

# ExIGE : Évaluation multi-environnementale de blé tendre « exotique » - prédictions génomiques GxE et génétique d'association

Justin BLANCON<sup>1</sup>, Jonathan KITT<sup>1</sup>, Pauline LASSERRE-ZUBER<sup>1</sup>, David ALVAREZ<sup>1</sup>, Sibille PERROCHON<sup>1</sup>, Jean-Noël THAUVIN<sup>2</sup>, Celine DUQUE<sup>3</sup>, Sylvie DUTRIEZ<sup>4</sup>, Pascal GIRAUDEAU<sup>5</sup>, Ellen GOUDEMANT-DUGUE<sup>6</sup>, Emmanuel HEUMEZ<sup>7</sup>, Christophe MICHELET<sup>2</sup>, Patrice SENELLART<sup>8</sup>, Alexis OGER<sup>9</sup>, François BALFOURIER<sup>1</sup>, Catherine RAVEL<sup>1</sup>, Sophie BOUCHET<sup>1\*</sup>

1 - INRAE - Université Clermont-Auvergne, UMR 1095, GDEC, 5 chemin de Beaulieu 63100 Clermont-Ferrand, FRANCE

2 - RAGT, Rue Emile Singla, BP 3331 12033 Rodez Cedex 9, FRANCE

3 - Limagrain Europe, Centre de recherche de Chappes, 63720 Chappes, FRANCE

4 - Lidea Seeds, avenue Gaston Phoebus, 64230 Lescar, FRANCE

5 - Secobra Recherches, Centre de Bois-Henry, 78580 Maule, FRANCE

6 - FLORIMOND DESPREZ VEUVE & FILS, 59242 Cappelle-en-Pévèle, FRANCE

7 - INRAE - UE GCIE, 2 chaussée Brunehaut, Estrées Mons, BP50136, 80203 Peronne Cedex, FRANCE

8 - KWS MOMONT SAS, 7 Rue de Martinval, 59246 Mons-en-Pévèle, FRANCE

9 - Syngenta France SA, 1228 Chemin de l'Hobit, 31790 Saint-Sauveur, FRANCE

\*Coordinateur : Sophie BOUCHET, [sophie.bouchet@inrae.fr](mailto:sophie.bouchet@inrae.fr)

## I. Prédictions génomiques du rendement avec des modèles intégrant les interactions GxE

### 1 Introduction

Alors que la compétition pour les terres arables s'accroît, le rendement des principales céréales (blé, maïs, riz) stagne en Europe et dans près de 40 % des surfaces allouées à ces cultures (Ray *et al.* 2012). Parallèlement, l'urgence écologique et climatique conduit à de nouvelles pratiques à faibles niveaux d'intrants, qui diminuent le potentiel de rendement et la standardisation des conditions de cultures. Dans le contexte de croissance démographique et d'instabilité géopolitique actuelle, il est cependant nécessaire d'augmenter la production alimentaire. La sélection de variétés adaptées à ces nouvelles conditions environnementales pourrait donc être un levier efficace pour relever ce défi (Parent *et al.* 2018; Asseng *et al.* 2019; Dias *et al.* 2022).

L'efficacité de la sélection dépend fondamentalement de la diversité disponible (Cobb *et al.* 2019). Cependant, la diversité génétique dans le matériel moderne est limitée (Reif *et al.* 2005; Gerke *et al.* 2015) et il est donc nécessaire d'élargir la base génétique en sélection (McCouch *et al.* 2013). Cependant les ressources génétiques sont peu valorisées en sélection (Glaszmann *et al.* 2010; Wang *et al.* 2017) car il est difficile d'évaluer ce genre de matériel exotique peu adapté (Longin and Reif 2014). De plus, la multiplicité des conditions environnementales constitue un second frein à la sélection, car les géotypes les plus performants dans un environnement ne sont souvent pas les plus performants dans un autre. En effet, l'importance des interactions GxE (géotype-environnement) pour le rendement entraîne un reclassement des géotypes lorsque les conditions environnementales varient. Pour y faire face, les sélectionneurs phénotypent à plusieurs reprises les mêmes géotypes dans des réseaux d'essais multi-environnementaux afin d'estimer une performance globale pour chaque individu. Cependant, cette approche est limitée par des contraintes économiques

et techniques (quantité de semences) et l'évaluation de tous les géotypes dans tous les environnements est impossible.

La sélection génomique peut permettre de lever ce verrou en prédisant la performance d'individus sans avoir besoin de les phénotyper, à partir de leur profil génotypique. En effet, en calibrant un modèle de prédiction génomique grâce à une population d'entraînement, il est possible d'évaluer la valeur d'un individu issu d'une population de sélection rapidement et à moindre coût par rapport à une évaluation au champ. La sélection génomique permet donc d'augmenter le gain génétique par unité de temps et de coût (Bhat *et al.* 2016). Cependant, plutôt que d'estimer une valeur génétique globale, il serait plus efficace d'identifier les individus les plus performants dans chaque environnement. Pour cela, il faut identifier la base génétique et environnementale des interactions GxE pour pouvoir les prédire.

Dans cet objectif, il est nécessaire de disposer de données de géotypage et de phénotypage, mais aussi de descripteurs fins des conditions environnementales. Mais aujourd'hui, alors qu'il est facile et bon marché de géotyper des centaines de géotypes à haut-débit (Elshire *et al.* 2011), et que de nombreuses plateformes et outils de phénotypage haut débit voient le jour, la caractérisation des environnements de sélection reste rare et superficielle. Il est donc crucial d'envirotyper finement les essais de sélection. La caractérisation des conditions pédoclimatiques constitue notamment une source d'information majeure permettant d'expliquer une partie des interactions GxE (Li *et al.* 2018, 2021; Rincint *et al.* 2019; Guo *et al.* 2020; Piepho and Blancon 2023). Les modèles écophysio-logiques permettent quant à eux d'intégrer l'impact de l'ensemble des conditions environnementales mesurées sur les processus physiologiques à la base de la construction du rendement, et génèrent des indicateurs écophysio-logiques particulièrement pertinents pour prédire les interactions GxE (Ly *et al.* 2017).

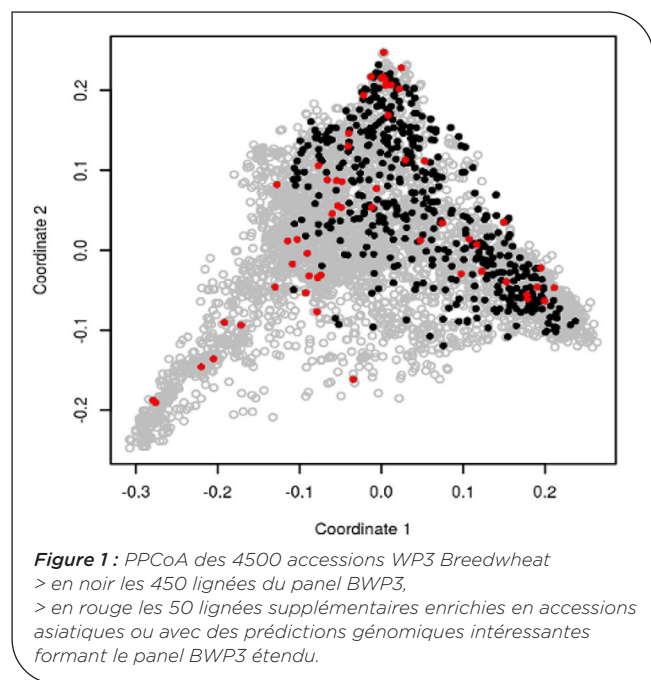


Le projet ExIGE s'appuiera sur des données génotypiques et phénotypiques recueillies sur un panel de diversité large chez le blé tendre d'hiver au cours du Projet Investissement d'Avenir BreedWheat et complètera ce jeu de données grâce à 12 nouveaux essais. Les objectifs de ce projet pour la prédiction du rendement sont de i) Décrire les conditions environnementales à l'échelle du réseau d'essais ii) Calibrer un modèle original de prédiction génomique qui intègre les interactions GxE, iii) Evaluer les performances de ce modèle dans différents scénarios de prédiction, notamment pour de nouveaux génotypes dans de nouveaux environnements.

## 2 Matériel et méthode

### ► Matériel végétal

Les panels étudiés dans le projet ExIGE, nommés BWP3 et BWP3ext, sont constitués de 450 (485) accessions de blé tendre d'hiver de type landraces ou cultivars anciens, mais aussi de variétés modernes (inscription jusqu'en 2004). Ces accessions ont été choisies au sein du centre de ressources biologique « Céréales à pailles » d'INRAE Clermont-Ferrand pour être représentatives de la diversité mondiale (Fig. 1), avec une précocité, une hauteur et une qualité de multiplication adaptées.

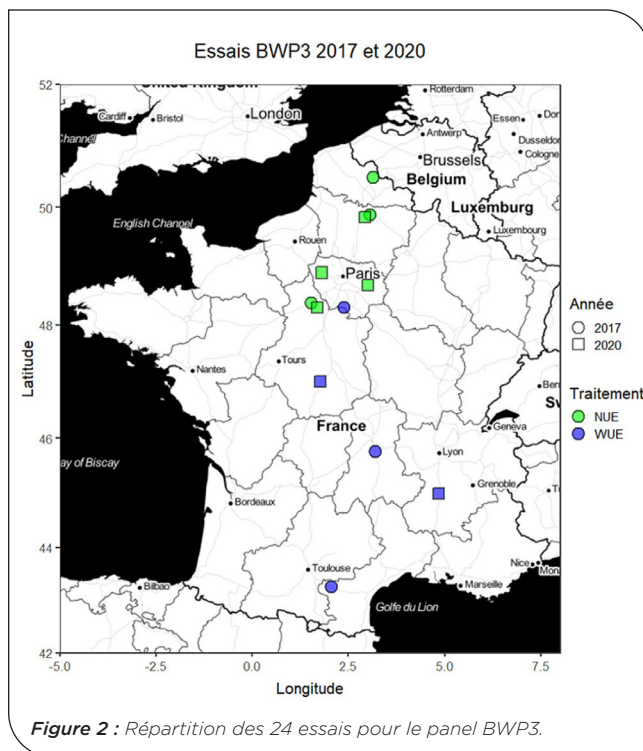


### ► Caractérisation phénotypique du panel

Le rendement en grain ajusté à 15% d'humidité a été mesuré dans 12 nouveaux essais (combinaison lieu x traitement) en 2020 sur le panel BWP3ext. Dans chaque lieu deux traitements ont été appliqués : un traitement pluvial et un traitement irrigué pour Chateauneuf-sur-Isère et Liniez, et un traitement faible azote et fort azote pour Estrées-Mons, Andrezel, Louville et Maule. Les essais ont été semés selon un design p-rep avec 10 à 20% de lignées répétées et 4 témoins. Dans chaque essai seule une moitié représentative du panel a été observée, sauf à Chateauneuf-sur-Isère et Estrées-Mons où l'ensemble du panel a été semé. La densité de

semis et la fertilisation azotée optimum ont été réduites d'environ 30% pour réduire le risque de verse sur ce matériel sensible.

Ces données s'ajoutent aux 12 essais (Fig. 2) réalisés en 2017 sur le panel BWP3 dans le cadre du PIA BreedWheat au sein desquels l'ensemble de ce panel avait été évalué.



### ► Génotypage et matrice d'apparement

L'ensemble du panel BWP3ext a été génotypé sur la puce BreedWheat 410K (Axiom TaBW410k (Rimbert *et al.* 2018 ; Kitt *et al.* 2021). Après élimination des marqueurs présentant plus de 20% de données manquantes ou une fréquence de l'allèle mineur inférieur à 5%, la corrélation entre paires de marqueurs a été analysée de manière à ne conserver que les marqueurs présentant une corrélation inférieure à 0.9. Une matrice d'apparement a ensuite été estimée grâce aux marqueurs restant selon la méthode de VanRaden (2008).

### ► Simulation d'indices environnementaux à l'aide du modèle écophysique CHN

Le développement du couvert a été simulé à l'aide du modèle écophysique CHN (Soenen *et al.* 2016) pour chacun des 24 essais (12 essais du projet ExIGE et des 12 essais du projet BreedWheat) grâce aux données climatiques, à la description du sol et de l'itinéraire technique. Grâce à ces simulations et à la date d'épiaison moyenne du panel, des indices environnementaux décrivant les conditions environnementales perçues par le panel ont été générés. Ces 192 indices, calculés sur différentes périodes de développement (semis-stade épi 1 cm, tallage-épiaison, remplissage-maturité, par exemple) résument soit des variables agroclimatiques (températures, ETP, etc...), soit des variables écophysiques (confort hydrique ou azoté) dans chaque environnement.

	$Y_{ij} = \mu + e_j + G_i + (G \times E_{ij}) + \epsilon_{ij}$
Modèle E+G [1]	-
Modèle E+G+GxE [2]	$G \times E_{ij} - N(0, \sigma_{ge}^2 - K \otimes I_e)$
Modèle E+G+GxW <sub>all</sub> [3]	$G \times E_{ij} - N(0, \sigma_{ge}^2 - K \otimes W_{all})$
Modèle E+G+GxW <sub>selpred</sub> [4]	$G \times E_{ij} - N(0, \sigma_{ge}^2 - K \otimes W_{selpred})$
Modèle E+G+G <sub>AMMI</sub> × W <sub>all</sub> [5]	$G \times E_{ij} - N(0, \sigma_{ge}^2 - K_{AMMI} \otimes W_{all})$
Modèle E+G+GxW <sub>AMMI</sub> [6]	$G \times E_{ij} - N(0, \sigma_{ge}^2 - K \otimes W_{AMMI})$
Modèle E+G+(GxW) <sub>AMMI</sub> [7]	$G \times E_{ij} - N(0, \sigma_{ge}^2 - K_{AMMI} \otimes W_{AMMI})$
Modèle E+G+ (GxW) <sub>AMMIPred</sub> [8]	$G \times E_{ij} - N(0, \sigma_{ge}^2 - K_{AMMIPred} \otimes W_{AMMIPred})$

**Table 1 :** Liste des 8 modèles comparés et modélisation de la variance d'interaction GxE correspondante. Le modèle de référence est le GBLUP (Modèle [1]).

### ► Prédiction génomique des interactions GxE

Un modèle de prédiction génomique prenant en compte les interactions GxE a été calibré sur le panel BWP3 (commun aux deux années) pour le rendement. L'approche utilisée est celle proposée par Rincet *et al.* (2019), qui se base sur une analyse AMMI (Additive Main Effets and Multiplicative Interaction, (Gollob 1968)) pour modéliser les interactions GxE. Cette approche tire parti d'une matrice de kinship pour prédire les scores génétiques de l'AMMI par GBLUP et d'une matrice de covariables environnementales pour prédire les scores environnementaux par régression Partial Least Squares. Ces scores permettent ensuite respectivement l'estimation d'une matrice de covariance d'interactivité génétique et environnementale.

Plusieurs modèles de prédiction génomique ont été comparés dans différents scénarios de cross-validation : un scénario oGoE (old Genotype & old Environment, 5 folds et 10 répétitions), un nGoE (new Genotype & old Environment, 5 folds et 10 répétitions), un oGnE (old Genotype & new Environment, 12 folds et 10 répétitions) et un nGnE (new Genotype & new Environment, 5x12=60 folds et 10 répétitions).

Le modèle de référence est le GBLUP (modèle [1]). Les différents modèles prenant en compte les interactions GxE sont présentés en **Table 1**. Différentes matrices ont été utilisées pour modéliser la variance d'interaction GxE :

- Pour la matrice d'interactivité génétique :
  - K : matrice de Kinship,
  - K<sub>AMMI</sub> : matrice de covariance calculée à partir des 8 premiers vecteurs singuliers gauches de l'AMMI sur le jeu de données entier,
  - K<sub>AMMIPred</sub> : identique à K<sub>AMMI</sub> mais calculée à partir du jeu de données de calibration seulement et prédite pour les nouveaux individus.
- Pour la matrice d'interactivité environnementale :
  - I : matrice identité pour le modèle GxE classique, ne permettant pas de prédire de nouveaux environnement,
  - W<sub>all</sub> : matrice de covariance calculée à partir de l'ensemble des indices environnementaux,
  - W<sub>selpred</sub> : matrice de covariance calculée à partir d'un sous-ensemble d'indices environnementaux, sélectionné pour maximiser la corrélation avec W<sub>AMMI</sub>, sur le jeu de calibration uniquement,

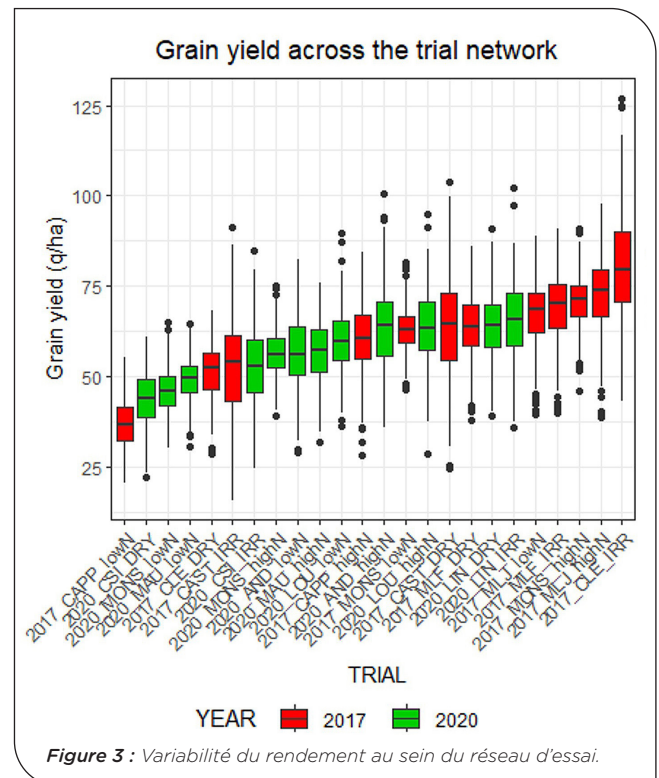
- W<sub>AMMI</sub> : matrice de covariance calculée à partir des 8 premiers vecteurs singuliers droits de l'AMMI sur le jeu de données entier,
- W<sub>AMMIPred</sub> : identique à W<sub>AMMI</sub> mais calculée à partir du jeu de données de calibration seulement et prédite pour les nouveaux environnements.

Dans chaque scénario et pour chaque modèle, la predictive ability a été estimée comme la moyenne du coefficient de corrélation entre les données observées et prédites dans chacun des environnements.

## 3 Résultats

### ► Caractérisation phénotypique du panel

La diversité des conditions environnementales et les différents traitements appliqués au sein du réseau d'essais BreedWheat (année 2017) et ExIGE (année 2020) a permis d'observer une gamme de rendement variant de 37 à 81 q.ha<sup>-1</sup> (**Fig. 3**). La qualité des données collectées est satisfaisante avec une héritabilité généralisée (Cullis *et al.* 2006) entre 0.54 et 0.95.



**Figure 3 :** Variabilité du rendement au sein du réseau d'essai.

### ► Caractérisation environnementale

Les données climatiques, la description du sol et de l'itinéraire technique ont permis d'utiliser le modèle écophysologique CHN pour estimer des indices agroclimatiques sur chaque phase de développement du couvert, ainsi que des cinétiques de confort hydrique et azoté. L'analyse du confort hydrique (Fig. 4) montre que la plupart des essais a subi un stress hydrique nul, faible ou tardif, à l'exception de CLE\_2017\_DRY, MAU\_2020 et étonnamment MONS\_2017 et MONS\_2020, ce qui est malgré tout cohérent avec les observations de l'expérimentateur. Le stress azoté a également été très limité dans le réseau et principalement présent sur les essais MONS\_2017\_lowN et MONS\_2020\_lowN. La classification des environnements sur la base des cinétiques de confort hydrique et/ou azoté n'a pas permis de révéler d'effets clairs sur le rendement, ce qui pourrait s'expliquer soit par l'effet complexe de plusieurs facteurs limitants au sein du réseau, soit par une imprécision de l'estimation des indices environnementaux.

### ► Analyse des interactions GxE

Les données de rendement issues des 24 essais considérés ont fait l'objet d'une analyse conjointe des interactions GxE. Les corrélations entre environnements sont toutes positives ( $r$  de 0,13 à 0,88, Fig. 5), avec un  $r^2$  d'environ 0,25 en moyenne. L'analyse des corrélations semble montrer des interactions GxE prépondérantes mais peu structurées. Ceci est confirmé par l'analyse de la variance (Table 2, méthode proposée par Smith *et al.* (2001)) avec une variance GxE près de deux fois plus importante que la variance génétique, et du même ordre que la variance environnementale. De même, une classification des environnements sur la base de la matrice GxE n'explique que 10% de la variance GxE, ce qui va dans le sens d'un réseau peu structuré globalement.

	Variance	Erreur standard
Génotype	136.52	10.22068
Environnement	321.3613	140.98888
GxE	300.9044	4.73078

Table 2 : Analyse de variance du rendement dans le réseau de 24 essais.

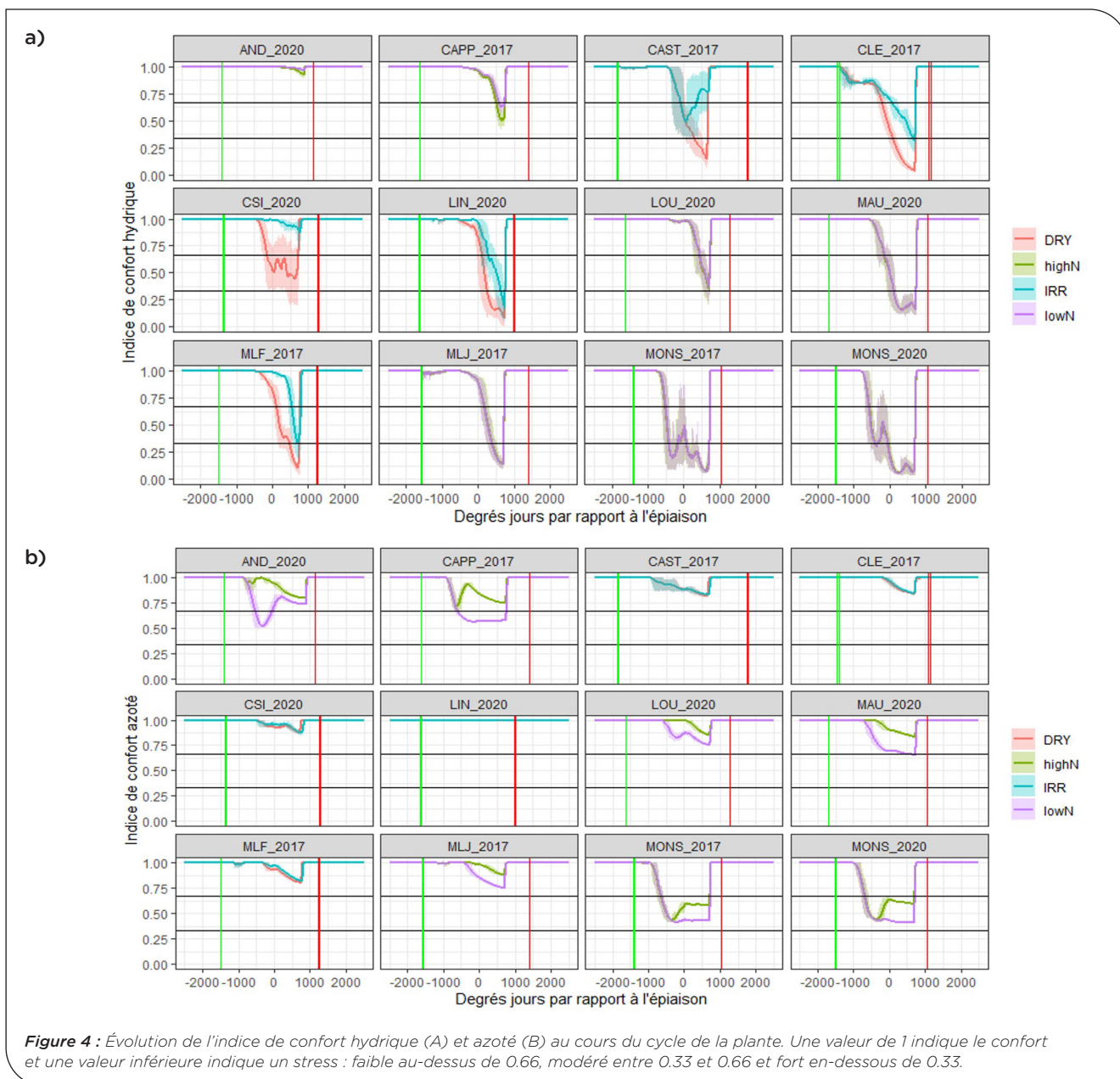
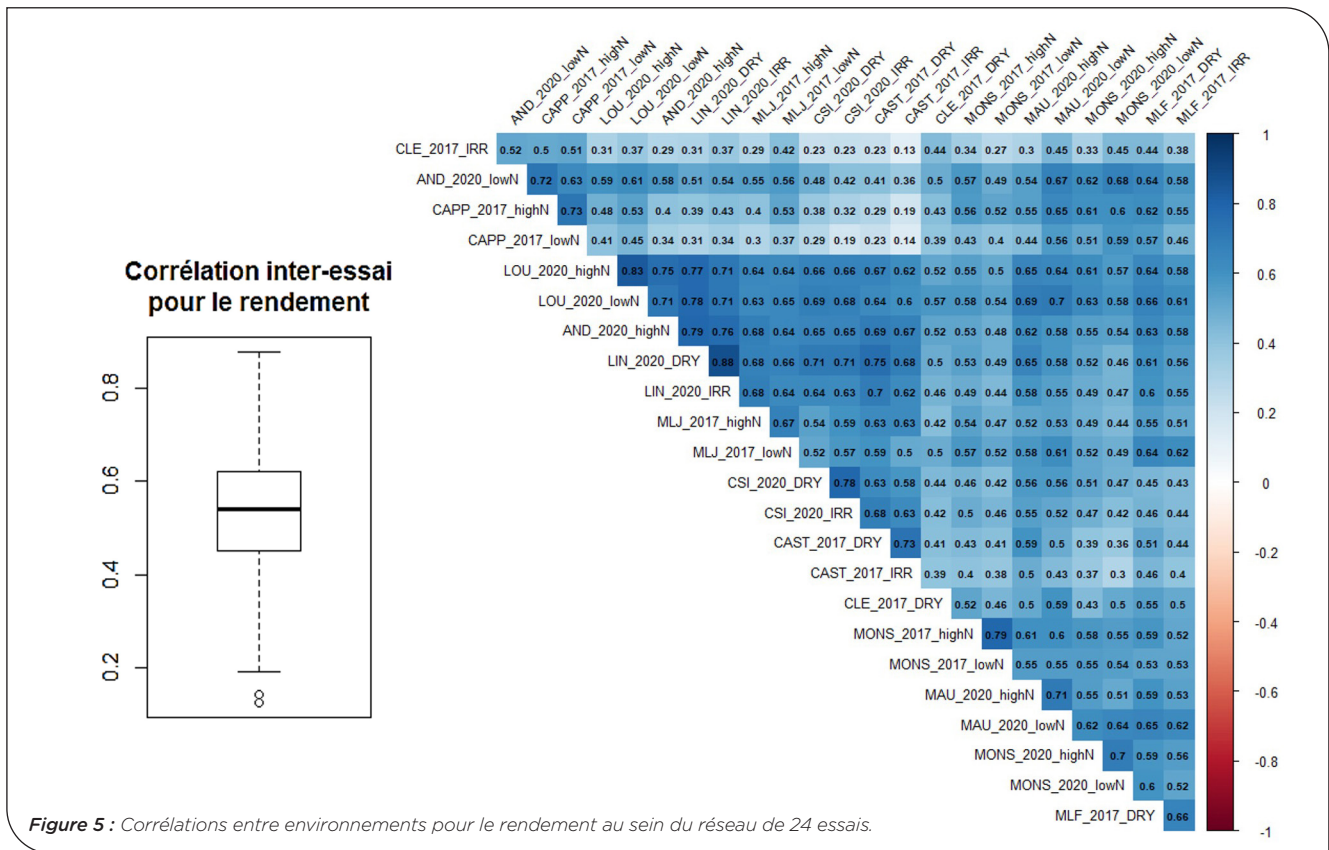


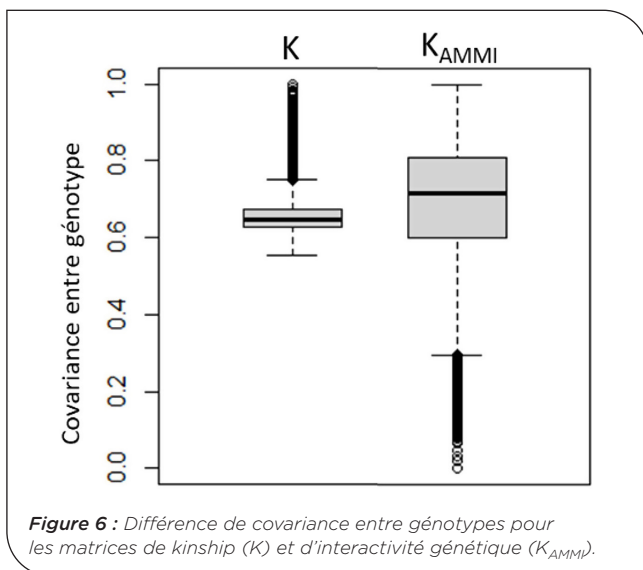
Figure 4 : Évolution de l'indice de confort hydrique (A) et azoté (B) au cours du cycle de la plante. Une valeur de 1 indique le confort et une valeur inférieure indique un stress : faible au-dessus de 0.66, modéré entre 0.33 et 0.66 et fort en-dessous de 0.33.



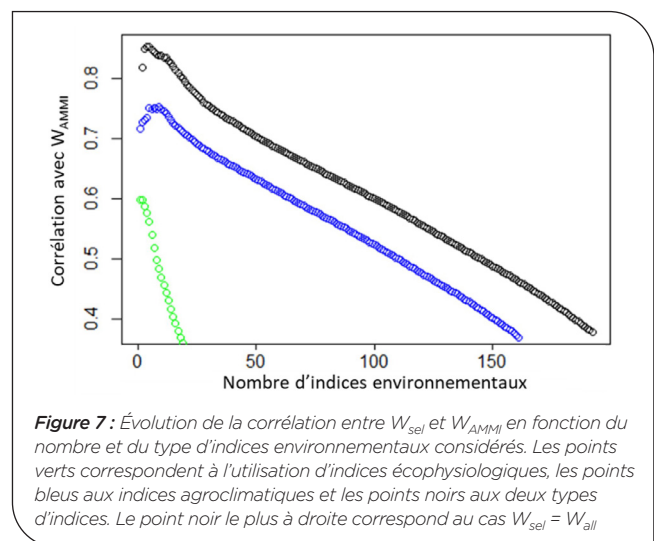
► **Prédiction génomique GxE**

**Analyse AMMI**

Les 8 premiers axes de l'AMMI captent 73% de la variance résiduelle dans l'effet d'interaction GxE. Les 8 premières colonnes de la matrice des scores environnementaux permettent de calculer la matrice de covariance de l'interactivité environnementale  $W_{AMMI}$  et celles de la matrice des scores génétiques permettent de calculer la matrice de covariance de l'interactivité génétique  $K_{AMMI}$ . Une comparaison entre la matrice de kinship et la matrice  $K_{AMMI}$  montre que l'information portée par  $K_{AMMI}$  est différente (Fig. 6), et qu'il semble pertinent de tirer parti de cette information lors de la modélisation des interaction GxE dans le modèle de prédiction génomique.



La corrélation entre la matrice d'interactivité environnementale et une matrice de covariance environnementale calculée à partir de tous les indices environnementaux disponibles ( $W_{all}$ ) est également faible ( $r=0.38$ , Fig. 7). Une sélection séquentielle d'un sous-ensemble de variables environnementales permet d'améliorer la corrélation entre la matrice de covariance ainsi construite ( $W_{sel}$ ) et  $W_{AMMI}$  pour atteindre  $r=0.83$  avec 6 indices dont 5 indices agroclimatiques et 1 indice écophysiological. Les indices écophysiological apportent donc une information complémentaire à celle des indices agroclimatiques. Un point important à noter, qui diffère de l'étude de Rincet *et al.* (2019), est que les indices agroclimatiques sont plus informatifs que les indices écophysiological. Cela pourrait indiquer une moins bonne qualité de la modélisation écophysiological dans ce projet.



## Prédiction génomique de l'AMMI

Les résultats de la cross-validation montrent une predictive ability allant de 0.46 à 0.67 pour le GBLUP (modèle [1], **Table 3**). La prise en compte des interactions GxE permet une amélioration de la predictive ability dans tous les scénarios. Le modèle [3] permet une amélioration seulement dans les scénarios oGoE et nGoE, un peu plus importante que celle du modèle [2]. Le modèle [4] en revanche présente de moins bonnes performances malgré une plus forte corrélation avec  $W_{AMMI}$  de  $W_{selpred}$  comparé à  $W_{all}$ . Ceci montre que la sélection du sous-ensemble de covariables environnementales n'est pas robuste. Le modèle [7], basé sur les matrices d'interactivité de l'AMMI sur le jeu de données complet, est le plus performant. Il permet une nette amélioration de la predictive ability avec un gain de  $r$  de +0.13 à +0.28 suivant les scénarios. Ce modèle ne permet pas de prédiction en pratique puisqu'il utilise l'ensemble du jeu de calibration et de validation pour ajuster le modèle, cependant il permet d'estimer la borne supérieure de performance atteignable avec ce type d'approche. De même les modèles [5] et [6] ne permettent pas de prédiction dans la pratique, mais confirment respectivement que la matrice d'interactivité génétique issue de l'AMMI est pertinente pour améliorer la qualité des prédictions dans les scénarios avec de nouveaux génotypes, et la matrice d'interactivité environnementale dans ceux avec de nouveaux environnements. Le modèle [8], basé sur la prédiction de ces deux matrices d'interactivité présente des performances similaires au modèle [3], les matrices prédites ne sont donc pas plus informatives que la matrice de kinship et que la matrice  $W_{all}$ .

Modèles	Scénarios de Cross-Validation			
	oGoE	nGoE	oGnE	nGnE
E+G [1]	0.671	0.461	0.656	0.439
E+G+GxE [2]	0.726	0.522	-	-
E+G+Gx $W_{all}$ [3]	0.753	0.532	0.650	0.435
E+G+Gx $W_{selpred}$ [4]	0.750	0.514	0.626	0.396
E+G+G $_{AMMI}$ x $W_{all}$ [5]	0.871	0.740	0.642	0.481
E+G+Gx $W_{AMMI}$ [6]	0.762	0.501	0.732	0.487
E+G+(GxW) $_{AMMI}$ [7]	0.871	0.740	0.788	0.662
E+G+(GXW) $_{AMMipred}$ [8]	0.753	0.530	0.644	0.432

**Table 3 :** Qualité de prédiction pour les 8 modèles testés dans 4 scénarios de cross-validation.

## 4 Discussion

Le projet ExIGE a permis l'évaluation d'une large diversité de blé tendre d'hiver au champ dans des conditions de stress abiotiques variées. La caractérisation de ces conditions environnementales à l'aide d'un modèle écophysologique a nécessité la collecte de données météorologiques, de composition

du sol ainsi que le suivi précis des itinéraires techniques appliqués. L'expérience acquise lors du projet montre que cette description des conditions d'expérimentation est fastidieuse et souvent lacunaire. Malgré ces difficultés nous avons pu décrire plus finement les conditions environnementales à l'aide du modèle écophysologique CHN ce qui nous a permis une meilleure compréhension du réseau expérimental.

L'approche utilisée pour prédire le rendement dans chaque environnement n'a pas permis un gain de performance important par rapport à l'approche standard (GBLUP). Nos résultats sont similaires à ceux de Rincet *et al.* (2019) pour  $W_{AMMipred}$  mais pas pour  $K_{AMMipred}$ , pour laquelle ils avaient obtenu un gain par rapport à K. Nous explorons cependant une diversité génétique beaucoup plus large car ils ont travaillé sur un panel élite de 156 variétés alors que nous avons utilisé un panel de diversité de 450 individus, ce qui est peut-être la raison de la difficulté à prédire  $K_{AMMI}$ . Il apparaît d'ailleurs plus difficile de prédire de nouveaux génotypes que de nouveaux environnements avec tous les modèles, sauf le modèle [5] qui tire parti de la matrice GAMMI mais pas de  $W_{AMMI}$ . De même, ce sont dans les scénarios nGoE et nGnE que les gains de prédiction sont les plus importants avec le modèle [7], ce qui montre l'intérêt majeur d'améliorer l'estimation de  $K_{AMMipred}$ . Les scores génétiques permettant le calcul de  $K_{AMMipred}$  ont été prédits par GBLUP pour les nouveaux génotypes, l'utilisation de modèles bayésiens supposant une architecture génétique plus complexe comme le BayesA ou BayesB permettrait peut-être une meilleure estimation. En effet, des résultats intermédiaires semblent montrer la prédominance d'un nombre restreint de déterminants génétiques pour la prédiction de certains scores génétiques. Concernant  $W_{AMMipred}$ , il est possible que la composante environnementale de l'interactivité ne soit pas suffisamment forte dans notre réseau d'essai pour quelle puisse être précisément estimée, avec notamment une faible structuration du réseau d'essai. Il est aussi possible que la qualité de la modélisation écophysologique permettant l'estimation des indices environnementaux ne soit pas suffisante. Enfin, la taille du réseau d'essai et sa diversité peuvent également être limitantes avec un réseau composé de seulement 24 essais répartis sur 2 années et 11 lieux majoritairement au Nord de la France.

La modélisation des interactions GxE en prédiction génomique par l'approche AMMI semble prometteuse (modèle [7]), bien qu'au cours de ce projet, la précision de prédiction des matrices d'interactivité génétique et environnementale n'a pas été suffisante pour permettre un gain de predictive ability notable. Dans le futur, l'augmentation de la taille et de la diversité du jeu de calibration environnementale, une meilleure qualité de modélisation écophysologique, et une meilleure prise en compte de l'architecture génétique de l'interactivité pourrait cependant permettre d'améliorer les capacités de prédictions du rendement. De même, la méthode utilisée dans ce travail consiste en de multiples étapes de prédictions qui entraînent des pertes d'information successives, une méthode single-step permettrait peut-être de mieux exploiter le potentiel prédictif de ce genre d'approche.

## II. Études d'association pour les caractères agronomiques

### 1 Introduction

Le changement climatique conduit à une stagnation des rendements du blé tendre en Europe (Brisson *et al.* 2010). Pour augmenter la production alimentaire, il est nécessaire d'optimiser l'efficacité des schémas de sélection. L'apport d'une nouvelle diversité pourrait permettre d'augmenter le gain génétique et ainsi de limiter l'effet négatif des aléas environnementaux dus au changement climatique.

Dans ce cadre, le réservoir constitué par les allèles originaux présents dans les ressources génétiques est une source de diversité susceptible de permettre de relever ces défis. On peut en effet intégrer rapidement une nouvelle diversité dans le matériel élite par sélection assistée par marqueurs. Pour cela, il est nécessaire d'identifier de nouveaux allèles favorables pour les caractères agronomiques majeurs (rendement et composantes, précocité, hauteur, teneur en protéines) dans les ressources génétiques.

L'objectif de ce projet est d'identifier de nouveaux allèles favorables dans des ressources exotiques, absents du matériel élite. Nous utiliserons une approche de génétique d'association afin de détecter les déterminants génétiques des caractères agronomiques majeurs au sein d'un panel de 485 lignées représentatives de la diversité génétique mondiale.

### 2 Matériel et méthode

#### ► Caractérisation phénotypique

Le réseau d'essai est celui décrit dans la partie « I. Prédiction génomique du rendement avec des modèles intégrant les interactions GxE ». Différents caractères agronomiques ont été mesurés dans chaque essai (**Table 1**) sur le panel BWP3 en 2017 (projet BreedWheat) et sur le panel BWP3ext en 2020 (projet ExIGE) : la date d'épiaison (HD), la hauteur de plante (PH), le rendement ajusté à 15% d'humidité (GY15) Le nombre d'épis par m<sup>2</sup> (SA), le nombre de grains par épi (GPS), le nombre de grains par m<sup>2</sup> (GPA), le poids de mille grains ajusté à 15% d'humidité (TKW15), la surface, la longueur et la largeur d'un grain (GA, GL, GW), la teneur en protéine (GPC) et l'écart à la régression rendement protéine (GPD).

#### ► Génétique d'association

Pour identifier les déterminants génétiques associés à chaque caractères, l'effet de chaque marqueur de la puce TaBW410K a été testé dans chaque environnement à l'aide du modèle K (Yu *et al.* 2006) selon la méthode LOCO (Rincent *et al.* 2014).

Les marqueurs présentant une p-value inférieure à 10<sup>-6</sup> ont été déclarés significatifs puis agrégés en QTL pour chaque caractère au sein de chaque essai sur la base du déséquilibre de liaison, d'après l'approche de Cormier *et al.* (2014). Chaque QTL a été décrit par son marqueur le plus significatif et sa part de variance expliquée estimée par la méthode de Sun *et al.* (2010).

	HD	PH	GY15	SA	GPS	GPA	TKW15	GA	GL	GW	GPC	GPD
2017_CAPP_highN	448	447	452	441	452	451	0	0	0	0	449	445
2017_CAPP_lowN	450	450	449	442	451	449	0	0	0	0	448	450
2017_CAST_DRY	454	452	453	445	448	450	452	0	0	0	450	451
2017_CAST_IRR	454	453	454	448	451	452	452	0	0	0	449	448
2017_CLE_DRY	310	310	310	306	303	307	306	302	300	304	311	311
2017_CLE_IRR	309	312	308	240	239	306	304	300	300	306	309	306
2017_MLF_DRY	450	450	447	0	0	446	444	444	441	439	443	446
2017_MLF_IRR	450	421	442	0	0	445	447	445	443	439	438	446
2017_MLJ_highN	449	445	446	449	450	446	450	0	0	0	447	445
2017_MLJ_lowN	449	445	449	444	450	448	448	0	0	0	444	450
2017_MONS_highN	454	452	451	438	449	450	450	0	0	0	450	452
2017_MONS_lowN	454	452	451	440	449	451	452	0	0	0	450	451
2020_AND_highN	275	274	277	0	0	0	277	0	0	0	271	273
2020_AND_lowN	275	274	277	0	0	0	276	0	0	0	274	275
2020_CSI_DRY	242	239	480	0	0	0	478	0	0	0	479	477
2020_CSI_IRR	239	238	478	0	0	0	475	0	0	0	472	474
2020_LIN_DRY	239	241	240	0	0	0	238	0	0	0	240	240
2020_LIN_IRR	239	241	238	0	0	0	240	0	0	0	240	237
2020_LOU_highN	263	262	258	0	0	0	256	0	0	0	241	243
2020_LOU_lowN	263	261	245	0	0	0	247	0	0	0	243	242
2020_MAU_highN	261	261	261	0	0	0	263	0	0	0	259	260
2020_MAU_lowN	238	238	233	0	0	0	239	0	0	0	238	239
2020_MONS_highN	485	480	480	0	0	0	479	0	0	0	476	478
2020_MONS_lowN	484	475	483	0	0	0	480	0	0	0	479	482

Table 1 : Effectif phénotypé pour chaque caractère dans chaque essai du réseau.

### 3 Résultats

Au total, 272 QTL ont été détectés, répartis sur l'ensemble des chromosomes (Fig. 1). Des QTL ont été détectés pour chacun des 12 caractères analysés, avec un maximum pour HD et PH (75 et 90 QTL respectivement) et un minimum pour les caractères de taille de grains GW, GL et GA (3, 5 et 7 QTL). Pour le rendement, 9 QTL ont été identifiés.

#### ► Colocalisation avec PPD-D1

Sur le chromosome 2D (Fig. 2), on observe une colocalisation très importante de QTL, principalement pour HD (65 QTL), dont l'emplacement correspond au gène de précocité PPD-D1. Ces QTL de précocité sont détectés dans les 24 essais. Ces associations sont particulièrement fortes, avec un  $-\log_{10}(p\text{-value})$  et un PEV (Proportion of Explained Variance) moyen (maximum) de 18 (48) et 20% (43%) respectivement. L'effet moyen associé à l'allèle minoritaire est de 39 degrés-jour.

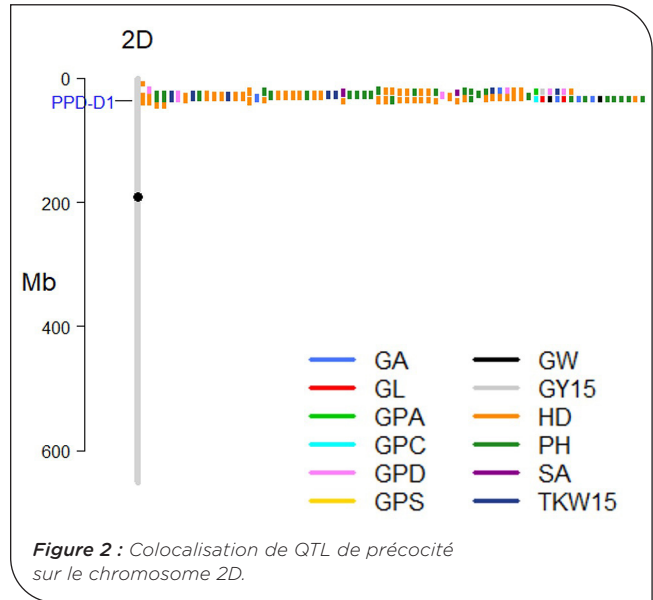


Figure 2 : Colocalisation de QTL de précocité sur le chromosome 2D.

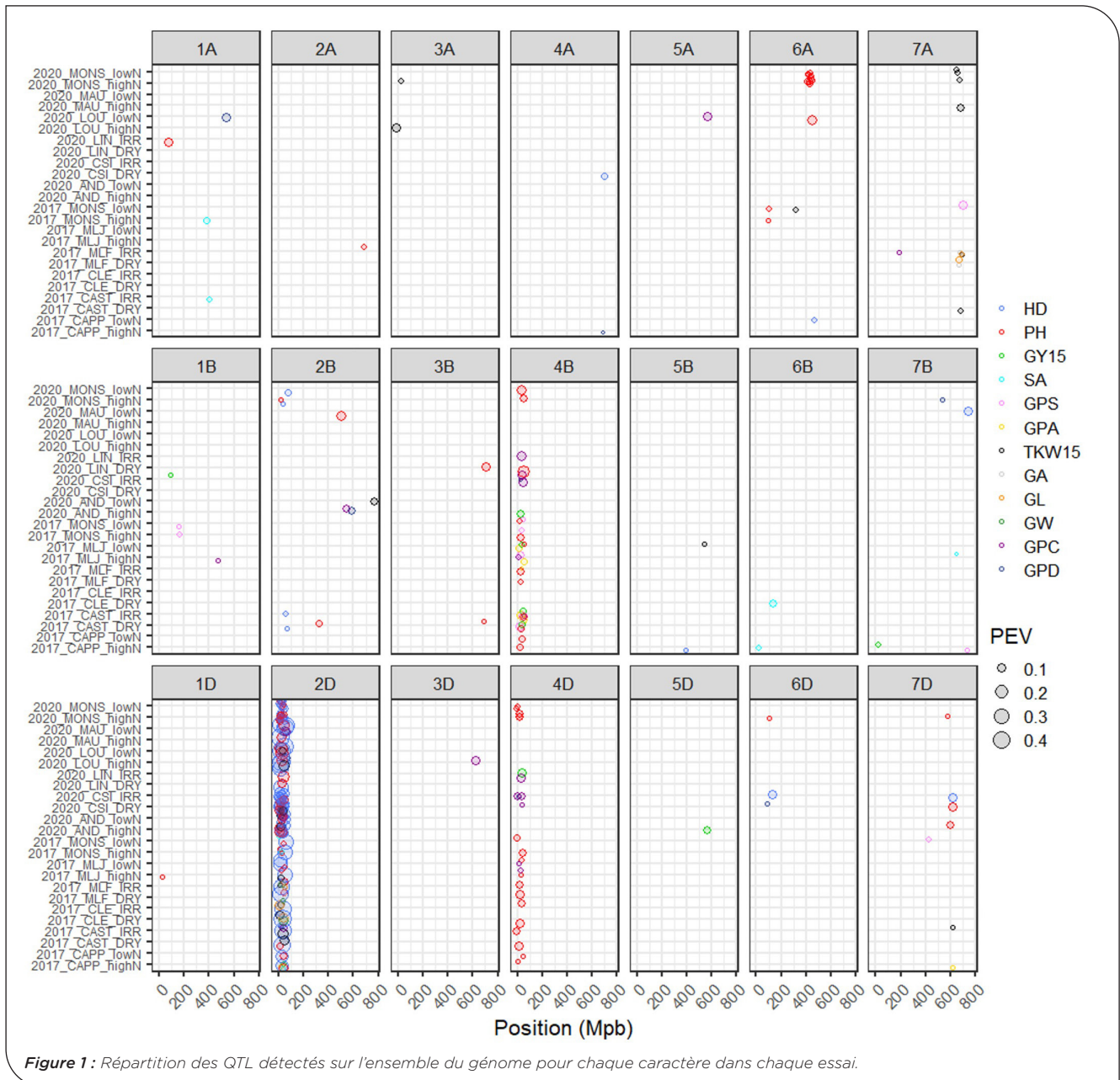
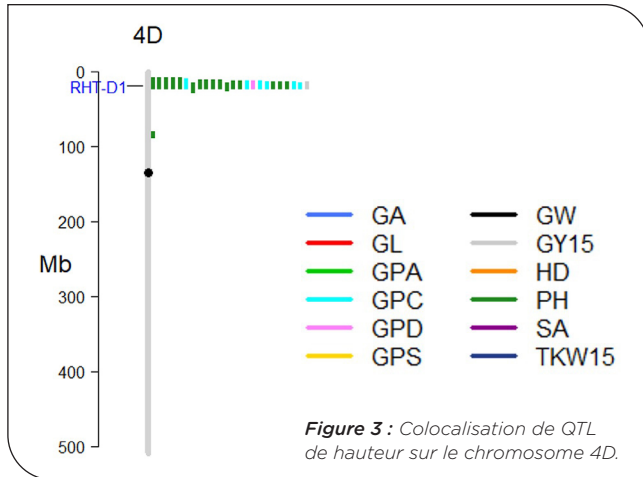


Figure 1 : Répartition des QTL détectés sur l'ensemble du génome pour chaque caractère dans chaque essai.



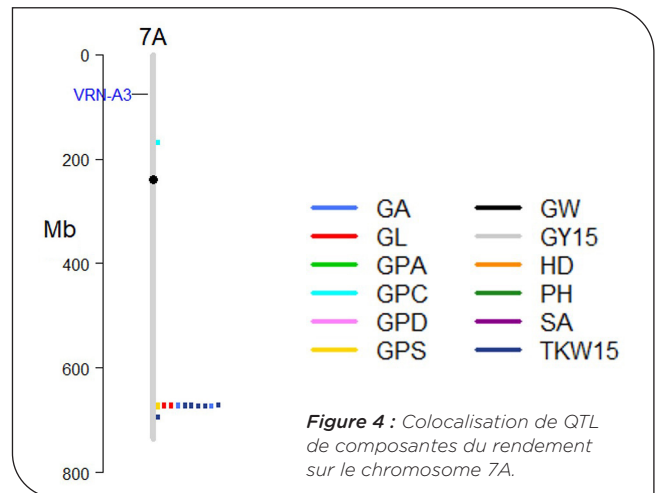
### ► Colocalisation avec RHT-D1

On observe également une importante colocalisation de QTL avec le gène de hauteur RHT-D1 sur le chromosome 4D (**Fig. 3**). Ce pic regroupe 16 QTL de hauteur, ainsi que 6 QTL de GPC, 1 QTL de GPD et 1 QTL de rendement. Pour la hauteur, la moyenne et le maximum sont de 8.5 et 11.7 pour  $-\log_{10}(p\text{-value})$  et de 7.8% et 10.8% pour le PEV, avec un effet moyen de 31 cm pour l'allèle minoritaire.



### ► Colocalisation de QTL de composantes du rendement

Sur le chromosome 7A, on observe une colocalisation de QTL de composantes du rendement, sans QTL liés à la précocité, à la hauteur ou à la teneur en protéines (**Fig. 4**). Ce pic est composé de 5 QTL de TKW15, de 2 QTL de longueur de grain, de 2 QTL de surface de grain et d'un QTL de nombre de grains par épi. Ces QTL ont été détectés dans 5 essais différents, regroupant 3 lieux, 2 années et les 4 traitements expérimentés (highN, lowN, IRR et DRY), et expliquent en moyenne 6.7% de la variance. L'allèle minoritaire diminue le nombre de grains et augmente leur taille et leur poids.



## 4 Discussion

L'objectif était d'identifier de nouveaux allèles favorables dans ce matériel exotique pouvant être intéressants à introduire dans le matériel en sélection. L'analyse de génétique d'association réalisée est de bonne qualité car on retrouve d'importante colocalisation sur des gènes connus pour la précocité et la hauteur de plante. D'autres QTL pourrait se révéler intéressants et mériteraient d'être analysés plus en détails, comme les QTL de rendement, de ses composantes, ou de teneur en protéines. La région la plus prometteuse semble être la colocalisation de QTL de composantes de rendement sur le chromosome 7A. Il serait intéressant d'étudier les allèles présents à ce locus dans le matériel élite pour déterminer s'il existe un nouvel allèle dans le matériel exotique et si oui, déterminer quel est son effet dans un fond élite.

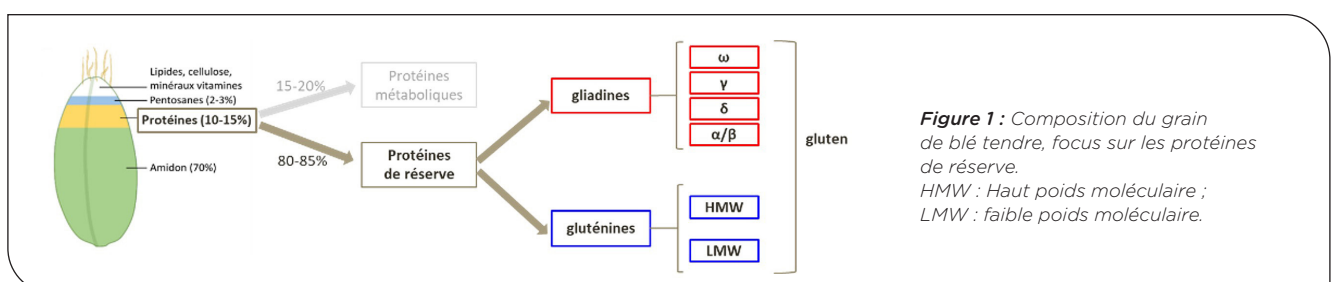
Enfin, à l'issu du projet FSOV Débat, de nouvelles données de génotypage de type présence-absence, issu de RNAseq, sont disponibles pour ce panel. Cette description différente de la diversité pourrait permettre de découvrir de nouvelles zones associées.

## III. Séquençage gluténines et qualité boulangère

### 1 Introduction

Le blé est à la base de l'alimentation humaine. Consommé généralement sous forme de produits transformés et essentiellement sous forme de pain pour le blé tendre (*Triticum aestivum*), il couvre environ 1/5 de nos besoins en énergie et en protéines. Les principaux constituant du grain de blé sont l'amidon (80-85%) et les protéines (15-20%). Les protéines de

réserve (ou protéines du gluten) sont majoritaires et représentent 80-85% des protéines totales du grain. Elles sont constituées par les gluténines et les gliadines, ces deux fractions étant subdivisées en plusieurs sous-classes (**Figure 1**). On distingue les gluténines de hauts et faibles poids moléculaires, respectivement HMWG qui représentent environ 10% des protéines de réserve, et LMWG qui elles représentent 1/3 des protéines du grain et environ 60% des gluténines totales. Les gliadines comprennent les  $\alpha/\beta$ ,  $\gamma$ - et  $\omega$ -gliadines.



Après hydratation et pétrissage de la farine, les protéines de réserve s'agrègent en un réseau tridimensionnel, le gluten. Rapidement, les gluténines se lient par des ponts disulfures et forment le squelette du gluten sur lequel se fixent par des liaisons faibles les gliadines (Johansson *et al.* 2013). Le gluten est à l'origine des propriétés de viscoélasticité des pâtes. Les gluténines apportent la ténacité, les HMWG contrôlent l'élasticité (Johansson *et al.* 2013), les gliadines sont responsables de la viscosité. La qualité boulangère résulte elle d'un équilibre entre viscosité, élasticité et ténacité.

Le déterminisme génétique des protéines de réserve est aujourd'hui bien établi. Les gènes codant pour ces protéines font partie de familles multigéniques organisées en clusters et s'héritant en blocs. Les HMWG sont codées par trois *loci* homéologues (*Glu-A1*, *Glu-B1* et *Glu-D1*) situés sur le bras long des chromosomes du groupe 1 (Payne 1987). Chaque locus comprend 2 gènes liés (*Glu-A1-1* et *Glu-A1-2* au locus *Glu-A1* par exemple) qui codent pour les sous-unités x et y de la protéine. (<http://www.aaccnet.org/initiatives/definitions/Pages/Gluten.aspx>). Les clusters de LMWG et de gliadines comprennent de nombreux gènes. Les *loci* de LMW, *Glu-A3*, *Glu-B3* et *Glu-D3*, sont situés sur le bras court des chromosomes 1 ; les gliadines sont-elles codées par des clusters sur les bras courts des chromosomes des groupes 1 et 6.

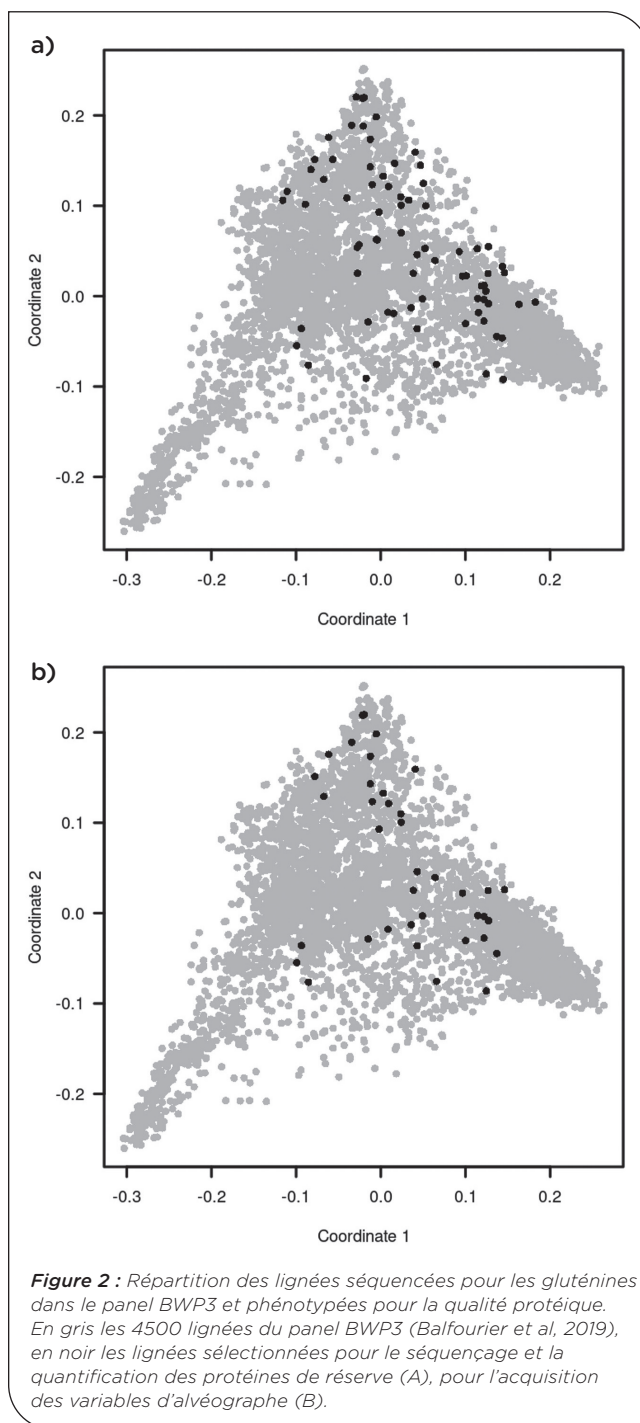
Les protéines de réserve sont connues pour leur niveau important de polymorphisme. Les allèles de gluténines ont été classés d'après leur effet sur la qualité boulangère (Branlard and Dardevet 1985; Payne *et al.* 1987). L'allèle d de *GluD1-1* qui donne l'haplotype protéique Dx5 + Dy10 et l'allèle b de *Glu-B1* (haplotype protéique Bx7 + Bx8) ont été associés à une très bonne qualité boulangère (Pirozi *et al.* 2008). C'est ce qui explique la prédominance de l'haplotype Dx5 + Dy10 observée dans les variétés élites qui expliquerait leur forte ténacité et leur manque d'élasticité que le meunier doit corriger par des mélanges avec des blés améliorants élastiques.

Disposer d'outils facilitant l'amélioration de la qualité boulangère est un enjeu majeur pour le sélectionneur. Développer de tels outils implique de déterminer le polymorphisme des gluténines, principales responsables de la qualité boulangère, au niveau du génome, puis de mettre au point des marqueurs moléculaire associés aux qualités de la farine. Cependant une difficulté de ce travail provient de la structure des gènes de protéines de réserve qui sont riches en motifs répétés que l'on retrouve non seulement au niveau d'un même gène mais aussi entre les différents gènes. Cette particularité explique la faible qualité de l'assemblage des génomes dans ces régions, expliquant que les gènes de protéines de réserve soient incomplets sur la pseudo-molécule de Chinese-Spring V2.1. Un séquençage long-read de ces régions pourraient corriger cette situation.

## 2 Matériel et méthode

### ► Matériel Végétal

Nous avons échantillonné 73 lignées représentatives du panel BWP3 (Figure 2A) pour séquencer les gènes de protéines de réserve. Certains gènes de gluténines pour ces 73 lignées ont été séquencés.



**Figure 2 :** Répartition des lignées séquencées pour les gluténines dans le panel BWP3 et phénotypées pour la qualité protéique. En gris les 4500 lignées du panel BWP3 (Balfourier *et al.*, 2019), en noir les lignées sélectionnées pour le séquençage et la quantification des protéines de réserve (A), pour l'acquisition des variables d'alvéographe (B).

### ► Phénotypage

Une mouture complète a été obtenue à partir des grains récoltés à pour chaque condition de culture à Estrées-Mons et à Chateaufort sur Isère (récolte 2020) pour les 73 lignées provenant du panel BWP3 étendu. Les protéines totales de ces échantillons ont été mesurées selon la méthode Dumas avec un analyseur élémentaire. Après extraction séquentielle des protéines des farines complètes, chaque classe de protéines de réserve a été dosée par RP-HPLC.

Quarante-cinq lignées représentatives de ces 73 lignées, ont été échantillonnées (Figure 2B) pour mesurer les variables de l'alvéographe de Chopin (Qualtech).

La liste des variables liées aux protéines de réserve est décodée dans la **Table S1**.

### ► Séquençage

L'annotation de 18 gènes de gluténines et 44 gènes de gliadines a été faite chez Chinese-Spring V2.1 (Table 1) et Renan.

Des amorces ont été développées sur Chinese-Spring pour amplifier des régions d'une longueur allant de 4500 à 8500 pb. Seules les amorces de gluténines ont fonctionné en PCR long range. Au final, 18 gènes de gluténines ont été séquencés (6 gènes HMWG et 12 gènes de LMWG) avec la technologie PACBIO HiFi.

Gène (1 <sup>er</sup> SNP)	Chr	CS V2.1 (pb)
Omega_gli_A-1	1A	4 012 862
Gamma_gli_A-1	1A	4 764 891
LMW1A_i_1_159	1A	5 085 939
LMW1A_m_3_7545	1A	7 714 484
Omega_gli_A-2	1A	17 808 422
HMW1Ax_60	1A	510 222 163
HMW1Ax_7073	1A	510 229 176
HMW1Ay_0	1A	510 416 929
HMW1Ay_9706	1A	510 426 635
Gamma_gli_B-1	1B	5 300 534
Omega_gli_B-1	1B	5 727 410
LMW1B_m_4_1	1B	6 886 293
Omega_gli_B-2	1B	22 656 686
LMW1B_m_5_10767	1B	23 030 274
HMW1Bx_12	1B	562 457 564
HMW1Bx_8030	1B	562 465 582
HMW1By_6	1B	562 634 324
HMW1By_8511	1B	562 642 829
Omega_gli_D-1	1D	2 740 455
Gamma_gli_D-1	1D	2 886 542
LMW1D_m_1_68	1D	3 171 203
LMW1D_m_8_7034	1D	7 592 250
HMW1Dx_222	1D	414 602 943
HMW1Dx_6232	1D	414 608 953
HMW1Dy_34	1D	414 658 794
HMW1Dy_7404	1D	414 666 164

**Table 1 :** Principaux gènes sur les chromosomes 1A, B, D. Seul le premier gène et le premier SNP d'un cluster de gènes est référencé

Pour chaque gène amplifié, des séquences consensus ont été produites avec SMRTTools (css) pour les amplicons présentant 3 lectures minimum. Les lignées ont été démultiplexées avec SMRTTools (lima). Puis les gènes ont été clusterisés avec pbaa en utilisant la séquence de Chinese-Spring comme guide, avec un seuil de 1000 lectures minimum par individu et par gène. Un alignement multiple a été obtenu avec ClustalOmega v1.2.4 pour chaque gène. Un calling de SNP a ensuite été produit sur les 18 alignements multiples (outil msa2vcf de jvarkit).

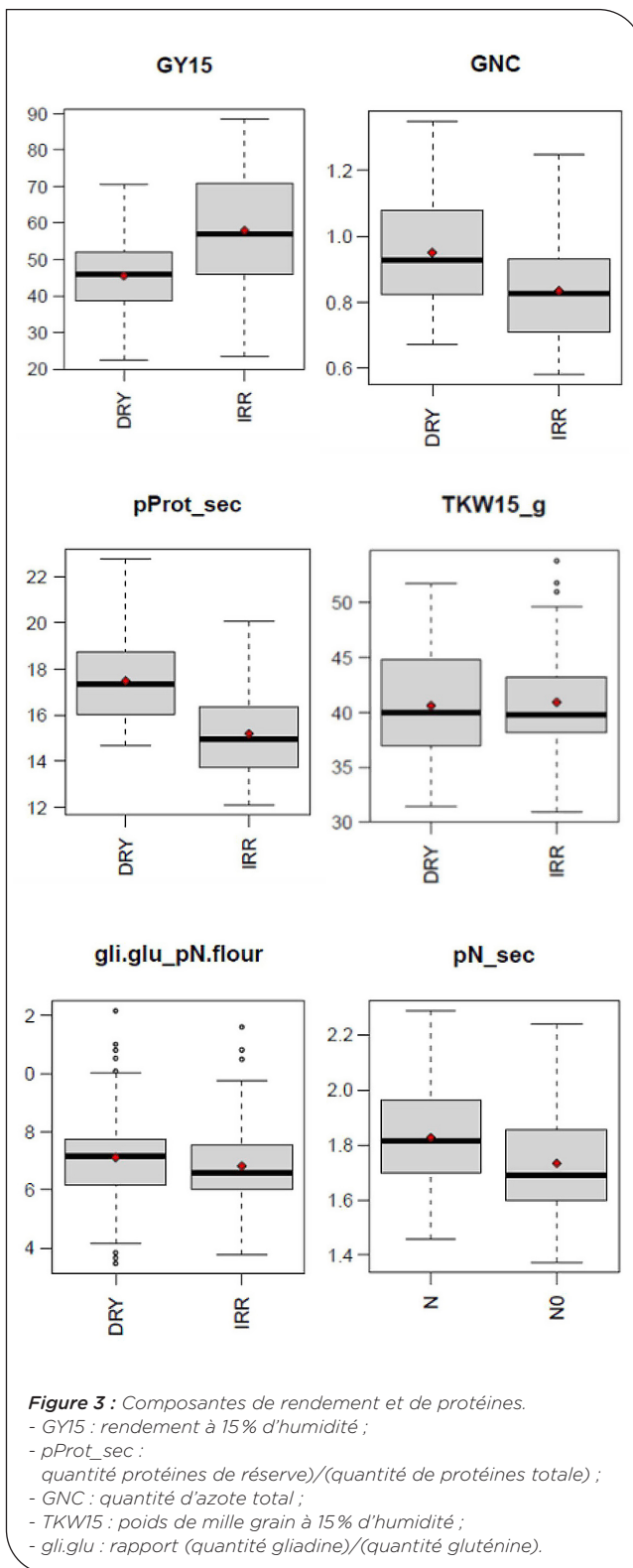
### ► Associations

Les études d'associations ont été effectuées avec le même modèle que celui décrit dans la partie I rendement. Elles ont été effectuées sur les données TaBW410K, les données d'exome de 2000 gènes candidats Breedwheat et les SNP de gluténines produites dans ce projet.

## 3 Résultats

### ► Données de phénotypage

Comme attendu, les rendements étaient supérieurs dans les essais en condition optimale (DRY ou N) et les taux de protéine étaient supérieurs dans les essais stressés (IRR ou NO). Nous n'avons pas observé de différence significative pour les poids de mille grains ou le rapport gliadines/gluténines du grain (Figure 3).



**Figure 3 :** Composantes de rendement et de protéines.

- GY15 : rendement à 15% d'humidité ;
- pProt\_sec : quantité protéines de réserve/(quantité de protéines totale) ;
- GNC : quantité d'azote total ;
- TKW15 : poids de mille grain à 15% d'humidité ;
- gli.glu : rapport (quantité gliadine)/(quantité glutenine).

Les quantités de gliadines et de gluténines sont corrélées à la quantité d'azote total (GNC) et à la quantité de protéines de réserve (Figure 4).

La force boulangère est très corrélée pour les essais d'un même lieu, en condition optimale et stressée (DRY vs IRR ou N vs NO). Elle est significativement corrélée entre essais d'environnements différents (DRY/IRR vs N/NO) (Figure 5).

#### ► Données de séquençage

Parmi les 2102 SNP, 1309 ont une MAF>0.05 et 53 sont en équilibre de liaison.

On remarque que l'allèle « absent » est un allèle fréquent à considérer dans les associations avec les phénotypes (Table 2).

On observe du DL intra-gène. Au total, 25 marqueurs permettent de résumer la diversité des HMW et 29 pour les LMW. Chaque gène possède en moyenne 4 haplotypes.

Gène	#SNP	#hap	PAV
HMW1Ax	2	3	0
HMW1Ay	2	4	1
HMW1Bx	9	4	0
HMW1By	5	3	1
HMW1Dx	2	3	0
HMW1Dy	5	5	1
LMW1Ai1	2	4	1
LMW1Am3	7	5	1
LMW1Bm4	7	6	1
LMW1Bm5	2	4	1
LMW1Dm1	1	2	0
LMW1Dm3	4	5	1
LMW1Dm4	3	5	1
LMW1Dm7	1	3	1
LMW1Dm8	2	3	0

Table 2 : SNP permettant de résumer la diversité nucléotidique des gluténines.

#SNP : nombre de SNP en équilibre de liaison dans le gène.

#hap : nombre d'haplotypes dans le gène.

PAV : absence du gène chez certains individus constituant un allèle supplémentaire.



Figure 4 : Corrélation quantité gliadines et gluténines / quantité d'azote total.

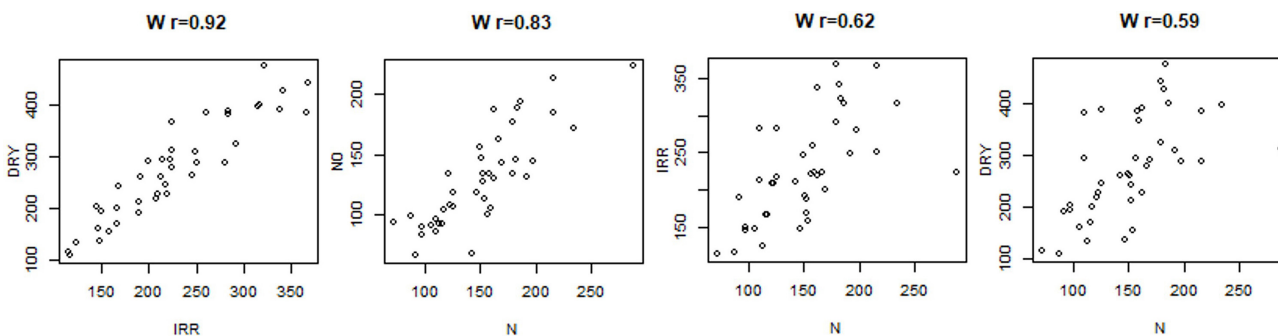


Figure 5 : Corrélation de la force boulangère (W) entre essais.

## ► Associations

On trouve des QTL impliqués dans le taux de protéine au début des chromosomes 1A, 1B et 1D (gliadines et gluténines HMW) (Annexe 1). Les associations les plus fortes ne se situent pas au niveau des marqueurs développés intra-gènes pour ce projet mais sur des marqueurs TaBW410K.

Des associations sur l'ensemble du bras court du chromosome 1B signent la présence de l'introggression 1RS/1BL dans le panel.

Les QTL impliqués dans les protéines expliquent entre 30 et 50% de la variance du caractère (**Table 3**). Ils sont en général associés à plusieurs variables/ environnements.

Quelques manhattan plots sont disponibles en Annexe 1 pour chaque caractère.

## 4 Discussion

L'objectif est d'établir une correspondance entre allèles PAGE et SNP et d'identifier des marqueurs liés aux

variables de l'alvéographe, l'élasticité en particulier. Les QTLs détectés restent à être confirmés sur des panels plus grands. La puissance de détection est en effet très faible pour des panels de 45 accessions (variables d'alvéographes) ou 73 accessions (variables liées aux protéines de réserve). Le jeu de données le plus complet dont nous disposons avec ces phénotypes est une core collection de 366 lignées génotypées sur TaBW410K. Nous allons typer les allèles PAGE sur les 75 lignées de ce projet, imputer les SNP et PAV obtenus dans les gènes de gluténines séquencés dans ce projet sur la core collection et associer les haplotypes obtenus avec les phénotypes.

La région des gliadines étant très complexe, il nous faut poursuivre nos efforts de séquençage long fragments de la région avec la technologie adaptive sampling pour développer des marqueurs pour ces gènes.

Nous ferons un répertoire de l'ensemble des allèles, même rares, du panel BWP3. Nous effectuerons une correspondance entre allèles SNP et PAGE. Nous proposerons ensuite un kit de marqueurs Kaspar pour faire un génotypage en routine des SNP associés aux variables liées à la qualité boulangère.

Trait	#Chr	Chr	Position (CS V2.1, bp)	Pvalue	R2LR	QTL
le	1	1A	509 017 352	9.40E-07	0.49	Glu_HMW-A
L	5	2B	745 716 941	2.19E-07	0.56	
L	15	5D	548 216 820	4.84E-07	0.52	CWI-5D
pPR.ProtTot_IRR	2	1B	94 458 186	6.01E-09	0.37	1RS/1BL(gli/LMW)
pPR.ProtTot_IRR	2	1B	105 538 550	6.01E-09	0.37	1RS/1BL(gli/LMW)
part.w5_DRY	3	1D	578 541	3.33E-09	0.39	gli/LMW
pN_gluTot.flour1_DRY	7	3A	724 660 462	6.02E-07	0.29	TGW6
pN_g.Nflour1_NO	8	3B	373 613 528	4.50E-07	0.32	TaLyc-B1
pN_g.Nflour1_NO	8	3B	774 409 784	4.32E-07	0.32	Red-3B
pN_g.Nflour1_NO	10	4A	263 086 190	4.72E-07	0.32	introggression dicoccoides
part.w5_DRY	11	4B	1 135 117	6.22E-09	0.38	Lox-B1
part.W1.2_IRR	13	5A	476 434 666	6.77E-07	0.30	nac27
pN_w5.Nflour1_N	14	5B	641 046 239	8.76E-07	0.30	SHP
part.w5_DRY	14	5B	703 647 253	7.25E-09	0.38	
part.w5_DRY	17	6B	10 754 025	7.96E-09	0.39	
qteN_w5.gr1_DRY	17	6B	514 027 602	4.38E-07	0.30	SAD

**Table 3 :** QTL liés aux protéines de réserve.

R2LR : part de variance du phénotype expliquée par le SNP.

Les caractères sont décodés dans la **Table S1**.

## IV. Études d'association phénologie/maladies

### 1 Introduction

Les maladies les plus impactantes en France sont la rouille, la septoriose et la fusariose. Deux types de gènes de résistance ont été mis en évidence, les gènes de résistance plantule (SR : seedling resistance) et les gènes de résistance stade adulte (APR : adult plant resistance). Les gènes SR confèrent une résistance qualitative (monogénique) à une ou quelques souches par une réponse d'hypersensibilité (mort cellulaire). Les gènes APR confèrent une résistance quantitative à de nombreuses souches et sont par nature plus durables.

La pression de rouille jaune (*Puccinia striiformis f. Tritici*; Pst) (Bouvet *et al.*, 2022a, 2022b) augmente depuis 2011 en France avec l'apparition du pathotype d'origine exotique Warrior 1. Les contournements sont rapides et on ne connaît pas les QTLs/gènes responsables des résistances actuelles.

Les attaques de rouille brune (*Puccinia triticina*) sont plus tardives que la rouille jaune au champ et elle peut être difficile à noter sur des plantes sensibles à la rouille jaune. A ce jour, 80 gènes Lr ont été identifiés dans le blé (Aktar-Uz-Zaman *et al.*, 2017; Kumar *et al.*, 2021) dont la plupart sont des SR contournés (Babayants *et al.*, 2015; Safavi and Afshari, 2013). Au moins 14 gènes Lr sont présents dans les variétés cultivées en France, seuls ou en combinaison (Goyeau and Lannou, 2011).

La fusariose (FHB; *Fusarium head blight*) causée par *Fusarium graminearum* affecte le rendement ainsi que la qualité du grain par l'accumulation de mycotoxines dans les grains non infectés (Dweba *et al.*, 2017). Plusieurs centaines de QTLs distribués sur tous les chromosomes ont été reportés (Bai *et al.*, 2018; Liu *et al.*, 2009; Löffler *et al.*, 2009).

Pour la septoriose (*Septoria tritici blotch*, STB), depuis quelques années, quelques variétés comme Cheignon,

LG\_Absalon, KWS\_Extase, RGT\_Cesario se démarquent par leur très bon niveau de résistance durable. Mais même si les contournements sont plus progressifs que pour les rouilles, on observe une érosion du niveau de résistance relativement rapide pour certaines variétés (Mutic, Filon, Unik, Cellule, Rubisko) contournés quelques années après leur déploiement. La plupart des souches décrites actuellement sont virulentes sur les 23 gènes Stb connus (Battache *et al.*, 2022; Saintenac *et al.*, 2021, 2018). Certaines sont avirulentes sur un seul gène. Ceci suggère que la résistance due à un ou plusieurs gènes Stb est rapidement contournée (Brown *et al.*, 2015).

Mais la résistance peut aussi être due à l'accumulation d'allèle à effet faible. En effet, des centaines de QTLs ont été identifiés en population bi-parentale ou panel de diversité pour la rouille brune (Azzimonti *et al.*, 2014, 2013; Dweba *et al.*, 2017; Goyeau *et al.*, 2010; Niks *et al.*, 2015; Rollar *et al.*, 2021), la septoriose (Brown *et al.*, 2015) et la fusariose (Zheng *et al.*, 2021).

Le moyen de détection de sources résistantes le plus sûr actuellement reste la notation de la maladie au champ. Et la détection d'allèles favorables la plus solide se fait dans des descendances de croisements entre parents sensibles et résistants aux souches actuelles. Les études d'association sont intéressantes pour améliorer la résolution des QTLs par rapport aux populations bi-parentales mais peuvent produire une grosse proportion de faux positifs. Et elles ne permettent pas de détecter un effet significatif au niveau d'allèles rares, ce qui est limitant.

## 2 Matériel et méthode

Les maladies majeures ont été évaluées dans plusieurs essais répartis entre partenaires (Table 1).

Afin de différencier les QTL d'évitement des QTL de résistance, les études d'association ont été conduites sur le phénotype et les résidus du phénotype d'un modèle qui inclut précocité et hauteur. Le modèle utilisé est le même que celui décrit dans la partie I rendement.

## 3 Résultats

### ► Présence de barbe

Le QTL de la présence de barbe explique 87% du phénotype sur le chromosome 5A.

### ► Date d'épaison/Hauteur de plante

A part Ppd-D1 qui explique 36% de la date d'épaison, les autres QTL expliquent entre 8 et 10% de la variance (Table 2). Les QTL de hauteur expliquent entre 9 et 15% de la variance (Rht-B1, Rht-D1 et Rht8).

### ► Maladies

Pour la fusariose, la septoriose et la rouille jaune, le QTL le plus fort correspond à la région impliquée dans la photosensibilité et la hauteur Ppd-D1/Rht8 sur le chromosome 2D.

Lorsqu'on utilise un modèle avec la hauteur et la précocité en covariables, les QTL les plus intéressants ont été trouvés pour la rouille jaune, sur les chromosomes 6A et 2B. Il semble qu'il y ait une introgression sur le 2B présente dans le panel visible sur le bras court.

Les QTL maladie expliquent entre 4 et 11% de la variance du caractère (Table 3).

La plupart des QTL septoriose Ex-IGE colocalisent avec des QTLs publiés. Le QTL le plus important n'est pas décrit dans la littérature. Il se situe sur le chromosome 3A.

La plupart des QTL de fusariose ont été trouvés dans l'essai Limagrain 2019.

Quelques manhattan plots sont présentés en Annexe 1.

## 4 Discussion

Les données de génotypage seront imputées pour augmenter la puissance des études d'association.

Une méta-analyse des QTL de septoriose est en cours dans l'équipe INRAE GDEC MDC (Maladies Des Céréales). Il serait intéressant de faire une méta-analyse pour la rouille jaune également.

	Programme			Notations réalisées			Inoc. SEP	Inoc. FUS
	R17	R19	R20	2016-2017	2018-2019	2019-2020		
BASF								
Limagrain				SEP, YR, LR, PM			17	
Lidea	SEP	LR		SEP, LR, PM	BR		17	
Florimond Desprez		SEP	SEP		(SEP), YR	SEP, YR, BR		
Limagrain		FUS			FUS, YR			19
INRAE						YR (hors programme)		
KWS	FUS	SEP		FUS, PM	SEP, YR			17
R2N		SEP		SEP, YR	YR		17	
Secobra	FUS	YR	SEP	FUS	X	SEP		17
Syngenta		YR	YR		X	YR		

Table 1 : Liste des essais maladie.

Trait	#Chr	Chr	Position	Pvalue	R2LR	QTL
AWN	13	5A	698 508 163	6.70E-240	0.87	
HD	2	1B	676 950 510	3.73E-10	0.08	
HD	4	2A	576 652 403	9.31E-10	0.08	
HD	5	2B	52 122 752	5.98E-09	0.07	Ppd-B1
HD	6	2D	34 676 118	1.33E-50	0.36	Ppd-D1
HD	8	3B	52 457 571	9.37E-12	0.10	
HD	14	5B	511 291 951	9.91E-11	0.09	

Trait	#Chr	Chr	Position	Pvalue	R2LR	QTL
HD	15	5D	561 638 264	4.62E-10	0.08	
HD	20	7B	50 358 038	6.48E-11	0.09	
HD	20	7B	633 843 543	5.69E-11	0.09	
PH	6	2D	34 676 118	1.14E-18	0.15	Ppd-D1, Rht8
PH	11	4B	40 214 576	1.26E-11	0.09	Rht-B1
PH	12	4D	18 738 996	6.90E-14	0.11	Rht-D1

Table 2 : QTL phénologie. AWN : présence de barbes, HD : date d'épaison, PH : hauteur de plante

Une approche bi-parentale qui croise un parent résistant avec un parent récemment contourné semble la meilleure approche pour détecter les QTL importants pour résister aux races actuelles pour la rouille jaune et la septoriose. Il serait intéressant de faire un croisement entre deux parents possédant des allèles sensibles et résistants sur le QTL de septoriose du chromosome 3A par exemple pour le valider et les deux QTL rouille jaune 2B et 6A.

D'après le caractère très quantitatif de la fusariose, une approche de prédiction génomique devrait être testée.

Un jeu de données de qualité de génotypes et de phénotypes a été produit sur un large panel de diversité (485). Des analyses poussées ont été faites sur les données rendement. Les analyses sont encore en cours pour les

autres caractères. Un travail d'imputation et de méta-analyses avec d'autres jeux de données sera nécessaire pour valider les QTL, en particulier pour les caractères où nous avons évalué seulement 45 (alvéographe) ou 75 (protéines de réserve) individus. L'objectif est de détecter des géniteurs intéressants et de constituer un set de marqueurs diagnostiques pour optimiser le pre-breeding.

### Accessibilité des données

Toutes les données de génotypage, de phénotypage et les résultats d'études d'association sont disponibles sur ce dépôt :

<https://entrepot.recherche.data.gouv.fr/dataset.xhtml?persistentId=doi:10.57745/VA8HMV>

Trait	#Chr	Chr	Position	Pvalue	R2LR	QTL
FUS	2	1B	587 089 444	1.29E-12	0.11	
FUS	5	2B	52 122 856	4.25E-13	0.11	
FUS	6	2D	34 676 118	1.30E-23	0.20	Ppd-D1, Rht8
FUS	8	3B	545 047 717	7.24E-11	0.09	
FUS	10	4A	600 976 627	8.78E-12	0.10	
FUS	14	5B	511 291 951	9.95E-13	0.11	
FUS	16	6A	551 943 682	2.89E-12	0.11	
FUS	17	6B	90 568 781	1.80E-12	0.11	
FUS	17	6B	717 891 974	4.32E-11	0.10	
FUS	20	7B	633 677 649	8.69E-12	0.10	
PM	1	1A	502 825 832	1.04E-12	0.11	
PM	13	5A	472 169 023	3.73E-12	0.10	
RJ	1	1A	468 407 947	5.65E-07	0.05	
RJ	3	1D	3 228 464	2.96E-07	0.06	Lr21
RJ	5	2B	68 374 877	3.13E-08	0.06	Ppd-B1
RJ	5	2B	137 932 382	1.95E-08	0.06	
RJ	5	2B	423 910 387	1.47E-07	0.04	Sr36
RJ	5	2B	692 714 733	1.94E-08	0.05	Yr7
RJ	5	2B	715 969 646	1.16E-08	0.05	YrSP
RJ	6	2D	33 955 687	1.54E-09	0.08	Ppd-D1
RJ	6	2D	547 423 316	3.29E-07	0.05	
RJ	9	3D	43 023 115	5.36E-07	0.05	
RJ	11	4B	521 608 690	2.12E-07	0.06	
RJ	13	5A	19 176 283	6.30E-07	0.06	
RJ	13	5A	487 696 937	9.77E-07	0.05	
RJ	13	5A	623 378 074	6.93E-07	0.05	
RJ	14	5B	43 638 059	2.74E-08	0.06	
RJ	14	5B	381 077 636	9.88E-07	0.05	Ne1
RJ	14	5B	440 210 891	7.08E-07	0.07	
RJ	16	6A	526 369 232	4.86E-08	0.06	Sr26
RJ	16	6A	613 075 235	6.10E-10	0.08	Sr13
RJ	18	6D	23 360 660	4.93E-08	0.06	
ST	2	1B	146 004 631	6.22E-07	0.08	wmc0626, Kollers <i>et al.</i> (2013)
ST	3	1D	361 475 043	9.15E-07	0.05	TraesCS1D01G265800, TraesCS1D01G266000, Brennan <i>et al.</i> (2020)
ST	5	2B	589 099 322	7.87E-07	0.05	QStb.ihar-2B.5, Piaszkowska <i>et al.</i> (2021)
ST	5	2B	681 047 323	4.23E-08	0.07	QStb.teagasc-2B.1, Stadlmeier <i>et al.</i> (2019)
ST	5	2B	756 026 847	1.77E-08	0.10	Stb9
ST	6	2D	34 676 118	3.04E-09	0.07	Ppd-D1
ST	7	3A	146 598 677	5.59E-07	0.08	QSTB.16, Mekonnen <i>et al.</i> (2021)
ST	7	3A	467 773 523	5.53E-07	0.05	QStb.wai.3A.2, Yang <i>et al.</i> (2022)
ST	7	3A	534 986 495	5.05E-10	0.08	
ST	7	3A	653 840 752	4.84E-08	0.07	
ST	9	3D	395 007 366	3.27E-08	0.10	
ST	13	5A	472 485 773	7.81E-07	0.06	QStb.cim-5AL-1, Dreisigacker <i>et al.</i> (2015)
ST	14	5B	439 993 368	2.63E-07	0.06	QTL-5BL, Mergoum <i>et al.</i> (2013), PGV_25, Yates <i>et al.</i> (2019)
ST	14	5B	596 724 405	2.24E-07	0.07	QStb.teagasc-5B.1, Riaz <i>et al.</i> (2020)
ST	16	6A	14 301 074	4.26E-08	0.10	QStb.ihar-6A, Radecka-Janusik <i>et al.</i> (2014)
ST	17	6B	77 872 165	5.99E-08	0.06	MQTL21, Goudemand <i>et al.</i> (2013)
ST	19	7A	76 126 109	2.82E-07	0.09	MQTL24, Goudemand <i>et al.</i> (2013)
ST	19	7A	729 488 814	3.80E-07	0.08	pQStb.wai.7A.2, Yang <i>et al.</i> (2022)
ST	20	7B	625 716 792	1.50E-08	0.10	QTL_7B_2 Karlstedt <i>et al.</i> (2019)
ST	20	7B	711 287 379	9.06E-07	0.05	Stb8? 721-753Mb, Adhikari <i>et al.</i> (2003); Qstb.renan-7B, Langlands-perry <i>et al.</i> (2022); wPt-2356, Arraiano et Brown (2017)
ST	21	7D	113 154 201	2.57E-07	0.09	Stb5? 50-98Mb, Arraiano <i>et al.</i> (2001)
ST	21	7D	636 138 983	1.25E-07	0.06	MQTL27, Goudemand <i>et al.</i> (2013)

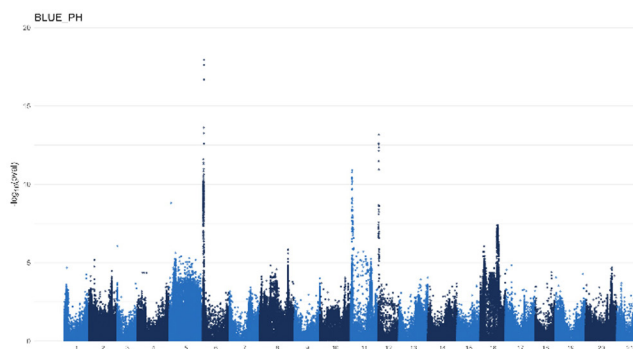
Table 3 : QTL maladies. FUS : fusariose, RJ : rouille jaune, ST : septoriose, PM : oïdium.

# Annexe 1 : Manhattan plots

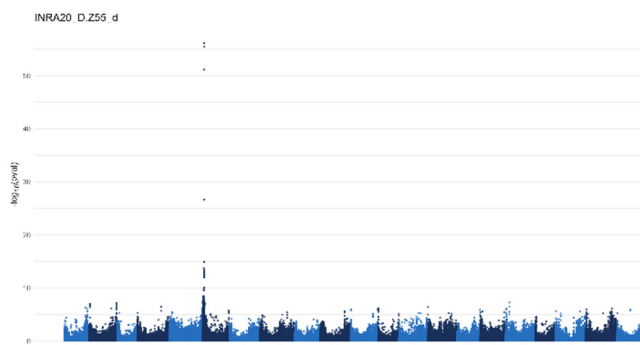
### Présence de barbes



### Hauteur de plante

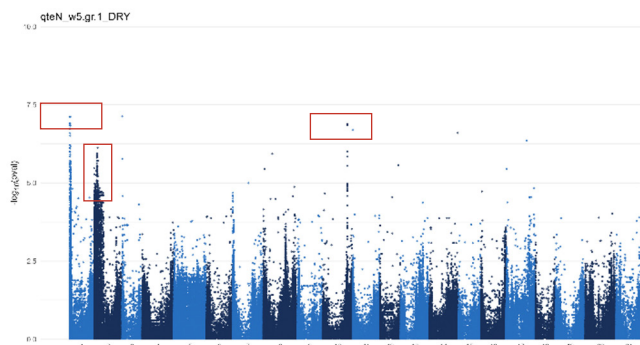


### Date d'épiaison

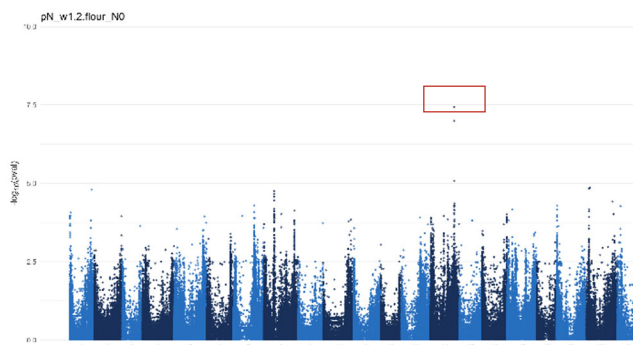


## Protéines de réserve

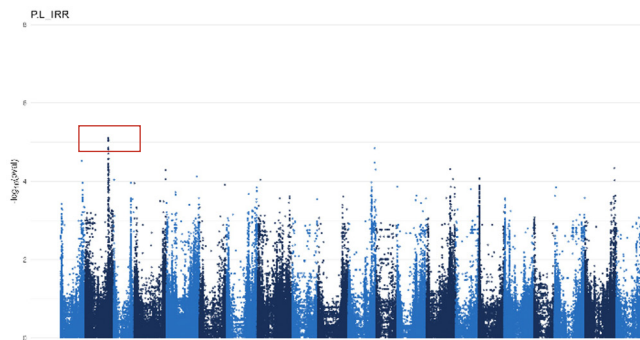
### quantité de gliadines omega5 dans le grain



### quantité de gliadines omega 1.2 dans le grain



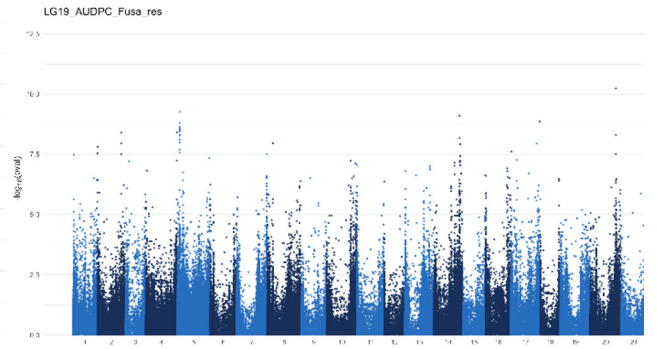
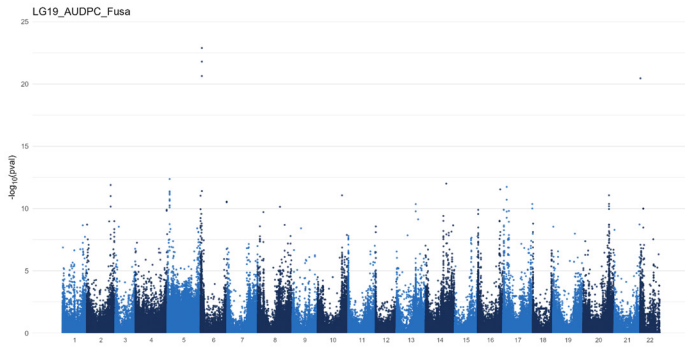
### rapport P/L



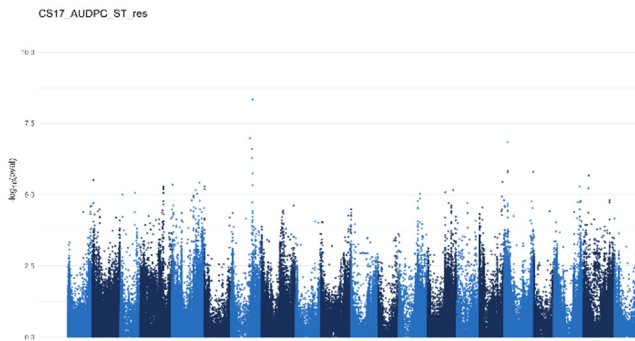
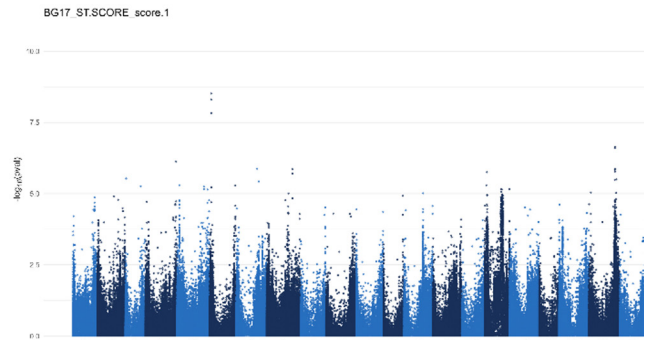
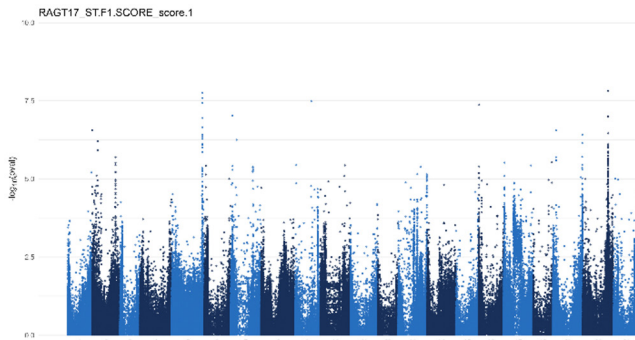


## Fusariose

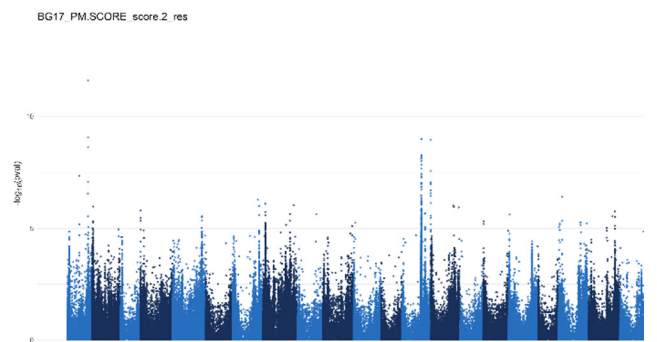
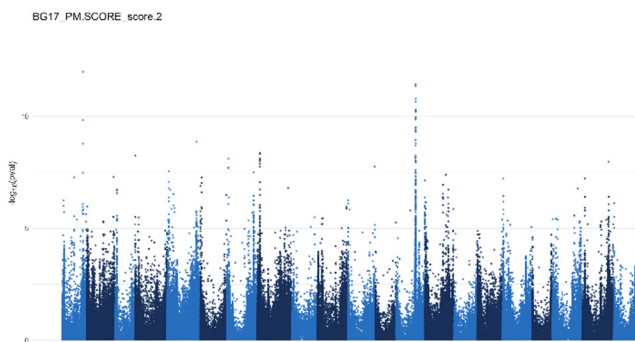
res : résidus d'un modèle avec hauteur et précocité en covariables



## Septoriose

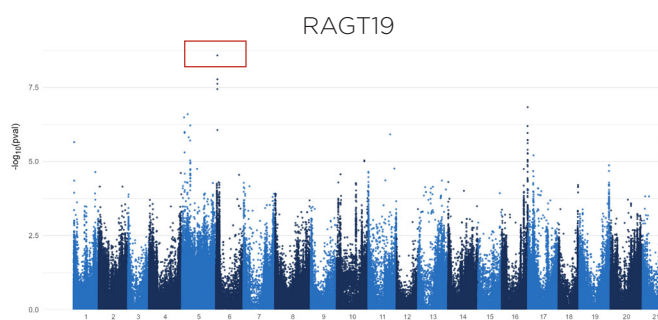
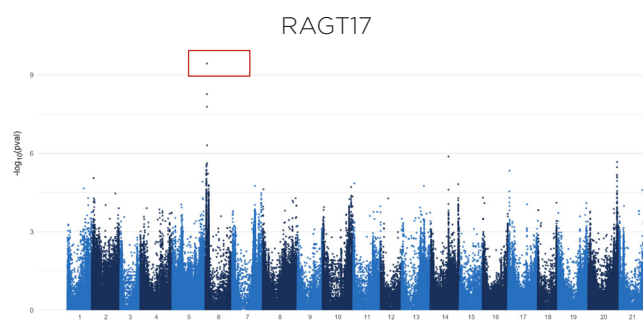


## Oïdium

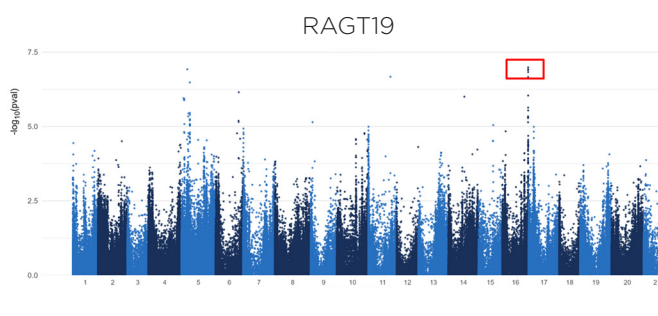
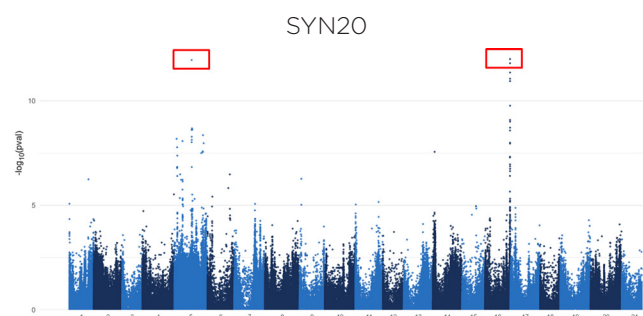


## Rouille jaune

Ppd-D1 sur chromosome 2D



res : résidus de modèle avec précocité et hauteur en covariable



**Table S1 : Variables liées aux protéines de réserve mesurées par RP-HPLC**

Abbréviation	Caractère
<b>HMW-LMW</b>	rapport HMW/LMW
<b>part.HMW</b>	pourcentage de l'air des HMW sur l'air total des Glu
<b>part.LMW</b>	pourcentage de l'air des LMW sur l'air total des Glu
<b>pN_HMW.flour</b>	pourcentage d'azote correspondant aux HMW dans la farine
<b>pN_LMW.flour</b>	pourcentage d'azote correspondant aux LMW dans la farine
<b>pN_gluTot.flour</b>	pourcentage d'azote correspondant aux gluténines dans la farine
<b>pN_HMW.Nflour</b>	pourcentage d'azote de HMW par rapport à l'azote total de la farine
<b>pN_LMW.Nflour</b>	pourcentage d'azote de LMW par rapport à l'azote total de la farine
<b>pN_gluTot.Nflour</b>	pourcentage d'azote de gluténines par rapport à l'azote total de la farine
<b>qteN_HMW.gr</b>	quantité d'azote allouée aux HMW dans un grain moyen
<b>qteN_LMW.gr</b>	quantité d'azote allouée aux LMW dans un grain moyen
<b>qteN_glu.gr</b>	quantité d'azote allouée aux gluténines dans un grain moyen
<b>part.w5</b>	pourcentage de l'air des w5 sur l'air total des Gli
<b>part.w1.2</b>	pourcentage de l'air des w1,2 sur l'air total des Gli
<b>part.ab</b>	pourcentage de l'air des ab sur l'air total des Gli
<b>part.g</b>	pourcentage de l'air des g sur l'air total des Gli
<b>pN_w5.flour</b>	pourcentage d'azote de w5 par rapport à l'azote total de la farine
<b>pN_w1-2.flour</b>	pourcentage d'azote de w1,2 par rapport à l'azote total de la farine
<b>pN_ab.flour</b>	pourcentage d'azote de α par rapport à l'azote total de la farine
<b>pN_g.flour</b>	pourcentage d'azote de γ par rapport à l'azote total de la farine
<b>pN_gliTot.flour</b>	pourcentage d'azote de gliadines par rapport à l'azote total de la farine

Abbréviation	Caractère
<b>pN_w5.Nflour</b>	pourcentage d'azote correspondant aux w5 dans la farine
<b>pN_w1-2.Nflour</b>	pourcentage d'azote correspondant aux w1,2 dans la farine
<b>pN_ab.Nflour</b>	pourcentage d'azote correspondant aux α dans la farine
<b>pN_g.Nflour</b>	pourcentage d'azote correspondant aux γ dans la farine
<b>pN_gliTot.Nflour</b>	pourcentage d'azote correspondant aux gliadines dans la farine
<b>qteN_w5.gr</b>	quantité d'azote allouée aux w5 dans un grain moyen
<b>qteN_w1-2.gr</b>	quantité d'azote allouée aux w1,2 dans un grain moyen
<b>qteN_ab.gr</b>	quantité d'azote allouée aux α dans un grain moyen
<b>qteN_g.gr</b>	quantité d'azote allouée aux γ dans un grain moyen
<b>qteN_gliTot.gr</b>	quantité d'azote allouée aux gliadines dans un grain moyen
<b>gli-glu_pN.flour</b>	rapport Gli/Glu calculé avec les données de pN.flour
<b>gli-glu_pN.Nflour</b>	rapport Gli/Glu calculé avec les données de pN.Nflour
<b>gli-glu_qteN.gr</b>	rapport Gli/Glu calculé avec les données de qteN.gr
<b>pPR.ProtTot</b>	pourcentage des protéines de réserve en fonction des protéines totales corrigé par les rendement d'extraction
<b>pGlu.PR</b>	pourcentage des Gluténines en fonction des protéines de réserve corrigé par les rendement d'extraction
<b>pGli.PR</b>	pourcentage des Gliadines en fonction des protéines de réserve corrigé par les rendement d'extraction
<b>L</b>	extensibilité
<b>W</b>	force boulangère
<b>P/L</b>	ténacité/extensibilité
<b>le</b>	élasticité

## Références bibliographiques

### **Diversité blé tendre**

**Balfourier, F., Bouchet, S., Robert, S., De Oliveira, R., Rimbart, H., Kitt, J., Choulet, F., Paux, E.,** 2019. Worldwide phylogeography and history of wheat genetic diversity. *Sci Adv* 5, eaav0536. <https://doi.org/10.1126/sciadv.aav0536>

### **Prédictions génomiques et GxE**

**Asseng S, Martre P, Maiorano A, et al** (2019) Climate change impact and adaptation for wheat protein. *Global Change Biology* 25:155-173.

**Bhat JA, Ali S, Salgotra RK, et al** (2016) Genomic Selection in the Era of Next Generation Sequencing for Complex Traits in Plant Breeding. *Frontiers in Genetics* 7:221

**Cobb JN, Juma RU, Biswas PS, et al** (2019) Enhancing the rate of genetic gain in public-sector plant breeding programs: lessons from the breeder's equation. *Theor Appl Genet* 132:627-645.

**Cullis BR, Smith AB, Coombes NE** (2006) On the design of early generation variety trials with correlated data. *JABES* 11:381.

**Dias KOG, dos Santos JPR, Krause MD, et al** (2022) Leveraging probability concepts for cultivar recommendation in multi-environment trials. *Theor Appl Genet* 135:1385-1399.

**Elshire RJ, Glaubitz JC, Sun Q, et al** (2011) A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLOS ONE* 6:e19379.

**Gerke JP, Edwards JW, Guill KE, et al** (2015) The Genomic Impacts of Drift and Selection for Hybrid Performance in Maize. *Genetics* 201:1201-1211.

**Glaszmann J, Kilian B, Upadhyaya H, Varshney R** (2010) Accessing genetic diversity for crop improvement. *Current Opinion in Plant Biology* 13:167-173.

**Gollob HF** (1968) A statistical model which combines features of factor analytic and analysis of variance techniques. *Psychometrika* 33:73-115.

**Guo T, Mu Q, Wang J, et al** (2020) Dynamic effects of interacting genes underlying rice flowering-time phenotypic plasticity and global adaptation. *Genome Research* 30:gr.255703.119.

**Kitt J, Danguy Des Déserts A, Bouchet S, et al** (2021) Genotyping of 4,506 bread wheat accessions with the TaBW410K SNP array. Zenodo.

**Li X, Guo T, Mu Q, et al** (2018) Genomic and environmental determinants and their interplay underlying phenotypic plasticity. *Proceedings of the National Academy of Sciences* 115:6679-6684.

**Li X, Guo T, Wang J, et al** (2021) An integrated framework reinstating the environmental dimension for GWAS and genomic selection in crops. *Molecular Plant* 14:874-887.

**Longin CFH, Reif JC** (2014) Redesigning the exploitation of wheat genetic resources. *Trends in Plant Science* 19:631-636.

**Ly D, Chenu K, Gauffreteau A, et al** (2017) Nitrogen nutrition index predicted by a crop model improves the genomic prediction of grain number for a bread wheat core collection. *Field Crops Research* 214:331-340.

**McCouch S, Baute GJ, Bradeen J, et al** (2013) Feeding the future. *Nature* 499:23-24.

**Parent B, Leclere M, Lacube S, et al** (2018) Maize yields over Europe may increase in spite of climate change, with an appropriate use of the genetic variability of flowering time. *PNAS* 201720716.

**Piepho H-P, Blancon J** (2023) Extending Finlay-Wilkinson regression with environmental covariates. *Plant Breeding* 142:621-631.

**Ray DK, Ramankutty N, Mueller ND, et al** (2012) Recent patterns of crop yield growth and stagnation. *Nature Communications* 3:1293.

**Reif JC, Zhang P, Dreisigacker S, et al** (2005) Wheat genetic diversity trends during domestication and breeding. *Theor Appl Genet* 110:859-864.

**Rimbart H, Darrier B, Navarro J, et al** (2018) High throughput SNP discovery and genotyping in hexaploid wheat. *PLOS ONE* 13:e0186329.

**Rincent R, Malosetti M, Ababaei B, et al** (2019) Using crop growth model stress covariates and AMMI decomposition to better predict genotype-by-environment interactions. *Theor Appl Genet* 132:3399-3411.

**Smith A, Cullis B, Gilmour A** (2001) Applications: The Analysis of Crop Variety Evaluation Data in Australia. *Australian & New Zealand Journal of Statistics* 43:129-145.

**Soenen B, Le Bris X, Laberdesque M, et al** (2016) « CHN », a crop model to jointly manage water and nitrogen on winter wheat

**VanRaden PM** (2008) Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science* 91:4414-4423.

**Wang C, Hu S, Gardner C, Lübberstedt T** (2017) Emerging Avenues for Utilization of Exotic Germplasm. *Trends in Plant Science* 22:624-637.

### **GWAS - caractères agronomique**

**Cormier F, Le Gouis J, Dubreuil P, et al** (2014) A genome-wide identification of chromosomal regions determining nitrogen use efficiency components in wheat (*Triticum aestivum* L.). *Theor Appl Genet* 127:2679-2693. <https://doi.org/10/f6qgfm>

**Rincent R, Moreau L, Monod H, et al** (2014) Recovering Power in Association Mapping Panels with Variable Levels of Linkage Disequilibrium. *Genetics* 197:375-387. <https://doi.org/10.1534/genetics.113.159731>

**Sun G, Zhu C, Kramer MH, et al** (2010) Variation explained in mixed-model association mapping. *Heredity* 105:333-340. <https://doi.org/10.1038/hdy.2010.11>

**Yu J, Pressoir G, Briggs WH, et al** (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet* 38:203-208. <https://doi.org/10/d8npfr>

### **Qualité boulangère**

**Bordes J, Branlard G, Oury FX, Charmet G, Balfourier F** (2008) Agronomic characteristics, grain quality and flour rheology of 372 bread wheats in a worldwide core collection. *J Cereal Sci* 48 : 569-579.

**Bordes J, Ravel C, Le Gouis J, Lapierre A, Charmet G, Balfourier F** (2011) Use of a global wheat core collection for association analysis of flour and dough quality traits. *Journal of Cereal Science* 54 : 137-147.

**Branlard G, Dardevet M** (1985) Diversity of grain protein and bread wheat quality .2. Correlation between high molecular-weight subunits of glutenin and flour quality characteristics. *J Cereal Sci* 3:345-354

**Branlard G, Dardevet M, Saccomano R, Lagoutte F, Gourdon J** (2001) Genetic diversity of wheat storage proteins and bread wheat quality. *Euphytica* 119 : 59-67.

**Charmet G, Robert N, Branlard G, Linossier L, Martre P, Triboui E** (2005) Genetic analysis of dry matter and nitrogen accumulation and protein composition in wheat kernels. *Theoretical and Applied Genetics* 111 : 540-550.

**Johansson E, Malik AH, Hussain A, Rasheed F, Newson WR, Plivelic T, Hedenqvist MS, Gällstedt M, Kuktaite R** (2013). Wheat Gluten Polymer Structures: The Impact of Genotype, Environment, and Processing on Their Functionality in Various Applications. *Cereal Chemistry*, 90:367–376. <https://doi.org/10/f22djp>

**Payne PI, Nightingale MA, Krattiger AF, Holt LM** (1987) The relationship between HMW glutenin subunit composition and the bread-making quality of British-grown wheat varieties. *J Sci Food Agric* 40 (1):51-65

**Pirozi M, Margiotta B, Latiandra D, MacRitchie F** (2008) Composition of polymeric proteins and bread-making quality of wheat lines with allelic HMW-GS differing in number of cysteines. *J Cereal Sci* 48: 117-122

**Plessis A, Ravel C, Duchateau N, Dardevet M, Bordes J, Rhazi L, Balfourier F, Martre P** (2013) Association study of wheat grain protein composition reveals that gliadin and glutenin composition are trans-regulated by different chromosome regions. *J Exp Bot*

**Oury, F.-X., Chiron, H., Faye, A., Gardet, O., Giraud, A., Heumez, E., Rolland, B., Rousset, M., Trottet, M., Charmet, G., et al.** (2010). The prediction of bread wheat quality: joint use of the phenotypic information brought by technological tests and the genetic information brought by HMW and LMW glutenin subunits. *Euphytica* 171, 87-109.

**Ravel C, Faye A, Ranoux M, Dardevet M, Dupuits C, Exbrayat F, Poncet C, Sourdille P, Branlard G** 2020 SNP markers for early identification of high molecular weight glutenin subunits (HMW-GSs) in bread wheat. *Theor Appl Genet* 133: 751-770. <https://doi.org/10.1007/s00122-019-03505-y>

#### **GWAS - Maladies**

**Adhikari T. B., Anderson J. M., et Goodwin S. B.**, « Identification and molecular mapping of a gene in wheat conferring resistance to *Mycosphaerella graminicola* », *Phytopathology*, vol. 93, no 9, p. 1158-1164, 2003, doi: <https://doi.org/10.1094/PHTO.2003.93.9.1158>.

**Arraiano L. S. et Brown J. K.**, « Sources of resistance and susceptibility to *Septoria tritici* blotch of wheat », *Mol. Plant Pathol.*, vol. 18, no 2, p. 276-292, 2017, doi: <https://doi.org/10.1111/mp.12482>.

**Aktar-Uz-Zaman, M., Tuhina-Khatun, M., Hanafi, M.M., Sahebi, M.**, 2017. Genetic analysis of rust resistance genes in global wheat cultivars: an overview. *Biotechnology & Biotechnological Equipment* 31, 431-445. <https://doi.org/10.1080/13102818.2017.1304180>

**Azzimonti, G., Lannou, C., Sache, I., Goyeau, H.**, 2013. Components of quantitative resistance to leaf rust in wheat cultivars: diversity, variability and specificity. *Plant Pathology* 62, 970-981. <https://doi.org/10.1111/ppa.12029>

**Azzimonti, G., Marcel, T.C., Robert, O., Paillard, S., Lannou, C., Goyeau, H.**, 2014. Diversity, specificity and impacts on field epidemics of QTLs involved in components of quantitative resistance in the wheat leaf rust pathosystem. *Mol Breeding* 34, 549-567. <https://doi.org/10.1007/s11032-014-0057-8>

**Babayants, O., Babayants, L., Gorash, A., Vasilev, A., Traskovetskaya, V., Galaev, A.**, 2015. Physiologic specialization of *Puccinia triticina* Erikss. and effectiveness of Lr-genes in the south of Ukraine during 2013-2014. *Chilean journal of agricultural research* 75, 443-450. <https://doi.org/10.4067/S0718-58392015000500009>

**Bai, G., Su, Z., Cai, J.**, 2018. Wheat resistance to *Fusarium* head blight. *Canadian Journal of Plant Pathology* 40, 336-346. <https://doi.org/10.1080/07060661.2018.1476411>

**Battache, M., Lebrun, M.-H., Sakai, K., Soudière, O., Cambon, F., Langin, T., Saintenac, C.**, 2022. Blocked at the Stomatal Gate, a Key Step of Wheat Stb16q-Mediated Resistance to *Zymoseptoria tritici*. *Front Plant Sci* 13, 921074. <https://doi.org/10.3389/fpls.2022.921074>

**Bouvet, L., Holdgate, S., James, L., Thomas, J., Mackay, I.J., Cockram, J.**, 2022a. The evolving battle between yellow rust and wheat: implications for global food security. *Theor Appl Genet* 135, 741-753. <https://doi.org/10.1007/s00122-021-03983-z>

**Bouvet, L., Percival-Alwyn, L., Berry, S., Fenwick, P., Mantello, C.C., Sharma, R., Holdgate, S., Mackay, I.J., Cockram, J.**, 2022b. Wheat genetic loci conferring resistance to stripe rust in the face of genetically diverse races of the fungus *Puccinia striiformis* f. sp. tritici. *Theoretical and Applied Genetics* 1-19.

**Brennan et al.** 2020 <https://doi.org/10.3389/fpls.2020.00433>

**Brown, J.K.M., Chartrain, L., Lasserre-Zuber, P., Saintenac, C.**, 2015. Genetics of resistance to *Zymoseptoria tritici* and applications to wheat breeding. *Fungal Genetics and Biology, Septoria tritici* blotch disease of wheat: Tools and techniques to study the pathogen *Zymoseptoria tritici* 79, 33-41. <https://doi.org/10.1016/j.fgb.2015.04.017>

**Dreisigacker S., Wang X., Martinez Cisneros B. A., Jing R., et Singh P. K.**, « Adult-plant resistance to *Septoria tritici* blotch in hexaploid spring wheat », *Theor. Appl. Genet.*, vol. 128, no 11, p. 2317-2329, 2015. Dweba, C.C., Figlan, S., Shimelis, H.A., Motaung, T.E., Sydenham, S., Mwadzingeni, L., Tsilo, T.J., 2017. *Fusarium* head blight of wheat: Pathogenesis and control strategies. *Crop Protection* 91, 114-122. <https://doi.org/10.1016/j.cropro.2016.10.002>

**Goudemand E. et al.**, « Association mapping and meta-analysis: two complementary approaches for the detection of reliable *Septoria tritici* blotch quantitative resistance in bread wheat (*Triticum aestivum* L.) », *Mol. Breed.*, vol. 32, no 3, p. 563-584, 2013, doi: <https://doi.org/10.1007/s11032-013-9890-4>. Goyeau, H., Lannou, C., 2011. Specific resistance to leaf rust expressed at the seedling stage in cultivars grown in France from 1983 to 2007. *Euphytica* 178, 45-62. <https://doi.org/10.1007/s10681-010-0261-5>

**Goyeau, H., Lannou, C., Pariaud, B., Azzimonti, G., Du Cheyron, P., Couleaud, G., Maumené, C.**, 2010. Blé et résistance variétale, le cas de la rouille brune. La résistance génétique du blé à la rouille brune causée par *Puccinia triticina* : comment évaluer sa durabilité pour mieux la gérer ?

- Karlstedt F., Kopahnke D., Perovic D., Jacobi A., Pillen K., et Ordon F.**, « Mapping of quantitative trait loci (QTL) for resistance against *Zymoseptoria tritici* in the winter spelt wheat accession HTRI1410 (*Triticum aestivum* subsp. *spelta*) », *Euphytica*, vol. 215, no 6, p. 1 15, 2019, doi: <https://doi.org/10.1007/s10681-019-2432-3>. **Kollers et al.**, « Genetic architecture of resistance to *Septoria tritici* blotch (*Mycosphaerella graminicola*) in European winter wheat », *Mol. Breed.*, vol. 32, no 2, p. 411 423, 2013, doi: <https://doi.org/10.1007/s11032-013-9880-6>.
- Langlands-Perry C. et al.**, « Resistance of the Wheat Cultivar 'Renan' to *Septoria* Leaf Blotch Explained by a Combination of Strain Specific and Strain Non-Specific QTL Mapped on an Ultra-Dense Genetic Map », *Genes*, vol. 13, p. 100, 2022, doi: <https://doi.org/10.3390/genes13010100>.
- Kumar, S., Bhardwaj, S.C., Gangwar, O.P., Sharma, A., Qureshi, N., Kumaran, V.V., Khan, H., Prasad, P., Miah, H., Singh, G.P., Sharma, K., Verma, H., Forrest, K.L., Trethowan, R.M., Bariana, H.S., Bansal, U.K.**, 2021. Lr80: A new and widely effective source of leaf rust resistance of wheat for enhancing diversity of resistance among modern cultivars. *Theor Appl Genet* 134, 849–858. <https://doi.org/10.1007/s00122-020-03735-5>
- Liu, S., Hall, M.D., Griffey, C.A., McKendry, A.L.**, 2009. Meta-Analysis of QTL Associated with Fusarium Head Blight Resistance in Wheat. *Crop Science* 49, 1955–1968. <https://doi.org/10.2135/cropsci2009.03.0115>
- Löffler, M., Schön, C.-C., Miedaner, T.**, 2009. Revealing the genetic architecture of FHB resistance in hexaploid wheat (*Triticum aestivum* L.) by QTL meta-analysis. *Mol Breeding* 23, 473–488. <https://doi.org/10.1007/s11032-008-9250-y>
- Mergoum et al.** (2013) <https://doi.org/10.1556/CRC.2013.0005>
- Mekonnen T. et al.**, « Genome-wide association study reveals novel genetic loci for quantitative resistance to *septoria tritici* blotch in wheat (*Triticum aestivum* L.) », *Front. Plant Sci.*, p. 1970, 2021, doi: <https://doi.org/10.3389/fpls.2021.671323>.
- Niks, R.E., Qi, X., Marcel, T.C.**, 2015. Quantitative Resistance to Biotrophic Filamentous Plant Pathogens: Concepts, Misconceptions, and Mechanisms. *Annual Review of Phytopathology* 53, 445–470. <https://doi.org/10.1146/annurev-phyto-080614-115928>
- Piaskowska et al.** (2021) <https://doi.org/10.3390/agronomy11061108>
- Radecka-Janusik M. et Czembor P. C.** « Genetic mapping of quantitative trait loci (QTL) for resistance to *Septoria tritici* blotch in a winter wheat cultivar Liwilla », *Euphytica*, vol. 200, no 1, p. 109 125, 2014, doi: <https://doi.org/10.1007/s10681-014-1157-6>.
- Riaz et al.** « Genetic analysis using a multi-parent wheat population identifies novel sources of *septoria tritici* blotch resistance », *Genes*, vol. 11, no 8, p. 887, 2020, doi: <https://doi.org/10.3390/genes11080887>.
- Rollar, S., Serfling, A., Geyer, M., Hartl, L., Mohler, V., Ordon, F.**, 2021. QTL mapping of adult plant and seedling resistance to leaf rust (*Puccinia triticina* Eriks.) in a multiparent advanced generation intercross (MAGIC) wheat population. *Theor Appl Genet* 134, 37–51. <https://doi.org/10.1007/s00122-020-03657-2>
- Safavi, S.A., Afshari, F.**, 2013. Virulence factors of *Puccinia triticina* on wheat and effectiveness of Lr genes for leaf rust resistance in Ardabil. *Archives of Phytopathology and Plant Protection* 46, 1246–1254. <https://doi.org/10.1080/03235408.2013.764055>
- Saintenac, C., Cambon, F., Aouini, L., Verstappen, E., Ghaffary, S.M.T., Poucet, T., Marande, W., Berges, H., Xu, S., Jaouannet, M., Favery, B., Alassimone, J., Sánchez-Vallet, A., Faris, J., Kema, G., Robert, O., Langin, T.**, 2021. A wheat cysteine-rich receptor-like kinase confers broad-spectrum resistance against *Septoria tritici* blotch. *Nat Commun* 12, 433. <https://doi.org/10.1038/s41467-020-20685-0>
- Saintenac, C., Lee, W.-S., Cambon, F., Rudd, J.J., King, R.C., Marande, W., Powers, S.J., Bergès, H., Phillips, A.L., Uauy, C.**, 2018. Wheat receptor-kinase-like protein *Stb6* controls gene-for-gene resistance to fungal pathogen *Zymoseptoria tritici*. *Nature genetics* 50, 368.
- Stadlmeier M., Jørgensen L. N., Corsi B., Cockram J., Hartl L., et Mohler V.**, « Genetic dissection of resistance to the three fungal plant pathogens *Blumeria graminis*, *Zymoseptoria tritici*, and *Pyrenophora tritici-repentis* using a multiparental winter wheat population », *G3 Genes Genomes Genet.*, vol. 9, no 5, p. 1745 1757, 2019, doi: <https://doi.org/10.1534/g3.119.400068>.
- Yang N., Ovenden B., Baxter B., McDonald M. C., Solomon P. S., et Milgate A.**, « Multi-stage resistance to *Zymoseptoria tritici* revealed by GWAS in an Australian bread wheat (*Triticum aestivum* L.) diversity panel », *bioRxiv*, 2022, doi: <https://doi.org/10.1101/2022.06.29.498182>.
- Yates S. et al.**, « Precision phenotyping reveals novel loci for quantitative resistance to *septoria tritici* blotch », *Plant Phenomics*, vol. 2019, 2019, doi: <https://doi.org/10.34133/2019/3285904>.
- Zheng, T., Hua, C., Li, L., Sun, Z., Yuan, M., Bai, G., Humphreys, G., Li, T.**, 2021. Integration of meta-QTL discovery with omics: Towards a molecular breeding platform for improving wheat resistance to *Fusarium* head blight. *The Crop Journal* 9, 739–749. <https://doi.org/10.1016/j.cj.2020.10.006>

# ExIGE : Evaluation multi-environnementale de blé tendre « exotique » - prédictions génomiques GxÉ et génétique d'association

Justin BLANCON<sup>1</sup>, Jonathan KITT<sup>1</sup>, Pauline LASSERRE-ZUBER<sup>1</sup>, David ALVAREZ<sup>1</sup>, Sibille PERROCHON<sup>1</sup>, Jean-Noël THAUVIN<sup>2</sup>, Céline DUQUE<sup>3</sup>, Sylvie DUTRIEZ<sup>4</sup>, Pascal GIRAudeau<sup>5</sup>, Ellen GOUEMAND-DUGUE<sup>6</sup>, Emmanuel HEUMEZ<sup>7</sup>, Christophe MICHELET<sup>2</sup>, Patrice SENELLART<sup>8</sup>, Alexis OGER<sup>9</sup>, François BALFOURIER<sup>1</sup>, Catherine RAVEL<sup>1</sup>, Sophie BOUCHET<sup>1\*</sup>

- 1 - INRAE Université Clermont-Auvergne, UMR 1095, GDEC, 5 chemin de Beaulieu, 63100 Clermont-Ferrand, FRANCE
- 2 - RAGT, Rue Emile Singla, BP 3331 120333 Rodez Cedex 9, FRANCE
- 3 - Limagrain Europe, Centre de recherche de Chappes, 63720 Chappes, FRANCE
- 4 - Lidea Seeds, avenue Gaston Phoebus, 64230 Lescar, FRANCE
- 5 - Secobra Recherches, Centre de Bois-Henry, 78580 Maule, FRANCE

- 6 - FLORIMOND DESPREZ VEUVE & FILS, 59242 Cappel-en-Pévèle, FRANCE
  - 7 - INRAE UE GCIE, 2 chaussée Brunehaut, Estrées Mons, BP50136,80203 Peronne Cedex, FRANCE
  - 8 - KWS MOMONT SAS, 7 Rue de Martinal, 59246 Mons-en-Pévèle, FRANCE
  - 9 - Syngenta France SA, 1228 Chemin de l'Hobit, 31790 Saint-Sauveur, FRANCE
- \*Coordinateur : Sophie BOUCHET, sophie.bouchet@inrae.fr

## Introduction

Face à l'instabilité et la stagnation des rendements de blé tendre, il est nécessaire de réintroduire une diversité utile en sélection pour améliorer la tolérance aux stress biotiques et abiotiques mais aussi la qualité boulangère. En raison de l'impact majeur des interactions GxÉ, ces réintroductions doivent s'opérer dans les environnements adéquats pour maximiser le gain génétique. Dans le projet ExIGE, nous avons caractérisé un panel de diversité mondiale enrichi en accessions asiatiques (Figure 4), au champ (rendement, maladies, variables liées à la qualité boulangère) et au laboratoire (SNP Breedwheat, RNaseq DeBAT). L'objectif principal était de calibrer un modèle de prédiction génomique GxÉ permettant de prédire de nouveaux individus dans de nouveaux environnements pour le rendement. Le deuxième objectif est d'identifier les déterminants génétiques des caractères d'intérêt majeur du blé tendre (rendement, maladies, qualité boulangère).

## Réseau d'essais

Au cours de ce projet nous avons complété un jeu de données phénotypiques issu de 12 essais menés lors du projet PIA BreedWheat en conditions fort vs faible azote ou pluvial vs irrigué, avec 12 nouveaux essais dans des conditions similaires (Figure 1). Le réseau constitué a permis d'observer une large gamme de rendement, avec des essais assez peu corrélés, mais sans structure d'interaction marquée.

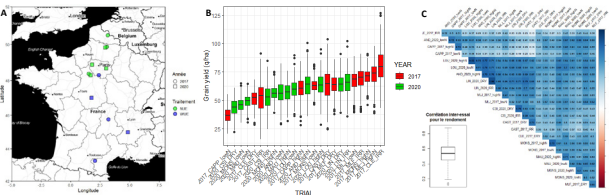


Figure 1 - (A) Répartition du réseau d'essais en France. La forme des points indique l'année et la couleur la contrainte abiotique appliquée ; (B) Gamme de variation pour le rendement. La couleur indique l'année ; (C) Matrice de corrélation entre chaque essai pour le rendement. Le boxplot montre la gamme de corrélation (r)

## Caractérisation environnementale

Les données météorologiques, de description du sol et d'itinéraire technique ont été utilisées pour simuler 192 des indices environnementaux à l'aide du modèle éco-physiologique CHN (Arvalis). Ces indices ont été estimés par phase de développement à l'aide des données mesurées pour l'épave. L'évolution du confort hydrique et azoté a été estimée tout au long du cycle (Figure 2) et montre que la plupart des essais a subi un stress hydrique nul, faible ou tardif, à l'exception de GLE\_2017\_DRY, MAU\_2020, MONS\_2017 et MONS\_2020. Le stress azoté a également été très limité dans le réseau et principalement présent sur les essais MONS\_2017\_lowN et MONS\_2020\_lowN.

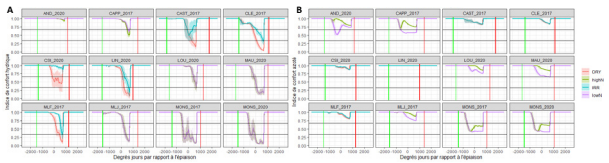


Figure 2 - Evolution de l'indice de confort hydrique (A) et azoté (B) au cours du cycle de la plante. Une valeur de 1 indique le confort et une valeur inférieure indique un stress : faible au-dessus de 0,66, modéré entre 0,66 et 0,33 et fort en-dessous

## Prédictions génomiques GxÉ

Une approche de prédiction génomique basée sur un modèle AMMI (Additive Main effect and Multiplicative Interaction, (Gollob 1968)) a été évaluée à l'aide de 8 modèles différents (Table 1). Ces modèles se différencient par la modélisation de la matrice d'interactivité génétique. Ils utilisent une matrice Kinship (modèles 1, 2, 3, 4 et 6), une matrice issue de l'approche AMMI (modèles 5, 6, 7 et 8), une matrice construite sur les indices environnementaux (modèles 3, 4 et 5). La predictive ability de ces modèles estimée par cross-validation dans différents scénarios de prédiction (Table 2) indique que l'approche AMMI semble prometteuse avec un gain théorique de prédictive ability de +0,2 en moyenne dans les 4 scénarios par rapport au GBLUP. Mais en pratique le gain reste minime (gain maximum de 0,08) car les matrices d'interactivité génétique et environnementale sont difficiles à prédire.

Modèles	ajusté	Sommaire de Cross-validation	ajusté
Modèle 1 (K)	0,672	0,662	0,656
Modèle 2 (K)	0,676	0,652	0,640
Modèle 3 (K)	0,713	0,702	0,690
Modèle 4 (K)	0,730	0,714	0,706
Modèle 5 (AMMI)	0,692	0,686	0,682
Modèle 6 (AMMI)	0,702	0,705	0,712
Modèle 7 (AMMI)	0,675	0,702	0,706
Modèle 8 (AMMI)	0,713	0,702	0,692

Table 1 (gauche) - Liste des 8 modèles comparés et modélisation de la variance d'interaction GxÉ correspondante

Table 2 (droite) - Qualité de prédiction pour les 8 modèles testés. Les scénarios sont oGoE (old Genotype & old Environment), nGoE (new Genotype & old Environment), oGoE (old Genotype & new Environment) et nGoE (new Genotype & new Environment)

## GWAS - Caractères agronomiques

Une analyse d'association a été réalisée avec le modèle K par l'approche LOCO (Rincent et al., 2014). Au total, 272 QTL ont été détectés, répartis sur l'ensemble des chromosomes (Figure 3). La qualité de l'analyse est confirmée par l'observation d'importantes colocalisations au niveau de gènes majeurs, comme PPD-D1 pour la précocité. Pour le rendement, 9 QTL ont été identifiés. Une colocalisation de QTL de composantes du rendement, indépendants de la précocité et de la hauteur semble intéressante sur le chromosome 7A et explique en moyenne 6,7% de la variance.

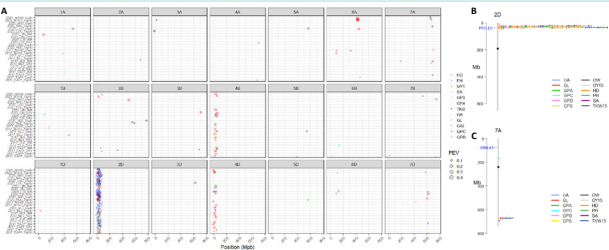


Figure 3 - (A) Répartition des QTL détectés sur l'ensemble du génome pour chaque caractère dans chaque essai ; (B) Colocalisation de QTL de précocité avec PPD-D1 ; (C) Colocalisation de QTL de composantes du rendement sur le 7A. Les caractères étudiés sont la précocité (PD), la hauteur (PH), le rendement à 15% d'humidité (GY15), le nombre d'épis par m<sup>2</sup> (SA), le nombre de grains par épi (GPE), le nombre de grains par m<sup>2</sup> (GPA), le poids de mille grains (TKW15), la surface, la longueur et la largeur du grain (GA, GL, GW), la teneur en protéine (GPC) et le Grain Protein Deviation (GPD).

## Phénotypage protéines de réserve / qualité boulangère

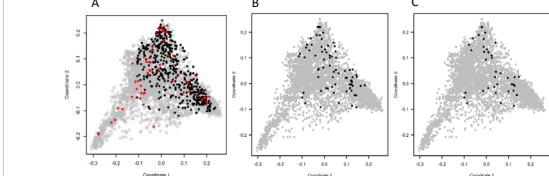


Figure 4 - (A) panel BWP3 étendu (485), (B) panel gluténines (73), (C) panel alvéographe (45) Parmi les 4500 accessions Breedwheat, 450 correspondent à BWP3 (noires sur Figure 4A), 20 asiatiques et 25 présentant des prédictions génomiques intéressantes ont été ajoutées (rouges sur Fig. 4A) pour former BWP3 étendu. Parmi elles, 73 accessions représentatives ont été dosées et séquencées pour les protéines de réserve (noires sur Figure 4B). Parmi elles, 45 ont été évaluées pour les variables d'alvéographe en conditions optimales et stressées (azote: N/NO, irrigation: DRY/IRR) (noires sur Figure 4C).

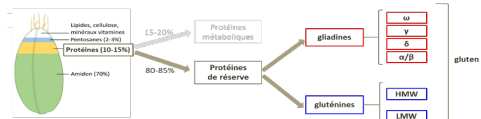


Figure 5 - Composition du grain de blé tendre, focus sur les protéines de réserve

## Séquençage des gluténines

L'annotation des principales protéines de réserve (Figure 5), 18 gènes de gluténines et 44 gènes de gliadines, a été réalisée chez Chinese-Spring V2.1. Des fragments de 4500 à 8500 pb ont été amplifiés. Seules les amorces de gluténines ont fonctionné et permis le séquençage de 6 gènes de gluténines de haut poids moléculaire (HMW en fin des chromosomes du groupe 1) et 12 de faible poids moléculaire (LMW en début des chromosomes du groupe 1). Parmi les 2102 SNP, 1309 ont une MAF>0.05 et 53 sont en équilibre de liaison.

On remarque que l'allèle « absent » est un allèle fréquent à considérer dans les associations avec les phénotypes (Table 2).

On observe du DL intra-gène. Au total, 25 marqueurs permettent de résumer la diversité des HMW et 29 pour les LMW. Chaque gène possède en moyenne 4 haplotypes (Figure 6).

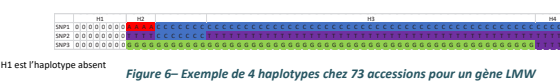


Figure 6 - Exemple de 4 haplotypes chez 73 accessions pour un gène LMW

## GWAS - Protéines de réserve

### Table 2 - Associations protéines de réserve

te: élasticité, L: extensibilité, protTot: protéine totale, glu: gluténine, gli: gliadine, N: azote total

Trait	#Chr	Chr	Position (CS V2.1, bp)	Pvalue	part variance	QTL
te	1	1A	509 017 352	9.40E-07	0.49	Glu_HMW-A
L	5	2B	745 716 941	2.19E-07	0.56	
L	15	5D	458 216 820	4.84E-07	0.52	CWi-5D
protTot	2	1B	94 458 186	6.01E-09	0.37	1RS/1BL(gli/LMW)
protTot	2	1B	105 538 550	6.01E-09	0.37	1RS/1BL(gli/LMW)
gli_w5	3	1D	578 541	3.33E-09	0.39	gli/LMW
gliTot	7	3A	724 660 462	6.02E-07	0.29	YGWS
Nfarine	8	3B	373 613 528	4.50E-07	0.32	TalcY-B1
Nfarine	8	3B	774 409 784	4.32E-07	0.32	Red-3B
Nfarine	10	4A	263 086 190	4.72E-07	0.32	introgression dicoccoides
gli_w5	11	4B	3 133 117	6.23E-09	0.38	Lev-B1
gli_w1.2	13	5A	476 434 666	6.77E-07	0.30	nac27
gli_w5	14	5B	641 046 239	8.76E-07	0.30	SHP
gli_w5	14	5B	703 647 253	7.25E-09	0.38	
gli_w5	17	6B	10 754 025	7.96E-09	0.39	
gli_w5	17	6B	514 027 602	4.38E-07	0.30	SAD

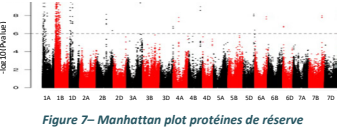


Figure 7 - Manhattan plot protéines de réserve

Toutes les variables liées aux protéines de réserve ont été projetées sur le même manhattan plot pour avoir une vision globale des QTL détectés

Au total, 16 QTL principaux ont été détectés (Table 2). Ils sont en général associés à plusieurs caractères dans plusieurs environnements et conditions. Ils expliquent entre 30 et 50% de la variance du phénotype. De nombreuses associations sont trouvées dans les régions des gliadines et LMW au début des chromosomes 1A, B et D (1,2 et 3) (Figure 7). L'introgression 1RS/1BL du seigle sur le chromosome 1B est présente dans le panel, ce qui explique l'étendue du QTL sur le bras court.

## GWAS - maladies

Au total, 54 QTL principaux ont été détectés sur la rouille jaune, la septoriose et la fusariose. Ils sont retrouvés dans plusieurs environnements et conditions. Ils expliquent entre 4 et 11% de la variance de la maladie. Les QTL les plus intéressants se situent sur les chromosomes 2B (5) et 6A (16) pour la rouille jaune (Figure 8). Certaines accessions présenteraient une introgression au niveau du chromosome 2B.

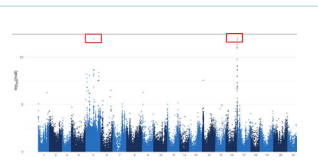


Figure 8 - Manhattan rouille jaune

L'association a été faite sur les résidus d'un modèle avec hauteur et précocité en variable explicative.

## Perspectives

Un jeu de données de qualité de génotypes et de phénotypes a été produit sur un large panel de diversité (485). Des analyses poussées ont été faites sur les données rendement. Les analyses sont encore en cours pour les autres caractères. Un travail d'imputation et de méta-analyses avec d'autres jeux de données sera nécessaire pour valider les QTL, en particulier pour les caractères où nous avons évalué seulement 45 (alvéographe) ou 75 (protéines de réserve) individus. L'objectif est de détecter des gènes intéressants et de constituer un set de marqueurs diagnostiques pour optimiser le pre-breeding.