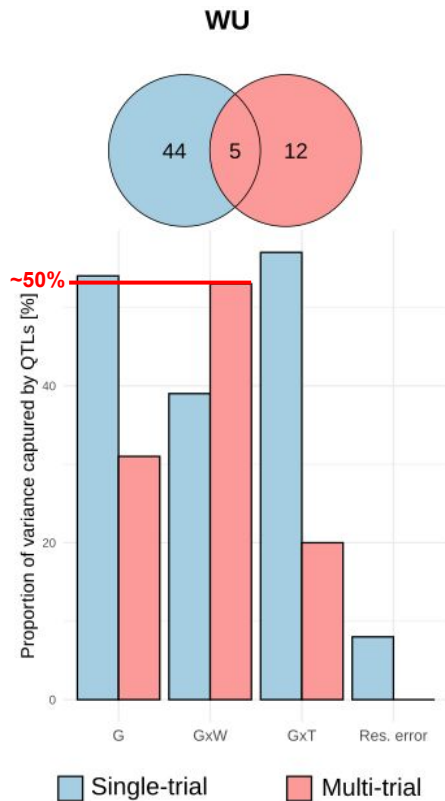


Les QTLs détecté avec le modèle GWAS multi-essai sont plus impliqués dans le GxW que le GxT contrairement aux QTLs détecté avec le modèle classique de GWAS

Découverte de nouveaux QTLs



Calcul d'indices de plasticité pour la détection de QTLs impliqué dans le GxW

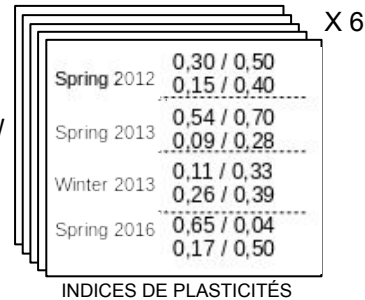


Problématique

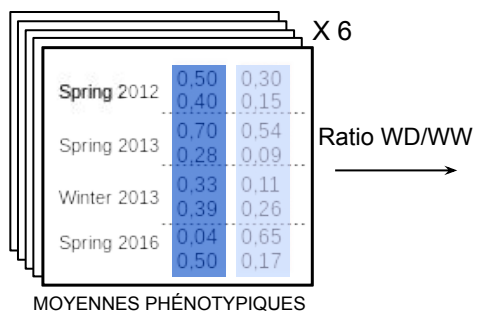
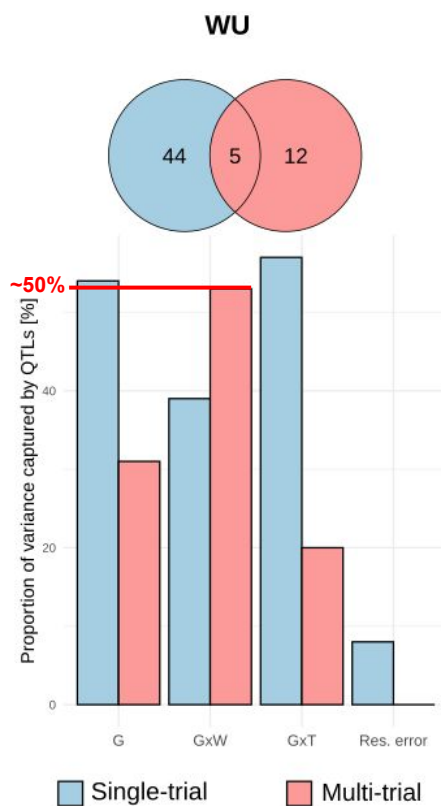
Une bonne moitié de la variance GxW n'est toujours pas expliquée

Solution

Utiliser des mesures plus direct de l'interaction GxW comme des indices de plasticités



Calcul d'indices de plasticité pour la détection de QTLs impliqués dans le GxW

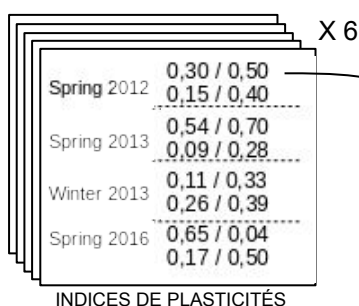


Problématique

Une bonne moitié de la variance GxW n'est toujours pas expliquée

Solution

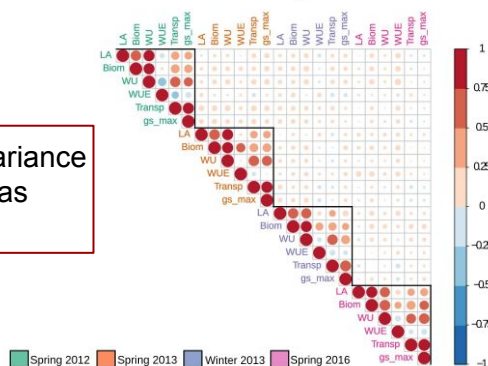
Utiliser des mesures plus directes de l'interaction GxW comme des indices de plasticités



GWAS multi-essai

$$Y_{gt} = \mu + T_t + \alpha X_g + \underline{G}_g + \underline{\varepsilon}_{gt}$$

Détection de 40 QTLs de plasticité



Quel est l'apport des QTLs de plasticité par rapport à ceux détectés sur les moyennes phénotypiques ?

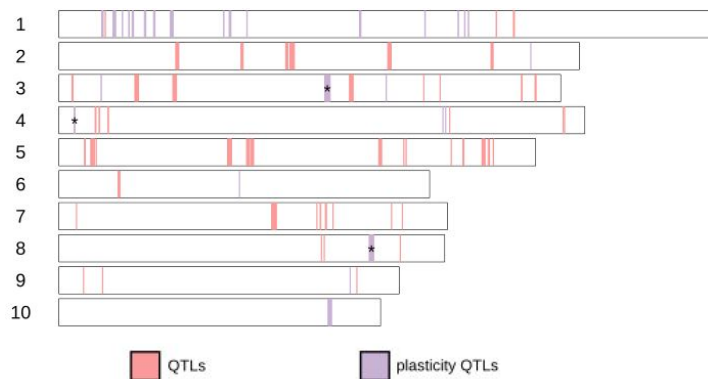
Les QTLs de plasticité capturent exclusivement une part de la variance GxW

GWAS multi-essai

$$Y_{gt} = \mu + T_t + \alpha X_g + \underline{G}_g + \underline{\varepsilon}_{gt}$$

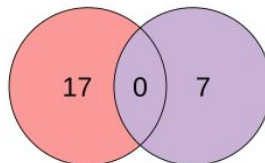


Détection de 40 QTLs de plasticité



Découverte de nouveaux QTLs

WU

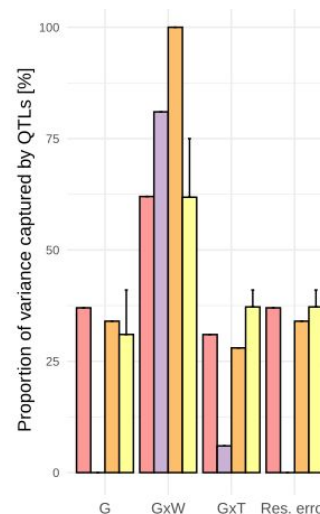


Modèle linéaire mixte de référence (sans QTLs)

$$Y_{gwt} = \mu + E_{wt} + PC_g + (PC \times E)_{gwt} + \underline{G}_g + \underline{(G \times W)}_{gw} + \underline{(G \times T)}_{gt} + \underline{\varepsilon}_{gwt}$$

Modèle linéaire mixte multi-locus (avec QTLs)

$$Y_{gwt} = \mu + E_{wt} + PC_g + (PC \times E)_{gwt} + \underline{G}_g + \underline{(G \times W)}_{gw} + \underline{(G \times T)}_{gt} + \underline{\varepsilon}_{gwt}$$



1. Les QTLs de plasticité expliquent exclusivement de la variance GxW
2. Les QTLs de plasticité et les QTLs détecté sur les moyennes phénotypiques sont complémentaire pour expliquer le GxW

Table des matières

- I. CONTEXTE
 - A. Implication socio-économique de la culture du maïs et des challenges associés
 - B. Détection de QTLs par les approches de génétique d'association
 - C. Les données omiques, un moyen de s'approcher de la complexité de la réponse au stress hydrique
- II. MATERIEL & DONNEES
 - A. Données de génotypage
 - B. Données de phénotypage
 - C. Objectifs de thèse
- III. G x E & PLASTICITE
 - A. Impact de la variabilité inter-essai sur la détection de QTLs impliqués dans la réponse au stress hydrique
 - B. Les GWAS multi-trait améliorent la détection de QTLs de réponse au stress hydrique
 - C. Les QTLs de plasticité sont spécifiquement impliqué dans la variance GxW
- IV. **INTÉGRATION DE DONNÉES OMIQUES**
 - A. Une étude de génétique des systèmes réalisé sur le même jeu de données
 - B. Un tier de la variance GxW n'est pas expliquée par les QTLs détectés sur les caractères
 - C. Apport fonctionnel des données de protéomique sur le GxW
 - D. Apport quantitative des données de protéomique sur le GxW
 - E. Inférence du réseau multi-échelle maximisant la part de variance capturé sur le GxW
- V. DISCUSSION
 - A. Interprétations des résultats obtenues
 - B. Perspectives

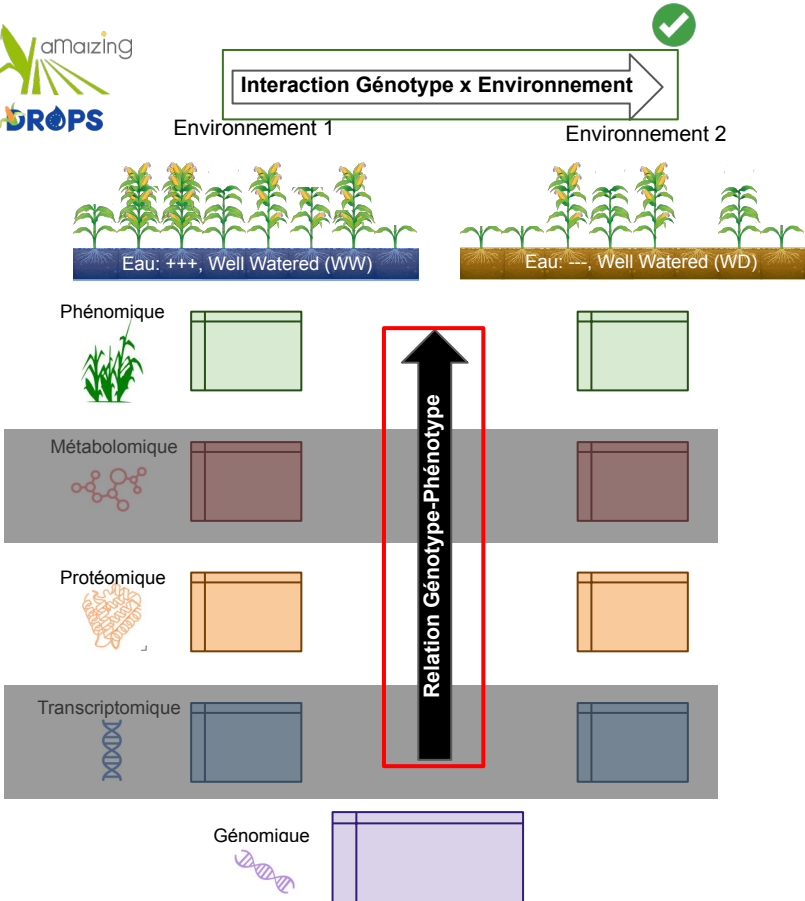
Table des matières

Comprendre les bases génétiques et moléculaires de la réponse du maïs à la sécheresse à l'aide d'une approche intégrative basée sur les données multi-omiques.

- ❖ Améliorer la détection des déterminants génétiques impliqués dans la réponse au stress hydrique en considérant les variations inter-essai et les indices de plasticités.
 - Prise en compte d'un effet "essai" dans les modèles GWAS
 - Décomposition de la variance GxE en Genotype x Disponibilité en eau (GxW) et Genotype x Essai (GxT)
 - Utilisation des indices de plasticité comme variable réponse dans les modèle GWAS

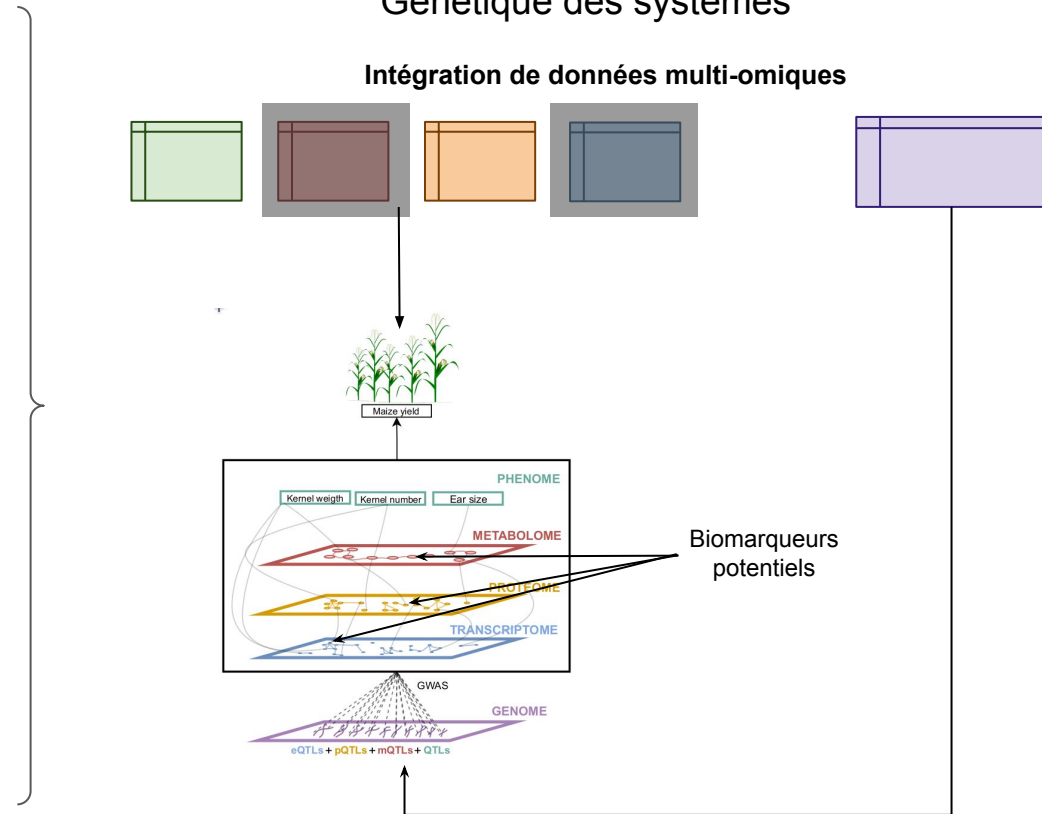
- ❖ Caractérisation génétique et moléculaire approfondie de la réponse du maïs au déficit hydrique en exploitant les données de phénomique, protéomique et génomique en suivant une approche de génétique des systèmes.
 - Utilisation d'un modèle multi-caractères pour l'identification du sous-ensemble de optimal de QTLs expliquant la variance GxW
 - Exploiter les données de protéomique pour fournir des réseaux d'interaction protéines-protéines (PPI) comme support fonctionnel aux QTLs
 - Exploiter les données de protéomique pour apporter de l'information génétique supplémentaire impliqué dans la variance GxW
 - Inférence d'un réseau multi-échelle contenant la totalité des entités biologique d'intérêt et ainsi mettre en valeur des biomarqueurs potentiels

Petit rappel et points de précisions



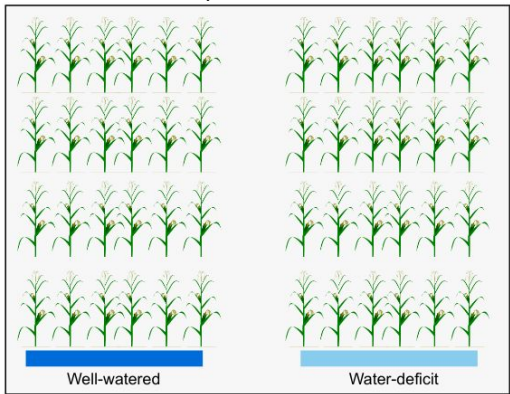
Génétique des systèmes

Intégration de données multi-omiques



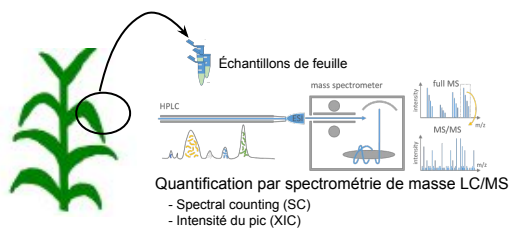
Première étude de génétique des systèmes réalisée sur ce jeu de données

254 hybrides
Plateforme de phénotypage (serre) Phenoarch,
Montpellier, France



Spring 2012

2,055 abondances de protéines



Quantification par spectrométrie de masse LC/MS
- Spectral counting (SC)
- Intensité du pic (XIC)



Research

A systems genetics approach reveals environment-dependent associations between SNPs, protein coexpression, and drought-related traits in maize

Blein-Nicolas *et al.*, 2020

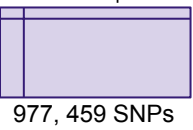


x 2,055

Modèle GWAS

$$Y_g = \mu + \alpha.X_g + G_g + \epsilon_g$$

Génomique



Détection de 22,000 pQTLs par abondance de protéine, par condition d'irrigation

Détection de 176 coQTLs par module de protéines, par condition d'irrigation

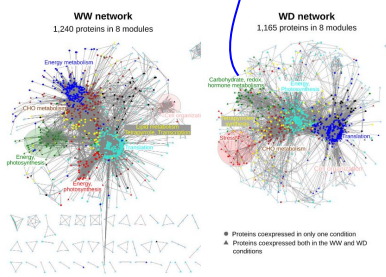


Phénomique



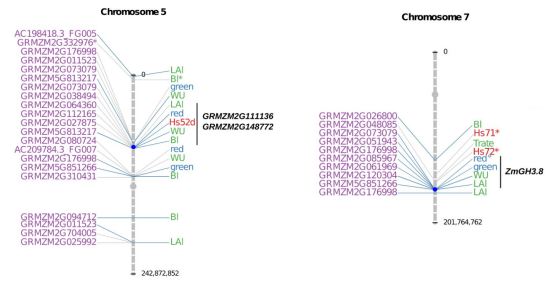
Alvarez Prado *et al.*, 2017

WGCNA



Réseaux de co-expression de protéines

Colocalization pQTLs, coQTLs, et QTLs



Première étude de génétique des systèmes réalisé sur ce jeu de données

Détection de 22,000 pQTLs par
abondance de protéine, par condition
d'irrigation

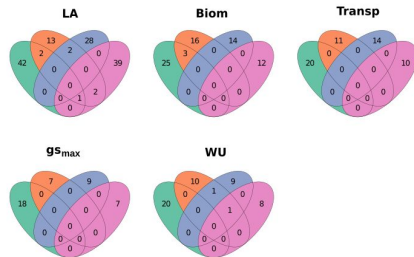
Détection de 176 coQTLs par module
de protéines, par condition d'irrigation

Phénomique



Alvarez Prado et al., 2017

Colocalization pQTLs, coQTLs, et **QTLs**



Problématique

Nous sommes pas certain que les régions de
colocalization sont explicatif du GxW

Solution

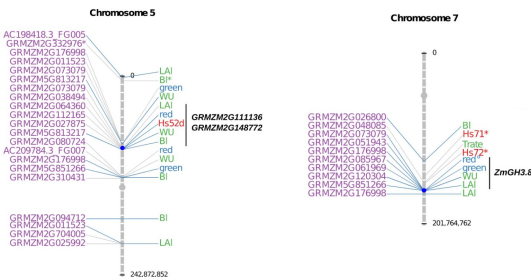
En se basant sur les QTLs et les QTLs de plasticité
détectés en partie III, utiliser les modèles linéaire
mixtes pour ne conserver seulement le
sous-ensemble de QTLs expliquant le maximum de
variance GxW

Modèle linéaire mixte

$$Y_{gwt} = \mu + E_{wt} + PC_g + (PC \times E)_{gwt} + \underline{G_g} + \underline{(G \times W)_{gw}} + \underline{(G \times T)_{gt}} + \underline{\epsilon_{gwt}}$$

Modèle linéaire mixte multi-caractère

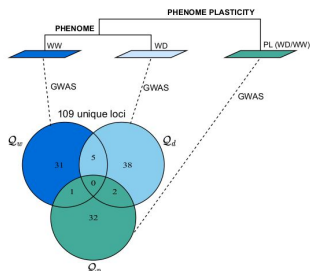
$$Y_{gwpt} = \mu + E_{wt} + \boxed{P_p} + (E \times P)_{wtp} + PC_g + (PC \times E)_{gwt} + (PC \times P)_{gp} \\ + \underline{G_g} + \underline{(G \times W)_{gw}} + \underline{(G \times P)_{gp}} + \underline{(G \times T)_{gt}} + \underline{\epsilon_{gwp}}$$



Identification du sous ensemble de QTLs maximisant la part de variance GxW capturée

$$Y_{gwpt} = \underbrace{\mu + E_{wt} + P_p + (E \times P)_{wtp} + PC_g + (PC \times E)_{gwt} + (PC \times P)_{gp}}_{X\beta} + \underbrace{G_g + (G \times W)_{gw} + (G \times P)_{gp} + (G \times T)_{gt} + \varepsilon_{gwpt}}_{Zu}$$

$M_0 : Y \sim \mathcal{N}(\bar{X}\beta, Zu)$
Modèle linéaire mixte multi-caractère
(référence, sans QTLs)



$$M_1 : Y \sim \mathcal{N}(X\beta + Q_w, Zu)$$

$$M_2 : Y \sim \mathcal{N}(X\beta + Q_d, Zu)$$

$$M_3 : Y \sim \mathcal{N}(X\beta + Q_w + Q_d, Zu)$$

$$M_4 : Y \sim \mathcal{N}(X\beta + Q_w, Zu)$$

$$M_5 : Y \sim \mathcal{N}(X\beta + Q_w + Q_p, Zu)$$

$$M_6 : Y \sim \mathcal{N}(X\beta + Q_d + Q_p, Zu)$$

$$M_7 : Y \sim \mathcal{N}(X\beta + Q_w + Q_p + Q_d, Zu)$$



Sélection des 71 loci associés au QTLs
(Qref) capturant près de 65% de la
variance GxW

FONCTIONNEL

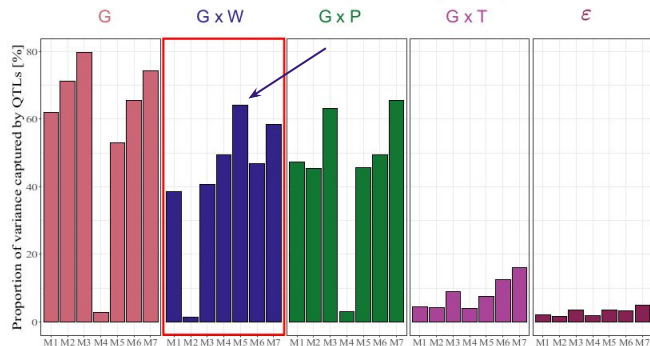
Ces 71 loci recouvrent 351 gènes mais aucun enrichissement fonctionnel n'a été trouvé

Peut-on exploiter les données de protéomique pour apporter une annotation fonctionnel de ces régions génomiques ?

QUANTITATIF

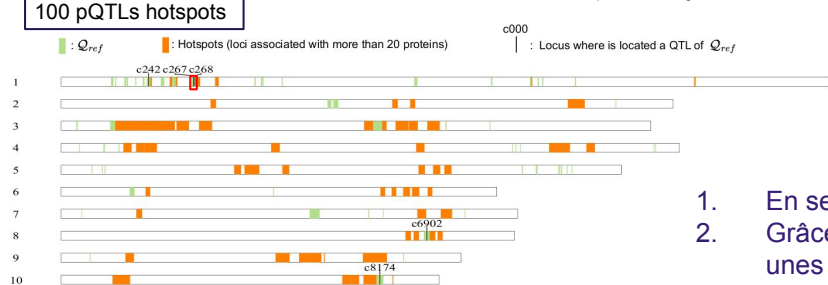
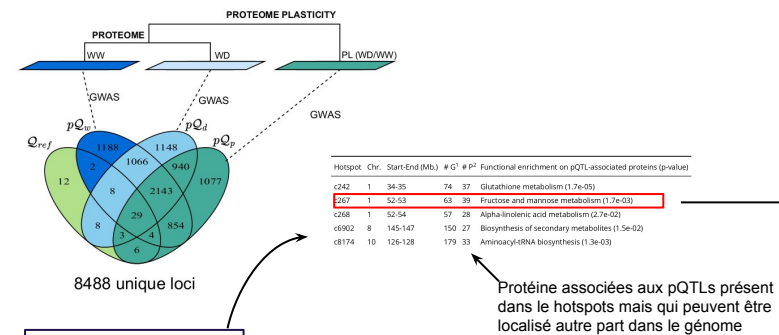
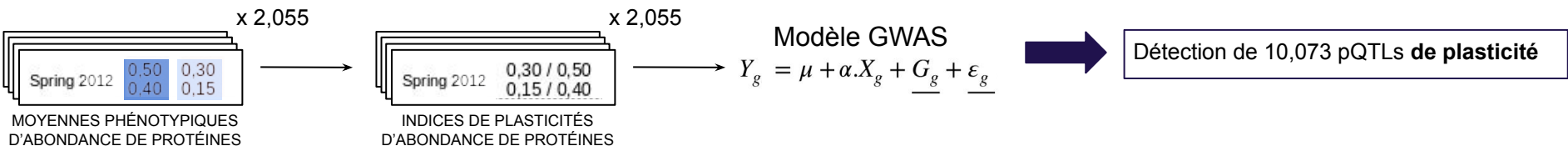
Encore 35% de la variance GxW n'a pas été capturé par l'ensemble Qref

Est-ce que les pQTLs peuvent permettre d'expliquer davantage de variance GxW ?

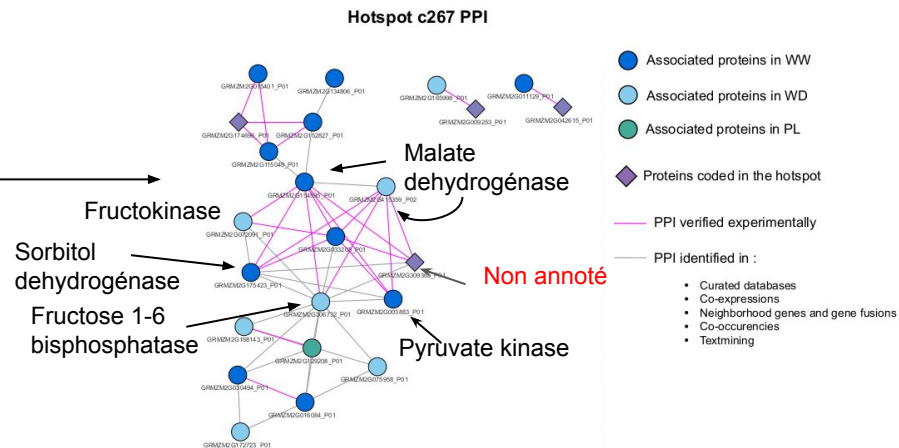


Les QTLs peuvent être intégrés dans des PPI par l'intermédiaire de la protéomique

FONCTIONNEL



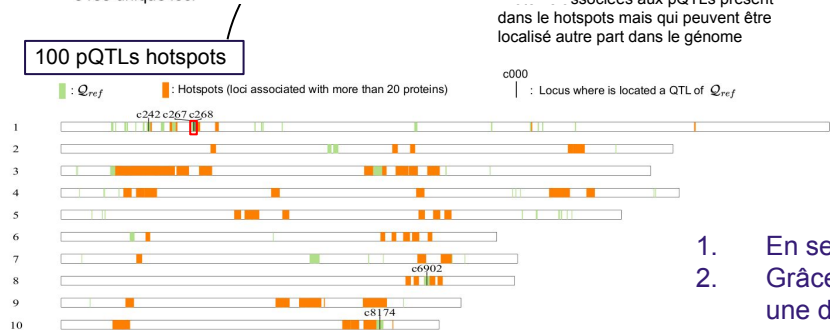
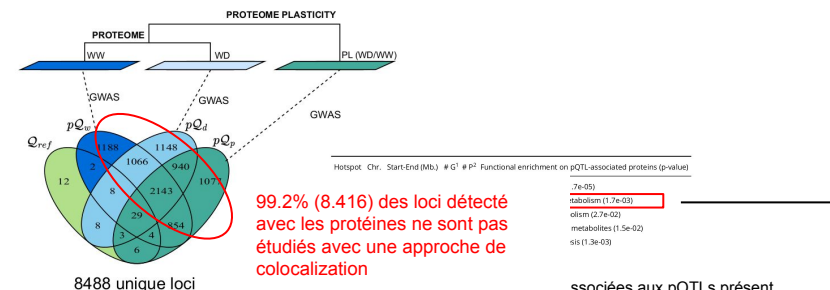
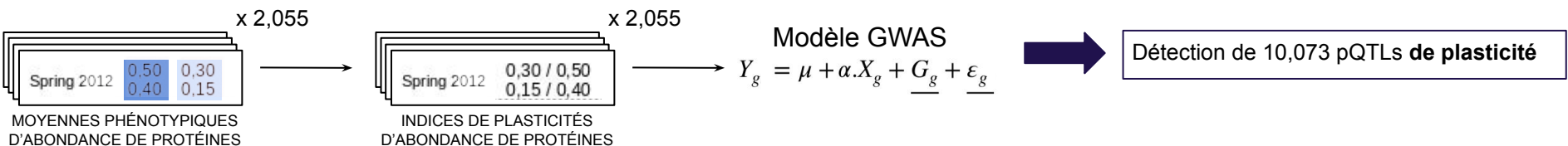
STRING



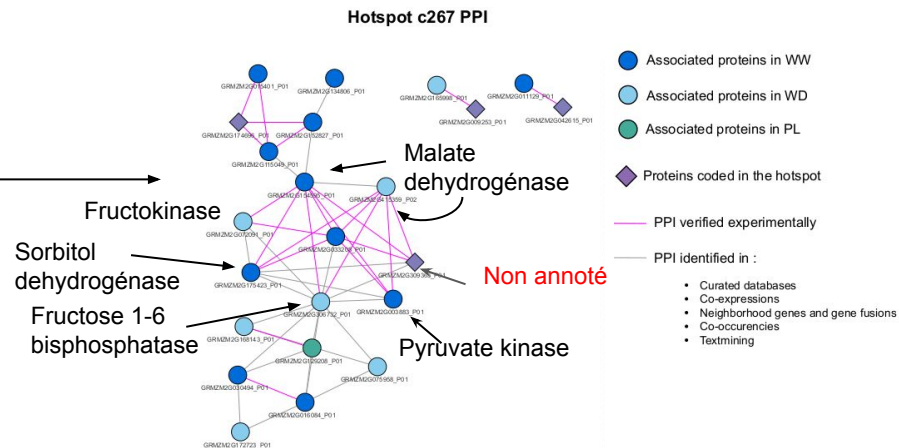
- En se basant sur les PPI nous pouvons déjà identifier des gènes candidats
- Grâce à l'étude du protéome on peut voir que les régions génétiques peuvent dépendre les unes des autres suivant le statut hydrique ou la plasticité de la plante

Les QTLs peuvent être intégrés dans des PPI par l'intermédiaire de la protéomique

FONCTIONNEL



STRING

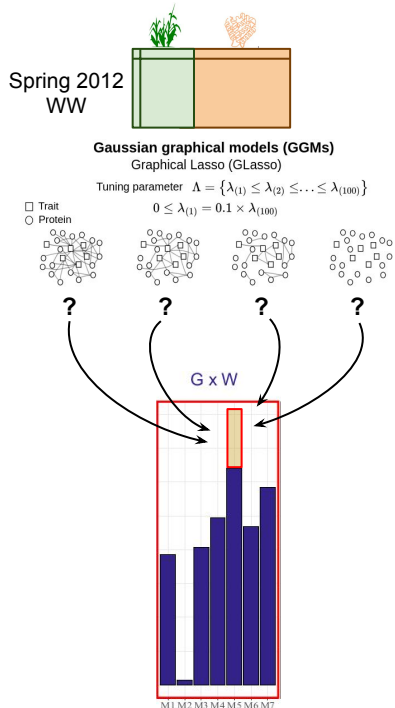


- En se basant sur les PPI nous pouvons déjà identifier des gènes candidats
- Grâce à l'étude du protéome on peut voir que les régions génétiques peuvent dépendre les unes des autres suivant le statut hydrique ou la plasticité de la plante

Les pQTLs peuvent être utilisés pour capturer de l'héritabilité manquante

QUANTITATIF

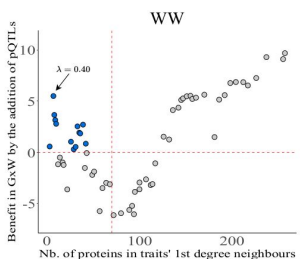
Réseau de dépendance conditionnelle



Pouvons-nous capturer davantage de variance en intégrant les différents réseaux (WW, PL, WD) de manière incrémental ?

7 protéines et 54 pQTLs
(dont 7 localisés à hotspots)

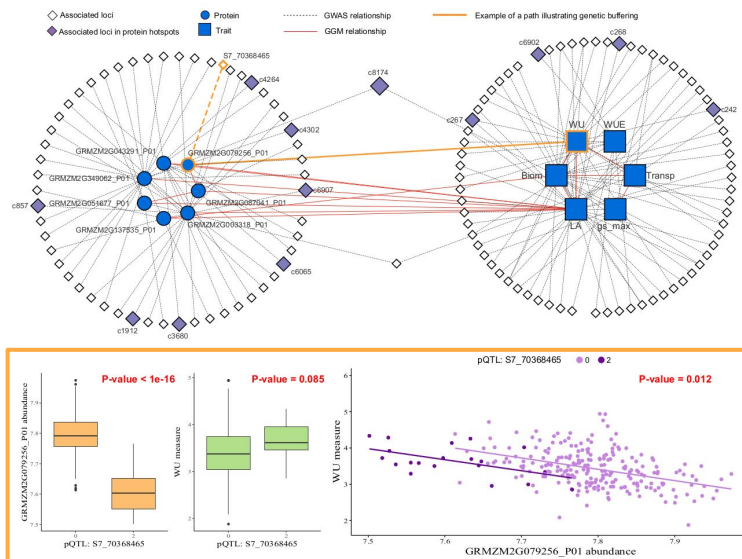
$$M_S : Y \sim \mathcal{N}(X\beta + Q_w + Q_p, Z_u)$$



2 CRITÈRES DE SÉLECTION

- 1) Maximise le gain dans l'explication de la variance GxW
- 2) Le réseau ne doit pas être composé de plus de 70 protéines

Réseau multi-échelle de la condition WW



L'ajout des pQTLs a permis d'expliquer davantage de la variance GxW: 70% avec le réseau WW, et 69% avec le réseau PL. Pas de réseau sélectionné en WD.

Modélisation complète du réseau associé à la réponse du maïs au stress hydrique

QUANTITATIF + FONCTIONNEL

Model	Incremental path	γ^1	# P^2	# pQTLs ³
M_8	WW	70	7	54
M_9	WW \rightarrow PL	78	48	271
M_{10}	WW \rightarrow PL \rightarrow WD	84	23	228
M_{11}	WW \rightarrow WD	71	30	272
M_{12}	WW \rightarrow WD \rightarrow PL	82	11	52
M_{13}	PL	69	45	258
M_{14}	PL \rightarrow WW	70	8	56
M_{15}	PL \rightarrow WW \rightarrow WD	81	23	228
M_{16}	PL \rightarrow WD	NA	NA	NA
M_{17}	PL \rightarrow WD \rightarrow WW	NA	NA	NA

63 protéines et 553 pQTLs et QTLs (dont 49 localisés dans des hotspots)

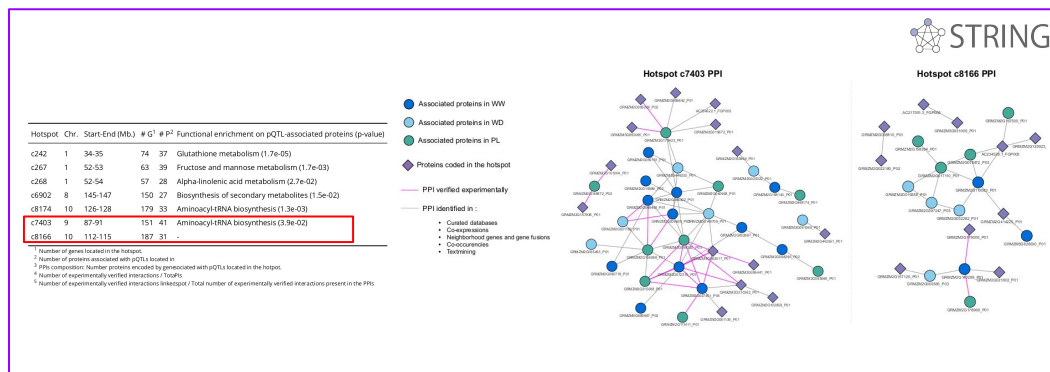
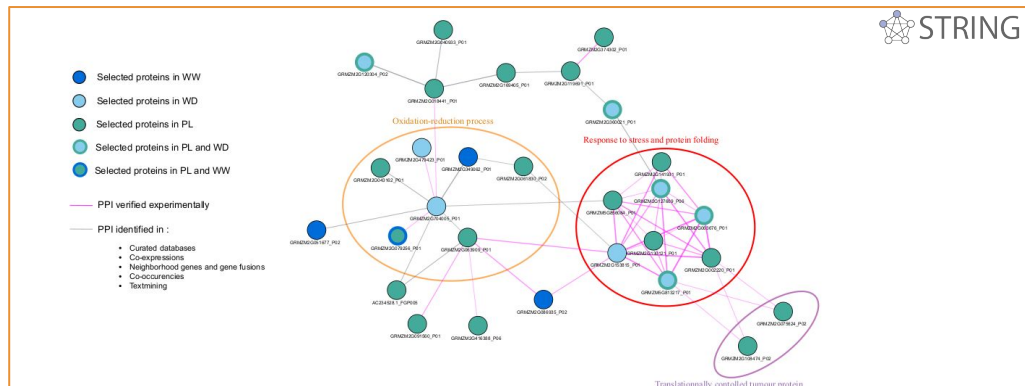
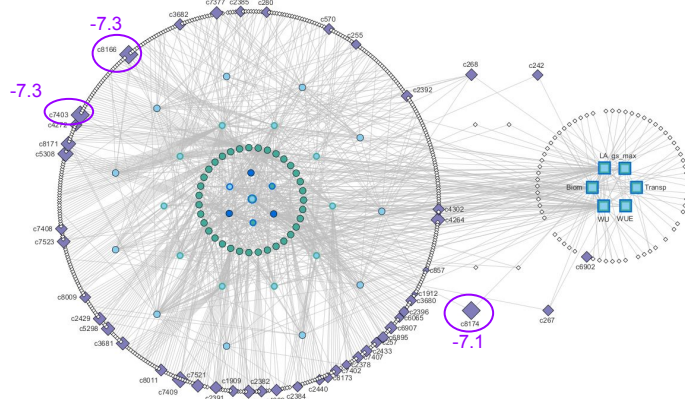


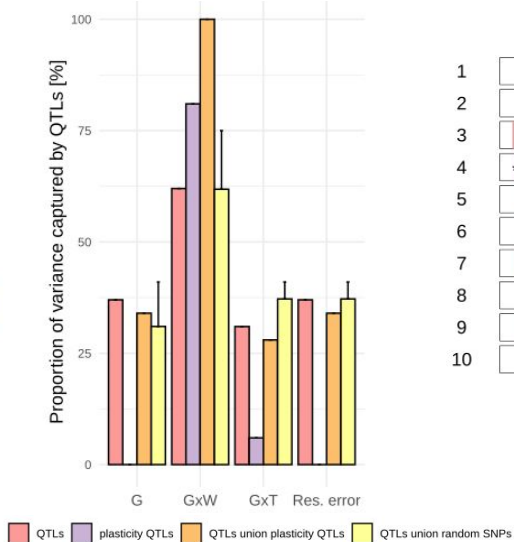
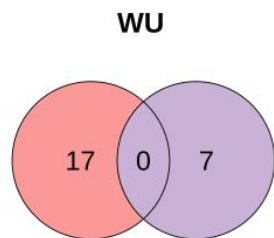
Table des matières

- I. CONTEXTE
 - A. Implication socio-économique de la culture du maïs et des challenges associés
 - B. Détection de QTLs par les approches de génétique d'association
 - C. Les données omiques, un moyen de s'approcher de la complexité de la réponse au stress hydrique
- II. MATERIEL & DONNEES
 - A. Données de génotypage
 - B. Données de phénotypage
 - C. Objectifs de thèse
- III. G x E & PLASTICITE
 - A. Impact de la variabilité inter-essai sur la détection de QTLs impliqués dans la réponse au stress hydrique
 - B. Les GWAS multi-trait améliorent la détection de QTLs de réponse au stress hydrique
 - C. Les QTLs de plasticité sont spécifiquement impliqué dans la variance GxW
- IV. INTÉGRATION DE DONNÉES OMIQUES
 - A. Une étude de génétique des systèmes réalisé sur le même jeu de données
 - B. Un tier de la variance GxW n'est pas expliquée par les QTLs détectés sur les caractères
 - C. Apport fonctionnel des données de protéomique sur le GxW
 - D. Apport quantitative des données de protéomique sur le GxW
 - E. Inférence du réseau multi-échelle maximisant la part de variance capturé sur le GxW
- V. **DISCUSSION**
 - A. Interprétations des résultats obtenues
 - B. Perspectives

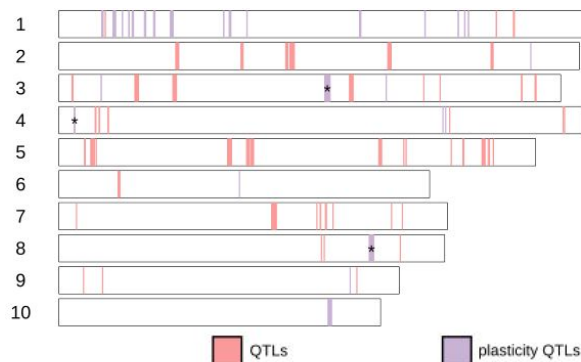
Discussion en lien avec la partie III

- ❖ Améliorer la détection des déterminants génétiques impliqués dans la réponse au stress hydrique en considérant les variations inter-essai et les indices de plasticités.
 - Prise en compte d'un effet "essai" dans les modèles GWAS
 - Décomposition de la variance GxE en Genotype x Disponibilité en eau (GxW) et Genotype x Essai (GxT)
 - Utilisation des indices de plasticité comme variable réponse dans les modèle GWAS

IMPORTANT DE LA PRISE EN COMPTE DES INDICES DE PLASTICITÉS DANS L'EXPLICATION DU GxE



MISE EN VALEUR DU MODÈLE DE RÉGULATION GÉNÉTIQUE EN LIEN AVEC L'EXPRESSION DE LA PLASTICITÉ (Scheiner *et al.* 1993, Via *et al.*, 1995, Kusmec *et al.*, 2017, Diouf *et al.*, 2018)



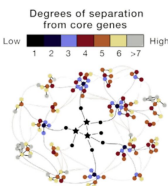
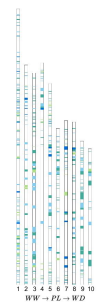
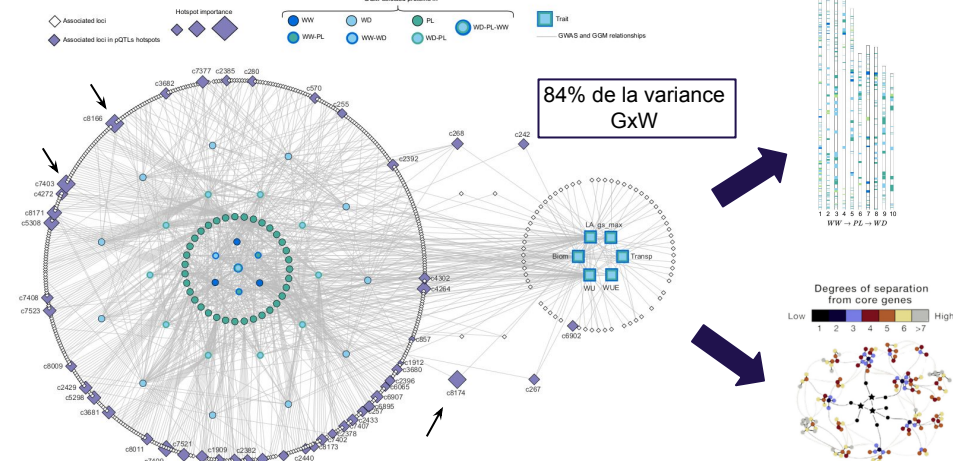
Stage de M1:
Analyses fonctionnelle de loci de caractères quantitatifs (QTLs) associés à la plasticité au stress hydrique chez le maïs

Discussion en lien avec la partie IV

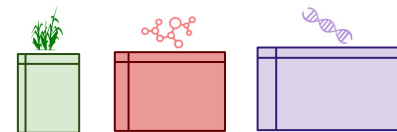
- ❖ Caractérisation génétique et moléculaire approfondie de la réponse du maïs au déficit hydrique en exploitant les données de phénotypique, protéomique et génomique en suivant une approche de génétique des systèmes.
 - Utilisation d'un modèle multi-caractères pour l'identification du sous-ensemble optimal de QTLs expliquant la variance GxW
 - Exploiter les données de protéomique pour fournir des réseaux d'interaction protéines-protéines (PPI) comme support fonctionnel aux QTLs
 - Exploiter les données de protéomique pour apporter de l'information génétique supplémentaire impliqué dans la variance GxW
 - Inférence d'un réseau multi-échelle contenant la totalité des entités biologique d'intérêt et ainsi mettre en valeur des biomarqueurs potentiels

GAIN QUANTITATIF ET FONCTIONNEL DE LA PROTÉOMIQUE

63 protéines et 553 pQTLs et QTLs (dont 49 localisés dans des hotspots)



QUID DE LA MÉTABOLOMIQUE ?

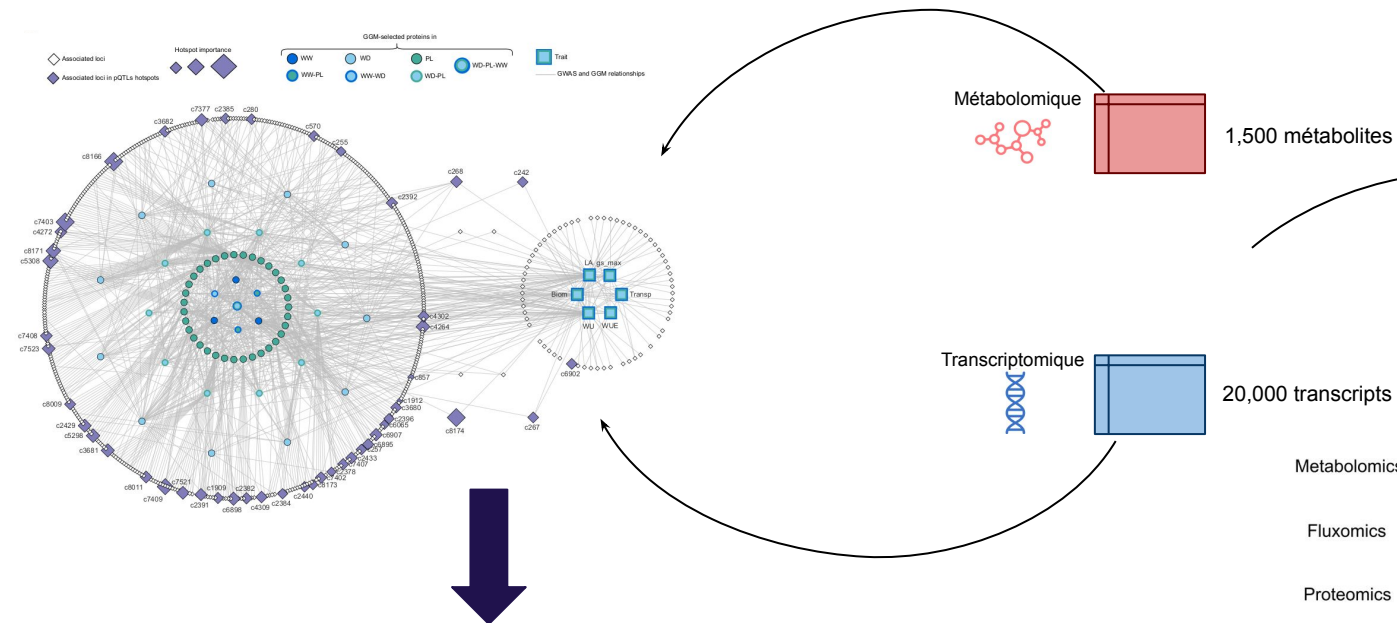


Stage de M1:

Application de la méthode d'inférence de réseaux multi-échelles sur les données de métabolomique

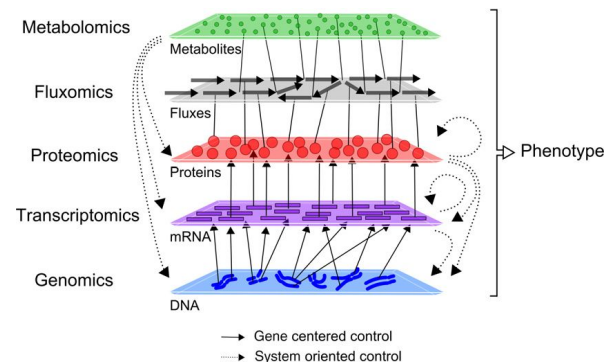
Perspectives

Continuer l'approche d'intégration avec les autres omiques



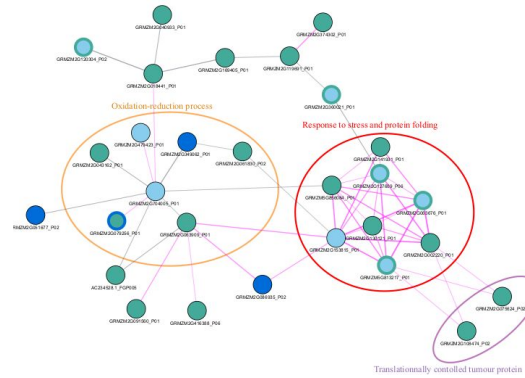
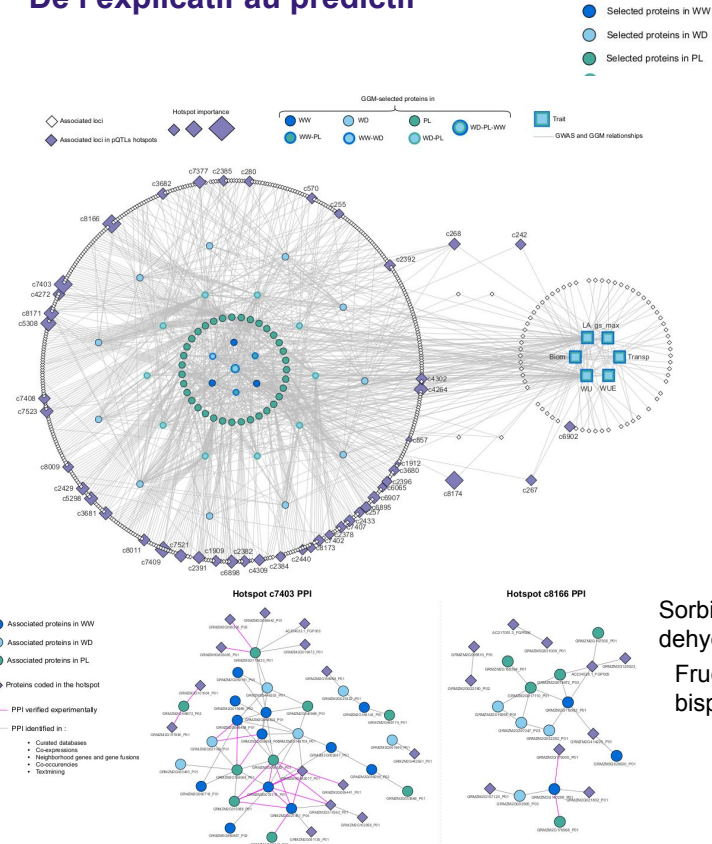
Quel sera le gain en y intégrant ces données ?

Comment adapter l'approche au regard de l'hyperdimensionnalité du jeu de données ?

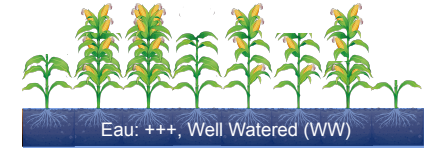
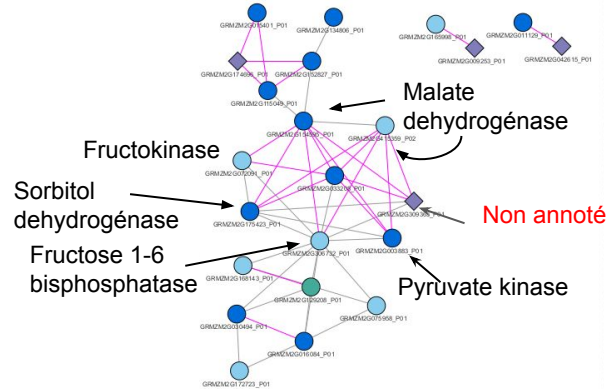


Perspectives

De l'explicatif au prédictif



Hotspot c267 PPI



Modèles de prédictions:
LMMS, NNs

