# Synthèse des programmes de recherche 2014

Jeudi 21 Mars 2019





# Fonds de Soutien à l'Obtention Végétale



Depuis sa création en 2001 jusqu'à aujourd'hui, le Fonds de Soutien à l'Obtention Végétale (FSOV) a soutenu 92 programmes de recherche et développement s'intéressant à toutes les céréales à paille : blé tendre, blé dur, orge, avoine, seigle, triticale, épeautre et riz.

Cela fait maintenant 18 ans que le FSOV contribue au développement d'innovations et à la création de nouvelles variétés performantes, économes en intrants, adaptées

à leur environnement et conduisant à une production de haute qualité technologique, conformément aux attentes de notre filière, du marché et de notre société.

Vous pourrez étudier les résultats des programmes retenus lors de l'appel à projets de 2014 au travers de ce recueil et constater l'impact de ces travaux pour le développement à court terme de nouveaux outils technologiques, moléculaires ou méthodologiques performants et innovants.

Dans un contexte politique et réglementaire où la solution aux défis de demain repose de plus en plus sur les progrès de la génétique et de l'amélioration variétale, le FSOV tient une place unique au cœur de notre filière. En effet, au-delà du rôle majeur joué par le Fonds dans l'innovation variétale, le FSOV symbolise aussi la volonté et l'engagement de notre filière depuis de nombreuses années déjà, à promouvoir le progrès génétique.

François JACQUES Président du Comité d'engagement du FSOV



Section Semences de Céréales à paille & Protéagineux 44 rue du Louvre - 75001 Paris Tél. : 0142 33 85 05 - E-mail : fsov@gnis.fr www.fsov.org

# SOMMAIRE

**FSOV 2014 A : [B-Dul]** Blé Durable de Qualité : Sélection de blés productifs de qualité sous contrainte de nutrition azotée

**FSOV 2014 B : [Pmg-TILL]** Identification des gènes impliqués dans le Poids de Mille Grains par une approche innovante (TILLING)

**FSOV 2014 C : [OWBM2]** Identification et validation de nouveaux marqueurs étroitement liés au gène Sm1 de résistance à la cécidomyie orange du blé tendre

- **FSOV 2014 D : [Microdochium spp.]** Vers une meilleure connaissance de l'occurrence, de l'épidémiologie du champignon et du comportement des variétés de blé tendre actuelles face à cette maladie
- **FSOV 2014 E : [NIL-N]** Caractérisation de régions chromosomiques pour augmenter l'efficacité d'utilisation de l'azote et la teneur en protéines

**FSOV 2014 J : [N-BTH]** Méthodes d'estimation des indicateurs d'efficacité de valorisation de l'azote par les nouvelles variétés de blé tendre

**FSOV 2014 K : [SG-Rendement]** Calibration et Implémentation d'un outil de Sélection Génomique Rendement dans un programme de sélection blé tendre

- **FSOV 2014 L : [HeatWheat]** Analyse de la diversité génétique de la réponse au stress thermique via phénotypage fin et génétique d'association
- **FSOV 2014 N : [GS-qualité]** Établissement d'un modèle de Sélection Génomique pour la Qualité Boulangère des blés
- **FSOV 2014 P : [WEAB]** Sélection Assistée par les Effecteurs fongiques de Résistances aux champignons pathogènes chez le blé

FSOV 2012 J: [Synthetics] Integrating New Diversity to improve Yield Stability (INDYS)

# [B-Dul] Blé Durable de Qualité : Sélection de blés productifs de qualité sous contrainte de nutrition azotée

Ellen GOUDEMAND-DUGUE<sup>1</sup>\*, Jérôme AUZANNEAU<sup>2</sup>, Pierre LEMEUNIER<sup>123</sup>, François-Xavier OURY<sup>3</sup>, Jacques BORDES<sup>3</sup>, Thierry DEMARQUET<sup>1</sup>, Davis ALVAREZ<sup>3</sup>, Sibille PERROCHON<sup>3</sup>, Mireille DARDEVET<sup>3</sup>, Catherine RAVEL<sup>3</sup>

1 - FLORIMOND DESPREZ VEUVE & FILS, BP41 59242 Cappelle-en-Pévèle

2 - AGRI OBTENTIONS, 78280 GUYANCOURT

3 - INRA, UMR 1095 INRA UCA GDEC, 5 chemin de Beaulieu 63100 Clermont-Ferrand

\* Coordinateur : Ellen GOUDEMAND-DUGUE, ellen.goudemand@florimond-desprez.fr

# 1. Introduction

Le blé tendre (Triticum aestivum L.) est une ressource essentielle pour l'alimentation humaine mondiale. La gualité d'usage de la farine de blé est définie par sa capacité à être transformée en une pâte aux propriétés rhéologiques appropriées, grâce à ses protéines de réserve, les gliadines et les gluténines qui forment le gluten. Les protéines de réserve, riches en acides aminés soufrés, à l'origine du gluten, constituent 80% des protéines totales présentes dans le grain (Shewry, 2009). Les gluténines représentent le plus couramment 35 à 50% des protéines du grain. Elles sont divisées en gluténines de hauts et faibles poids moléculaires. Les gliadines sont des monomères divisés en  $\omega$ 5-,  $\omega$ 1,2-,  $\alpha/\beta$ - et  $\gamma$ gliadines (Wieser, 2007). Elles représentent entre 30 et 40% des protéines du grain. Les protéines de réserve sont à l'origine du polymère de gluten. Les gluténines sont liées par des liaisons covalentes (des ponts disulfures). Les gliadines sont fixées sur ces polymères par des liaisons faibles (liaisons hydrogènes...). Les procédés de transformation secondaire du blé vont conduire à une restructuration de ce réseau : lors du mélange de la farine avec de l'eau, des liaisons disparaissent, de nouvelles se forment, ce qui donnera de nouvelles propriétés viscoélastiques (Schiedt et al., 2013 ; Ortolan et Steel, 2017). Les gluténines, et particulièrement les sous-unités de haut poids moléculaire, confèrent l'élasticité à la pâte alors que les gliadines apportent la viscosité (MacRitchie, 1999 ; Shewry et al., 2002). La qualité de la pâte résulte de la balance entre ces deux propriétés, due aux proportions de gliadines et gluténines (d'où l'importance du ratio gliadines/gluténines (gli/glu) pour la qualité technologique de la pâte). De plus, depuis 1979, le rôle clef des gluténines de haut poids moléculaire dans la force de la pâte a été démontré (Payne et al., 1979). Une augmentation des gliadines par rapport aux gluténines va réduire la force de la pâte, son élasticité, sa stabilité au cours du temps ainsi que le temps nécessaire pour qu'elle acquière ses propriétés optimales. Le plus souvent, des pâtes fortes (faible rapport gli/glu) sont souhaitées pour la boulangerie et des pâtes plus visqueuses (fort gli/glu) pour la fabrication de pâtisseries.

Si le rapport gliadines sur gluténines est important pour la qualité technologique de la farine, sa prise en compte dans les programmes de sélection n'est pas acquise. En effet, ce trait est difficile et couteux à mesurer car son phénotypage nécessite plusieurs étapes pour extraire et doser ces protéines (Plessis *et al.*, 2013). L'introduction de ce caractère en sélection repose donc sur des méthodes alternatives basée sur l'utilisation de marqueurs comme la sélection assistée par marqueurs (SAM) ou la sélection génomique (SG) qui permettrait de le prédire.

Quelques études ont été réalisées pour établir le déterminisme génétique (régions chromosomiques impliquées dans le caractère ou QTL et allèles favorables qui peuvent ensuite être sélectionnés). C'est le cas d'une étude de génétique d'association à l'échelle du génome (GWAS, Genome Wide Association Study) réalisée par Plessis *et al.* (2013). Ces auteurs ont montré que le QTL qu i contrôlait la plus grande variation du rapport gli/glu dans le grain expliquait seulement moins de 10% de la variabilité totale du caractère. Dans ce contexte, la SG serait une alternative plus adaptée pour prédire ce rapport, car capable de prendre en compte un nombre important de QTL à effet faible impliqués dans un caractère (Meuwissen, 2001). Les modèles de SG sont établis à partir d'une population d'entrainement (TP), génotypée et phénotypée pour le caractère d'intérêt. Ce modèle est ensuite appliqué pour prédire ce caractère au sein d'une population à sélectionner, uniquement génotypée. Ajouter certains QTL majeurs en effet fixe au modèle de SG peut améliorer sa précision (Bernardo, 2014 ; Spindel *et al.*, 2016 ; Michel *et al.*, 2018).

Dans ce projet, d'une part, des marqueurs moléculaires, facilement utilisables d'un laboratoire à un autre, ont été développés pour permettre aux sélectionneurs d'intégrer ces caractères dans leurs programmes. Ces marqueurs résultent de QTL identifiés par GWAS. D'autre part, la capacité de prédiction du rapport gli/glu a été étudiée en utilisant la SG. Différentes possibilités d'amélioration ont été étudiées comme l'optimisation de la population d'entrainement ou l'ajout de QTL majeur dans le modèle. C'est cette seconde partie du travail que nous avons choisie de présenter.

# 2. Matériel et méthode

#### Matériel végétal et expérimentation

Trois populations de blé tendre d'hiver ont été utilisées. Ces populations ont été phénotypées dans différents environnements (lieux et années). La première population, CC196, est composée de 196 génotypes issus de la core collection de 372 accessions établie par Balfourier et al. (2007). Ces 196 génotypes sont les plus courts et les plus productifs de la collection. La moitié d'entre eux ont été inscrits après 1960 (Bordes et al. 2013). Le panel CC196 a été phénotypé à Clermont-Ferrand en 2005-2006 et au Moulon en 2006-2007. La seconde population est constituée de 186 génotypes, issus de la population BWP3 du projet Breedhweat. La population BWP3 est une core collection comprenant les variétés les plus modernes sélectionnées au sein des 4600 accessions du projet Breedwheat représentant la diversité mondiale. La population BWP3 a été phénotypé en 2016-2017 à Clermont-Ferrand. Enfin, la troisième population BDUL est composée de 94 génotypes élites en cours de sélection fournis par les sociétés Agri-Obtentions et Florimond Desprez. Ces génotypes étaient en 2015 soit dans le cycle d'inscription, soit un ou deux ans avant l'inscription. Le panel BDUL a été phénotypé en 2015-2016 dans trois lieux (Clermont, Cappelle-en-Pévèle et Orsonville) dans le cadre du projet FSOV. Le panel BWP3 partage 8 génotypes avec le panel CC196 et 1 génotype avec BDUL. Les panels CC196 et BDUL ne partagent aucun génotype.

Dans tous les essais, les blés ont été cultivés dans des conditions représentatives de ce qui est fait en France.

Site	Year									
	2006	2007	2016	2017						
CF	CC196		BDul	BWP3						
LM		CC196								
Or			BDul							
CeP			BDul							

Tableau 1 : Répartition des collections dans les six environnements.

#### Quantification des gliadines et des gluténines

Une extraction séquentielle des gliadines et des gluténines a été réalisée à partir de 5mg de farine de chaque génotype selon la méthode décrite dans Plessis *et al.* (2013) pour le panel CC196. Pour les panels BWP3 et BDUL, le protocole d'extraction a été ajusté en réduisant le pH et la solution tampon pour améliorer l'extraction des gliadines. Une extraction par génotype et par environnement a été effectuée. La teneur en azote dans chaque fraction (gliadines puis gluténines) a été dosé avec un analyseur élémentaire de type DUMAS (Thermo Electron). Une multiplication par 5.7 de ces teneurs a permis de les « transformer » en % gliadines ou % de gluténines. Ces dernières ont servi à calculer le rapport gli/glu.

#### Données de génotypage

Chaque lignée a été génotypée avec la puce Affymetrix Axiom 420k (Rimbert *et al.* 2018). Les marqueurs OTV, les marqueurs avec une MAF (minor allele frequency) inférieure à 0.1, les marqueurs avec un taux de données manquantes supérieur à 0.2 et les marqueurs avec un taux d'hétérozygotie supérieur à 0.15 ont été supprimés. Les génotypes avec un taux de données manquantes supérieur à 0.1 et un taux d'homozygotie supérieur à 0.05 ont été supprimés. Après l'application de ces filtres, le jeu de données était composé de 166 génotypes dans CC196, 167 génotypes dans BWP3, 88 génotypes dans BDUL et 101 402 marqueurs. Les allèles de chaque marqueur ont été codés 1 ou -1 pour chaque allèle présent à l'état homozygote, et 0 pour les hétérozygotes. Les données manquantes ont été imputées par la moyenne de chaque marqueur.

#### Evaluation de la qualité du modèle prédictif

Les modèles de SG testés prédisent trois effets : génotype, environnement et génotype x environnement. Le but de cette étude est d'évaluer si les modèles proposés permettent une bonne prédiction de la valeur génétique du rapport gli/glu, quel que soit l'environnement. De ce fait, les valeurs prédites doivent être comparées avec les vraies valeurs génétiques des individus. Les modèles de SG testés ont été évalués selon deux critères : la précision et la RMSE (root mean square error). La précision est définie comme la corrélation entre les valeurs prédites par le modèle et les valeurs mesurées du caractère. La RMSE est la distance quadratique entre les valeurs mesurées et prédites. Ce critère reflète la capacité du modèle à correctement prédire la valeur de l'ensemble des individus, pas uniquement un classement comme c'est le cas de la précision.

Précision et les RMSE ont été estimés par validation croisée. CC196 et BW3 sont systématiquement utilisés dans la population d'entrainement pour calibrer les modèles. On y ajoute 2/3 de la population BDUL, qui contient les génotypes les plus récents. La population de validation sera constituée du dernier tier de la population BDUL. Le processus de validation croisée a été répété 60 fois afin de calculer une moyenne et une variance de la précision et de la RMSE.

#### Les modèles de sélection génomique testés

Le premier modèle de sélection génomique utilisé pour prédire le ratio gli/glu est le modèle G-BLUP. Les QTL influençant le ratio gli/glu ont été détecté par Zhang *et al.* (2011) et Plessis *et al.* (2013). Nous avons donc exploré une possible amélioration du modèle de prédiction en ajoutant les effets de QTL au modèle G-BLUP. Ce modèle a été intitulé G-BLUP + de novo GWAS comme dans la publication de Spindel *et al.* (2016) et la procédure utilisée est la suivante :

- Modèle GWAS : test des 101 402 SNP avec le modèle GBLUP établi avec la population d'entrainement.
- Sélection des SNP les plus significatifs (FDR =0.5%)
- Regroupement des SNP les plus significatifs en fonction de leur position sur les chromosomes. Les SNP distants de moins de 6000 kb sont considérés comme appartenant au même QTL.
- Sélection d'un SNP représentatis par QTL, celui avec la plus petite Pvalue.
- Ajout de ces SNP en effet fixe au modèle G-BLUP
- Calcul de la précision et de la RMSE de ce modèle G-BLUP + de novo GWAS.

#### Optimisation de la population d'entrainement

Plusieurs études ont démontré qu'un fort apparentement entre la population d'entrainement et la population à sélectionner permet d'obtenir de meilleures précisions de prédiction (Albrecht et al., 2011 ; Ly et al., 2013). Si une entreprise de sélection souhaite prédire un caractère par sélection génomique, elle doit construire une population d'entrainement avec des variétés issues de son propre programme de sélection. Si une entreprise possède une grande diversité de variétés génotypées et phénotypées pour le caractère d'intérêt alors, sélectionner seulement une partie d'entre elles peut permettre d'augmenter la capacité de prédiction du modèle. Pour tester cette hypothèse, nous avons comparé 5 méthodes d'optimisation de la population d'entrainement à une méthode de sélection aléatoire. Ces méthodes ont été testées en utilisant un modèle G-BLUP avec des populations d'entrainement de différentes tailles, allant de 20 génotypes au nombre maximum de génotypes disponibles, soit 382.

- La première méthode d'optimisation (OPT\_MEAN) consiste en la sélection des génotypes de la population d'entrainement les plus apparentés en moyenne aux génotypes de la population à sélectionner. Certains génotypes seront fortement apparentés et d'autres moins.
- La seconde méthode (OPT\_MAX) va sélectionner les génotypes avec l'apparentement maximal.
- La méthode OPT\_MIN va sélectionner des individus avec les plus hauts apparentements minimaux.
- La méthode d'optimisation CD\_MEAN va se baser sur le coefficient de détermination généralisé (CD) décrit par Laloë (1993), Rincent *et al.* (2012) et Rutkoski *et al.* (2015). Un CDmean proche de 1 signifie que les contrastes prédits entre la population d'entrainement et la population à sélectionner sont fortement précis et donc les précisions de prédiction seront bonnes. Afin d'optimiser la population d'entrainement avec le critère CDmean, nous avons utilisé le script proposé par Rincent *et al.* (2012) avec 3000 répétitions.
- La dernière méthode, OPT\_IND, optimise la population d'entrainement pour chaque génotype de la population à sélectionner en choisissant les individus les plus apparentés.

Les calculs ont été effectués sous R (R Core Team 2017). Le package Sommer a été utilisé pour la SG et la recherche de QTL pour la partie GWAS (Covarrubias-Pazaran, 2016).

### 3. Résultats

#### Apparentement entre les génotypes

L'apparentement entre les génotypes des différents panels est représenté sur l'analyse en composantes principales ci-dessous (Figure 1). L'axe 1 et l'axe 2 représentent 9.5% et 3.7% de la variance totale, respectivement. On peut clairement voir que l'apparentement entre les génotypes élites de BDUL est fort, avec une moyenne de 0.30. Les génotypes d'Agri-Obtentions sont plus apparentés aux core collections (BWP3 et CC196) que les génotypes Florimond Desprez. Les deux core collections sont relativement peu apparentées au panel BDUL, avec une moyenne de -0.09. Cependant, 9 et 3 génotypes des populations CC196 et BWP3 respectivement sont fortement apparentés à 1 ou plusieurs génotypes du panel BDUL (coefficient d'apparentement supérieur à 0.6).



Figure 1 : Analyse en composantes principales des génotypes des 3 collections.

#### Modèles de sélection génomique

Tout d'abord, les modèles G-BLUP et G-BLUP + de novo GWAS ont été comparés. De légères différences ont observées entre les deux modèles.

G-BLUP	Accuracy=0.543 (standard-deviar					
	(sd)=0.144), RMSE=	=0.0606 (sd=0.0043)				
G-BLUP + de	Accuracy=0.57	(sd=0.142),				
novo GWAS	RMSE=0.0600 (sd=	0.0042)				

Le modèle G-BLUP + de novo GWAS intègre entre 0 et 4 QTL détectés par GWAS. A part deux répétitions sur les 60, les QTL détectés été systématiquement les suivants :

- QTL sur le chromosome 1A, effet moyen = -0.0275, position =  $509\ 014\ 168\ pb$
- QTL sur le chromosome 1B, effet moyen = -0.0201, position entre 5 143 303 et 5 699 395 pb
- QTL sur le chromosome 7D, effet moyen = 0.0145, position = 571 785 471 pb

#### Méthodes d'optimisation

Cing méthodes d'optimisation de la population d'entrainement ont été testées et comparées à une méthode de sélection aléatoire. Le but était d'augmenter la capacité de prédiction en sélectionnant une population d'entrainement d'une taille limitée. Les précisions et les RMSE montrent que les méthodes d'optimisation de la population d'entrainement améliorent les capacités de prédiction, comparées à la méthode de sélection aléatoire (Figure 2). Toutes les méthodes de sélection montrent une amélioration de la capacité de prédiction lorsque la taille de la population d'entrainement augmente jusqu'à atteindre un seuil, minimisant la RMSE et maximisant la précision, qui se situe entre 80 et 140 génotypes. Une fois le maximum atteint, augmenter la taille de la population d'entrainement aboutit globalement à une diminution des capacités de prédiction jusqu'à atteindre une précision de 0.543 et une RMSE de 0.0606 lorsque le maximum d'individus est utilisé (382) (Figure 2).

La méthode de sélection aléatoire des individus présente une augmentation logarithmique de la précision (avec un minimum de 0.271 pour 20 génotypes dans la population d'entrainement) et une décroissance exponentielle de la RMSE (avec un maximum de 0.0818) avec la taille de la population. Les différences de capacité de prédiction entre les méthodes d'optimisation sont faibles. A 20 génotypes, la précision varie de 0.368 (OPT\_IND) à 0.482 (CD\_MEAN) mais la RMSE ne suit pas le même ordre en variant de 0.0723 (OPT\_MIN) à 0.0618 (OPT\_IND). La précision maximale (0.624) est obtenue avec la méthode OPT\_IND avec 80 génotypes mais les quatre autres méthodes montrent des résultats proches. La méthode CD-MEAN présente la particularité de conserver une capacité de prédiction constante de 60 à 280 génotypes. On observe également que la méthode OPT\_MAX présente également de faibles variations de la capacité de prédiction.



Figure 2 : Evolution de la précision (en haut) et du RMSE (en bas) selon la méthode de sélection de la population d'entrainement et sa taille. Méthode random (RDM) : les génotypes sont choisis au hasard.

# 4. Discussion

Le ratio gli/glu dans la farine de blé est un composant majeur de la qualité boulangère (Marchetti *et al.*, 2012 ; Barak *et al.*, 2014). Mais son phénotypage est cher et fastidieux. C'est pourquoi sa prédiction par sélection génomique pourrait être utilisée pour sélectionner plus précocement et plus précisément des blés de qualité.

# Le modèle G-BLUP est bien adapté pour prédire le ratio gliadine/gluténine

Le modèle G-BLUP qui a été appliqué prend en compte trois effets importants en sélection : le génotype, l'environnement et l'interaction génotype x environnement (Jarquin *et al.*, 2014). La covariance entre les effets des génotypes a été établie avec l'apparentement entre les individus, calculé avec les marqueurs. Cependant, aucune covariance n'a été considérée entre les environnements, qui aurait pu être calculée grâce à des covariables environnementales. Ces covariables influencent très certainement le ratio gli/glu et elles ajouteraient de l'information au modèle G-BLUP. Connaitre ces informations et les inclurent dans le modèle permettraient d'en augmenter la précision et donc la capacité de prédiction de l'effet génétique des individus.

La capacité de prédiction du modèle G-BLUP (précision de 0.543 et RMSE de 0.0606) était bonne comparée aux prédictions des caractères de qualité précédemment publiés. Avec un modèle RR-BLUP, Michel *et al.* (2018) ont obtenu des précisions entre 0.30 et 0.54 en fonction du caractère de qualité boulangère analysé. Cependant, la comparaison reste difficile puisque les structures des populations sont différentes.

Ajouter des QTL en effet fixe au modèle de sélection génomique peut améliorer la capacité de prédiction d'un caractère. Cela a déjà été démontré pour plusieurs caractères, comme les résistances aux maladies et la tolérance à la germination sur pied (Arruda et al., 2016; Moore et al. 2017). Cependant, Michel et al. (2018) et Liu et al. (2016) ont observé qu'ajouter des marqueurs en effet fixe au modèle de sélection génomique améliore très peu, voire pas du tout, la capacité de prédiction des caractères de qualité. Ces résultats sont confirmé dans notre étude avec notre modèle G-BLUP + de novo GWAS qui permet d'améliorer que très faiblement la précision du modèle G-BLUP, alors que la RMSE reste constante. Cela confirme les résultats obtenus par Bernardo (2014) qui a observé que les marqueurs doivent expliquer une part importante de la variance phénotypique (supérieure à 10%) pour être intégrés au modèle de sélection génomique et améliorer les prédictions. Le ratio gli/glu est un caractère fortement polygénique et donc les QTL trouvés ont des effets très faibles ce qui ne permet pas d'améliorer le modèle de sélection génomique.

Les modèles Bayesiens sont également une alternative au modèle G-BLUP testé. Cependant, ces modèles semblent plus efficaces que le modèle G-BLUP uniquement si des QTLs majeurs influencent le caractère (Clark *et al.* 2011 ; Heffner *et al.* 2011).

#### L'optimisation de la population d'entrainement permet d'aboutir à de meilleures capacités de prédiction

La grande diversité observée entre les génotypes des trois jeux de données implique que certains individus de la core collection sont génétiquement distants des génotypes modernes que nous souhaitions prédire. Inclure ces génotypes distants dans la population d'entrainement de notre modèle de prédiction peut diminuer la capacité à prédire la valeur des génotypes modernes. Pour contourner ce problème, nous avons testé cinq méthodes d'optimisation de la population d'entrainement. Toutes ces méthodes étaient basées sur l'apparentement entre les génotypes de la population à sélectionner et les individus de la population d'entrainement. Elles aboutissent à de meilleurs résultats que la sélection d'individus au hasard. La meilleure méthode d'optimisation sera celle qui sélectionnera une population d'entrainement avec la même diversité allélique que la population à sélectionner. On observe une diminution de la capacité de prédiction une fois que le maximum est atteint avec les méthodes d'optimisation. Cela s'explique par le fait que les derniers individus ajoutés sont trop distants génétiquement de la population à sélectionner, et ils apportent des allèles qui sont absent de la population à sélectionner. Isidro et al. (2015) suggéraient que capturer la plus grande partie de la variance phénotypique dans la population d'entrainement optimisée peut permettre d'augmenter la capacité de prédiction. Nos résultats vont à l'encontre de cette suggestion puisqu'ajouter des génotypes à la population d'entrainement augmente la variance phénotypique capturée mais diminue la capacité de prédiction. Ce qui certainement dû à notre objectif de prédire une population avec une faible diversité en comparaison avec la diversité de la population d'entrainement.

Les méthodes d'optimisation testées donnent des résultats similaires. Aucune n'apparait vraiment meilleure que les autres. On ne peut donc que conseiller d'utiliser la méthode la plus rapide en termes de calcul sachant que la sélection des individus et la demande computationnelle étaient différentes entre les méthodes (la mise en œuvre de OPT\_IND et CD-MEAN demande plus de 10 fois plus de temps que les autres méthodes). Rincent *et al.* (2017) ont testé les méthodes d'optimisation OPT\_MEAN et CD\_MEAN alors que lsidro *et al.* (2015 et Rutkoski *et al.* (2015) ont testé uniquement CD\_MEAN. Leurs résultats montrent que l'efficacité de la méthode d'optimisation est fortement dépendante du caractère et de la structure de la population. Notre objectif était de prédire du matériel élite avec un panel de diversité, ce qui est un cas particulier. Nos conclusions doivent donc être mises en perspective.

OPT\_IND aboutit à une relativement meilleure capacité de prédiction. Chaque génotype est prédit par son plus proche apparenté présent dans la population d'entrainement. La capacité de prédiction aurait été probablement bien meilleure que les autres méthodes si les génotypes de la population à sélectionner avaient été plus divers. Il n'en reste que cette méthode demande du temps de calcul.

Pour conclure, malgré le fait que nous ayons cherché à prédire du matériel élite avec des collection de diversité, le modèle G-BLUP développé montre des résultats satisfaisants ce qui montre l'intérêt de la sélection génomique pour la prise en compte des caractères de qualité dans les programmes de sélection.

### Références bibliographiques

Albrecht T, Wimmer V, Auinger HJ *et al.* (2011) Genome-based prediction of test - cross values in maize. Theor Appl Genet 123:339–350.

Arruda MP, Lipka AE, Brown PJ, et al. (2016) Comparing Genomic Selection and Marker-Assisted Selection for Fusarium Head Blight Resistance in Wheat (Triticum Aestivum L.). Molecular Breeding 36:84.

**Balfourier F, Roussel V, Strelchenko P et al.** (2007) A Worldwide Bread Wheat Core Collection Arrayed in a 384-Well Plate. Theor Appl Genet 114:1265-75.

Barak S, Mudgil D, Khatkar BS (2014) Influence of Gliadin and Glutenin Fractions on Rheological, Pasting, and Textural Properties of Dough. International Journal of Food Properties 17:1428-1438.

Bernardo R (2014) Genomewide Selection When Major Genes Are Known. Crop Science 54:68-75.

**Bordes J, Ravel C, Jaubertie JP et al.** (2013) Genomic Regions Associated with the Nitrogen Limitation Response Revealed in a Global Wheat Core Collection. Theoretical and Applied Genetics 126:805 22.

**Clark SA, Hickey JM, van der Werf JH** (2011) Different models of genetic variation and their effect on genomic evaluation. Genet Sel Evol 43:18.

Covarrubias-Pazaran G (2016) Genome-Assisted prediction of quantitative traits using the r package sommer. PLoS One 11:1–15.

Heffner EL, Jannink JL, Iwata H et al. (2011) Genomic Selection Accuracy for Grain Quality Traits in Biparental Wheat Populations. Crop Science 51:2597-2606. https://doi.org/10.2135/cropsci2011.05.0253

**Isidro J, Jannink JL, Akdemir D et al.** (2015) Training Set Optimization under Population Structure in Genomic Selection. Theor Appl Gene 128:145 158. https://doi.org/10.1007/s00122-014-2418-4

Jarquín D, Crossa J, Lacaze X, *et al.* (2013) A reaction norm model for genomic selection using high-dimensional genomic and environmental data. Theor Appl Genet 127:595-607. https://doi.org/10.1007/s00122-013-2243-1

Laloë D (1993) Precision and information in linear models of genetic evaluation. Genet Sel Evol 25:557-576.

Liu G, Zhao Y, Gowda M et al. (2016) Predicting Hybrid Performances for Quality Traits through Genomic-Assisted Approaches in Central European Wheat. PLOS ONE 11:e0158635. https://doi.org/10.1371/journal.pone.0158635

Ly D, Hamblin M, Rabbi I *et al.* (2013) Relatedness and Genotype × Environment Interaction Affect Prediction Accuracies in Genomic Selection: A Study in Cassava. Crop Sci 53:1312-1325.

MacRitchie F (1999) Wheat proteins: characterization and role in flour functionality. Cereal Foods World, 44:188-193

Marchetti L, Miguel C, Leda C, Cristina F (2012) Effect of Glutens of Different Quality on Dough Characteristics and Breadmaking Performance. Food Science and Technology 46:224-31. https://doi.org/10.1016/j.lwt.2011.10.002

Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819–1829

Michel S, Ametz C, Gungor H et al. (2016) Genomic selection across multiple breeding cycles in applied bread wheat breeding. Theor Appl Genet. 129:1179 89.

Michel S, Kummer S, Gallee M *et al.* (2018) Improving the baking quality of bread wheat by genomic selection in early generations. Theor Appl Genet. 131:477-93.

Moore J K, Manmathan HK, Anderson VA *et al.* (2017) Improving Genomic Prediction for Pre-Harvest Sprouting Tolerance in Wheat by Weighting Large-Effect Quantitative Trait Loci. Crop Sci 57:1315-1324.

**Ortolan F, Steel CJ** (2017) Protein Characteristics that Affect the Quality of Vital Wheat Gluten to be Used in Baking: A Review. Comprehensive reviews in food science and food safety 16:369-381.

**Oury FX, Godin C** (2007) Yield and Grain Protein Concentration in Bread Wheat: How to Use the Negative Relationship between the Two Characters to Identify Favourable Genotypes? Euphytica 157:45-57.

Payne PI, Corfield KG, Blackman JA (1979) Identification of HMW-subunit of glutenin whose presence correlates with bread-making quality in wheat of related pedigree. Theor appl genet. 55:153-159.

**Rimbert H, Darrier B, Navarro J et al.** (2018) High Throughput SNP Discovery and Genotyping in Hexaploid Wheat. PLOS ONE 13:e0186329.

**Rincent R, Laloë D, Nicolas S, et al.** (2012) Maximizing the reliability of genomic selection by optimizing the calibration set of reference individuals: comparison of methods in two diverse groups of Maize Inbreds (Zea mays L.). Genetics 192:715–728.

**Rutkoski J, Singh RP, Huerta-Espino J et al.** (2015) Efficient Use of Historical Data for Genomic Selection: A Case Study of Stem Rust Resistance in Wheat. Plant Genome 8:1.

**Plessis A, Ravel C, Bordes J et al.** (2013) Association study of wheat grain protein composition reveals that gliadin and glutenin composition are trans-regulated by different chromosome regions. J Exp Bot 64:3627–3644.

Schiedt B, Baumann A, Conde-Petit B, Vilgis TA (2013) Viscoelastic Properties during Dough Development. J Texture Stud 44:317-332.

**Shewry PR** (2009) Wheat Journal of Experimental Botany 60: 1537–1553.

Spindel JE, Begum H, Akdemir D et al. (2016) Genome-Wide Prediction Models That Incorporate de Novo GWAS are a powerful new tool for tropical rice improvement. Heredity 116:395-408.

**R Development Core Team** (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Wieser H (2007) Chemistry of gluten proteins. Food Microbiology 24:115 119.

Zhang Y, Tang J, Zhang Y *et al.* (2011) QTL Mapping for Quantities of Protein Fractions in Bread Wheat (Triticum Aestivum L.). Theor Appl Genet 122:971-987

# [Pmg-TILL] Identification des gènes impliqués dans le Poids de Mille Grains par une approche innovante (TILLING)

#### Valérie LAURENT<sup>1</sup>\*, Cristobal UAUY<sup>2</sup>, Laure DUCHALAIS<sup>3</sup>, Chris BURT<sup>3</sup>

1 - Florimond DESPREZ - BP41, 59242 Cappelle en pévèle

- 2 JOHN INNES CENTRE NR4 7UH, Colney, Norwich United Kingdom
- 3 R2N Route d'Epincy, 28150 Louville La Chenard

\* Coordinatrice : Valérie LAURENT, valerie.laurent@florimond-desprez.fr

# 1. Introduction

Selon la FAO, le blé représentait 21% de la production mondiale de céréales en 2012 (soit plus de 675 millions de tonnes). Une proportion croissante de blé est utilisée pour l'alimentation animale dans les pays industrialisés (45% de son usage total dans l'UE). L'utilisation de blé par habitant dans les pays en développement, essentiellement pour l'alimentation humaine, a continué d'augmenter, et la plupart de ces pays sont de plus en plus dépendants des importations. Le recours aux importations par les pays en développement (à l'exception de l'Argentine et de l'Uruguay, qui sont exportateurs) devrait continuer de s'intensifier et les importations nettes de blé devraient passer de 51 millions de tonnes par an en 2012 à 160 millions en 2030. Pour répondre à cette demande mondiale l'agriculteur français doit continuer à produire plus tout en adoptant des pratiques culturales différentes plus en respect avec l'environnement. L'amélioration d'une des composantes du rendement, le Poids de Mille Grains (PMG), est une solution permettant de répondre aux attentes des agriculteurs. Le PMG est un facteur complexe de type quantitatif, déterminant pour le rendement grain. Ce caractère influe également sur la qualité de l'utilisation finale du blé en affectant le contenu du grain en protéines et le rendement en farine. Un certain nombre d'études, en particulier dans le modèle céréales-riz, ont été menées afin de déterminer les fondements génétiques et moléculaires de ce caractère. Elles ont conduit à l'identification de plusieurs gènes qui régulent la taille du grain de riz, comme gSW5 (Shomura et al., 2008), DW5 (Weng et al., 2008), MIC 1 (Wang et al., 2008), Ghd7 (Xue et al., 2008) et GS3 (Fan et al., 2006 ; Takano et al., 2009). Le blé tendre (Triticum aestivum L.) possède un grand génome (17 000 Mo) avec trois groupes homéologues et d'abondantes séquences répétées. Actuellement, des dizaines de QTL liés au PMG ou au rendement ont été identifiés sur presque tous les 21 groupes de liaison dans différentes populations de cartographie (Kato et al., 2000; Borner et al., 2002; Groos et al., 2003; McCartney et al., 2005; Quarrie et al., 2005; Huang et al., 2006; Li SS et al., 2007; Gegas et al., 2010; Peleg et al., 2011). Récemment, Su et al. (2011) ont développé des marqueurs proches du gène TaGW2 (gène orthologue de OsGW2 chez le riz).

L'approche TILLING (Targeting Induced Local Lesions IN Genomes) permet de cibler plus facilement des gènes potentiellement impliqués dans l'expression d'un caractère; elle a déjà été utilisée avec succès par Slade *et al.* (2012) qui ont obtenu des grains de blé dur et de blé tendre fortement enrichis en amylose par mutations des gènes *SBElla* sur les génomes A, B et D. De même, Sestili *et al.* (2010) ont également eu une approche de TILLING pour les gènes (*Sgp-1* et Wx) impliqués dans la synthèse de l'amidon.

Des études préliminaires obtenues par le John Innes Centre ont permis de montrer que plusieurs mutations du gène *TaGW2* présentes sur le génome A de la variété Kronos permettaient d'augmenter en moyenne de 6,7% le PMG avec, dans certains cas une augmentation de plus de 10%. Ce projet vise à identifier des mutations permettant d'influencer favorablement la morphologie du grain à partir de populations de TILLING. Ces mutations ont été introduites dans 4 lignées élites françaises par backcross et les descendances ont été phénotypées en serre et au champ pour valider l'effet de ces mutations. Des mutants doubles et triples combinant des mutations de différents gènes ont également été produits pour cumuler les effets sur le PMG. Des marqueurs SNP liés aux mutations ont parallèlement été développés pour permettre aux sélectionneurs de disposer d'un outil d'identification des haplotypes favorables dans les variétés élites.

# 2. Matériel et méthode

#### Matériel

Deux populations de TILLING obtenues pour le blé tendre alternatif Cadenza et le blé dur Kronos par le John Innes Centre ont été utilisées pour la recherche de mutations impactant le PMG.

Les gènes mutés identifiés ont été introgressés par backross dans 4 fonds génétiques élites français : Cellule, Oregrain, Rubisko et RGT\_Mondio. Cellule a un petit PMG (cotation 3), Oregrain et RGT\_Mondio sont des variétés à PMG intermédiaire bas (cotation 4) et Rubisko a un assez gros PMG (cotation 6). Ces 4 variétés ont des origines génétiques différentes, Cellule et Oregrain proviennent du programme de sélection de Florimond Desprez et Rubisko et RGT\_Mondio proviennent de celui de RAGT.

Les donneurs des mutations sont un génotype BC5F3 de la variété Paragon porteur de la mutation du gène *TaGw2A* de Cadenza, le mutant Cadenza1441 pour la mutation de *TaGW2D* et le mutant Kronos T4-341 pour *TaGW2B*. Pour les mutations de *TaARF2.1*, les donneurs sont les mutants Cadenza0148 pour le gène *ARF2.1\_A*, Cadenza1403 pour *ARF2.1\_B* et Cadenza1274 pour *ARF2.1\_D*.

#### Méthodes

#### Identification des mutations

A partir des séquences de 1535 mutants de TILLING de Kronos et de 1200 mutants de Cadenza (www.wheat-tilling.com), la recherche de mutations a été réalisée *in silico*, par le JIC, pour les gènes *TaGW2* des génomes B et D et les gènes *DST* et *ARF2.1* des trois génomes L'objectif était d'identifier des mutations et des troncatures sur chaque génome homéologue. Les meilleurs mutants (codon stop, acide aminé responsable d'une modification de la fonction de la protéine, etc...) ont été sélectionnés et croisés avec les variétés élites.

#### Dévelopement des marqueurs

Une fois le polymorphisme de séquence de chaque gène homéologue identifié (pour les génomes A, B et D), des marqueurs SNP KASPar ont été développés afin de faciliter l'identification de leurs variants alléliques et permettre leur transfert dans les variétés élites. A chaque génération, le génotypage avec le marqueur de la mutation a permis de sélectionner les plantes porteuses de la mutation. Pour les 2 premières générations de rétrocroisements, le choix des plantes a également été effectué à partir de marqueurs du fonds génétique pour accélérer le retour vers le parent récurent.

#### Croisements

Les mutations ont été introduites par backcross dans les lignées élites à raison de 2 générations par an en serre jusqu'au stade BC2F2 et BC3F2.

Ce schéma a été et est en cours d'application pour les 6 allèles *TaGW2-A*, *TaGW2-B* et *TaGW2-D*. Ces trois mutations seront cumulées par deux puis par trois dans un même génotype.

#### Phénotypage

Afin de valider l'effet et l'intérêt de ces mutations lorsqu'elles sont présentes dans des fonds génétiques élites, l'effet des gènes introgressés (seuls et/ou cumulés) a été évalué sur la morphologie des grains en serre. Il sera évalué ensuite pour le rendement au champ.

Au stade BC2F3 et BC3F2, un phénotypage sur 15 épis pour chacune des populations d' individus homozygotes pour la mutation: la taille du grain a été mesurée à l'aide d'un Marvin grain analyser (GTA Sensorik GmbH, Allemagne) et le nombre d'épillets/épi ainsi que le nombre de grain/épi ont été évalués.

Trois descendances des plantes BC2F2 homozygotes pour le gène *TaGW2* muté ont été et seront phénotypées au champ en épi ligne. La hauteur et le rendement seront relevés

Les descendances BC2F4, BC3F3 et BC3F4 ont été et seront regroupées (en bulks) avec ou sans le gène *TaGW2* muté pour une étude des composantes du rendement au champ, dont le PMG. Les essais au champ des bulks BC2F4 et BC3F4 pour la mutation de *TaGW2A* ont été semés en automne 2018.

#### 3. Résultats

#### Identification et caractérisation des mutations responsables de la taille et du poids du grain

Des mutations des gènes *TaGW2*, *DST* et *ARF2.1* ont été recherchées sur la variété Cadenza sur les 3 génomes A, B et D.

#### Les mutations du gène TaGW2



Figure 1 : Représentation de la localisation des mutations du gène TaGW2A

La mutation originale du gène *TaGW2-A1* a été identifiée dans la lignée de TILLING T4-2235, elle correspond à une transition G>A au début de l'exon 5, au niveau du site d'épissage. L'allèle *TaGW2-B* vient du mutant de blé dur Kronos T4-0341, qui porte une transition C>T à la position 2557 et un codon stop prématuré. Une mutation stop du gène *TaGW2-D1* a été identifiée chez le mutant Cadenza 1441 au niveau du 7<sup>ème</sup> exon (cf Fig. 1).

Tous ces génotypes présentent des caractéristiques intéressantes sur la taille et le poids des grains par rapport au génotype d'origine (cf Fig. 2 et Tab. 1).



Figure 2 : Effet de l'allèle muté gw2-B du mutant de blé dur Kronos T4-0341 sur le PMG et la taille du grain

ID		Seed/Spike	Yield/ Spike	TGW(g)	ØArea	ØWidth	ØLength	Spike Length	Viable Spikelets	Seeds/ Spikelet
Cadenza		49.09	2.06	41.99	18.23	3.42	6.72	9.89	18.09	2.7121
C1441_1-1	C1441	41	2.53	61.71	22.35	4.14	6.87	9.8	16	2.5625
C1441_1-2	C1441	39	2.45	62.82	22.34	4.05	6.98	9.4	18	2.1667

 Tableau 1 : Effet de l'allèle muté gw2-D du mutant Cadenza sur le PMG, la taille du grain et les caractéritiques de l'épi.

#### Les mutations des gènes DST

Le gène *DST* régule l'expression de la cytokinine oxidase 2 (*OsCKX2*) chez le riz. Un allèle mutant tronqué dominant de ce gène chez le riz agit comme un régulateur négatif de l'*OsCKX2*, provoquant une augmentation de la production de grains (Li *et al.*, 2013). Les auteurs ont montré que l'expression ectopique de cet allèle mutant de riz dans le blé conduit, en serre, à une augmentation de la taille des épis et du nombre d'épillets.

Sur le génome A du blé, 24 mutations du gène *DST* ont été identifiées dont une est un codon stop pour le mutant T4-622. Sur le génome B, 30 mutations ont été découvertes dans le gène *DST* mais aucune ne provoquant de codon stop. Malheureusement, aucun grain de la lignée mutante T4-622 n'a donné une plante viable. Cette voie d'amélioration du PMG à l'aide des gènes *DST* est donc sans issue.

#### Les mutations des gènes ARF

Les gènes ARF des facteur de réponse à l'auxine agissent sur la taille du grain chez *Arabidopsis* et des résultats préliminaires semblent indiquer le même effet chez le blé dur. Les mutants de *TaARF2.1*, Cadenza0148 (*ARF2.1\_A*), Cadenza1403 (*ARF2.1\_B*) et Cadenza1274 (*ARF2.1\_D*) ont été obtenus.

#### Croisements mutants de TaGW2 x lignées élites :

#### Mutant du gène TaGW2-A1

Le génotype Parangon-*TaGW2A* a été croisé avec les parents receveurs. Pour 20 plantes BC1 par lignée récurrente, sélectionnées par marqueurs pour la présence du gène *TaGW2A*, la taille de l'introgression a été évaluée par marquage à l'aide de 7 marqueurs SSR du chromosome 6A. Trois et quatre plantes BC1 avec une introgression de moins de 30 cM autour du gène *TaGW2A* ont été retenues pour poursuivre les rétrocroisements. Les BC3 et BC2F1 ont produit des descendants avec la mutation à l'état homozygote. Ces individus ont été testés pour les gènes de développement *ppD-D1*, *VrnA1* et *VrnB1* ainsi que *Rht1* et *Rht2* de façon à ne retenir que les individus de type hiver. Le gène *VrnD1* n'a pas été testé car le donneur printemps Parangon porte l'allèle hiver *vrnD1* à ce gène.

#### Mutant du gène TaGW2-D1

Les plantes BC2 ont été sélectionnées pour leur retour vers le génotype du receveur pour 13 des 20 marqueurs SNP du chromosome 6D testés et pour leur homozygotie pour le type hiver (*vrnA1 et vrnB1*), *RhtB1* et l'insensibilité à la photopériode (*Ppd-D1*).

Le stade BC3F3 a été atteint en 2018 en serre. Les grains obtenus des plantes homozygotes mutées et homozygotes sauvages seront semés aux champs en 2019 pour évaluation de leur rendement.

#### Mutant du gène TaGW2-B1

Une tentative de croisement entre le mutant stop du gène *GW2-B1*, identifié dans le blé dur Kronos, et les lignées élites a été réalisée mais les hybrides n'ont pas germés ou n'ont pas survécu après la vernalisation (dans Oregrain cf. Fig. 3). La mutation a toutefois pu être stabilisée à l'état hétérozygote dans le blé tendre Paragon (à gauche sur la photo). Ces plantes ont été utilisées en croisement pour introduire la mutation dans les 4 variétés élites.



Figure 3 : Incorporation de l'allèle mutant gw2-B1 du blé dur dans les blés tendres Paragon (gauche) et Oregrain (au centre).

Sept individus BC2 portant la mutation *gw2B* à l'état hétérozygote ont été obtenus avec la variété élite Cellule. Un témoin homozygote sauvage sera également conservé pour comparaison de l'effet de la mutation au champ.

#### Mutants doubles et triples

Des plantes BC3F2 homozygotes pour la mutation *gw2-A1* ont été croisées avec des plantes hétérozygotes pour la mutation *gw2-D1* BC2 pour cumuler les mutations dans Cellule. Les doubles hétérozygotes ont été croisés avec les BC1 *gw2-B1* (hétérozygotes). Six individus cumulant les 3 mutations à l'état hétérozygote ont été obtenus. Ils sont en cours d'autofécondation pour essayer de fixer les mutations à l'état homozygote.

# Mutations gw2B obtenues dans Cellule suite au croisement 3 voies

F1	(gw2A/wt; gv	v2B/wt)	17 individus
F1	(gw2D/wt; gv	v2B/wt)	8 individus
- 1	(		

F1 (gw2A/wt; gw2D/wt; gw2B/wt) 6 individus

Des doubles mutants *gw2-A1 / gw2-B1* tétraploïdes sont en cours d'évaluation aux champs au JIC mais aussi aux USA et au Mexique. Ces évaluations permettront d'obtenir une estimation des effets d'un mutant nul pour ces deux gènes chez le blé.

#### Autres mutants

Les mutants de *TaARF2.1* ont été croisés avec les 4 lignées élites jusqu'au stade BC1.

#### ▶ Effet de la mutation du gène TaGW2-A1

#### NILs BC4 dans Paragon

Les NILS BC4 dans Paragon ont été testées en 5 répétitions d'essai rendement de 6 m<sup>2</sup> à Church Farm, Norwich UK (cf Fig. 4). On observe un effet significatif de la mutation sur le PMG (7.96%; P<0.001), sur la largeur du grain (2.03%; P<0.001) mais aussi sur sa longueur (2.66%; P<0.001).

Aucun effet sur le rendement n'est mis en évidence (-0.75%, P=0.44) ni sur le tallage, la date de maturité ou le Poids Spécifique.



Figure 4 : Effet de l'allèle mutant de gw2-A1 sur le rendement ajusté (Adj Yield), le PMG (TGW), la largeur et la longueur du grain dans les essais. Les NILs avec l'allèle mutant sont en bleu et les NILs avec l'allèle GW2-A1 sauvage fonctionnel sont en gris.

# Descendances BC2F2, BC3F3 et BC2F3 dans Cellule et Oregrain

La taille des grains des plantes BC2F2 dans Cellule et Oregrain, a été estimée par une mesure sur le MARVIN grain analyser.

ID	Main Seeds	Weight(g)	TGW(g)	ØArea	ØWidth	ØLength
PARAGON TAGW2A	50	1.05	21.00	13.98	2.87	5.88
CELLULE	55	1.61	29.27	14.33	2.97	5.98
WFS1-7-17-31	79	2.71	34.30	15.15	3.10	6.15
WFS1-7-9-72	339	8.83	26.05	13.67	2.81	6.18
WFS1-7-17-34	265	9.69	36.57	15.17	3.08	6.20
OREGRAIN	55	1.88	34.18	15.96	3.17	6.16
WFS150-42-34	161	9.1	56.52	20.12	3.70	6.92

 Tableau 2 : Effet de la mutation gw2-A1 dans les descendances BC2F2 de Cellule

 et Oregrain.

On observe bien une augmentation de la taille des grains par rapport au receveur élite (cf Tab. 2) ; mais comme nous n'avons pas de témoin négatif (avec le gène *TaGW2A* à l'état sauvage) il n'est pas possible de déterminer dans quelle mesure l'augmentation de taille est due au fond génétique résiduel du donneur ou bien au gène muté.

Les stades BC3F3 et BC2F3 ont été atteints. Ces stades ont la mutation à l'état homozygote, ils ont été semés au champ pour évaluer leur rendement. Des lignées BC3F3 avec le gène à l'état sauvage ont été évaluées en même temps de façon à pouvoir cibler plus précisément l'effet de la mutation. La date d'épiaison, la hauteur et le PMG ont été phénotypés. Sur 15 épis, le nombre de grains par épis, le nombre d'épillets par épis et la taille du grain ont également été évalués avec le Marvin grain analyser sur 50 grains.

On remarque, que la mutation du gène *TaGW2A* a bien produit une augmentation de la taille des grains et du PMG pour les 4 descendances portant la mutation *gw2-A1* (cf Tab.3). Les résultats pour les nombre d'épillets et de grains par épi ne sont pas homogènes.

ID												
Wf152001-9-72-1	BC2F3	gw2 -A1	Cellule	144	90	1158.5	47	18.47	52.27			
Wf152001-17-31-1	BC2F3	gw2 -A1	Cellule	141	95	139.8	43	19.07	52.73			
Wf152001-17-34-1	BC2F3	gw2 -A1	Cellule	143	95	1039	45	17.87	46.47			
Wf152002-42-34-1	BC2F3	gw2 -A1	Oregrain	144	95	321.3	42	18.53	41.33			
Wf162001a-M	BC3F2	GW2 - A1	Cellule	141	95	809.9	41	18.13	50.73	16.10	3.32	6.03
Wf162001b-M	BC3F2	gw2 -A1	Cellule	143	100	809.4	44	16.47	40.80	17.11	3.44	6.23
Wf162001c-M	BC3F2	GW2-A1	Cellule	142	95	711.3	40	17.64	43.00	16.29	3.35	6.07
Wf162001d-M	BC3F2	gw2 -A1	Cellule	142	95	790.4	45	18.00	49.07	17.12	3.43	6.24
Wf162003e-M	BC3F2	GW2-A1	Oregrain	140	90	408.2	41	20.60	58.33	17.02	3.29	6.38
Wf162003f-M	BC3F2	gw2 -A1	Oregrain	140	90	739.4	43	19.20	47.67	17.47	3.43	6.28
CELLULE	Temoin			143	90							
OREGRAIN	Temoin			143	90							

Tableau 3 : Effet de lamutation gw2-A1 dansles descendances BC2F3et BC3F3 de Cellule etOregrain.

#### Descendances BC2F3 dans RGT\_Mondio

Dans RGT\_Mondio, 7 lignées avec la mutation (*TaGW2*+) et 6 lignées sans (*TaGW2*-) ont été comparées en une parcelle de rendement traité en 2017/2018 à Louville-la-Chenard. L'ANOVA met en évidence un effet de la mutation *TaGW2-A1* sur le PMG, le rendement et sur le nombre de grains/épi dans RGT\_Mondio (cf Tab.4).

Y (numerical)	X (categorical)	p-value	FStat	S2Btwn	S2Wthn	dfBtwn	dfWthn	n
NB GRAINS/EPI	TaGW2	0.00173	15.44204251	318.72456	268.3206758	1	13	15
NB EPILLETS/EPI	TaGW2	0.34100	0.975701176	0.64421894	8.583413008	1	13	15
NB EPIS/M <sup>2</sup>	TaGW2	0.47600	0.538179079	1680.17143	40585.42857	1	13	15
RDT %T	TaGW2	0.01620	7.620070694	102.080185	174.1509301	1	13	15
Hauteur	TaGW2	0.30200	1.155555556	7.61904762	85.71428571	1	13	15
EE	TaGW2	0.86700	0.029082774	0.0297619	13.30357143	1	13	15
SPW	TaGW2	0.87000	0.027922008	0.0447344	20.82755893	1	13	15
PMG Récolte	TaGW2	0.01130	6.457525295	42.0803571	42.36	2	13	16

 Tableau 4 : Effet de la mutation gw2-A1 dans les descendances BC2F3 dans

 RGT\_Mondio

Comme attendu, la mutation de *TaGW2-A1* a un effet positif sur le PMG mais elle est également liée à une diminution du rendement qui s'explique par la diminution du nombre de grains/épi (cf Fig. 5).

#### Descendances BC3F3 dans RGT\_Mondio et Rubisko

Dans RGT\_Mondio, 11 lignées *TaGW2A*+ et 10 lignées *TaGW2A*ont été comparées en pépinière en 2017/2018 en 4 épis lignes. Contrairement aux résultats obtenus sur les BC2F3, aucun effet de ce gène n'a été mis en évidence par ANOVA sur les 3 composantes de rendement étudiées (Nb de grains/épi, nb d'épillets/épi et PMG).





Figure 5 : Effet de la mutation gw2-A1 sur le PMG, le rendement et le nombre d'épillets dans les descendances BC3F3 dans RGT\_Mondio. TaGW2+ : i ndivdus avec la mutation, TaGW2- : individus sans la mutation.

Dans Rubisko, la comparaison de 7 lignées *TaGW2A*+ et 6 lignées *TaGW2*- a mis en évidence par ANOVA un très léger effet de la mutation de *TaGW2-A1* sur le nombre d'épillets/épi (p-value=0.0171, Fstat= 7.87).

La mutation de *TaGW2-A1* a un effet négatif sur cette composante (cf Fig. 6).



Figure 6 : Effet de la mutation gw2-A1 sur le nombre dépillets dans les descendances BC3F3 dans Rubisko. TaGW2+ : i ndivdus avec la mutation, TaGW2- : individus sans la mutation.

#### Effet de la mutation du gène TaGW2-D1

Une population F2 a été développée entre le mutant Cadenza 1441 et le génotype Cadenza d'origine. Dix-neuf plantes F2 ont l'allèle muté et 24 l'allèle sauvage. Les F2 mutantes ont un PMG moyen de 55.9 g  $\pm$  0.9, ce qui représente une augmentation de 4.1% par rapport aux F2 sauvages (PMG 53.7 g  $\pm$  0.8). Cela suggère que cette mutation permet également d'augmenter le PMG (cf. figure 7).



# 4. Discussion

L'ensemble des résultats de phénotypage obtenus pour les gènes *TaGW2* semblent confirmer que l'action de ce gène cible spécifiquement le PMG mais pas le rendement. Il semble en effet que l'augmentation de la taille des grains se fait au détriment du nombre de grains par épi, ce qui induit une baisse de rendement , mise ici en évidence dans le fonds génétique RGT\_mondio. Un QTL majeur pour le rendement et le PMG a

Figure 7 : Distribution du PMG

dans la population F2 ségrégeant

pour la mutation gw2-D1. Les

plantes homozygotes pour l'allèle

mutant (rouge) ont un PMG

moyen 4.1% plus élevé que les plantes F2 homozygotes pour

l'allèle sauvage (vert) et que les

plantes hétérozygotes (bleu). Ces

variations sont cohérentes avec le dosage des allèles. été identifié dans la zone du gène *TaGW2A* à partir d'un croisement Spark x Rialto (Simmonds *et al.*, 2014). Intégré dans des NILs, ce QTL permet une augmentation du rendement de 5.5% et du PMG de 5.1%. Il semble toutefois qu'il ne s'agisse pas du gène *TaGW2A* car les 2 parents ont le même allèle pour ce gène *TaGW2A*. Cet autre gène, quand il sera identifié, pourrait peut-être permettre de cumuler augmentation du PMG et augmentation du rendement.

# 5. Conclusion

Malgré la multiplicité des fonds génétiques testés et des lieux, les premières années de phénotypage confirment que les mutations du gène *TaGW2A* si elles ont bien un effet positif sur le PMG, semblent plutôt avoir un effet négatif sur les composantes liées à l'épi : nombre d'épillets/épi, nombre de grains/épi et rendement.

L'évaluation des descendances constituées est poursuivie hors projet jusqu'en 2021 pour vérifier au niveau d'essais avec répétitions l'impact du cumul de ces mutations des gènes *TaGW2* sur le rendement. Une étude des descendances des mutants ARF2.1 sera également entreprise pour vérifier la nature de la relation entre la taille du grain, le PMG et le rendement.

### Références bibliographiques

Borner A, Schumann E, Furste A, Coster H, Leithold B, Roder M, Weber W (2002) Mapping of quantitative trait loci determining agronomic important characters in hexaploid wheat (Triticum aestivum L.). Theoretical and Applied Genetics 105: 921-936

Fan C, Xing Y, Mao H, Lu T, Han B, Xu C, Li X, Zhang QF (2006) GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. Theoretical and Applied Genetics 112: 1164-1171

Gegas VC, Nazari A, Griffiths S, Simmonds J, Fish L, Orford S, Sayers L, Doonan JH, Snape JW (2010) A genetic framework for grain size and shape variation in wheat. Plant Cell 22: 1046-1056

**Groos C, Robert N, Bervas E, Charmet G** (2003) Genetic analysis of grain protein-content, grain yield and thousand-kernel weight in bread wheat. Theoretical and Applied Genetics 106: 1032-1040

Huang XQ, Cloutier S, Lycar L, Radovanovic N, Humphreys DG *et al.* (2006) Molecular detection of QTLs for agronomic and quality traits in a doubled haploid population derived from two Canadian wheats (Triticum aestivum L.). Theor Appl Genet 113: 753-766.

**Kato K, Miura H, Sawada S** (2000) Mapping QTLs controlling grain yield and its components on chromosome 5A of wheat. Theoretical and Applied Genetics 101: 1114-1121

Li SS, Jia JZ, Wei XY, Zhang XC, Chen HM, Sun HY, Zhao XH, Lei TD, Xu YF, Jiang FS *et al.* (2007) An intervarietal genetic map and QTL analysis for yield traits in wheat. Molecular Breeding 20: 167-178.

Li S, Zhao B, Yuan D, Duan M *et al.* (2013) Rice zinc finger protein DST enhances grain production through controlling Gn1a/OsCKX2 expression." Proceedings of the National Academy of Sciences: 201300359

McCartney CA, Somers DJ, Humphreys DJ, Lukow O (2005) Mapping quantitative trait loci controlling agronomic traits in the spring wheat cross RL 4452 · AC 'Domain'. Genome 48: 870-883

Mir RR, Kumar N *et al.* (2012) Genetic dissection of grain weight in bread wheat through quantitative trait locus interval and association mapping." Molecular Breeding 29: 963-972

Peleg Z, Fahima T, Korol AB, Abbo S, Saranga Y (2011) Genetic analysis of wheat domestication and evolution under domestication. Journal of Experimental Botany 62: 5051-5061 Quarrie SA, Steed A, Calestani C *et al.* (2005) A high-density genetic map of hexaploid wheat (Triticum aestivum L.) from the cross Chinese Spring/SQ1 and its use to compare QTLs for grain yield across a range of environments. Theor Appl Genet 110: 865-880

**Sestili F, Botticella E et al.** (2010) Production of novel allelic variation for genes involved in starch biosynthesis through mutagenesis. Molecular Breeding 25: 145-154

Shomura A, Izawa T, Ebana K, Ebitani T, Konishi S, Yano M (2008) Deletion in a gene associated with grain size increased yields during rice domestication. Nature Genetics 40: 1023-1028

Simmonds J, Scott P,Leverington-Waite M, Turner AS, Brinton J *et al.* (2014) Identification and independent validation of a stable yield and thousand grain weight QTL on chromosome 6A of hexaploid wheat (Triticum aestivum L.). BMC Plant Biology 14:191

Slade A J, McGuire C *et al.* (2012) Development of high amylose wheat through TILLING. Bmc Plant Biology 12: 69-69

**Su Z, Hao C et al.** (2011) Identification and development of a functional marker of TaGW2 associated with grain weight in bread wheat (Triticum aestivum L.) Theoretical and Applied Genetics 122: 211-223

Takano KN, Jiang H, Kubo T, Sweeney M, Matsumoto T, Kanamori H, Padhukasahasram B *et al.* (2009) Evolutionary history of GS3, a gene conferring grain length in rice. Genetics 182:1323-1334

Wang E, Wang J, Zhu XD, Hao W, Wang LY, Li Q, Zhang LX, He W, Lu BR, Lin HX (2008) Control of rice grain-filling and yield by a gene with a potential signature of domestication. Nature Genetics 40: 1370-1374

Weng J, Gu S, Wan X, Gao H, Guo T, Su N, Lei C, Zhang X, Cheng Z, Guo X (2008) Isolation and initial characterization of GW5, a major QTL associated with rice grain width and weight. Cell Research 18: 1199-1209

Xue WY, Xing YZ, Weng XY, Zhao Y, Tang WJ, Wang L, Zhou HJ, Yu SB, Xu CG, Li XH (2008) Natural variation in Ghd7 is an important regulator of heading date and yield potential in rice. Nature Genetics 40: 761-767

Zhang, H. P., N. C. Turner, *et al.* (2010) Source-sink balance and manipulating sink-source relations of wheat indicate that the yield potential of wheat is sink-limited in high-rainfall zones. Crop & Pasture Science 61: 852-861

Zhang L, Zhao YL *et al.* (2012) TaCKX6-D1, the ortholog of rice OsCKX2, is associated with grain weight in hexaploid wheat. New Phytologist 195: 574-584

# [OWBM2] Identification et validation de nouveaux marqueurs étroitement liés au gène *Sm1* de résistance à la cécidomyie orange du blé tendre

Valérie LAURENT<sup>1</sup>\*, Cristobal UAUY<sup>2</sup>, Delphine Hourcade<sup>3</sup>, Agnès Tréguier<sup>3</sup>, Raphaël Ducerf<sup>3</sup>, Laure DUCHALAIS<sup>4</sup>, Patrice SENELLART<sup>5</sup>, Sébastien CAIVEAU<sup>5</sup>

1 - Florimond DESPREZ - BP41, 59242 Cappelle en pévèle

2 - JOHN INNES CENTRE - NR4 7UH, Colney, Norwich United Kingdom

3 - ARVALIS-Institut du végétal - Chemin de la côte vieille, 31450 Baziège

- 4 R2N Route d'Epincy, 28150 Louville La Chenard
- 5 SYNGENTA 2 Avenue Gustave Eiffel, 28000 Chartres

\* Coordinatrice : Valérie LAURENT, valerie.laurent@florimond-desprez.fr

# 1. Introduction

La cécidomyie orange, Sitodiplosis mosellana, est un parasite commun du blé responsable de pertes sévères de rendement chaque année dans l'hémisphère nord. Ses larves se nourrissent des grains en cours de formation causant ainsi leur échaudage et des problèmes de germination sur pied. De plus, la présence de ce parasite facilite l'attaque secondaire de champignons comme la fusariose et la septoriose. En 2004, lors d'une pandémie de cécidomyies au Royaume Uni, les pertes de rendement occasionnées par ce parasite furent estimées à 6% (soit environ 1 million de tonnes); ces pertes étaient accompagnées de la diminution de la qualité du grain, malgré l'épandage d'insecticides sur 500 000 ha de blé (Ellis et al., 2009). La cécidomyie orange a une distribution spatiale très inégale et les infestations varient d'année en année en fonction des conditions climatiques, ce qui rend difficile toute prédiction des risques. La principale raison de cette irrégularité est la nécessaire coïncidence entre trois éléments : la phase sensible du blé (début épiaison à fin floraison), la présence de l'insecte adulte en état de pondre, et les conditions climatiques favorables à l'activité de l'insecte à la tombée du jour (vent faible ou nul, températures douces). En plein champ, les dates d'épiaison entre les variétés les plus précoces et les variétés les plus tardives s'échelonnent sur 2 à 3 semaines; la probabilité d'observer des femelles de S. mosellana en état de pondre, mais aussi des conditions favorables aux pontes pendant toute cette période, est très faible. L'expérimentation de plein champ est donc assez difficile pour ce type de caractérisation.

La femelle de cécidomyie est un petit insecte qui peut rester caché dans le feuillage du couvert végétal et les larves sont cachées à l'intérieur de l'épi de blé, ce qui les rend difficile à atteindre par pulvérisation d'insecticide. Aussi, pour assurer un contrôle efficace, une application d'insecticide doit être appliquée sans délai avant que les larves ne s'enfoncent dans l'épi.

Un polymorphisme variétal de la résistance à la cécidomyie a été observé dans le matériel de différents pays. Cependant, il y a eu très peu d'études sur la génétique de ces sources de résistance. L'étude la plus importante démontre que la résistance à la cécidomyie des variétés canadiennes est conditionnée par la présence d'un gène majeur unique, *Sm1*, sur le chromosome 2BS du blé tendre. Le marqueur moléculaire PCR (Wm1) du gène *Sm1* (Thomas *et al.* 2005) n'est pas fiable pour l'ensemble du matériel génétique britannique (par exemple, Wm1 ne permet pas de diagnostiquer la présence ou l'absence de *Sm1* dans les variétés Robigus ou Shamrock; Oakley *et al.* 2005). Dans le cadre du précédent projet FSOV FSOV2010L (Robert *et al.* 2015), la colinéarité entre le blé, le riz et *Brachypodium* au niveau de l'intervalle de *Sm1* (entre les

marqueurs Wm1 et gwm210; Thomas *et al.* 2005) a permis de cartographier plus finement le gène *Sm1* entre les 2 marqueurs Kaspar OWBM7 et OWBM9 au niveau du marqueur OWBM6. L'haplotype sensible constitué des allèles de ces 3 marqueurs s'est révélé très prédictif des 97 génotypes français testés. Toutefois pour 3 lignées, les marqueurs ont révélé des faux négatifs (haplotype sensible mais génotype résistant).

Ce projet vise à obtenir des marqueurs totalement fiables du caractère de résistance qui permettront aux sélectionneurs de s'affranchir du phénotypage.

# 2. Matériel et méthode

#### Matériel

Dans le projet FSOV2010L, un ensemble de lignées quasiisogéniques (NILs) pour la région du chromosome 2BS entre Shamrock (-Sm1) et Robigus (+Sm1), a été développé par le John Innes Center (JIC) avec Robigus comme parent récurrent. Sept lignées recombinantes ont été retenues, parmi les 43 présentant une recombinaison au plus près du gène Sm1, pour la cartographie fine du gène.

Une nouvelle population de 2796 F2 issue du croisement Rubisko x RW21233, développée par RAGT, a été utilisée pour affiner la cartographie du gène.

Un panel d'Arvalis de 72 variétés de phénotype connu, la liste de 106 variétés recommandées en Angleterre et le panel de 92 variétés utilisé dans le projet FSOV 2010L ont été utilisés pour valider les marqueurs et 811 accessions de la collection Watkins de landraces de blé hexaploïdes (Miller *et al.*, 2001) et 479 variétés de la collection Gediflux (Reeves *et al.*, 2004) pour tester la présence du gène *Sm1*.

#### Méthodes

#### Phénotypage en champs

En 2015, 7 lignées NILs avec une recombinaison flanquant le gène *Sm1* ainsi que les 2 parents ont été phénotypés en plein champ en 3 lieux :

- à Ickleton (UK) à quelques kilomètres au sud-ouest de Cambridge en 6 répétitions semées à 2 dates différentes (3 répétitions le 08 octobre et 3 autres le 27 octobre) afin de maximiser les chances d'avoir un vol de cécidomyies lors de la floraison de ces lignées,
- à Orgerus (78) en 5 répétitions de 6 lignes,
- à Ouzouer-Le-Marché (41) sur une parcelle caractérisée pour son potentiel d'infestation et utilisée annuellement pour l'évaluation en 3 répétitions de la résistance à la cécidomyie des nouvelles inscriptions.

En 2016, à partir de la population Rubisko x RW21233, 32 descendances de 30 plantes F3 chacune ont été phénotypées au champ à lckleton. Ces plantes correspondent à des plantes F2 suspectées d'être recombinantes (hétérozygotie des marqueurs). En parallèle de leur phénotypage, elles ont été génotypées et les plantes homozygotes informatives pour le gène ont été envoyées à Arvalis pour le comptage des larves.

En 2017, pour affiner la résolution de la cartographie de *Sm1*, 31 plantes F3 de la population Rubisko x RW21233 dans la zone cible ont été phénotypées par le CRAW en conditions contrôlées et en 2 lieux au champ.

31 individus F3 et 2 témoins ont été implantés :

- en 2 répétitions sur le site d'Ouzouer-le-Marché.
- en 3 répétitions de 1 ligne de 1m50 à Theuville (28).

Les dates d'épiaison et de floraison ont été notées pour chaque entrée.

L'importance des vols de cécidomyies orange a été estimée grâce à un relevé de pièges (cuvette jaune) tous les 2 jours de mi-mai à mi-juillet. Les pièges à insectes sont des cuvettes remplies d'eau additionnée d'une dizaine de gouttes de Teepol. Deux cuvettes par essai ont été placées de façon à ce que le bord supérieur soit au niveau de la base des épis.

Dix épis par entrée ont été récoltés 3 semaines après que la lignée la plus tardive ait fleuri. Ces épis ont été immédiatement congelés puis envoyés à Arvalis pour comptage des larves. Un comptage détaillé des larves a été réalisé pour l'essai présentant le niveau d'infestation le plus élevé.

#### Phénotypage en conditions contrôlées

Afin de s'affranchir de l'irrégularité des émergences et des pontes de l'insecte au champ, une partie du phénotypage a été réalisé en conditions contrôlées par le Centre Wallon de Recherches Agronomiques de Gembloux.

Les mêmes 7 lignées de la descendance Robigus x Shamrock et 32 génotypes de la population Rubisko x RW21233, phénotypés au champ, ont été évalués en 2 répétitions en conditions contrôlées en 2015 et 2017 respectivement.

Après la vernalisation, les plantes ont été repiquées dans des conteneurs disposés à l'extérieur puis à l'approche de l'épiaison, en serre et couverts de voiles. Les lâchers d'insectes ont eu lieu tous les jours de l'épiaison des premiers épis dans les variétés les plus précoces jusqu'à la fin de la floraison des variétés les plus tardives. Au total, plus de 40.000 individus (mâles et femelles) ont été lâchés, ce qui représente près de 4 cécidomyies pour un épi.

Les larves de 20 épis de chaque génotype par répétition ont été comptées.

#### Génotypage et développement de marqueurs moléculaires

Les SNPs utilisés sont des KASPar dérivés de marqueurs présents sur la puce 90K iSelect.

Suite à la collaboration du JIC avec les laboratoires canadiens de Curt McCartney (Morden Research and Development Centre, Agriculture and Agri-Food Canada,) et Curtis Pozniak (Crop Development Centre, University of Saskatchewan) mise en place durant le projet FSOV2010L (Kassa *et al.* 2016), des marqueurs kwm canadiens cartographiés sur une descendance 99B60-EJ2D x Thatcher ont également été utilisés.

Le génotypage a été réalisé pour affiner la cartographie de la population Robigus x Shamrock, réaliser la cartographie de la zone de *Sm1* pour la population Rubisko x RW21233 et effectuer l'étude de génétique d'association (GWAS). Les meilleurs marqueurs ont également été génotypés sur les 3 panels de validations et 2 panels à caractériser pour la résistance.

# Résultats

#### Phénotypage

En 2015, en plein champ, les vols se sont avérés généralement relativement faibles voire nuls à lckleton et le vol principal de cécidomyies est arrivé trop tard, après la floraison des lignée à Orgerus. Le pic de vol des cécidomyies a coïncidé avec la phase de sensibilité des variétés (épiaison) seulement sur le site d'Ouzouer (cf Fig. 1). Sur cette figure, les pluies inductrices déclenchant l'émergence des cécidomyies adultes sont mises en évidence.



Figure 1: Relevés des conditions climatiques 2015 à Ouzouer en relation avec le nombre de cécidomyies capturées et la période d'épiaison. La période d'épiaison des lignées étudiée est représentée par la trait jaune.

En conditions contrôlées, les lâchers d'insectes ont eu lieu tous les jours de l'épiaison des premiers épis dans les variétés les plus précoces (19/05/2015) jusqu'à la fin de la floraison des variétés les plus tardives (15/06/2015).

A Ouzouer, au champ, le contraste entre le nombre de larves dans le témoin sensible et le parent résistant Robigus est de 3.1/0. A Orgerus, l'essai présente une infestation plus importante avec un pic de 27 cécidomyies par cuvette et par jour en période post floraison, mais le pic de vol des cécidomyies, plus tardif par rapport à la phase de sensibilité, a fait que seulement 0.3 larve en moyenne a été observée dans la variété sensible.

L'expérimentation en conditions contrôlées a soumis les génotypes à une pression beaucoup plus importante de cécidomyies oranges, conduisant à des niveaux moyens d'infestation très élevés, voisins de 100 larves par épi dans les variétés les plus sensibles. Dans ces conditions, Soissons et Bermude, utilisées comme références sensibles, et le parent sensible Shamrock renfermaient beaucoup de larves, alors que Robigus, parent résistant, n'en renfermait que très peu (cf Fig. 2).



Figure 2 : Comparaison du nombre de larves par épi en 2015 au champ à Ouzouer et en conditions contrôlées au CRA-W.

En plein champ, les NILs RS33, RS10, RS17 confirment leur tolérance à la cécidomyie orange et RS12, RS36 et RS38 leur sensibilité. La résistance des RILs RS17, RS33 et RS10- et la sensibilité des lignées RS20, RS36, RS38 sont confirmées en conditions contrôlées.

Les résultats sont concordants pour toutes les variétés sauf pour RS12+. La NIL RS12+ (sensible à Rothwell en 2014 et chez Arvalis en 2013) est notée résistante pour le test en conditions controlées. Une erreur du lot de semences est envisagée.

En 2017, du fait des conditions climatiques en France, sur les 2 sites de notations, il y a eu un décalage entre la phase d'émergence des larves et la période de sensibilité des variétés provoquant une pression relativement faible pendant l'expérimentation. Il y a en effet eu une période de faible pluviométrie entre le 20 mars et le 20 avril durant la phase d'émergence des adultes, qui ont besoin d'humidité pour que le processus soit favorisé. Cette faible pluviométrie peut expliquer les niveaux d'infestation très faibles observés sur les 2 sites français.

En effet le relevé des cuvettes met en évidence une moyenne de 2 cécidomyies/cuvette au moment de l'épiaison des variétés et à 3 cécidomyies en post épiaison. Le nombre moyen de larves par épi s'élève à 2 larves par épis, confirmant le faible taux d'infestation relevé avec les cuvettes. A Orgerus, le nombre moyen de larves par épi diminue à 0.13, mettant en évidence que l'infestation était quasi inexistante sur ce site en 2017.

Ce décalage entre période d'épiaison et vol de cécidomyie montre que la pression de l'insecte a atteint son maximum après le pic des épiaisons.

En conditions contrôlées, à partir du 16 mai, au début des premières épiaisons, les génotypes ont été mis en contact sous un filet insectproof avec des cécidomyies prêtes à pondre, chaque jour et jusqu'à la fin de la floraison, le 3 juin.

Le classement des génotypes entre résistance et sensibilité en conditions contrôlées (cf Fig. 3) est le même que celui obtenu par les mesures en conditions naturelles sur le site d'Ouzouer malgré le faible taux d'infestation. On a pu ainsi identifier 15 individus F4 sensibles, et 16 individus F4 résistants.

#### Amélioration de la résolution de la cartographie du gène de résistance Sm1

La réduction de l'intervalle de marqueurs cartographiés autour du gène *Sm1* lors du projet FSOV2010L sur la population Robigus x Shamrock a été poursuivie par cartographie à l'aide de marqueurs situés entre 19 et 32 cM sur le chromosome 2B de la carte de Wang *et al.* 2014 (Wheat\_2014\_90KSNP). L'intervalle a diminué de 1.46 cM à 0.72 cM. De plus, 7 nouveaux marqueurs coségrégeant avec *Sm1* ont été identifiés. Ces marqueurs ont été transformés en marqueurs KASPar.

Comme, il n'y avait plus de lignées recombinantes du croisement Robigus x Shamrock, la réduction de l'intervalle de marqueurs cartographiés autour du gène *Sm1* a été poursuivie sur la population Rubisko x RW21233 de 2798 individus F2. Ces individus ont été génotypés avec les marqueurs flanquant le gène *Sm1*, 128838 (distal) et 184137 (kwm29; proximal).

La carte Rubisko x RW21233 obtenue est cohérente avec celle de la population Robigus x Shamrock. Cent-cinquante-deux individus recombinants ont été identifiés entre les marqueurs 128838 et 184137, ils couvrent une zone de 2.63 cM qui correspond parfaitement à la zone de 1.1 cM de la population Robigus x Shamrock (cf Fig. 4).

Figure 3 : Nombre moyen de larves par épi en conditions controlées au CRA-W en 2017





Figure 4 : Représentation de l'avancement de la cartographie de Robigus x Shamrock par ajout des marqueurs canadiens (kwm) et des marqueurs définis sur la population Rubisko x RW21233

La carte obtenue est également concordante avec la carte canadienne 99B60-EJ2D x Thatcher.

Soixante-dix plantes F2 intéressantes ont été identifiées à partir des données de cartographie et une quinzaine de plantes F3 correspondant à 38 de ces F2 ont été semées à lckleton et génotypées de façon à identifier les plantes homozygotes pour le gène *Sm1*.



Figure 5 : Localisation des marqueurs SNP liés à Sm1 par cartographie et GWAS. Les chiffres en bleu représentent le nombre de recombinants entre 2 marqueurs.

En se concentrant autour du gène Sm1, 23 recombinaisons d'un côté du gène et 15 recombinaisons de l'autre côté ont été identifiées (cf Fig. 5). Trente-deux individus recombinants supplémentaires douteux, dû à la dominance de marqueurs, ont également été retenus pour une caractérisation phénotypique au champ et génotypique. Huit individus recombinants homozygotes supplémentaires dans la région, sont ainsi venus compléter les 23 plantes homozygotes autour du gène Sm1 sélectionnées en année 1. Les 31 plantes correspondantes ont été autofécondées pour effectuer un phénotypage multilocal en champ et en conditions contrôlées en année 3 et affiner la cartographie.

#### Identification de marqueurs par génétique d'association

Kwm1140

Les marqueurs de cartographie flanquant Sm1 (2.63 cM) ont permis de délimiter une zone physique située entre 3.7 Mpb à

17 Mpb sur la version de génome assemblé IWGSC WGA v0.4 (NRGene DeNovoMAGIC publié en juin 2016).

Une analyse d'association a été réalisée avec 22 lignées anglaises séquencées (8 lignées avec Sm1 et 14 sans Sm1), et 1564 SNP provenant de cet intervalle de 14 Mpb (cf Fig. 6).



Figure 6 : Répartition le long de la zone cible de 13.3 Mpb du gène Sm1 des SNP utilisés en génétique d'association. Vert : SNP, bleu les gènes NB-LRR, rouge : SNP parfaitement associés, marron : SNP non associés

Trois marqueurs SNP (BS000128837, BS00160539 et BS00022126) présentent une association parfaite avec la résistance. Seuls les 2 premiers SNP se trouvent également sur la carte Robigus x Shamrock de part et d'autre de *Sm1*. Vingt et un marqueurs SNP ont une association plus lâche.

Vingt-sept gènes de résistance de type NB-LRR, organisés en 6 clusters de gènes, ont été trouvés dans la région de *Sm1*. Les 2 marqueurs parfaits sont situés aux environs de 10 Mpb et 14 Mpb sur le chromosome 2BS; ils ont été génotypés sur le panel de 96 lignées du 1<sup>er</sup> projet FSOV 2010L pour valider leur fiabilité.

#### Recherche de marqueurs parfaits

Les marqueurs 164086 et 16856919 coségrègent parfaitement avec la résistance des 2 populations Robigus x Shamrock et Rubisko x RW21233 (cf Tab. 1).

**Tableau 1 :** Génotypage de 47 recombinantsde la population Rubisko x RW21233

SNP	128838	181603		16856919	CRAW-2017	164086		133621	176043	23831	5158921-1	182447	18413
Ru (A)	A:A	A:A		A:A	R	A:A		A:A	A:A	A:A	A:A	A:A	A:A
RW (B)	B:B	B:B		B:B	5	B:B		B:B	B:B	B:B	B:B	B:B	B:B
303	B:8	B:B	x	A:A	R	A:A		A:A	A:A	A:A	A:A	A:A	A:A
1512	B:8	B:B	х	A:A	R	A:A		A:A	A:A	A:A	A:A	A:A	A:A
1466	B:B	B:B	×		R	A:A		A:A	A:A	A:A	A:A	A:A	A:A
2295	B:B	B:B	×	A:A	R	A:A		A:A	A:A	A:A	A:A	A:A	A:A
1420	B:B	B:B	x	A:A	R	A:A		A:A	A:A	A:A	A:A	A:A	A:A
892	B:B	B:B		B:B	s	B:B	x	A:A	A:A	A:A	A:A	A:A	A:A
910	B:B	B:B		B:B	S	B:B	x	A:A	A:A	A:A	A:A	A:A	A:A
970	B:8	8:8		B:B	5	B:B	x	A:A	A:A	A:A	A:A	A:A	A:A
1812	B:B	B:B		B:B	S	B:B	x	A:A	A:A	A:A	A:A	A:A	A:A
2583	B:B	B:B		B:B		B:B	x	A:A	A:A	A:A	A:A	A:A	A:A
502	B:B	B:B		B:B	S	B:B	x		A:A	A:A	A:A	A:A	A:A
195	B:B	B:B			s	B:B	x	A:A	A:A	A:A	A:A	A:A	A:A
348	B:B	B:B		B:B	S	B:B	×	A:A	A:A	A:A	A:A	A:A	A:A
677	B:8	B:B		B:B	s	B:B	x	A:A	A:A	A:A	A:A	A:A	A'A
1077	B:B	B:B		B:B	s	B:B	x	A:A	A:A	A:A	A:A	A:A	A:A
1107	B:B	B:B		B:B		B:B	x	A:A	A:A	A:A	A:A	A:A	A:A
1496	B:B	B:B		B:B	s	B:8	x	A:A	A:A	A:A	A-A	A:A	A:A
1569	B:8	•		B:B	S	8:8	x	A:A	A:A	A:A	A:A	A:A	A:A
2436	B:B	B'B		B'B	5	B-B	×	A:A	A:A	A-A	A-A	A'A	A-A
2234	B·B	B·B		B-B	-	B·B	-	B-B	A:A	A-A	A-A	A-A	A-A
2763	B-B	R'R		B-B		R·R		R-R	A-A	۵.۵	A-A	۵.۵	A-A
1028	B-B	B-B		B-B		B-B		B-B	A-A	A-A	A-A	A-A	A-A
1948	B-B	B-B		B-B		B-B	-	B-B	B-B	A-A	A-A	A-A	A-A
1710	0.0	0.0		0.0		0.0	-	0.0	D:D	A-A	A-A	A:A	A-A
389	B-B	B-B		R-R		B-B		B-B	B-B	8-8	B-B	A-A	A-A
868	B-B	B-B		B-B		B-9	-	B-B	B-B	B-B	8-8	A.A	A-A
1179	R-R	B-B		B-B		B-B	-		B-B	8.8	B-B	A-A	A-A
1486	B-B	R-R		R R		R-R		B-R	B-B	R-R	R-B	A:A	0.0
1705	D.0	D.D D.D		DiD		0.0	-	0.0	0.0	0.0	8.0	A.A.	A.A.
2270	B-0	0.0		0.0		0.0		0.0	B-0	0.0	0.0	A-A	A-A
22/0	0.0	0:0		0.0		0.0	-	0.0	0.0	0.0	0:0	AJA	A.A.
2720	D.D D.D	D.D D.D		0.0		D.D D.D	-	D.D D.D	0.0	0.0	D.0	A:A	A:A
710	0.0	0.0		0.0		0.0		0.0	0.0	0.0	0.0	ALA	A.A.
1120	D.D.	0.0		0.0		0.0	-	0.0	0.0	0.0	DiD	A:A	A.A.
1130	0.0	B:B		8,8		0.0	-	0.0	0.0	0.0	B:8	0,0	A.A.
121	ACA	ACA	×	8:6	0	8:8		0.0	8:8	8.8	8:8	B:B	0.0
150	A-A	A:A	H	A.A.	P	A:A	×	0,0	0.0	0.0	D:0	D,D D-D	D,D
203	A:A	0:0	H	A:A	P	0:0	×	0.0	8.0	0.0	8-9	B-D	D.D
202	ALA	A:A	H	A:A	n D	A:A	×	0.0	0.0	0.0	8:0	B:0	0:0
264	ALA	ALA	H	A:A	n D	AA	×	0.0	0:0	0.0	8:8	B;D	0.0
368	AA	AA	H	A:A	R	ACA	X	0.0	0.0	0.0	0.0	0.0	0.0
501		A:A	H	A:A	R	ALA	×	8:8	8:8	8:8	8:8	8:8	8:8
746	A:A	A:A		A:A	ĸ	ACA	X	8:8	8:8	8:8	B:B	8:8	8:8
989	A:A	A:A	H	A:A	В	A:A	X	8:8	8:8	8:8	B:8	B:B	B:B
2341_	A:A	A:A			R	A:A	X	8:8	8:8	8:8	B:8	B:B	B:B
1655	A:A	A:A	H	A:A	R	A:A	X	B:B	8:8	8:8	B:B	B:8	B:B
41_	A:A	A:A		12112	R	A:A	X	B:8	8:8	B:B	B:B	B:B	B:B
2560	A:A	A:A		A:A		A:A		A:A	A:A	B;B	B:8	B:B	B;B

#### Validations des marqueurs diagnostiques

Afin de valider les marqueurs diagnostiques du gène *Sm1* identifiés, un panel de 72 variétés élites constitué sur la base d'évaluations phénotypiques menées par Arvalis a été génotypé avec les 2 marqueurs identifiés, celui du FSOV, BS00164086 et celui des canadiens, BS16856919. Le résultat du génotypage a été comparé au phénotypage (cf Tab. 2).

BS00164086	Marqueur R	Marqueur S	total	
R phénotype	37	0	37	
S phénotype	19	16	35	
BS16856919	Marqueur R	Marqueur S	total	
R phénotype	37	0	37 35	
Sabónotuno	0	35		

Tableau 2 : Comparaison de la fiabilité du génotypage des 2 marqueurs BS00164086 et BS16856919.

Le marqueur BS00164086 présente 33% de faux positifs (variété annoncée résistante alors que le gène est absent) alors que le génotypage obtenu par le marqueur canadien BS16856919 est en adéquation parfaite avec la résistance au champ. Ce marqueur a été positionné sur le génome de référence IWGSC refseq1, il est situé dans le R-gene de type NBS-LRR-like resistance, TraesCS2B01G033700.

L'apparition de faux positifs avec le marqueur BS00164086 est confirmée sur 2 autres panels de génotypes: la liste de variétés recommandées en Angleterre (106 UK Recommended genotypes) et le panel utilisé dans le projet FSOV 2010L.

BS00164086	Marqueur R	Marqueur S	total	
R phénotype	37	0	37	UK Recommanded List
S phénotype	2	67	69	
BS00164086	Marqueur R	Marqueur S	total	
BS00164086 R phénotype	Marqueur R 54	Marqueur S 3	total 57	panel FSOV2010L

**Tableau 3 :** Validation complémentaire du marqueur BS00164086 sur les panels de variétés UK Recommanded List et FSOV2010L.

Les 6 individus faux positifs (RGT illustrious, Paragon et K3909) et faux négatifs (As de Cœur, Zobel et LGB OBM 11-08) de ces

2 panels (cf Tab. 3) sont en cours d'étude au Canada en chambre de culture pour confirmer/ infirmer leur singularité et valider le marqueur BS16856919.

#### Présence de Sm1 dans des panels de diversité

Le marqueur BS00164086 du gène *Sm1* a été génotypé sur 811 accessions de la collection Watkins de landraces de blé hexaploïdes (Miller et al, 2001). Seulement 1.7% des landraces (13 génotypes) présentent l'allèle résistant. Aucune origine géographique de la résistance ne semble prédominante, l'allèle résistant semble avoir été introduit faiblement à partir d'origines multiples : 10 à 15% parmi les accessions de Chine, Hongrie, Roumanie et dans une moindre mesure, moins de 5%, parmi les accessions du Maroc, d'URSS et d'Inde.

Le marqueur a également été testé sur 479 variétés de la collection Gediflux. La collection Gediflux est constituée de variétés de blé d'hiver cultivées en Angleterre et au nord de l'Europe de l'Ouest entre les années 1940s et 1980s (Reeves et al, 2004). De nouveau, l'allèle résistant est peu présent, 1.4% (6 variétés dont Renan).

L'ensemble de ces résultats suggère que le gène *Sm1* est récemment apparu en Europe, sa présence étant validée dans moins de 2% de la collection Gediflux et jusqu'à plus de 35% de la liste de recommandation britannique de 2005 à 2018.

#### 4. Conclusions

Durant ce projet, par l'étude d'individus recombinants supplémentaires, nous sommes parvenus à diminuer l'intervalle de confiance autour du gène de résistance à la cécidomyie orange, *Sm1*, de 6.2 cM sur la carte consensus de blé (à la fin du projet FSOV 2010L) à 1.1 cM grâce à la nouvelle population F2 Rubisko x RW21233. Sur le génome de référence RefSeq, cela représente une diminution de 5.7 Mb à 1.2 Mb. Nous avons ainsi obtenu un marqueur parfaitement lié avec l'expression de la résistance au champ grâce à une collaboration avec des chercheurs canadiens mise en place durant le précédent projet FSOV2010L. Ce marqueur sera très utile aux sélectionneurs pour identifier et suivre la résistance dans les lignées sans être tributaire d'une infestation naturelle aléatoire.

#### Références bibliographiques

**Ellis SA, Bruce TJA, Smart LE et al.** (2009) Integrated management strategies for varieties tolerant and susceptible to wheat blossom midge. HGCA Poject Report 451.

Kassa MT, Haas S, Schliephake E, Lewis C, You, FM, Pozniak CJ *et al.* (2016) A saturated SNP linkage map for the orange wheat blossom midge resistance gene *Sm1* Theor Appl Genet 129 : 1507-1517

**Miller TE, Ambrose MJ, Reader SM** (2001) Linnean special issue, the Watkins collection of landrace derived wheats. Wheat taxonomy: the legacy of John Precival. The Linnean Society of London, special issue 3 : 113–120.

**Oakley JN, Ellis SA, Wathling M** *et al.* (2005) Responses of summer cereal aphid populations to reduced rate aphicide applications in field plots of winter wheat. Agrigultural and Forest Entomology 7: 211-218 **Reeves J, Chiapparino E, Donini P, Ganal M et al.** (2004) Changes over time in the genetic diversity of four major European crops: a report from the GEDIFLUX Framework 5 Project. Genetic variation for plant breeding. Proceedings of the 17th EUCARPIA general congress, 8–11 September 2004: 3–7

Robert O, Uauy C, Taupin, P, Duchalais L, Senellart P et Satragliati, J (2015) Etude et identification de facteurs de résistance à la cécidomyie chez le blé tendre. Synthèse des programmes de Recherche FSOV 8 janvier 2015 : 38-41

Thomas J, Fineberg N, Penner G *et al.* (2005) Chromosome location and markers of Sm1: a gene of wheat that conditions antibiotic resistance to orange wheat blossom midge. Molecular Breeding 15: 183-192.

Wang S, Wong D, Forrest K *et al.* (2014) Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. Plant Biotechnology Journal 12:787-796.

[Microdochium spp.] Vers une meilleure connaissance de l'occurrence, de l'épidémiologie du champignon et du comportement des variétés de blé tendre actuelles face à cette maladie

Delphine Taillieu<sup>1 \* α</sup>, Valerie Cadot<sup>3 α</sup>, Benoit Foucault<sup>1 α</sup>, Olivier Robert<sup>1 †</sup>, Laure Duchalais<sup>1</sup>, Sébastien Caiveau<sup>1</sup>, Clément Debiton<sup>1</sup>, Pascal Giraudeau<sup>1</sup>, Sylvie Dutriez<sup>1</sup>, Thierry Bouthillier<sup>1</sup>, Jérôme Auzanneau<sup>1</sup>, Céline Duque<sup>1</sup>, Cindy Vitry<sup>2</sup>, Stéphanie Le Prieur<sup>2</sup>, Florian Dauthieux<sup>2</sup>, Marlène Faure<sup>3</sup>, Thomas Baldwin<sup>3</sup>, Clémence Galon<sup>3</sup>, Isabelle Serandat<sup>3</sup>, Jean-Philippe Maigniel<sup>4</sup>, Romain Valade<sup>2 α</sup>

2 - ARAVLIS - 3 Rue Joseph et Marie Hackin, 75016 PARIS

3 - GEVES - 25 rue Georges Moral, 49071 BEAUCOUZE

4 - GEVES - Domaine de l'Anjouère - La Pouëze, 49370 Erdre-en-Anjou - France

\* Coordinateur : Delphine Taillieu, delphine.taillieu@florimond-desprez.fr, 06 75 42 96 40

 $\alpha$  Co auteur Article et Poster : Romain Valade, Valerie Cadot , Benoit Foucault, Delphine Taillieu

† Décédé en Février 2017 Olivier Robert

### 1. Introduction

La fusariose de l'épi est une maladie causée par un complexe d'espèces fongiques toxinogène (*Fusarium*) et non toxinogène (*Microdochium*). Actuellement, la sélection variétale est principalement axée contre les espèces productrices de fusariotoxines notamment *F. graminearum*. Néanmoins, depuis plusieurs années la présence de deux espèces de *Microdochium*, *M. nivale* et *M. majus* ne cesse d'être significative sur les céréales françaises (Observatoire CartoFusa, Bayer ; Gourdain *et al.*, 2016) et leur rôle dans les pertes de rendements devient de plus en plus préoccupant. De plus, ces deux espèces ont la particularité d'attaquer non seulement l'épi mais également les autres organes de la plante : des racines aux feuilles, la gaine et les nœuds.

L'émergence avérée de ces agents pathogènes, couplée au manque de connaissances à la fois sur l'épidémiologie de ces champignons et sur le comportement des variétés de blé actuelles, traduisent l'enjeu du projet.

De plus l'apparition de résistance à certains produits phytosanitaires dans la population de *Microdochium*, notamment les strobilurines (Walker *et al.*, 2009) et les benzimidazoles, rend la lutte contre ces agents pathogènes plus difficile et met en évidence l'importance du levier variétale. C'est dans ce contexte que s'inscrit cette étude afin de mieux connaitre l'occurrence des *Microdochium* en France, de mieux comprendre leur développement au champ et de mettre au point des outils d'évaluation de la résistance variétale à *Microdochium*. Ces outils seront utilisables par toute la filière, notamment pour la création de variétés résistantes, pour l'inscription au Catalogue français et pour la post inscription.

Le projet s'articule autour de trois axes de travail, l'occurrence de *Microdochium* sur le territoire français, son épidémiologie et enfin l'évaluation du comportement des variétés face à ce pathogène notamment grâce à des outils moléculaires et à un outil basé sur l'imagerie multispectrale de détermination rapide du taux de grains fusariés, le Videometer.

# 2. Occurrence de Microdochium sur le territoire et facteurs agronomiques associés

#### Matériel & Méthodes

Au travers d'un réseau d'enquêtes menées par Arvalis Institut du végétal, 150 échantillons (120 échantillons de blé tendre et 30 échantillons de blé dur), représentatifs des variétés actuelles et

sélectionnés afin de couvrir le territoire national, ont été récupérés au cours de chaque année du projet (2015, 2016 et 2017). L'identification et la fréquence de Microdochiumont été réalisée de trois manières différentes afin d'éviter tout biais lié à l'analyse. Chaque échantillon de grains a été divisé en deux lots. Une analyse de flore par culture microbiologique a été réalisée en prestation externe par le laboratoire EQUASA afin de déterminer l'occurrence du genre *Microdochium* au sein de la flore fongique présente dans chaque lot. A partir de ces analyses, des isolats de *Microdochium* spp ont été transmis au laboratoire de pathologie végétale d'Arvalis pour y être purifiés par isolement monospore puis l'espèce précise a été caractérisée par PCR. Une sélection de ces isolats a également été mise en collection dans un but d'utilisation ultérieure comme inoculum artificiel. En parallèle, le second aliquot de chaque échantillon a été analysé par PCR quantitative afin de déterminer directement la proportion de chaque espèce de Microdochium, mais également de Fusarium graminearum. Ainsi, outre la détermination fiable de la présence et de la composition de la population du Microdochium, l'équilibre entre Microdochium et F. graminearum sur le grain a pu être appréhendé.

Dans le cadre de l'enquête, chaque échantillon est associé à une fiche technique recensant les points importants de l'itinéraire cultural suivi, notamment le lieu précis du prélèvement, le précédent et l'antéprécédent cultural, le travail du sol réalisé, la variété de blé tendre et la date de semis. Ces informations ont été mises en relation avec les quantités de *Microdochium* retrouvées dans chaque lot ainsi qu'avec la composition en espèce afin de déterminer les facteurs déterminants dans la présence de *Microdochium*.

#### Résultats

Les analyses microbiologiques et moléculaires réalisées sur les grains issus des enquêtes agriculteurs ont permis de mettre en évidence une fréquence très variable des espèces entre les trois années avec des niveaux d'occurrence très différents (Figure 1). 4.6%, 49% et 6.41% de grains ont été identifiés, respectivement en 2015, 2016 et 2017, comme contaminés par au moins une espèce du complexe de la fusariose des épis.

Ce résultat est en adéquation avec la pression fusariose observée lors de ces trois années, notamment 2016, qui a été une des années les plus marquantes de cette dernière décennie pour la forte pression maladie notamment sur épis et grains.

*Microdochium* est le taxon prédominant isolé de grains lors de ces 3 années (75% des thalles observés) devant *F. graminearum* (13%) et *F. poae* (7%). Les 3 autres taxa significativement identifiés lors

<sup>1 -</sup> UFS - 17 Rue du Louvre, 75001 PARIS

de ce projet sont *F. tricinctum, F. avenaceum* et *F. langsethiae* (**Figure 1**). La très forte majorité de *Microdochium* est notamment liée à l'année 2016 pour laquelle *Microdochium* a représenté 85% de la flore des grains infectés contre 8% pour *F. graminearum*. Néanmoins, même en 2015, année où la pression a été très faible, *Microdochium* est le taxon prédominant (38% de la flore identifiée). *F. graminearum* a été l'espèce prédominante en 2017, devant *F. poae* et *Microdochium*.



Figure 1 : Fréquence des différentes espèces de Fusarium et Microdochium identifiées par analyse microbiologique pour chaque année (2015, 2016, 2017) et en global.

Les analyses qPCR spécifiques de *F. graminearum*, *M. majus* et *M. nivale* réalisées sur ces mêmes lots de grains ont permis de mettre en évidence également une prédominance de *Microdochium* pour ces années (**Figure 2**).



Figure 2 : Moyenne de la biomasse fongique par taxa (en pg/ng d'ADN total) pour les 3 années du projet.

Les corrélations entre les analyses microbiologiques et les qPCR pour *Microdochium* et *F. graminearum* sont respectivement de R2=0.75 et R2=0.6. D'ailleurs, la moyenne de la biomasse fongique pour *Microdochium* (2.03 pg/ng d'ADN total) est environ 4.5 fois supérieure à celle obtenue pour *F. graminearum* (0.45) ce qui est sensiblement le même facteur obtenu avec les analyses microbiologiques (5.7x).

Les analyses qPCR ont permis d'identifier et quantifier chaque espèce de *Microdochium*. Ainsi, l'espèce majoritairement quantifiée en moyenne est *M. nivale* avec 1.49 pg/ng d'ADN total contre 0.6 pg/ng d'ADN total pour *M. majus*. Cette différence est principalement liée à l'année 2016 qui semble avoir favorisé *M. nivale* par rapport à *M. majus*.

Ces résultats ont pu être mis en corrélation avec des données obtenues dans des projets précédents (ANR Don&Co et CASDAR Ecofusa). Cette analyse globale permet de montrer que la pression *Microdochium* a été la plus forte en 2016 pour les 9 années d'étude et que la pression est significative 1 année sur 2 (**Figure 3**). Aucune n'exclusion n'a été observée entre *F. graminearum* et *Microdochium* et une corrélation forte existe entre les deux espèces de *Microdochium*.



Figure 3 : Biomasse de Microdochium spp. et F. graminearum au cours des 10 dernières années.

La pression en 2015 et 2017 étant assez faible vis-à-vis de *Microdochium* dans l'ensemble des enquêtes, il n'a pas été possible d'identifier des facteurs de risque liés à l'itinéraire technique dans le cadre du projet. En 2016, un échantillonnage complémentaire au projet a permis d'analyser, par qPCR, 149 échantillons prélevés pendant le remplissage. L'analyse des résultats montre que le travail du sol et le précédent n'ont pas d'effet significatif sur la quantité de *Microdochium* contrairement à *F. graminearum* qui est, comme attendu, significativement plus abondant après un précédent maïs non labouré (**Figure 4**).



Figure 4 : Quantité d'ADN de Microdochium et F. graminearum en fonction du précédent et du travail du sol (149 échantillons prélevés pendant remplissage).

L'analyse des enquêtes de 2016 a mis en évidence de potentielles différences entre les variétés prélevées mais le nombre d'échantillons par rapport aux autres variables n'est pas assez important pour analyser statistiquement cette réponse.

# 3. Epidemiologie de Microdochium spp.

Le phénotypage reste la base quant à la recherche de génotype répondant aux critères de sélection, comme la tolérance aux maladies. La détermination du stade d'infection optimal en condition semi contrôlées et des conditions météorologiques favorisant l'installation de *Microdochium* spp. est un sujet majeur pour les acteurs de la sélection. Optimiser et définir une méthode d'inoculation au champ est important afin de pouvoir discriminer le matériel végétal en sélection. Outre cet élément, il nous permet de mieux comprendre le mode d'infection de ce champignon.

#### Matériel & Méthodes

Les essais ont été menés sur 2 années consécutives (2015 et 2016) afin d'optimiser le phénotypage au travers de dispositifs type pépinières irriguées dans des conditions climatiques naturellement différentes : un site Nord (Site Florimond Desprez, Cappelle en Pévèle 59), et deux sites dans le Centre (Site ARVALIS, Boigneville 91 et le site RAGT2n, Louville-la-Chenard 28). Quatre variétés de blé tendre de comportements différents face à la fusariose ont été testées, OREGRAIN, CELLULE, PREMIO et AREZZO. Le dispositif est commun aux 3 sites, en 2 répétitions de 6 lignes sous 5 modalités d'inoculation différentes (**Tableau 1**).

Stade d'infection	Objectifs
Dernière Feuille Etalée (DFE)	Appréhender l'apparition de symptômes sur feuilles
Mi Epiaison	Décrit comme stade optimal pour l'infection par Microdochium
Mi Floraison	Stade critique pour la fusariose
DFE + Mi Floraison	Inoculation double pour observer des symptômes sur les organes de la plante
NON inoculé	TEMOIN

Tableau 1 : Modalités d'inoculation.

Les essais ont été inoculés à partir d'un inoculum de *M. nivale*, espèce tendant à dominer les populations naturelles depuis ces

3 dernières années. L'inoculation artificielle a eu lieu par pulvérisation de spores à une concentration de 106 spores/ml.

L'observation de l'apparition des symptômes a été déterminante pour le déclenchement des premières notations. 15 à 20% de symptômes sur feuilles observés sur notre témoin de référence, Premio, était synonyme des premières notations pour chaque modalité.

Le but de cet essai a été de suivre l'évolution des symptômes sur feuilles jusqu'aux épis via plusieurs notations visuelles, grâce à l'échelle utilisée classiquement en sélection variétale, note de 1 (=résistant) à 9 (=sensible). Outre la notation, la biomasse fongique a été mesurée pour chaque notation réalisée à l'aide d'outils moléculaire (cf. point 2). Ainsi, nous avons pu suivre la cinétique du complexe fusarien dans les différents prélèvements de feuilles, épis et grains (de 48h, 2 semaines, 1 mois après inoculation et à récolte variable selon l'organe) et relier ces données aux notations.

#### Résultats

Les résultats ont été analysés selon l'organe prélevé : feuilles, épis et grains. Les corrélations et les analyses statistiques ont été réalisées sur le jeu de données de 2016 en rassemblant les 3 lieux d'essai : Cappelle en Pévèle, Louville la Chenard et Boigneville. Seuls les résultats de l'année 2016 seront présentés, en effet l'année de 2015 n'a pas été favorable au développement de *Microdochium* malgré l'apport d'inoculum.

#### Analyses des Feuilles

Une assez bonne corrélation est observée entre la notation sur feuilles réalisée par l'expérimentateur et la quantité de *M. nivale*, soit R= 0.89 (**Figure 5**). L'analyse de variance ne démontre aucune différence significative selon les différentes modalités d'inoculation.



Figure 5 : Matrice de corrélation de Nécroses Feuilles Vs Quantité de Microdochium spp. (3 sites regroupés).

#### Analyses des épis

Sur épi, une plus grande quantité de *M. nivale* a été observée comparativement à celle sur feuilles (quantité de *M. nivale* en pg/ng d'ADN total). Selon les modalités des différences significatives ont été observées et confirmées par l'analyse de variance (**Figure 6**).

Deux groupes statistiques ont été identifiés. Les modalités Mi épiaison et Mi floraison présentent les moyennes ajustées les plus élevées (**Figure 7**). Cependant la modalité Témoin et la modalité DFE + Mi Floraison ne sont pas statistiquement différentes. En effet l'année 2016 a été marquée par une forte présence de *Microdochium* spp., sur le territoire français. La présence de celuici dans les témoins non inoculés, démontre que la contamination naturelle a été aussi importante dans les sites d'essais. Cependant les essais permettent de mettre en évidence qu'une inoculation entre le stade mi-Epiaison et mi-Floraison serait toutefois plus favorable à l'expression de symptômes et donc du champignon.



Figure 6 : Résultats qPCR - Quantité de M.nivale (pg/ng ADN Total) sur épi selon les modalités d'inoculations (3 sites regroupés).

\* Dernière Feuille Etalée.

Modalite	Ismean	SE	df	lower.CL	upper.CL	. grou	<b>,</b>
DFE	10.25227	2.660480	139	4.992025	15. 51251	A	
Temoin	16.54319	2.660480	139	11.282946	21.80343	AB	
DFE + MI FLO	19.02941	2.660480	139	13.769166	24.28965	AB	
Mi floraison	23.70125	3.072057	139	17.627245	29.77525	В	
Mi épiaison	24.94038	3.072057	139	18.866377	31.01438	в	
Confidence lev P value adjust significance	vel used: tment: tub level used	0.95 key metho d: alpha	d fo = 0.	r comparin 05	g a famil	y of 5	estimates

Figure 7 : Analyses Variance - ANOVA - Effet de la modalité d'inoculation (date) sur la quantité de M. nivale (3 sites regroupés).

La corrélation entre la présence de symptômes de *Microdochium* spp, et la biomasse fongique sur l'épi, est assez moyenne : R =0.57 (**Figure 8**). Cette corrélation est plus faible que celle observée sur feuilles.



Figure 8 : Matrice de corrélation entre la notation épis (Nécrose épillets) vs la quantité de Microdochium spp. (3 sites regroupés).

Ce résultat démontre la difficulté pour l'expérimentateur à noter et appréhender les symptômes de *Microdochium* spp. sur épi par rapport aux feuilles. En effet l'expérience des notations au champ a prouvé la difficulté à phénotyper sur épi ce champignon, difficulté augmentée également par la présence de *F. graminearum* dans les essais (données non montrées). Sur feuilles les symptômes sont certes proches de ceux de la septoriose mais sont plus faciles à apprécier que les symptômes sur épis.

#### Analyses sur grains

L'ensemble des dispositifs ont été récoltés, puis ont été analysés par qPCR pour quantifier *M. nivale* dans les lots de grains récoltés pour chaque modalité. Aucun effet significatif n'a été mis en évidence pour les différentes modalités d'inoculation.

# 4. Evaluation du comportement des variétés face à *Microdochium* spp.

Cet axe du projet comportait plusieurs objectifs, tout d'abord la mise en application d'un protocole de contamination de *Microdochium* spp. pour une caractérisation phénotypique optimale au champ. Le deuxième était l'optimisation du protocole de phénotypage à la résistance aux *Microdochium* spp. Enfin la caractérisation du comportement des variétés de blé françaises à ces agents pathogènes afin de mieux orienter la sélection variétale. Le panel de variétés comportait 45 variétés dont 5 témoins.

#### Matériel & Méthodes

Ce panel a été sous-divisé en 2 sous-panels en fonction de la gamme de précocité et a été testé sur 2 ans répartis dans 12 lieux différents en France. Le design expérimental était sous forme de pépinière maladie irriguée avec 3 modalités en 2 répétitions ; 1 modalité sans inoculum (Témoin), une modalité avec M. nivale et enfin une modalité avec *M. majus*. Au cours de ces deux années d'expérimentations, les deux protocoles d'inoculation ont consisté en deux ou trois inoculations. L'efficacité de l'infection et la pression maladie ont été mesurées par 2 méthodes : visuelle et moléculaire. La première basée sur l'observation visuelle des feuilles lorsque le témoin Premio était touché à 15-20% avec une échelle de 1 à 9 selon le pourcentage de nécroses foliaires, 1 étant l'absence de symptôme. Pour les épis, des notations visuelles ont été réalisées respectivement à 450 et 550°C jours après Mi-Floraison sur épis, de 1 à 9 comme une notation classique de fusariose. Des prélèvements ont été effectués pour réaliser des analyses qPCR sur grains et sur feuilles pour les sites les plus infestés. Au total 9 sites ont été retenus et 4 sites ont été analysés entièrement par gPCR.

#### Résultats

Sur les témoins (Apache, Renan, Premio, Tremie et Bermude) révélateurs, l'analyse statistique a mis en évidence des effets variétés et des effets souches significatifs sur toutes les mesures, excepté pour les feuilles (**Tableau 2**). Pour autant, nous n'avons jamais eu d'effet significatif de la modalité d'inoculation.

Factor	Trait	Df	Pr(>F)
Variété	qPCR_M.nivale	46	8.12e-09***
Variété	qPCR_M.majus	46	4.93e-08***
Variété	qPCR_F.grami	46	2.17-05***
Variété	qPCR_Total_Fungi	46	5.20e-11***
Variété	NOTE_GLOBALE_FEU	46	5.85e-09***
Variété	NOTE_GLOBALE_EPI	46	1.63e-14***
Souche	qPCR_M.nivale	3	2.75e-17***
Souche	qPCR_M.majus	3	1.19e-15***
Souche	qPCR_F.grami	3	0.003**
Souche	qPCR_Total_Fungi	3	0.050*
Souche	NOTE_GLOBALE_FEU	3	7.39e-11***
Souche	NOTE_GLOBALE_EPI	3	2.60e-05***

Tableau 2 : Effets variétaux et modalités sur les témoins de l'essai (2016 - 2018).

En prenant en compte, la totalité du panel variétés, des effets significatifs des souches et des variétés ont été mis en évidence pour différents caractères mesurés (**Tableau 3**).

Factor	Trait	Df	Pr(>F)
Variete	qPCR_M.nivale	4	0.0004***
Variete	qPCR_M.majus	4	0.0545*
Variete	qPCR_F.grami	4	1.780e-06***
Variete	qPCR_Total_Fungi	4	7.079e-08***
Variete	NOTE_GLOBALE_FEU	4	0.1142
Variete	NOTE_GLOBALE_EPI	4	6.317e-06***
Souche	qPCR_M.nivale	3	1.254e-05***
Souche	qPCR_M.majus	3	3.384e-08***
Souche	qPCR_F.grami	3	0.1309
Souche	qPCR_Total_Fungi	3	0.0194**
Souche	NOTE_GLOBALE_FEU	3	2.716e-05***
Souche	NOTE_GLOBALE_EPI	3	0.0246*

Tableau 3 : ANOVA sur les variétés du panel (40 Variétés).

Les deux graphiques ci-dessous (**Figure 9**) représentent les quantités de *M. nivale* et de *M. majus* en pg/ng d'ADN total, retrouvées dans les grains post récolte. Le graphique est illustré en associant la note délivrée par le CTPS en post inscription pour la sensibilité à la DON ?, mycotoxines produites notamment par Fusarium graminearum.

Les analyses qPCR sur les grains récoltés dans les différents essais ont permis de mettre en évidence des différences variétales vis-à-vis de *Microdochium*. Elles ont également permis de mettre en évidence une réponse des variétés sensiblement identiques entre *M. majus* et *M. nivale*, même si ce résultat reste à confirmer avec plus de sites et de données.

Les 2 années furent très différentes en termes de conditions climatiques et n'ont pas permis de mettre en avant la réussite d'une méthode d'inoculation. Sans l'outil qPCR, il est très compliqué de dissocier les maladies sur épis dues à *Microdochium* spp. ou à Fusarium spp. Un effet variétal existe dans la résistance aux *Microdochium* spp. et il peut s'avérer différer de celui à la Fusarium gramineaum, mettant en évidence des mécanismes de résistances potentiellement différents contre ces agents pathogènes.

Figure 9 : Quantité de M.majus et M.nivale en pg/ng d'ADN total sur grains post récolte.



# 5. Outil d'imagerie multispectrale du taux de grains attaqués par *Microdochium*

Un des objectifs de ce projet était de tester la capacité d'un imageur multispectral, à différencier des grains présentant des symptômes de *Microdochium* spp. ou de Fusarium graminearum afin de pouvoir réaliser un classement de la résistance variétale à ces deux bioagresseurs. La capacité de l'imagerie multispectrale à classer les grains en 3 groupes : grains contaminés par Fusarium spp, grains contaminés par *Microdochium* spp. et grains non contaminés par ces agents pathogènes ; a été testée pendant 2 ans.

#### Matériel & Méthodes

L'imageur multispectral utilisé par le GEVES est le VideometerLab3, possédant 20 LEDs de longueurs d'onde différentes, allant de l'ultra-violet à l'infra-rouge (de 375 nm à 970 nm). L'image multispectrale est obtenue en éclairant successivement l'échantillon avec ces 20 LEDs (**Figure 10**).

A partir de 6 essais contaminés par ce complexe fusariose en 2016 et 2017, une base de données de référence de 8090 grains a été constituée, en couplant les données multispectrales, les analyses visuelles et microbiologiques (Figure 11).

Une bibliothèque de référence appelée BlobCollection a été créée : à chaque image de grain, appelé Blob, est associé sa classe d'appartenance (**Figure 12**).

Vingt-sept caractéristiques décrivant les grains ont été définies à l'aide de l'outil BlobTool, comprenant 12 caractéristiques spectrales issues de la classification des grains en 2 groupes par Analyse Discriminantes Canonique normalisée (nCDA), mais aussi 15 caractéristiques morpho-métriques (surface, forme, texture du grain).

Le modèle de classification utilisé est un Extra Trees Classifier basé sur des forêts d'arbres de décisions. Le GEVES a développé, sous Matlab, de nombreux modèles de classification en modifiant la base d'apprentissage (enrichissement ou équilibration des données) et/ou les caractéristiques utilisées.

Figure 10 : VideometerLab3 : images multispectrales d'un grain sain et d'un grain fusarié





Figure 11 : Schéma de la construction des données de références.



Figure 12 : Exemple de BlobCollection.

#### Résultats

Le choix d'un modèle équilibré, avec un même nombre de grains pour chaque classe, en utilisant 50% des grains comme base d'apprentissage, a permis d'obtenir un taux moyen de **81.51%** de bonne classification (**Figure 13**). Cependant, Le taux de bonne classification n'est que de **63.03%**, lors de la validation externe sur les 4046 grains non utilisés lors de la phase d'apprentissage (**Figure 14**). Le modèle à 3 classes présente des difficultés pour différencier les grains sains des grains contaminés par *Microdochium* spp (**Figure 15**).

		Modèle						
		Sain	Fusa	Micro				
Analyse sanitaire	Sain	87%	2%	11%				
	Fusa	5%	87%	8%				
	Micro	20%	6%	75%				



Figure 13 : Matrice de confusion sur l'ensemble des données, soit 8090 grains.

Figure 14 : matrice de confusion sur l'ensemble des données externes, soit 4046 grains.



Figure 15 : Echantillon de grains sains (pastilles vertes) et de grains Microdochium (pastilles bleues) issus de la collection de référence

Dans les essais avec de la fusariose sans *Microdochium* spp., l'algorithme Fusaspectral blé tendre à 2 classes (grains fusariés, grains indemnes), développé par le GEVES en 2015 demeure la référence (R<sup>2</sup>=0.95) par rapport à l'algorithme à 3 classes (R<sup>2</sup>=0.74). Pour améliorer la classification à 3 classes, deux pistes pourraient être poursuivies : élargir la base de référence et modifier les longueurs d'onde présentes dans le Videometer. Une étude hyperspectrale pourrait être menée pour déterminer les longueurs d'onde les plus discriminantes pour la classification des grains en 3 groupes. Les longueurs d'onde retenues pourraient ensuite être introduites dans une nouvelle version du Videometer.

# 6. Discussion

Nous avons, au cours de ces 3 années de travaux, constitué une base de données riche à partir de laquelle nous avons pu apporter des réponses aux principales interrogations à l'initiative de ce projet. Les données acquises confirment l'importance de *M. majus* et *M. nivale*, plus fréquemment retrouvés que *F. graminearum* pendant les 3 années du projet. Souvent confondu et mal interprété entre la septoriose sur feuilles ou les fusarioses sur épi, l'impact de *Microdochium* sur les cultures céréalières en France a pu être sous-estimé par la sélection variétale.

Selon les conditions environnementales et climatiques, les niveaux de pression ont été très variables. En effet, dans les deux sites Centre, les symptômes et les quantités de *Microdochium* spp. ont été plus marqués que dans le site Nord, à Cappelle en Pévèle (59). De plus, au cours des 3 années de recherche, l'effet année était un facteur important sur le développement de *Microdochium* spp. Une pression forte a été observée sur tout le territoire en 2016, comparativement aux années 2015 et 2017, peu favorables au développement des champignons *Microdochium*. L'année 2016 a été caractérisée par de fortes pluies, un rayonnement lumineux insuffisant associés à des températures fraiches entre l'épiaison et la fin du remplissage. Ces critères sont décrits dans la littérature comme étant plus favorables à *Microdochium* spp. et sont en adéquations avec les notations et résultats obtenus en 2016.

L'ensemble des résultats suggère fortement que *Microdochium* spp., est plus sensible aux conditions climatiques qu'aux conditions agronomiques. En effet, nous n'avons pas mis en évidence de conditions agronomiques favorables à *Microdochium* contrairement à *F. graminearum* qui est plus fréquent après un précédent maïs en non labour si les conditions climatiques sont favorables.

Les résultats des essais afin de constituer un protocole d'inoculation pour les acteurs de la sélection en conditions semi contrôlées, ont révélé et mis en évidence les difficultés de la mise en place d'un protocole d'inoculation pour un phénotypage discriminant. Celui-ci s'est avéré plus difficile sur épis que sur feuilles. La présence sur épi de Fusarium graminearum et de *Microdochium* spp., est impossible à distinguer de façon claire et rigoureuse à l'œil nu. Seule l'analyse moléculaire permet de faire la différence. Parallèlement, l'étude via le Videometer a, elle aussi démontré des difficultés quant à la possibilité de différencier des grains présentant des symptômes de *Microdochium* spp. et de F.graminearum dans un objectif de pouvoir réaliser un classement de la résistance variétale à ces deux bioagresseurs.

La forte pression naturelle de *Microdochium* en 2016 ne nous a pas permis mettre en évidence le stade optimal pour inoculer artificiellement le champignon au champ. De même, en 2015 et 2017, l'inoculation artificielle n'a pas permis un développement significatif du champignon quelle que soit la modalité d'inoculation ou une interaction trop forte avec *F. graminearum* a empêché une notation visuelle spécifique de *Microdochium*. Néanmoins, des tendances ont été observées suggérant que plusieurs applications entre Mi- Epiaison et Fin-Floraison pourraient être efficaces. De plus, des ajustements méthodologiques sont encore possibles pour réussir à maîtriser ces inoculations artificielles (meilleure maitrise de la brumisation, gestion de la pression exercée par *F. graminearum*, bâchage...).

Concernant la résistance variétale, les essais menés sous inoculations artificielles et grâce à la pression naturelle de 2016, démontrent qu'il existe des différences variétales dans la résistance. Ainsi, du progrès génétique est possible pour lutter contre *Microdochium* même si aucune résistance totale n'a été observée. Nous n'avons pas mis en évidence de différences marquantes pour les variétés testées entre *M. majus* et *M. nivale*. Dans l'ensemble, les variétés sensibles à *F. graminearum* le sont aussi à *Microdochium* spp. Cependant, les résultats suggèrent des différences pour certaines variétés, pouvant ainsi mettre en évidence des mécanismes de résistance différents. Enfin, certains résultats suggèrent que les variétés présentent des niveaux de résistance différents selon l'organe attaqué. Ainsi, certaines variétés semblent plus sensibles sur feuilles que sur épis (ou l'inverse).

L'imagerie multispectrale, avec une vingtaine de longueurs d'onde a montré ses limites pour différencier finement des symptômes proches entre Fusarium et *Microdochium*. Mais l'imagerie hyperspectrale avec plus de longueurs d'onde (>200) et sur une gamme plus étendue, du visible au proche infrarouge, permettrait d'identifier finement les longueurs d'onde spécifiques de *Microdochium* et de les comparer à celles déjà identifiées pour Fusarium graminearum, dans le programme CASDAR IRIGAM. L'ensemble de ces résultats sur la résistance variétale reste à être confirmé par des essais complémentaires.

#### 7. Conclusion et perspectives

Ainsi, cette étude a mis en évidence l'importance de travailler cette thématique et a démontré que le levier génétique pourrait

être un levier utile dans la lutte contre *Microdochium*. Il a également permis de mettre en évidence des verrous méthodologiques notamment dans la maîtrise des inoculations artificielles. Les limites identifiées sont en partie liées aux manques de connaissances sur la biologie de ces champignons. Ainsi, des efforts de recherche doivent être menés en axant les travaux autour de la compréhension et la maîtrise des infections de *Microdochium* spp. afin de fournir à la filière céréalière française, des méthodes et outils pour optimiser la sélection de variétés tolérantes et lutter ainsi plus efficacement contre la microdochiose. Ces travaux seront menés dans le cadre du nouveau projet FSOV : FSOV RESISTAMICRO.

# Références bibliographiques

**Observation CartoFusa**, 2013. BayerCropScience, Communication personnelle

Gourdain E., Batina H., Du Cheyron P., Fourrey A., Gélisse S., Grignon G., Laval V., Maumené C., Méléard B. et Valade R., 2016. Lutte contre les fusarioses des épis de blés : quantification des espèces du complexe fusarien, facteurs de risque et méthodes de lutte, Innovations agronomiques, vol. 49, pp. 133-145.

Walker A.S., Auclair C., Gredt M., Leroux P., 2009. First occurrence of resistance to strobilurin fungicides in Microdochium nivale and Microdochium majus from France naturally infected wheat grains. Pest management Science, vol. 65 (8), pp 906-915 Parry D.W., Rezanoor H.N., Pettitt T.R., Hare M.C., Nicholson P., 2008. Analysis of Microdochium nivale isolates from wheat in the UK during 1993. Annuals of Applied Biology, vol 126 (3), pp 449-455

**ANR DON&CO** (2010-2013) - Mycotoxinogénèse chez le blé : de la diversité de la microflore fusarienne à la toxicologie. INRA MycSA (coordinateur) Arvalis-Institut du végétal, INRA Bioger), INRA Agroécologie, INRA Toxalim, ENVT, INRA BIOGECO.

CASDAR ECOFUSA (2010-2013) - Lutte contre les fusarioses des épis de blés : de l'utilisation raisonnée des fongicides aux méthodes de luttes alternatives. ARVALIS – Institut du végétal Coordinateur), INRA BIOGER, INRA MycSA.

# [NIL-N] Caractérisation de régions chromosomiques pour augmenter l'efficacité d'utilisation de l'azote et la teneur en protéines

Mickaël THROUDE<sup>1</sup>\*, Jacques LE GOUIS<sup>2</sup>, Jérôme SALSE<sup>2</sup>, Caroline PONT<sup>2</sup>, Séverine ROUGEOL<sup>2</sup>, Katia BEAUCHENE<sup>3</sup>, Jérémy DERORY<sup>4</sup>, Céline DUQUE<sup>5</sup>, Didier TROPEE<sup>6</sup>, Emmanuel HEUMEZ<sup>7</sup>, Alain Chassin<sup>8</sup>, Nadine DURANTON<sup>1</sup>, Stéphane LAFARGE<sup>1</sup>, Sébastien Praud<sup>1</sup>

1 - BIOGEMMA - Centre de Recherche de Chappes 63720 Chappes, France

- 2 INRA, UCA UMR 1095 GDEC 63039 Clermont-Ferrand, France
- 3 Arvalis-Institut du Végétal 45 Voie Romaine 41240 Ouzouer-le-marché, France
- 4 Limagrain Europe Centre de Recherche de Chappes 63720 Chappes, France
- 5 Limagrain Europe Ferme de l'Etang 77390 Verneuil L'Etang, France
- 6 INRA, UPS, CNRS, AgroParisTech UMR 320 GQE 91190 Gif sur Yvette, France
- 7 INRA UE 972 GCIE 80203 Péronne cedex, France
- 8 INRA UE 1375 PHACC 63039 Clermont-Ferrand, France
- \* Coordinateur : Mickaël THROUDE, mickael.throude@biogemma.com

#### 1. Introduction

Notre projet a été motivé par l'attente de l'agriculture française et européenne de pouvoir répondre aux besoins d'augmentation et d'optimisation de la production des céréales tout en maintenant la qualité des produits, en limitant les coûts de production et en réduisant les effets négatifs sur l'environnement. L'azote est un élément nutritif essentiel qui influence fortement le rendement en grains et la teneur en protéines. Cependant, La production, le transport et l'application des engrais azotés sont très consommateurs d'énergie et le prix de ces engrais est indexé au prix du pétrole. Ainsi le prix des engrais simples azoté a très fortement varié dans les dix dernières années (INSEE). Cela se traduit pour l'agriculteur par un coût équivalent à environ 25-30% des charges opérationnelles. Un des moyens de diminuer ces charges est d'améliorer l'efficacité d'utilisation par la culture et notamment l'efficacité d'absorption. En effet seuls environ 60% de l'azote disponible dans le sol seront absorbés par le blé (Gaju et al., 2011 ; Cormier et al. 2016). Une partie est perdue par lessivage entrainant une pollution des eaux de surface et des nappes phréatiques. Une autre partie est volatilisée sous forme d'oxyde nitreux, un gaz à effet de serre supérieur à celui du gaz carbonique. Ainsi, les engrais azotés seraient responsables tout au long de leur cycle d'environ 50% des émissions de gaz à effet de serre d'origine agricole. Pour limiter ces impacts négatifs, il peut être envisagé une diminution des apports azotés. Mais cela doit se faire tout en maintenant le rendement et la teneur en protéines. Il semble donc essentiel de favoriser la sélection de variétés de blé plus efficace dans leur utilisation de l'azote pour produire du rendement et des protéines.

L'efficacité d'utilisation de l'azote peut être définie comme la quantité de grain produit par unité d'azote disponible pour la plante. Pour tenir compte de la corrélation négative entre production de grain et concentration en protéines, il est intéressant d'utiliser la déviation à la régression linéaire entre ces deux caractères. Monaghan *et al.* (2001) ont proposé ce critère pour la sélection qu'ils ont dénommé GPD, pour *Grain Protein Deviation*.

Il a été montré qu'il existait des différences génotypiques significatives pour les composantes de l'efficacité d'utilisation de l'azote (e.g. Van Sanford et MacKown 1986; Fossati *et al.* 1993) et pour le GPD (e.g. Oury et Godin, 2007; Bogard *et al.*, 2010). Cela a permis de conduire des études génétiques pour identifier les régions chromosomiques impliquées dans la variabilité observée (Laperche *et al.*, 2007; Bogard *et al.*, 2011). Laperche

*et al.* (2006 ; 2007) ont particulièrement utilisé une population de lignées recombinantes issue du croisement entre les variétés Arche et Recital qui sont contrastées pour leur capacité à utiliser l'azote (Le Gouis *et al.*, 2000). Ainsi, cette population a permis d'identifier 37 régions chromosomiques principales (appelées QTL ou Quantitative Trait Locus) ayant un impact sur l'utilisation de l'azote (Laperche *et al.*, 2007). De son côté, les sociétés Biogemma et Limagrain ont construit une population de type diallèle nommée Metapop et composée de 5 parents élites contrastés pour leur teneur en protéines, composantes du rendement et efficience d'utilisation de l'azote. Cette population, évaluée de 2005 à 2007 sur 5 lieux et 35 environnements dans le cadre des projets ANR WheatPerformance et WheatGrowth, a permis de mettre en évidence 82 régions chromosomiques d'intérêt pour les caractères étudiés.

En parallèle de ces études conduites sur le blé tendre, des analyses ont été réalisées aux Etats-Unis chez le blé dur pour identifier des régions influençant la concentration en protéines du grain. Un QTL a été identifié sur le bras court du chromosome 6B à partir d'une population de lignées recombinantes (Joppa et al., 1997). Le clonage positionnel du QTL a permis d'identifier un facteur de transcription dénommé NAM-B1, responsable du QTL GPC-B1, augmentant la teneur en protéines sans diminuer le rendement (Uauy et al., 2005), ce qui correspond à la définition du GPD. Il a par ailleurs été montré que ce gène accélère la sénescence des feuilles durant le remplissage du grain, augmente la remobilisation d'azote et conduit à une meilleure allocation de l'azote vers le grain (Waters et al., 2009). Par la suite, il a été montré que ce QTL n'est pas seulement associé à une accélération de la sénescence durant le remplissage du grain mais aussi à une accélération du rythme de développement avant floraison (Lacerenza et al., 2010). L'allèle favorable de NAM-B1 a notamment été utilisé en sélection au Canada pour augmenter la concentration en protéines de blé de printemps (dePauw et al., 2007). Pour cela l'allèle de l'accession israélienne FA15-3 a été transféré au blé hexaploïde Glupro puis à la variété inscrite Lillian, avec une augmentation de concentration en protéines sans effet négatif apparent sur le rendement. Ce gène, qui n'a pas encore été testé dans du matériel de type hiver adapté aux conditions nordeuropéennes pourrait présenter un intérêt pour améliorer l'efficience d'utilisation de l'azote.

Enfin, des études australiennes ont montré l'intérêt de limiter l'étendue du tallage dans des situations de déficit en eau. Une réduction du tallage limite le développement foliaire, limite la transpiration et donc la consommation d'eau qui peut être utilisée plus tardivement durant le remplissage du grain. Pour cela des lignées possédant un allèle limitant le tallage (allèle Rtin pour restricted tillering) ont été obtenues à partir de différentes sources. Ainsi en conditions de contraintes hydriques des lignées à faible tallage montraient un rendement supérieur aux variétés classiques (+11%) alors que leur rendement était plus faible en conditions non stressées (Mitchell et al., 2013). Le gène Tin, situé en partie distale du chromosome 1A, est capable d'induire une réduction du tallage de plus de 50% (Spielmeyer et al., 2004). Cet allèle pourrait présenter un intérêt dans des systèmes à faible niveau d'azote en permettant une économie d'azote durant la phase pré-floraison qui pourrait se traduire par de plus fortes concentrations en protéines grâce à l'absorption postfloraison.

Une approche couramment utilisée pour étudier ce type de caractère quantitatif est le recours au développement de lignées quasi-isogéniques (NIL pour *Near Isogenic Lines*). En effet le matériel NIL est un outil idéal pour les études génétiques, l'exploration des effets géniques, le criblage de marqueurs moléculaires liés à un locus d'intérêt, l'analyse de son expression et le clonage du gène cible. Depuis les années 90, les NIL ont été largement utilisés pour caractériser et cloner des QTL chez les plantes (Miralles *et al.*, 1997 ; Jonnala *et al.*, 2010 ; Wang *et al.*, 2018).

L'ensemble de ces résultats tendent à démontrer l'intérêt de ces régions chromosomiques ou de ces gènes en sélection, pour augmenter l'efficacité d'utilisation de l'azote et la teneur en protéines. Le projet mené de 2014 à 2018 visait à allier les efforts des partenaires du projet (Biogemma, Limagrain, INRA et ARVALIS) pour évaluer l'effet des gènes majeurs NAM-B1 et *Tin*, ainsi que les allèles portés par plusieurs QTL détectés dans les populations recombinantes Arche x Recital et Metapop. Ce travail s'est appuyé sur l'utilisation de matériel isogénique NIL construit en amont et apporté au projet par les partenaires. Le projet a permis d'évaluer l'effet de ces régions et gènes via la mise en place d'un réseau multilocal en champ, comportant différentes stratégies de fertilisation azotée ou de densité de plantes. Les lignées démontrant un impact majeur ont ensuite été expertisées plus finement dans un réseau d'essai plus étendu et représentatif du marché français, visant à mieux évaluer le potentiel des gènes travaillés pour la sélection et tenter d'appréhender les interactions Génotype Х Environnement. Enfin, toujours dans l'objectif de faciliter l'utilisation des gènes et du matériel par les sélectionneurs, nous avons caractérisé génétiquement les lignées NIL, évalué la taille des introgressions et développé de nouveaux marqueurs moléculaires pour suivre ces régions.

# 2. Matériel et méthode

2.1 - Caractérisation écophysiologique et agronomique des régions chromosomiques d'intérêt

#### Matériel végétal

La validation de l'effet de régions chromosomiques d'intérêt pour augmenter l'efficacité d'utilisation de l'azote et la teneur en protéines a été réalisée par l'évaluation de lignées quasiisogéniques (NIL) crées en amont du projet. Ces lignées NIL ont été obtenues par au moins trois rétrocroisements successifs du parent donneur (allèle positif) avec des lignées receveuses (NIL BC pour Backcross) ou par autofécondations successives de lignées hétérozygotes pour l'allèle d'intérêt (lignées NIL HIF pour *Heterogenous Inbred Families*). La construction de ce matériel a été pilotée par l'utilisation de marqueurs moléculaires encadrant les QTL ou les gènes d'intérêt. Quel que soit le type de construction utilisé, une dernière étape d'autofécondation des plantes homozygotes au locus d'intérêt a permis de générer des couples de lignées homozygotes (couples NIL) pour chacun des deux allèles positifs et négatifs. Les évaluations au champ ont été réalisées sur ces couples NIL provenant de 4 populations ou gènes décrits ci-dessous.

#### Lignées GPC Arche x Recital, 12 couples NIL BC :

Deux QTL ont été identifiées en conditions contrôlées à faible niveau d'azote dans la population Arche x Recital (Laperche *et al.*, 2006). Le QTL sur le chromosome 1B joue sur la longueur racinaire, celui sur le 7A sur la capacité d'absorption d'azote. Les couples étudiés correspondent à ces deux régions chromosomiques, 1B et 7A, pour lesquelles 3 séries de rétrocroisements ont été effectuées dans chacun des fonds parentaux Arche et Recital.

#### Lignées Metapop, 7 couples NIL BC et 21 couples NIL HIF :

La Metapop est une population de type diallèle incomplet comprenant 551 HD (haploïde doublée) créées à partir de 5 variétés de blé d'hiver élites commerciales : Apache, Aztec, Autan, Cezanne et Uli. Douze QTL influençant le rendement et la teneur en protéine du grain ont été identifiés dans cette population. Ces QTL ont été l'objet de construction de matériel isogénique, 21 couples NIL HIF et 7 couples NIL BC par 3 rétrocroisements dans les fonds Apache ou Cezanne. Les lignées isogéniques sont listées dans le tableau ci-dessous.

QTL	Chrom.		Couples NIL HIF ou NIL BC (grisé)										
Q5	1A	5039_Q5	A P UL_Q5										
Q28	3B	A P A U_Q28	A P A Z_Q28	UL A Z_Q28									
Q29	3B	A P C E_Q29	A P UL_Q29										
Q30	3B	A P A U_Q30	A P A U_Q30_2	A P C E_Q30	A P C E_Q30_2								
Q36	4A	A U A Z_Q36	A UA Z_Q36_4	A UA Z_Q36_3	A Z C E_Q36	A Z C E_Q36_2							
Q54	6A	5108_Q54	5150_Q54_2										
Q59	6B	A U C E_Q59											
Q60	6B	5197_Q60											
Q65	6D	5013_Q65	5013_Q65_2										
Q76	7A	A P A U_Q76	A P A U_Q76_3	A P UL_Q76									
Q81	7D	A P A Z_Q81	A P A Z_Q81_2										
Q82	7D	5025_Q82											

#### Lignées NAM-B1, 12 couples NIL BC :

L'allèle favorable du gène *NAM-B1* a été transféré dans quatre fonds génétiques différents : les variétés Premio et Skerzzo qui montrent un écart positif à la relation négative entre le rendement et la concentration en protéines (GPD+) et les variétés Arlequin et Recital qui montrent un écart négatif (GPD-). Trois donneurs ont été utilisés (Yecora Rojo, Kern et Anza) ; un couple isogénique correspondant à une combinaison receveur-donneur. Neuf de ces combinaisons ont été testées (deux couples correspondent à une même combinaison receveur-donneur).

#### Gène TIN-1A, 3 couples NIL BC :

L'allèle du gène *Tin* associé à la perte de fonction du gène et induisant une réduction du nombre de talles a été introgressé à partir du génotype Oligoculm, dans les 3 fonds génétiques Apache, Caphorn et Korelli.

#### Réseau expérimental

Un réseau expérimental de 20 essais en champ a été mis en place entre 2015 et 2017 en France sur 10 lieux (cf. figure ci-dessous).



Chaque couple de lignée isogénique a été évalué sur une, deux ou trois années, avec pour objectif de valider l'effet des QTL et d'évaluer leur potentiel (volet 1), puis d'effectuer une caractérisation écophysiologique et agronomique plus fine (volet 2). En fonction des traits travaillés, une partie des essais a visé spécifiquement la caractérisation du tallage et les composantes de rendement en modulant la densité de semis (densité optimale régionale, densité -60% et densité +50%). L'autre partie des essais était focalisée sur les dynamiques d'absorption/remobilisation de l'azote en faisant varier les modalités d'apport azoté (dose bilan X, dose X-80 kg N / ha et sur certains lieux X+80 kg / ha).

Le plan parcellaire utilisé était organisé en sous essais, en fonction de l'origine génétique du matériel (Metapop, Arche x Recital, *NAM-B1* et *TIN-1A*, incluant les témoins parentaux associés), puis par type d'essai (Tallage-rendement ou Azote). Enfin chaque sous essai a fait l'objet d'une randomisation de type « imbriquée totale », c'est-à-dire que la randomisation conservait la proximité des deux lignées d'un couple isogénique.

#### Phénotypage

Sur chaque expérimentation, nous avons mesuré le nombre de plantes par m<sup>2</sup> à la levée (PPA) et le nombre d'épis par m<sup>2</sup> à maturité (SA) comptés pour les deux mesures sur les mêmes 2 placettes de 2 mètres linéaires, la date d'épiaison (Z55), la hauteur des plantes (PH), le poids spécifique (PS), le rendement à 15 % d'humidité (GY15), le poids de mille grains à 15% d'humidité (PMG) et la concentration en protéines (GPC). A partir de ces variables ont ensuite été calculés le nombre de grains par m<sup>2</sup> (GPA), le nombre de grains par épis (GPS), le nombre d'épis par plante (SPP) et l'écart à la régression concentration en protéine / rendement nommé GPD. Sur une partie des essais, des variables complémentaires ont été acquises, le stade épi 1 cm (Z30), la longueur et largeur moyennes du grain (GL, GW), l'indice de nutrition azotée à floraison (INN\_flo), la durée de sénescence (dursen) et en fonction de leur apparition des notes de verse ou de maladies.

#### Analyse des effets

Un script d'analyse des couples NILs a été développé durant le projet. Ce script permet l'automatisation des tests statistiques, la production de rapports de synthèse et des visualisations graphiques des effets observés (sous forme de graphe radar). Pour chaque lieu et année, les lignées isogéniques ont été analysées couple par couple pour chaque modalité azotée ou densité de semis, puis pour toutes les modalités confondues. Pour chaque variable, nous avons réalisé un test basé sur la méthode Fisher LSD (*Fisher's Least Significant Difference* test) avec une p-value fixée à 5% pour détecter les différences significatives entre les couples isogéniques. L'analyse multilocale et multi-année a été réalisée à l'aide du package R Metafor (rma.mv) qui permet de caractériser un effet et son risque associé (seuil également fixé à 5%). Lors de ces analyses, les essais présentant des défauts d'implantation (nombre de plantes par m<sup>2</sup> significativement différents entre lignées lors du test LSD avec une p-value à 5%) ont été exclus.

#### 2.2 - Caractérisation génétique des lignées NIL

#### Analyse du fond génétique et de la taille des introgressions

La totalité du matériel végétal isogénique disponible a été génotypée par la méthode GBC (Génotypage par Capture et séquençage). La capture a été réalisée par la technologie « NimbleGen *Sequence Capture* » en phase liquide et le séquençage par Illumina MiSeq par un protocole déjà mis en place par Biogemma. Nous avons utilisé 9 278 marqueurs : 2 723 marqueurs provenant de la puce de génotypage 90K, 3 240 marqueurs provenant de la puce de génotypage Axiom 420K, 3 305 marqueurs Biogemma et 10 marqueurs INRA. Ces marqueurs ont été sélectionnés sur la base de leur répartition homogène sur tout le génome du blé avec un effort de densification de la couverture dans les zones couvertes par les QTL ou gènes suivis.

A partir des informations générées par ces marqueurs, nous avons calculé un index de similarité qui permet d'évaluer le niveau d'homologie au sein de chaque couple isogénique. En parallèle, une visualisation graphique de chaque couple isogénique a été mise en place afin de confirmer l'état d'homogénéité de chaque chromosome mais également de valider la présence des différents QTL dans leur intégralité dans le matériel. Ces visualisations ont été réalisées à l'aide du logiciel Spotfire.

Ancrage physique des régions d'intérêt et développement de nouveaux marqueurs

A partir des marqueurs encadrant les QTL, un ancrage physique a été réalisé sur l'assemblage IWGSC RefSeq V1.0 de la variété de blé hexaploïde Chinese Spring (IWGSC 2018). A partir de cette position physique, il a été possible de sélectionner de nouveaux marqueurs SNP dans plusieurs QTL de la Metapop. Les SNP polymorphes ont été génotypés par la méthode KASPar (KBiosciences, Herts, UK) sur les 551 individus de la population. Les oligonucléotides KASPar ont été dessinés avec le logiciel Primer picker (KBioscience), les amplifications PCR ont été réalisées sur un hydrocycleur (LGC genomics), sur 50 cycles à 57°C et les lectures ont été effectuées sur ABI PRISM 7900 HT.

#### 3. Résultats et discussions

#### 3.1 - Caractérisation écophysiologique et agronomique des lignées isogéniques

#### Caractérisation des lignées GPC, Arche x Recital

Douze couples de lignées isogéniques ont été testés. Ils correspondent à deux régions chromosomiques (1B et 7A) pour lesquelles des rétrocroisements ont été effectuées dans chacun des fonds parentaux (Arche et Récital). En 2015, les 12 couples de lignées isogéniques (ARE) ont été implantés dans un réseau de 5

essais à trois doses d'azote. En 2016, quatre couples et les deux parents ont été expérimentés à deux doses d'azote en cinq lieux.

Une analyse globale a été réalisée en comparant pour chacune des deux régions chromosomiques, l'effet des deux allèles. Globalement, nous n'observons pas d'effet des deux régions sur le rendement, la concentration en protéines, le nombre de plantes / m<sup>2</sup>, le nombre d'épis par m<sup>2</sup> et le nombre de grains / m<sup>2</sup>. Pour les deux régions, un effet significatif est observé pour la date d'épiaison et la hauteur des plantes. Pour la région du chromosome 1B, l'allèle provenant de Arche rend les plantes plus tardives (+1.2 j) et plus hautes (+1.6 cm). C'est le contraire pour la région sur le chromosome 7A, l'allèle provenant de Arche rendant les plantes plus précoces (-0.8 j) et plus courtes (-1.8 cm). Enfin, un effet de la région du chromosome 7A est détecté sur le poids de mille grains, l'allèle provenant de Arche l'augmentant de +0.5 g (cf figure ci-dessous).



MON = INRA Estrées-Mons, CHA= Arvalis Châlons, VER= Limagrain Verneuil, IDF= Arvalis Ile-de-France, CLF = INRA Clermont-Ferrand, LML = INRA Le Moulon, suivi de l'année d'expérimentation.

Dans tous les cas, l'effet de l'environnement est très significatif et aucune interaction allèle x environnement n'est mise en évidence.

#### Caractérisation des lignées NAM-B1

Dix couples de lignées isogéniques pour la région sur le chromosome 6B contenant *NAM-B1* ont été testés durant le projet. Ils correspondent à quatre fonds parentaux : (i) Arlequin et Récital, deux variétés à faible concentration en protéines (GPD-), (ii) Premio et Skerzzo, deux variétés à forte concentration en protéines (GPD+). En 2016, les 10 couples ont été testés à deux doses d'azote en quatre lieux. En 2017, trois couples ont été testés à une seule dose d'azote (dose bilan) en cinq lieux et un couple a été expérimenté sur la plateforme PhenoField d'Ouzouer-le-marché d'Arvalis à deux niveaux de disponibilité en eau (irrigué et sec) et deux doses d'azote (dose bilan et 0).

Une analyse globale a été réalisée sur les deux années d'expérimentation en considérant chacune des 9 combinaisons receveur-donneur de façon séparée. L'effet de l'allèle est significatif pour le rendement dans 5 combinaisons sur 9. Dans 2 de ces cas (fond Premio), c'est allèle non fonctionnel (nul) qui augmente le rendement, et dans 3 cas (fonds Arlequin et Recital) c'est l'allèle fonctionnel (GPC). L'effet de l'allèle est significatif pour la concentration en protéines dans 8 cas sur 9. Dans tous les cas l'allèle fonctionnel augmente la concentration en protéines. L'interaction entre l'effet de l'allèle et de l'environnement (combinaison année-lieu-traitement) est presque toujours significative (8 cas sur 9) indiquant que l'impact de la région NAM-B1 dépend fortement du milieu dans lequel elle est testée. Une partie des effets peut provenir en fait de la présence dans les donneurs du gène Yr36 de résistance à la rouille jaune (Uauy et al. 2005), maladie particulièrement marquée en 2016. Il est difficile de conclure sur l'impact du receveur sur l'effet de l'allèle favorable. Mais le plus fort effet sur la teneur en protéines est observé dans le fond Premio (Pre-KE) et le plus faible effet dans le fond Arlequin (Arl-AZ), un GPD+ et un GPD- respectivement.

En moyenne sur toutes les combinaisons, la différence entre les deux allèles est de +0.2 q / ha pour le rendement et +0.6 % pour la concentration en protéines en faveur de l'allèle fonctionnel (cf figure ci-dessous).



On remarque en tendance, une diminution de l'impact de l'allèle fonctionnel quand le rendement augmente. Ceci pourrait être dû à une limitation de la quantité d'azote remobilisable, quantité sur laquelle l'allèle fonctionnel à un impact. En effet, la quantité supplémentaire qui doit être remobilisée pour augmenter la concentration en protéines de 0.6% à 80 q / ha est supérieure à celle qui doit être remobilisée à 40 q / ha.

L'impact du gène sur les autres caractères est significatif mais faible. Les résultats sont variables en fonction des combinaisons receveur-donneur, mais en moyenne, les lignées possédant l'allèle fonctionnel sont un peu plus précoces à l'épiaison (-1.1 jour), un peu plus courtes (-1.1 cm). Il n'y a pas d'effet moyen sur le PMG.

#### Caractérisation des lignées TIN 1A

Trois couples de lignées isogéniques dans les fonds génétiques Apache, Caphorn et Korelli ont été implantés en 2017 dans un réseau de 2 essais à trois doses de semis.

La lignée dans le fond Caphorn a montré dans les deux lieux un très faible niveau de rendement probablement dû à un très bas niveau de tolérance au froid. Elle a été retirée de l'analyse empêchant la caractérisation de l'effet du gène dans ce fond génétique.

L'analyse de variance par lieu montre un effet génotype toujours significatif et un effet densité de semis généralement significatif, sauf pour le rendement. L'effet densité est très fort pour le nombre de plantes et le nombre d'épis, mais les compensations existantes entre ces composantes font que l'effet est moins fort sur le nombre de grains et n'est plus significatif pour le rendement. L'analyse de l'effet de l'allèle dans les deux lieux montre un effet significatif pour l'ensemble des variables à l'exception de la date d'épiaison à Clermont-Ferrand. L'allèle *Tin* limite le nombre d'épis (-76 épis / m<sup>2</sup>). Il augmente le poids de mille grains et la hauteur des plantes et les rend plus tardives à l'INRA d'Estrées-Mons. L'impact moyen de cet allèle sur le rendement est négatif (-11.3 q / ha) et en contrepartie positif sur la concentration en protéines (+0.7).

Pour plusieurs variables, il existe une interaction de l'effet de l'allèle *Tin* et du fond génétique. Dans tous les cas l'effet de l'allèle *Tin* est plus fort dans le fond Korelli que dans le fond Apache. Par exemple pour le nombre de grains par m<sup>2</sup>, l'allèle *Tin* entraine une diminution de 2 907 grains / m<sup>2</sup> en moyenne pour Apache et de 7 482 grains / m<sup>2</sup> pour Korelli. Une seule interaction entre l'effet de l'allèle et la densité de semis a été détectée à Estrées-Mons pour le rendement. Dans ce cas, la différence de rendement entre les deux allèles a tendance à être plus faible pour la densité basse (-5.3 q / ha) que pour la densité haute (-10.1 q / ha).

Cette analyse valide dans deux des trois fonds génétiques l'effet fort sur le tallage du gène *TIN1*. Dans les conditions testées, l'effet est négatif sur le rendement mais ce matériel pourra être très pertinent à tester dans d'autres environnements où par exemple un développement végétatif réduit peut être intéressant, par exemple sous contrainte thermique ou hydrique.

#### Caractérisation des lignées GPC et composantes du rendement, Metapop

Vingt-huit couples de lignées isogéniques ont été testés entre 2015 et 2017. Ce matériel correspond à 12 régions chromosomiques décrites par des QTL identifiés dans la population Metapop. Sur la base des résultats des essais ainsi que la caractérisation génétique du matériel, 4 couples ont été évalués une seule année puis stoppés, 17 couples ont été évalués durant deux années et 7 durant 3 années. Le tableau ci-dessous présente la répartition des couples isogéniques dans les différents essais : essais de caractérisation agronomique fine (volet2) ciblés sur la remobilisation de l'azote (notés « Azote ») ou sur le tallage et composantes du rendement (notés « Densité ») ou essais d'évaluation du potentiel des QTL effectués dans le volet 1 (notés « volet 1 »).

QTL	Chrom.	Couple NIL	essai 2015	essai 2016	essai 2017
Q5	1A	5039_Q5	Densité		Densité
Q5	1A	APUL_Q5	Densité	volet 1	Densité
Q28	3B	APAU_Q28	volet 1	Densité	
Q28	3B	APAZ_Q28	volet 1	volet 1	Densité
Q28	3B	ULAZ_Q28	Azote	volet 1	
Q29	3B	APCE_Q29	volet 1	Densité	
Q29	3B	APUL_Q29	volet 1	Stop	
Q30	3B	APAU_Q30	volet 1	volet 1	
Q30	3B	APAU_Q30_2	volet 1	Densité	
Q30	3B	APCE_Q30	volet 1	Stop	
Q30	3B	APCE_Q30_2	volet 1	volet 1	
Q36	4A	AUAZ_Q36	volet 1	Stop	
Q36	4A	AUAZ_Q36_4	volet 1	Azote	
Q36	4A	AUAZ_Q36_3	Azote	volet 1	Azote
Q36	4A	AZCE_Q36	volet 1	Stop	
Q36	4A	AZCE_Q36_2	volet 1	volet 1	
Q54	6A	5108_Q54	Densité	volet 1	
Q54	6A	5150_Q54_2	Densité	volet 1	
Q59	6B	AUCE_Q59	volet 1	Azote	
Q60	6B	5197_Q60	Densité	volet 1	Azote
Q65	6D	5013_Q65	volet 1	Densité	
Q65	6D	5013_Q65_2	volet 1	volet 1	
Q76	7A	APAU_Q76	volet 1	Densité	Densité
Q76	7A	APAU_Q76_3	Azote	volet 1	Azote
Q76	7A	APUL_Q76	volet 1	volet 1	
Q81	7D	APAZ_Q81	volet 1	volet 1	Azote
Q81	7D	APAZ_Q81_2	volet 1	Azote	
Q82	7D	5025_Q82	Azote	volet 1	

En 2015, 19 couples de lignées isogéniques ont été évalués, sur un réseau de 5 lieux et trois doses d'azote (volet 1). Une caractérisation plus fine (volet 2) a été effectuée sur 9 couples isogéniques sur un lieu, découpé en deux sous essais faisant varier 2 doses d'Azote ou 3 densités de plantes. Ces 9 couples NIL ont été sélectionnés à partir des résultats d'évaluation menés en 2014 par Limagrain et Arvalis en amont du projet. En 2016, 15 couples ont été évalués sur un réseau de 5 lieux et trois doses d'azote dans le volet 1. Une caractérisation plus fine de 8 couples NIL a été menée dans le volet

2 sur 1 lieu, selon les modalités décrites en 2015, à raison de 5 couples sur l'essai « Densité » et 3 couples sur l'essai « Azote ». En 2017, un réseau plus étendu de 7 lieux nous a permis d'étudier la «pénétrance» de 8 couples NIL, dans différents environnements représentatifs du marché français, et tenter d'appréhender les effets d'interaction Génotype x Environnement aux loci étudiés. Pour les essais « Azote », 4 couples NIL ont été testés à une seule dose d'azote (dose bilan) et 4 répétitions sur 5 lieux ; deux doses d'azote et 3 répétitions sur un lieu et 2 couples NIL sur la plateforme Phenofield d'Arvalis à deux niveaux d'irrigation et deux doses d'azote. Pour les essais « Densité », 4 couples NIL ont été évalués sur 3 lieux, selon 3 modalités de densité de plantes et 3 répétitions ; et sur 3 autres lieux avec deux densités de plantes et 3 répétitions.

L'analyse des essais a été réalisée par lieu et par année (données non présentées). Les résultats des années 2015 et 2016 nous ont permis de sélectionner les couples isogéniques et QTL les plus prometteurs pour réaliser l'étude de leurs effets dans le réseau d'essai plus large déployé en 2017. L'analyse globale regroupant la totalité des essais réalisées en 2015, 2016 et 2017 a été effectuée à l'aide d'un script développé dans le cadre du projet. Pour chaque couple de lignées isogéniques, une fiche descriptive des effets observés au sein des essais réalisés pour chaque année, pour chaque modalité testée ou toutes modalités confondues, a été produit. A titre d'exemple, la figure présentée page suivante présente la fiche de synthèse du couple isogénique 5039\_Q5.

Dans cette fiche, nous pouvons retrouver dans la partie A, le nom du couple isogénique testé (la lignée portant les allèles du donneur 5039\_Q5\_2.2D contre la lignée isogénique portant les allèles du receveur 5039\_Q5\_2.5R) ainsi que les effets attendus sur le QTL étudié (le QTL5 qui diminue le nombre d'épis par m<sup>2</sup> à maturité et augmente le poids de mille grains). Dans le tableau C, nous retrouvons les effets observés pour chacune des variables mesurées en 2015 et 2017 sur les 7 lieux testés pour les modalités de densité optimale (D), réduite (D-) et dense (D+) ainsi que toutes modalités densité confondues (toute\_D). Le tableau D, présente les résultats de la méta-analyse, de tous les lieux et années, pour chaque modalité dans l'unité de mesure de chaque trait. Enfin, la figure de la partie B, est une représentation graphique des effets de la méta-analyse dans le sens donneur/receveur, exprimés en pourcentage de la moyenne du trait.



D : NDEN_15AR NDEN_15ARVI]- NDEN_15ARVVI]- NDEN_15ARVVI]- NDEN_17ARVI NDEN_17ARVI NDEN_17ARVI NDEN_17ARVI NDEN_17I N	RVvil_v2tt D Vvil_v2tt D Vvil_v2tt D vizt toute_D ARVrottt D ARVrottt D ARVvitt D ARVvitt D ARVvitt D ARVvitt D ARVvitt D ARVvitt A ARVvitt D ARVvitt D	GL 000000000000000000000000000000000000	GPC -0.2 -0.2 0 -0.21 0 0 0 -0.17 -0.02 0.6 0.13 -0.02 0.6 0.13 -0.16 0.18 -0.2 0.03 0.03 0.03 0.05 0 0.05	GPD 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	GW 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	GY15 -3.41 -4.37 -6.09 -4.63 1.83 -0.83 0.5 -3.08 5.07 1.35 1.66 -1.57 1.87 -1.23 -0.29 -1.87 -1.87 -1.87 -0.93 -0.5 -0.72	HR -0.17 0.05 -0.05 -0.37 0.37 0.03 0.07 0.03 0.07 0.03 0.07 0.03 0.07 0.00 0.00	LS 00 -0.93 -0.47 00 00 00 00 00 00 00 00 00 00 00 00 00	Ndum 0 -0,63 -0.33 -0.48 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	PH 0.85 0.16 0.87 0.87 0.87 0.87 0.87 0.87 0.87 0.87 0.07 0.00 0.	PPA 7.62 3.49 -6.03 -5.7 -15.97 -14.3 -0.67 -24.80 -10.33 -0.67 -24.80 -10.33 -24.80 -10.33 -24.80 -10.33 -24.80 -10.33 -24.80 -10.33 -24.80 -14.33 -24.80 -14.33 -28.67 -14.33 -28.67 -14.5 -14.5	PS 000000000000000000000000000000000000	SA -47.62 -97.48 -94.29 -59.79 -58.73 -58.73 -58.7 -39.939 -770.33 -87.67 -71 -44.2 -71 -61 -58 -60.67 -72 -44.83 -48.33 -721.17	SPP -0.54 -0.94 -0.39 -0.03 -0.17 -0.1 0 0 0 0 0 0 0 0 0 0 0 0 0	TKW 2.36 -0.02 -0.02 -1.85 1.4 0 0 1.47 1.53 1.23 1.24 -0.13 1.24 -0.085 -0.02 1.03 -0.33 0.97 0.73 0.85	YR 000000000000000000000000000000000000	<b>Z30_calc</b> -3.22 -3.22 -0.96 -2.48 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	255 -1.37 -0.67 -0.67 -0.33 -0.33 -0.33 -0.33 -0.33 -0.33 -0.33 -0.33 -0.22 -0.33 -0.22 -0.31 -0.22 -0.33 -0.53 -0.55 -0.5
C: GL Tout -0.03 D D- D+	GPC -0.02 -0.06 0 0.06	GPD -0.05 -0.22 0.11	GW 0.06	GY15 -0.48 -0.87 -0.93 0.44	0.11 -0.04 0.03	e   0 -0.	LS No 82 -	<b>dum</b> 0.39	PH 0.3 0.23 0.72 -0.05	PPA -0.96 -2.93 2.16 -5.03	PS 0.06 0.05 0.04	SA -52.53 -54.65 -34.91 -74.07	SPP -0.3 -0.28 -0.49 -0.25	<b>TKW</b> 0.91 1.14 0.26 1.48	98 0.12	<b>z z</b> 30	0_calc -2.22	<b>Z55</b> -0.37 -0.35 -0.42 -0.43

Fiche de synthèse des effets observés au sein du couple de NIL 5039\_Q5. Les effets sont exprimés dans le sens donneur vs receveur, les écarts statistiquement significativement différents (seuil de 5%) sont représentés en couleur (effets positifs en rouge et négatifs en vert). A: nom du couple isogénique et effets attendus au QTL, B: radar plot des effets exprimés en pourcents. C: effets observés pour les 7 lieux pour les modalités D-,D,D+ et toute\_D (toutes modalités).D: effets pour tous les lieux confondus pour les modalités D-,D, D+ et Tout (toute modalité).

Les résultats des essais des 7 couples NIL retenus en 2017 ont été synthétisés dans les deux tableaux ci-dessous. Ils présentent l'analyse toute modalité confondue lieu et année, pour les essais « azote » et pour les essais « densité ».

D'une façon générale, nous observons très nettement que les allèles positifs des 4 couples isogéniques évalués sur les essais «azote» induisent une augmentation significative des valeurs GPC et GPD sans altérer le rendement. De la même façon, 3 sur 4 des couples isogéniques évalués sur les essais «densité», démontrent une baisse significative du nombre d'épis par m<sup>2</sup> ainsi qu'une augmentation du poids de mille grains sans altération du rendement et de la teneur en protéines du grain. De façon plus approfondie, en exploitant tous les résultats des essais réalisés entre 2015 et 2017, nous pouvons caractériser les effets des 6 QTL de tallage et/ou d'efficience d'utilisation de l'azote étudiés en 2017.

#### Le QTL5

Ce QTL, situé sur le chromosome 1A, a été étudié via l'utilisation de deux couples de lignées isogéniques évalués

sur l'essai « densité » : le couple NIL BC 5039\_Q5 (effet de l'allèle Autan dans le fond Apache) dont les effets ont déjà été décrits ci-dessus et le couple NIL HIF APUL\_Q5 (effet de l'allèle Uli3 versus l'allèle Apache). Ces deux couples démontrent un effet significatif du QTL5 sur le tallage avec un effet des allèles Autan et Uli3 qui induisent une réduction de -8% à -10% du nombre d'épis par m<sup>2</sup> à maturité par rapport à l'allèle Apache, toutes modalités densité de semis confondues. Dans les deux couples, nous pouvons constater que plus la densité de semis est forte, plus l'effet sur le tallage est conséquent (de -13 à -15 % pour la modalité D+). Cette réduction du nombre d'épis par m<sup>2</sup> induit une augmentation du poids de mille grains (TKW) d'environ +1 gramme, sans induire d'effet négatif significatif sur le rendement et la teneur en protéines du grain (GPC). Cependant, chez les deux couples isogéniques, le rendement et le GPC sont augmentés dans la modalité D+. On peut également noter que dans le couple HIF APUL\_Q5, l'allèle de Uli3 induit une augmentation de la hauteur des plantes de +2,7 cm, qui n'est pas observée pour l'allèle Autan.

Couple NIL	GL	GW	GPC	GPD	GY15	PH	SA	SPP	TKW	PS	Z55	dursen	INN_FLO
5197_Q60	-0.12	0.06	0.35	0.70	0.36	7.46	12.76	0.11	1.45	4.05	-0.56	NA	NA
APAU_Q76_3	0,00	0.00	0.33	0.23	0.19	-0.32	11.97	0.20	0.79	0.32	-0.57	-4.47	0.04
APAZ_Q81	0.00	0.09	0.27	0.42	0.44	-0.44	10.40	-0.01	0.56	2.06	-1.64	NA	NA
AUAZ_Q36_3	0.05	0.04	0.35	0.26	0.03	-0.02	-8.41	0.02	1.19	2.19	-1.57	-32.72	-0.04

Synthèse des effets observés au sein des familles isogéniques Metapop « azote » sur tous les lieux, toutes modalités azote confondues pour les variables : longueur et largeur du grain (GL, GW) uniquement évalués sur un lieu (Clermont-Ferrand), concentration en protéines (GPC) et écart à la régression de la relation protéine/rendement (GPD), rendement à 15 % d'humidité (GY15), hauteur des plantes (PH), nombre d'épis à maturité (SA), nombre d'épis par plante (SPP), poids de mille grains (TKW), poids spécifique (PS), stade épi 1 cm (Z30), date d'épiaison (Z55), durée de sénescence (dursen), l'indice de nutrition azotée à floraison (INN\_flo). Les effets sont exprimés dans l'unité de mesure des variables, dans le sens de l'effet de l'allèle positif au QTL de la lignée donneuse versus l'allèle de la lignée receveuse. Les effets significativement différents (analyse Metafor, pvalue à 5%) sont en gras, les effets positifs sont surlignés en rouge et les effets négatifs en vert.

Couple NIL	GL		GW		GPC	GP	PD	GY15	1	PH		SA	SPP		TKW	1	PS		Z30		Z55	
5039_Q5		-0.04		0.06	-0.02		-0.05		0.48		0.30	-52.53		-0.30		0.91		0.06		-2 22		-0.38
APAU_Q76		0.02	(	0.01	0.33		0.32		0.22		-0.22	-33.66		-0.18		0.59		0.33		-6 74		1.20
APAZ_Q28		0.11		0.01	0.19		0.18		-0.98		0.56	-5.06	8	0.00		1.10		-0.18	NA			-0.11
APUL Q5		0.07	(	0.03	0.09		0.11		0.92		2.75	-37.03		-0.24		1.41		0.50		2 53		-0.01

Synthèse des effets observés au sein des familles isogéniques Metapop « densité » sur tous les lieux, toutes modalités densité de semis confondues.

#### Le QTL28

Ce QTL, situé sur le chromosome 3B, a été étudié sur l'essai «densité» par le couple NIL HIF APAZ\_Q28 qui permet de mesurer l'effet de l'allèle Aztec (réduction du tallage et augmentation du TKW) versus l'allèle Apache. Ce couple isogénique n'a pas démontré d'effet significatif sur le tallage pour ce QTL. Toutefois, nous démontrons un effet significatif de +3 % sur le TKW pour les modalités D et D+ associé à une augmentation de la teneur en protéines du grain de +1,5 %. L'augmentation du TKW est accompagnée d'une augmentation de la longueur et largeur du grain, cependant nous observons une réduction du rendement d'environ 1 q / ha.

#### Le QTL76

Ce QTL, situé sur le chromosome 7A, a été étudié à la fois sur les essais «azote» et «densité» via deux couples isogéniques HIF APAU\_Q76 et APAU\_Q76\_3 permettant tous les deux de quantifier l'effet de l'allèle Autan versus l'allèle Apache. Les deux constructions démontrent clairement un effet positif de l'allèle Autan sur le GPC de +3 % et sur le GPD. Ce QTL a un effet positif sur le TKW qui n'est pas corrélé négativement au rendement. Sur la modalité N-, le QTL a un effet moyen sur le rendement de + 1,5 q / ha mais qui n'est pas statistiquement significatif. Il faudrait réaliser davantage d'essais sous conditions d'azote réduit pour confirmer cet effet.

#### Le QTL60

Ce QTL situé sur le chromosome 6B, a été évalué sur l'essai «azote» via l'utilisation du couple NIL BC 5197\_Q60 qui permet d'étudier l'effet de l'allèle Uli3 par rapport à l'allèle Cezanne. L'allèle Uli3 induit une augmentation moyenne de la taille des plantes de + 7,5 cm avec une augmentation du TKW de 1 à 1,7 g, via une augmentation de la taille du grain de + 0,11 mm (la largeur ne semble pas affectée). Pour la modalité azote optimal, l'augmentation du TKW n'induit pas de perte du rendement, de plus nous observons une augmentation des valeurs GPC et GPD. Pour la modalité azote réduit, nous pouvons observer une augmentation du rendement de +1,5 q / ha, accompagnée d'un indice GPD de + 1,06 points.

#### Le QTL36

Ce QTL est situé sur le chromosome 4A. Il a été évalué sur l'essai «Azote» via l'utilisation du couple NIL HIF AUAZ\_Q36\_3 qui permet d'étudier l'effet de l'allèle Autan versus Aztec. L'allèle Autan est significativement plus précoce de -1,5 jour et induit une amélioration du TKW et des valeurs GPC et GPD. Cependant, le QTL semble induire une réduction du rendement pour les modalités N et N+ et une amélioration de +4,4 q / ha pour la modalité N-. Il est important de noter que sur les essais de 2016 (année avec forte pression maladie), l'allèle Autan s'est comporté différemment avec une augmentation significative du rendement qui a impacté l'effet positif sur le GPC observé en 2015 et 2017.

#### Le QTL81

Ce QTL situé sur le chromosome 7D a été évalué sur l'essai «azote» par le couple NIL HIF APAZ\_Q81 qui permet d'étudier l'effet de l'allèle Apache versus l'allèle Aztec. Nous observons un effet de l'allèle Apache sur la précocité de -1,6 jours, ainsi que sur le poids spécifique et sur le TKW qui est accompagné d'un effet positif sur le GPC et GPD. Cependant, ces effets s'accompagnent d'une perte de rendement qui est non significative sur l'analyse multi année-lieu-modalité mais qui présente de fortes variations positives et négatives selon les lieux et années testés (fortes interactions génotype x environnement), notamment sur les essais de 2016. Dans ces six QTL, étudiés au travers de 8 couples de lignées isogéniques, nous avons pu confirmer les effets attendus et caractériser plus finement l'impact de ces locus sur les différentes variables mesurées. Les QTL retenus à l'issue des trois années d'essais démontrent clairement des effets positifs sur le rendement, les composantes du rendement (notamment le tallage et le poids de mille grains) ainsi que des effets sur l'efficience d'utilisation de l'azote avec des amélioration de la teneur en protéine du grain et de l'indice GPD. Des essais complémentaires ou des réductions des intervalles de confiance pourront s'avérer nécessaires pour certains QTL, en fonction de leur intérêt pour la sélection. Nous avons noté des effets de compensation des variables et des interactions avec l'environnement sur lesquels des analyses complémentaires pourront être effectuées à partir du jeu de données généré.

#### 3.2 - Caractérisation génétique des lignées NIL

Une caractérisation génétique des lignées isogéniques a été menée afin de contrôler la qualité du matériel construit et de faciliter son utilisation en sélection. Nous avons génotypé 185 lignées NIL ainsi que leurs lignées parentales (47 NIL GPC Arche x Recital, 40 NIL *NAM B1*, 82 NIL Metapop, 16 NIL *TIN1A*). Sur les 9802 marqueurs génotypés par la technique GBC, 9278 SNP (94.7%) se sont avérés polymorphes sur ce matériel. Le nombre de données manquantes est faible avec moins de 1.8% d'échecs. Afin de comparer les couples NIL, les marqueurs ont été ancrés sur le génome. 82.9% des marqueurs (7642 SNP) ont pu être positionnés physiquement sur le référentiel IWGSC RefSeq V1.0 et 72.4 % (6674 SNP) ancrés génétiquement sur la carte consensus INRA, développée dans le projet Breedwheat (Rimbert *et al.*, 2018).

L'analyse des données de génotypage a permis de mesurer la taille des fragments d'intérêt introgressés par rapport à la taille des QTL ciblés. Cette information permet de contrôler l'entrainement de locus portant des allèles négatifs ou inadaptés qui pourraient diminuer, voire annuler l'effet des allèles positifs importés dans les lignées élites. La taille moyenne des fragments introgressés était comprise entre 100% et 150% de la taille de la zone ciblée. Enfin, nous avons évalué la pureté du matériel isogénique en calculant l'indice de similarité des couples NIL BC ou NIL HIF sur la base des marqueurs polymorphes entre les deux lignées parentales des couples. Cette analyse a démontré des niveaux de similarité élevés entre les couples NIL GPC Arche x Recital, NAM B1 et NIL Metapop avec des variations d'indice compris entre 49.0 % et 99.4 % et une moyenne située à 97.5%. Quatre couples NIL Metapop inférieurs à 85.0% d'identité ont été exclus pour la suite des évaluations. Les lignées NIL TIN1A se sont avérées moins bien fixées avec un indice moyen à 55.0% d'identité. Les 3 couples les mieux fixés ont été retenus pour les évaluations agronomiques.

Afin d'identifier plus précisément les chromosomes portant des résidus provenant des lignées donneuses, nous avons réalisé des visualisations graphiques présentant l'homologie des couples NIL (cf. figure ci-contre).

Ce travail a permis, d'orienter au mieux le choix des lignées en vue de leur caractérisation dans le volet 2. A terme, l'identification des nouveaux marqueurs dans les zones d'intérêt pourra être utilisée pour le suivi ou le clonage des QTL. Enfin, les couples isogéniques pourront être retravaillés par les sélectionneurs qui disposeront d'informations et de marqueurs pour suivre les zones présentant des défauts de fixation.

Nous avons initié un travail de réduction des intervalles de confiance de 6 QTL issus de la Metapop, priorisés sur la base



Représentation graphique de l'homologie d'un couple isogénique (exemple d'un NIL sur le QTL 7A Arche x Recital). En haut, la lignée portant les allèles négatifs (Arche) au QTL et en bas la lignée portant les allèles positifs (Recital). En ordonnée, les 21 chromosomes et en abscisse les marqueurs SNP (barres colorées : rouge=allèle Arche ; bleu= allèle Recital). Le QTL est matérialisé par les deux barres verticales rouges sur le chromosome 7A. On peut observer un fragment résiduel de la lignée donneuse Arche sur le chromosome 5B.

des résultats de caractérisation agronomiques des NIL. En effet, bien que ces QTL étaient initialement décrits avec des positions génétiques de quelques dizaines de centimorgans, leur ancrage physique a mis en évidence des zones étendues allant de 24 à 511 Mb. Il est donc apparu pertinent de génotyper de nouveaux marqueurs dans ces QTL en vue d'une nouvelle détection de QTL sur la population d'origine et tenter de réduire les intervalles de confiance. Afin d'atteindre cet objectif, nous avons sélectionné 168 nouveaux marqueurs SNP en fonction de leur polymorphisme au sein des parents de la population Metapop et de leur répartition homogène au niveau des zones ciblées (tableau ci-dessous).

QTL	Chrom.	Taille (pb)	SNP sélectionnés
QTL_5	chr1A	391100000	59
QTL_28	chr3B	24	
QTL_36	chr4A	24649000	16
QTL_60	chr6B	511668000	25
QTL_76	chr7A	49595000	19
QTL_81	chr7D	42451000	25
		Total :	168

Tableau des 6 QTL sélectionnés pour la densification en marqueurs. De gauche à droite : nom du QTL, chromosome, taille de l'intervalle en paires de bases sur le référentiel Chinese Spring IWGSC RefSeq V0.1, nombre de SNP sélectionnés.

Ces SNP ont été génotypés sur l'ensemble des 551 individus de la population Metapop. La figure ci-contre présente un exemple de résultat de densification en marqueurs sur le QTL81 situé sur le chromosome 7D. Nous constatons que les marqueurs développés comblent les zones peu couvertes de la carte initiale (en rouge).



Représentation graphique de la carte physique du QTL81 avant densification en marqueurs (barre rouge) et après (barre bleue). Les liens entre les marqueurs communs des deux cartes sont matérialisés par des traits bleus.

33

La nouvelle détection de QTL avec la carte enrichie en marqueurs n'a pas été effectuée, mais elle devrait permettre à terme de réduire l'intervalle de confiance de ces 6 QTL. Ce travail peut s'avérer important en vue de potentielles étapes de clonage des QTL ou leur utilisation par les sélectionneurs, en limitant les pollutions alléliques engendrées par l'introgression de zones QTL trop étendues.

# 4. Conclusions - Perspectives

La collaboration entre Biogemma, Limagrain, INRA et ARVALIS nous a permis d'explorer un large catalogue de QTL et d'accéder à un vaste réseau de plateformes expérimentales. Le projet est également un bel exemple de continuité entre des projets ANR Génoplante, à l'origine de la détection de certains QTL ainsi que de la construction de plusieurs couples NIL. Cette synergie s'illustre également au travers de l'utilisation de ressources produits par le projet d'Investissement d'Avenir ANR BreedWheat (ANR-10-BTBR-03).

D'un point de vue scientifique, bien que des analyses ou expérimentations complémentaires pourraient s'avérer nécessaires dans certains cas, nous sommes parvenus à démontrer des effets significatifs pour plusieurs QTL qui confirment les effets attendus par les loci ou allèles positifs ciblés. Ces objectifs ont pu être atteints grâce à l'utilisation de lignées isogéniques, couplée à la mise en place d'un réseau d'essais étendu permettant d'étudier différents environnements au travers de combinaisons lieux, années, et modalités de culture. En parallèle, le projet a permis de mettre en place des échanges entre les partenaires sur les dispositifs expérimentaux et les méthodes d'analyse de ce type de matériel.

Le projet a permis de réunir un jeu de données conséquent sur lequel de nouvelles analyses pourront être envisagées pour étudier plus finement les interactions Allèle x Environnement ou encore les phénomènes de compensation entre les variables en utilisant par exemple des modèles de type *path analysis* (Dhungana *et al.* 2007).

Ce projet permet de fournir un catalogue de QTL et de lignées isogéniques susceptibles d'être rapidement exploités dans des programmes de sélection variétale. A titre d'exemple, des lignées isogéniques pour le gène *NAM-B1* ont été choisies et introduites dans le programme de croisements de l'INRA. Les descendants sont en cours d'étude. Le travail de caractérisation génétique des lignées NIL ainsi que la densification en marqueurs de plusieurs QTL, devrait faciliter la valorisation du matériel. Enfin, une publication est envisagée sur l'effet du gène *NAM-B1* sur la concentration en protéines et le rendement en utilisant les données du réseau d'expérimentation du projet.
## Références bibliographiques

Bogard M, Allard V, Brancourt-Hulmel M, Heumez, E., Machet JM, Jeuffroy MH, Gate P, Martre P, and Le Gouis J (2010) Deviation from the grain protein concentration - grain yield negative relationship is highly correlated to post-anthesis N uptake in winter wheat. Journal of Experimental Botany 61, 4303-4312

Bogard M, Jourdan M, Allard V, Martre P, Perretant MR, Ravel C, Heumez E, Orford S, Snape J, Griffiths S, Gaju O, Foulkes J, and Le Gouis J (2011) Anthesis date mainly explained correlations between post-anthesis leaf senescence, grain yield and grain protein concentration in a winter wheat population segregating for flowering time QTL. Journal of Experimental Botany 62, 3621-3636

Cormier F, Faure S, Dubreuil P, Heumez E, Beauchêne K, Lafarge S, Praud S, and Le Gouis J (2013) A multienvironmental study of recent breeding progress on nitrogen use efficiency in wheat (Triticum aestivum L.). Theoretical and Applied Genetics 126, 3035-3048

Cormier F, Foulkes J, Hirel B, Gouache D, Moënne-Locco Y, Le Gouis J (2016) Breeding for increased nitrogen-use efficiency: a review for wheat (T. aestivum L.). Plant Breeding 135:255-278

DePauw RM, Knox RE, Clarke FR, Wang H, Fernandez MR, Clarke JM, McCaig TN (2007) Shifting undesirable correlations. Euphytica 157:409-415

Dhungana P, Eskridge KM, Baenziger PS, et al (2007) Analysis of Genotype-by-Environment Interaction in Wheat Using a Structural Equation Model and Chromosome Substitution Lines. Crop Science 47:477

**Fossati D, Fossati A, Feil B** (1993) Relationship between grain yield and grain nitrogen concentration in winter triticale. Euphytica 71:115-123

Gaju O, Allard V, Martre P, Snape J, Heumez E, Le Gouis J, Moreau D, Bogard M, Griffiths S, Orford S, Hubbart S, and Foulkes J (2011) Identification of traits to improve the nitrogen-use efficiency (NUE) of wheat genotypes. Field Crop. Res. 123, 139-152

Jonnala RS, MacRitchie F, Smail VW, et al (2010) Protein and Quality Characterization of Complete and Partial Near-Isogenic Lines of Waxy Wheat. Cereal Chemistry Journal 87:538-545

Joppa LR, Du CH, Hart GE, Hareland GA (1997) Mapping gene(s) for grain protein in tetraploid wheat (Triticum turgidum L.) using a population of recombinant inbred chromosome lines. Crop Sci 37:1586-1589

Lacerenza JA, Parrott DL, Fischer AM (2010) A major grain protein content locus on barley (Hordeum vulgare L.) chromosome 6 influences flowering time and sequential leaf senescence. Journal of Experimental Botany 61:3137-3149 Laperche A, Brancourt-Hulmel M, Heumez E, Gardet O, Hanocq E, Devienne-Barret F, Le Gouis J (2007) Using genotype x nitrogen interaction variables to evaluate the QTL involved in wheat tolerance to nitrogen constraints. Theoretical and Applied Genetics 115:399-415

Laperche A, Devienne-Barret F, Maury O, Le Gouis J, Ney B (2006) A simplified conceptual model of carbon and nitrogen functioning for QTL analysis of winter wheat adaptation to nitrogen deficiency. Theoretical and Applied Genetics 113:1131-1146

Le Gouis J, Béghin D, Heumez E, Pluchard P (2000) Genetic differences for nitrogen uptake and nitrogen utilisation efficiencies in winter wheat. European Journal of Agronomy 12:163-173

Miralles DJ, Slafer GA, Lynch V (1997) Rooting patterns in nearisogenic lines of spring wheat for dwarfism. Plant and Soil 197:79-86

**Mitchell JH, Rebetzke GJ, Chapman SC, Fukai S** (2013) Evaluation of reduced-tillering (*tin*) wheat lines in managed, terminal water deficit environments. J Exp Bot 64:3439-3451

Monaghan JM, Snape JW, Chojecki AJS, Kettlewell PS (2001) The use of grain protein deviation for identifying wheat cultivars with high protein concentration and yield. Euphytica 122:309-317

**Oury FX, Godin C** (2007) Yield and protein concentration in bread wheat: how to use the negative relationship between the two characters to identify favourable genotypes. Euphytica 157:45-57

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

**Spielmeyer W, Richards RA** (2004) Comparative mapping of wheat chromosome 1AS which contains the tiller inhibition gene (*tin*) with rice chromosome 5S. Theor Appl Genet 109: 1303-1310

Uauy C, Brevis JC, Chen XM, Khan I, Jackson L, Chicaiza O, Distelfeld A, Fahima T, Dubcovsky J (2005) High-temperature adult-plant (HTAP) stripe rust resistance gene Yr36 from Triticum turgidum ssp dicoccoides is closely linked to the grain protein content locus Gpc-B1. Theoretical and Applied Genetics 112:97-105

Van Sanford DA, MacKown CT (1986) Variation in nitrogen use efficiency among soft red winter wheat genotypes. Theor Appl Genet 72:158-163

Wang X, Zhang Y, Zhang B, *et al.* (2018) Comparison of quality properties between high-molecular-weight glutenin subunits 5 + 10 and 2 + 12 near-isogenic lines under three common wheat genetic backgrounds. Cereal Chemistry 95:575-583

Waters BM, Uauy C, Dubcovsky J, Grusak MA (2009) Wheat (Triticum aestivum) NAM proteins regulate the translocation of iron, zinc, and nitrogen compounds from vegetative tissues to grain. J Exp Bot 60:4263-4274

# [N-BTH] Méthodes d'estimation des indicateurs d'efficacité de valorisation de l'azote par les nouvelles variétés de blé tendre

Jean-Pierre COHAN<sup>\*1</sup>, François-Xavier OURY<sup>2</sup>, Aurélie MAILLIARD<sup>3</sup>, Josiane LORGEOU<sup>\*4</sup>, Marie-Hélène BERNICOT<sup>3</sup>, Christine LE SOUDER<sup>4</sup>, Adeline STREIFF<sup>4</sup>, Sonia GEOFFROY<sup>4</sup>, Philippe LEREBOUR<sup>5</sup>, Thierry MOITTIE<sup>5</sup>, Patrice SENELLART<sup>5</sup>, Olivier DRUELLE<sup>6</sup>, Corentin BONNARD<sup>6</sup>, François GUION<sup>7</sup>, Jacques LE GOUIS<sup>2</sup>

1 - ARVALIS-Institut du végétal - Station expérimentale de La Jaillière - La Chapelle St Sauveur - 44370 Loireauxence

- 2 INRA-UCA UMR GDEC 5, chemin de Beaulieu 63000 Clermont-Ferrand
- 3 GEVES Domaine de l'Anjouère 49370 La Poueze
- 4 ARVALIS-Institut du végétal Station expérimentale 91720 Boigneville
- 5 UFS 17, rue du Louvre 75001 Paris
- 6 Semences de France Ferme du Corbeau 91490 Milly la Forêt
- 7 ANMF 66, rue de la Boétie 75008 Paris
- \* Coordinateur : Jean-Pierre COHAN (jp.cohan@arvalis.fr), Josiane LORGEOU (j.lorgeou@arvalis.fr)

#### 1. Introduction

Notre projet a été motivé par l'attente de l'agriculture française et européenne de pouvoir répondre aux besoins d'augmentation et d'optimisation de la production des céréales tout en maintenant la qualité des produits, en limitant les coûts de production et en réduisant les effets négatifs sur l'environnement. L'azote est un élément nutritif essentiel qui influence fortement le rendement en grains et la teneur en protéines. Cependant, La production, le transport et l'application des engrais azotés sont très consommateurs d'énergie et le prix de ces engrais est indexé au prix du pétrole. Ainsi le prix des engrais simples azoté a très fortement varié dans les dix dernières années (INSEE). Cela se traduit pour l'agriculteur par un coût équivalent à environ 25-30% des charges opérationnelles. Un des moyens de diminuer ces charges est d'améliorer l'efficacité d'utilisation par la culture et notamment l'efficacité d'absorption. En effet seuls environ 60% de l'azote disponible dans le sol seront absorbés par le blé (Gaju et al., 2011 ; Cormier et al. 2016). Une partie est perdue par lessivage entrainant une pollution des eaux de surface et des nappes phréatiques. Une autre partie est volatilisée sous forme d'oxyde nitreux, un gaz à effet de serre supérieur à celui du gaz carbonique. Ainsi, les engrais azotés seraient responsables tout au long de leur cycle d'environ 50% des émissions de gaz à effet de serre d'origine agricole. Pour limiter ces impacts négatifs, il peut être envisagé une diminution des apports azotés. Mais cela doit se faire tout en maintenant le rendement et la teneur en protéines. Il semble donc essentiel de favoriser la sélection de variétés de blé plus efficace dans leur utilisation de l'azote pour produire du rendement et des protéines.

Comme pour la plupart des cultures, l'azote est un facteur de production essentiel pour le blé tendre d'hiver (Jensen et al. 2011). Le maintien d'un état de nutrition azotée non limitant est donc un enjeu majeur pour assurer une collecte de blé répondant aux objectifs de production quantitative et qualitative pour les débouchés intérieurs et export, particulièrement exigeants en matière de teneur en protéines du grain (Méléard 2007, ARVALIS 2013). En parallèle, les pertes d'azote réactif dans différents compartiments environnementaux (atmosphère, aquifère...) sont dans certains cas à l'origine d'impacts négatifs sur la qualité de l'air, la qualité de l'eau et le bilan gaz à effet de serre des filières de production agricoles (Galloway et al. 2003, Sutton et al. 2011). Enfin, l'instabilité des cours du prix des engrais azotés, avec néanmoins une tendance haussière sur le long terme, fait que le recours aux intrants azotés de synthèse pèse de plus en plus lourd dans le bilan économique des

exploitations françaises (Laurent et Leveau 2014). Pour toutes ces raisons, l'amélioration de l'efficience des apports d'azote reste une des priorités de l'optimisation des itinéraires culturaux du blé tendre d'hiver (Hawkesford 2014). Ajoutons enfin que les contraintes réglementaires (directe Nitrates, directive NEC, plans de protection de l'atmosphère) et/ou économiques ainsi que des choix volontaires d'orientation des systèmes de culture vont probablement générer de plus en plus de situations où le blé tendre d'hiver sera conduit en état de nutrition azotée sousoptimale. Il devient donc nécessaire de s'intéresser également au comportement des variétés dans ces situations de carence subies ou volontaires.

La réponse à ces enjeux fait nécessairement intervenir plusieurs leviers qui ont tous bénéficié d'importants travaux de recherches et développements depuis 30 ans : l'élaboration de la méthode du bilan prévisionnel (Rémy and Hébert, 1977; Rémy and Viaux, 1983), le développement d'outils de pilotage de la fertilisation azotée (Soenen et al. 2017), l'amélioration des pratiques d'apports des engrais (Bouthier 1997; Limaux et al., 1999; Cohan et Bouthier 2010) et l'adoption des formes d'engrais les plus efficaces (Chambers et Dampney 2009, Sylvester-Bradley et al. 2014). Associé à tous ces leviers agronomiques, l'obtention (par les généticiens et les sélectionneurs), l'évaluation (par les agronomes) puis l'adoption (par les agriculteurs) de variétés affichant la meilleure efficience de valorisation de l'azote possible, est un moyen d'action majeur (Good et al., 2004; Hirel et al. 2007; Foulkes et al., 2009, Cormier et al. 2016). Actionner ce levier nécessite néanmoins de disposer de traits à sélectionner clairement identifiés et d'une caractérisation fine sur le critère « efficience de l'azote apporté » des variétés mises sur le marché. En 2014, seuls deux indicateurs étaient utilisés en routine. La Grain Protein Deviation (GPD, Oury et Godin 2007) représente l'écart à la relation négative rendement-teneur en protéines attribuable à une variété. Elle est utilisée dans le processus d'inscription français piloté par le CTPS-GEVES depuis 2007 et est maintenant prise en compte dans la plupart des programmes de sélection. Elle n'a néanmoins pas d'utilisation pratique dans le cadre de la gestion de la fertilisation azotée par l'agriculteur. Le besoin unitaire en azote par unité de rendement (coefficient b, Le Souder et Bernicot 1993) est couramment utilisé dans les outils de gestion de la fertilisation azotée pour estimer le besoin global en azote de la culture (par multiplication avec l'objectif de rendement). Des valeurs par variété sont diffusées tous les ans par ARVALIS-Institut du végétal et relayées par les organismes agricoles économiques et de développement. Le coefficient b présente néanmoins deux inconvénients : ses bases scientifiques ont été définies dans les années 90 avec un paysage variétal différent du panorama actuel de la diversité génétique à disposition des agriculteurs et il ne prend en compte qu'un objectif de production quantitatif, sans considérer les objectifs spécifiques de teneur en protéines du grain imposés par les marchés. A contrario, la littérature scientifique propose de nombreux indicateurs susceptibles d'être utilisés. En particulier, les travaux en génétique se reposent beaucoup sur la notion de Nitrogen Use Efficiency (NUE, Moll et al., 1982) et de ses composantes la Nitrogen Uptake Efficiency (NupE) et la Nitrogen Utilization Efficiency (NutE) (Hirel et al., 2007; Sylvester-Bradley et Kindred 2009). Les travaux en agronomie, eux, ont le plus souvent recours à la notion d'Apparent Fertilizer Recovery (AFR, ou Coefficient Apparent d'Utilisation-CAU en français) (Foulkes et al., 1998; Limaux et al., 1999) et à la notion de besoin unitaire en azote (Le Souder et Bernicot 1993).

Afin de mettre au point de nouveaux indicateurs d'efficience de l'azote pour différents usages (sélection, préconisation et pilotage des variétés) et critères recherchés (rendement, teneur en protéines, propriétés de panification), le projet N-BT a été lancé en 2014 réunissant 6 partenaires : ARVALIS-Institut du végétal, INRA, GEVES, UFS, Semences de France et ANMF. Il a poursuivi trois objectifs :

- Elaborer de nouveaux indicateurs d'efficacité de valorisation de l'azote des nouvelles variétés de blé tendre d'hiver, y compris des indicateurs de tolérance à une carence azotée.
- Evaluer les conditions de leur mise en pratique en routine par le sélectionneur et l'évaluateur de variétés.
- Mettre à disposition des résultats de recherches permettant d'expliciter les différences d'efficience de l'azote entre variétés, à destination des sélectionneurs et du monde de la recherche en génétique et en agronomie.

Le présent article a pour objet de présenter la synthèse des travaux réalisés au cours de ce projet conduit d'octobre 2014 à mars 2018.

## 2. Matériel et méthode

Trois réseaux expérimentaux complémentaires ont été mis en place et valorisés : le nombre de variétés allant croissant à mesure que l'intensité analytique et le nombre de régimes azotés testés allaient décroissants. Ils ont chacun fait l'objet d'une procédure d'analyse de données adaptée à leurs caractéristiques (nombre de régimes de fertilisation azotée, variables mesurées).

#### Réseau expérimental n°1

Le réseau expérimental n°1 a regroupé des essais méthodologiques à haute intensité analytique, destinés à étudier de manière détaillée les interactions entre les variétés et les régimes de fertilisation azotée. Il a aussi été le support de plusieurs procédures de validation de méthodes élaborées sur les autres réseaux. Le protocole expérimental mis en place sur le réseau n°1 a consisté à évaluer 6 à 12 variétés de blé tendre d'hiver soumises à des doses croissantes d'apport d'engrais azoté minéral (dispositif dit en « courbe de réponse »). La courbe de réponse était constituée de 6 doses totales d'engrais apportées en 3 apports (stades tallage, début montaison, fin montaison) sous forme d'ammonitrate. Ces 6 doses étaient réparties de la manière suivante : témoin sans apport, X-80, X-40, X, X+40 et X+80 kg N.ha<sup>-1</sup>, X représentant la dose totale prévisionnelle d'engrais azoté calculée à la sortie de l'hiver. Pour les plus petites doses totales, un fractionnement en 2 apports a pu être adopté dans certains essais en privilégiant les apports tardifs pour éviter d'induire un biais dans la contribution des apports à la teneur en protéines du grain. Les essais ont tous été réalisés en microparcelles par plusieurs stations expérimentales d'ARVALIS-Institut du végétal selon un plan expérimental en criss-cross ou en splitplot. 6 essais ont été implantés sur la campagne 2013-2014, 6 essais en 2014-2015 et 4 essais en 2015-2016. Un essai de la campagne 2013-2014 n'a pas été conservé pour l'analyse en raison d'importants dégâts de grêle survenus en fin de cycle. De même, un essai de la campagne 2014-2015 a été retiré en raison d'une trop grande hétérogénéité de sol affectant la précision des résultats. Enfin, sur la campagne 2015-2016, un essai n'a pas été mené à son terme en raison de problème d'implantation, et un essai a été retiré de l'analyse en raison d'un comportement non interprétable dans le contexte climatique atypique de la récolte 2016. Au final, 12 essais sur les 16 implantés ont été valorisés dans le cadre du projet, représentant une diversité de sol et de climat correspondant aux principaux bassins de production du blé tendre en France. Les essais ont tous fait l'objet d'une caractérisation détaillée du sol (analyse physico-chimique complète sur 3 horizons 0-30 cm/30-60 cm/60-90 cm), ainsi que de l'enregistrement quotidien des données agro-météorologiques de base (températures, pluie, rayonnement global, ETP...). Afin d'étudier finement les interactions de la variété avec le régime de fertilisation azotée sur de multiples variables et indicateurs agronomiques et physiologiques, de nombreuses mesures ont été réalisées : production et teneur en azote des grains et des pailles, teneur en protéines du grain, stock d'azote minéral du sol à la récolte (reliquat post-récolte), indices de qualité de panification. La quantité d'azote prélevée par la plante entière (grains+pailles+racines) a été estimée en multipliant la quantité d'azote présente dans les grains et les pailles par 1.25.

Les données élémentaires (variables mesurées et calculées sur chaque variété à chaque dose d'azote) ont fait l'objet d'une analyse de variance en modèle mixte permettant de tester l'effet des facteurs variété, azote et essai de manière directe et en interaction entre eux. Le coefficient apparent d'utilisation de l'azote (CAU) de chaque variété dans chaque essai tel que défini dans Limaux et al. 1999, ainsi que le NUE et ses composantes telles que définies dans Hirel et al. (2007) et Sylvester-Bradley et Kindred (2009), ont été calculés sous chaque régime de fertilisation et étudiés dans l'analyse de variance citée ci-dessus. Ensuite, la réponse des différentes variables aux doses croissantes d'engrais azotés (dispositif en courbe de réponse) a été analysée en ajustant successivement plusieurs modèles statistiques. Nous avons ainsi pu déterminer pour chaque variété le rendement optimal et la dose d'azote nécessaire pour l'atteindre, la teneur en protéines obtenue à l'optimum et la quantité d'azote absorbé nécessaire pour l'atteindre. Ces variables de base ont servi à calculer le besoin d'azote par unité de rendement (coefficient « b ») tel que défini par Le Souder et Bernicot (1993). En prenant comme référence un objectif de teneur en protéines de 11.5%, les différents modèles de courbes de réponses ont permis de déterminer l'éventuel besoin unitaire complémentaire pour atteindre cet objectif (coefficient « bc »), et donc le besoin unitaire total pour combiner optimum de rendement et 11.5% de teneur en protéines du grain (coefficient « bq »). Une évaluation de l'impact de l'utilisation de ce coefficient sur le reliquat azoté post-récolte a aussi été réalisée. Enfin, nous avons calculé le coefficient apparent d'utilisation de l'azote (CAU) de chaque variété conduite à son optimum dans chaque essai tel que défini dans Foulkes et al. (1998) et Cohan et al. (2018), ainsi que la NUE à l'optimum et ses composantes telles que définies dans Hirel et al. (2007) et Sylvester-Bradley and Kindred (2009). Les liens entre les différents indicateurs d'efficience de l'azote sont explicités dans la figure 1. L'ensemble des variables calculées sur la base des courbes de réponses a aussi fait l'objet d'une analyse de variance en modèle mixte pour déterminer l'impact de la variété et de l'essai sur leurs variations.



Figure 1 : Liens entre les principaux indicateurs d'efficience d'absorption et de valorisation de l'azote calculés sur le réseau 1. Adapté de Cohan et al. 2017.

#### Réseau expérimental n°2

Le réseau 2 représente le cœur du dispositif expérimental utilisé pour caractériser les variétés vis-à-vis de leur tolérance à une diminution de la fertilisation azotée et vis-à-vis de leur capacité à valoriser l'azote mis à disposition par le sol ou sous forme d'engrais. Il reprend les résultats des sites d'inscription du protocole Evaluation des Variétés Economes en Azote du CTPS. Ces essais CTPS représentent six jeux de données : deux zones (Nord et Sud) x trois couples d'années (2013-2014, 2014-2015 et 2015-2016). Ces essais ont tous été conduits en microparcelles selon différents dispositifs statistiques (split-plot, crisscross, blocs adjacents) et sont décrits de manière détaillée dans un document du GEVES (CTPS/GEVES 2017). En résumé, le protocole consiste à tester les listes variétales soumises au processus d'inscription à 3 régimes de fertilisation azotée : X-80, X et X+40 kg N.ha<sup>-1</sup>, X représentant la dose totale prévisionnelle d'engrais azoté calculée à la sortie de l'hiver. L'engrais azoté a été apporté en 4 apports (stades tallage, début et mi montaison, montaison), les différences de régimes portant fin essentiellement sur les apports à début montaison. Ils ont été complétés par une expérimentation spécifique du projet N-BT. conduite selon le même protocole sur 12 sites de postinscription comparant un tronc commun de variétés inscrites en octobre 2014 et bénéficiant de résultats antérieurs du protocole spécifique du CTPS. Plusieurs aléas climatiques et d'expérimentation ont conduit à ne pouvoir valoriser que huit sites (quatre en 2015, et quatre en 2016). Ces huit sites représentent une diversité d'environnements représentative de la zone de culture du groupe des variétés de blé tendre de précocité intermédiaire. Tous ces essais ont été conduits par le GEVES, l'INRA et les semenciers membres du partenaire UFS. Le rendement, la teneur en protéines et le poids de 1000 grains (PMG) ont été mesurés pour toutes les variétés présentes dans les essais du réseau 2. Une caractérisation du sol et une mesure de la quantité d'azote absorbé dans les pailles et le grain sur une variété témoin ont aussi été réalisées. Enfin, des mesures de qualité technologique du grain (panification, alvéogramme de Chopin, Gluten Humide et Gluten Index) ont été effectuées sur certaines modalités de certains essais.

Dans le cas spécifique de l'étude des indicateurs de tolérance à la carence azotée, nous avons fait appel à des jeux de données supplémentaires. Il s'agit de quatre jeux de données issus de deux projets FSOV 2003A et 2010F « azote » antérieurs (FSOV 2004-2005 et FSOV 2011-2012, ce dernier ayant donné lieu à deux expérimentations distinctes, l'une pour des génotypes précoces, notée FSOV-P, et l'autre pour des génotypes tardifs, notée FSOV-T), et du programme PIA BreedWheat ANR 2010-

BTBR-003 (essais conduits en 2012, 2013 et 2014). Ces jeux de données supplémentaires ont permis de gagner en puissance statistique pour l'étude des propriétés des indicateurs de tolérance à une carence en azote. Pour ces quatre dispositifs, le protocole mettait en œuvre deux niveaux d'apports d'engrais azoté: une dose X pour potentiellement atteindre l'objectif de rendement prévisionnel et une dose X-100 kg N.ha<sup>-1</sup> pour induire une carence azotée significative.

Les données élémentaires ont fait l'objet d'une analyse de variance en modèle mixte pour évaluer l'impact de la variété, de la dose d'azote, de l'essai (lieu et année) et de leurs interactions sur le rendement et la teneur en protéines du grain.

Les modalités X-80 et X ont été le support de l'étude de 2 indicateurs de tolérance à la carence azotée sur le rendement et la teneur en protéines du grain. Le premier correspond à la perte de rendement de la modalité X-80 par rapport à la modalité X, exprimée en pourcentage du rendement à la dose X, soit la formule (rdt=rendement) :

indicateur de tolérance pour le rendement = 
$$\frac{rdt_x - rdt_{x-80}}{rdt_x}$$

Le second correspond à l'écart à la régression (relation positive) entre le rendement de la modalité X et le rendement de la modalité X-80. Comme pour la GPD, sont utilisés les résidus standardisés de la régression pour identifier les génotypes qui, pour un niveau de rendement à la dose X donné, enregistrent une perte de rendement significativement plus faible (ou significativement plus forte) que les autres lorsqu'ils sont cultivés à la dose X-80.

Les 3 modalités X-80, X et X+40 ont été le support de l'étude de **2 indicateurs d'efficience de valorisation de l'azote** dits « simples » car ne recourant pas à des mesures d'azote absorbé plante entière (seuls les jeux de données CTPS et du projet ont été utilisés car les jeux de données provenant d'autres projets ne comportaient pas de modalité X+40). Ils sont basés sur la régression (calculée pour chaque génotype) entre la quantité d'azote exportée dans les grains (qN grains, calculée à partir du rendement et de la teneur en protéines du grain) et la dose d'azote apportée. On obtient ainsi une ordonnée à l'origine, qu'on peut considérer comme la capacité de la variété à valoriser l'azote du sol, et une pente qui peut être assimilée à sa capacité à valoriser l'apport d'engrais. Une 1<sup>ère</sup> approche méthodologique pour exploiter les 3 modalités sous la forme d'un dispositif en courbe de réponse a aussi été abordée.

#### Réseau expérimental n°3

Le réseau 3 a été constitué par des extraits de deux réseaux d'expérimentations préexistants au projet : le réseau postinscription ARVALIS (France entière) et le Variétoscope de Semences de France-In Vivo (liste Centre-Précocité intermédiaire). Au total, 115 essais ont été valorisés (39 provenant de Semences de France et 76 provenant d'ARVALIS-Institut du végétal), de 2010 à 2016. Le protocole a consisté dans la mise en essais de longues listes variétales soumises à un seul régime de fertilisation azotée (a priori à l'optimum), dans des dispositifs statistiques en bloc de micro-parcelles. Tous les essais ont bénéficié de mesures de rendement et de teneurs en protéines du grain. Les essais réalisés dans la période temporelle couverte par le projet ont aussi bénéficié de mesure complémentaire permettant de mieux caractériser les sites d'essais en terme de dynamique de l'azote (caractérisation du sol, reliquat sortie hiver, reliquat post-récolte).

L'analyse de données du réseau 3 s'est concentrée sur l'élaboration d'une méthode simplifiée de calculs des coefficients b, bc et bq sur la base de validation d'hypothèses simplificatrices à partir des données des réseaux 1 et 2.

## 3. Résultats et discussion

#### Etude détaillée des facteurs impactant l'efficience du blé tendre vis à vis de l'azote (réseau 1).

L'analyse des données obtenues sur le réseau n°1 indique clairement que la variété, la dose d'engrais azotée et leur interaction ont un effet significatif sur le rendement, la teneur en protéines, la quantité d'azote absorbé par le blé tendre d'hiver et le reliquat azoté post-récolte (Tableau 1). La variété présente aussi un effet significatif sur les ratios à la récolte, mais seul le NHI est impacté par la dose d'azote. Tous les indicateurs d'efficience de l'azote testés (efficience globale NUE, efficience d'absorption de l'azote NupE-CAU et efficience de valorisation de l'azote absorbé en grain NutE) sont significativement influencés par la variété et la dose d'azote et, à l'exception du CAU, par leur interaction.

L'analyse des données obtenues sur le réseau n°1 indique clairement que la variété, la dose d'engrais azotée et leur interaction ont un effet significatif sur le rendement, la teneur en protéines, la quantité d'azote absorbé par le blé tendre d'hiver et le reliquat azoté post-récolte (Tableau 1). La variété présente aussi un effet significatif sur les ratios à la récolte, mais seul le NHI est impacté par la dose d'azote. Tous les indicateurs d'efficience de l'azote testés (efficience globale NUE, efficience d'absorption de l'azote NupE-CAU et efficience de valorisation de l'azote absorbé en grain NutE) sont significativement influencés par la variété et la dose d'azote et, à l'exception du CAU, par leur interaction.

	Effets fixes			Effets aléatoires						
	V	N	$VAR \times N$	E×V	E×N	E				
	Variables de production									
RDT	***	***	***	***	***	***				
PROT	***	***	***	***	***	***				
QNT	***	***	***	***	***	***				
RPR	***	***	**	***	***	***				
HI	***	ns	ns	***	***	***				
NHI	***	***	ns	***	***	***				
		Indic	ateurs d'effi	cience N						
NUE	***	***	***	***	***	***				
NupE	***	***	***	***	***	***				
NutE	***	***	**	***	***	***				
CAU	*	*	ns	***	***	***				

\*\*\* (p <0.001) ; \*\* (0.001<p<0.01) ; \* (0.01<p<0.05) ; ns = non significatif

**Tableau 1**: Analyse de variance sur les données élémentaires du réseau 1. V = variété,

 N = dose d'engrais N, E = essai. RDT = rendement en grain, PROT = teneur en

 protéines du grain, QNT = quantité N totale absorbé à la récolte (grain+paille+racine),

 HI = indice de récolte (Harvest Index), NHI = indice de récolte azote (Nitrogen Harvest

 Index), CAU = coefficient apparent d'utilisation de l'azote, NUE = Nitrogen Use

 Efficiency, NupE = Nitrogen Uptake Efficiency, NutE = Nitrogen Utilization Efficiency.

L'analyse des résultats obtenus à l'optimum de fertilisation azotée calculé a posteriori sur les courbes de réponse à l'azote du réseau 1 est aussi riche d'enseignements (Tableau 2). La variété a un effet significatif sur le rendement, la teneur en protéines et la quantité d'azote totale absorbé à la récolte à l'optimum mais pas sur la dose d'azote optimale, ni sur le reliquat post-récolte correspondant. Quand on se donne le double objectif d'atteindre l'optimum de rendement et une teneur en protéines égale à 11.5 %, la variété présente un effet significatif sur la dose d'azote nécessaire pour assurer ce double objectif, mais pas sur la quantité d'azote totale absorbé et le reliquat post-récolte correspondant. Il est intéressant de noter que la variété ne

présente pas d'effet significatif sur la quantité d'azote absorbé pour la conduite « témoin sans apport d'engrais », ce qui reviendrait à dire qu'il n'y a pas de différence variétale sur la capacité à absorber l'azote fourni par le sol. Il faut néanmoins souligner que la précision des mesures sur ce caractère peut dans certains cas être faible et que nous ne pouvons donc pas écarter l'hypothèse de l'effet de cette imprécision sur l'absence de significativité statistique mise en évidence. Pour finir, la variété présente un effet significatif sur les ratios HI et NHI à la récolte et sur tous les indicateurs d'efficience de l'azote étudiés.

	V (effet fixe)	E (effet aléatoire)				
Variables de production						
RDTopt	***	***				
DNopt	ns	***				
PROTopt	***	***				
QNTopt	*	***				
RPRopt	ns	***				
QNT-0N	ns	***				
DNprot	**	***				
QNTprot	ns	***				
RPRprot	ns	***				
	Ratios à la récolte					
Hlopt	***	***				
NHlopt	***	***				
	Indicateurs d'efficience	e N				
NUEopt	***	***				
NupEopt	*	***				
NutEopt	***	***				
CAU	***	***				
Coefficient b	***	***				
Coefficient bc	***	***				
Coefficient bq	**	***				

\*\*\* (p <0.001) ; \*\* (0.001<p<0.01) ; \* (0.01<p<0.05) ; ns = non significatif

**Tableau 2 :** Analyse de variance sur les données du réseau 1 calculées à l'optimum de fertilisation azotée. V = variété, E = essai. RDT = rendement en grain, DN = dose d'engrais N, PROT = teneur en protéines du grain, QNT = quantité N totale absorbé à la récolte (grain+paille+racine), RPR = reliquat azoté post-récolte, HI = indice de récolte (Harvest Index), NHI = indice de récolte azote (Nitrogen Harvest Index), CAU = coefficient apparent d'utilisation de l'azote, NUE = Nitrogen Use Efficiency, NupE = Nitrogen Uptake Efficiency, NutE = Nitrogen Utilization Efficiency. L'indice « opt » se rapporte à un calcul à la dose N optimale pour le rendement. L'indice « prot » se rapporte à un calcul à la dose N optimale pour le rendement et pour atteindre une teneur en protéines du grain égale à 11.5%. Le suffixe « ON » se rapporte à la modalité sans apport d'engrais N.

Ainsi, à l'optimum de fertilisation azotée, les variétés testées dans le réseau 1 présentent des différences significatives quant à leur capacité à absorber l'azote de l'engrais apporté (Figure 2 pour une illustration avec le CAU), puis à valoriser cet azote absorbé en rendement et en teneur en protéines répondant à la demande des marchés (Figure 3 pour une illustration avec les coefficients b, bc et bq). Ces résultats sont concordants avec ceux décrits dans de nombreuses publications qui mettent en évidence l'effet de la variété sur la NUE et la NutE (Ortiz-Monasterio et al. 1997, Foulkes et al 1998, Le Gouis et al 2000, Brancourt Hulmel et al. 2003, Gooding et al. 2005, Sadras et Lawson 2013, Cormier et al. 2013). Concernant la NupE ou le CAU, l'effet variété est aussi souvent mis en évidence mais pas systématiquement (Cormier et al. 2013). La plupart de ces études montrent, comme la nôtre, une interaction entre la dose d'azote et la variété sur les indicateurs d'efficience de l'azote. En l'occurrence, étudier ces indicateurs à l'optimum spécifiquement calculé pour chaque situation permet de s'affranchir de ce phénomène pour mettre en évidence l'impact

de la variété seule. Notons aussi que l'effet de la variété sur le NHI est un phénomène connu dans la littérature (Ortiz-Monasterio *et al.* 1997, Le Gouis et al 2000, Cormier *et al.* 2013). Ce phénomène doit nous inciter à la prudence quand on souhaite utiliser des valeurs moyennes de NHI afin de calculer des indicateurs d'efficience dans des dispositifs expérimentaux ne disposant pas de mesure des quantités totales d'azote absorbé pour chacune des variétés testées (cf. analyses des réseaux 2 et 3). Enfin, il faut souligner que 1) la dernière étude de ce type réalisée sur des variétés disponibles en France date du début des années 90 (Le Souder et Bernicot 1993) et 2), que nous retrouvons globalement les mêmes résultats alors que le paysage variétal a largement évolué depuis.



Figure 2 : Coefficient Apparent d'Utilisation de l'azote (CAU) calculés sur le réseau 1. Moyennes ajustées par variété. Les barres verticales représentent les écart-types.



Figure 3 : Coefficients b, bc et bq calculés sur le réseau 1. Moyennes ajustées par variété. Les barres verticales représentent les écart-types calculés pour le coefficient bq (b+bc).

Le reliquat post-récolte est influencé par la dose d'engrais azoté (Tableau 1). La figure 4 illustre la nature de ce lien. Il apparaît que le reliquat post-récolte est relativement stable pour une dose d'engrais inférieure ou égale à la dose d'engrais azoté optimale. Ce résultat est cohérent avec les études déjà conduites sur ce sujet (Comifer 1997, Comifer 2013). L'accroissement du reliquat postrécolte devient vraiment significatif au-delà d'une limite aux alentours de 35 kg N.ha<sup>-1</sup> au-dessus de l'optimum. Pour un objectif de rendement de 90 q.ha<sup>-1</sup>, cela signifie qu'un besoin unitaire complémentaire (bc) pour atteindre une teneur en protéines égale à 11.5% ne doit pas dépasser 0.4 kg N.q<sup>-1</sup>.



Figure 4 : Relation entre le reliquat azoté post-récolte (RPR) et la dose d'engrais azotée (Dose N) élaborée à partir des données du réseau 1. Axes centrés sur la valeur calculée à l'optimum de fertilisation azotée. La ligne pointillée représente la limite au-delà de laquelle une dose d'azote complémentaire à la dose N optimale pour le rendement induit une progression significative du reliquat post-récolte.

#### Résultats élémentaires obtenus sur le réseau 2

La figure 5 représente un exemple de distributions du rendement et de la teneur en protéines sur un jeu de données représentatif du réseau 2. Le rendement baisse fortement lorsqu'on passe de la dose X à la dose X-80, mais le rendement change peu entre la dose X et X+40. En revanche, pour la teneur en protéines, l'évolution apparaît linéaire : la baisse est également forte entre la dose X et la dose X-80, mais la teneur en protéines continue à augmenter entre X et X+40. Ces résultats sont cohérents avec la plupart des autres études publiées sur le sujet (Foulkes *et al.* 1998, Sylvester-Bradley et Kindred 2009).



Figure 5 : Boxplots représentant les distributions des valeurs de rendement et de teneur en protéines pour les six essais validés de l'expérimentation CTPS zone Sud de 2013-2014 (réseau 2).

Sur tous les jeux de données disponibles, la réduction de la fertilisation azotée provoque une baisse des valeurs moyennes de rendement et de teneur en protéines. Cependant les analyses de variance réalisées montrent que, dans la plupart des cas, il y a une interaction « génotype \* conduite N » significative, ce qui indique que les différents génotypes ne répondent pas tous de la même façon et qu'il est donc pertinent de s'intéresser à des indicateurs variétaux liée à l'efficience de l'azote apporté au blé tendre. Ces résultats sont cohérents avec ceux obtenus sur le réseau 1.

#### Elaboration d'indicateurs de tolérance à la carence azotée (réseau 2)

Les premières analyses réalisées ont tout d'abord consisté à étudier le lien entre les 2 types d'indicateurs testés. De fait les deux types d'indicateurs de tolérance apparaissent très fortement corrélés : sur les 11 jeux de données, les R<sup>2</sup> des régressions entre les indicateurs de type 1 (écarts pondérés) et les indicateurs de type 2 (écarts à la régression) s'échelonnent entre 0.68 et 0.99, et la moyenne des ces R<sup>2</sup> est de 0.95 pour l'indicateur rendement et de 0.92 pour l'indicateur protéines. Comme les deux types d'indicateurs sont pratiquement redondants, dans la suite nous ne considèrerons plus que l'indicateur de type 1. La figure 6 présente un exemple des indicateurs de tolérance pour le rendement et la teneur en protéines, selon deux types d'agrégation : calculés lieu par lieu, puis moyennés à l'échelle du réseau d'essai, ou calculés directement à partir des moyennes multi-locales. Il ressort que ces deux méthodes donnent des résultats très fortement corrélés : sur les 11 jeux de données, les R<sup>2</sup> des régressions entre les deux types de valeurs s'échelonnent entre 0.96 et 1, et la moyenne des ces R<sup>2</sup> est de 0.99 pour l'indicateur rendement et de 1 pour l'indicateur protéines. Par souci de cohérence avec le mode de calcul de la GPD actuellement pratiqué dans le processus d'inscription, nous avons opté pour le mode de calcul utilisant les moyennes multi-locales.



Figure 6: Indicateurs de tolérance pour le rendement et la teneur en protéines calculés pour l'expérimentation CTPS zone Nord 2014-2015 (réseau 2), soit lieu par lieu pour les quatre essais validés (graphiques du haut), soit à partir des moyennes multi-locales (graphiques du bas)

Les analyses de variance réalisées sur les indicateurs de tolérance montrent qu'il y a un effet « génotype » significatif quel que soit le jeu de données dans le cas de l'indicateur « protéines », et significatif dans neuf cas sur 11 dans le cas de l'indicateur « rendement » (Tableau 3). Ceci indique qu'on peut effectivement discriminer des génotypes à l'aide de ces indicateurs de tolérance à une diminution de la fertilisation azotée.

	Α	V	L (A)	$A \times V$			
Indicateur de tolérance rendement							
CTPS Nord 13-14	***	***	***	ns			
CTPS Sud 13-14	***	*	***	ns			
FSOV 15-16	***	ns	***	ns			
CTPS Nord 14-15	**	**	***	*			
CTPS Sud 14-15	***	(*)	***	ns			
CTPS Nord 15-16	***	***	***	ns			
CTPS Sud 15-16	***	ns	***	ns			
FSOV 04-05	***	*	***	*			
FSOV-P 11-12	*	*	***	(*)			
FSOV-T 11-12	ns	**	***	ns			
BW 12-14	***	***	***	***			
Indic	ateur de tol	érance prote	éines				
CTPS Nord 13-14	***	***	***	ns			
CTPS Sud 13-14	***	***	***	ns			
FSOV 15-16	***	***	***	*			
CTPS Nord 14-15	ns	**	***	ns			
CTPS Sud 14-15	***	**	***	ns			
CTPS Nord 15-16	***	***	***	**			
CTPS Sud 15-16	***	*	***	ns			
FSOV 04-05	***	***	***	ns			
FS0V-P 11-12	***	***	***	*			
FSOV-T 11-12	***	***	***	ns			
BW 12-14	***	***	***	ns			

\*\*\*\* (p <0.001) ; \*\* (0.001<p<0.01) ; \* (0.01<p<0.05) ; (\*) (0.05<p<0.1) ; ns = non significatif

**Tableau 3 :** Analyse de variance sur les indicateurs de tolérance rendement et protéines calculés sur le réseau 2 pour les 11 jeux de données. A = effet année, V = effet variété, L(A) = effet lieu (année).

Une relation négative significative s'observe entre les indicateurs rendement et protéines sur les 11 jeux de données (R<sup>2</sup> de régression moyen de 0.30, figure 7 pour un exemple). Cette liaison négative signifie que, globalement, les lignées qui perdent le moins en rendement entre les doses X et X-80, sont celles qui subissent les pertes les plus importantes en teneur en protéines ; et qu'à l'inverse, celles qui perdent le moins en rendement, car c'est directement la conséquence de la forte relation négative qui existe entre rendement et teneur en protéines (Oury *et al.* 2003).

Les indicateurs de tolérance sont construits à partir de caractères (le rendement et la teneur en protéines) pour lesquels il y a de fortes interactions « génotype \* milieu », et ils sont donc estimés avec un niveau d'imprécision assez élevé. Ceci entache aussi l'estimation de la position de la droite de régression entre ces deux indicateurs (R<sup>2</sup> moyen est seulement de 0.3 sur l'ensemble des 11 jeux de données). De ce fait, il apparaît illusoire d'utiliser les écarts à la régression (comme on le fait par exemple pour les GPDs, mais dans un cas où le R<sup>2</sup> moyen des régressions « rendement -teneur en protéines » est plutôt de l'ordre de 0.5), pour rechercher des lignées réalisant un bon compromis entre leur perte en rendement et leur perte en teneur

en protéines lorsqu'on passe de la dose X à la dose X-80. Cependant, dans les situations où le manque d'azote durant la montaison est prévisible (choix d'itinéraire cultural voulu ou contraint par le milieu ou la réglementation), les indicateurs de tolérance peuvent constituer des critères importants pour le choix variétal. En effet, en fonction de l'objectif prioritaire, qui peut être soit la minimisation des pertes en rendement, soit la minimisation des pertes en teneur en protéines, ces indicateurs permettent de définir les variétés les plus favorables pour atteindre cet objectif. Concernant justement la GPD, il a été mis en évidence une bonne corrélation entre les valeurs de cet indicateur calculées à X-80 et celles calculées à X (R<sup>2</sup> de régression moyen de 0.62) et que, a contrario, aucune corrélation significative n'a été mise en évidence entre les indicateurs de tolérance (rendement et protéines) et la GPD, quelle que soit la conduite azotée. Il n'y a donc pas de redondance entre ces critères, qui apportent bien des informations différentes sur les variétés.



Figure 7 : Régression entre l'indicateur de tolérance pour le rendement et l'indicateur de tolérance pour la teneur en protéines dans le cas de l'expérimentation CTPS zone Sud 2014-2015 (réseau 2).

Pour qu'ils soient utiles dans la description des variétés, les indicateurs de tolérance ne doivent pas être trop fortement soumis aux interactions «génotype x milieu», et donc présenter une certaine stabilité malgré les fluctuations environnementales. Afin d'étudier ce point, nous avons comparé les valeurs des indicateurs de tolérance obtenues sur chacune des deux années consécutives des différents dispositifs (figure 8 pour un exemple). L'effet « année » apparaît très généralement significatif (10 cas sur 11 pour les différents jeux de données, qu'il s'agisse de l'indicateur «rendement» ou de l'indicateur «protéines») dans les analyses de variance (tableau 3). Cependant, le fait qu'il y ait un important effet « année » n'est pas gênant en soi : ce qui compte c'est le niveau de l'interaction « génotype x année » qui apparaît assez peu fréquemment significatif dans les analyses de variance (tableau 3), et que l'on peut estimer à travers la corrélation entre les valeurs des indicateurs obtenues pour chacune des deux années (figure 8). Les 11 jeux de données présentent la même tendance : le R<sup>2</sup> des régressions simples entre les valeurs des indicateurs « rendement » obtenues sur les deux années consécutives est en moyenne de 0.11, mais on passe à une valeur moyenne de 0.40 en régression robuste (méthode qui exclut les données à résidus extrêmes). L'indicateur « protéines » apparaît plus stable que l'indicateur «

rendement » puisque, en moyenne sur les 11 jeux de données, le R<sup>2</sup> en régression simple est de 0.24, et monte à 0.47 en régression robuste. Ces valeurs de R<sup>2</sup> sont du même ordre de grandeur que celles obtenues lorsqu'on étudie la stabilité interannuelle de la GPD. Or nous avons assez de recul dans l'utilisation de la GPD pour savoir qu'elle est un critère suffisamment stable pour être utile en sélection. Bien que calculés à partir de caractères assez fortement soumis aux interactions « génotype x milieu », les indicateurs de tolérance à une diminution de la fertilisation azotée semblent donc présenter un niveau de stabilité qui permet leur utilisation, que ce soit pour le rendement ou, avec encore moins de réserves, pour la teneur en protéines. Enfin, nous avons également regardé la stabilité des indicateurs de tolérance à une diminution de la fertilisation azotée en comparant les valeurs calculées pour deux dispositifs différents, lorsqu'il y avait des génotypes communs à ces deux dispositifs. Cette comparaison inter-dispositifs conduit à la même conclusion : les indicateurs de tolérance à une diminution de la fertilisation azotée présentent un niveau de stabilité qui, au moins pour l'indicateur «protéine», est tout à fait comparable à celui des GPDs, et qui semble suffisant pour qu'ils soient utiles pour la

caractérisation des variétés (même si l'indicateur « rendement » apparaît un peu moins stable).

Les valeurs prises par les indicateurs de tolérance peuvent être très variables d'un ieu de données à l'autre. Ceci est dû aux effets environnementaux (« année » et « lieu »), qui apparaissent très souvent hautement significatifs dans l'analyse de variance (Tableau 3). Il n'est donc pas possible d'utiliser directement des valeurs « seuils » pour discriminer les génotypes présentant des comportements contrastés pour leur tolérance à une diminution de la fertilisation azotée, puisqu'il faudrait faire varier ces seuils en fonction du jeu de données. Par contre, on peut s'inspirer de la méthode utilisée dans les essais CTPS pour attribuer des bonus/malus en fonction de la valeur de l'écart de rendement entre conduites « traité » et « non traité » avec des fongicides. L'application de cette approche sur chacun des jeux de données met en évidence une certaine constance du comportement de plusieurs génotypes (tolérant ou sensible) sur des jeux de données indépendants. En faisant le parallèle avec l'expérience acquise sur la GPD qui présente le même phénomène, ceci renforce nos conclusions sur la possibilité d'utiliser les indicateurs de tolérance à la carence azotée pour discriminer des variétés.



Figure 8 : Indicateurs de tolérance pour le rendement de l'expérimentation CTPS Nord 2015-2016 (réseau 2), calculés avec les moyennes pluriannuelles des rendements obtenus à dose X-80 et à dose X (en haut à gauche), ou avec les moyennes des années 2015 et 2016 (en haut à droite). Régression simple entre les valeurs obtenues pour chacune des deux années (en bas à gauche), et régression robuste entre ces mêmes valeurs (en bas à droite).

#### Elaboration d'indicateurs « simples » de l'efficience de valorisation de l'azote (réseau 2)

L'objet de cette partie de l'étude est d'élaborer des indicateurs variétaux d'efficience de l'azote apporté (sol et/ou engrais) dits « simples ». Nous entendons par « simple » un indicateur qui ne nécessite que les deux mesures réalisées en routine dans les essais d'inscription et de post-inscription, soient le rendement et la teneur en protéines du grain, par opposition aux indicateurs étudiés dans le réseau 1 nécessitant d'avoir accès à la quantité d'azote absorbé totale à la récolte (grains+pailles+racines). En effet, le coût analytique de telles mesures à l'échelle d'un réseau d'évaluation variétale est prohibitif. De plus, l'option choisie dans le projet de faire ces mesures uniquement sur un témoin «azote» et donc de disposer d'un NHI moyen par essai résout le problème de coût mais repose sur une hypothèse d'extrapolation de cette mesure à toutes les variétés de l'essai. Or, nous avons démontré avec le réseau 1 que la variété a un effet significatif sur le NHI, quelle que soit la dose d'engrais azoté apportée (et donc y compris à l'optimum). Ignorer cet effet en prenant un NHI moyen pour toutes les variétés est problématique dans un processus dont l'objectif est justement de mettre en évidence des différences de comportement variétal vis-à-vis de la nutrition azotée.

La figure 9 présente quatre exemples de régression « qN grains - dose d'azote ». On peut voir sur ces exemples que les pentes sont estimées de manière précise car les trois points relatifs aux doses X-80, X et X+40 sont bien alignés et les R<sup>2</sup> sont donc très élevés. En général, il est délicat d'extrapoler une relation linéaire en dehors de l'intervalle où elle a été mise en évidence. Cependant, les courbes de réponse à l'azote établies avec une gamme de doses plus large que celle utilisée pour le réseau 2 (réseau 1), montrent que la réponse est linéaire sur l'ensemble du domaine de variation de la dose d'azote. On peut donc considérer que les ordonnées à l'origine obtenues avec un dispositif à seulement trois doses peuvent être des estimations des qN grains des variétés en l'absence de fertilisation azotée.



Figure 9 : Régression entre qN grains et dose d'azote pour quatre génotypes de l'expérimentation CTPS zone Nord 2015-2016 (réseau 2). Filon et 1040127 ont des pentes très similaires, mais des ordonnées à l'origine très différentes. Chevignon et 1040163 ont à la fois des ordonnées à l'origine et des pentes très contrastées.

A l'image des indicateurs de tolérance, les valeurs des indicateurs d'efficience peuvent être obtenues soit comme la moyenne des valeurs de pente et d'ordonnée à l'origine calculées sur chacun des sites du réseau d'expérimentation, soit en calculant la régression à partir des moyennes multi-locales. Ces deux méthodes donnent des résultats très fortement corrélés : sur les sept jeux de données, les R<sup>2</sup> des régressions entre les deux types de valeurs sont comprises entre 0.84 et 0.99 pour la pente, et entre 0.74 et 0.98 pour l'ordonnée à l'origine. Pour des raisons de cohérence globale entre indicateurs, nous avons opté pour la méthode de calcul à partir des moyennes multi-locales.

Les analyses de variance montrent que l'effet « variété » n'est significatif que dans trois cas sur sept dans le cas de la pente (Tableau 4), et encore moins fréquemment significatif pour ce qui est de l'ordonnée à l'origine (deux cas sur sept avec une significativité faible). Pour cette dernière, c'est sans doute en partie dû à l'imprécision du calcul, car une faible erreur sur l'estimation de la pente peut conduire à de grandes variations dans le calcul des ordonnées à l'origine, puisqu'elles sont estimées loin du domaine d'établissement de la régression « qN grains - doses d'azote ». Dans les deux cas il y a sans doute aussi des interactions « génotype \* lieu » qui interviennent, même si nous n'avons pu les tester dans notre modèle d'ANOVA (les interactions « génotype \* année » sont, quant à elles, souvent non significatives). Quoiqu'il en soit, le calcul à partir des moyennes multi-locales permet probablement une certaine stabilisation des indicateurs d'efficience simples. Il est donc intéressant d'utiliser d'autres méthodes statistiques que l'ANOVA pour déterminer si ces indicateurs peuvent permettre de discriminer des génotypes pour la valorisation de l'azote apporté.

	Α	V	L (A)	$A \times V$			
Pente							
CTPS Nord 13-14	ns	**	***	(*)			
CTPS Sud 13-14	***	ns	***	ns			
FSOV 15-16	***	ns	***	ns			
CTPS Nord 14-15	***	**	***	*			
CTPS Sud 14-15	***	ns	***	ns			
CTPS Nord 15-16	***	ns	***	ns			
CTPS Sud 15-16	***	***	***	ns			
	Ordonnée	à l'origine					
CTPS Nord 13-14	***	ns	***	ns			
CTPS Sud 13-14	**	ns	***	ns			
FSOV 15-16	(*)	ns	***	ns			
CTPS Nord 14-15	***	(*)	***	*			
CTPS Sud 14-15	***	ns	**	ns			
CTPS Nord 15-16	***	(*)	***	ns			
CTPS Sud 15-16	***	ns	***	ns			

\*\*\* (p <0.001) ; \*\* (0.001<p<0.01) ; \* (0.01<p<0.05) ; (\*) (0.05<p<0.1) ; ns = non significatif

**Tableau 4 :** Analyse de variance sur les indicateurs d'efficience simple calculés sur le réseau 2 pour les 7 jeux de données. A = effet année, V = effet variété, L(A) = effet lieu (année)

Etant donné que l'effet « année » est très significatif (Tableau 4), nous avons procédé à une étude de stabilité des indicateurs d'efficience simple selon les mêmes méthodes employées pour les indicateurs de tolérance : corrélation par régression robuste entre années consécutives au sein du même jeu de données, puis comparaison du comportement de génotypes communs entre jeux de données indépendants. Les deux approches confirment le fait que les indicateurs d'efficience simples présentent un niveau de stabilité qui, au moins pour la pente, est tout à fait comparable à celui des GPDs, et qui peut donc être suffisant pour qu'ils soient utiles pour la caractérisation des variétés.

Tous les jeux de données présentent un niveau de corrélation fort entre l'indicateur « pente » et l'indicateur « ordonnée à l'origine » (Figure 10 pour un exemple).



Figure 10 : Régression entre les deux indicateurs d'efficience simples pour le jeu de données CTPS Sud 2013-14 (réseau 2).

Par construction, la pente et l'ordonnée à l'origine d'une régression sont liées négativement (lorsqu'on fait pivoter une droite de régression autour de son barycentre, une augmentation de pente conduit mécaniquement à une diminution de l'ordonnée à l'origine). Cependant, comme l'illustre la figure 9, il reste une part importante de la variabilité de l'ordonnée à l'origine qui n'est pas expliquée par la pente. Cet aspect a aussi été confirmé par l'analyse des données du réseau 1. Comme par ailleurs la position de la droite de régression entre les deux indicateurs d'efficience simples est établie de manière assez précise (sur les 7 jeux de données, le R<sup>2</sup> moyen est de 0.52, valeur tout à fait comparable à ce qu'on trouve en moyenne pour la relation « rendement - teneur en protéines »), on peut envisager d'utiliser les écarts à la régression pour mettre en évidence les génotypes présentant un bon compromis pour les valeurs des deux indicateurs.

L'analyse du lien statistique entre les indicateurs de tolérance et les indicateurs d'efficience simples indique une corrélation très faible entre les 2 familles d'indicateurs, quel que soit le jeu de données considéré. La capacité à supporter une baisse de fertilisation azotée semble donc assez indépendante de l'efficience d'utilisation de l'azote apporté par le sol ou par les engrais. Par contre, une analyse des relations entre les indicateurs d'efficience simple et les GPD par régression simple et PLS indiquent que les indicateurs d'efficience simples sont en fait deux composantes essentielles des GPDs. De manière plus précise en distinguant les situations X-80 et X, il apparait que l'efficience de valorisation de l'azote vue à travers la GPD, correspond plutôt à la capacité du génotype à valoriser l'azote du sol (et donc à l'ordonnée à l'origine de la régression « qN grains - dose d'azote ») lorsqu'on est en situation de carence azotée (dose X-80), et plutôt à la capacité du génotype à valoriser l'apport d'engrais (et donc à la pente de la régression) lorsqu'on est en fertilisation azotée normale (dose X).

La méthode employée pour calculer les indicateurs d'efficience simples a recours aux résultats obtenus sur trois conduites de fertilisation azotée (X-80, X, X+40). Leur calcul étant mathématiquement possible avec seulement les modalités X-80 et X, nous avons étudié la robustesse de cette approche en vue d'optimiser les coûts des réseaux expérimentaux. La figure 11 présente les corrélations entre les pentes mesurées en faisant varier le nombre de modalités de fertilisation azotée utilisées à partir des données acquises sur le réseau 1. Il apparaît une forte corrélation entre les méthodes. Ce résultat n'est pas étonnant étant donné le caractère fortement linéaire de la relation « gN grains-doses d'azote ». Des résultats similaires (avec même des corrélations légèrement plus élevées) ont été obtenus pour l'ordonnée à l'origine. Le même type d'analyse de sensibilité a été réalisée sur les jeux de données du réseau 2 (comparaison entre les indicateurs calculés avec X-80/X/X+40 et ceux calculés avec seulement X-80/X). Les corrélations obtenues sont également assez fortes (leur valeur moyenne est de 0.77, aussi bien pour la pente que pour l'ordonnée à l'origine) et nous retrouvons pour les indicateurs calculés sur 2 modalités l'ensemble des propriétés des indicateurs calculés sur 3 modalités (notamment le lien avec les GPDs). En première approximation, il semble donc possible de se contenter de dispositifs expérimentaux avec deux doses d'azote pour obtenir des indicateurs d'efficience simples utilisables.

Au regard du lien « explicatif » entre les indicateurs d'efficience simple et les GPDs, leur valorisation dépendra du besoin de l'utilisateur final (chercheur, sélectionneur, préconisateur, agriculteur) en terme d'explicitation des sources de variation des GPDs calculées.



**Figure 11 :** Corrélations entre 3 modes de calculs de l'indicateur d'efficience simple « pente » testés sur les données du réseau 1. « Pente.2pts » = calcul avec les points X-80 et X ; « Pente.3pts » = calcul avec les points X-80, X et X+40 ; « Pente.CRN » = calcul avec les points de la courbe de réponse à l'azote.

#### Conception d'indicateurs « élaborés » de l'efficience de valorisation de l'azote (réseau 2)

Une première approche méthodologique a été réalisée pour déterminer si l'ensemble des indicateurs étudiés sur le réseau 1 étaient calculables sur un dispositif ayant les caractéristiques du réseau 2. La 1<sup>ère</sup> spécificité du réseau 1 est son dispositif d'évaluation des variétés en courbe de réponse complète à la dose d'engrais azoté, permettant la caractérisation des variétés à l'optimum de fertilisation azotée pour le rendement et un niveau donné de teneur en protéines du grain. Nous avons donc tout d'abord testé différents mode de calage d'un modèle statistique de courbe de réponse du rendement à la dose d'engrais N à partir d'un dispositif en 3 doses. Testé sur les données du réseau 1 en comparaison des résultats obtenus sur un modèle ajusté sur 6 doses, il apparaît que l'estimation du rendement optimal est satisfaisante (Figure 12, R<sup>2</sup> de la régression = 0.92), que celle de la

45

teneur en protéines est moins bonne (R<sup>2</sup> de la régression = 0.58) et que celle de la dose d'azote optimale est très mauvaise (R<sup>2</sup> de la régression = 0.02). A partir de cette approche et l'extrapolation des NHI des variétés témoins à l'ensemble des variétés du même essai, nous avons calculé l'ensemble des indicateurs testés sur le réseau 1 sur le réseau 2. Pour les variétés présentes dans les deux réseaux, les résultats obtenus sur le réseau 2 sont globalement cohérents avec ceux obtenus sur le réseau 1. Néanmoins, l'imprécision du mode de calage des courbes de réponses à l'azote et le recours à une hypothèse forte d'emploi d'un NHI moyen pour toutes les variétés dans un dispositif d'évaluation variétale nous a fait écarter cette approche, sous réserve d'investigations futures plus poussées.



Figure 12 : Lien entre le calcul du rendement optimal avec une courbe de réponse azote complète (QP) et une courbe de réponse azote à 3 points (X-80/X/X+40, 3pts). Calculs réalisés sur le réseau 1.

#### Effet variété et dose d'azote sur la qualité technologique des blés (réseau 1 et réseau 2)

Des tests de panification de type pain courant français (NF V03-716) ont été effectués sur des farines issues des récoltes 2014, 2015 et 2016 des essais du réseau 1. Il en ressort que, dans la plage 9-13% de teneurs en protéines, l'impact de la dose d'azote sur la note totale de panification est faible (Figure 13) alors que des interactions « variétés-essais » ressortent comme les plus fortes du fait d'effets lieux sur la note d'allongement. Les situations qui génèrent de l'extensibilité accentuent le défaut des variétés à profil extensible, mais favorisent les variétés à profil tenace (pâtes plus courtes). A l'inverse, les situations qui génèrent de la ténacité profitent aux variétés à profil extensibles. Dans certains cas, l'accroissement de la dose d'azote sur un lieu donné peut légèrement accentuer le raccourcissement des pâtes.

Les analyses dont nous disposons sur le profil des protéines analysé par HPLC complètent la première approche. Pour rappel, les gluténines de haut poids moléculaire (fraction F1) jouent un rôle essentiel sur les propriétés de ténacité. Par conséquent, plus le rapport %F1/%F2 est élevé plus la pâte est tenace. L'analyse des données issues des récoltes 2014, 2015 et 2016 a permis d'obtenir les résultats suivants :

- Le rapport %F1/%F2 n'évolue pas avec la dose d'azote, mais il varie en fonction du lieu et de l'année.
- Les situations avec un rapport %F1/%F2 faible, génèrent de l'extensibilité et accentuent le défaut des variétés à profil extensible, mais favorisent celles à profil court. Au contraire, celles avec un rapport %F1/%F2 plus élevé génèrent des pâtes courtes ce qui profite aux variétés à profil extensible.

Les variations dans le comportement en panification des variétés semblent donc attribuables à un effet « qualité des protéines » et non à effet « quantité des protéines ». Cependant, la qualité des protéines n'explique pas à elle seule le comportement en panification.



Figure 13 : Effet de la dose d'engrais azoté appliquée (à gauche) et de la teneur en protéines (à droite) sur la note de panification des variétés RUBISKO et SOISSONS. Essais du réseau 1 2014-2016. Doses d'engrais azotées exprimées en écart à la dose optimale pour le rendement calculée a posteriori sur la base d'une courbe de réponse. Des analyses supplémentaires ont été effectuées sur le réseau 1 récoltes 2015 et 2016, à savoir :

- Un alvéogramme de Chopin (NF EN ISO 27971) : la force boulangère (W) augmente avec la teneur en protéines, qui elle-même augmente avec la dose des apports d'azote. Le rapport P/L diminue globalement avec la teneur en protéines et la dose d'azote.
- Des mesures de Gluten Humide et Gluten Index (NF EN ISO 21415-2) : le gluten humide augmente avec la teneur en protéines et la dose d'azote.

Le classement des variétés sur le W, le P/L et le Gluten humide est peu modifié en fonction de la dose d'azote apportée dans une plage de X-80 à X+80 kg N.ha<sup>-1</sup>. De la même façon, les notes totales de panification ne ressortent pas sensibles à la dose d'azote, avec peu de variation de classement des variétés sur leur note de panification, excepté dans les situations à teneurs en protéines inférieures à 9% ou supérieures à 13 %.

Des analyses de qualité technologiques ont aussi été réalisées sur des échantillons des modalités X et X-80 kg N.ha-1 de 21 variétés de 3 essais du réseau 2 2015 et 2016. Les analyses technologiques sont les tests de l'alvéogramme de Chopin (NF EN ISO 27971) et de panification normalisée « pain courant français » (NF V03-716). Compte tenu du faible nombre d'essais étudiés et afin de vérifier la conformité et la représentativité des 3 essais retenus, les données ont été comparées avec celles obtenues dans le cadre des épreuves de post-inscription en 2015 (provenant du réseau CTPS 2) pour les témoins Apache et Arezzo. Les W, P/L et notes de panification des variétés témoins à la dose X des 3 essais sont confirmés comme étant cohérents. Pour les variétés inscrites en 2015 et présentes dans le réseau 2, le jeu de données a été complété avec des données acquises dans le cadre des épreuves d'inscription (réseaux CTPS 1 2013 et CTPS 2 2014) et de Post-Inscription. L'étude des critères de qualité technologique, montre que :

- Pour le W, le P/L et la note totale de panification, le classement des variétés expérimentées dans le réseau 2 en 2015 et en 2016 est peu modifié en fonction de la dose d'azote apportée dans une plage de X à X+80.
- Pour chaque variété testée individuellement, la note totale de panification est peu modifiée en fonction de la dose d'azote.
- L'interaction variétés-lieux est plus significative que l'effet de l'azote ou teneur en protéines (Figure 14).

#### Indicateur d'efficience de valorisation rendement et protéines des variétés inscrites (réseau 3)

Afin de mettre à jour le mode de calcul des coefficients d'efficience de valorisation de l'azote des variétés inscrites en France (besoin unitaire en azote à l'optimum de rendement - coefficient « b » destiné au calcul de la dose d'azote prévisionnelle par les agriculteurs) et d'y inclure un objectif spécifique d'atteinte d'une teneur en protéines de 11.5% (coefficients « bc » et « bq »), nous avons procédé à une étude méthodologique complète basée sur le réseau 3, en faisant appel aux données des réseaux 1 et 2 afin de valider certaines hypothèses de calculs inhérentes à la nature des essais utilisés (longues listes variétales testées sous un seul régime de fertilisation azoté supposé optimal dans de larges réseaux multi-locaux et pluriannuels). Le résultat de cette étude et la méthode de calcul des coefficients « b », « bc » et « bq » sont résumés dans la figure 15.

L'étape 1 passe par le calcul de la quantité d'azote présente dans le grain à la récolte (QNG) pour toutes les variétés dans tous les essais. C'est le même calcul qui est à la base des indicateurs d'efficience simples. Les résultats acquis sur le réseau 1 nous permettent d'élaborer une équation d'estimation de la quantité d'azote totale prélevée à la récolte (QNT) à partir du QNG (étape 2).



Figure 14 : comportement moyen des variétés en notes de pâte, de pain, de mie et totale de panification en fonction de la dose d'azote et des lieux. 3 essais du réseau 2 2015 et 2016.





A la différence d'indicateurs à utiliser dans les processus de sélection ou d'inscription, nous nous autorisons ici à négliger l'impact de la variété sur le NHI car 1) l'effet de la variété sur le NHI est significatif mais d'impact mineur sur le classement des variétés relatif aux coefficients b, bc et bg (analyse de sensibilité réalisée sur le réseau 1) et 2), l'utilisation des coefficients b, bc et bg est réservée à la définition de la dose d'azote optimale à appliquer par les agriculteurs dont le calcul est influencé par de nombreux autres facteurs de variation d'impact plus conséquent (climat, pratiques d'apports des engrais...). Par définition, le coefficient « b » est calculé à l'optimum de fertilisation azotée pour le rendement (Le Souder et Bernicot 1993). Les essais du réseau 3 ne comportant qu'une seule dose N, l'hypothèse que cette dernière correspond bien à la dose N optimale est forte. Afin de l'évaluer, nous avons procédé à une analyse de sensibilité du classement variétal « b », « bc » et « bq » selon la dose d'azote appliquée sur les réseaux 1 et 2. Il s'avère que le classement reste globalement inchangé quand on se place au voisinage de la dose N optimale ± 40 kg N.ha<sup>-1</sup>. A l'aide du modèle CHN (Soenen et al. 2016), nous avons procédé à

une sélection des essais du réseau 3 selon leur statut azoté à floraison et leur niveau de rendement et de teneur en protéines (étape 3). Sur cette base, le coefficient « b » est calculé pour toutes les variétés dans tous les essais sélectionnés (étape 4). En utilisant les moyennes ajustées des teneurs en protéines à l'échelle du réseau, le coefficient « bc » (étape 5) puis « bq » (étape 6) sont calculés pour l'ensemble des variétés. A partir de cette méthode cette fois appliquée à l'ensemble des essais d'évaluation de postinscription conduits par ARVALIS depuis de nombreuses année, nous avons pu produire dès 2016, un classement innovant de l'ensemble des variétés de blé tendre d'hiver disponibles en France vis-à-vis de l'efficience de valorisation de l'azote pour le rendement et la teneur en protéines du grain. La figure 16 présente la version de ce classement paru en décembre 2016 (et mis à jour tous les ans depuis). Notons que les préconisations de compléments de besoin « bc » s'accompagnent de préconisations spécifiques sur le fractionnement des apports d'engrais afin que ce complément bénéficie en priorité à l'accumulation de protéines dans le grain.

## 4. Conclusion et valorisation

ARVAI

Les résultats produits par le projet FSOV N-BT ont permis de répondre aux objectifs fixés initialement et de proposer plusieurs valorisations opérationnelles. En 1 er lieu, à partir du réseau 1, nous avons produit des résultats originaux et détaillés sur l'efficience d'absorption et de valorisation de plusieurs variétés de blé tendre d'hiver récemment développées en France. Grâce à un dispositif expérimental combinant les méthodes des études agronomiques (dispositif en courbe de réponse azote) et génétiques (décomposition de la NUE), l'impact significatif des variétés sur les 2 composantes de la NUE a pu être mis en évidence à leur

optimum de fertilisation respectif. Le projet fournit donc une base de référence sur l'efficience « azotée » de variétés élites encore largement utilisées par les sélectionneurs et les chercheurs. Les résultats issus de ce dispositif ont pu aussi servir de données de base pour valider plusieurs hypothèses de travail dans le développement d'indicateurs plus adaptés à des dispositifs expérimentaux plus simples (réseau 2 et réseau 3). Afin de permettre la valorisation de ces résultats dans de futurs projets de recherches, les résultats acquis sur le réseau 1 ont fait l'objet d'une publication scientifique (en cours de soumission au moment de la rédaction de cet article scientifique). En 2ème lieu, le projet a permis l'élaboration d'indicateurs de tolérance à la carence azotée des variétés de blé tendre d'hiver, utilisables pour le rendement et la teneur en protéines du grain dans différents dispositifs d'évaluation variétale (sélection, inscription, post-inscription) à condition de disposer de 2 conduites de fertilisation azotée bien discriminées dans les essais. Pendant la durée du projet, ces indicateurs de tolérance à la carence azotée ont été directement valorisés par le CTPS-GEVES car ils sont publiés à titre indicatif depuis 2017 à partir des inscriptions 2014 (CTPS/GEVES 2017, Mailliard et al. 2017). Leur passage au statut de variables rentrant dans le calcul des cotations est actuellement en cours de discussion au sein de la section céréales à paille du CTPS. En parallèle, nous avons élaboré des indicateurs d'efficience « simple » de l'azote fourni (par le sol ou l'engrais). Ceux-ci peuvent être calculés sur le même type de dispositif que les indicateurs de tolérance et sont à même de discriminer les variétés avec la même puissance statistique que la GPD. Néanmoins, leur forte corrélation avec cette dernière pose la question de la redondance des indicateurs entre eux, à moins de souhaiter fournir une explication physiologique à la GPD en la

## CLASSEMENT DES VARIÉTÉS SELON LEUR BESOIN EN AZOTE (COEFFICIENTS b ET bq<sub>11.5%</sub>) (en kgN/q)



LASSES		CLASSES	Modai respec	ités de fractionnement à ter en utilisant bg11.5%
DE b	VARIETES	DE bq11.5%	bc11.5%	Mise en réserve minimale conseillée pour la fin de montaison
2.8	Addict, Adhoc, Advisor, Aigle, Ambition, Arlequin, Armada, Atoupic, Basmati, Bermude, Boisseau, Complice, Costello, Creek, Diderot, Fairplay, Folklor, Garcia, Granamax, Hybello, Hybery, Hybiza, Hyclick, Hydrock, Hyguardo, Hyking, Hystar, Hysun, Hyteck, Hywin, JB Diego, Kundera, Lear, Lithium, Lyrik, Modern, Popeye, RGT Mondio, RGT Texaco, Salvador, Sokal, Stadium, Stereo, Trapez, Tremie, Viscount, Zephyr	3	0.2	60 kg N (40*+20)
	Glasgow, Istabraq, Sobred, Torp	3.2	0.4	70 kg N (40*+30)
3	Accor, Alhambra, Allez Y, Altigo, Andino, Apache, Apanage, Aplomb, Aprilio, Arezzo, As De Coeur, Aubusson, Bagou, Bonifacio, Boregar, Brentano, Buenno, Calabro, Calcio, Calisol, Calumet, Cellule, Cezanne, Chevalier, Comilfo, Compil, Descartes, Diamento, Distinxion, Ephoros, Euclide, Fluor, Forblanc, Foxyl, Galactic, Galopain, Goncourt, Gotik, Hyfi, Hyxo, Hyxpress, Illico, Interet, Isengrain, Kalystar, Koreli, Lavoisier, LG Abraham, LG Absalon, LG Altamont, Memory, Musik, Nucleo, Numeric, Oregrain, Paledor, Pibrac, Prevert, Reciproc, RGT Ampiezzo, RGT Cesario, RGT Kilimanjaro, RGT Tekno, RGT Velasko, RGT Venezio, Sollario, Solognac, Solveig, Sothys CS, Sponsor, Starway, Syllon, Vyckor	3	0	40* kg N
	Accroc, Alixan, Andalou, Aristote, Arkeos, Ascott, Auckland, Barok, Belepi, Bergamo, Chevron, Collector, Expert, Fructidor, Gallixe, Grapeli, Hyxtra, Ionesco, Laurier, Matheo, Milor, Nemo, Oxebo, Pakito, Pr22r58, RGT Celesto, RGT Libravo, RGT Sacramento, Ronsard, Sherlock, SY Mattis, SY Moisson, System, Terroir, Thalys, Tobak, Triomph, Valdo, Waximum	3.2	0.2	60 kg N (40*+20)
3.2	Aerobic, Altamira, Ambello, Athlon, Attlass, Bienfait, Camp Rémy, CCB Ingenio, Centurion, Exelcior, Exotic, Falado, Graindor, Hendrix, Lazaro, Lukullus, Manager, Nogal, Scipion, Soissons, Sorrial, Tulip	3.2	0	40* kg N

Figure 16 : Classement « b », « bc » et « bq » des variétés de blé tendre d'hiver disponibles en France fin 2016. Valeurs élaborées à partir de la méthode mise au point dans le cadre du projet et appliquée à l'ensemble des données de post-inscription obtenues dans les réseaux expérimentaux d'ARVALIS-Institut du végétal.

décomposant selon que la nutrition azotée est majoritairement assurée par l'azote fourni par le sol ou par l'engrais. En 3ème lieu, les analyses de qualité technologique « panification » mettent en évidence l'effet prépondérant de la variété et du lieu d'essai (interaction avec l'environnement) sur ces critères. La dose d'azote ne joue qu'un très faible rôle dans le classement d'aptitude à la panification des variétés, bien qu'elle ait des conséquences sur la teneur en protéines du grain. Des dispositifs spécifiques « azote » pour évaluer les propriétés des variétés de blé tendre d'hiver vis-à-vis de la panification ne semblent pas nécessaires. Enfin, l'ensemble des résultats issus des réseaux 1, 2 et 3 ont permis de valider les hypothèses nécessaires à l'élaboration d'un nouveau référentiel de besoins unitaires en azote des variétés de blé tendre d'hiver disponibles en France. Celui-ci, élaboré à partir de l'ensemble des données d'évaluation post-inscription issues des essais conduits par ARVALIS-Institut du végétal, fournit les informations nécessaires pour calculer le besoin en azote du blé

tendre d'hiver, indispensable au calcul de la dose d'azote prévisionnelle par les agriculteurs, pour différents objectifs du couple rendement/teneur en protéines du grain. Fin 2016, la 1<sup>ère</sup> version de ce nouveau référentiel a été mise à la disposition des agriculteurs et des techniciens qui les accompagnent et est mis à jour annuellement depuis (www.arvalis.fr).

Étant donné la grande richesse des jeux de données produits, la valorisation des résultats se poursuivra sous d'autres formes pour répondre à d'autres questionnements ou pour approfondir certains points insuffisamment développés dans la durée du projet.

#### Remerciements :

Les partenaires du projet tiennent à remercier le GNIS-FSOV pour son soutien, l'ensemble des équipes techniques ayant réalisé les essais et les projets Breedwheat (ANR-2010-BTBR-003), FSOV 2003A « azote » 2004-2005 et FSOV 2010F « azote » 2011-2012 pour la mise en commun des jeux de données « réseau 2 ».

## Références bibliographiques

**ARVALIS-Institut du végétal** (2013). Teneur en protéines des blés : relever le double défi agronomique et économique. 6p, ARVALIS Editions.

**Bouthier A.** (1997). Efficacité d'un apport d'azote sur blé tendre : la pluie est rarement limitante courant montaison. Perspectives Agricoles 220, 56-59.

**Chambers B.J., Dampney P.M.R.** (2009). Nitrogen efficiency and ammonia emissions from urea-based and ammonium nitrate fertilizers. IFS proceedings 657.

**Cohan J.P, Bouthier A.** (2010). Engrais azoté - Pluie après apport : un facteur d'efficacité déterminant. Perspectives Agricoles 354, 38-39.

Cohan J.P., Oury F.X., Mailliard A., Lorgeou J., Le Souder C., Praud S., Bernicot M.H., Le Gouis J. (2017). Effect of variety on *Nitrogen Use Efficiency* of bread-wheat : from breeding programs to farm practice, in: Proceedings of the International Fertiliser Society. Presented at the Agronomic conference of the International Fertiliser Society, Cambridge - UK. December 12th 2017.

Cohan J.P., Hannon C., Houilliez B., Gravoueille J.M., Geille A., Lampaert E., Laurent F. (2018). Effects of Potato Cultivar on the Components of *Nitrogen Use Efficiency*. Potato Research 61, 231-246.

Cormier F., Faure S., Dubreuil P., Heumez E., Beauchêne K., Lafarge S., Praud S., Le Gouis J. (2013). A multi-environmental study of recent breeding progress on *Nitrogen Use Efficiency* in wheat (Triticum aestivum L.). Theoretical and Applied Genetics 126, 3035-3048.

Cormier F., Foulkes J., Hirel B., Gouache D., Moënne-Loccoz Y., Le Gouis J. (2016). Breeding for increased nitrogen-use efficiency: a review for wheat (T. aestivum L.). Plant Breed 135, 255-278.

**CTPS/GEVES** (2017). Résultats de l'évaluation du comportement des variétés de blé tendre d'hiver vis à vis de l'azote réalisée pour le CTPS dans le cadre de l'inscription des variétés au catalogue national. Document CTPS/GEVES Janvier 2017, pp 22. Disponible à l'adresse https://cat.geves.info/CAT\_WEB/Data/ COMPORTEMENTAZOTEVARBTH201420152016.pdf.

Foulkes M.J., Sylvester-Bradley R., Scott R.K. (1998). Evidence for differences between winter wheat cultivars in acquisition of soil mineral nitrogen and uptake and utilization of applied fertilizer nitrogen. The Journal of Agricultural Science 130, 29-44.

Foulkes M.J., Hawkesford M.J., Barraclough P.B., Holdsworth M.J., Kerr S., Kightley S., Shewry P.R. (2009). Identifying traits to improve the nitrogen economy of wheat: Recent advances and future prospects. Field Crops Research 114, 329-342.

Galloway J.N., Aber J.D., Erisman J.W., Seitzinger S.P., Howarth R.W., Cowling E.B., Cosby B.J. (2003). The Nitrogen Cascade. BioScience, 53, 341-356.

**Good A.G., Shrawat A.K., Muench D.G.** (2004). Can less yield more? Is reducing nutrient input into the environment compatible with maintaining crop production? Trends in Plant Science 9, 597-605.

Hawkesford M.J. (2014). Reducing the reliance on nitrogen fertilizer for wheat production. Journal of Cereal Science, Cereal Science for Food Security, Nutrition and Sustainability 59, 276-283.

**Hirel B., Le Gouis J., Ney B., Gallais A.** (2007). The challenge of improving *Nitrogen Use Efficiency* in crop plants: towards a more central role for genetic variability and quantitative genetics within integrated approaches. J Exp Bot 58, 2369-2387.

Jensen L. S., Schjoerring J. K., Van der Hoek K. W., Poulsen H.D., Zevenbergen J. F., Pallière C., Brentrup F., Jongbloed A. W., Willems J., Van Grinsven H. (2011). Benefits of nitrogen for food, fibre and industrial production. The European Nitrogen Assessment, ed. Mark A. Sutton, Clare M. Howard, Jan Willem Erisman, Gilles Billen, Albert Bleeker, Peringe Grennfelt, Hans van Grinsven and Bruna Grizzetti, Cambridge University Press, pp 32-61.

Laurent F., Leveau V. (2014). La grande culture face aux (r)évolutions des marchés et politiques agricoles. In Fertilisation et environnement. Quelles pistes pour l'aide à la décision ? S. Pellerin, F. Butler, C. Guiard-Van Laethem coord. Editions QUAE, pp 59-72.

Le Souder C., Bernicot M.H. (1993). Blé tendre et bilan azoté -Faut-il fertiliser de la même façon toutes les variétés? Perspectives Agricoles 179, 67-73.

Limaux F., Recous S., Meynard J.M., Guckert A. (1999). Relationship between rate of crop growth at date of fertiliser N application and fate of fertiliser N applied to winter wheat. Plant and Soil 214, 49-59.

Mailliard A., Bernicot M.H., Lorgeou J., Cohan J.P. (2017). Tolérance du blé tendre à la carence azotée - des variétés aux profils opposés. Perspectives Agricoles 445, 10-13.

Méléard B. (2007). Choisir une variété - Mettre ses atouts qualité au service d'un débouché. Perspectives Agricoles, 395 (décembre), 18-19.

Moll R.H., Kamprath E.J., Jackson W.A. (1982). Analysis and interpretation of factors which contribute to efficiency of nitrogen utilization. Agronomy Journal 74, 562.

Oury F.X., Bérard P., Brancourt-Hulmel M., Depatureaux C., Doussinault G., Galic N., Giraud A., Heumez E., Lecomte C., Pluchard P., Rolland B., Rousset M., Trottet M. (2003) Yield and grain protein concentration in bread wheat: a review and a study of multi-annual data from a French breeding program. J. Genet. & Breed. 57: 59-68.

**Oury F.X., Godin C.** (2007). Yield and grain protein concentration in bread wheat: how to use the negative relationship between the two characters to identify favourable genotypes? Euphytica 157, 45-57.

**Rémy J.C., Hébert J.** (1977). Le devenir des engrais dans le sol. Compte Rendu de l'Académie d'Agriculture Française 63, 700-710.

Rémy, J.C., Viaux P. (1983). La fertilisation azotée du blé tendre en système intensif en France. Perspectives Agricoles 67, 26-34.

Soenen B., Le Bris X., Laberdesque M., Cohan J.P., Laurent F., Bouthier A., Gouache D., Garcia C. (2016). "CHN", a crop model to jointly manage water and nitrogen on winter wheat. Presented at the International Crop Modelling Symposium, Berlin.

Soenen B., Closset M., Bonnard A., Le Bris X. (2017). Validation of a new nitrogen management tool on winter wheat based on remote sensing diagnostic and agronomic prognosis: 'QN-method' - FARMSTAR. Proceedings of 'Innovative Solutions for Sustainable Nitrogen Management', June 26-28 2017, Aarhus, Denmark.

Sutton M.A., Oenema O., Erisman J.W., Leip A., Grinsven H. van, Winiwarter W. (2011). Too much of a good thing. Nature 472, 159-161.

Sylvester-Bradley R., Kindred D.R. (2009). Analysing nitrogen responses of cereals to prioritize routes to the improvement of *Nitrogen Use Efficiency*. Journal of Experimental Botany 60, 1939-1951.

Sylvester-Bradley R., Kindred D.R., Wynn S.C., Thorman R.E., Smith K.E. (2014). Efficiencies of nitrogen fertilizers for winter cereal production, with implications for greenhouse gas intensities of grain. The Journal of Agricultural Science 152, 3-22.

## [SG-Rendement] Calibration et Implémentation d'un outil de Sélection Génomique Rendement dans un programme de sélection blé tendre

#### Laure DUCHALAIS \*, Simon TEYSSEDRE<sup>1</sup>, Bruno CLAUSTRES<sup>1</sup>, Christophe MICHELET<sup>1</sup>, Matthieu BOGARD<sup>2</sup>

1 - RAGT 2n SAS Rue Emile SINGLA, Site de Bourran - B.P. 3336 - 12033 RODEZ Cedex 9

2 - ARVALIS- Institut de Végétal, 6 chemin de la côte vieille, 31450 BAZIEGES

\* Coordinateur : Laure DUCHALAIS, Iduchalais@ragt.fr

## 1. Introduction

La situation des producteurs français n'a jamais été aussi difficile. Le prix de la matière première est très variable et le prix des intrants ne cesse d'augmenter, les revenus agricoles sont donc largement impactés.

Le monde agricole est en constante évolution grâce aux nouvelles pratiques et aux nouveaux outils de gestion culturale. Cependant, même si ces évolutions tendent à minimiser les dépenses des agriculteurs, elles ne leurs garantissent pas un revenu approprié. Dans un tel environnement économique, il est donc indispensable de garantir aux agriculteurs des rendements élevés et stables. Malheureusement, une stagnation des rendements de blé tendre a été constatée, depuis 1996. Les gains génétiques et autres améliorations technologiques ne semblent pas suffire à compenser les aléas climatiques observés au cours des deux dernières décennies. La sélection de variétés de blé adaptées aux contraintes futures (climatiques, économiques...) et à l'environnement ciblé est devenue un facteur essentiel pour la durabilité de nos secteurs agricoles.

La disponibilité de milliers de marqueurs moléculaires à l'échelle du génome a rendu l'utilisation de la sélection génomique (SG) possible chez les plantes et les animaux (Meuwissen *et al.*, 2001 ; Bernado et Yu, 2007; de Los Campos et al., 2009, 2010; Crossa *et al.*, 2010, 2011; Pérez et al., 2010). Afin de pouvoir utiliser la SG, des données de phénotypage et de génotypage doivent être collectées sur une population dite d' « entraînement ». Ces données sont ensuite utilisées pour estimer les effets des allèles, et créer ainsi un modèle de prédiction. Ce modèle de prédiction peut ensuite être appliqué aux génotypes des candidats de la population cible afin de calculer leur « Genomic Estimated Breeding Values » (GEBVs). La sélection s'effectue alors sur la base de ces GEBVs.

Les sélectionneurs font face à différents types de problèmes de prédiction. Un premier problème consiste à prédire les valeurs génétiques des lignées nouvellement développées en début de cycle. Ces lignées n'ont généralement que très peu voire aucune donnée de phénotypage. La précision de la prédiction dépend alors beaucoup du degré d'apparentement entre les candidats de sélection et les individus utilisés pour former l'équation de prédiction. Un deuxième problème de prédiction se rapporte à la prévision des performances des lignées avancées qui ont été évaluées dans un ou plusieurs environnements. Ici, la tâche principale est de prédire les performances dans les environnements où la lignée n'a pas encore été évaluée.

Une piste d'amélioration des modèles de SG pour des caractères complexes comme le rendement en grains qui sont sous l'influence de nombreux facteurs génétiques, environnementaux et de leurs interactions, est la prise en compte d'effets génétiques non additifs. En effet, la plupart des études sur la SG font implicitement l'hypothèse que les caractères étudiés sont sous l'influence de QTLs dont les effets s'additionnent. Or, il est raisonnable de penser que des interactions entre QTLs existent et ont un impact sur ces caractères.

Dans la perspective d'étudier la stabilité des performances des génotypes, il apparait nécessaire d'introduire dans les modèles de SG des composantes pour prédire les effets environnementaux et les interactions GxE afin d'être en mesure de prédire les performances génotypiques dans des environnements non testés. Des études sur ce sujet ont été réalisées (Jarquin *et al.*, 2014) et ont permis de proposer de tels modèles en utilisant des covariables calculées sur la base de données météorologiques et agronomiques.

Ce projet vise à démontrer l'efficacité d'utilisation d'un outil de SG rendement dans un programme de sélection de blé tendre et à améliorer les modèles de SG par la prise en compte des effets environnementaux et des interactions GxE.

## 2. Matériel et méthode

2.1 - Application d'un outil de SG rendement dans un programme de sélection de blé tendre

#### Matériel végétal

#### Population d'entraînement R2n

La population d'entraînement R2n se compose de 264 lignées : 11 variétés commerciales et 253 lignées avancées dans le programme de sélection R2n.

#### Population cible n°1

La population cible n°1 rassemble 500 lignées au stade Fn au cours de la saison 2014/2015. Ces lignées ont été choisies afin que chacun des croisements réalisés par R2n soit représenté. Pour les croisements aux effectifs les plus faibles, seule une lignée a été choisie. Pour les croisements aux effectifs plus importants, plusieurs lignées ont été choisies aléatoirement en fonction de leur effectif total. Enfin, la totalité des lignées des deux croisements majoritaires (41 lignées pour l'un et 46 lignées pour l'autre) a été sélectionnée afin de pouvoir étudier les prédictions de la SG intra-croisement.

#### Population cible n°2

La population cible n°2 se compose de 482 lignées au stade Fn+1 au cours de la saison 2015/2016. De précocité intermédiaire à tardive, elles représentent une grande majorité de la diversité disponible à ce stade dans le programme de sélection blé tendre R2n.

#### Génotypage

Les extractions d'ADN des lignées appartenant aux différentes populations ont été réalisées au laboratoire de biologie moléculaire chez RAGT Seeds. L'ensemble des échantillons a ensuite été envoyé à l'université de Bristol où ils ont été génotypés avec la puce Axiom 35 K. Les individus et marqueurs avec plus de 20% de données manquantes ont été supprimés. Les marqueurs ayant une fréquence de l'allèle minoritaire (MAF) < 1% ont également été supprimés.

#### Phénotypage

#### Population d'entraînement R2n

Les 264 lignées constituant la population d'entraînement R2n ont été au moins phénotypées pour le rendement sur 5 lieux en 2 répétitions traitées en 2014 et en 2015. Les conduites culturales ont été menées à l'optimum afin de bien évaluer le potentiel de productivité des différentes lignées.

#### Population cible n°1

Les lignées de la population cible n°1 ont été évaluées pour le rendement traité sur une seule micro-parcelle à Louville-la-Chenard (28) en 2015. Elles ont été implantées sous forme d'essais à 30 entrées dont 4 témoins répétés.

#### Population cible n°2

Les lignées de la population cible n°2 ont été évaluées pour le rendement traité sur un site sans répétition en 2015 et sur 5 sites différents en 2 répétitions en 2016. Elles ont été implantées sous forme d'essais à 30 entrées dont 4 témoins répétés.

Au sein de cette population, 30 individus ont pu être évalués pour le rendement traité en 2017 sur 13 lieux communs français en 2 répétitions.

#### Ajustement des données

Dans chaque essai, les effets terrains ont été ajustés selon cette formule :

#### $Y_i = mean + fixed(variety) + fixed(Rep) + e$

Pour les modèles de SG sans prise en compte des effets GxE, une moyenne ajustée des variétés sur l'ensemble des lieux a été calculée à partir de tous les  $\hat{Y}_i$  précédents et du modèle :

 $\hat{Y}$  = mean + fixed(Loc) + fixed (variety) + random(variety x loc) + e Avec :

Variety x loc ~ 
$$N(0, I\sigma_{ge}^2)$$

Test de différents modèles de SG sur la population cible n°1 et application du meilleur modèle sur cette dernière

#### Modèles testés

Différents modèles de sélection génomique ont été testés sur la population cible n°1 : un modèle G-BLUP prenant en compte uniquement les effets d'additivité, un modèle G-BLUP prenant en compte à la fois les effets d'additivité et d'épistasie, un modèle RR-BLUP utilisant GS3 (Legarra *et al.*, 2013), un modèle RKHS utilisant BGLR (50 000 itérations et 20 000 burnin période) et un modèle bayésien utilisant BayesR avec 4 classes (Rvarg\*c (0.0001, 0.001, 0.01) ; 50 000 itérations avec 20 000 burnin périodes) ont été appliqués.

#### Validation croisée

Afin de valider l'efficacité des différents modèles testés, une validation croisée de type k-fold a été réalisée avec k=10% des données. Cette validation croisée a permis d'identifier le meilleur modèle de SG qui a ensuite été appliqué pour réaliser les prédictions.

#### Calculs des GEBV<sub>2014</sub> et des GEBV<sub>2014+2015</sub> à l'aide du meilleur modèle défini sur la population cible n°1

Une première GEBV dite GEBV<sub>2014</sub> a été calculée à partir de toutes les données de phénotypage disponibles sur la population d'entraînement R2N avant la récolte 2014 (récolte 2014 incluse) puis une seconde GEBV dite GEBV<sub>2014+2015</sub> a été calculée à partir de toutes les données de phénotypage disponibles avant la récolte 2015 (récolte 2015 incluse) sur la population cible n°1.

#### Application d'une sélection phénotypique classique sur la population cible n°1

Les sélectionneurs de R2n ont appliqué une sélection classique sur la première population cible. Les lignées trop sensibles, pas assez productives ou n'ayant pas été appréciées par les sélectionneurs, ont été éliminées du programme de sélection. A l'inverse, les lignées retenant leur attention ont été conservées et à nouveau phénotypées sur 5 lieux en 2 répétitions en 2016 pour le rendement traité.

#### Application du meilleur modèle de SG de la population cible n°1 sur la population cible n°2

Le meilleur des 5 modèles testés sur la population cible n°1 a été retenu pour être appliqué sur la population cible n°2.

#### Calculs des GEBV<sub>2014</sub> et des GEBV<sub>2014+2015</sub> à l'aide du meilleur modèle sur la population cible n°2

Une première GEBV dite GEBV<sub>2014</sub> a été calculée à partir de toutes les données de phénotypage disponibles sur la population d'entraînement R2N avant la récolte 2014 (récolte 2014 incluse) puis une seconde GEBV dite GEBV<sub>2014+2015</sub> a été calculée à partir de toutes les données de phénotypage disponibles avant la récolte 2015 (récolte 2015 incluse).

#### Validation de l'utilisation d'un modèle de SG rendement dans un programme de sélection blé tendre

#### Etude des corrélations entre GEBVs et rendements traités

Les coefficients de détermination ont été calculés entre GEBVs et moyenne des rendements traités 2015 pour les populations cibles n°1 et n°2, entre GEBVs et moyenne des rendements traités 2016 pour la population cible n°2, entre GEBVs et moyenne des rendements traités 2017 pour les 30 individus de la population cible n°2 ayant été poursuivis l'année suivante.

#### Comparaison de l'efficacité de la sélection génomique et de la sélection phénotypique ou des 2 combinées

Pour la population cible n°1, les pools de lignées suivants ont été créés afin de pouvoir valider ou invalider l'utilisation d'un outil de SG rendement dans un programme de sélection :

- Un pool « Bottom 50 GS- » correspondant aux lignées possédant une valeur de rendement 2015 intermédiaire et une valeur de GEBV<sub>2014</sub> se situant dans les 50 GEBVs les plus mauvaises.
- Un pool « Bottom 50 PS- » correspondant aux lignées possédant une valeur de GEBV<sub>2014</sub> intermédiaire et se situant dans les 50 plus mauvaises lignées pour le rendement traité en 2015 qui ont été sélectionnées pour passer dans l'année supérieure.
- Un pool « Top 50 GS+ » correspondant aux lignées possédant une valeur de rendement 2015 intermédiaire et une valeur de GEBV<sub>2014</sub> se situant dans les 50 GEBVs les plus élevées.
- Un pool « Top 50 PS+ » correspondant aux lignées possédant une valeur de GEBV<sub>2014</sub> intermédiaire et se situant dans les 50 meilleures lignées pour le rendement traité en 2015 qui ont été sélectionnées pour passer dans l'année supérieure.
- Un pool « Top 50 PS+ and Top 50 GS+ » correspondant aux lignées se situant à la fois dans les 50 meilleures lignées en terme de GEBV<sub>2014</sub> et dans les 50 meilleures lignées en terme de rendement traité 2015.
- Un pool « Bottom 50 PS- and 50 GS- » correspondant aux lignées se situant à la fois dans les 50 plus mauvaises lignées en terme de GEBV<sub>2014</sub> et dans les 50 moins bonnes lignées en terme de rendement traité 2015 ayant été sélectionnées pour passer dans l'année supérieure.
- Un pool « Bottom 50 PS- and 50 GS+ » correspondant aux

lignées se situant à la fois dans les 50 meilleures lignées en terme de GEBV<sub>2014</sub> et dans les 50 moins bonnes lignées en terme de rendement traité 2015 ayant été sélectionnées pour passer dans l'année supérieure.

- Un pool « Top 50 PS+ and Bottom 50 GS- » correspondant aux lignées se situant à la fois dans les 50 moins bonnes lignées en terme de GEBV<sub>2014</sub> et dans les 50 meilleures lignées en terme de rendement traité 2015.

Pour la population cible n°2, les pools de lignées suivants ont été créés afin de pouvoir valider ou invalider l'utilisation d'un outil de SG rendement dans un programme de sélection :

- Un pool « GS- » correspondant aux lignées possédant une valeur de rendement 2015 intermédiaire et une valeur de GEBV<sub>2014+2015</sub> se situant dans les 100 GEBVs les plus mauvaises.
- Un pool « PS- » correspondant aux lignées possédant une valeur de GEBV<sub>2014+2015</sub> intermédiaire et se situant dans les 100 plus mauvaises lignées pour le rendement traité en 2015 qui ont été sélectionnées pour passer dans l'année supérieure.
- Un pool « GS+ » correspond aux lignées possédant une valeur de rendement 2015 intermédiaire et une valeur de GEBV<sub>2014+2015</sub> se situant dans les 100 GEBVs les plus élevées.
- Un pool « PS+ » correspondant aux lignées possédant une valeur de GEBV<sub>2014+2015</sub> intermédiaire et se situant dans les 100 meilleures lignées pour le rendement traité en 2015 qui ont été sélectionnées pour passer dans l'année supérieure.
- Un pool « PS+ et GS+ » correspondant aux lignées se situant à la fois dans les 100 meilleures lignées en terme de GEBV<sub>2014+2015</sub> et dans les 100 meilleures lignées en terme de rendement traité 2015.
- Un pool « PS- et GS- » correspondant aux lignées se situant à la fois dans les 100 plus mauvaises lignées en terme de GEBV<sub>2014+2015</sub> et dans les 100 moins bonnes lignées en terme de rendement traité 2015 ayant été sélectionnées pour passer dans l'année supérieure.
- Un pool « PS- et GS+ » correspondant aux lignées se situant à la fois dans les 100 meilleures lignées en terme de GEBV<sub>2014+2015</sub> et dans les 100 moins bonnes lignées en terme de rendement traité 2015 ayant été sélectionnées pour passer dans l'année supérieure.
- Un pool « PS+ et GS- » correspondant aux lignées se situant à la fois dans les 100 moins bonnes lignées en terme de GEBV<sub>2014+2015</sub> et dans les 100 meilleures lignées en terme de rendement traité 2015.
- Un pool « Empty » correspondant aux lignées présentant un comportement intermédiaire à la fois pour la GEBV<sub>2014+2015</sub> et pour le rendement traité 2015, pool par rapport auquel il sera intéressant de comparer tous les autres pools pour conclure sur l'efficacité de l'une ou l'autre des 2 méthodes de sélection.
- 2.2 Prise en compte des effets environnementaux et des interactions GxE dans les modèles de SG

#### Caractérisation des sites expérimentaux et calcul de co-variables environnementales

Sur chacun des sites où les populations cibles et d'entraînement ont été implantées, des analyses de sol ont été réalisées afin de connaître la composition du sol (%limon, %argile, % sable, % de Matière Organique, taux de cailloux).

Des indicateurs pédoclimatiques ont été calculés pour les différentes

lignées utilisées dans ce projet à l'aide du modèle écophysiologique développé par ARVALIS Institut du végétal. Dans un premier temps, un rattachement des lignées à une variété paramétrée pour le modèle a été effectué. Pour cela, la date d'épiaison moyenne de chaque lignée a été calculée puis chaque lignée a été rattachée au témoin présentant la date d'épiaison la plus proche.

Enfin, des calculs d'indicateurs (somme de températures moyennes, nombre de jours de gel, avec des températures échaudantes ou un déficit hydrique...) ont été réalisés pour différentes périodes de développement. Ces indicateurs ont été utilisés comme co-variables dans les modèles de GS incorporant un terme d'interaction GxE (Jarquin *et al.* 2014).

#### Introduction d'interaction GxE dans nos modèles de sélection génomique

Le modèle G-BLUP, classiquement utilisé en sélection génomique (MA), sert à évaluer la valeur génétique additive d'une lignée. Lorsque l'on souhaite estimer la valeur génétique totale d'une lignée dans un environnement particulier, la valeur génétique additive n'est alors pas suffisante car d'autres termes interviennent : les déviations dues aux effets d'épistasie ; les interactions avec l'environnement (GxE). Si l'extension du modèle classique (MA) à la prise en compte de l'épistasie (MAI) est directement possible via l'utilisation des marqueurs, l'extension à la prise en compte des interactions GxE (MAIE) nécessite la caractérisation de nos lieux d'essais par des co-variables environnementales.

En utilisant les co-variables décrites précédemment, nous nous sommes intéressés à la capacité de prédiction des modèles MA, MAI et MAIE pour prédire : a) la valeur génétique moyenne totale des lignées ; b) la valeur génétique totale des lignées dans un environnement spécifique. Une validation croisée sur la population d'entraînement R2n avec 50 répétitions a été réalisée pour évaluer la capacité de prédiction des différents modèles, en supprimant aléatoirement 15% des lignées à chaque répétition et en prédisant leur valeur génétique pour le a) ou leur valeur génétique dans un environnement spécifique pour le b).

#### 2.3 - Calibration de modèles de SG spécifiques à différents scénarios environnementaux

#### Données de phénotypage et de génotypage

Les données de phénotypage et de génotypage (TaBW420K ; Rimbert *et al.* 2018) pour un panel de 426 variétés de blé tendre testées dans 46 environnements (lieu × année × conduite) en France ont été utilisées pour évaluer l'intérêt de calibrer des modèles de SG spécifiques à une population d'environnements donnée. Ces essais visaient à évaluer la réponse des variétés à des conditions climatiques et agronomiques variées (fertilisation azotée optimale versus limité, condition irriguée versus pluviale). Les données de génotypage ont été analysées afin d'éliminer les marqueurs génétiques présentant un taux de données manquantes > 20%, un taux d'hétérozygotie > 10% et une fréquence de l'allèle minoritaire < 5%. La matrice de génotypage ainsi obtenue comprenait 170 000 marqueurs SNP.

#### Caractérisation et regroupement des essais

Le modèle de culture CHN développé par Arvalis a été utilisé afin de calculer des indices de stress azoté et hydrique dans chaque essai en tenant compte des données climatiques, de l'itinéraire technique (date et densité de semis, apport d'azote, irrigation) et des caractéristiques du sol (texture, volume exploitable par les racines). L'indice de stress azoté est en lien direct avec l'INN tandis que l'indice de stress hydrique dépend de la fraction d'eau transpirable du sol. Ces deux indices journaliers varient entre zéro (stress maximal) et un (absence de stress) et réduisent la production de biomasse et l'augmentation de la surface foliaire.

L'ensemble des simulations ont été réalisées pour la variété Soissons présente dans l'ensemble des essais. Les indices de stress azoté et hydrique journaliers ont été centrés sur la date d'épiaison et des valeurs moyennes ont été calculées sur des fenêtres de 100°C jours de -1000 à +600°C jours autour de la date d'épiaison. La dynamique de ces indices a ensuite été analysée et les essais ont été regroupés en k groupes correspondant à différents scénarios environnementaux à l'aide d'un algorithme des k-médianes selon la méthode proposée par Chenu *et al.* (2011).

#### Modèle de prédiction

Une matrice d'apparentement a été calculée à partir de la matrice de génotypage avec le package rrBLUP (Endelman, 2011) dans R (R Core Team, 2016). Le modèle de prédiction utilisé correspondait au modèle 5 de Jarquin *et al.* (2014) et peut être formulé ainsi :

$$Y_{ij} = \mu + w_{ij} + g_i + gw_{ij} + \varepsilon_{ij}$$

Avec  $Y_{ij}$  le rendement en grains du génotype i dans l'essai j,  $w_{ij}$  l'effet des co-variables environnementales pour le génotype i dans l'essai j,  $g_i$  l'effet du génotype i, et  $gw_{ij}$  l'interaction entre le génotype i et les co-variables environnementales dans l'essai j et  $\varepsilon_{ij}$  l'erreur résiduelle. Dans ce modèle, les termes w, g, gw,  $\varepsilon$  suivent une loi normale et sont indépendants. En revanche, seuls les résidus sont distribués de manière aléatoire et les autres termes présentent une structure dérivée de matrices de similarités. Ces dernières permettent de prédire les performances des variétés dans des environnements non testés.

#### Evaluation des capacités prédictives par validation croisée

Une procédure de validation croisée de type "leave-one-out" a été utilisée afin d'estimer les capacités prédictives du modèle calibré sur différents jeux de données pour prédire les performances des variétés dans des environnements nontestés. Trois jeux de données ont été considérés :

- Le set 1 correspond à l'ensemble des essais excepté l'essai j appartenant au groupe k
- Le set 2 correspond à l'ensemble des essais excepté les essais appartenant au groupe k
- Le set 3 correspond à l'ensemble des essais appartenant au groupe k excepté l'essai j

Dans chaque cas, les rendements en grains prédits pour les variétés présentes dans l'essai j ont été calculés puis la corrélation entre rendements prédits et observés a été estimée. Cette procédure a été répétée pour les essais de 1 à j. Enfin, les coefficients de corrélations obtenus à l'aide du modèle calibré sur ces trois jeux de données ont été comparés pour chaque groupe d'essais.

Dans ce cadre, le set 1 correspond à la situation où aucune caractérisation environnementale et aucun regroupement d'essai n'a été effectué et où l'ensemble des données disponibles sont utilisées pour calibrer le modèle de SG. Le set 2 correspond à la situation où la base de données de calibration ne contient pas d'essai correspondant au scénario environnemental k (par exemple, absence d'essai en condition de stress hydrique ou de fertilisation azotée limitante). Le set 3 quant à lui représente la situation où les performances des variétés dans l'essai k sont prédites à l'aide d'un modèle calibré spécifiquement pour le scénario environnemental k. Les calibrations ont été réalisées à l'aide du package BGLR (Perez et De los Campos, 2014) dans R.

## 3. Résultats et Discussion

3.1 - Application d'un outil de SG rendement dans un programme de sélection de blé tendre

#### Génotypage

Après suppression des individus avec plus de 20% de données manquantes : 246 individus de la population d'entraînement R2N ont été génotypés à l'aide de 19 740 SNPs polymorphes de la puce 35 K, 481 individus de la population cible n°1 ont été génotypés à l'aide de 19 740 SNPs polymorphes de la puce 35 K et 479 individus de la population cible n°2 ont été génotypés à l'aide de 18 705 SNPs polymorphes de la puce 35 K.

Après suppression des marqueurs ayant une MAF trop faible ou trop de données manquantes, 13 793 SNPs ont été considérés pour les analyses de SG.

#### Test de différents modèles de sélection génomique sur la population cible n°1

Pour les 5 modèles testés, les coefficients de détermination entre GEBVs et moyennes ajustées par validation croisée se sont avérés relativement faibles ( $0.14 < r^2 < 0.19$ ). C'est néanmoins le modèle avec le  $r^2$  maximal qui a été retenu pour la suite des analyses, c'est-à-dire un modèle G-BLUP prenant en compte à la fois les effets d'épistasie et les effets d'additivité.

#### Validation de l'utilité d'utiliser un modèle de sélection génomique rendement dans un programme de sélection blé tendre

#### Etude des corrélations entre GEBVs et rendements traités

#### Sur la population cible n°1

Le graphe ci-dessous montre la relation entre GEBV<sub>2014</sub> (en abscisse) et rendements obtenus en 2015 (en ordonnée) pour l'ensemble des individus de la population cible n°1.



Le r<sup>2</sup> entre GEBV<sub>2014</sub> et le rendement mesuré en 2015 est alors nul (r<sup>2</sup> = 0.001).

Les prédictions par SG intra-croisement ont ensuite été étudiées. Pour le premier croisement, composé de 41 lignées, le r<sup>2</sup> entre GEBV<sub>2014</sub> et le rendement mesuré 2015 est également très bas avec une valeur de 0.008. Pour le second croisement composé de 46 lignées, le r<sup>2</sup> chute alors à 0.

Avec ce modèle et les données de phénotypage disponibles après la récolte 2014, il est donc impossible de prédire le potentiel de rendement traité d'une lignée que ce soit aussi bien au sein d'un croisement ou entre lignées issues de plusieurs croisements différents.

Dans un second temps, les valeurs de rendement obtenues en

2015 et les GEBV\_{2014+2015} sur l'ensemble de la population cible n°1 ont été comparées.

Le graphe ci-dessous montre la relation entre GEBV<sub>2014+2015</sub> (en abscisse) et rendements obtenus en 2015 (en ordonnée) pour l'ensemble des individus de la population cible  $n^{\circ}1$ .



Le r<sup>2</sup> entre GEBV<sub>2014+2015</sub> et le rendement mesuré en 2015 est alors de 0.003.

Pour le croisement contenant 41 lignées, le r<sup>2</sup> s'améliore de manière significative puisqu'il atteint alors une valeur de 0.101. Pour le croisement contenant 46 lignées, le r<sup>2</sup> reste trop faible (r<sup>2</sup> = 0.024). Une année de phénotypage supplémentaire de la population d'entraînement n'a donc pas permis l'amélioration de notre modèle de prédiction mise à part pour les prédictions intra croisement de celui composé de 41 lignées.

Les individus de la population cible n°1 ont été phénotypés pour le rendement traité que sur une seule parcelle d'expérimentation en 2015, et donc nécessairement que dans un seul environnement. Cela peut expliquer la difficulté à prédire le rendement par SG.

#### Sur la population cible n°2

Dans un premier temps, les valeurs de rendement obtenues en 2015 et 2016 ont été comparées aux GEBV<sub>2014</sub> sur l'ensemble de la population cible n°2.

Dans tous les graphiques ci-dessous, les individus de la  $2^{eme}$  population cible seront représentés par 3 couleurs différentes. En bleu, les individus qui ont été choisis pour passer au stade Fn+1 en raison de leur faible GEBV<sub>2014</sub> pour le rendement indépendamment de toute sélection phénotypique (pool GS-); en vert, les individus qui ont été choisis pour passer au stade Fn+1 en raison de leur bonne GEBV<sub>2014</sub> pour le rendement indépendamment de toute sélection phénotypique (pool GS+) et en violet les 328 individus restant de la population cible n°2 qui ont été choisis pour passer au stade Fn+1 uniquement sur la base de leur phénotypage.

La corrélation entre GEBV<sub>2014</sub> et rendement 2015 est alors du même niveau que celle obtenue sur la 1<sup>ère</sup> population cible. Le R<sup>2</sup> est de 0.006 (graphe ci-dessous).



Une fois encore cette absence de prédiction peut s'expliquer par le fait que la population cible n°2 a été phénotypée sur une seule parcelle d'expérimentation en 2015 et donc nécessairement que dans un seul environnement.

Comme le montre ensuite le graphe ci-dessous, la corrélation entre GEBV<sub>2014</sub> et rendement 2016 augmente significativement par rapport à celle obtenue entre GEBV<sub>2014</sub> avec les rendements 2015 tout en demeurant assez faible ( $r^2 = 0.034$ ).





Le graphe ci-dessus, montre une différence significative (p =  $1.39 \times 10^{-7}$ ) entre les 3 pools. Le pool vert « GS+ » est significativement meilleur que le pool violet « Fn+1 » qui est lui-même significativement meilleur que le pool GS-.

Malgré l'absence de prédiction de la SG sur le rendement 2016, cette figure montre qu'il aurait été plus performant d'utiliser les valeurs de GEBV<sub>2014</sub> plutôt que les valeurs de rendement 2015 (pour rappel : 1 seule parcelle d'expérimentation) pour sélectionner le matériel à phénotyper en 2016.

Le graphe ci-dessous montre la corrélation entre GEBV<sub>2014+2015</sub>. La corrélation est bonne (r<sup>2</sup> = 0.571) mais grâce au jeu de couleur nous pouvons tout de même voir que certains individus à faible GEBV<sub>2014</sub> ne sont plus si mauvais en GEBV<sub>2014+2015</sub> et inversement. Il est donc intéressant de voir quelles sont les nouvelles corrélations entre GEBV<sub>2014+2015</sub> et rendement 2015 puis rendement 2016.



Les valeurs de rendement obtenues en 2015, 2016 et 2017 ont

donc été comparées aux GEBV $_{2014+2015}$  sur l'ensemble de la population cible n°2.

Pour l'année 2015, la corrélation entre GEBV<sub>2014+2015</sub> et rendement est quasi-nulle ( $r^2 = 0.007$ ) comme le montre le graphe ci-dessous.



Malgré le phénotypage supplémentaire réalisé en 2015 sur la population d'entraînement R2n, les prédictions ne s'améliorent pas. Une fois encore cette absence de prédiction peut s'expliquer par le fait que la population cible n°2 a été phénotypée sur une seule parcelle d'expérimentation en 2015.

Pour l'année 2016, la corrélation entre GEBV<sub>2014+2015</sub> et rendement est assez élevée ( $r^2 = 0.290$ ) comme le montre le graphe ci-dessous.



Dans ce cas, la multiplication des environnements dans lesquels la population cible n°2 a été testée et l'ajout d'environnements supplémentaires dans lesquels la population d'entraînement R2n a été phénotypée a un impact très intéressant sur les qualités de prédiction et le r<sup>2</sup> obtenu est largement satisfaisant.

En ne s'intéressant qu'aux individus aux comportements extrêmes identifiés sur la base des GEBV<sub>2014+2015</sub>, on peut alors obtenir un r<sup>2</sup> relativement fiable (r<sup>2</sup> = 0.651).



Ces résultats confirment donc la possibilité de pouvoir identifier les individus les plus prometteurs sur la base de leur GEBV et de pouvoir éliminer les individus les plus mauvais sur la base de leur GEBV sans prendre de risque.

Le graphe ci-dessous montre la corrélation entre GEBV<sub>2014+2015</sub> et rendement traité 2017 pour 30 lignées de la population cible conservées en 2017.



Contrairement à 2016, le r<sup>2</sup> entre rendement traité 2017 et GEBV<sub>2014+2015</sub> est relativement faible (r<sup>2</sup> = 0.029 en 2017 vs r<sup>2</sup> = 0.306 en 2016). Le faible effectif (30 individus en 2017 vs 482 en 2016) peut expliquer une variation de cette importance.

Le tableau ci-dessous montre les corrélations entre année pour le rendement traité pour ces 30 lignées.

	YLD_%TT_2015	YLD_%TT_2016	YLD_%TT_2017
YLD_%TT_2015	1		
YLD_%TT_2016	- 0.113015425	1	
YLD_%TT_2017	0.090454022	0.097875168	1

Il n'existe donc aucune corrélation entre les 3 années où des données de rendement ont pu être collectées.

L'effet année est donc très important et le classement variétal est complètement chamboulé d'une année sur l'autre comme le montre le graphe ci-dessous.



L'année 2016 a été impactée par un effet microdochium très important, cela peut donc expliquer les différences de comportement variétal observées. Par contre, l'absence de corrélation entre 2015 et 2017 ne semble pouvoir être expliquée que par le faible nombre d'individus et par le fait que le matériel avait été phénotypé sur un seul lieu en 2015 vs 13 en 2017.

Comparaison de l'efficacité de la sélection génomique vs celle de la sélection phénotypique ou des 2 combinées

#### Sur la population cible n°1

Le graphe ci-dessous montre l'impact de l'appartenance des lignées de la population cible n°1 aux différents pools sur le rendement traité 2016.



Le nombre de lignées par pool est dans l'ensemble assez faible. En raison d'un effectif trop faible, nous ne pourrons pas conclure sur les 2 pools marron et bleu ciel : « Top 50 PS+ and Bottom 50 GS - » et « Top 50 PS+ and Top 50 GS+ ». Par contre, pour les autres pools une légère tendance se dessine. Le pool orange « Top 50 GS+ » et le pool gris « Top 50 PS+ » ont des médianes équivalentes à 104% des témoins et des moyennes assez proches avec un léger avantage pour le pool « Top 50 PS+ » (103.9% vs 102.7%). Le pool jaune « Bottom 50 PS- and 50 GS -» arrive en dernière position à 92% des témoins en moyenne. Le pool rose « Bottom 50 GS- » donne une large gamme de réponse sur le rendement traité 2016. Ce pool est en effet un des plus mauvais en terme de moyenne (97.5% des témoins) pour le rendement 2016 mais un certain nombre de lignées performe très bien. Sur les 30 lignées appartenant à ce pool, 5 d'entre elles ont été gardées pour passer dans l'année supérieure indépendamment de leur faible GEBV mais elles ont des valeurs de rendement 2016 toutes inférieures à 104% des témoins et même une valeur de 81% des témoins pour la plus mauvaise. Les 4 lignées au dessus de 105% des témoins sont particulièrement intéressantes et il aurait été dommage de les éliminer sur la base des résultats obtenus par SG. Elles n'auraient néanmoins pas été sélectionnées non plus sur la base du phénotypage. Il n'aurait donc pas été commis d'erreur trop importante en éliminant les individus aux GEBV<sub>2014</sub> les plus faibles.

Enfin, le pool bleu foncé « Bottom 50 PS- and 50 GS+ » est lui aussi assez informatif. Une lignée se détache vers le bas à 84% des témoins et plusieurs individus se détachent au-dessus de 105% des témoins. Sans la sélection génomique, la totalité de ces lignées auraient été éliminée alors que certaines se révèlent intéressantes en termes de potentiel de rendement traité. Se servir des 2 méthodes de sélection aurait donc été un plus puisque ce matériel n'aurait pas été semé en prenant compte des résultats de SG pour ces lignées.

Sur la base de ces résultats, appliquer une stratégie de sélection génomique aurait été plus performant qu'appliquer une simple stratégie de sélection phénotypique dans un seul environnement.

#### Sur la population cible n°2

Le graphe ci-dessous montre l'impact de l'appartenance des lignées de la population cible n°2 aux différents pools sur le rendement traité 2016.



Le nombre de lignées par pool est dans l'ensemble satisfaisant afin de pouvoir comparer les pools les uns par rapport aux autres.

Trois pools se distinguent pour leurs bons résultats et permettent de mettre en évidence l'efficacité de la sélection génomique pour distinguer les individus les plus prometteurs. Les pools jaune « GS+ » (moyenne à 104.3% des témoins), violet « GS+ et PS-» (moyenne à 104.4% des témoins) et orange « GS+ et PS+ » (moyenne à 105.5% des témoins) se distinguent de tous les autres et sont significativement bien meilleurs que le pool gris « Empty » de référence (moyenne à 100.2% des témoins). En terme de moyenne, le pool « GS+ et PS+ » est légèrement meilleur que les 2 autres, ceci confirme l'intérêt de travailler en combinant à la fois sélection phénotypique et SG.

Deux pools sont équivalents au pool de référence gris « Empty ». Il s'agit des pools bleu clair « PS- » et vert clair « PS+ » avec des moyennes proches de 99-100% des témoins. Ces résultats mettent une fois de plus en évidence la faible corrélation obtenue entre rendements 2015 et 2016 et donc un fort effet année.

Enfin, 3 pools se distinguent pour leurs faibles moyennes de rendement traité en 2016 et permettent de mettre en évidence l'efficacité de la sélection génomique pour distinguer les individus les moins prometteurs. Il s'agit des pools bleu foncé « GS- » (moyenne à 96.3% des témoins), vert foncé « GS- et PS- » (moyenne à 92.2% des témoins) et rouge « GS- et PS+ » (moyenne à 96.6% des témoins). Le pool vert foncé « GS- et PS- » est significativement plus mauvais que les 2 autres pools. Une fois encore, combiner SG et sélection phénotypique présente donc un intérêt.

Sur la base de ces résultats, appliquer une stratégie de sélection génomique aurait été plus performant qu'appliquer une simple stratégie de sélection phénotypique. Combiner Sélection génomique et Sélection phénotypique semble être un très bon compromis pour identifier les individus les plus prometteurs et pour éliminer ceux qui le sont le moins.

#### 3.2 - Prise en compte des effets environnementaux et des interactions GxE dans les modèles de SG

Le tableau ci-dessous montre les composantes de la variance et de l'héritabilité pour chacun des modèles testés.

Modèle	Additif	Epistasie	Genetic	GxEC	GxE	Résiduelle	r	r_env	H <sup>2</sup>
MA	10.49	NA	11.47	NA	11.09	63.92	0.13	0.26	0.6
MAI	6.08	3.61	10.1	NA	10.88	63.75	0.12	0.25	0.58
MAIE	4.93	3.33	8.26	19.84	0.14	59.38	0.09	0.32	0.5

Les résultats des variances estimées par les différents modèles montrent une répétabilité du caractère étudié (à savoir le rendement) autour de 0.10, suggérant une corrélation au carré entre une seule observation phénotypique (1 plot) et la vraie valeur génétique totale (qui est inconnue) de 0.1. De la même manière, on trouve une répétabilité dans un environnement donné de 0.25 suggérant une corrélation au carré entre une seule observation phénotypique et la valeur génétique totale dans l'environnement testé de 0.25. Enfin, l'héritabilité estimée au sens large du dispositif est de l'ordre de 0.55 suggérant une corrélation au carré entre la moyenne des observations phénotypiques d'une lignée et la vraie valeur génétique totale de 0.55.

Concernant les différences entre les modèles, le modèle MAIE, qui intègre les co-variables environnementales, capture mieux les effets GxE que les modèles MA et MAI (une résiduelle plus faible et un effet GxE total plus important).

En termes de prédiction des valeurs génétiques moyennes, cas du a), il semble que l'ajout d'épistasie dans le modèle améliore les prédictions mais pas l'ajout du GxE comme le montre la figure ci-dessous.



Ce qui semble normal même si on aurait pu imaginer que l'ajout du GxE dans le modèle entraine une non-orthogonalité des effets et donc une précision réduite vis-à-vis de la valeur génétique totale. A noter qu'une prédiction d'une lignée non phénotypée mais génotypée avec le meilleur modèle, MAI, donne une précision équivalente à un phénotypage seul dans le dispositif actuel, soit l'héritabilité (le résultat de MAI est de 0.3, qui divisée par H<sup>2</sup> donne 0.3 / 0.55 = 0.55, équivalent à H<sup>2</sup>).

En ce qui concerne les prédictions des valeurs génétiques dans un lieu donné, cas du b), il semble que l'introduction des covariables environnementales dans le modèle MAIE aient été bénéfiques comme le montre la figure ci-dessous.



Ce résultat signifie que la prédiction d'une lignée génotypée mais pas phénotypée avec MAIE dans un lieu donné est supérieure à une évaluation purement phénotypique dans ce même lieu. Cela vient du fait que la corrélation au carré entre la valeur génétique prédite dans un lieu donné et la vraie valeur dans ce même lieu est égale à (0.31<sup>2</sup>) divisé par la répétabilité dans un lieu (r\_env = 0.25, qui est la précision de la valeur génétique dans un lieu donné), ce qui donne ~0.1/0.25 = 0.4 > r\_env. En d'autres termes, une prédiction génomique est plus précise qu'une observation phénotypique seule dans un lieu donné.



Dynamique des indices de stress hydrique et azoté (0 = stress maximal, 1 = pas de stress) calculés à l'aide d'un modèle de culture pour la variété Soissons dans 42 essais regroupés en 4 groupes (N1W1, absence de stress azoté ou hydrique ; N2W1, essais stress azoté ; N1W2, essais stress hydrique ; N2W2, essais combinant stress hydrique et stress azoté). Les lignes tiretées représentent la dynamique des indices de stress pour les différents essais. Les lignes pleines représentent les médianes des indices de stress pour chaque groupe d'essai.

#### 3.3 - Calibration de modèles de SG spécifiques à différents scénarios environnementaux

#### Regroupement des essais

La méthode appliquée a permis de regrouper les essais en quatre groupes correspondant à quatre scénarios environnementaux :

- N1W1 : absence de stress hydrique ou azoté
- N1W2 : absence de stress azoté / stress hydrique de fin de cycle
- N2W1 : stress azoté au cours du cycle / absence de stress hydrique
- N2W2 : présence de stress hydrique et azoté

Le groupe N1W1 comprenait 28 essais, dont 17 essais en condition de fertilisation optimale, deux essais en condition de fertilisation limitante, quatre essais en condition pluviale et cinq essais « date de semis tardive ». Le groupe N1W2 comprenait 11 essais, dont six essais en condition de fertilisation optimale, un essai en condition de fertilisation limitante, un essai en condition de fertilisation pluviale, deux essais « date de semis tardive » et un essai « densité de semis réduite ». Le groupe N2W1 comprenait trois essais en condition de fertilisation limitante. Enfin, le groupe N2W2 comprenait quatre essais en condition de fertilisation limitante et un essai « date de semis tardive ».

#### Comparaison des capacités prédictives

L'analyse des capacités prédictives des modèles calibrés sur différent jeux de données montre que l'utilisation de modèles spécifiquement calibrés pour un scénario environnemental donné semble permettre un léger gain en termes de prédiction. C'est le cas pour trois des quatre groupes d'environnements testés avec N1W2 (r médian égal à 0.55 avec le set 3 contre 0.50 avec le set 1), N2W1 (r médian égal à 0.76 avec le set 3 contre 0.67 avec le set 1) et N2W2 (r médian égal à 0.56 avec le set 3 contre 0.49 avec le set 1). En revanche, pour le groupe N1W1, la calibration sur des données spécifiques à ce groupe conduit à une diminution des capacités prédictives (r médian égal à 0.51 avec le set 3 contre 0.57 avec le set 1). Il est à noter que ce groupe est celui qui contient le plus d'essais et compte tenu de la nature quantitative des dynamiques de stress, la classification des essais dans ce groupe (comme dans les autres) reste relativement sujette à imprécision. Comme attendu, les capacités prédictives des modèles calibrés sur le set 2 sont systématiquement plus faibles que celles obtenues avec les modèles calibrés sur les sets 1 et 3 (r médian égal à 0.44 et 0.30 pour les groupes N1W1 et N1W2 avec le set 2 contre 0.57 et 0.50 pour le set 1). Néanmoins les différences sont parfois relativement faibles (r médian égal à 0.67 et 0.51 pour les groupes N2W1 et N2W2 avec le set 2 contre 0.67 et 0.49 pour le set 1).

Même si les différences entre le set 1 et le set 3 sont relativement faibles, il est notable que leurs capacités prédictives ont été obtenues avec un nombre d'essais de calibration très différent (45 pour le set 1 et deux à 27 pour le set 3 suivant le scénario environnemental considéré). Il semble donc qu'il soit possible d'atteindre des capacités prédictives similaires pour un scénario environnemental donné avec un nombre d'essais restreint, à condition de caractériser les essais et d'inclure dans la base de données de prédiction des essais dans lesquels les conditions environnementales correspondent au scénario ciblé.



Boxplots des corrélations intra-essais après validation croisée de type « leave-one-out » entre valeurs prédites et observées pour le rendement en grains. Les prédictions ont été obtenues à l'aide d'un modèle intégrant un effet génotype (basé sur les données de génotypage), un effet environnement (basé sur l'utilisation de co-variables agroclimatiques) et un effet d'interaction GxE (interactions marqueurs × co-variables agroclimatiques). Ce modèle a été calibré en utilisant trois jeux de données différents : **set1** l'ensemble des essais sauf l'essai j (appartenant au groupe d'essais k) pour lequel les prédictions étaient réalisées, **set2** l'ensemble des essais sauf les essais appartenant au groupe k, **set3** l'ensemble des essais appartenant au groupe k sauf l'essai j.

## 4. Conclusions

Les résultats de prédiction obtenus varient énormément d'une année sur l'autre ( $0 < r^2 < 0.3$ ) et confirment les difficultés du sélectionneur à anticiper les effets année.

Quoi qu'il en soit, il ressort de ces analyses que le sélectionneur aurait moins de risque de se tromper en appliquant un modèle de sélection génomique rendement sur son matériel plutôt que de se fier à une seule parcelle rendement dans un environnement donné. De même, combiner à la fois les résultats obtenus par sélection génomique et par sélection phénotypique permettrait au sélectionneur d'éliminer sans risque les individus les plus mauvais et de mettre en évidence rapidement les plus prometteurs. Un modèle classique G-BLUP prenant en compte à la fois les effets d'additivité et d'épistasie ressort comme étant le modèle le plus efficace pour prédire la valeur génétique moyenne d'un individu de blé tendre. Introduire les interactions GxE dans ce modèle après calculs de co-variables environnementales permet d'améliorer considérablement les prédictions génétiques dans un environnement non-testé et est susceptible d'apporter des informations sur la stabilité des lignées et de complémenter les données de réseaux d'essais multi-environnementaux qui sont souvent déséquilibrés. Travailler par groupes d'environnements, après caractérisation fine de ces derniers, pourrait donc être une solution afin d'améliorer considérablement les modèles de sélection génomique rendement par prise en compte de des interactions GxE.

## Références bibliographiques

**Bernado, R., and J. Yu.** 2007. Prospects for genome-wide selection for quantative traits in maize. Crop Sci. 47:1082:1090. Doi:10.2135/cropsci2006.11.0690

**Burgueño, J., G. de los Campos, K. Weigel, and J. Crossa.** 2012. Genomic prediction of breeding values when modeling genotype x environment interaction using pedigree and dense molecular markers. Crop Sci. 52:707-719.

Chenu K., Cooper M., Hammer G.L., Mathews K.L., Dreccer M.F., Chapman S.C., 2011. Environment characterization as an aid to wheat improvement: interpreting genotype environment interactions by modelling water deficit patterns in north eastern Australia. Journal of Experimental Botany 62: 1743–1755.

**Comstock, R.E., and R.H. Moll.** 1963. Genotype-environment intersections. p. 164-196. In: W.D. Hanson and H.F. Robinson (ed.) Statistical genetics and plant breeding. NAS-NRC Publ. 982.

Crossa, J., G. de los Campos, P.Pérez, D. Gianola, S. Dreisigacker, J. Burgueño, J.L. Araus, D. Makumbi, J. Yan, R. Singh, V. Arief, M. Banziger, and H.-J. Braun. 2010. Prediction of genetic values of quantitative traits in plant breeding using pedigree and molecular markers. Genetics 186:713-724. Doi:10.1534/genetics.110.118521

Crossa, J., P. Pérez, G. de los Campos, G. Mahuku, S. Dreisigacker, and C. Magorokosho. 2011. Genomic selection and prediction in plant breeding. J. Crop. Improv. 25:239-261. Doi:10.1080/15427528.2011.558767

Crossa, J., P. Pérez, J. Hickey, J. Burgueño, L. Ornella, J. Cerón-Rojas, X. Zhang, S. Dreisigacker, R. Babu, Y. Li, D. Bonnett and K. Mathews. 2013. Genomic prediction in CIMMYT maize and wheat breeding programs. Heredity (2013), 1-13.

de los Campos, G., H. Naya, D. Gianola, J. Crossa, A. Legarra, E. Manfredi, K. Weigel and J.M. Cotes. 2009. Predicting quantitative traits with regression models for dense molecular markers and pedigree. Genetics 182:375-385. Doi10.1534/genetics.109.101501

de los Campos, G., D. Gianola, G.J.M. Rosa, K.A. Wiegel, and J. Crossa. 2010. Semi-parametric genomic-enabled prediction of genetic values using reproducing kernel Hilbert spaces methods. Genet. Res. 92:295-308. Doi10.1017/ S0016672310000285

**Endelman, J.B.** 2011. Ridge regression and other kernels for genomic selection with R package rrBLUP. Plant Genome 4:250-255. Doi:10.3835/plantgenome2011.08.0024

Jarquin, D., J. Crossa, X. Lacze, P. D. Chevron, J. Daucourt *et al.* 2014. A reaction norm model for genomic selection using high-dimensional genomic and environmental data. TAG 127:595-607.

Legarra, A., A. Ricard, O. Filangi. 2013. GS3 software. Available at: http://snp.toulouse.inra.fr/~ãlegarra/. Version 6. INRA, Toulouse.

Long, N., D. Gianola, G. J. M. Rosa, S. Weigel and S. Avendano. 2007. Machine learning classification procedure for selecting SNPs in genomic selection: application to early mortality in broilers. J. Anim. Breed. Genet. 124:377-389

Meuwissen, T.H.E., B.J. Hayes, and M. E. Goddard. 2001. Prediction of total genetic values using genome-wide dense marker maps. Genetics 157 :1819-1829.

**Pérez, P., G. de los Campos, J. Crossa and D. Gianola.** 2010. Genomic-enabled prediction based on molecular markers and pedigree using the Bayesian Linear Regression package in R. Plant genome 3(2):106-116.

**Pérez, P., G. de los Campos, J.** 2014. Genome-wide Regression and Prediction with the BGLR statistical package. Genetics 198 (2): 483-495.

**R Core Team.** 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.

Tier, B., J. Cavanagh, R. Crump, M. Khatar, G. Moster, 2007. Genome wide selection : experiences from the Australian Dairy Industry. The 3rd International Conference on Quantitative Genetics, August 2007, Hangzhou, China.

Woolaston, A.F., B. Tier and R.D. Murison. 2007. Principal Component regression of SNP to predict genetic merit. Papers and Abstracts From the Workshop on QTL and Marker-Assisted Selection, March 22-23, 2007, Toulouse, France, edited by A. Legarra.

## [HeatWheat] Analyse de la diversité génétique de la réponse au stress thermique via phénotypage fin et génétique d'association

Stéphane Lafarge \* 1, Gaetan Touzy 3, JC Deswarte 3, Katia Beauchene 3, Laurent Maunas 3, Christine Girousse 2, Jane Roche 4, Agnés Piquet 2, Said Mouzeyar 4, Jacques Le Gouis 2

1 - Biogemma, Centre de Recherche de Chappes, F-63720 Chappes.

2 - INRA-UBP, UMR1095 Genetic, Diversity & Ecophysiology of Cereals, F-63100 Clermont-Ferrand.

3 - ARVALIS-Institut du végétal, 3 rue Joseph et Marie Hackin F-75116 Paris

\* Coordinateur : Stéphane Lafarge, stephane.lafarge@Biogemma.com

## 1. Introduction

La stagnation du rendement constatée chez le blé depuis les dix dernières années est fortement corrélée à des épisodes climatiques. Lobell et al. (2011) ont chiffré l'impact du stress thermique à une perte mondiale de rendement de 5.5%, soit une perte totale de 35 M de tonnes. En Europe, des études ont été menées (Semenov et al. 2010 ; Gouache et al. 2012) ; l'analyse de l'évolution climatique ainsi que celle des rendements lors des vingt dernières années met en évidence une augmentation de l'impact du stress thermique en Europe sur les cultures de blé (Brisson et al. 2010). Les pertes de rendement en blé en Europe seront probablement plus importantes suite à des stress thermiques que suite à des déficits hydriques estivaux, qui seront évités du fait d'une maturation plus précoce des blés (Semenov et Shewry, 2011). Il reste cependant à clairement identifier les types de stress les plus courants en France en fonction de la spécificité des régions.

Ces données doivent orienter la recherche vers une compréhension des mécanismes de résistance et de sensibilité des variétés à ce type de stress, à une identification de matériel pertinent pour la sélection et d'outil moléculaire pour le suivi des allèles favorables dans le germplasm. Ce renforcement du germplasm français en allèles de tolérance aux futurs épisodes climatiques permettra d'améliorer la stabilité des rendements sous contraintes abiotiques et de répondre aux besoins de production en prenant en compte les nouvelles contraintes liées aux variations climatiques.

Sur le plan international, l'investissement sur la recherche d'une adaptation aux températures élevées du blé existe (Ortiz 2008 ; Reynolds 2007 ; Cossani 2012 et Reynolds 2012). Des études montrent que le blé est sujet fréquemment à des stress thermiques dans sa période de floraison et de remplissage du grain (Stone et Nicolas 1994). Durant la période reproductive ce genre de stress a un impact direct sur la fertilité et les composantes du rendement (Wollenweber et al. 2003). L'impact sur la capacité de fécondation est important si le stress se situe lors de la floraison (stérilité pollinique, avortement des grains etc...) (Tashiro et Wardlaw 1990). L'impact d'un stress lors du remplissage du grain se caractérise par un impact sur la volumétrie du grain (poids, taille) et sur sa composition (Calderini et al. 1999). Des traits physiologiques à travailler en sélection ont donc été identifiés pour une sélection spécifique de la résistance aux fortes températures, notamment lors du remplissage (Reynolds et al. 2007, Cossani et al. 2012), notamment la température de la canopée et la mesure de la concentration en sucres solubles. Des QTL pour certains de ces traits adaptatifs ont été identifiés (Suzuky Pinto et al. 2010). Des études physiologiques comparatives ont été menées sur différents génotypes présentant des réponses au stress différentes (Dhyani et al. 2013). Elles mettent en évidence des différences tant au niveau physiologique que métabolique. La mise en place d'un protocole de stress en conditions contrôlées et l'analyse d'un groupe de variétés sélectionnées permettrait d'accélérer l'identification de génotypes résistants au stress thermique et de ce fait, leur utilisation en sélection.

Des approches de génétique quantitative ont aussi été menées et des QTL ont pu être détectés (Suzuky-Pinto et al. 2010, Bennett et al. 2012, Rebetzke et al. 2013, Bonneau 2013). L'ensemble de ces études a permis de mettre en évidence des zones du génome liées à la résistance au stress thermique, notamment les zones situées sur les chromosomes 1B, 2B, 3B, 4A, 6B et 7A/B confirmant leur valeur générique sur ce type de contraintes.

Afin d'aller plus loin dans la compréhension de ces mécanismes et dans l'identification des facteurs génétiques impliqués dans la sensibilité/résistance au stress thermique, des ressources génomiques sont aussi disponibles comme des données d'expression sur le blé (Qin et al. 2008). Ceci a amené une meilleure compréhension sur les gènes potentiellement impliqués dans la réponse au stress thermique (Mittler et al. 2012) notamment avec l'implication de la famille des Heat Shock Factor (HSF) (Liu et al. 2013).

Au bilan, il existe dans la littérature de nombreux éléments sur l'impact du stress thermique chez blé. Cependant, les stress appliqués sont généralement des stress aigus et les variétés utilisées ne permettent pas d'imaginer une utilisation rapide en sélection en France. De ce fait, HeatWheat avait pour objectif d'identifier les sources potentielles de résistance au stress thermique dans le matériel à disposition du sélectionneur et aussi de créer des outils permettant une caractérisation variétale fine par utilisation de marqueurs moléculaires.

#### 2. Matériel et méthode

#### 2.1 - Bilan agro-climatologique et disponibilités des données historiques

Une étude préliminaire a été menée pour tenter d'identifier, dans des données expérimentales de plein champ, des situations où un stress thermique a eu lieu, afin de qualifier la séquence climatique responsable de la perte de rendement. Les données d'un essai de type « observatoire » ont donc été mobilisées. Les éléments à disposition étaient : données météorologiques journalières (températures min/max, précipitations, rayonnement), composantes de rendement (rendement grain, densité d'épis, PMG, indice de récolte), stades (épi 1 cm, épiaison, floraison) et caractérisation sommaire du sol. L'essai retenu sur la base d'une identification *a priori* par un agronome de terrain était conduit en 2005, sur le plateau du Neubourg (Eure), sur un sol de limons profonds. 4 variétés de précocités différentes étaient suivies pour leur phénologie et l'élaboration du rendement. La séquence climatique observée en 2005 en Haute-Normandie a ensuite été

comparée à d'autres situations pour valider le lien entre température et altération des composantes.

Une analyse agro-climatique nationale a ensuite été réalisée pour qualifier les conditions thermiques et hydriques dans lesquelles a lieu la phase de remplissage des grains du blé tendre en France. Pour cela, les données météorologiques de 232 stations météorologiques ont été exploitées ; ces stations présentaient des données sur la période 1994-2013 et étaient situées à une altitude inférieure à 500m. Pour décrire précisément les conditions climatiques impactant les cultures, des indicateurs phénoclimatiques ont été calculés : chaque station météorologique était rattachée à un cas-type (précocité variétale et date de semis), et un modèle phénologique était utilisé pour calculer des stades clés (stade épi 1 cm, floraison, maturité physiologique). Les indicateurs retenus sont résumés dans le tableau 1 ci-dessous. La simulation consistait à déterminer pour chaque station une valeur movenne par indicateur (pas de détermination de la variabilité interannuelle). Une analyse en composantes principales suivie d'une classification hiérarchique ascendante ont permis d'identifier des groupes de stations présentant les mêmes tendances phénoclimatiques. Pour chaque groupe, une station représentative a été identifiée, et fait l'objet d'analyses complémentaires (analyse de données horaires lors de séquences climatiques historiques).

Critère	Paramètre	Seuil	Période de calcul
Tmax	nombre	>25°C	floraison-maturité
Tmin	nombre	>15°C	floraison-maturité
Tmax	cumul	>25°C	floraison-grain laiteux
Tmax	cumul	>25°C	grain laiteux-maturité
Tmax	date du 1 <sup>er</sup> évè.	>25°C	épi 1cm-maturité
ETP	nombre	>5mm/j	floraison-maturité
ETP	nombre	>7mm/j	floraison-maturité
Tmax	nombre	>25°C	6 créneaux consécutifs de 150°Cj entre Z65-300°Cj et Z65+600°Cj
TMin	nombre	>15°C	6 créneaux consécutifs de 150°Cj entre Z65-300°Cj et Z65+600°Cj
ETP	cumul		floraison-maturité
P-ETP	cumul		épi 1cm-floraison
P-ETP	cumul		floraison-maturité

Tableau 1 : liste des indicateurs retenus

#### 2.2 - Analyse de la diversité génétique en conditions contrôlées

Le matériel génétique utilisé est le panel BreedWheat composé de 200 lignées de blé hiver élites inscrites dans la période 2000/2012. Celui-ci a été phénotypé en conditions controlées en différenciant une modalité de fortes températures (« heat stress ») à une modalité témoin, selon le protocole suivant : chaque plante devant subir un stress thermique était placée dans un compartiment stressant 3 jours après floraison pour une durée de 10j, puis était remise en condition optimale pour le reste du cycle.

Chaque génotype était représenté par 6 plantes en pot individuel avec 3 plantes en conditions optimales pour l'ensemble du cycle et 3 subissant le stress thermique comme indiqué ci-dessus. L'ensemble du dispositif représente 1200 plantes implantées en split-plot dans un compartiment de serre. Un second compartiment est utilisé pour la condition stressante.

En terme de phénotypage, pour chaque plante ont été mesurés notamment : la date de floraison du brin maitre, le contenu en chlorophylle par Dualex (Force-A, Orsay, France) (une fois par semaine avec cinq mesures sur la feuille drapeau). En fin d'expérimentation, la récolte a été effectuée sur le brin maitre et sur les talles de façon séparée et le nombre d'épis compté. Pour l'épi du brin maitre et ceux des talles, les grains récoltés ont été comptés et analysés pour déterminer les longueur/ largeur/ surface. Les grains pesés ont permis de calculer le PMG.

Deux années d'expérimentations ont été réalisées sur les sites de Montardon (Arvalis) et Chappes (Biogemma) permettant ainsi de générer un jeu de données phénotypiques pour la génétique d'association.



Figure 1 : Représentation des cycles en condition optimale et stress thermique

## 2.3 - Génétique d'association (GWAS)

Les données de génotypage utilisées sont ceux issus de la puce Affymetrix Axiom 420k produites dans le cadre du projet BreedWheat (ANR-2010-BTBR-03). Seuls les SNP polymorphes ont été utilisés dans cette analyse. Tous les SNP ont été physiquement cartographiés sur la séquence de référence du génome de Chinese Spring RefSeq V1.0 (IWGS et al. 2018). Les locus hétérozygotes ont été considérés comme des données manquantes. Les SNP monomorphes, les SNP avec plus de 10% de données manquantes et les SNP avec une fréquence d'allèle mineure (MAF) inférieure à 5% ont été écartés. Au final 219k SNPs ont été utilisés dans les modèles d'association décrits ci-dessous.

Les mêmes modèles expérimentaux ayant été utilisés dans les deux années d'expériences, nous avons donc décidé d'utiliser la même base pour modéliser l'interaction génétique par traitement (GxT). Le modèle de base a été adapté à l'aide d'un modèle mixte écrit en R utilisant le package ASReml-R (Butler et al. 2009). L'interaction GxT a été déclarée comme suit:

#### $Yijkl = \mu + Rj + Tk + Tk(Rj) + Bl(RjTk) + Gi + Gi \times Tk + \varepsilon ijkl$

Pour une analyse combinée des 2 années d'essais, le modèle suivant a été utilisé :

$$\begin{split} Yijkl = & \mu + Am + Tk + Tk \times Am + Am(Rj) + Tk(AmRj) + Bl(AmRjTk) \\ & + Gi + Gi \times Tk + Gi \times Am + \varepsilon ijkl \end{split}$$

#### 2.4 - Phénotypage physiologique et moléculaire

En complément du phénotypage effectué à l'échelle de la plante entière en serre, il a été proposé d'effectuer une expérimentation spécifique en chambre de culture, afin d'obtenir un phénotypage fin à l'échelle du grain avec des approches écophysiologique et transcriptomique. Ont donc été retenus 2 génotypes, contrastés pour leur tolérance aux températures élevées post-anthèse, de manière à identifier des mécanismes potentiels de tolérance/sensibilité au stress thermique lors du développement du grain et des gènes majeurs impliqués dans ces mécanismes.

Sur la base de l'évaluation de la diversité génétique réalisée en conditions de serre, deux génotypes contrastés pour leur réponse aux températures élevées post-anthèse ont été sélectionnés : Altigo, classé comme résistant et Attlass classé comme sensible. Ces deux génotypes présentent des dates de floraison similaires et des PMG en conditions témoins proches ; cette classification a été confirmée par les tests réalisés au stade plantule de CMT (Cell membrane thermostability).

Après semis et vernalisation (8 semaines à 5°C), 128 plantes par génotype ont été conduites jusqu'à floraison en conditions contrôlées (21/15°C jour/nuit ; 16/8h de photopériode ; arrosage non limitant par solution nutritive). A floraison, sur chaque plante, des épis homogènes en termes de nombre d'épillets ont été repérés et la date de floraison des florets des épillets centraux de ces épis notée. Le stade de développement des grains a été calculé en temps thermique (°Cd, base 0°C) à partir de cette date de floraison. Deux jours après floraison, la moitié des plantes de chaque génotype a été transférée en conditions de températures élevées (31/25°C jour/nuit soit une température moyenne journalière de 29°C) pendant environ 250°Cd, correspondant approximativement à la fin de la phase précoce de développement des grains. Les plantes restantes ont été laissées en conditions témoins (21/15°C jour/nuit soit une température moyenne journalière de 19°C). Ces conditions de stress sont identiques à celles appliquées en serre (expérimentation sur la diversité génétique). Pour chaque génotype et chaque traitement thermique, selon leur utilisation entre 5 et 12 épis ont été prélevés à différents stades de développement pour les analyses écophysiologique et transcriptomique.

Pour l'analyse morphométrique des grains, pour chaque génotype/condition thermique, 5 épis ont été prélevés et par épi les 2 grains basaux de l'épillet central à 10 stades de développement (0, 80, 120, 180, 220, 300, 500, 600, 700, 800°Cd). Pour tous les stades, la masse sèche des grains a été mesurée (d'autres variables caractérisant la croissance dimensionnelle et l'évolution hydrique du grain ont été mesurées mais seuls les résultats de croissance en masse sèche sont détaillés ici) ; au stade 300°Cd (en sortie de stress thermique et début de la phase de remplissage au sens strict), les grains ont été disséqués en vue d'isoler l'albumen et d'en dénombrer les cellules (estimateur de la force de puits des grains). Pour chaque génotype et traitement thermique, les cinétiques de croissance en matière sèche ont été ajustées par une fonction de croissance (logistique 3 paramètres) permettant d'extraire les paramètres caractéristiques de croissance des grains (valeurs maximales, vitesse et durée des processus).

L'analyse transcriptomique a été réalisée sur 3 répétitions biologiques prélevées par génotype/stade/condition ; et à 6 stades de développement du grain de blé (0, 80, 120, 180, 220 et 500°C.d). L'expression de 40 gènes a été suivie en qRT-PCR à tous les stades, pour les 2 génotypes, pour les 2 conditions. Ces 40 gènes ont été sélectionnés dans la littérature comme répondant à l'élévation de températures. En outre, des résultats antérieurs laissaient supposer que l'effet des températures élevées s'exprimait en priorité sur le métabolisme carboné. Aussi, 2 groupes de gènes ont été sélectionnés : l'un sur le métabolisme carboné et l'autre sur des voies de signalisation permettant de définir des gènes marqueurs de stress thermique potentiellement liés au niveau de tolérance au stress thermique.

Les réactions de PCR quantitative en temps réel (RT-qPCR) ont été réalisées sur la plateforme de génotypage à haut débit GENTYANE (UMR 1095 GDEC, http://gentyane.clermont.inra.fr). La distribution du mix de réaction dans les plaques qPCR 384 puits a été effectuée à l'aide du robot Microlab Hamilton Star puis la réaction de RT-qPCR

a été réalisée dans un thermocycleur LightCycler Roche LC480 selon le programme donné par les recommandations de l'appareil. Toutes les réactions ont été réalisées en duplicata.

Trois gènes de ménage ont été testés (Ta2776, Ta2291 et Ta54227 tirés de la publication de Paolacci et al. 2009). Parmi eux, le gène Ta2291 a été choisi pour normaliser les données de RT-qPCR entre plaques car son niveau moyen d'expression était similaire à ceux des gènes testés et qu'il n'est pas impacté par l'application d'un stress abiotique (Paolacci et al. 2009). L'expression relative des gènes d'intérêt a été déterminée selon la méthode  $2-\Delta\Delta$ Ct (Livak et Schmittgen, 2001).

Enfin, des amorces permettant l'amplification d'une séquence intronique du gène Proteine Disulfide Isomérase (PDI) chez le blé (Paolacci et al, 2009) ont été utilisées afin de vérifier l'absence de contamination par de l'ADN génomique. Une analyse en composantes principales a été réalisée à l'aide des données d'expression normalisées de l'ensemble des 40 gènes de tous les échantillons et du package R ggfortify (Tang et al. 2016). Pour chaque gène, un test non paramétrique de Kruskal-Wallis (homoscédasticité non respectée) a été réalisé pour mettre en évidence les différences entre les moyennes des échantillons.

## 3. Résultats

## 3.1 - Bilan agro-climatologique et disponibilités des données historiques

Les composantes de rendement mesurées en 2005 sur l'essai près du Neubourg (Eure) sont présentées ci-dessous :

Cv	Z55	Rend.	PMG	Dens Ep	Grains / épi	IR
Apache	21-mai	87.0	41.8	534	39	0.45
Caphorn	23-mai	101.6	40.7	500	50	0.48
Shango	18-mai	75.8	32.6	489	48	0.36
Soissons	30-mai	86.7	38.5	544	41	0.44

 Tableau 2 : liste des composantes retenues. Rendement en q/ha, exprimé à 15%

 de teneur en eau ; PMG en g ; Densité d'épis en épis/m².

On peut donc constater un écart significatif de rendement final, avec notamment une forte variation du PMG et de l'indice de récolte, et ceci malgré des biomasses totales élevées. La variété la plus impactée semble être Shango, qui s'avère également être la plus tardive. Sur ce site, une séquence chaude est apparue au cours de la seconde moitié du mois de juin (coïncidant avec le stade Grain Laiteux de la variété la plus tardive), avec des températures maximales dépassant ponctuellement 30°C. Il est donc supposé que cet évènement chaud tardif soit à l'origine des comportements différenciés au sein de cet essai, et notamment l'altération des composantes de rendement de la variété la plus tardive.



Figure 2 : Evolution des paramètres climatiques au cours du cycle

Des conditions climatiques similaires ont été recherchées dans d'autres secteurs géographiques, afin d'explorer le comportement des cultures dans de telles situations. Un essai a été identifié en 2005, dans le département de Seine-et-Marne (77). Les températures relevées (Tmax et Tmin) sur ce second site ont été effectivement plus élevées autour du stade Grain Laiteux, avec notamment un pic à 35°C.



Figure 3 : Evolution des paramètres climatiques au cours du cycle

Les composantes de rendement mesurées dans l'essai ne présentent pas d'altération nette selon la précocité (variété ou date de semis).

Cv	Date sem.	Z55	Rend.	PMG	Dens Ep	Grains / épi	IR
Apache	D2	22-mai	98.9	39.0	814	31.2	0.49
Apache	D3	25-mai	92.6	36.4	826	30.8	0.49
Caphorn	D2	24-mai	105.5	39.6	634	42.0	0.52
Caphorn	D3	27-mai	100.5	38.4	591	44.3	0.53
Lancelot	D1	25-mai	95.3	39.1	634	38.4	0.45
Lancelot	D2	28-mai	91.1	33.1	708	38.9	0.46

**Tableau 3 :** liste des composantes obtenues. Rendement en q/ha, exprimé à 15%de teneur en eau ; PMG en g ; Densité d'épis en épis/m².

Il semble donc qu'il n'est pas possible de définir un lien stable entre l'occurrence d'une séquence chaude et une pénalisation du PMG, toutes variétés confondues. L'absence de constance de réponse à des températures supérieures à 30°C laisse supposer qu'un déterminisme multifactoriel est à l'origine de l'altération du fonctionnement de la plante en fin de cycle, ou que les variétés se comportent de manière très différente (Shango non présent dans l'essai en Seine-et-Marne) ?

La classification ascendante hiérarchique basée sur l'analyse de valeurs moyennes par station a permis d'identifier 5 groupes climatiques définis par des critères thermiques et hydriques autour du remplissage, et représentant des scénarii distincts pour l'élaboration du rendement.

Sc.	Caractéristiques phénologiques et climatiques
1	Arrivée tardive à Z30, montaison et remplissage plus rapide; moyen- nement échaudant, mais Tmin élevées; bilan hydrique plutôt favorable
2	Précocité médiane; températures chaudes rencontrées précocement dans le cycle; alimentation hydrique plutôt bonne
3	Précocité médiane, peu de températures élevées ; conditions plutôt sèches
4	Phénologie précoce permettant de l'échappement ; occurrence forte de journées à ETP élevée ; déficit en eau pendant dès montaison; températures nocturnes pas trop élevées
5	Phénologie plutôt tardive avec montaison et remplissage longs, peu de températures élevées, peu d'ETP élevées; éventuellement un peu sec en montaison

Tableau 4 : liste des 5 scenarii identifiés



Figure 4 : Positionnement des stations météorologiques étudiées, réparties en 5 scenarii climatiques

Pour chacun de ces groupes, un parangon a été identifié et a permis de mieux préciser les conditions climatiques de fin de cycle ; pour cela, les critères climatiques ont été calculés sur 20 campagnes cultures (récoltes 1994 à 2013). De plus, des données horaires ont été mobilisées lorsqu'elles étaient disponibles.

Pour le scénario 4 (caractérisé par sa précocité et des demandes évaporatives élevées), le parangon retenu est la station de Misérieux (Ain). Plusieurs éléments ont été analysés :

- Les valeurs de températures minimales lors de journées
   « échaudantes » (par convention, journée où Tmax>25°C)
   s'avèrent plus élevées.
- Les ETP lors de ces mêmes journées sont en moyenne de près de 6mm/j et peuvent dépasser 7 voire 8 mm/j
- Les séquences où les Tmax dépassent 25°C sont souvent courtes : dans près de 50% des cas, il s'agit d'une journée isolée ou de 2 jours consécutifs. Les évènements longs restent peu présents lors du remplissage sur la période étudiée.

Pour le scénario 5, caractérisé par un cycle long et des conditions peu stressantes, le parangon retenu est la station du Neubourg (Eure). Les données horaires collectées en 2005 pendant la phase de remplissage des grains d'un essai supposément touché par un stress thermique ont présentés les caractéristiques suivantes :

- Des amplitudes journalières qui peuvent atteindre 15 à 20°C
- Des hausses de température horaires qui peuvent atteindre 4°C pendant les premières heures du jour
- Des températures supérieures à 25°C qui se maintiennent, pendant les journées les plus chaudes, pendant 8 à 11h consécutives
- Des niveaux de VPD très limités malgré les températures : entre 0 et 2 kPa.

#### 3.2 - Analyse de la diversité génétique en conditions contrôlées

Lors du stress thermique appliqué à chaque plante 3 jours après l'anthèse, aucun impact significatif du traitement sur la période de floraison n'a été observé dans les expériences de 2016 et 2017. En revanche, un léger effet du traitement a été observé sur le nombre d'épi par plante. Le nombre de graines n'a pas été affecté sur le brin maitre, mais a été affecté sur les talles: légèrement en

2016 avec un indice de stress (IS) de 8,3% (0,05 <valeur-P <0,1) et fortement en 2017 avec un IS à 13,9% (0,001 <P -value <0,01). Le traitement par stress thermique a un impact sur la morphologie du grain dans les deux expériences, peu importe si le grain a été mesuré sur le brin maitre ou sur les autres talles. Pour le PMG du brin maître, la moyenne était de 46,8 g et de 42 g dans le traitement optimal alors qu'elle était de 34,9 g et de 32,9 g dans le traitement optimal et stressé en 2016 et 2017, respectivement. Le traitement thermique affecte le PMG du brin maitre de 25,7% et 21,9% en 2016 et 2017, respectivement. En ce qui concerne le maintien du stay-green, seuls les axes X du point d'inflexion (XPI) et fin de la sénescence en 2016 étaient significativement affectés. L'ensemble de ces résultats montre que le stress obtenu est bien celui attendu : durant le remplissage, car sans modification (ou très faible) du nombre de grains par épi et une réduction significative du PMG. Nous disposons donc d'un jeu de phénotypique adéquat pour une utilisation en GWAS.

La décomposition de la variance a révélé des interactions GxT significatives pour le PMG, la surface des grains, la longueur et la largeur. Peu importe si elle a été mesurée sur le brin maître ou sur des talles en 2016 ou en 2017. L'interaction GxT était la plus importante pour le PMG et la largeur du grain sur le brin maître, représentant respectivement 22,4 et 25,6% de sa variance génétique en 2016 et 2017.

Ce jeu de données phénotypiques a aussi permis d'identifier 18 variétés contrastées (9 résistantes (en rouge dans le graphique ci-dessous) et 9 sensibles (en bleu)) qui ont été utilisées pour une validation aux champs en 2018 sur 6 lieux.



Figure 5 : Boxplot du PMG en condition optimale (NS) et stressée (S) sur les expérimentations 2016 et 2017

Malheureusement, les conditions climatiques de l'année 2018 n'ont pas permis d'obtenir des essais avec une contrainte thermique aux champs et ont ainsi rendu impossible la conclusion sur cette classification réalisée en conditions contrôlées.



Figure 6 : Données des PMG en conditions optimales (NS) ou stressées (S) pour les 18 génotypes (rouge/tolérant ; bleu/sensible) lors de l'expérimentation 2016

#### 3.3 - Génétique d'association

Au final 219k SNP ont été traités en GWAS sur 24 caractères regroupant des traits agronomiques (floraison, PMG, longueur/largeur des grains) et des traits physiologiques liés à la sénescence (contenu en chlorophylle, *stay-green*, vitesse de sénescence etc...). Les résultats d'association sont donc disponibles sous les 2 modèles : effet du SNP et interaction SNP x Traitement et est visualisable notamment sous forme de Manhattan Plot (Figure 7). Concernant le nombre de SNP associés, on peut noter :

 Effet SNP : 448 associations significatives SNP/Traits représentant 361 SNPs uniques

Trait	SNP/trait association
DOF	72
Grain.Area	2
Grain.Length_TALLES	1
Grain.Width	3
MARVIN.GW	33
MARVIN.GW_TALLES	31
MARVIN.NBSEEDS	22
MARVIN.NBSEEDS_TALLES	10
MARVIN.TKW	3
MARVIN.TKW_TALLES	22
SEN.a	13
SEN.D	2
SEN.END	18
SEN.fEND	1
SEN.START	2
SEN.VIT	1
SEN.XPI	5

Tableau 5 : tableau de relation traits et nombre de SNPs associés

- Effet interaction SNP x Traitement : 196 associations significatives SNP/Traits x Traitement représentant 103 SNPs uniques :

Trait	SNP/trait association/Traitement
Grain.Area	35
Grain.Area_TALLES	8
Grain.Length	26
Grain.Length_TALLES	9
Grain.Width	11
Grain.Width_TALLES	4
MARVIN.GW_TALLES	1
MARVIN.TKW	30
MARVIN.TKW_TALLES	7
SEN.a	1
SEN.D	8
SEN.fEND	7
SEN.r	7
SEN.XPI	4
SEN.YPI	2
AUCK	36



Un regroupement des SNP sur la base du déséquilibre de liaison (DL) a été effectué afin de définir 41 zones d'intérêts au sein du génome, notamment une positionnée sur le chromosome 4B

L'ensemble de ces marqueurs associés sera une ressource pour la sélection afin de suivre les loci d'intérêt dans le germplasm actuel.



Figure 7 : Manhattan plot des résultats GWAS pour le PMG

#### 2.4 - Phénotypage physiologique et moléculaire

# 2.4.a - Effets des températures élevées post-anthèse sur la masse finale des grains et sur leurs caractéristiques de croissance (durée et vitesse)

Quel que soit le génotype les températures élevées post-anthèse ont entraîné une diminution statistiquement significative (P<0.05) de la masse finale des grains, respectivement de 11.3 % et 30.7% pour les génotypes Altigo et Attlass (Figure 8).



Figure 8 : Effets de températures élevées post-anthèse sur les cinétiques d'accumulation en matière sèche des grains pour les génotypes Altigo (orange) et Attlass (vert). Les points et les cercles représentent les valeurs observées et les lignes continues et brisées les valeurs ajustées par une fonction logistique (3 paramètres). Les valeurs et les flèches rouges indiquent les diminutions de matière sèche finale des grains entre les traitements témoin (C) et températures élevées (HT). Les flèches bleues indiquent pour chaque génotype et chaque traitement thermique les durées d'accumulation de matière sèche dans les grains.

Ces diminutions sont du même ordre que celles observées en conditions de serre pour ces deux génotypes. Ces diminutions ne sont que partiellement liées à l'effet des températures élevées qui ont entraîné une diminution non significative statistiquement de 21.5 % du nombre de cellules maximum dans l'albumen en fin de phase de division cellulaire pour les deux génotypes (Figure 9).





Cette diminution s'est accompagnée de diminutions significatives de la matière sèche des grains dès le stade 180°Cd de 20% et de 27% respectivement pour les génotypes Altigo et Attlass. L'analyse des cinétiques d'accumulation en matière sèche des grains (Figure 8) montre que quel que soit le génotype, les températures élevées ont conduit à une diminution non significative des vitesses moyennes et maximales de ce processus (e.g. respectivement de 6.9% et 2.4 % pour les génotypes Altigo et Attlass pour la vitesse maximale). Les températures élevées ont conduit à une diminution de la durée (exprimée en jours calendaires) de l'accumulation de matière sèche dans les grains, respectivement de 10% (soit environ 4 jours de remplissage) pour le génotype Altigo (différence non significative au seuil 5%) et de 30% (soit environ 10 jours) pour le génotype Attlass (différence significative statistiquement au seuil 5%). Notons que les températures élevées ont des effets similaires sur les autres variables dimensionnelles (diminution des vitesses

des différents processus ; diminutions non significatives voire augmentations significatives des durées pour le génotype Altigo (engendrant une compensation incomplète vitesse/durée) mais diminution systématique des durées pour le génotype Attlass).

#### 2.4.b - Effets des températures élevées post-anthèse sur l'expression des gènes

L'analyse statistique de l'ensemble des données (ACP) (Figure 10) montre que, dans les conditions de stress thermique testées en chambre de culture, l'effet global du stress thermique est négligeable. Seuls les stades 180 et 220°Cd présentent un léger décalage dans l'expression des 40 gènes sélectionnés entre génotype représentant un effet génotype certain.

Les clusters d'expression des gènes (Figure 11) confirment le fait que le comportement des gènes est variable selon les conditions.



Figure 10 : Analyse en Composantes Principales (ACP) de l'expression des 40 gènes sélectionnés dans le grain du génotype Altigo (tolérant) et Attlas (sensible) en réponse au stress (S) et dans les conditions contrôles (NS) pour les 6 stades de développement du grain (0-80-120-180-220 et 500°Cj)



Figure 11 : Clusters d'expression des 40 gènes sélectionnés en fonction du génotype Attlas (AS) et Altigo (AG) et du stade (80, 120, 180, 220 et 500°Cj). Le vert représente un gène sous-exprimé et le rouge un gène sur-exprimé. Par exemple, l'expression de certains gènes diminue en conditions de stress thermique par rapport aux conditions contrôles comme l'enzyme IIB de ramification de l'amidon (Figure 12A) impliquée dans le métabolisme carboné, qui pourrait être lié à la diminution de la matière sèche (et de volume) du grain observé.

Il apparait cependant que l'expression d'autres gènes, au contraire, augmente, comme par exemple une protéine induite par l'auxine (Figure 12C). De plus, l'expression du stress est visible principalement dès 180°Cd et à 220°Cd, comme le montre la figure 12B présentant l'expression de HSP90, ou celle de la petite sous-unité de l'ADP-glucose pyrophosphorylase (TraesCS7A01G287400), ou encore l'enzyme IIB de ramification de l'amidon à nouveau. Ces données transcriptomiques concordent avec les résultats des données écophysiologiques à l'échelle de l'organe.

Aucune différence nette ne ressort en ce qui concerne la réponse différentielle de l'expression des gènes. Parmi les 40 gènes, deux gènes présentent toutefois une réponse différente selon le génotype : un gène lié à la différentiation cellulaire chez *Oryza brachyantha* RCD1 impliqué dans le manque d'azote

chez les eucaryotes (Okasaki et al. 1998) ; ou encore un gène impliqué dans la voie Ubiquitine Protéasome 26S (comme la F-Box TraesCS6B01G409000.1) qui présente une expression accentuée à 220°Cd sous l'effet du stress thermique qui est différente entre les 2 génotypes (Figure 13A).

De nombreux gènes liés à la signalisation cellulaire avaient été sélectionnés issus de travaux en cours comme par exemple 3 gènes impliqués dans la régulation transcriptionnelle appartenant aux facteurs de transcription de la famille NAC (un exemple est donné pour le TaNAC, TraesCS7A01G194700; mais le même profil est obtenu pour les gènes TaNAC TraesCS5A01G245900, TraesCS7A01G569100). Les résultats montrent que leur expression est également différente entre génotypes, particulièrement à 120°Cj, en réponse à la contrainte thermique appliquée. Ces résultats sont en accord avec les résultats obtenus par Girousse et al. (2018) et ceux de Guérin et al. (2019) qui ont montré que ces gènes sont grain-spécifique et fortement impliqués dans le développement du grain en réponse à une contrainte thermique similaire à celle de la présente étude, mais aussi en réponse à un stress de type sécheresse (Figure 12B).



Figure 12 : Cinétique d'expression relative des gènes codant pour (A) l'enzyme IIb impliquée dans la ramification de l'amidon (Starch branching enzyme IIb, TraesCS2A01G310300), (B) une protéine Heat Shock (HSP90, TraesCS2A01G033700), (C) une protéine induite par l'auxine (Auxin-induced protein 5NG4, TraesCS5A01G220400), et (D) la petite sousunité de l'ADP-glucose pyrophosphorylase (TraesCS7A01G287400) en fonction des °C.jours après floraison dans les grains d'Altigo (AG) et d'Attlas (AS) en condition de stress thermique (S)



**Figure 13 :** Log de Fold Change (FC) du ratio d'expression du TaNAC TraesCS7A01G194700 (A) et cinétique d'expression relative des gènes codant une F-Box (TraesCS6B01G409000) (B) . AG = Altigo, AS = Attlas, NS = non stressé, S = stress thermique

## 4. Discussion & Conclusions

Le stress thermique est un stress abiotique difficile à appréhender au champ car il existe peu de méthodes fiables pour piloter un tel stress en extérieur et il est d'autre part difficile d'avoir une condition contrôle en parallèle sur un site dédié. De ce fait, l'approche d'analyse de données historiques permettant d'identifier des années répondant aux critères de stress thermique et l'utilisation de conditions contrôlées pour générer du matériel et des données phénotypiques a permis de faire avancer les connaissances sur ce stress de plus en plus prévalant en France.

Au niveau des données historiques, l'analyse a permis de mettre en évidence 5 scénarii différents suite au traitement des données phénoclimatiques sur la période 1994-2013, la description fine des 5 classes devrait permettre de mieux analyser les sorties de la classification. Ceci met en évidence la difficulté d'identifier des lieux propices au stress thermique en France, reproductible et permettant de réaliser des essais robustes pour caractériser et mieux étudier ce stress. De ce fait, disposer d'un environnement contrôlé permettant de piloter un tel stress est une nécessité. L'utilisation de simulations phéno-climatiques pourrait par ailleurs apporter 3 informations supplémentaires : la représentativité des conditions expérimentales testées, la co-occurrence de stress climatiques (simultanéité de stress hydriques et thermiques) et l'évolution des conditions de remplissage des grains dans un contexte de changement climatique, avec ou sans adaptation de la phénologie.

La mise en place d'un protocole d'expérimentation pour le pilotage d'un stress thermique en conditions contrôlées a ainsi permis de mettre en évidence une réponse contrastée des lignées testées et de produire des données phénotypiques sur différents caractères d'intérêts. Cette ressource a permis d'identifier des lignées plus résistantes au stress thermique et de ce fait, utilisables en sélection et aussi d'identifier des marqueurs moléculaires liés à la réponse au stress thermique de façon spécifique. En effet, aux champs, il est très difficile de dé-corréler stress hydrique et stress thermique, de ce fait, les conditions en serre nous ont permis d'individualiser ces 2 stress très liés et, ici, d'identifier spécifiquement des loci impliqués dans la réponse au stress thermique. Une méta-analyse avec les données GWAS stress hydrique générées dans le cadre de BreedWheat (Touzy et al. 2019) permet d'identifier ces loci spécifiques au stress thermique et ceux, à l'inverse, plus généraux de réponse aux stress abiotiques. De plus, le protocole utilisé a ciblé le stade remplissage du grain, il est aussi possible d'adapter ce protocole et de cibler un autre stade, notamment floraison, afin de déterminer les loci similaires et spécifiques de ces 2 stress thermiques différents (même si celui-ci n'est pas (encore) très présent en Europe). Au niveau variétal, 18 lignées ont été sélectionnées pour le caractère discriminant (Sensible/résistant) au stress thermique en conditions contrôlées, celles-ci ont été implantées en 2018 dans un réseau d'essais afin d'évaluer aux champs leur comportement. Malheureusement, l'année 2018 n'a pas été propice au stress thermique (ceci démontrant l'utilité du travail en conditions contrôlées dans la 1ère partie du projet) et ne nous a pas permis de conclure sur la caractérisation de ces lignées.

Une analyse plus fine de ces génotypes contrastés a permis de montrer que, quel que soit le génotype considéré, le stress thermique impacte fortement la croissance du grain diminuant essentiellement la durée d'accumulation de matière sèche dans le grain, en concordance avec l'expression des gènes du métabolisme carboné (gènes du métabolisme de l'amidon en particulier). Ces gènes pourraient servir de marqueurs spécifiques du stress thermique. Toutefois, si à l'échelle du grain (masses et dimensions), l'effet génotypique a pu être mis en évidence, les études transcriptomiques n'ont pas permis de mettre en évidence un effet aussi marqué, cela pouvant être lié au trop faible nombre de gènes étudiés qui a pu être limitant dans cette étude. En particulier, le nombre de gènes sélectionnés au cours de cette étude ne permet pas de rendre compte de la différence de durée d'accumulation de matière sèche (amidon) entre génotypes ; il conviendrait alors d'élargir le spectre de gènes liés au métabolisme carboné. Compte tenu des observations faites à l'échelle du grain sur la croissance des autres variables (volume, masses d'eau du grain), il serait intéressant de regarder l'expression de gènes liés à l'état hydrique du grain (e.g. aquaporine).

Au final, HeatWheat a délivré les résultats attendus en terme de protocole de criblage du matériel vis-à-vis de la réponse au stress thermique, de caractérisation de lignées sensibles et résistantes et de connaissance sur les mécanismes physiologiques et moléculaires de réponse mis en place par la plante pour que son métabolisme ne soit pas impacté.

## Références bibliographiques

**Barah P, Jayavelu ND, Mundy J, Bones AM.** Genome scale transcriptional response diversity among ten ecotypes of Arabidopsis thaliana during heat stress. Front Plant Sci. 2013 Dec 26 ; 4:532

Bennett D, Reynolds M, Mullan D, Izanloo A, Kuchel H, Langridge P, Schnurbusch T. Detection of two major grain yield QTL in bread wheat (Triticum aestivum L.) under heat, drought and high yield potential environments. Theor Appl Genet. 2012 Nov;125(7):1473-85

**Bita CE, Gerats T.** Plant tolerance to high temperature in a changing environment: scientific fundamentals and production of heat stress-tolerant crops. Front Plant Sci. 2013 Jul 31;4:273

Bonneau J, Taylor J, Parent B, Bennett D, Reynolds M, Feuillet C, Langridge P, Mather D. Multi-environment analysis and improved mapping of a yield-related QTL on chromosome 3B of wheat. Theor Appl Genet. 2013 Mar;126(3):747-61

Brisson, Philippe Gate, David Gouache, Gilles Charmet, François-Xavier Oury, Frédéric Huard. Why are wheat yields stagnating in Europe ? Field Crops Research 2010

Calderini, D.F., Abeledo, L.G., Savin, R., Slafer, G.A. Effect of temperature andcarpel size during pre-anthesis on potential grain weight in wheat. J. Agric. Sci.1999, 132, 453-459.

**Calderini, D.F., Abeledo, L.G., Savin, R., Slafer, G.A.** Final grain weight in wheatas affected by short periods of high temperature during pre- and post-anthesisunder field conditions. Aust. J. Plant Physiol. 1999. 26, 453-458.

Calderini, D.F., Reynolds, M.P., Slafer, G.A. Source-sink effects on grain weightof bread wheat, durum wheat and triticale at different locations. Aust. J. Agric.Res. 2006. 57, 227-233.

**Chauhan H.** A Seed Preferential Heat Shock Transcription Factor from Wheat Provides Abiotic Stress Tolerance and Yield Enhancement in Transgenic Arabidopsis under Heat Stress Environment. PLoS ONE, 2013, 8(11): e79577

Cossani CM, Reynolds MP. Physiological traits for improving heat tolerance in wheat. Plant Physiol. 2012 Dec;160(4):1710-8

**Dhyani K.** Comparative physiological response of wheat genotypes under terminal heat stress. Plant Signaling & Behavior 2013 8:6, e24564

Girousse C, Roche J, Guérin C, Le Gouis J, Balzegue S, Mouzeyar S, Bouzidi MF. 2018. Coexpression network and phenotypic analysis identify metabolic pathways associated with the effect of warming on grain yield components in wheat. PLoS One. 2018 Jun 25;13(6):e0199434. doi: 10.1371/journal.pone.0199434.

Gouache D, X Le Bris, M Bogard, O Deudon, C. Page, P. Gate. Evaluating agronomic adaptation options to increasing heat stress under climate change during wheat grain filling in France. Europ. J. Agronomy 39 (2012) 62-70

Guérin C, Roche J, Allard V., Ravel C, Mouzeyar S, Bouzidi MF. 2019 PLoS One publication acceptée, en cours de publication. Genome-wide analysis, expansion and expression of the NAC family under drought and heat stresses in bread wheat (T. aestivum L.).

Liu H. Common and Distinct Functions of Arabidopsis Class A1 and A2 Heat Shock Factors in Diverse Abiotic Stress Responses and Development. Plant Physiology, 2013, Vol. 163, pp. 276-290 Mittler R. How do plants feel the heat? Trends in Biochemical Sciences March 2012, Vol. 37, No. 3

**Morgan JM.** Osmotic Components and Properties Associated With Genotypic Differences in Osmoregulation in Wheat. Australian Journal of Plant Physiology 19(1) 67 - 76

Okazaki N, Okazaki K, Watanabe Y, Kato-Hayashi M, Yamamoto M, Okayama H, 1998. Novel factor highly conserved among eukaryotes controls sexual development in fission yeast. Mol Cell Biol. Feb; 18(2):887-95.

**Ortiz R.** Wheat genetic resources enhancement by the International Maize and Wheat Improvement Center (CIMMYT). Genet Resour Crop Evol. 2008. 55: 1095-1140

**Qin D.** Heat stress-responsive transcriptome analysis in heat susceptible and tolerant wheat (Triticum aestivum L.) by using Wheat Genome Array. BMC Genomics 2008, 9:432

**Rebetzke G.** Genomic regions for canopy temperature and their genetic association with stomatal conductance and grain yield in wheat. Functional Plant Biology, 2013, 40, 14-33

Reynolds, M.P., Balota, M., Delgado, M.I.B., Amani, I., Fischer, R.A. Physiological and morphological traits associated with spring wheat yield under hot, irrigatedconditions. Aust. J. Plant Physiol. 1994. 21, 717-730.

**Reynolds MP.** Evaluating potential genetic gains in wheat associated with stress-adaptive trait expression in elite genetic resources under drought and heat stress. Crop Sci. 2007. 47: S-172-S-189

Semenov MA, Shewry PR. Modelling predicts that heat stress, not drought, will increase vulnerability of wheat in Europe. Sci Rep. 2011;1:66

**Stone, P.J., Nicolas, M.E.** Wheat cultivars vary widely in their responses of grain-yield and quality to short periods of postanthesis heat-stress. Aust. J. PlantPhysiol. 1994 21, 887-900.

**Suzuky Pinto.** Heat and drought adaptive QTL in a wheat population designed to minimize confounding agronomic effects. Theor Appl Genet (2010) 121:1001-10

Tashiro T, Wardlaw IF (1990) The effect of high temperature at different stages of ripening on grain set, grain weight and grain dimensions in the semi-dwarf wheat'banks'. Annals of Botany 65: 51-61

Wang X, Cai J, Liu F, Jin M, Yu H, Jiang D, Wollenweber B, Dai T, Cao W (2012) Pre-anthesis high temperature acclimation alleviates the negative effects of post-anthesis heat stress on stem stored carbohydrates remobilization and grain starch accumulation in wheat. Journal of Cereal Science 55: 331-336

Wollenweber, B., Porter, J.R., Schellberg, J. Lack of interaction betweenextreme high-temperature events at vegetative and reproductive growth stagesin wheat. J. Agron. Crop Sci. 2003 189, 142-150.

Xu, Q.A., Paulsen, A.Q., Guikema, J.A., Paulsen, G.M. Functional and ultrastruc-tural injury to photosynthesis in wheat by high-temperature during maturation. Environ. Exp. Bot. 1995. 35, 43-54

Yu E, Fan C, Yang Q, Li X, Wan B, Dong Y, Wang X, Zhou Y. Identification of Heat Responsive Genes in Brassica napus Siliques at the Seed-Filling Stage through Transcriptional Profiling. PLoS One. 2014 Jul 11;9(7)
## [GS-qualite] Établissement d'un modèle de Sélection Génomique pour la Qualité Boulangère des blés

#### Bruno POUPARD<sup>1</sup>\*, Sophie BOUCHET<sup>2</sup>, François Guion<sup>3</sup>, Matthieu BOGARD<sup>4</sup>

1 - LIMAGRAIN EUROPE - CS50005 ST BEAUZIRE, 63360 GERZAT

- 2 INRA 147 rue de l'université, 75007 PARIS
- 3 ANMF 66 rue La Boétie, 75008 PARIS
- 4 ARVALIS Institut du végétal 3 rue Joseph et Marie Hackin, 75016 PARIS

\* Coordinateur : Bruno POUPARD, bruno.poupard@limagrain.com

## 1. Introduction et objectifs du projet

#### 1.1 - Introduction

L'offre variétale de blé tendre actuelle répond globalement bien aux besoins des utilisateurs. C'est la conséquence d'une volonté de la filière traduite par les critères d'inscription des variétés au CTPS. Le niveau de tous les critères technologiques (Poids Spécifique, force boulangère et Note Totale de Panification) a augmenté depuis 15 ans (Méléard, 2010). Toutefois, le progrès génétique, notamment en matière de Note Totale de Panification, pourrait notablement être amélioré. En effet, la complexité de l'élaboration de cette note, le besoin important en grain qu'elle requiert, ainsi que son coût, impliquent que ce critère n'est évalué dans les programmes de sélection que dans les étapes finales, juste avant le positionnement pour l'inscription au catalogue.

Dans une publication fondatrice, Lande and Thompson (1990) ont introduit la théorie de la Sélection Assistée par Marqueurs (SAM). Un score moléculaire est calculé comme la somme des effets aux allèles des QTL (Quantitative Trait Loci) détectés. Son intérêt pour l'amélioration de caractères quantitatifs chez les plantes a été démontré expérimentalement (ex. Eathington et al. 2007, Blanc et al. 2008) et ces méthodes sont aujourd'hui largement utilisées par les grandes entreprises de sélection. Néanmoins, l'efficacité de cette SAM peut être limitée par la première étape de détection de QTL, dont la puissance est faible pour les QTL d'effets faibles avec les tailles usuelles des populations de sélection. En effet, pour des caractères complexes comme le rendement, dont on peut supposer que les QTLs à effets fort ont déjà été sélectionnés et fixés, le contrôle génétique est vraisemblablement assuré par un très grand nombre de gènes à effets très faibles (modèle « infinitésimal »). Comme ceux-ci sont difficiles à détecter, la variance expliquée par les QTL retenus dans le modèle est souvent faible. C'est pourquoi, les premiers, Whittacker et al. (2000) ont suggéré d'inclure dans le modèle de prédiction tous les marqueurs, sans étape préalable de détection de QTL. Comme le nombre de margueurs est le plus souvent très supérieur au nombre d'observations, les méthodes classiques de régression multiple ne peuvent pas être utilisées. En conséquence, Whittacker et al. (2000) ont suggéré l'utilisation de régressions pénalisées (« Ridge Regression ») pour résoudre ce problème de sur-paramétrage. Cette approche, aujourd'hui communément nommée Sélection Génomique (GS) a été largement développée en génétique animale (Meuwissen et al. 2001), pour la prédiction des valeurs d'élevage (Breeding value). Cette estimation requière des informations sur les génotypes aux marqueurs et les phénotypes dans une population de calibration pour calculer les estimations génomiques des valeurs génétiques (Genomic Estimation of Breeding Value, ou GEBV). Ces GEBV peuvent ensuite être estimées, à l'aide des seuls

marqueurs, dans n'importe quelle population cible, où la sélection peut donc être conduite sur les GEBV plutôt que sur les phénotypes.

Le principal intérêt de la Sélection Génomique est de pouvoir remplacer la sélection phénotypique quand l'évaluation du caractère est longue et/ou coûteuse. C'est bien le cas pour le test de panification, qui est cher (>100€), nécessite plusieurs kilos de grain, et n'est donc effectué que sur quelques lignées de générations avancées. L'efficacité relative de la GS par rapport à la sélection phénotypique dépend de la précision des prédictions (GEBV).

À notre connaissance, aucune étude n'a encore porté sur la valeur boulangère telle qu'estimée par le test BIPEA. Pour d'autres caractères liés à la qualité comme la teneur en protéine ou le rendement en farine, Heffner *et al.* (2011b) rapportent des précisions variant de 0.45 à 0.76. À ce jour, la seule publication sur la prédiction du test BIPEA à l'aide des gluténines (Oury *et al.* 2010) rapporte des précisions de prédiction (r, coefficient de corrélation) de l'ordre de 0.4, qui montent à 0.55 si on inclue les tests indirects (teneur en protéines, dureté, alvéographe) dans l'équation de prédiction.

Des études plus récentes concernent des paramètres physiques de la pâte qui ne sont pas utilisés dans le système d'inscription français (Michel *et al.*, 2017, 2018, Hayes *et al.*, 2017).

#### 1.2 - Objectifs de ce projet

Notre hypothèse était que le développement d'un outil de Sélection Génomique permettrait de sélectionner pour la Note Totale de Panification plus en amont dans les programmes de sélection, et donc d'augmenter la réponse à la sélection pour ce caractère. Les objectifs de ce projet étaient les suivants :

- Construire une base de calibration de Sélection Génomique permettant de prédire la Note Totale de Panification, ainsi que ses composantes, et la teneur en protéines,
- Evaluer la précision de ces prédictions et déterminer par simulation les facteurs influençant cette précision,
- Évaluer la « portabilité » des prédictions d'un jeu de données à l'autre,
- Tester l'effet de l'inclusion dans les modèles GS des variables phénotypiques de l'alvéographe,
- Sur la base des résultats obtenus, proposer un protocole pour la construction d'une base de calibration de Sélection Génomique pour la qualité boulangère,
- Réaliser un test d'associations pour identifier les zones génomiques (QTL) ayant un effet sur la qualité boulangère et la teneur en protéine.

Afin de réaliser cette étude nous avons utilisé un panel de lignées de sélection du programme de Limagrain et utilisé les données d'essais post-inscription d'ARVALIS - Institut du végétal.

## 2. Matériel et méthode

#### 2.1 - Matériel végétal du Panel Recherche

Le matériel végétal de ce panel était constitué de 2318 lignées de sélection, issues du programme de recherche de Limagrain dont l'objectif est la création de lignées commerciale pour le marché français. Les parents de ces croisements étaient soit des variétés commerciales françaises, soit des lignées de sélection performantes pour le marché français. Au total, 74% (1725) de ces lignées étaient des Haploïdes Doublés (HD), le reste (593) étant constitué des lignées obtenues par Single Seed Descend (SSD) au stade F3:5 (la dernière plante unique est une F3, mais les plantes cultivées sur les micro-parcelles d'essai sont des F5).

L'année précédant leur mise en essai, ces lignées ont été sélectionnées sur des critères agronomiques (épiaison, hauteur, résistances aux stress...) de manière à être adaptées au marché français, mais aucune sélection n'a été effectuée sur leur qualité ou leur rendement.

#### 2.2 - Dispositif expérimental du Panel Recherche

Le dispositif expérimental initial prévoyait 12 environnements (4 lieux d'essais x 3 années), avec dans chaque environnement 192 lignées de sélection ainsi que 4 variétés témoins (observation sans répétition). Les conditions climatiques ayant eu un impact négatif sur la qualité boulangère pour certains environnements (germination sur pieds observée lors de la campagne 2013-2014, conditions climatiques très atypiques au Nord de la Loire lors de la campagne 2015-2016), nous avons dû adapter notre dispositif. Le dispositif final est décrit dans le Tableau 1. Chaque lignée de ce panel a donc été observée une seule fois, tandis que les témoins ont été répétés d'un environnement à l'autre et au sein de chaque environnement.

Lieu	Commune (Département)	2014	2015	2016	Total
CVCV	Chartainvilliers (Eure-et-Loir)		184 APACHE (2), AREZZO (2), CELLULE (2), RUBISKO (2)		184 (8)
CVTA	Aigrefeuille-d'Aunis (Charente-Maritime)	388 APACHE (2), AREZZO (2), SOLEHIO (2), SY MOISSON (2)	157 APACHE (2), AREZZO (2), CELLULE (2), RUBISKO (2)	396 AREZZO (2), CELLULE (2), RUBISKO (2), SOLEHIO (2), SY MOISSON (1)	941 (25)
CYRV	Revel (Haute-Garonne)	380 APACHE (2), AREZZO (2), SOLEHIO (2), SY MOISSON (2)		387 AREZZO (2), CELLULE (2), RUBISKO (2), SOLEHIO (1), SY MOISSON (1)	767 (16)
VRPE	Péronne (Somme)		209 APACHE (3), AREZZO (2), CELLULE (2), RUBISKO (2)		209 (9)
VRVR	Verneuil l'Etang (Seine-et-Marne)		217 APACHE (1), AREZZO (2), CELLULE (2), RUBISKO (2)		217 (7)
Total		768 (16)	767 (32)	783 (17)	2318 (65

**Tableau 1 :** Dispositif expérimental du Panel Recherche. Dans chaque cellule est indiqué le nombre de lignées observées (en gras), ainsi que les témoins présents et leur nombre de répétitions (entre parenthèses).

Dans chaque environnement, les lignées et les témoins ont été cultivés sur des micro-parcelles de 6m<sup>2</sup> selon un itinéraire technique non-limitant (fertilisation azotée, traitements fongiques...).

#### 2.3 - Test de panification du Panel Recherche et transformation des données

Les tests de panification ont été réalisés par le laboratoire Qualtech sur des échantillons issus de la récolte des microparcelles d'essai, selon le test BIPEA pain courant (NF V 03-716, Décembre 2008). L'Annexe 1 présente le détail des variables de ce test. Le teneur en protéine des farines a été mesurée par spectrométrie infrarouge.

Les notes composantes du test de panification BIPEA suivent un barème non linéaire : les scores -1, -4, -7, indiquent un déficit (-1 étant le plus extrême), tandis que les scores 7, 4 et 1 indiquent un excès (1 étant le plus fort). Le score de 10 traduit le niveau optimal de chaque variable. Pour des raisons arithmétiques, il a été nécessaire de linéariser selon la méthode présentée dans le Tableau 2.

SCORE BIPEA	SIGNIFICATION	<b>RECODAGE ECHELLE 1-10</b>
-1	Fort déficit	1
-4	Déficit	2.5
-7	Déficit léger	4
10	Optimal	5.5
7	Excès léger	7
4	Excès	8.5
1	Fort excès	10

 Tableau 2 : Recodage des scores du test BIPEA. Exemple : le score BIPEA de -7,

 qui correspond à un déficit léger par rapport à l'optimum, est recodé en 4. Ainsi, sur

 la nouvelle échelle, plus la valeur augmente et plus on passe d'un déficit à un excès.

## 2.4 - Matériel végétal du Panel Variétal et dispositif expérimental

Un panel de 192 variétés commerciales (« Panel Variétal ») a aussi été utilisé dans ce projet. Il s'agit de variétés étudiées dans le cadre de l'activité d'évaluation post-inscription d'ARVALIS -Institut du végétal. Ce panel comprend 3 des 6 témoins du Panel Recherche (APACHE, ARREZO et SOLEHIO). Les noms de variétés de ce panel sont présentés en Annexe 2.

Le dispositif expérimental de ce panel est fortement déséquilibré et les conditions environnementales et de culture sont extrêmement diverses. Les données de ce panel ont été acquises entre 2003 et 2013 avec des variétés communes entre les années. En moyenne, 57 variétés ont été testées par an. Le jeu de données comporte 158 lieux de récolte différents avec en moyenne 13 variétés par lieux. Les modalités de culture testées dans ces essais étaient principalement le niveau de fertilisation azotée et l'irrigation (essai pluvial versus irrigué). Le Tableau 3 donne plus de détails sur ce dispositif. Au final, les variétés ont été observées en moyenne 20 fois (médiane à 15), et sur 1 à 10 environnements (3.6 en moyenne).

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
2003	33										
2004	30	65									
2005	15	49	63								
2006	6	35	43	81							
2007	5	19	22	51	79						
2008	5	18	16	35	59	83					
2009	8	19	19	33	54	68	106				
2010	6	17	15	22	27	38	68	71			
2011	7	16	16	24	25	18	45	44	51		
2012	4	10	11	19	22	17	23	21	25	27	
2013	1	8	10	17	20	16	21	20	19	19	23

Tableau 3 : Occurrence des variétés testées en fonction des années pour la Panel Variétal. Les chiffres représentent le nombre de variétés communes entre les deux années (ligne, colonne). Les chiffres sur la diagonale représentent les variétés présentes dans l'année correspondante.

Les caractères mesurés sur ce panel sont les variables d'alvéographe (W, G, P, L, le), les notes de panification et leurs composantes, le taux de protéines, la dureté, le taux d'hydratation et le volume du pain. Les notes de panification sont attribuées selon la norme NF V03-716 en vigueur lors de la réalisation du test. Toutes les mesures ont été effectuées au laboratoire d'ARVALIS Institut du végétal de Boigneville (91).

Les scores du test BIPEA ont été linéarisés selon la méthode présentée dans le paragraphe 2.3

#### 2.5 - Ajustement des données phénotypiques

Pour les 2 panels, les données phénotypiques ont été ajustées de manière à décomposer les effets génétiques des effets environnementaux. Pour ce faire, les différents facteurs environnementaux (facteurs « année » et « lieu » pour le Panel Recherche, facteurs « année », « conduite » et « lieu » pour le Panel Variétal) ont été combinés pour former un facteur « Environnement ».

Puis les moyennes ajustées ont été calculées via un modèle mixte en considérant les facteurs « Génotype » et « Environnement » comme aléatoires. Le modèle pour la variété i et l'essai k s'écrit ainsi :

$$Y_{ik} = \mu + G_i + E_k + \varepsilon_{ik}$$

Avec  $Y_{ik}$  la valeur phénotypique du Génotype i dans l'Environnement K,  $\mu$  la moyenne générale du trait étudié, G l'effet Génotype, E le facteur Environnement et  $\epsilon$  la résiduelle.

L'héritabilité (h²) a été calculée de la façon suivante :

$$h^2 = \frac{VG}{VG + VR}$$

avec VG la variance génétique et VR la variance résiduelle.

#### 2.6 - Génotypage et nettoyage des données

L'ensemble du matériel (2318 lignées de sélection du Panel Recherche, 192 variétés du Panel Variétal, 3 témoins non compris dans le Panel Variétal) a été analysé au laboratoire de Génotypage de Limagrain (Chappes, 63), au moyen d'une puce de haute densité développée par Limagrain et comprenant 14592 marqueurs SNP. 97.8% de ces marqueurs ont une position de cartographie génétique (carte consensus interne à Limagrain, la répartition des SNP sur les chromosomes est présentée dans le Tableau 4).

Chromosome	Α	В	D	Inconnu
1	876	948	251	
2	742	1034	285	
3	891	1150	337	
4	659	595	307	
5	945	1038	286	
6	577	859	221	
7	1004	990	276	
Inconnu				321
Total	5694	6614	1963	321

Tableau 4 : Répartition des 14592 marqueurs SNP utilisés sur le génome

Les données de génotypage ont été nettoyées de la façon suivante : les marqueurs monomorphes, ayant un taux de données manques supérieur à 30%, un taux d'hétérozygotie supérieur à 10% ou une fréquence de l'allèle minoritaire inférieure à 5% ont été écartés. Les individus ayant plus de 10% de données manquantes ont aussi été écartés.

#### 2.7 - Modèle de Sélection Génomique, cross validation et portabilité des prédictions

Le modèle de prédiction génomique utilisé dans notre étude est un GBLUP (Meuwissen *et al.*, 2001) qui s'écrit de la façon suivante :

$$Y_i = \mu + G_i + \varepsilon_i$$

Avec  $Y_i$  la moyenne ajustée du génotype i,  $\mu$  la moyenne générale, G l'effet Génétique avec G~N (0, K $\sigma^2$ g), K la matrice d'apparentement et  $\epsilon$  la résiduelle.

Les calculs ont été effectués à l'aide de la fonction kin.blup du package 'rrBLUP' (Endelman, 2011). La matrice d'apparentement a été calculée avec la fonction A.mat du même package. Cette fonction calcule une matrice d'apparentement qui estime une IBS (Identity By State, Endelman and Jannink, 2012).

Les capacités prédictives des modèles GS ont été estimées par validation croisée de type k-fold avec k=10. Cette procédure a consisté à séparer le jeu de données en k échantillons. Le modèle a été calibré sur k-1 échantillons puis les valeurs prédites pour l'échantillon k ont été calculées et comparées aux valeurs observées. Pour chaque échantillon k, le coefficient de corrélation entre valeurs prédites et observées a été calculé. La capacité prédictive de chaque modèle a été définie comme la moyenne des coefficients de corrélations des k échantillons.

La « portabilité » des équations de prédiction génomique obtenues sur ces 2 panels a été évaluée en réalisant des prédictions croisées. Autrement dit, les moyennes ajustées du Panel Recherche ont été prédites à l'aide du modèle calibré sur les données du Panel Variétal et vice versa.

#### 2.8 - Influence des caractéristiques de la base de calibration et utilisation de covariables

L'influence du nombre de marqueurs pris en compte dans la calibration, ainsi que l'influence du nombre de génotypes, ont été testées sur le Panel Recherche. Pour cela, un sous échantillon de taille croissante (Nombre de Marqueurs = 50, 100, 1000, 5000, 8000, 1000 et Nombre de Génotypes = 50, 250, 500, 1000, 1500, 2000) a été tiré aléatoirement et une cross-validation a été effectuée selon la procédure définie au paragraphe 2.7. Cette procédure a été répétée 30 fois de manière à calculer la précision moyenne ainsi que son écart-type.

L'influence du nombre d'observations par génotype, quant à elle, a été testée sur le Panel Variétal selon la même procédure en prenant Nombre d'Observations = 1, 2, 3, 4, 5. Pour chaque valeur d'Observation la cross-validation a été itérée 10 fois.

L'intérêt d'inclure dans les modèles de prédiction génomique des covariables phénotypiques liées à la qualité a également été testé. Les covariables utilisées sont celles mesurées par l'alvéographe. Pour ce faire, les variables ont été ajoutées au modèle GBLUP décrit dans le paragraphe 2.7, et une procédure de cross-validation similaire a été utilisée. Les résultats obtenus ont été comparés aux résultats du GBLUP ainsi qu'aux résultats d'un modèle de régression linéaire (variable BIPEA = f(variable alvéographe). L'effet de ces covariables n'a pu être testé que sur le Panel Variétal car elles n'ont pas été mesurées sur le Panel Recherche.

#### 2.9 - Tests d'associations

Un test d'associations a été réalisé pour l'ensemble des variables sur chacun des 2 panels. La méthode utilisée correspondait à la méthode MLMM proposée par Segura *et al.* (2012). Il s'agit d'une méthode itérative consistant à tester pour chaque SNP le modèle suivant :

#### $Y = \mu + \alpha X + \beta SNP + \gamma G + \varepsilon$

Avec Y la valeur phénotypique (moyenne ajustée),  $\mu$  la moyenne générale, aX l'effet des SNP inclus en effets fixes,  $\beta SNP$  l'effet du SNP testé, G l'effet du fond génétique avec G~N(0, K $\sigma^2$ g), K la matrice d'apparentement et  $\varepsilon$  la résiduelle.

A chaque itération, le marqueur le plus significatif est inclus en effet fixe dans le modèle et la procédure se termine lorsque plus aucun marqueur ne présente une p-value inférieure ou égale à un seuil prédéfini (0.0001 dans le cas présent).

## 3. Résultats - Discussion

#### 3.1 - Données moléculaires et structuration

Après nettoyage des données de génotypage, le Panel Variétal comprend 188 variétés et 10482 marqueurs, et le Panel Recherche 2300 lignées et 10759 marqueurs.

La Figure 1 présente les résultats de l'Analyse en Composantes Principales réalisée sur les résultats de génotypage.



Figure 1 : ACP sur les données Génotypiques

Il apparait que les 2 panels sont peu structurés (Variance expliquée par les 2 premiers axes : 8.2%). Par ailleurs, le Panel Recherche semble bien recouvrir le Panel Variétal. L'inverse n'est cependant pas totalement vrai puisque l'on constate que la majorité des variétés de ce panel se situent dans la partie gauche du graphique (proche du témoin APACHE). Les témoins quant à eux semblent se positionner dans ces nuages de points, mais en périphérie.

#### 3.2 - Données Phénotypiques

La Figure 2 montre que les 2 panels ont des performances moyennes très similaires pour la teneur en protéines et les Notes de Pâte, de Pain, de Mie et de Panification Totale. Cette dernière tourne autour de 250 sur 300 pour les 2 panels, avec de façon surprenante une variabilité moins importante sur le Panel Recherche, pourtant non sélectionné sur ce critère.

La Note de Pain, avec des valeurs moyennes comprises entre 65 et 68 pour les 2 panels, est la composante la plus faible et la plus variable de la Note Totale de Panification.

La Note de Pâte, dont la moyenne tourne de 85 pour les 2 panels, est la deuxième composante impactant le plus la Note Totale de Panification. Il est à noter que les variables entrant dans le calcul de la Note de Pâte présentent des corrélations fortes, notamment les variables liées à l'extensibilité, fortement corrélées entre elles et négativement corrélées aux variables d'élasticité (résultats non montrés).

La Note de Mie présente des valeurs très fortes (>95) et peu variables, de même que ses composantes (résultats non montrés). Ces variables ont donc souvent été écartées des analyses suivantes, à part la couleur de la mie (COUM). Un certain nombre d'autres variables (Collant au façonnage et à la mise au four, déchirement à l'apprêt...), elles aussi, ne présentent pas ou peu de variabilité et ont été écartées des analyses suivantes.

L'importance des effets environnementaux n'a pas été étudiée en détail, mais il est rappelé que nous avons écarté 4 environnements de cette étude à cause de conditions climatiques impactant négativement et de façon non désirée les résultats de panification. Pour le Panel Recherche, le Tableau 5 permet de voir que la Note Totale de Panification était au-dessus de 250 pour tous les environnements sauf deux où elle chute fortement (Aigrefeuille-d'Aunis 2015 = 238, Revel 2016 = 224) du fait de faibles Notes de Pain.



**Figure 2 :** Comparaison des distributions de la teneur en protéines, de la note de pain (NPAI), de mie (NMIE), de pâte (NPAT) et de la note totale de panification entre les panels Variétal et Recherche.

ANNEE	LIEU	PROT	VOL	NPAT	NPAI	NMIE	NPANI
2014	CVTA	10.80 (0.77)	1561 (193)	80.2 (9.6)	69.7 (13.0)	99.0 (3.7)	249.3 (19.6)
2014	CYRV	11.67 (0.85)	1617 (202)	81.5 (10.1)	72.1 (12.5)	99.1 (3.5)	253.1 (19.1)
2015	CVCV	11.53 (0.65)	1519 (143)	86.7 (12.0)	68.3 (17.9)	95.0 (4.4)	250.3 (28.6)
2015	CVTA	12.46 (0.74)	1501 (143)	84.4 (15.4)	58.0 (21.1)	95.6 (5.0)	238.5 (35.7)
2015	VRPE	10.99 (0.71)	1453 (160)	86.7 (11.4)	67.3 (16.4)	96.7 (4.1)	251.0 (26.9)
2015	VRVR	11.32 (0.66)	1511 (120)	87.8 (10.1)	70.6 (13.7)	96.0 (4.3)	254.7 (20.4)
2016	CVTA	11.72 (0.65)	1435 (131)	89.7 (10.0)	67.9 (16.5)	97.9 (3.4)	255.4 (23.7)
2016	CYRV	12.57 (0.71)	1412 (162)	81.1 (17.4)	46.8 (26.2)	96.6 (3.7)	224.4 (40.9)
GLO	BAL	11.63 (0.94)	1502 (179)	84.3 (12.7)	64.9 (19.7)	97.4 (4.1)	246.8 (29.5)

**Tableau 5 :** Moyenne et écartype (entre parenthèses) pour chaque Environnement du Panel Recherche. Les variables présentées sont la Teneur en Protéine (%, PROT), le Volume (cm3, VOL), la Note de Pâte (NPAT), la Note de Pain (NPAI), la Note de Mie (NMIE) et la Note Totale de Panification (NPANI).

Quant aux interactions Génotype x Environnement, elles n'ont pas été formellement étudiées. Cependant, leur importance par rapport à la Variance Génétique peut être approchée via le calcul d'héritabilité effectué sur le Panel Variétal (voir Tableau 6). Si les Notes de Panification, la Teneur de Protéines et le Volume ont des héritabilités similaires (autour de 0.4), les variables affectant les Notes de Mie, de Pâte et de Pain ont des héritabilités plus faibles (autour de 0.3), bien plus faibles que les héritabilités des variables de l'alvéographe (0.7).

#### 3.3 - Précision des modèles de GS

Les résultats des validations croisées sur les 2 panels sont présentés dans le Tableau 6. Pour le Panel Variétal, la précision des prédictions pour les critères de panification varie de -0.01 pour ELAM (élasticité de la mie) à 0.56 pour COUM (couleur de la mie), avec une précision moyenne de r=0.21, ce qui est globalement faible. La précision est en revanche meilleure pour les variables de l'alvéographe (r moyen = 0.41). Comme attendu, la précision de prédiction au carré (r<sup>2</sup>) d'un caractère est toujours inférieure à son héritabilité.

Pour le Panel Recherche, la précision des prédictions pour les critères de panification varie de 0.13 pour SOUM (souplesse de la mie) à 0.56 pour EXTF (extensibilité au façonnage), avec une moyenne à r=0.31, ce qui est meilleur que pour la Panel Variétal mais reste faible.

La teneur en protéines et le Volume sont prédits avec des précisions similaires sur les 2 panels (r~0.4), tandis que la prédiction de la Note Totale de Panification est nettement meilleure avec le Panel Recherche (0.35) qu'avec le Panel Variétal (0.14). Il est à noter que les variables liées à l'extensibilité et à l'élasticité au pétrissage et au façonnage (EXTP, ELAP, EXTF, ELAF) ainsi qu'à la couleur de la mie (COUM) semblent plutôt bien prédites par rapport aux autres variables du test de panification.

#### 3.4 - Portabilité des modèles de Sélection Génomique

Les résultats montrent que les capacités prédictives des modèles calibrés sur le Panel Variétal sont faibles (-0.11 < r < 0.13). Elles sont, en revanche, plus élevées pour les modèles calibrés sur le Panel Recherche (-0.28 < r <0.43, voir le Tableau 7). Les prédictions obtenues avec les modèles calibrés sur le Panel Recherche apparaissent donc plus robustes bien que le niveau de précision soit relativement modéré. Ces différences sont à mettre en relation avec la taille des jeux de données de calibration (2300 lignées pour le Panel Recherche contre 190 variétés inscrites pour le Panel Variétal), facteur connu comme ayant un impact significatif sur les capacités prédictives des modèles et leur robustesse. Par ailleurs, bien que le niveau de précision soit relativement modéré, ces modèles peuvent néanmoins trouver une application pratique en sélection étant donné que les caractères analysés résultent de tests relativement coûteux et qui ne peuvent être mis en œuvre que tardivement dans les schémas de sélection.

Catégorie	Variables	Héritabilité Panel Variétal	Validation croisée Panel Variétal	Validation croisée Panel Recherche
	Note de panification	0.41	0.14	0.35
	Note de mie	0.30	0.09	
	COUM	0.52	0.56	0.50
	SOUM	0.24	0.00	0.13
	ELAM	0.11	-0.01	
	EPAL	0.37	0.12	
	Note de pâte	0.46	0.16	0.32
	LISP	0.37	0.26	0.32
	COLP	0.45	0.26	0.33
	EXTP	0.31	0.32	0.48
A	ELAP	0.23	0.16	0.45
IPE	RELP	0.40	0.14	0.24
tB	REPO	0.34	0.15	
tes	EXTF	0.61	0.43	0.56
oles	DECF	0.22	0.24	0.15
arial	ELAF	0.43	0.35	0.51
N <sup>2</sup>	ACTP	0.34	0.17	0.23
	DECA	0.24	0.11	
	COMF	0.22	0.18	
	TENU	0.35	0.17	0.23
	COLF			0.22
	Note de pain	0.33	0.22	0.35
	SECT	0.45	0.34	0.30
	COUP	0.39	0.23	0.23
	DVCL	0.33	0.11	0.31
	RECL	0.15	0.14	0.16
	DECL	0.24	0.38	0.21
	W	0.68	0.37	
pe s	G	0.63	0.39	
ble	Р	0.79	0.42	
aria	P/L	0.67	0.35	
Valv	L	0.60	0.40	
	Ie	0.79	0.53	
	Dureté	0.85	0.25	
bles	Hydratation	0.51	0.32	
Aut	Volume	0.41	0.41	0.37
V.S.V	Protéine (%)	0.37	0.41	0.43

Tableau 6 : Héritabilité des variables mesurées sur le Panel Variétal et capacités prédictives (r) estimées par validation croisée des modèles de sélection génomique

Variable	Calibration Panel Variétal	Calibration Panel Recherche
Teneur en protéines	0.13	0.38
Note totale de panification	0.07	0.26
Note de pain	0.08	0.33
Note de pâte	0.11	0.28
Volume	0.16	0.43
LISP	-0.11	-0.28
COLP	0.09	0.14
EXTP	0.05	0.18
ELAP	0.08	0.17
EXTF	0.09	0.27
DECF	0.1	0.25
DVCL	0.07	0.27
DECL	0.08	0.28
COUM	0.04	0.08

 Tableau 7 : Portabilité des calibrations GS. La capacité prédictive des équations de prédiction génomique calibrées sur le Panel Variétal a été évaluée en calculant la corrélation (r) entre valeurs prédites et observées sur le Panel Recherche, et vice versa

#### 3.5 - Influence des caractéristiques de la base de calibration sur la précision des modèles

Les Figures 3 et 4 illustrent l'effet du nombre de génotypes et de marqueurs sur la précision (r) des prédictions dans le Panel Recherche. On observe sur la Figure 3 que plus on augmente le nombre de génotypes dans la base de calibration, plus les prédictions sont précises. Il semble que même avec 2000 génotypes dans la base ont n'ai pas encore atteint de plateau dans les prédictions.

A contrario, la Figure 4 montre que l'on atteint un plateau dans les prédictions avec 2000 marqueurs pour la Note de Pain et la Note de Panification, et 5000 marqueurs pour la teneur en Protéine. Ces chiffres sont donc bien inférieurs au nombre de marqueurs disponibles dans cette étude (10000) ou disponibles

75

sur des puces couramment utilisées en blé (9K, 35K, 420K...).

Seules 3 variables ont été présentées pour des questions de lisibilité mais ces tendances se retrouvent pour tous les critères étudiés. Par ailleurs, ces observations sont en accord avec ce qui est reporté dans la littérature et avec des résultats obtenus sur un 3° Panel (Panel INRA-AO, résultats non inclus dans cet article).



Figure 3 : Capacité de prédiction (r) en fonction du Nombre d'Individus pris en compte dans la Calibration. Les variables présentées sont la teneur en protéine (%Proteines), la Note de Pain (Note pain) et la Note Totale de Panification (Note panif).



compte dans la Calibration. Les variables présentées sont la teneur en protéine (%Proteines), la Note de Pain (Note pain) et la Note Totale de Parification (Note parif).

La Figure 5, quant à elle, montre un important gain de précision lorsque les génotypes sont évalués dans 2 ou 3 environnements plutôt qu'un seul. Les environnements supplémentaires apportent un gain beaucoup moins important. Peu de génotypes ont été mesurées dans plus de trois environnements dans ce panel, d'où un léger décrochement de la courbe. Ces résultats sont conformes à la littérature et corroborés par les résultats obtenus sur le Panel INRA-AO (résultats non présentés).

Il semble donc qu'au global le Panel Recherche permette une calibration plus précise et plus robuste que le Panel Variétal, et ce alors que le set de marqueurs est le même pour les 2 panels, et qu'il y a beaucoup plus de données phénotypiques pour la Panel Variétal que pour le Panel Recherche (près de 3800 observations contre 2300). Comme les données Phénotypiques sont vraisemblablement plus précises pour le Panel Variétal (du fait du nombre d'observations par génotype) que le pour le Panel Recherche, notre hypothèse est que la précision supérieure du Panel Recherche vient du fait qu'il est constitué de 10 fois plus de génotypes que le Panel Variétal (2300 contre près de 200).



Figure 5 : Précision de prédiction (r) en fonction du Nombre d'Observations (Environnements) par génotype dans la Calibration du Panel Variétal, toutes variables confondues.

#### 3.6 - Apport des covariables d'alvéographe dans les modèles GS

Sur le Panel Variétal, les résultats obtenus montrent que les capacités prédictives du modèle GBLUP sont relativement limitées pour les caractères principaux (Teneur de Protéines, Note de Pain, Note de Pâte, Note de Mie et Note Totale de Panification, 0.10 < r < 0.40). Cependant, il a été possible de tester sur ce panel l'effet de l'ajout dans le modèle GBLUP des covariables mesurées par l'alvéographe. Cet ajout a permis d'augmenter significativement les capacités prédictives des modèles (0.40 < r < 0.70, Figure 6).



Figure 6 : Apport des covariables d'alvéographe à la capacité prédictive des modèles GS. Les capacités prédictives ont été estimées par validations croisées sur le Panel Variétal pour la teneur en protéines, la note de pain, la note de pâte, la note de mie et la note totale de panification. Les 2 modèles comparés sont (1) un modèle GBLUP ("gs") et (2) un modèle GBLUP auquel est ajouté les variables issues du test alvéographe ("gs+alveo"). La ligne tiretée correspond à un coefficient de corrélation r=0.70.

Un tel accroissement des capacités prédictives grâce à l'utilisation de mesures indirectes de la qualité boulangère peut s'expliquer par le fait que ces mesures capturent des effets génétiques mais aussi des effets environnementaux. Le modèle GBLUP, qui utilise uniquement les données de génotypage, ne peut pas prendre en compte ces effets environnementaux. Ce résultat parait intéressant du point de vue du sélectionneur puisque la quantité de grains nécessaire à la réalisation de ce type de test est relativement réduite et pourrait permettre d'appliquer ce type de modèle de manière relativement précoce dans les schémas de sélection.

#### 3.7 - Tests d'associations

Au total, 49 associations significatives ont été identifiées sur le Panel Recherche et 57 sur le Panel Variétal (voir le détail en Annexe 3). La plupart des marqueurs identifiés étaient situés sur les chromosomes 1D (respectivement 14 et 10 marqueurs pour les Panels Variétal et Recherche), 1B (respectivement 8 et 5 marqueurs pour les Panels Variétal et Recherche) et 6B (respectivement cinq et quatre marqueurs pour les Panels Variétal et Recherche). L'analyse des colocalisations pour les deux panels a permis de mettre en évidence :

- une zone située entre 31 et 33cM sur le chromosome 1B associée au collant de la pâte au pétrissage et à l'extensibilité au façonnage.
- une zone située à 66cM sur le chromosome 1D associée au collant de la pâte au pétrissage, au déchirement et à l'extensibilité au façonnage.
- une zone située à 34cM sur le chromosome 6B associée au volume de la pâte.

Les zones situées sur les chromosomes 1B et 1D correspondent à la localisation des gènes de gluténines de haut poids moléculaire. Concernant la zone sur le 6B, un candidat possible est le gène de gliadine Gli-B2 bien que cela n'ait pu être démontré dans cette étude. De manière générale, la plupart des zones identifiées dans un panel diffèrent de celles identifiées dans l'autre panel, à l'exception des trois cas mentionnés cidessus.

# 4. Conclusion : protocole de calibration GS pour la qualité boulangère

A l'issue de ce projet, nous sommes en mesure de proposer le protocole suivant pour créer et utiliser une calibration pour la qualité boulangère des blés :

- Taille de la base de calibration : ce point est crucial pour s'assurer d'une bonne calibration. Nos résultats montrent que plus de 2000 génotypes n'est pas encore suffisant pour prédire précisément les caractères principaux,
- Génotypage : quelques milliers de marqueurs SNP (2000 à 5000), bien répartis sur le génome, semblent suffisant.
- Phénotypage : ce point est aussi crucial pour s'assurer d'une bonne calibration. Il est conseillé de s'intéresser à tous les critères du test BIPEA (et pas seulement aux notes globales), car tous peuvent présenter un intérêt et être plus ou moins bien prédits. Compte tenu des risques d'accidents climatiques et de l'importance des interactions GxE, un réseau d'essai multilocal et multi-annuel doit être envisagé, en prenant en compte un fort risque de perte d'essais. Contrairement au choix fait dans cette étude, nous recommandons d'observer au moins 2 à 3 fois chaque génotype (dans des environnements différents). Enfin, il faut s'assurer d'un bon niveau de variabilité pour les caractères, et donc ne pas utiliser uniquement des individus « bons » en termes de qualité.
- Nous recommandons aussi d'inclure les résultats d'alvéographe, ou tout autre test indirect de la qualité boulangère, en tant que covariables quand cela est possible, car cela améliore significativement la précision des modèles de GS.

Enfin, quant à l'utilisation de ces prédictions dans les programmes de sélection, nous pensons que la précision observée dans cette étude ne permet qu'un tri grossier. Si ce tri intervient dans les stades précoces de la sélection, il est conseillé d'appliquer une pression de sélection modérée. Dans les stades plus avancés, les mesures indirectes de la qualité boulangère, tel l'alvéographe, peuvent être combinées aux données de génotypage pour obtenir une prédiction beaucoup plus précise de la qualité boulangère.

Tous nos remerciements aux partenaires, stagiaires et collègues ayant contribué à ce projet : Gilles Charmet, Magali Thierry, Pierre Colin, Antonin Galien, Nicolas Perardel, Benoît Méléard, Céline Duque, Sébastien Deshayes, Valérie Herteman et David Gouache. Annexe 1 : Liste des composantes de la note de panification selon l'étape dans la fabrication du pain et selon la catégorie (définie par la note principale correspondante), et détail du calcul des notes.

Catégorie (Note Principale)	Туре	Observations	Abréviations
		Rapidité lissage	LISP
		Collant	COLP
	DETDISCACE	Consistance	CONP
	PETRISSAGE	Extensibilité	EXTP
		Elasticité	ELAP
		Relâchement	RELP
	POINTAGE	Détente relâchement	REPO
Note de pâte		Extensibilité	EXTF
50 C	EACONNAGE	Déchirement	DECF
	FACONNAGE	Elasticité	ELAF
		Collant	COLF
	ADDRET	Activité fermentaire	ACTP
	APPRET	Déchirement	DECA
	MISE ALL FOUR	Collant	COMF
	MISE AU FOUR	Tenue	TENU
		Section	SECT
	ASPECT DU DAIN	Couleur	COUP
	ASPECT DU PAIN	Epaisseur	EPAP
Note de pain		Croustillant	CROP
		Développement	DVCL
	COUP DE LAME	Régularité	RECL
		Déchirement	DECL
	ASPECT DE LA MIE	Couleur	COUM
		Souplesse	SOUM
	TEXTURE	Elasticité	ELAM
Note de mie		Collant	COLM
		Régularité	REAL
	ALVEOLAGE	Epaisseur	EPAL
		Saveur et arôme	FLAV

Formules de calcul des notes principales à partir de composantes. Le coefficient est attribué en fonction de la note du critère pondérant (en gras dans la formule).

#### Note de pâte =

(0.5×LISP + **COLP**×0.5 + EXTP×0.5 + ELAP×0.5 + **RELP**+0.5)×coef(\*) + REPO + (**EXTF**×0.5 + DECF×0.5 + ELAF×0.5 + COLF)×coef(\*) + (ACTP×0.5 + DECA×0.5) + (**COMF** + **TENU**×2)×coef(\*)

#### Note de pain =

(SECT + COUP×2 + EPAP×0.5 + CROP×0.5 + **DVCL** + RECL + DECL)×coef (\*) + VOL\*\* Note de mie =

```
(SOUM×2 + ELAM + COLM) + (REAL + EPAL×2 + FLAV×3)
```

*Note Totale (0-300) =* 

Note de pâte(0-100) + Note de pain(0-100) + Note de mie(0-100)

Note du critère pondérant (en gras)	10	7	4	1
coef(*) correspondant	1	0.75	0.5	0.25

\*\* Un score basé sur le Volume du Pain (VOL) est utilisé dans ce calcul.

#### Annexe 2 : Liste des 192 variétés du Panel Variétal.

ABAQUE, ACCOLADE, ACCOR, ACCROC, ACIENDA, ACOUSTIC, ADAGIO, ADEQUAT, ADHOC, AEROBIC, AGAPE, AGRESTIS, AGUILA, AKRATOS, ALCAZAR, ALDRIC, ALIGATOR, ALIXAN, ALIZEO, ALLEZY, ALTAMIRA, ALTIGO, ALVAREZ, AMADOR, AMBELLO, AMBITION, AMUNDSEN, ANDALOU, ANDINO, ANTONIUS, APACHE, APRILIO, ARACK, ARAMIS, ARDELOR, AREZZO, ARISTOTE, ARKEOS, ARLEQUIN, AROBASE, ASDECOEUR, ASTUCE, ATHLON, ATTITUDE, ATTLASS, AUBUSSON, AUDI, AUTAN, AUTENTIC, AVANTAGE, AXIMA, AZIMUT, AZZERTI, AZZURO, BAGATELLE007, BAROK, BATTANT, BERMUDE, BIANCOR, BOISSEAU, BOKARO, BOREGAR, BOSPHOR, BRENTANO, BUENNO, CAMPERO, CAPHORN, CARIBOU, CARRE, CCBINEDIT, CCBINGENIO, CCBPREFERENCE, CELESTIN, CENTENAIRE, CEZANNE, CHAGALL, CHARGER, CHEVALIER, CHEVRON, CIENTO, COMODOR, COMPIL, CONTREFOR, CORDIALE, COURTOT, CROISADE, DIALOG, DINOSOR, EMERALD, EPHOROS, EPIDOC, EUCLIDE, EXELCIOR, EXOTIC, EXPERT, FARINELLI, FARMEUR, FIORETTO, FLAMENKO, FLAUBERT, FLUOR, FOLKLOR, FORBAN, FORBLANC, FOURMI, GALACTIC, GALIBIER, GALOPAIN, GARANTUS, GARCIA, GAUGAIN, GLADIATOR, GONCOURT, GRAINDOR, GRETHEL, GULLIVER, HATTRICK, HAUSSMANN, HEKTO, HONNOR, HOURRA, HYBERY, HYMACK, HYSCORE, HYSTAR, HYSUN, HYXO, ILLICO, INCISIF, INNOV, INOUI, INSPIRATION, INSTINCT, INTACT, INTERET, IRIDIUM, ISENGRAIN, ISTABRAQ, JBASANO, JBDIEGO, KALANGO, KALYSTAR, KARILLON, KETCHUM, KLEBER, KORELI, LANCELOT, LIMES, LORD, MANAGER, MARCELIN, MATRIX, MAXWELL, MAXYL, MAYFAIR, MELKIOR, MENDEL, MENESTREL, MERCATO, MINOTOR, MIROIR, MUSIK, NIRVANA, NOGAL, NUAGE, NUCLEO, OAKLEY, OCTET, OLIVART, ORVANTIS, OXEBO, PAKITO, PALADAIN, PALEDOR, PAROLI, PEPIDOR, PERFECTOR, PERICLES, PHARE, PICADOR, PLAINEDOR, PLAYER, POTENZIAL, PR22R20, PR22R58, PREMIO, PREVERT, QUEBON, RACINE, RAGLAN, RAISON, RASPAIL, RAZZANO, RICHEPAIN, RIMBAUD, BIZE, BOCHEORT, RODRIGO, BOSARIO, ROYSSAC, RUSTIC, SAINTEX, SAMURAI, SANKARA, SCENARIO, SCOR, SEBASTO, SEI EKT, SEYRAC, SIRTAKI, SOBBEL, SOGOOD, SOISSONS, SOKAL, SOLEHIO, SOLLARIO, SOLUTION, SOPHYTRA, SORRIAL, SPONSOR, SUBLIM, SUBTIL, SUMO, SWKUNGSJET, SWEET, SWINGGY, SYALTEO, SYMATTIS, TAMARO, TIAGO, TIMBER, TIMING, TITLIS, TOISONDOR, TRAPEZ, TULIP, USKI, VALODOR, VELOURS, VERLAINE, VOLONTAIRE, WALLABY, WELLAND

Annexe 3 : Description des associations significatives détectées dans les 2 Panels.

ACT         WCI222199         5B         115.5         5.03         AA         TT         0.25         0.01         Rethreshe           COLF         WCI224844         1D         66.2         6.52         AA         GG         0.17         0.02         Rethreshe           COLF         WCI224844         1D         66.2         6.52         AA         GG         0.03         0.01         Rethreshe           COLF         WCI224874         AA         4.4         4.82         0.02         Rothorehe         Month           COLP         WCI21873         1B         28.1         4.36         CC         TT         0.96         0.11         Rethreshe           COLP         WCI224874         1B         33.8         6.07         7.9         AA         GG         0.35         0.02         Rethreshe         -           COLP         WCI224864         1D         0.62         11.07         AA         G6         0.33         0.03         Rethreshe         -           COLP         WCI224863         1B         93.9         4.32         AA         G6         0.11         Rethreshe         -           COLP         WCI224864         1D	Variable	Marqueur	Chrom.	Position (cM)	-log(p)	Allèle référence	Allèle alternatif	Effet Allélique	R2	Panel	Colocalisation
COLF         W02218089         18         92.7         4.86         CC         TI         0.14         0.01         Recherche           COLF         W02221802         4A         98.1         4.78         AA         CC         -0.17         0.02         Recherche           COLF         W02221802         4A         98.1         4.78         AA         CC         -0.15         0.01         Recherche           COLP         W02218131         A         4.4         4.82         CC         TI         -0.86         0.11         Warld           COLP         W02218765         18         0.38         6.07         66         TI         -0.86         0.11         Warld         *           COLP         W0222483         18         93.9         4.32         AA         CC         TI         0.86         0.11         Warld         *           COLP         W0222483         18         93.9         4.32         AA         CC         TI         0.86         0.11         Warld         *           COLP         W0222483         18         93.9         4.32         AA         CC         TI         0.32         0.31         Retherabe	ACTP	WC0222309	5B	115.5	5.03	AA	Π	0.25	0.01	Recherche	
COLF         WC022484         10         66.2         6.5.2         AA         66         0.17         0.02         Recherche           COLF         WC0222483         6A         72         4.15         AA         CC         0.17         0.02         Recherche           COLP         WC022483         1A         4.4         4.82         CC         66         0.20         0.01         Recherche           COLP         WC022483         1A         4.4         4.82         CC         66         0.20         0.02         Recherche           COLP         WC0224863         1B         33.8         6.07         66         11         0.98         0.1         Wariela         *           COLP         WC0224863         1B         33.9         4.32         AA         0.62         -0.16         0.01         Recherche         *           COLP         WC0224863         1B         93.9         4.32         AA         0.62         0.13         Wariela         *           COLP         WC0224863         1B         0.52         5.18         AA         0.62         0.17         0.01         Recherche           COLP         WC02226867 <t< th=""><th>COLF</th><th>WC0216089</th><th>1B</th><th>92.7</th><th>4.86</th><th>CC</th><th>Π</th><th>0.14</th><th>0.01</th><th>Recherche</th><th></th></t<>	COLF	WC0216089	1B	92.7	4.86	CC	Π	0.14	0.01	Recherche	
COLF         WC0221928         4A         98.1         4.78         AA         CC         -0.15         0.01         Recherche           COLF         WC0218313         1A         4.4         4.82         CC         66.6         0.09         0.01         Recherche           COLP         WC0218765         1B         30.3         67.7         AA         CC         0.11         Variability         *           COLP         WC0218767         1B         33.3         7.79         AA         CC         0.16         0.01         Recherche         *           COLP         WC02218674         1B         33.3         7.79         AA         CC         0.16         0.01         Recherche         *           COLP         WC02224663         1D         66.2         6.89         CC         TT         0.66         0.11         Variability         *           COLP         WC0222468         1D         66.2         18.89         CC         TT         0.36         0.31         Resherate         *           COMP         WC0222463         24.8         4.48         4.60         0.71         0.35         0.31         Marietal           COMP	COLF	WC0224844	1D	66.2	8.52	AA	GG	0.17	0.02	Recherche	
COLF         W02022489         6A         72         4 15         AA         6G         0.01         Perchandre           COLF         W02018779         1B         26.1         4.38         CC         T         0.86         0.12         Variable           COLP         W02018776         1B         26.1         4.38         CC         T         0.86         0.12         Variable           COLP         W02018765         1B         30.3         6.07         GG         TT         0.86         0.11         Variable         *           COLP         W02224664         1D         66.2         11.07         AA         CG         -0.16         0.01         Restructure         *           COLP         W02224669         1D         66.2         7.16         7.66         CC         TT         0.06         0.11         Variable           COMF         W02224669         1D         66.2         7.66         CC         TT         0.93         0.31         Variable           COMF         W02221665         36         0.03         4.22         AA         0.66         0.11         Variable           COMF         W02221657         3A         <	COLF	WC0221092	4A	98.1	4.78	AA	CC	-0.15	0.01	Recherche	
COLP         W02021873         1A         4.4         4.82         CC         GG         0.02         Rescrict           COLP         W02018776         1B         38.8         6.07         GG         1T         0.96         0.12         Variabil           COLP         W02018764         1B         33.3         7.79         AA         GG         0.36         0.01         Variabil           COLP         W02224684         1D         66.2         11.07         AA         GG         0.03         Resherche         *           COLP         W02224684         1D         66.2         1.16         0.66         0.11         Variabil         *           COLP         W02224689         1D         66.2         7.65         CC         TT         0.38         0.01         Variabil           COMF         W0222169         A         2.48         4.80         AA         GG         0.33         0.01         Variabil           COMF         W02221697         3.4         140.8         4.80         AA         TT         0.97         0.05         Variabil           COMF         W0221692         2.A         21.9         4.80         GG <td< th=""><th>COLF</th><th>WC0222489</th><th>6A</th><th>72</th><th>4.15</th><th>AA</th><th>GG</th><th>0.09</th><th>0.01</th><th>Recherche</th><th></th></td<>	COLF	WC0222489	6A	72	4.15	AA	GG	0.09	0.01	Recherche	
COLP         W00218765         1B         26.1         4.38         CC         TT         -0.88         0.12         Warktal           COLP         W00218676         1B         30.8         607         66         TT         0.06         0.11         Warktal           COLP         W00224681         1B         93.9         4.32         AA         CC         -0.16         0.01         Recherche           COLP         W00224868         1D         66.2         6.88         CC         TT         0.68         0.11         Warktal           COLP         W00224869         1D         66.2         7.85         CC         TT         0.68         0.11         Warktal           COMF         W00224869         1D         66.2         7.85         CC         TT         0.93         0.31         Warktal           COMF         W00228163         5.4         10.18         4.50         AA         TG         0.19         Varietal           COMF         W0022865         38         60.9         4.32         AA         66         0.11         Pacherbe           COUM         W0021802         7A         25.4         AA         66         0.1	COLP	WC0218313	1A	4.4	4.82	CC	GG	0.20	0.02	Recherche	
COLP         W0021876         1B         30.8         6.07         6.66         TT         0.96         0.1         Vanielal           COLP         W00218674         1B         33.3         7.79         AA         66         0.35         0.02         Recherche         •           COLP         W00224463         1B         93.3         4.32         AA         CC         -0.16         0.01         Recherche         •           COLP         W00224683         1D         66.2         6.89         CC         TT         0.32         0.11         Vanietal         •           COMF         W00228175         6A         24.8         4.58         AA         6G         -0.11         Vanietal           COMF         W00228175         6A         24.8         4.60         AA         TT         0.57         O.99         Vanietal           COMF         W00228172         2A         21.9         4.80         6G         TT         0.11         Vanietal           COMM         W00218202         2A         21.9         4.80         6A         TT         0.23         Vanietal           COUM         W00217292         74         4.82         A	COLP	WC0218779	1B	26.1	4.36	CC	Π	-0.86	0.12	Variétal	
COLP         WC0218674         18         93.3         7.79         AA         GG         0.35         0.02         Pacherche           COLP         WC0224843         1B         93.9         4.32         AA         CC         -0.16         0.01         Pacherche         *           COLP         WC0224849         1D         66.2         6.89         CC         TT         0.66         0.11         Variatial         *           COLP         WC0224889         1D         66.2         7.85         CC         TT         0.32         0.14         Variatial           COMF         WC0224889         1D         66.2         7.85         CC         TT         0.32         0.14         Variatial           COMF         WC0225855         38         60.9         4.32         AA         66         0.61         0.09         Variatial           COMM         WC0225857         3A         140.8         4.80         AA         T         0.28         0.01         Recherche           COUM         WC022769375         3A         4.32         AA         CG         17         0.28         0.01         Recherche           COUM         WC02274877	COLP	WC0218765	1B	30.8	6.07	GG	Π	0.96	0.1	Variétal	*
COLP         W00224484         1D         66.2         11.07         AA         CC         -0.16         0.01         Recherche           COLP         W00224869         1D         66.2         11.07         AA         6G         0.23         0.03         Recherche         *           COLP         W00224869         1D         66.2         7.65         CC         TT         0.66         0.11         Warfelal           COMF         W00224869         1D         66.2         7.65         CC         TT         0.32         0.14         Variefal           COMF         W00225687         38         60.9         4.32         AA         6G         0.61         0.09         Variefal           COMP         W00215655         38         60.9         4.32         AA         6G         0.61         0.09         Variefal           COUM         W00215920         2A         21.9         4.90         GG         TT         0.23         0.01         Recherche           COUM         W00217329         7A         158.7         5.84         AA         6G         0.11         Netherche           COUM         W00217003         7B         83.7	COLP	WC0218674	1B	33.3	7.79	AA	GG	0.35	0.02	Recherche	*
COLP         WC0224844         10         66.2         11.07         AA         6G         0.23         0.03         Pecherche         *           COLP         WC0224899         1D         66.2         5.8         96.2         5.8         AA         6G         0.11         Varietal         *           COMF         WC0224899         1D         66.2         7.65         CC         TT         0.32         0.14         Varietal           COMF         WC0228176         6A         24.8         4.58         AA         6G         0.05         Varietal           COMF         WC0228178         6A         101.8         4.30         CC         TT         0.05         Varietal           COMP         WC0228587         3A         140.3         4.60         AA         TT         0.23         0.01         Pectrothe           COUM         WC0217329         7A         158.7         5.44         AA         CG         0.19         0.01         Pectrothe           COUM         WC0217329         7A         158.7         5.44         AA         CG         0.28         0.01         Pectrothe           COUM         WC021737         7B	COLP	WC0224463	1B	93.9	4.32	AA	CC	-0.16	0.01	Recherche	
COLP         W00224899         1D         662         5.18         AA         CC         TT         0.86         0.11         Varietal           COMF         W00224899         1D         662         7.65         CC         TT         0.32         0.14         Varietal           COMF         W00223488         6A         101.8         4.30         CC         TT         0.19         0.65         Varietal           COMF         W00223687         3A         140.8         4.60         AA         TT         0.57         0.09         Varietal           COMF         W0021655         3B         60.9         4.32         AA         6G         0.61         0.09         Varietal           COUM         W0021655         3B         60.9         4.32         AA         CG         0.10         Recherche           COUM         W00217329         7A         158.7         5.84         AA         CG         0.12         0.01         Recherche           COUM         W0021703         7B         83.7         14.13         AA         CG         -0.28         0.01         Recherche           COUM         W00214073         3B         14.31	COLP	WC0224844	1D	66.2	11.07	AA	GG	0.23	0.03	Recherche	*
COLP         W00227883         58         96.2         5.18         AA         GG         1.11         0.11         Varietal           COMF         W00228175         6A         24.8         4.58         AA         GG         0.32         0.14         Varietal           COMF         W00228186         GA         101.8         4.30         CC         TT         0.19         0.05         Varietal           COMF         W00228687         3A         140.8         4.30         CC         TT         0.17         0.05         Varietal           COMP         W00219202         2A         21.9         4.90         GG         TT         0.23         0.01         Recherche           COUM         W00217320         7A         45.6         4.03         AA         TT         0.28         0.01         Recherche           COUM         W00217329         7A         158.7         5.84         AA         GG         0.11         Recherche           COUM         W0021703         7B         83.7         14.13         AA         GG         0.01         Recherche           COUM         W00216424         3D         37.9         5.71         AA	COLP	WC0224869	1D	66.2	6.89	CC	Π	0.66	0.11	Variétal	*
COMF         W002248175         6Å         24.8         4.58         AA         GG         -0.35         0.11         Varietal           COMF         W00228175         6Å         101.8         4.30         CC         TT         -0.19         0.05         Varietal           COMF         W00225867         3Å         140.8         4.60         AA         TT         -0.19         0.05         Varietal           COMP         W0022687         3Å         140.8         4.60         AA         TT         -0.23         0.01         Recherche           COUM         W00221902         2A         21.9         4.90         GG         TT         -0.23         0.01         Recherche           COUM         W00217329         7A         158.7         5.84         AA         GG         0.12         0.01         Recherche           COUM         W00217329         7A         158.7         5.84         AA         GG         -2.52         0.13         Varietal           COUM         W00214377         38         14.13         AA         GG         -2.52         0.13         Varietal           COUP         W00216424         30         3.7.9         5	COLP	WC0227693	5B	96.2	5.18	AA	GG	1.11	0.11	Variétal	
COMF         WC0228148         6A         101.8         4.30         CC         TT         -0.19         0.05         Variétal           COMF         WC0228148         6A         101.8         4.30         CC         TT         -0.19         0.05         Variétal           COMP         WC0226687         3A         140.8         4.80         GG         TT         0.23         0.01         Recherche           COUM         WC0228894         7A         45.6         4.03         AA         CC         -0.22         0.01         Recherche           COUM         WC0217329         7A         158.7         5.84         AA         GG         0.12         0.01         Recherche           COUM         WC0217329         7A         158.7         5.84         AA         GG         0.13         Variétal           COUM         WC0214377         38         143.1         AA         GG         -2.52         0.03         Recherche           COUP         WC0216424         30         3.7         5.71         AA         GG         2.89         0.09         Variétal           COUP         WC0221688         5.A         117.7         4.56         AA<	COMF	WC0224869	1D	66.2	7.65	CC	Π	0.32	0.14	Variétal	
COMF         WC0228687         3A         140.8         4.80         AA         TT         -0.19         0.05         Variètal           CONP         WC0226687         3A         140.8         4.60         AA         GG         0.61         0.09         Variètal           COUM         WC0216555         3B         60.9         4.32         AA         GG         0.61         0.09         Variètal           COUM         WC02179202         2A         21.9         4.90         GG         TT         0.23         0.01         Recherche           COUM         WC0217032         7A         158.7         5.84         AA         GG         0.01         Recherche           COUM         WC0217032         7B         13.7         14.13         AA         GG         -0.28         0.03         Recherche           COUM         WC0214377         3B         143.1         4.56         CC         GG         -2.52         0.13         Variètal           COUP         WC0214377         3B         143.1         4.56         CC         GG         -2.52         0.13         Variètal           COUP         WC0214377         3B         143.1         4.5	COMF	WC0228175	6A	24.8	4.58	AA	GG	-0.35	0.1	Variétal	
CONP         WC0226857         3A         140.8         4.60         AA         TT         0.57         0.09         Varietal           CONP         WC0226555         3B         60.9         4.32         AA         GG         0.61         0.09         Varietal           COUM         WC02269202         22.4         21.9         4.90         GG         TT         0.28         0.01         Recherche           COUM         WC0227829         7A         158.7         5.84         AA         GG         0.01         Recherche           COUM         WC0217329         7A         158.7         5.84         AA         GG         0.01         Recherche           COUM         WC0217329         7A         158.7         5.84         AA         GG         0.01         Recherche           COUM         WC0214377         3B         143.1         4.56         CC         GG         -2.52         0.01         Recherche           COUP         WC0216424         3D         37.9         5.71         AA         GG         2.69         0.09         Varietal           COUP         WC0216424         3D         37.9         5.71         AA         GG <th>COMF</th> <th>WC0228148</th> <th>6A</th> <th>101.8</th> <th>4.30</th> <th>CC</th> <th>Π</th> <th>-0.19</th> <th>0.05</th> <th>Variétal</th> <th></th>	COMF	WC0228148	6A	101.8	4.30	CC	Π	-0.19	0.05	Variétal	
COMP         WO0216555         38         60.9         4.32         AA         GG         0.61         0.09         Variétal           COUM         WC0219202         2A         21.9         4.90         GG         TT         0.23         0.01         Recherche           COUM         WC0217442         7A         128         4.20         AA         CC         -0.20         0.01         Recherche           COUM         WC0217429         7A         158.7         5.84         AA         GG         0.19         0.01         Recherche           COUM         WC021703         7B         83.7         14.13         AA         GG         -0.28         0.03         Recherche           COUP         WC0216424         3D         37.9         5.71         AA         GG         3.11         0.12         Variétal           COUP         WC0218314         1A         7.8         4.67         CC         TT         0.10         0.01         Recherche           DECF         WC02218446         1D         66.2         12.07         AA         GG         -0.41         0.06         Variétal           DECF         WC0224844         1D         66.2	CONP	WC0225687	ЗA	140.8	4.60	AA	Π	0.57	0.09	Variétal	
COUM         WC0219202         2A         21.9         4.90         GG         TT         0.23         0.01         Recherche           COUM         WC021742         TA         45.6         4.03         AA         TT         0.28         0.01         Recherche           COUM         WC0217429         TA         158.7         5.84         AA         GG         0.19         0.01         Recherche           COUM         WC0217329         TA         158.7         5.84         AA         GG         0.19         0.01         Recherche           COUM         WC021703         7B         83.7         14.13         AA         GG         -0.28         0.03         Recherche           COUP         WC0216424         3D         37.9         5.71         AA         GG         3.11         0.12         Varietal           COUP         WC021668         5A         117.7         4.56         AA         GG         2.69         0.03         Recherche           DECF         WC0224846         1D         66         10.12         GG         TT         -0.71         0.25         Varietal           DECF         WC0224845         5D         60.4	CONP	WC0216555	3B	60.9	4.32	AA	GG	0.61	0.09	Variétal	
COUM         WC0228994         7A         45.6         4.03         AA         TT         -0.28         0.01         Recherche           COUM         WC0217342         7A         128         4.20         AA         CC         -0.20         0.01         Recherche           COUM         WC0217329         7A         158.7         5.84         AA         GC         -0.20         0.01         Recherche           COUM         WC0217003         7B         83.7         14.13         AA         GG         -0.28         0.03         Recherche           COUP         WC0216424         30         37.9         5.71         AA         GG         -2.82         0.13         Variétal           COUP         WC0216424         30         37.9         5.71         AA         GG         2.69         0.09         Variétal           COUP         WC021568         5A         117.7         4.56         AA         GG         2.69         0.09         Variétal           DECF         WC0224844         1D         66.2         12.07         AA         GG         -0.20         0.03         Recherche         *           DECF         WC0224844         1D <th>COUM</th> <th>WC0219202</th> <th>2A</th> <th>21.9</th> <th>4.90</th> <th>GG</th> <th>Π</th> <th>0.23</th> <th>0.01</th> <th>Recherche</th> <th></th>	COUM	WC0219202	2A	21.9	4.90	GG	Π	0.23	0.01	Recherche	
COUM         WC0217442         7A         128         4.20         AA         CC         -0.20         0.011         Recherche           COUM         WC0217329         7A         158.7         5.84         AA         GG         0.11         Recherche           COUM         WC0221703         7B         83.7         14.13         AA         GG         -0.12         0.01         Recherche           COUP         WC0214377         3B         143.1         4.56         CC         GG         -2.52         0.13         Varietal           COUP         WC0216424         3D         37.9         5.71         AA         GG         3.11         0.12         Varietal           DECF         WC021568         5A         11.7.7         4.56         AA         GG         3.11         0.10         0.01         Recherche           DECF         WC0218642         1D         66         10.12         GG         1T         -0.71         0.25         Varietal           DECF         WC0228484         1D         66         10.12         AA         GG         -0.41         0.06         Varietal           DECF         WC0228484         1D         66.2 <th>COUM</th> <th>WC0228994</th> <th>7A</th> <th>45.6</th> <th>4.03</th> <th>AA</th> <th>Π</th> <th>-0.28</th> <th>0.01</th> <th>Recherche</th> <th></th>	COUM	WC0228994	7A	45.6	4.03	AA	Π	-0.28	0.01	Recherche	
COUM         WC0217329         7A         158.7         5.84         AA         GG         0.19         0.01         Recherche           COUM         WC0223874         7B         79.4         4.12         CC         TT         0.12         0.01         Recherche           COUM         WC0217003         7B         83.7         14.13         AA         6G         -0.28         0.03         Recherche           COUP         WC0217033         7B         83.7         14.13         AA         6G         -0.28         0.03         Recherche           COUP         WC02176244         3D         37.9         5.71         AA         6G         2.69         0.99         Variétal           DECF         WC0218424         1D         66         10.12         GT         T         0.10         0.01         Recherche           DECF         WC0218446         1D         66.2         12.07         AA         6G         -0.41         0.06         Variétal           DECF         WC02224844         1D         66.2         12.07         AA         6G         -0.41         0.06         Variétal           DECF         WC02224845         5D         60.	COUM	WC0217442	7A	128	4.20	AA	CC	-0.20	0.01	Recherche	
COUM         WC0223874         7B         79.4         4.12         CC         T1         0.12         0.01         Recherche           COUM         WC0217003         7B         83.7         14.13         AA         GG         -0.28         0.03         Recherche           COUP         WC0214377         3B         14.13         4.56         CC         GG         -2.52         0.13         Varietal           COUP         WC0216424         3D         37.9         5.71         AA         GG         3.11         0.12         Varietal           DECF         WC0221568         5A         117.7         4.56         AA         GG         2.69         0.09         Varietal           DECF         WC0221844         1D         66.2         12.07         AA         GG         -0.20         0.03         Recherche           DECF         WC0224844         1D         66.2         12.07         AA         GG         -0.41         0.06         Varietal           DECF         WC0224853         5D         60.4         4.27         AA         GG         -0.41         0.06         Varietal           DECL         WC02218078         1B         26.3	COUM	WC0217329	7A	158.7	5.84	AA	GG	0.19	0.01	Recherche	
COUM         WC0217003         7B         83.7         14.13         AA         GG         -0.28         0.03         Recherche           COUP         WC0214377         3B         143.1         4.56         CC         GG         -2.52         0.13         Variétal           COUP         WC0216424         3D         37.9         5.71         AA         GG         3.11         0.12         Variétal           COUP         WC02163246         1D         66         117.7         4.56         AA         GG         2.59         0.09         Variétal           DECF         WC0224846         1D         66.2         12.07         AA         GG         -0.71         0.25         Variétal           DECF         WC0224844         1D         66.2         12.07         AA         GG         -0.20         0.03         Recherche         *           DECF         WC0228494         2D         66.2         12.07         AA         GG         -0.48         0.08         Variétal           DECF         WC0228433         6B         76.2         4.28         GG         TT         0.20         0.01         Recherche           DECL         WC0218778 </th <th>COUM</th> <th>WC0223874</th> <th>7B</th> <th>79.4</th> <th>4.12</th> <th>CC</th> <th>Π</th> <th>0.12</th> <th>0.01</th> <th>Recherche</th> <th></th>	COUM	WC0223874	7B	79.4	4.12	CC	Π	0.12	0.01	Recherche	
COUP         WC0214377         38         143.1         4.56         CC         GG         -2.52         0.13         Variétal           COUP         WC0216424         30         37.9         5.71         AA         GG         3.11         0.12         Variétal           COUP         WC021683         5A         117.7         4.56         AA         GG         2.69         0.09         Variétal           DECF         WC0218314         1A         7.8         4.67         CC         TT         0.10         0.01         Recherche           DECF         WC0224844         1D         66.2         12.07         AA         GG         -0.41         0.06         Variétal         *           DECF         WC0224844         1D         66.2         12.07         AA         GG         -0.41         0.06         Variétal           DECF         WC0224845         5D         60.4         4.27         AA         GG         -0.48         0.08         Variétal           DECL         WC021809         1D         65.7         9.91         CC         GG         0.21         Variétal           DECL         WC021803         68         33.9	COUM	WC0217003	7B	83.7	14.13	AA	GG	-0.28	0.03	Recherche	
COUP         WC0216424         3D         37.9         5.71         AA         GG         3.11         0.12         Variétal           COUP         WC021568         5.A         117.7         4.56         AA         GG         2.69         0.09         Variétal           DECF         WC0218314         1A         7.8         4.67         CC         TT         0.10         0.01         Recherche           DECF         WC0224844         1D         66.2         12.07         AA         GG         -0.71         0.25         Variétal         *           DECF         WC022445         5D         60.4         4.27         AA         GG         -0.41         0.06         Variétal           DECF         WC0228383         68         76.2         4.28         GG         TT         1.07         0.12         Variétal           DECL         WC0218778         1B         26.3         5.40         AA         TT         1.07         0.12         Variétal           DECL         WC021878         1B         26.3         5.40         AA         TT         1.07         0.02         Recherche           DECL         WC0214309         1D	COUP	WC0214377	3B	143.1	4.56	CC	GG	-2.52	0.13	Variétal	
COUP         WC0221568         5A         117.7         4.56         AA         GG         2.69         0.09         Variétal           DECF         WC0218314         1A         7.8         4.67         CC         TT         0.10         0.01         Recherche           DECF         WC0224846         1D         66         10.12         GG         TT         -0.71         0.25         Variétal         *           DECF         WC0224844         1D         66.2         12.07         AA         GG         -0.20         0.03         Recherche         *           DECF         WC0224845         5D         60.4         4.27         AA         GG         -0.41         0.06         Variétal           DECF         WC0218778         1B         26.3         5.40         AA         TT         1.07         0.12         Variétal           DECL         WC0218778         1B         26.3         5.40         AA         TT         1.07         0.12         Variétal           DECL         WC0218778         1B         26.3         5.40         AA         TT         0.72         0.09         Variétal           DECL         WC0218703	COUP	WC0216424	3D	37.9	5.71	AA	GG	3.11	0.12	Variétal	
DECF         WC0218314         1A         7.8         4.67         CC         TT         0.10         0.01         Recherche           DECF         WC0224846         1D         66         10.12         GG         TT         -0.71         0.25         Variétal         *           DECF         WC0224844         1D         66.2         12.07         AA         GG         -0.20         0.03         Recherche         *           DECF         WC0224845         5D         60.4         4.27         AA         GG         0.41         0.06         Variétal           DECF         WC0228383         6B         76.2         4.28         GG         TT         0.20         0.01         Recherche           DECL         WC0218778         1B         26.3         5.40         AA         TT         1.07         0.12         Variétal           DECL         WC021809         1D         65.7         9.91         CC         GG         0.26         0.02         Recherche           DECL         WC0221803         6B         33.9         5.62         AA         GG         -0.33         0.12         Variétal           DECL         WC021377	COUP	WC0221568	5A	117.7	4.56	AA	GG	2.69	0.09	Variétal	
DECF         WC0224846         1D         66         10.12         GG         TT         -0.71         0.25         Variétal         *           DECF         WC0224844         1D         66.2         12.07         AA         GG         -0.20         0.03         Recherche         *           DECF         WC022485         5D         60.4         4.27         AA         GG         0.48         0.08         Variétal           DECF         WC0224853         6B         76.2         4.28         GG         TT         0.20         0.01         Recherche           DECL         WC0218778         1B         26.3         5.40         AA         TT         1.07         0.12         Variétal           DECL         WC0218809         1D         65.7         9.91         CC         GG         0.26         0.02         Recherche           DECL         WC0218093         6B         33.9         5.62         AA         GG         -0.31         0.01         Recherche           DECL         WC0213174         6A         9         4.41         CC         TT         -1.07         0.08         Variétal           DECL         WC0214377	DECF	WC0218314	1A	7.8	4.67	CC	Π	0.10	0.01	Recherche	
DECF         WC0224844         1D         66.2         12.07         AA         GG         -0.20         0.03         Recherche         *           DECF         WC0219669         2B         51.1         4.25         AA         GG         -0.41         0.06         Variétal           DECF         WC0228435         5D         60.4         4.27         AA         GG         0.48         0.08         Variétal           DECF         WC0228383         6B         76.2         4.28         GG         TT         0.20         0.01         Recherche           DECL         WC0218778         1B         26.3         5.40         AA         TT         1.07         0.12         Variétal           DECL         WC0218809         1D         65.7         9.91         CC         GG         0.26         0.02         Recherche           DECL         WC0218703         4B         40.6         4.85         GG         TT         0.72         0.09         Variétal           DECL         WC0213174         6A         9         4.41         CC         TT         1.07         0.08         Variétal           DECL         WC0214377         3B	DECF	WC0224846	1D	66	10.12	GG	TT	-0.71	0.25	Variétal	*
DECF         WC0219669         2B         51.1         4.25         AA         GG         -0.41         0.06         Variétal           DECF         WC0222445         5D         60.4         4.27         AA         GG         0.48         0.08         Variétal           DECF         WC0228383         6B         76.2         4.28         GG         TT         0.20         0.01         Recherche           DECL         WC0218778         1B         26.3         5.40         AA         TT         1.07         0.12         Variétal           DECL         WC021809         1D         65.7         9.91         CC         GG         0.26         0.02         Recherche           DECL         WC0228492         2D         140.3         4.71         AA         GG         -0.31         0.01         Recherche           DECL         WC0213174         6A         9         4.41         CC         TT         -1.07         0.08         Variétal           DECL         WC0213174         6A         9         4.41         CC         TT         1.09         0.06         Variétal           DECL         WC0213773         3B         143.1	DECF	WC0224844	1D	66.2	12.07	AA	GG	-0.20	0.03	Recherche	*
DECF         WC0222445         5D         60.4         4.27         AA         GG         0.48         0.08         Variétal           DECF         WC0228383         6B         76.2         4.28         GG         TT         0.20         0.01         Recherche           DECL         WC0218778         1B         26.3         5.40         AA         TT         1.07         0.12         Variétal           DECL         WC021809         1D         65.7         9.91         CC         6G         0.26         0.02         Recherche           DECL         WC0225492         2D         140.3         4.71         AA         GG         -0.31         0.01         Recherche           DECL         WC0214174         6A         9         4.41         CC         TT         -1.07         0.08         Variétal           DECL         WC0215903         6B         33.9         5.62         AA         GG         -0.93         0.12         Variétal           DECL         WC0214377         3B         143.1         5.07         CC         GG         -14.08         0.12         Variétal           DUR         WC0216424         3D         37.9	DECF	WC0219669	2B	51.1	4.25	AA	GG	-0.41	0.06	Variétal	
DECF         WC0228383         GB         76.2         4.28         GG         TT         0.20         0.01         Recherche           DECL         WC0218778         1B         26.3         5.40         AA         TT         1.07         0.12         Variétal           DECL         WC0218809         1D         65.7         9.91         CC         GG         0.26         0.02         Recherche           DECL         WC0225492         2D         140.3         4.71         AA         GG         -0.31         0.01         Recherche           DECL         WC0221413         4B         40.6         4.85         GG         TT         0.72         0.09         Variétal           DECL         WC0213174         6A         9         4.41         CC         TT         -1.07         0.08         Variétal           DECL         WC0214377         3B         143.1         5.07         CC         GG         -T1         1.09         0.06         Variétal           DUR         WC0214377         3B         143.1         5.07         CC         GG         2.49         0.09         Variétal           DUR         WC0216424         3D	DECF	WC0222445	5D	60.4	4.27	AA	GG	0.48	0.08	Variétal	
DECL         WC0218778         1B         26.3         5.40         AA         IT         1.07         0.12         Variétal           DECL         WC0218809         1D         65.7         9.91         CC         GG         0.26         0.02         Recherche           DECL         WC0225492         2D         140.3         4.71         AA         GG         -0.31         0.01         Recherche           DECL         WC0221413         4B         40.6         4.85         GG         TT         0.72         0.09         Variétal           DECL         WC0213174         6A         9         4.41         CC         TT         -1.07         0.08         Variétal           DECL         WC0215903         6B         33.9         5.62         AA         GG         -0.93         0.12         Variétal           DECL         WC0228288         6B         52.2         4.09         GG         TT         1.09         0.06         Variétal           DUR         WC0216424         3D         37.9         10.50         AA         GG         25.79         0.22         Variétal           DVCL         WC02165903         6B         33.9	DECF	WC0228383	6B	76.2	4.28	GG	Π	0.20	0.01	Recherche	
DECL         WC0218809         1D         65.7         9.91         CC         GG         0.26         0.02         Recherche           DECL         WC0225492         2D         140.3         4.71         AA         GG         -0.31         0.01         Recherche           DECL         WC0221413         4B         40.6         4.85         GG         TT         0.72         0.09         Variétal           DECL         WC0213174         6A         9         4.41         CC         TT         -1.07         0.08         Variétal           DECL         WC0215903         6B         33.9         5.62         AA         GG         -0.93         0.12         Variétal           DECL         WC0228288         6B         52.2         4.09         GG         TT         1.09         0.06         Variétal           DUR         WC0214377         3B         143.1         5.07         CC         GG         -14.08         0.12         Variétal           DVCL         WC0221622         5B         120.4         4.28         CC         GG         2.49         0.09         Variétal           DVCL         WC0228764         7A         92.1	DECL	WC0218778	1B	26.3	5.40	AA	Π	1.07	0.12	Variétal	
DECL         WC0225492         2D         140.3         4./1         AA         GG         -0.31         0.01         Recherche           DECL         WC0221413         4B         40.6         4.85         GG         TT         0.72         0.09         Variétal           DECL         WC0213174         6A         9         4.41         CC         TT         -1.07         0.08         Variétal           DECL         WC0215903         6B         33.9         5.62         AA         GG         -0.93         0.12         Variétal           DECL         WC0228288         6B         52.2         4.09         GG         TT         1.09         0.06         Variétal           DUR         WC0214377         3B         143.1         5.07         CC         GG         -14.08         0.12         Variétal           DUR         WC0216424         3D         37.9         10.50         AA         GG         2.49         0.09         Variétal           DVCL         WC022162         5B         120.4         4.28         CC         GG         2.49         0.09         Variétal           DVCL         WC0215903         6B         33.9	DECL	WC0218809	10	65.7	9.91	CC	GG	0.26	0.02	Recherche	
DECL         WC0221413         4B         40.6         4.85         GG         11         0.72         0.09         Varietal           DECL         WC0213174         6A         9         4.41         CC         TT         -1.07         0.08         Varietal           DECL         WC0215903         6B         33.9         5.62         AA         GG         -0.93         0.12         Varietal           DECL         WC0228288         6B         52.2         4.09         GG         TT         1.09         0.06         Varietal           DUR         WC0214377         3B         143.1         5.07         CC         GG         -14.08         0.12         Varietal           DUR         WC0216424         3D         37.9         10.50         AA         GG         25.79         0.22         Varietal           DVCL         WC022162         5B         120.4         4.28         CC         GG         2.49         0.09         Varietal           DVCL         WC0215903         6B         33.9         4.06         AA         GG         -0.17         0.01         Recherche           ELAF         WC0218731         1B         30.8	DECL	WC0225492	2D	140.3	4./1	AA	GG	-0.31	0.01	Recherche	
DECL         WC0213174         6A         9         4.41         CC         IT         -1.07         0.08         Varietal           DECL         WC0215903         6B         33.9         5.62         AA         GG         -0.93         0.12         Varietal           DECL         WC0228288         6B         52.2         4.09         GG         TT         1.09         0.06         Varietal           DUR         WC0214377         3B         143.1         5.07         CC         GG         -14.08         0.12         Varietal           DUR         WC0216424         3D         37.9         10.50         AA         GG         25.79         0.22         Varietal           DVCL         WC0216424         3D         37.9         10.50         AA         GG         2.49         0.09         Varietal           DVCL         WC0215903         6B         33.9         4.06         AA         GG         1.89         0.11         Varietal           DVCL         WC0218731         1B         30.8         4.72         CC         TT         -3.15         0.06         Varietal           ELAF         WC0218674         1B         33.3	DECL	WC0221413	4B	40.6	4.85	GG		0.72	0.09	Varietal	
DECL         WC0215903         6B         33.9         5.62         AA         GG         -0.93         0.12         Varietal           DECL         WC0228288         6B         52.2         4.09         GG         TT         1.09         0.06         Varietal           DUR         WC0214377         3B         143.1         5.07         CC         GG         -14.08         0.12         Varietal           DUR         WC0216424         3D         37.9         10.50         AA         GG         25.79         0.22         Varietal           DVCL         WC022162         5B         120.4         4.28         CC         GG         2.49         0.09         Varietal           DVCL         WC0215903         6B         33.9         4.06         AA         GG         1.89         0.11         Varietal           DVCL         WC0218731         1B         30.8         4.72         CC         TT         -3.15         0.06         Varietal           ELAF         WC0218731         1B         30.8         4.09         CC         TT         -1.97         0.09         Varietal           ELAP         WC0218674         1B         33.3	DECL	W60213174	6A	9	4.41		11	-1.07	0.08	Varietal	
DECL         WC0222288         6B         52.2         4.09         GG         11         1.09         0.06         Varietal           DUR         WC0214377         3B         143.1         5.07         CC         GG         -14.08         0.12         Varietal           DUR         WC0216424         3D         37.9         10.50         AA         GG         25.79         0.22         Varietal           DVCL         WC022162         5B         120.4         4.28         CC         GG         2.49         0.09         Varietal           DVCL         WC0215903         6B         33.9         4.06         AA         GG         1.89         0.11         Varietal           DVCL         WC0218731         1B         30.8         4.72         CC         TT         -3.15         0.06         Varietal           ELAF         WC0218731         1B         30.8         4.72         CC         TT         -2.55         0.1         Varietal           ELAF         WC0218731         1B         30.8         4.09         CC         TT         -1.97         0.09         Varietal         *           ELAP         WC0218674         1B	DECL	WC0215903	6B CD	33.9	5.62	AA	<u> </u>	-0.93	0.12	Variétal	
DUR         WC0214377         3B         143.1         3.07         CC         GG         -14.06         0.12         Valietal           DUR         WC0216424         3D         37.9         10.50         AA         GG         25.79         0.22         Variétal           DVCL         WC0222162         5B         120.4         4.28         CC         GG         2.49         0.09         Variétal           DVCL         WC0215903         6B         33.9         4.06         AA         GG         1.89         0.11         Variétal           DVCL         WC0228764         7A         92.1         4.28         AA         GG         -0.17         0.01         Recherche           ELAF         WC0218731         1B         30.8         4.72         CC         TT         -3.15         0.06         Variétal           ELAF         WC0218674         1B         30.8         4.09         CC         TT         -1.97         0.09         Variétal         *           ELAP         WC0218674         1B         33.3         9.64         AA         GG         0.61         0.02         Recherche         *           ELAP         WC0228096	DEGL	WC0214277	08	142.1	4.09	GG	11	14.09	0.00	Varietal	
DUR         WC0216424         3D         37.9         10.30         AA         GG         23.79         0.22         Validation           DVCL         WC0222162         5B         120.4         4.28         CC         GG         2.49         0.09         Variétal           DVCL         WC0215903         6B         33.9         4.06         AA         GG         1.89         0.11         Variétal           DVCL         WC0228764         7A         92.1         4.28         AA         GG         -0.17         0.01         Recherche           ELAF         WC0218731         1B         30.8         4.72         CC         TT         -3.15         0.06         Variétal           ELAF         WC0218731         1B         30.8         4.09         CC         TT         -1.97         0.09         Variétal           ELAP         WC0218674         1B         33.3         9.64         AA         GG         0.61         0.02         Recherche         *           ELAP         WC0218674         1B         33.3         9.64         AA         GG         0.61         0.02         Recherche         *           ELAP         WC022869	DUR	WC0214377	JD D	143.1	0.07 10.50	00	66	-14.00	0.12	Variétal	
DVCL       WC0222102       3B       120.4       4.28       CC       GG       2.49       0.09       Validation         DVCL       WC0215903       6B       33.9       4.06       AA       GG       1.89       0.11       Variétal         DVCL       WC0228764       7A       92.1       4.28       AA       GG       -0.17       0.01       Recherche         ELAF       WC0218731       1B       30.8       4.72       CC       TT       -3.15       0.06       Variétal         ELAF       WC0218731       1B       30.8       4.72       CC       TT       -3.15       0.06       Variétal         ELAF       WC0218731       1B       30.8       4.09       CC       TT       -1.97       0.09       Variétal         ELAP       WC0218674       1B       33.3       9.64       AA       GG       0.61       0.02       Recherche       *         ELAP       WC022869       1D       66.2       7.83       CC       TT       0.39       0.02       Recherche       *         ELAP       WC0228096       6A       104       4.15       AA       CC       0.79       0.04       Variétal	DUR	WC02210424	3D 5P	37.9	10.00	AA	GG	20.79	0.22	Variétal	
DVCL         WC0213903         OB         33.3         4.00         AA         Odd         1.05         0.11         Validation           DVCL         WC0228764         7A         92.1         4.28         AA         GG         -0.17         0.01         Recherche           ELAF         WC0218731         1B         30.8         4.72         CC         TT         -3.15         0.06         Variétal           ELAF         WC0218731         1B         30.8         4.09         CC         TT         2.25         0.1         Variétal           ELAP         WC0218674         1B         33.3         9.64         AA         GG         0.61         0.02         Recherche         *           ELAP         WC0218674         1B         33.3         9.64         AA         GG         0.61         0.02         Recherche         *           ELAP         WC0224869         1D         66.2         7.83         CC         TT         0.39         0.02         Recherche           ELAP         WC0228096         6A         104         4.15         AA         CC         0.79         0.04         Variétal           ELAP         WC022869	DVCL	WC0222102	6R	22.0	4.20	00	GG	1.90	0.09	Variótal	
ELAF         WC0224869         1D         66.2         4.72         CC         TT         -3.15         0.06         Variétal           ELAF         WC0218731         1B         30.8         4.72         CC         TT         -3.15         0.06         Variétal           ELAF         WC0224869         1D         66.2         4.72         CC         TT         2.25         0.1         Variétal           ELAP         WC0218731         1B         30.8         4.09         CC         TT         -1.97         0.09         Variétal         *           ELAP         WC0218674         1B         33.3         9.64         AA         GG         0.61         0.02         Recherche         *           ELAP         WC0224869         1D         66.2         7.83         CC         TT         0.39         0.02         Recherche           ELAP         WC0228096         6A         104         4.15         AA         CC         0.79         0.04         Variétal           EPAL         WC0224869         1D         66.2         4.21         CC         TT         0.34         0.05         Variétal           EPAL         WC0219824         <	DVCL	WC0213903	70	02.1	4.00		GG	-0.17	0.11	Recherche	
ELAR         W60210731         TD         30.0         4.72         CC         TT         50.10         Varietal           ELAF         WC0224869         1D         66.2         4.72         CC         TT         2.25         0.1         Varietal           ELAP         WC0218731         1B         30.8         4.09         CC         TT         -1.97         0.09         Varietal         *           ELAP         WC0218674         1B         33.3         9.64         AA         GG         0.61         0.02         Recherche         *           ELAP         WC0224869         1D         66.2         7.83         CC         TT         0.39         0.02         Recherche         *           ELAP         WC0228096         6A         104         4.15         AA         CC         0.79         0.04         Variétal           EPAL         WC0224869         1D         66.2         4.21         CC         TT         0.34         0.05         Variétal           EPAL         WC0224869         1D         66.2         4.21         CC         TT         0.67         0.16         Variétal	FLAE	WC0218731	1R	30.8	4.20		TT	-3.15	0.01	Variátal	
ELAR         W00224003         ID         00.2         4.72         000         II         2.23         0.1         Validation           ELAP         WC0218731         1B         30.8         4.09         CC         TT         -1.97         0.09         Variétal         *           ELAP         WC0218674         1B         33.3         9.64         AA         GG         0.61         0.02         Recherche         *           ELAP         WC0224869         1D         66.2         7.83         CC         TT         0.39         0.02         Recherche         *           ELAP         WC0228096         6A         104         4.15         AA         CC         0.79         0.04         Variétal           EPAL         WC0224869         1D         66.2         4.21         CC         TT         0.34         0.05         Variétal           EPAL         WC0224869         1D         66.2         4.21         CC         TT         0.67         0.16         Variétal	ELAI ELAE	WC0210751	1D 1D	66.2	4.72	00	TT	2.25	0.00	Variétal	
ELAP         WC0218674         1B         33.3         9.64         AA         GG         0.61         0.02         Recherche         *           ELAP         WC0218674         1B         33.3         9.64         AA         GG         0.61         0.02         Recherche         *           ELAP         WC0224869         1D         66.2         7.83         CC         TT         0.39         0.02         Recherche           ELAP         WC0228096         6A         104         4.15         AA         CC         0.79         0.04         Variétal           EPAL         WC0224869         1D         66.2         4.21         CC         TT         0.34         0.05         Variétal           EPAL         WC0219824         2D         24.6         7.16         CC         TT         0.67         0.16         Variétal	FI ΔP	WC0218731	1R	30.2	4.00	00	ТТ	-1 07	0.1	Variétal	*
ELAP         WC0224869         1D         66.2         7.83         CC         TT         0.39         0.02         Recherche           ELAP         WC0224869         1D         66.2         7.83         CC         TT         0.39         0.02         Recherche           ELAP         WC0228096         6A         104         4.15         AA         CC         0.79         0.04         Variétal           EPAL         WC0224869         1D         66.2         4.21         CC         TT         0.34         0.05         Variétal           EPAL         WC0219824         2D         24.6         7.16         CC         TT         0.67         0.16         Variétal	FI ΔP	WC0218674	1B	33.3	9.64		GG	0.61	0.03	Recherche	*
ELAP         WC0228096         6A         104         4.15         AA         CC         0.79         0.04         Variétal           EPAL         WC0224869         1D         66.2         4.21         CC         TT         0.34         0.05         Variétal           EPAL         WC0224869         1D         66.2         4.21         CC         TT         0.67         0.16         Variétal	FIΔP	WC0224869	1D	66.2	7.83	00	Π	0.39	0.02	Recherche	
EPAL         WC0224869         1D         66.2         4.21         CC         TT         0.34         0.05         Variétal           EPAL         WC0219824         2D         24.6         7.16         CC         TT         0.67         0.16         Variétal	FIΔP	WC0228096	64	104	4 15	ΔΔ	CC.	0.00	0.02	Variétal	
FPAI         WC0219824         2D         24.6         7.16         CC         TT         0.67         0.16         Variated	FPAI	WC0224869	1D	66.2	4.21	CC	Π	0.73	0.05	Variétal	
	EPAL	WC0219824	2D	24.6	7.16	CC	Π	0.67	0.16	Variétal	

Variable	Marqueur	Chrom.	Position (cM)	-log(p)	Allèle référence	Allèle alternatif	Effet Allélique	R2	Panel	Colocalisation
EPAP	WC0224560	1B	86.4	4.24	AA	CC	-1.19	0.09	Variétal	
EPAP	WC0220154	ЗA	19.7	6.04	AA	CC	1.32	0.16	Variétal	
EXTF	WC0224841	1D	55.5	4.27	CC	TT	-0.29	0	Recherche	
EXTF	WC0224874	1D	65.2	5.72	GG	TT	1.69	0.12	Variétal	*
EXTF	WC0224869	1D	66.2	8.91	CC	TT	-0.50	0.02	Recherche	*
EXTF	WC0225282	2B	95.6	5.09	AA	GG	-0.73	0.01	Recherche	
EXTF	WC0227751	5B	36.6	4.98	AA	GG	2.45	0.13	Variétal	
G	WC0224846	1D	66	5.28	GG	TT	-1.45	0.13	Variétal	
G	WC0227324	5A	80.2	4.06	CC	TT	1.33	0.11	Variétal	
le	WC0224869	1D	66.2	13.13	CC	TT	4.81	0.21	Variétal	
LISP	WC0224869	1D	66.2	13.80	CC	TT	-0.31	0.05	Recherche	
MIE	WC0222070	5B	43.2	4.87	CC	TT	-2.17	0.1	Variétal	
NPAI	WC0216750	2B	14.9	4.84	CC	TT	-3.32	0.01	Recherche	
NPAI	WC0215903	6B	33.9	4.78	AA	GG	1.99	0.01	Recherche	
NPAI	WC0228733	7A	92.2	5.11	GG	TT	2.56	0.01	Recherche	
NPAI	WC0229499	7D	8.9	5.40	AA	GG	3.11	0.01	Recherche	
NPANI	WC0218674	1B	33.3	5.27	AA	GG	3.74	0.01	Recherche	
NPANI	WC0224844	1D	66.2	5.06	AA	GG	2.14	0.01	Recherche	
NPANI	WC0216750	2B	14.9	5.20	CC	TT	-5.08	0.02	Recherche	
NPANI	WC0229502	7D	7.9	4.62	GG	TT	-2.60	0.01	Recherche	
P.L	WC0224846	1D	66	4.70	GG	TT	0.30	0.1	Variétal	
PANIF	WC0224869	1D	66.2	4.30	CC	TT	8.36	0.05	Variétal	
PANIF	WC0219824	2D	24.6	5.84	CC	TT	14.08	0.12	Variétal	
PANIF	WC0221946	5A	22.5	4.43	AA	GG	-10.73	0.1	Variétal	
PATE	WC0218806	1B	26.9	5.12	AA	GG	-8.73	0.09	Variétal	
PATE	WC0224869	1D	66.2	7.26	CC	TT	6.28	0.12	Variétal	
PROT	WC0214404	2A	7	4.58	CC	GG	-0.13	0.01	Recherche	
PROT	WC0226542	4A	94.8	4.36	CC	TT	0.18	0.01	Recherche	
PROT	WC0215996	4A	141.5	5.43	GG	TT	-0.21	0.02	Recherche	
PROT	WC0226868	4B	54.5	12.32	AA	GG	0.04	0	Recherche	
PROT	WC0213051	4D	25.5	12.44	AA	GG	0.14	0.03	Recherche	
PROT	WC0227190	4D	30.2	4.72	AA	GG	-0.23	0.17	Variétal	
PROT	WC0227400	5A	175.8	4.43	CC	GG	0.13	0.02	Recherche	
PROT	WC0228421	UNM	49	4.44	CC	TT	0.03	0	Recherche	
REAL	WC0222969	6B	57.3	4.06	CC	GG	-0.91	0.12	Variétal	
RELP	WC0220871	4A	141.5	5.23	CC	TT	-0.92	0.13	Variétal	
RELP	WC0227399	5A	73.3	4.51	CC	TT	-0.24	0.01	Variétal	
REPO	WC0224869	1D	66.2	8.92	CC	TT	0.60	0.17	Variétal	
REPO	WC0225626	2D	24.6	5.93	CC	TT	-0.56	0.11	Variétal	
SECT	WC0218731	1B	30.8	4.66	CC	TT	-2.80	0.07	Variétal	
SECT	WC0218809	1D	65.7	4.89	CC	GG	1.92	0.12	Variétal	
SECT	WC0226602	4A	47.5	4.37	AA	CC	1.44	0.06	Variétal	
TENU	WC0224846	1D	66	7.49	GG	П	2.27	0.17	Variétal	
TENU	WC0219824	2D	24.6	4.34	CC	Π	2.13	0.11	Variétal	
VOL	WC0213718	3B	64.6	4.74	AA	GG	31.43	0.02	Kecherche	
VOL	WC0215903	6B	33.9	7.71	AA	GG	43.45	0.03	Recherche	*
VOL	WC0215903	6B	33.9	5.98	AA	GG	96.96	0.23	Variétal	*
VOL	WC0228764	/A	92.1	4.93	AA	GG	-38.79	0.02	Recherche	
VOL	WC0229499	/D	8.9	4.66	AA	GG	5/.8/	0.02	Kecherche	
VOLG	WC0216017	10	63.6	4.10	GG		11.56	0.01	Recherche	
VOLG	WC021/247	2B	34.3	4.63	CC		15.36	0.02	Kecherche	
VOLG	WC0215903	6B	33.9	9.09	AA	GG	22.43	0.04	Kecherche	
W	WC0225377	2B	42.6	4.12	AA	GG	-52.95	0.11	Variétal	

## Références bibliographiques

Blanc G., Charcosset A., Veyrieras J. B., Gallais A., and Moreau L. (2008). Marker-assisted selection efficiency in multiple connected populations: a simulation study based on the results of a QTL detection experiment in maize. Euphytica 161, 71–84.

Eathington, S. R., T. M. Crosbie, M. D. Edwards, R. S. Reiter, and J. K. Bull. (2007). Molecular Markers in a Commercial Breeding Program. Crop Sci. 47(Suppl3):S-154-S-163.

**Endelman, J.B.** (2011) Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. The Plant Genome Journal 4, 250.

Endelman J. B. and J.-L. Jannink (2012) Shrinkage estimation of the realized relationship matrix. G3 Genes Genomes Genet. 2: 1405–1413

Hayes B., J. Panozzo, C. Walker, A. Choy, S. Kant, et al. (2017) Accelerating wheat breeding for end-use quality with multi-trait genomic predictions incorporating near infrared and nuclear magnetic resonance-derived phenotypes. Theor. Appl. Genet. 130: 2505–2519.

Heffner E.L., Sorrells M.E., Jannink J-L (2009) Genomic Selection for Crop Improvement. Crop Science 49:1-12

Heffner E.L., Lorenz A.J., Jannink J-L. Sorrells M.E. (2010). Plant breeding with genomic selection: gain per unit time and cost. Crop Sci 50: 1681-1690.

Heffner E.L., Jannink J-L. Sorrells M.E. (2011a) Genomic Selection Accuracy using Multifamily Prediction Models in a Wheat Breeding Program. The Plant Genome 4:65–75.

Heffner E.L., Jannink J-L., Iwata H., Souza E., Sorrells M. (2011b) Genomic selection accuracy for grain quality traits in biparental wheat populations. Crop Sci 51:2597–2606.

Lande R, Thompson R (1990) Efficiency of Marker-Assisted Selection in the Improvement of Quantitative Traits. Genetics 124 : 743-756

**Méléard B.** (2010). Quelles évolutions qualitatives dans l'offre française de blé tendre ? Bilan de 15 années d'amélioration. Industries des Céréales, 170: 14-19.

Meuwissen THE, Hayes B, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps; Genetics 157:1819-1829

Michel S., M. Gallee, F. Löschenberger, H. Buerstmayr, and C. Kummer (2017) Improving the baking quality of bread wheat using rapid tests and genomics: The prediction of dough rheological parameters by gluten peak indices and genomic selection models. J. Cereal Sci. 77: 24–34.

Michel S., C. Kummer, M. Gallee, J. Hellinger, C. Ametz, et al. (2018) Improving the baking quality of bread wheat by genomic selection in early generations. Theor. Appl. Genet. 131: 477–493.

Oury FX, Chiron H, Faye A, gardet O, Giraud A, Heumez E, Rolland B, Rousset M, Trottet M, Charmet G (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.

Segura V., Vilhjálmsson B.J., Platt A., Korte A., Seren Ü., Long Q., Nordborg M. (2012) An efficient multi-locus mixed model approach for genome-wide association studies in structured populations. Nat Genet 44, 825–830.

Whittaker, J.C., R., Thompson, M.C. Denham. (2000) Markerassisted selection using ridge regression. Genet. Res. 75:249-252.

## [WEAB] Sélection Assistée par les Effecteurs fongiques de Résistances aux champignons pathogènes chez le blé

Marc-Henri LEBRUN<sup>\*1</sup>, Elsa BALLINI<sup>2</sup>, Guylaine BESNIER-HEBERT<sup>3</sup>, Ludovic BONHOMME<sup>4</sup>, Ruth BRYANT<sup>5</sup>, Florence CAMBON<sup>4</sup>, James COCKRAM<sup>6</sup>, Jean-Michel DELHAYE<sup>7</sup>, Aurélie DUCASSE<sup>2</sup>, Rowena DOWNIE<sup>6</sup>, Laure DUCHALAIS<sup>8</sup>, Sylvie DUTRIEZ<sup>9</sup>, Benoit FOUCAULT<sup>10</sup>, Pascal GIRAUDEAU<sup>11</sup>, Lilian GOUT<sup>1</sup>, Bruno GREZEZS-BESSE<sup>3</sup>, Delphine HOURCADE<sup>12</sup>, Gert KEMA<sup>13</sup>, Thomas KROJ<sup>2</sup>, Stéphane LAFARGE<sup>3</sup>, Thierry LANGIN<sup>4</sup>, Valérie LAURENT<sup>14</sup>, Jean-Benoit MOREL<sup>2</sup>, Richard OLIVER<sup>15</sup>, Jan PANEK<sup>16</sup>, Cyrille SAINTENAC<sup>4</sup>, Charles SNIJDERS<sup>17</sup>, Kar-Chun TAN<sup>15</sup>, Romain VALADE<sup>12</sup>

- 1 INRA AgroParis Tech BIOGER, Thiverval-Grignon, France
- 2 INRA BGPI, Montpellier, France
- 3 BIOGEMMA, Clermont-Ferrand, France
- 4 INRA GDEC, Clermont-Ferrand, France
- 5 RAGT UK, Ickleton, UK
- 6 NIAB, Cambridge, UK
- 7 Lemaire-Deffontaines, Auchy-les-Orchies, France
- 8 RAGT, Louville la Chenard, France
- 9 Caussade Semences, Caussade, France
- 10 CETAC et KWS-Momont Recherche, Monts en Pelée, France
- 11 SECOBRA, Maule, France
- 12 Arvalis, Thiverval-Grignon, France
- 13- WUR, Wageningen, The Netherlands
- 14 Florimond Desprez, Cappelle en Pévèle, France
- 15 CCDM, Curtin University, Perth, Australia
- 16 RAGT Czech, Branišovice, Czech Republic
- 17 ASUR Plant Breeding, Estrées-Saint-Denis, France
- \* Coordinateur : Marc-Henri LEBRUN, marc-henri.lebrun@inra.fr

## 1. Introduction

Les maladies fongiques du blé sont un problème sanitaire majeur qui affecte aussi bien les rendements que la qualité des grains. Parmi les agents pathogènes plus dommageables, Fusarium graminearum (Fg) et Zymoseptoria triticii (Zt) provoquent en France des épidémies récurrentes et importantes sur les épis et les feuilles du blé respectivement. Les pertes occasionnées par Zt peuvent atteindre de 15 qx/ha à 50 qx/ha (Arvalis), rendant nécessaire l'utilisation récurrente de traitements fongicides, avec un coût moyen de 40 €/ha chaque année. Parastagonospora nodorum (Pn) est un champignon pathogène des feuilles du blé qui était dominant en Europe entre 1950 et 1980 (Bearchell et al, 2005). Si l'incidence de cette maladie a fortement déclinée en France à partir de 1980, elle provoque encore des dommages en Suisse, en Europe de l'Est et en Europe du Nord (Tommasini et al, 2007 ; Gilon et al. 2014). En France, cette maladie reste responsable d'épidémies récurrentes sur le blé dur, les triticales et l'orge (Arvalis). Les isolats de Pn collectés sur blé dur sont aussi pathogènes du blé tendre (CASDAR Septodur). Ainsi, les épidémies de Pn sur le blé dur représentent un réservoir potentiel de souches capables de déclencher des épidémies sur le blé tendre. Pyrenophora tritici-repentis (Ptr) est une maladie fongique du blé observée en France dans les situations de monocultures de blé.

En dehors des traitements fongicides, la principale méthode de lutte contre les champignons pathogènes du blé, est l'utilisation de cultivars génétiquement résistants. Cependant, ces résistances sont souvent quantitatives et contrôlées par des déterminismes génétiques complexes (Brown *et al.*, 2015 ; Phan *et al.*, 2018). Enfin, les gènes/QTLs actuellement disponibles ne sont pas toujours suffisants pour contrôler totalement ces maladies. Il apparaît donc nécessaire de rechercher de nouvelles sources de résistance pour lutter efficacement contre ces maladies fongiques. Une approche prometteuse est de développer des méthodes d'identification et de criblage rapide de ces résistances en utilisant les connaissances acquises sur les mécanismes mis en place par ces agents pathogènes pour attaquer les plantes.

L'identification et l'étude d'effecteurs ayant des rôles clefs dans le pouvoir pathogène des champignons phytopathogènes est un domaine de recherche en plein essor (Lo Presti et al., 2015 ; De Wit et al., 2016). Les effecteurs fongiques sont en général de petites protéines sécrétées par le champignon dans les tissus végétaux lors de l'infection. Certains de ces effecteurs sont capables d'entrer dans les cellules végétales, où ils pertubent des fonctions importantes pour la plante (Toruño et al. 2016, Khan et al. 2018). Certains effecteurs sont capables d'induire une mort cellulaire spécifiquement chez certains cultivars sensibles, comme les effecteurs nécrotiques cultivar-spécifiques de Pn (Oliver et al, 2012 ; McDonald et al., 2018). D'autres effecteurs sont reconnus directement ou indirectement par les produits des gènes de résistance de la plante (effecteurs d'avirulence, Stergiopoulos et al, 2009). Les connaissances acquises sur les effecteurs fongiques ouvrent des possibilités d'applications innovantes pour la sélection de cultivars résistants aux maladies (Vleeshouwers et al. 2015). La sélection assistée par effecteurs offre de nombreux avantages par rapport à la sélection classique. En particulier, les effecteurs peuvent être utilisés rapidement à grande échelle, et accélérer la détection de nouvelles sources de résistance aux agents pathogènes. Ces méthodes ont permis d'identifier des mécanismes de résistance originaux, principalement chez la tomate et la pomme de terre (Zhu et al., 2012 ; Vleeshouwers et al. 2015), mais aussi chez le blé dans le cas des effecteurs nécrotiques de Pn (McDonald et al., 2018).

Le projet (WEAB) a pour objectif d'évaluer des méthodes de sélection de résistances aux champignons pathogènes chez le blé à l'aide d'effecteurs fongiques. Nous avons choisi d'utiliser des effecteurs produits par *Ptr* et *Pn*, capables d'induire des nécroses spécifiques chez des cultivars sensibles à ces champignons. Les effecteurs nécrotiques spécifiques de *Pn* (ToxA, Tox1 et Tox3)

peuvent être produits sous la forme de protéines recombinantes qui sont infiltrées dans des feuilles de blé. Ces méthodes permettent d'identifier les cultivars insensibles et résistants à ces effecteurs et elles sont en cours d'utilisation dans plusieurs programmes internationaux (Tan et al. 2015; Downie et al., 2018; Phan et al., 2018). Le projet WEAB comporte quatre axes, (1) la recherche d'effecteurs de Zt et Fg toxiques pour le blé (2) l'identification de cultivars de blé insensibles aux trois effecteurs nécrotiques de Pn connus (ToxA, Tox1 et Tox3) parmi les cultivars élites du panel Breedwheat, (3) l'évaluation en conditions contrôlées et au champ de ces mêmes cultivars pour leur résistance à Pn et Ptr, et (4) la cartographie des insensibilités aux effecteurs nécrotiques de Pn et de la résistance à Pn et Ptr, par des méthodes de génétique d'association. Ce projet a ainsi pour objectif d'évaluer l'efficacité des méthodes de criblage rapide à l'aide d'effecteurs fongiques chez le blé.

## 2. Matériel et méthode

Les effecteurs de Zt, Fg et Pn sélectionnés ont été produits sous la forme de protéines recombinées, soit dans Pichia pastoris (plateforme 3PE, UMR BFF, Marseille, France), soit dans Escherichia coli (CBS, Montpellier, France). Les protéines recombinantes ont été conservées à 4°C à une concentration d'environ 1-0,1 mg/mL. Tox1, Tox3 et ToxA ont été produits par l'Université de Curtin (Perth, Australie). Ces effecteurs ont été infiltrés dans les feuilles de blé à l'aide d'une seringue sans aiguille simplement pressée contre la feuille (Figure 1). Cette méthode permet d'infiltrer dans le mésophile des feuilles de blé environ 100 microlitres de la solution de protéine testée. Les symptômes de nécroses apparaissent sur les feuilles de cultivars sensibles en 4 à 5 jours. Ces méthodes sont celles déjà utilisées avec succès pour la détection de cultivars sensibles aux effecteurs nécrotiques de Pn (Tox1, Tox2, ToxA; Tan et al. 2015 ; Downie et al, 2018 : Phan et al., 2018).



Figure 1 : Infiltration d'effecteurs fongiques dans les feuilles de blé.

Un panel de 219 cultivars élites issu du programme ANR Breedwheat, a été utilisé pour évaluer la sensibilité/insensibilité de ces cultivars aux effecteurs nécrotiques Tox A, Tox 1 et Tox3 de *Pn* en chambre de culture. Les feuilles de ces cultivars ont été infiltrées dans le cadre d'essais multi-locaux réalisés par les différents partenaires du projet. Les symptômes provoqués par ces effecteurs nécrotiques de *Pn* ont été notés sur une échelle de 0 (pas de nécrose) à 4 (large nécrose, voir Figure 2). Ce panel a été aussi utilisé pour évaluer la résistance des cultivars de blé à *Pn* ou à *Ptr* dans des essais aux champs. Les essais pour *Pn* ont été réalisés en République Tchèque (station RAGT), et en Norvège (NMBU) avec des sur-inoculation d'isolats locaux de *Pn*. Les essais pour *Ptr* ont été réalisés en France avec un inoculum naturel dans une région où cette maladie est récurrente (station SECOBRA, 2015-2016). Les données phénotypiques obtenues par le programme WEAB ont permis de réaliser des analyses d'association en utilisant les données de génotypage produites par le PIA Breedwheat (197K SNPs). Ces analyses d'associations ont été réalisées avec le package GenABEL de la version 3.2 du logiciel statistique R (Aulchenko Y.S. *et al.*, 2007). Un modèle mixte linéaire (MLM) a été utilisé comme présenté par Yu J. *et al.* (2006). L'effet polygénique a été pris en compte en utilisant une matrice d'apparentement (Kinship). Cette matrice est estimée par la fonction IBS (identity-by-state) en utilisant un sous-ensemble de 8078 SNPs distribués tout le long du génome. Pour réduire le taux de faux-positifs, les associations marqueur-phénotype sont considérées comme significatives lorsque le LOD score est supérieur à 3.

## 3. Résultats

#### 3.1 - Identification, production et caractérisation d'effecteurs des champignons Zt, Fg et Pn

La recherche d'effecteurs candidats dans les génomes de Fg et Zt a été réalisée à l'aide de logiciels d'analyse spécialisés, tels qu'EffectorP (Sperschneider et al, 2003). Les données obtenues sont comparables à celles déjà publiées (Zt : Morais do Amaral et al. 2012 ; Fg : Brown et al. 2012). Les profils d'expression de ces effecteurs candidats ont été analysés à l'aide des données publiées (Zt : Palma-Guerrero et al. 2017; Fg: Brown et al. 2017). Seuls les effecteurs candidats de *Zt* et *Fg* exprimés au cours de l'infection du blé ont été retenus pour le projet WEAB. Dans le cadre du projet CASDAR Septodur (2012-2015), des souches de Pn ont été isolées en France à partir de cultivars de blé dur et de triticale infectés. Une seule souche de Pn a été isolée d'un cultivar de blé tendre (Arvalis). Deux autres isolats de Pn provenant de blé tendre sont issus de collections mycologiques françaises. L'isolat récent de blé tendre est pathogène pour le blé tendre, et pour la majorité des six cultivars de blé dur testés. Les isolats provenant de blé dur sont tous pathogènes pour le blé tendre. Un échantillon de 12 isolats de Pn français provenant de blé tendre, blé dur et triticale a été étudié pour la diversité des gènes encodant les effecteurs nécrotiques Tox1, Tox3 et ToxA. La plupart des isolats de Pn français ne possèdent pas ToxA, mais ils possèdent Tox1 et Tox3. Certains allèles de Tox1 et Tox3 sont différents de ceux connus chez Pn (Phan et al., soumis).

Huit des dix effecteurs de Zt sélectionnés ont pu être produits sous la forme de protéines recombinantes. La protéine Zt-MAX1 est similaire aux protéines de la famille MAX qui comporte des protéines toxiques pour le blé, comme ToxB (de Guillen et al. 2015). Les protéines Zt-NEP1 et Zt-NEP2 ont déjà été décrites comme toxiques pour le blé (Ben M'Barek et al. 2015). Deux protéines de Zt, apparentées aux cérato-platanines / expansines (Zt-Cer1, Zt-Expn1, Zt-Expn2, Zt-Expn3) toxiques pour les plantes (Pazzagli et al., 2014), ont été aussi produites. Aucune de ces protéines n'a induit de nécroses après infiltration de feuilles d'un cultivar de blé très sensible à Zt (Taichung 29), ou de cultivars de blé utilisés dans l'étude de Ben M'Barek et al. (2015), et ce, même à de fortes concentrations (10 µg/mL). Ces expériences ont toutefois permis de produire les protéines Zt-MAX1, et Zt-NEP1 en quantités suffisantes pour étudier leurs structures tridimensionnelles (de Gillem et Padilla, CBS, Montpellier, France).

Six des dix effecteurs de *Fg* sélectionnés ont pu être produits en quantités suffisantes sous la forme de protéines recombinantes. La protéine Fg-XYL1 a déjà été décrite comme toxique pour le blé

(Sella *et al.*, 2013). Une protéine de *Fg* apparentée aux cératoplatanines / expansines (Fg-CER1) connues pour être toxiques pour les plantes, et importantes pour l'infection du blé par *Fg* (Quarantin *et al.*, 2016) a également été produite. Les 4 autres effecteurs candidats de *Fg* correspondent à des protéines possédant soit des domaines associés à une activité toxique (HR inducing), soit des profils d'expression précoces au cours de l'infection du blé. Pour évaluer leur toxicité, ces effecteurs ont été pulvérisés sur l'épi de blé et déposés directement dans les fleurs. Aucune de ces protéines n'a induit de nécroses des fleurs de blé, même à de très fortes concentrations (100 µg/mL).

#### 3.2 - Identification de cultivars de blé sensibles ou insensibles aux effecteurs nécrotiques de Pn

Les feuilles de blé ont été infiltrées par les effecteurs nécrotiques de Pn, Tox1, Tox3 et ToxA. Les symptômes provoqués par ces effecteurs ont été notés sur une échelle de 0 (pas de nécrose) à 4 (nécrose, Figure 2). Les notes intermédiaires correspondent à des chloroses (1, 2) ou un mélange de chlorose et de nécrose (3, Figure 2).



Figure 2 : Échelle de notation des symptômes provoqués sur les feuilles de blé par les effecteurs nécrotiques de Pn.

Les effecteurs Tox1, Tox3 et ToxA ont été infiltrés dans des feuilles de blé de différents cultivars connus pour leurs réactions vis-à-vis de ces effecteurs. Les cultivars Cadenza (témoin de sensibilité, Downie *et al.* 2018) et Soissons sont extrêmement sensibles à ToxA (Figure 3), mais aussi à Tox 1 et Tox 3. Ils font partie des rares cultivars de blé d'hiver sensibles aux trois effecteurs nécrotiques de *Pn*. Ces deux cultivars ont été utilisés comme témoins de réussites des infiltrations des effecteurs nécrotiques de *Pn* dans les différentes expérimentations du projet. Le cultivar Euclide est insensible à ToxA.

Ces effecteurs ont été aussi infiltrés dans les feuilles des 219 cultivars de blé du panel Breedwheat. Ces expériences ont montré que la plupart des cultivars (46%) sont insensibles aux trois effecteurs nécrotiques Tox1, Tox3 et ToxA (Tableau 1). Seuls les quatre cultivars, XI19, Soissons, Isengrain et Icarda-4 sont sensibles à ces trois effecteurs (2%). La plupart des autres cultivars sont sensibles à un seul effecteur, majoritairement à Tox3 (39%).

Ces résultats étaient inattendus, car la totalité des blés australiens sont sensibles à au moins un des trois effecteurs Tox1, Tox3 et ToxA, et il n'existe pas de cultivars insensibles à ces trois effecteurs (Tan *et al.*, 2015 ; Phan *et al.*, 2018). Les cultivars australiens sont des blés de printemps, tandis que ceux du panel Breedwheat sont des blés d'hiver (212/219). Nous avons donc analysé la sensibilité des cultivars de blé de printemps européens aux effecteurs Tox1, Tox3 et ToxA (Tableau 2). La fréquence de cultivars sensibles aux trois effecteurs Tox1,

Tox3 et ToxA parmi les blés de printemps est six fois supérieure (0,13) à celle observée parmi les blés d'hiver européens (0,02). Ce résultat a été aussi observé par Downie *et al.* (2018) lors de la comparaison entre d'autres cultivars de blé d'hiver et de printemps européens.



**Figure 3**. Symptômes provoqués par l'effecteur nécrotique ToxA sur les feuilles de blé des cultivars Cadenza (note 4 : sensible), Soissons (note 4 : sensible) et Euclide (note 0 : insensible).

Pays	S ToxA	S Tox3	S Tox1	N	R 3 tox	S 3 tox
AUT	0,00	0,75	0,25	4	0,00	0,00
BEL	0,00	0,33	0,00	3	0,67	0,00
DEU	0,00	0,43	0,14	7	0,43	0,00
DNK	0,00	0,00	0,00	3	1,00	0,00
FRA	0,14	0,37	0,08	153	0,50	0,01
ITA	0,27	0,55	0,09	11	0,36	0,00
MOR	0,50	0,50	0,00	2	0,00	0,00
SPA	0,00	0,25	0,00	4	0,75	0,00
SYR	0,40	0,67	0,40	5	0,00	0,20
UK	0,04	0,48	0,22	23	0,30	0,04
Autre	0,00	0,25	0,25	4	0,50	0,00
Total	0,13	0,39	0,11	219	0,46	0,02

 Tableau 1 : Fréquence des cultivars de blé sensibles aux effecteurs nécrotiques

 Tox1, Tox3 et ToxA de Pn (Panel Breedwheat)

S : sensible à l'effecteur X ; N : nombre de cultivars testés ; AUT: Autriche, BEL: Belgique; DEU: Allemagne; DNK: Danemark; FRA : France ; ITA : Italie ; MOR : Maroc ; SPA : Espagne ; SYR : Syrie ; UK : Grande-Bretagne. Vert : fréquence 4% <, Violet : fréquence > 40%.

	<b>S</b> ToxA	S Tox3	S Tox1	N	R 3 tox	S 3 tox
Breedwheat	0.12	0.20	0.10	212	0.47	0.02
Hiver	0,12	0,38	0,10	212	0,47	0,02
Europe	0.65	0.57	0.26	22	0.12	0.12
Printemps	0,65	0,57	0,26	33	0,13	0,13

 Tableau 2 : Fréquence des cultivars de blé sensibles aux effecteurs nécrotiques

 Tox1, Tox3 et ToxA de Pn

S : sensible à l'effecteur X ; Breedwheat Hiver : Panel de blés d'hiver du programme Breedwheat ; Europe pintemps : panel de blés de printemps européens. Vert : fréquence 4% <, Violet : fréquence > 40%

#### 3.3 - Evaluation de la sensibilité au champ des cultivars de blé du Panel Breedwheat à Pn et à Ptr

Les niveaux de sensibilité des cultivars de blé du panel Breedwheat à *Pn* et *Ptr* ont été évalués dans le cadre d'expérimentations au champ, réalisées dans différents pays européens où cette maladie provoque des épidémies sur blé tendre. Ces expérimentations ont permis d'obtenir des informations sur les niveaux de sensibilité des 219 cultivars du panel Breedwheat aussi bien à *Pn* qu'à *Ptr*. Les cultivars présentent des niveaux de sensibilité foliaires à *Pn* très différents (Figure 4), allant d'une sensibilité très forte (Courtot, Laser, Mercato, XI19), a une résistance importante (Azzerti, Bergamo, Musik, Player, Sankara). La plupart des cultivars ont des niveaux intermédiaires de sensibilité. La sensibilité du cultivar XI19 et la résistance des cultivars Bergamo et Player ont été confirmées par des inoculations contrôlées de jeunes plants de blé par une souche française de *Pn* isolée de blé tendre (voir 2.1).

## 3.4 - Identification des loci impliqués dans la sensibilité aux effecteurs nécrotiques de Pn et à la résistance à Pn/Ptr

L'ensemble des données phénotypiques obtenues par le projet WEAB (voir 2.2, 2.3) a été utilisé pour réaliser des études

d'association sur l'ensemble du génome du blé. En effet, le lien statistique entre ces données phénotypiques et les données de génotypage de ces 219 cultivars obtenus dans le cadre du PIA Breedwheat (197K SNPs), permet de mettre en évidence les régions du génome impliquées dans l'insensibilité aux effecteurs nécrotiques de Pn, et dans la résistance au champ à Pn et Ptr. Avec les données sur les insensibilités à Tox1, nous avons pu identifier le gène SNN1 localisé à une extrémité du chromosome 1B (2354630-2357674) qui est impliqué dans la sensibilité à Tox1 (Figure 5). Ce résultat attendu montre que les données phénotypiques et génotypiques utilisées par WEAB permettent de détecter un locus associé à la sensibilité à Tox1. Les gènes TSN1 (chromosome 5B: 546825379-546833139) et SNN3-B1 (chromosome 5B: 6654131-6654201) impliqués respectivement dans la sensibilité du blé à ToxA et Tox3, ont été aussi détectés dans ces analyses. De nombreux loci impliqués dans la réponse des blés à ces trois effecteurs nécrotiques différents de TSN1, SNN1 et SNN3-B1 ont aussi été détectés, en particulier pour ToxA.

Les analyses de génétique d'association réalisées avec les données phénotypiques provenant des des essais aux champs (*Pn, Ptr*) ont permis de détecter des 96 loci associés à la résistance à *Ptr*, 82 loci pour *Pn*. Dans le cas de la résistance à *Pn*, ces loci sont différents de ceux détectés avec les effecteurs nécrotiques Tox1, Tox3 et ToxA.



## 4. Discussion

L'utilisation d'effecteurs fongiques pour la sélection de cultivars résistants aux agents pathogènes est une stratégie qui paraît très prometteuse. Cependant, elle nécessite des connaissances approfondies sur les interactions entre les champignons et leurs plantes hôtes. Dans le cas de Zt et Fg, pour lesquels il n'est pas encore bien établi quels sont leurs effecteurs toxiques pour le blé, cette stratégie en est sans doute encore à son début. En effet, le criblage de 14 candidats n'a pas permis d'identifier d'effecteur toxique pour le blé. Cette approche n'a pas non plus abouti dans le cas des expériences réalisées par Kettles et al. (2017a). Ces auteurs n'ont trouvé que des effecteurs toxiques pour Arabidopsis ou le tabac (14/63 testés), à l'exception d'un effecteur toxique pour le blé, encodant une ribonucléase. Cet effecteur qui est aussi toxique pour les bactéries et les champignons, sauf Zt (Kettles et al., 2017b), n'est produit qu'au début et à la fin de l'infection. Il pourrait permettre à Zt d'éliminer les microorganismes présents à la surface de la feuille. Ces premières études fonctionnelles sur les effecteurs de Zt montrent la nécessité d'étudier soit un plus grand nombre d'effecteurs (71 effecteurs testés sur au moins 300 candidats), soit d'utiliser d'autres méthodes de criblage pour espérer identifier des effecteurs fongiques toxiques pour le blé.

Dans le cas des effecteurs de *Pn*, les connaissances sur les effecteurs sont bien plus avancées que chez *Zt* et *Fg*. En particulier, il est bien établi que *Pn* produit un certain nombre (> 9) d'effecteurs nécrotiques dont trois ont été caractérisés (Tox1, Tox3, ToxA; Mac Donald *et al.* 2018). L'analyse de la sensibilité des cultivars de blé d'hiver du panel Breedwheat à Tox1, Tox3 et ToxA, a permis d'identifier un grand nombre de cultivars européens insensibles à ces trois effecteurs (n : 100, Tableau 1). Ce résultat est inattendu, car les travaux réalisés jusqu'à présent avaient montré l'absence de ce type de cultivars dans les collections de blé australiens (Tan *et al.*, 2015; Phan *et al.*, 2018). En parallèle de notre projet, d'autres études ont montré que de tels cultivars étaient présents dans les collections de blé d'hiver.

Cependant, la fréquence de ces cultivars semble inférieure (17% du panel russe : Phan et al., 2018; 10 % du panel européen analysé par Downie et al., 2018) à celle observée dans le panel Breedwheat (39 %). Ce génotype pourrait conférer une certaine résistance à Pn qui aurait été sélectionnée au cours des années 1960-1980 en Europe, lorsque cette maladie était importante. Cette hypothèse a également été suggérée pour expliquer l'existence de tels cultivars dans le panel russe, y compris dans des cultivars russes anciens (Phan et al., 2018). Cette information est importante, car elle indique qu'il faut éviter de perdre ces gènes d'insensibilité dans les programmes de sélection du blé actuels, depuis que la résistance à Pn n'est plus prise en compte. Ainsi, la généalogie des cultivars issus de Cadenza en Grande Bretagne montre que ce cultivar a conduit au maintien de gènes de sensibilité à Pn, car il a fréquemment été utilisé comme géniteur (Downie et al. 2018). C'est le cas du cultivar XI19 qui est issu de Cadenza et qui est à la fois sensible à ToxA, Tox1 et Tox3, et extrêmement sensible à Pn dans tous les essais que nous avons réalisés.

Les approches de génétique d'association ont une nouvelle fois démontré leur efficacité pour identifier les loci impliqués dans la sensibilité aux effecteurs toxiques de Pn, en particulier à l'aide d'un panel de cultivars élites tel que celui développé dans le PIA Breedwheat. En effet, nous avons pu facilement détecter les loci TSN1, SNN1 et SNN3, impliqués dans la sensibilité à ToxA, Tox1 et Tox3, respectivement. Des QTLs mineurs impliqués dans la réponse à ces effecteurs ont également été détectés. L'analyse de la sensibilité de ces cultivars à Pn et Ptr, réalisée à l'aide de ces méthodes, a aussi permis d'identifier des loci associés à une résistance quantitative à Ptr et Pn dans ce panel, indépendante des loci TSN1, SNN1 et SNN3. Ce résultat suggère qu'il existe d'autres loci de sensibilité du blé sans doute en interaction avec de nouveaux effecteurs de Pn, dont les allèles d'insensibilité confèrent une résistance aux populations de Pn responsables d'épidémies en Europe, comme cela a été suggéré pour d'autres populations de Pn (Tan et al. 2018; Phan et al. 2018).

## Références bibliographiques

Bearchell SJ, Fraaije BA, Shaw MW, Fitt BD (2005) Wheat archive links long-term fungal pathogen population dynamics to air pollution. Proc Natl Acad Sci USA. 102(15):5438-42.

Ben M'Barek S, Cordewener JH, Tabib Ghaffary SM, van der Lee TA, Liu Z, Mirzadi Gohari A, Mehrabi R, America AH, Robert O, Friesen TL, Hamza S, Stergiopoulos I, de Wit PJ, Kema GH. (2015) FPLC and liquid-chromatography mass spectrometry identify candidate necrosis-inducing proteins from culture filtrates of the fungal wheat pathogen Zymoseptoria tritici. Fungal Genet Biol. 79:54-62.

**Brown NA, Antoniw J, Hammond-Kosack KE.** (2012) The predicted secretome of the plant pathogenic fungus *Fusarium graminearum*: a refined comparative analysis. PLoS One. 7(4):e33731.

Brown JK, Chartrain L, Lasserre-Zuber P, Saintenac C. (2015) Genetics of resistanceto Zymoseptoria tritici and applications to wheat breeding. Fungal Genet Biol. 79:33-41.

**Brown NA, Evans J, Mead A, Hammond-Kosack KE.** (2017) A spatial temporal analysis of the *Fusarium graminearum* transcriptome during symptomless and symptomatic wheat infection. Mol Plant Pathol. 18(9):1295-1312.

**De Guillen K, Ortiz-Vallejo D, Gracy J, Fournier E, Kroj T, Padilla A.** (2015) Structure Analysis Uncovers a Highly Diverse but Structurally Conserved Effector Family in Phytopathogenic Fungi. PLoS Pathog. 11(10):e1005228.

De Wit PJ, Testa AC, Oliver RP. (2016) Fungal Plant Pathogenesis Mediated by Effectors. Microbiol Spectr. 4(6).

Downie RC, Bouvet L, Furuki E, Gosman N, Gardner KA, Mackay IJ, Campos Mantello C, Mellers G, Phan HTT, Rose GA, Tan KC, Oliver RP, Cockram J. (2018) Assessing European Wheat Sensitivities to Parastagonospora nodorum Necrotrophic Effectors and Fine-Mapping the Snn3-B1 Locus Conferring Sensitivity to the Effector SnTox3. Front Plant Sci. 9:881.

Kettles GJ, Bayon C, Canning G, Rudd JJ, Kanyuka K. (2017a) Apoplastic recognition of multiple candidate effectors from the wheat pathogen Zymoseptoria tritici in the nonhost Nicotiana benthamiana. New Phytol. 213(1):338-350.

Kettles GJ, Bayon C, Sparks CA, Canning G, Kanyuka K, Rudd JJ. (2017b) Characterization of an antimicrobial and phytotoxic ribonuclease secreted by the fungal wheat pathogen Zymoseptoria tritici. New Phytol. 217(1):320-331.

Khan M, Seto D, Subramaniam R, Desveaux D. (2018) Oh, the places they'll go! A survey of phytopathogen effectors and their host targets. Plant J. 93(4):651-663.

**Gilon M, Arseniuk E.** (2014) Natural field infections of wheat and triticale by fungi from the complex of fungi Stagonospora nodorum / Septoria tritici under climatic conditions of Poland. Commun Agric Appl Biol Sci. 79(4):216-27.

Lo Presti L, Lanver D, Schweizer G, Tanaka S, Liang L, Tollot M, Zuccaro A, Reissmann S, Kahmann R. (2015) Fungal effectors and plant susceptibility. Annu Rev Plant Biol. 66:513-45.

McDonald MC, Solomon PS. (2018) Just the surface: advances in the discovery and characterization of necrotrophic wheat effectors. Curr Opin Microbiol. 46:14-18. Morais do Amaral A, Antoniw J, Rudd JJ, Hammond-Kosack KE. (2012) Defining the predicted protein secretome of the fungal wheat leaf pathogen Mycosphaerella graminicola. PLoS One. 7(12):e49904.

**Oliver RP, Friesen TL, Faris JD, Solomon PS.** (2012) Stagonospora nodorum: from pathology to genomics and host resistance. Annu Rev Phytopathol. 50:23-43.

Palma-Guerrero J, Ma X, Torriani SF, Zala M, Francisco CS, Hartmann FE, Croll D, McDonald BA. (2017) Comparative Transcriptome Analyses in Zymoseptoria tritici Reveal Significant Differences in Gene Expression Among Strains During Plant Infection. Mol Plant Microbe Interact. 30(3):231-244.

Pazzagli L, Seidl-Seiboth V, Barsottini M, Vargas WA, Scala A, Mukherjee PK. (2014) Cerato-platanins: elicitors and effectors. Plant Sci. 228:79-87.

Phan HTT, Rybak K, Bertazzoni S, Furuki E, Dinglasan E, Hickey LT, Oliver RP, Tan KC. (2018) Novel sources of resistance to Septoria nodorum blotch in the Vavilov wheat collection identified by genome-wide association studies. Theor Appl Genet. 131(6):1223-1238.

Phan HTT, Jones D, Rybak K, Dodhia K, Lopez-Ruiz FJ, Valade R, Gout R, Lebrun M-H, Oliver RP, Tan KC. (Submitted) Periodic Pyrrhic victories define the Septoria nodorum blotch interaction.

#### blotch interaction

**Quarantin A, Glasenapp A, Schäfer W, Favaron F, Sella L.** (2016) Involvement of the *Fusarium graminearum* cerato-platanin proteins in fungal growth and plant infection. Plant Physiol Biochem. 109:220-229.

Sella L, Gazzetti K, Faoro F, Odorizzi S, D'Ovidio R, Schäfer W, Favaron F. A (2013) *Fusarium graminearum* xylanase expressed during wheat infection is a necrotizing factor not essential for virulence. Plant Physiol Biochem. 64:1-10.

Sperschneider J, Gardiner DM, Taylor JM, Hane JK, Singh KB, Manners JM. (2013) A comparative hidden Markov model analysis pipeline identifies proteins characteristic of cereal-infecting fungi. BMC Genomics. 20;14:807.

Tan KC, Phan HT, Rybak K, John E, Chooi YH, Solomon PS, Oliver RP. (2015) Functional redundancy of necrotrophic effectors - consequences for exploitation for breeding. Front Plant Sci. 6:501.

Tommasini L, Schnurbusch T, Fossati D, Mascher F, Keller B. (2007) Association mapping of Stagonospora nodorum blotch resistance in modern European winter wheat varieties. Theor Appl Genet. 115(5):697-708.

**Toruño TY, Stergiopoulos I, Coaker G.** (2016) Plant-Pathogen Effectors: Cellular Probes Interfering with Plant Defenses in Spatial and Temporal Manners. Annu Rev Phytopathol. 4;54:419-41.

Vleeshouwers VG, Oliver RP. (2015) Effectors as Tools in Disease Resistance Breeding Against Biotrophic, Hemibiotrophic, and Necrotrophic Plant Pathogens. Mol Plant Microbe Interact. (1):40-50.

**Zhu S, Li Y, Vossen JH, Visser RG, Jacobsen E.** (2012) Functional stacking of three resistance genes against Phytophthora infestans in potato. Transgenic Res. 21(1):89-99. Phil HOWELL<sup>1</sup>, Richard HORSNELL<sup>1</sup>, Pascal GIRAUDEAU<sup>2</sup>, Pierre PIN<sup>2</sup>, Stephen SUNDERWITH<sup>3</sup>, Benoît FOUCAULT<sup>3</sup>, Patrice SENNELART<sup>3</sup>, Philippe MOMONT<sup>3</sup>, Sylvie DUTRIEZ<sup>4</sup>, Volker LEIN<sup>5</sup>, Thierry BOUTHILLIER<sup>5</sup>, Jean-Michel DELHAYE<sup>6</sup>, Anne-Charlotte LOMBART<sup>6</sup>

- 1 NIAB Huntingdon Road, Cambridge, CB3 0LE, UK
- 2 SECOBRA Recherches Centre de Bois Henry, 78580 Maule, France
- 3 KWS Momont 7 rue de Martinval, 59246 Mons-en-Pevele
- 4 Caussade Semences ZI de Meaux, BP 109, 82303 Caussade cedex
- 5 ASUR Plant Breeding 163 Ter avenue de Flandre BP6 60190 Estrées St Denis
- 6 Lemaire Deffontaines 180 rue du Rossignol, 59310 Auchy-les-Orchies
- \* Coordinator: Pascal GIRAUDEAU, pascal.giraudeau@secobra.com

## 1. Introduction

Agriculture is facing a major challenge of how to grow more from less. It has been predicted that the global population will exceed 10 billion by 2055 (United Nations, 2017), with a rising demand for food, energy, water and land. Coupled with more frequent extreme weather events caused by climate change, this scenario has been described as a 'Perfect Storm' (Beddington, 2009). It is clear that plant breeders must increase the current rate of genetic gain, and hence varietal improvement, within their programmes. This will require an increase in absolute yield potential: the maximum yield when grown under optimal conditions. It will also need an improvement in yield stability: the likelihood that yield performance will remain regular across different locations and years.

Genetic gain can be increased through a combination of factors: i) more accurate phenotyping; ii) shorter breeding cycles; iii) more intensive selection; and iv) making use of increased genetic variation. Much research is targeted at improving these, such as: i) high-throughput phenotyping, remote phenotyping and image analysis (Araus and Cairns, 2014); ii) doubled haploid and speed breeding technologies (Kalinowska et al., 2018; Watson et al., 2018); iii) marker-assisted selection approaches, including genomic selection (Jannink et al., 2010) and iv) generating greater genetic diversity through pre-breeding (Moore, 2015).

Several countries have recently made significant public investments into crop pre-breeding, with high profile programmes in wheat such as BreedWheat (France), Wheat Improvement Strategic Programme / Designing Future Wheat and Wheat Genetic Improvement Network (WISP/DFW and WGIN, UK), ProWeizen (Germany) and Seeds of Discovery (Mexico), plus international efforts such as the International Wheat Yield Partnership (IWYP). Whilst these are primarily academic programmes, commercial breeders have an influence on the areas of research and are generally granted free access to project outputs. These can include genomic information, genetic markers to assist with selection for key production traits, and improved germplasm to use within private breeding programmes.

Bread wheat (*Triticum aestivum* L.) is an allohexaploid, with three related, but distinct, genomes termed A, B and D. It arose in the wild through two successive inter-specific hybridisation events, each of which was followed by chromosome doubling, in the Fertile Crescent. The first event, which occurred around 100 000 years ago, was between the diploid grasses *T. urartu* (donor of the A genome) and *Aegilops speltoides* (B genome and cytoplasm donor). This resulted in the tetraploid wild emmer (*T. turgidum dicoccoides*; AABB), from which several other

tetraploid wheats have been domesticated (eg durum wheat, cultivated emmer, Khorasan wheat, Rivet wheat). This was followed by a second event around 10 000 years ago, in which a tetraploid was pollinated by a third diploid grass, the wild goatgrass *Aegilops tauschii* (D genome donor), resulting in in the progenitor of hexaploid bread wheat (*T. aestivum*; AABBDD; Feldman 2001).

Molecular genetics suggests that hexaploid bread wheat has experienced substantial post-domestication gene flow with its tetraploid relatives, but not with diploid goatgrass. This is reflected by the fact that genetic maps typically display much higher levels of polymorphism across the A- and B-genomes than across the D-genome (Wang et al., 2014). Breeders are keen to access novel genetic variation, especially for the D-genome, as part of their attempts to increase the rate of genetic gain.

The second of wheat's evolutionary hybridisations can be recreated through controlled interspecific crossing, a process referred to as resynthesis. Here, an emasculated tetraploid (typically durum wheat, *T. turgidum durum*) is pollinated by *Ae. tauschii*. Embryo rescue is carried out on immature seeds, and after chromosome doubling, stable Synthetic Hexaploid Wheat lines (SHWs) can be recovered (Mujeeb-Kazi et al, 2001). As SHWs can be readily crossed with elite bread wheat varieties, they provide an excellent means of exploring the value of novel D-genome variation.

CIMMYT, the International Maize and Wheat Improvement Centre, carried out a large programme of resynthesis in the 1980s-90s. These CIMMYT SHWs have shown great breeding value especially under low-input regimes (reviewed by van Ginkel and Obgonnaya, 2007). Material derived by crossing CIMMYT SHWs into UK varieties has also shown promise in higher-input situations (Howell et al, 2017), leading to renewed interest from breeders in exploring the potential of SHWs. Following this success based on CIMMYT SHWs, NIAB began its own resynthesis within WISP, the UK public-sector wheat pre-breeding programme (Howell et al, 2017).

This FSOV project 'Integrating novel functional diversity for improved yield stability' (INDYS) is a pre-breeding collaboration between five French breeding companies grouped in the CETAC (ASUR Plant Breeding, Caussade Semences, KWS-Momont, Lemaire Deffontaines and SECOBRA Recherches) and NIAB, a UK academic partner. INDYS exploits novel genetic variation already produced by NIAB as part of the UK's WISP /DFW project, and has transferred this into French-adapted backgrounds, generating thousands of pre-breeding lines. These have been screened in the field by breeders across several locations in France, with the best selections taken forwards to multi-location yield testing and detailed genotyping.

## 2. Materials and Methods

#### Pre-screening of elite French varieties

NIAB's previous work demonstrated that most elite UK wheat varieties were vulnerable to hybrid necrosis in crosses with CIMMYT SHWs; F1s from these crosses showed classical hybrid necrosis symptoms of foliar lesions, reduced tillering, small ear size and even juvenile lethality. A small pre-project crossing programme (French elite varieties x CIMMYT SHWs) was initiated to generate a series of F1s which were assessed for hybrid necrosis symptoms, to help with the selection of two varieties to use as recurrent backcross parents for the INDYS project.

Thirty elite French varieties were selected by the CETAC breeders for this pre-screen (Table 1). Eight seeds per variety were sown at NIAB, vernalised for 8 weeks (8 hour daylength, 5 °C) then potted up and grown on in a lit, heated glasshouse (16 hour daylength; 22 °C daytime, 14 °C night time). Five CIMMYT SHWs (SHW-052, -109, -144, -181, -441; Mujeeb-Khazi and Hettel, 1995) were grown in the same fashion, except that these were potted up across several weeks, to provide the best chance of synchronous flowering and successful crossing. Prior to anthesis, ears of elite lines were emasculated and placed into a labelled cellophane crossing bag. After two days, these were pollinated using detached SHW ears. Four weeks after crossing, mature seed was harvested from successful crosses, threshed by hand, and planted as F1 families. F1 seedlings were visually scored during vernalisation for hybrid necrosis symptoms.

#### Population development

This was split across several successive cycles, to harmonise with the generation of new SHWs from NIAB's WISP project. All plants were grown as for the pre-screening experiment, with staggered potting of the recurrent parents to maximise synchronous flowering. Populations were developed through three successive glasshouse generations: i) elite x SHW  $\rightarrow$ F1; ii) elite x F1  $\rightarrow$  BC1 and F1 $\rightarrow$ F2; iii) BC1  $\rightarrow$  BC1F2, before moving on to field testing (Figure 1).



Figure 1: Schematic showing planned flow of material through the project

Seed of BC1F2 populations (self seed harvested from individual BC1F1 plants) was threshed into individual packets and sent to Secobra and KWS Momont in time for autumn planting for the 2014-15, 2015-16 and 2016-17 seasons.

#### Breeder selection: BC1F2 onwards

In 2014-15, 29 crosses, representing 17 SHWs and 13 different D-genome donors, were sown in France in autumn 2014 as BC1F2 families, tracing over 761 different BC1 plants. All families were grown at Maule (78, Secobra Recherches), and families with sufficient seed for two locations were also grown at Allonnes (28, KWS Momont). This resulted in a total of around 60 000 plants at Maule and 35 000 at Allonnes. Families were sown as 5m rows (1 to 3 rows per family depending on the number of seeds), with 10cm between plants. Observation nurseries received 1-3 fungicide treatments, depending on yellow rust pressure.

A pre-selection was made in the Secobra nursery to identify the best families, based on tiller number, ear size and shape, plant height, disease resistance, earliness and maturity. These families were then explored in more detail at both locations and single plants were selected, harvested and threshed. Due to the tenacious glume character inherited from the SHW parents, there were difficulties in threshing out clean seed samples and some further lines were discarded as a result. Selections were distributed across the full CETAC 5 network, for planting in multilocation BC1F3 nurseries.

In 2015-16, a further 23 crosses (547 BC1F2 families representing 15 SHWs and 14 D-genomes) were drilled at Maule (36 000 plants) and Allonnes (40 000). These have been grown in a similar fashion to the 2014-15 nurseries with at least 1 fungicide application. Single plants were selected at Maule whereas at Allonnes bulks from the best BC1F2 families were harvested.

In 2016-2017, the last 13 crosses (136 BC1F2 families representing 10 SHWs and 10 D-genomes) were grown at Maule (20 000 plants) and Allonnes as plant plots; again, individual plants were selected at Maule, and bulks from the best BC1F2 plant plots were harvested at Allonnes

#### Stage 2: multi-location evaluation of BC1F3 material

In 2015-16, each of the 4488 BC1F3 selections was grown at three different locations. Therefore, each breeder had 2680 different lines to experiment. Locations and breeders are indicated below:

- Reclainville 28 (Caussade Semences, CS)
- Allonnes 28 (KWS Momont, KM)
- Auchy les Orchies 59 (Lemaire Deffontaines, LD)
- Estrees St Denis 60 (ASUR Plant Breeding, SUR)
- Maule 78 (Secobra Recherches, SC)

The design for the experiments consisted of 1.5m double rows per line. On each site, earliness to heading was scored on all lines, either with the CTPS scale (5=Robigus; 5.5=Terroir; 6=Bermude; 6.5=Bastide; 7=Arezzo; 7.5=Solehio; 8=Nogal) or with the heading date. To verify that the identity of lines evaluated was the same in each location, several DUS characters were assessed: presence/absence of awns, plant glaucosity, and plant height. Diseases were scored when present on the site; since 2016 was very favourable to diseases, a general disease score (green leaves score) was made on several sites. Differential lodging was also scored at three sites (LD, KWS, and CS).

For selection purposes, each breeder gave each line a global score from 1 (very good) to 9 (very bad); all scores were gathered to make the final selection, centred and reduced to give the same weight to the selection scores of each breeder.

However, due to high disease pressure on all sites (especially *Microdochium spp.*) and poor weather conditions during grainfilling and harvest, yields and grain quality were not good enough to provide sufficient seed for 2016-17 yield trials as planned. Instead, all selected lines were replanted at all five locations in 2016-17, using BC1F4 seed harvested either from Secobra or KWS Momont, and the opportunity taken to make additional notation on phenotypes. Each breeder harvested bulk BC1F5 seed from their own nurseries to use for trials in 2017-18.

In 2016-17 and 2017-18, BC1F3 from the BC1F2 plants harvested at Maule the previous year were planted at Maule only, using a similar experimental design to the 2015-16 multilocation evaluation, with 1094 (2016-17) and 581 (2017-18) lines, respectively. This nursery received one fungicide treatment in 2016-17 due to early high yellow rust levels, whereas no fungicide was applied the following year. For both years, earliness of each line was assessed using the CTPS scale, a negative selection for diseases was applied, and the final selection took into account ear type, height, tillering and general disease resistance. The bulks of the best BC1F1 progenies harvested at Allonnes in 2016 and 2017 were distributed among the 5 breeders of CETAC; each breeder will then be able to proceed with its own selection on the following years and the advanced lines produced this way will be evaluated in a common network and made available for crossing to the others.

#### Stage 3: multi-location yield testing of BC1F4 and BC1F5 material

Selections from the 2016-17 evaluation were sown in a coordinated series of 2017- 18 yield trials across the CETAC 5 network (Figure 2): Caussade 82 (Caussade semences, CS), Allonnes 28 (KWS Momont, KM), Auchy les Orchies 59 (Lemaire Deffontaines, LD), Estrees St Denis 60 (ASUR Plant Breeding, SUR), Maule 78 (Secobra Recherches, SC).



Figure 2: location of the five trial locations, 2017-18

Trials were designed with 1 rep per location of each experimental line overlaid onto a grid of highly replicated controls. Each trial had a unique randomisation, overlaid onto a common structure. The primary control, 'Rubisko' was placed in every tenth plot, alternating between position 1 and position 6 moving down the field. (Figure 3). The secondary checks were the other three CTPS control varieties: Cellule, Fructidor and Triomph (KM, LD, SUR, SC trials); for the southern CS trial, Triomph was replaced by the southern CTPS check 'Filon'. The secondary checks were replicated seven to eight times across each trial. Each trial was grown to local practice, with different harvested plot areas (SUR 4m<sup>2</sup>; KM, LD 7m<sup>2</sup>; CS 9m<sup>2</sup>; SC 10 m<sup>2</sup>).



Figure 3: schematic of trial layout showing grid of Rubisko and randomised other lines.

The grid of Rubisko checks is shown in grey, with other CTPS controls in red. Unshaded plots contain experimental lines.

All trials received a full programme of fungicide and plant growth regulators, following the CTPS protocol for treated yield. Heading date, height and lodging were assessed on all trials. Specific weight, grain moisture, grain protein content and TKW were measured on all harvested plots. Yields were adjusted using spatial analysis to account for within-field variation of the control varieties, following the type 2 modified augmented design described by Lin and Poushinsky (1985). All phenotype information was collated from stage 2 and 3 observations and trials, and checked prior to submission for QTL analysis.

#### DNA extraction and SNP array

A representative seed sample (approx. 10 seeds) was sent to NIAB of each line tested in yield trials. One seed per line was sown to provide leaf tissue for DNA extraction and genotyping, together with parental reference samples.

Leaf tissue was harvested from young seedlings (four-leaf stage) and extracted using a modified Tanskley miniprep protocol (Fulton et al, 1995). DNA samples were quantified using Nanodrop and concentrations normalised to 30ng gDNA/µI. DNA samples were hybridised to the Axiom 35K Wheat Breeders Array (Affymetrix ThermoFisher), at the Bristol Genomics Facility (University of Bristol, UK).

#### Genotyping workflow

Quality control (QC) and SNP calling was carried out using the Axiom<sup>™</sup> Analysis Suite v4.0.1 software. QC followed a modified 'Best Practices Workflow' (see: Affymetrix 2011). Modifications included: a lowered sample quality of ≥95%; a heterozygous inbred penalty of 4 (experimental lines) or 8 (parental lines); a lowered SNP call rate minimum of 96%. Quality control was carried out on all material together.

Once this had been completed, material was split into three batches for genotyping, according to genetic background: Expert, Bastide or Robigus.

Within each batch, all progenies and relevant parents were loaded together for genotyping and analysis. Axiom Analysis Suite

software classified probes into the five standard classes: PolyHiRes (Poly), NoMinorHom (MNH), MonoHighRes (Mono), Call Rate Below Threshold (BT) OffTargetVariant (OTV) and Other (OTH). This information, plus the genotype calls, was exported for each batch and assembled into a spreadsheet. Genotype data for probes which were Poly, NMH, and OTV were retained; all other data was discarded. Probes were sorted into approximate map order based upon the annotated data provided by the software for Avalon x Cadenza and Savannah x Rialto maps.

Within each batch, parental replicate samples were compared and probes stripped out where genotype calls did not match. The genotype data was uploaded into FLAPJACK v1.18.06.29 (Milne et al, 2010) for analysis and visualisation. For each genetic background, an overview of all material was carried out, plus a cross-by-cross breakdown of related material according to pedigree information.

A second approach brought all three backgrounds together to use a common set of markers. Another quality control sift was carried out and markers were retained with missing values < 0.1, minor allele frequency > 0.05, heterozygotes < 0.1). This dataset was used for GWAS and QTL analysis.

#### QTL analysis

Genotypes were passed through multidimensional scaling analysis to identify any underlying structure before devising appropriate models for GWAS.

A relatively simple model was chosen: GLM with pedigree (F) and multidimensional structure (Q) as co-variates. For simplification, trait data was included as best linear unbiased estimates (BLUEs), rather than separate analyses for each location/year.

### 3. Résults

#### ▶ Hybrid necrosis and recurrent parent selection

F1 seed was successfully recovered from crosses involving 20 of the 30 elite parents chosen for pre-project screening. Hybrid necrosis symptoms (HN) were scored for all F1s from crosses against five CIMMYT SHWs (Table 1) and these data, together with breeder consideration, was used to select two varieties, 'Expert' and 'Bastide' to use as recurrent parents during population development. HN levels were greater for all potential elite parents than had been observed for 'Robigus' x SHW crosses within the WISP project.

Variety	Crosses attempted <sup>a</sup>	F1 seeds <sup>₀</sup>	HN score <sup>c</sup>
ACCROC	1	7	1
ADHOC	1	2	2
ALIXAN	1*	0	
ALLEZY	1	1	1
ALTIGO	1*	0	
AREZZO	0**		
ASCOTT	2	9	1/2
BAROK	1*	0	
BERMUDE	1	1	
BOISSEAU	0**		
CHEVRON	3	13	1/2/1
EXPERT	3	13	2/1/1
FANION	0**		
GARCIA	1*	0	
GONCOURT	1	6	2
JB ASANO	3	10	1/2/1
KORELI	0**		
MIROIR	2	7	2/3
PAKITO	1	3	1
PALEDOR	1	3	0
BASTIDE	1	5	1
PREVERT	2	10	1/2

RUBISKO	0**		
RUSTIC	2	5	1/2
SOKAL	1	4	3
SOLEIHO	1*	0	
SOLLARIO	1	1	0
SWEET	1	2	2
SY MOISSON	1	2	2
TRAPEZ	1	5	1

**Table 1:** Pre-project parental screening for hybrid necrosis in elite x SHW crosses

 <sup>a</sup> Number of crosses attempted: \*unsuccessful cross with no seed set; \*\* asynchronous

 flowering prevented crossing. / <sup>b</sup> Number of F1 seedlings germinated across all crosses. /

 <sup>c</sup> Visual HN score for the most affected plant in each F1, scored as 0: no symptoms;

 1: lesions on 1-2 leaves; 2: lesions on 3-4 leaves; 3: lesions on 5+ leaves.

#### Crossing cycles and population development

Material was generated across several successive cycles. Plant losses were high, as F1 and BC1 HN levels were even greater than during the parental screening experiment. For this reason, material based on the UK variety 'Robigus', previously known to show no levels of HN in crosses with SHWs, was also included in the project, alongside the 'Expert' and 'Bastide' material, to keep plant numbers high enough for effective selection. Population development is summarised in Table 2. Typically, 100 seeds were distributed for each BC1F2 family, split between Secobra (all families) and KWS Momont (if sufficient seed).

SHW	/ donor pare	ent	Red	current pa backgrour	irent Id
SHW	AABBa	DDb	EXPERT	BASTIDE	ROBIGUS
NIAB SHW 008	HOH 501	WX 895	1		
NIAB SHW 012	HOH 501	WX 325	3		
NIAB SHW 027	ICARDA	WX 325	29	56	19
NIAB SHW 029	SCULPTUR	WX 224	30	41	
NIAB SHW 033	DAKTER	WX 224		35	19
NIAB SHW 036	ARDENTE	JIC 2220007	13	22	
NIAB SHW 041	HOH 501	ENT 336			24
NIAB SHW 042	HOH 506	ENT 336		11	
NIAB SHW 043	HOH 501	ENT 088	19	37	12
NIAB SHW 051	HOH 501	JIC 2220007	20		
NIAB SHW 069	HOH 506	ENT 127			61
NIAB SHW 070	HOH 501	ENT 127	22		
NIAB SHW 071	HOH 501	ENT 077			92
NIAB SHW 073	HOH 501	ENT 242	16		61
NIAB SHW 075	HOH 501	ENT 404	11	18	45
NIAB SHW 076	HOH 501	ENT 405	19	30	69
NIAB SHW 077	HOH 501	ENT 402	40	18	63
NIAB SHW 078	HOH 501	ENT 431	2	13	10
NIAB SHW 080	HOH 501	ENT 270	14	11	
NIAB SHW 082	HOH 501	ENT 090	14	37	20
NIAB SHW 083	HOH 501	ENT 087	6	30	10
NIAB SHW 084	HOH 501	ENT 084	3	3	19
NIAB SHW 085	HOH 501	ENT 389	23	7	17
NIAB SHW 086	HOH 501	ENT 118	10	34	20
NIAB SHW 087	HOH 501	ENT 382		30	19
NIAB SHW 088	HOH 501	ENT 080			6
NIAB SHW 089	HOH 501	ENT 086			9
NIAB SHW 090	HOH 501	ENT 102	4	9	16
NIAB SHW 091	HOH 501	ENT 228		2	14
NIAB SHW 092	HOH 501	ENT 272			16
NIAB SHW 093	HOH 501	ENT 302			11
NIAB SHW 094	HOH 501	ENT 383		12	
NIAB SHW 095	HOH 501	ENT 078			5
NIAB SHW 099	HOH 501	ENT 430			16
NIAB SHW 100	HOH 501	ENT 141			16

 Table 2: BC1F2 families distributed to breeders by donor / recurrent combination

 a Tetraploid and <sup>b</sup> diploid components of each SHW. BC1F2 seed was distributed across three seasons: planting in 2014-15 (green), 2015-16 (yellow) and 2016-17 (blue).

In total, 1444 BC1F2 populations were developed, which captured 30 different D genomes across 35 different SHWs, in 66 donor / background combinations (Figure 4). Twenty SHW were captured in crosses with Expert (299 families), 20 with Bastide (456 families) and 26 with Robigus (689 families).



Figure 4: Venn diagram illustrating SHW crosses with Expert, Bastide and Robigus

#### ► Early generation field selection: stage 1 BC1F2 → BC1F3

Seed sent from NIAB was grown at two breeder locations (Secobra and KWS Momont). Breeders at each location selected individual plants, summarised in Table 3.

	201	4-15	201	5-16	201	6-17
	Secobra	KWS Momont	Secobra	KWS Momont	Secobra	KWS Momont
Families	761	656	547	530	136	136
sowna	(17)	(17)	(17)	(17)	(10)	(10)
Plants	2844	1644	1094	bulk	581	bulk
harvestedb	(15)	(16)	(17)	(17)	(10)	(10)

Table 3: Early generation field selection

<sup>a</sup> Numbers of families sown, representing the number of SHWs in parentheses; bnumber of selections made, representing the number of SHWs in parentheses

As expected, there was much more variation observed in the field than is typically seen in commercial breeding programmes. A full consortium meeting was held in June 2015, with most of the time spent in the field discussing selection criteria and the best way to deal with this diverse material, which was segregating for gross phenotypes such as ear and canopy glaucosity, height, and presence/absence of awns (Fig 5). After threshing individual plants, seed was split into several packets and distributed to the other breeders for stage 2 evaluation.

In subsequent years, Secobra continued plant-by-plant selection, and KWS Momont switched to a bulk selection system where the best plants within selected families were bulked together.

Around 190 000 BC1F2 plants were screened in total across the three years: 95 000 in 2014-15, followed by 76 000 (2015-16) and 20 000 (2016-17).

### ► Early generation selection: stage 2 BC1F3 → BC1F4 multi-location evaluation

Seed samples of BC1F3 families were grown by each of the five breeders in small observation plots, and scored for disease and agronomic characters during 2015-16. 343 lines were selected; due to the yield and quality issues at harvest, these did not go straight

into yield trials but were replanted at each location in 2016-17 and the exercise repeated on BC1F4 material. Additional lines were deselected and 308 lines passed on to Stage 3 (yield testing).



Figure 5: BC1F2 families grown as spaced plants at Maule (Secobra), June 2015.

New BC1F3 selections from the 2015-16 BC1F2 were grown at Secobra only; 116 of these were selected to move on to yield testing.

#### Stage 3: BC1F4/BC1F5 multi-location yield testing

Yield trials were carried on 424 experimental lines plus controls. These consisted of 308 BC1F5 lines (included in 1 rep at all five locations) plus 116 BC1F4 lines, included in 1 rep at 1-5 locations, depending on seed quantity.

The 116 BC1F4 lines were also grown in observation plots at three locations: Secobra, KWS Momont and Lemaire Deffontaines.

The major target traits recorded are:

- 1. Adjusted yield (ADJ\_YLD) in fungicide treated trials
- 2. Basic grain quality traits: specific weight (SW), grain protein content (PROT)
- 3. Agronomy traits: plant height (PH), heading date (HD), maturity (MAT), lodging (LODG)
- 4. Foliar disease resistance: yellow rust (YR), brown rust (BR), powdery mildew (MILD), septoria tritici blotch (SEPTO)
- 5. DUS characteristics: ear glaucosity (GLAU), presence/absence of awns (AWN)

Traits were analysed by fitting a mixed model with genotype as random effect to estimate variance components and calculate h<sup>2</sup>; with no replicates within a location, there was no possibility to get plot-level h<sup>2</sup>. A mixed model with genotypes as fixed effect was used to derive BLUE values.

The CETAC 5 are continuing to test material within their own programmes. Thirty-six lines were selected out of the yield trial, as a collective decision based on the balance of yield, diversity and other characteristics.

	Trait	h²
Yield	ADJ_YLD	0.400
Grain quality	PROT	0.684
urani quanty	SW	0.474
	PH	0.652
Agronomy	HD	0.845
	LODG	0.276
Dicoaco	YR	0.576
DISCASE	BR	0.280
DUS	AWN	0.972

Table 4: heritabilities for multi-location traits

The distribution of yields recorded is shown in Figure 6 and Table 5.



Figure 6: Adjusted yields, 2017-18.

	SC (78)	CS (82)	KM (28)	LD (59)	SUR (60)	5-loc avge
INDYS Min	47.9	29.8	28.5	60.9	68.8	53.3
INDYS Max	118.4	77	84.2	104	109.8	99.6
INDYS 36 Min	90	56	62	84.8	89	82.0
INDYS 36 Max	118.4	76.6	84.2	104	109.7	99.6
Std Min	97.6	65.3	75.6	92.7	95.4	87.4
Std Max	108.9	75.7	83.9	100.2	107.9	96.9
INDYS Avge	91.99	59.5	64.2	87.31	93.52	79.5
INDYS 36 Avge	100.8	66.9	72.8	94.05	101.2	88.0
Std Avge	102.6	72.2	79.3	96.91	102.2	93.0

Table 5: Breakdown of 2017-18 yields: performance of INDYS experimentals, INDYS selections and standards

Relative yields between INDYS and CTPS trials were compared for the CTPS North and South regions (Figure 7).



Figure 7: Relative yields of controls in INDYS and CTPS trials

Yields (expressed as % mean of standards) are shown for CTPS control varieties in (a) North and (b) South INDYS and CTPS trials Similar data analyses for other traits are shown in for the Supplementary Data as follows:

PROT (Figure S1, Table S1); SW (Figure S2, Table S2); HD (Figure S3, Table S3); PH (Figure S4, Table S4); MAT (Figure S5); LODG (Figure S6); YR (Figure S7); BR (Figure S8).

The disease traits powdery mildew and septoria tritic blotch were scored but showed little variation (data not shown). Data was also gathered for the DUS traits AWN and plant glaucosity; as these are essentially binary traits, the data is not shown.

#### Genotyping

Loss of information was very low during genotyping. Very few lines failed to germinate, or gave poor DNA samples; similarly, very few lines were discarded during genotyping quality control.

The breakdown of probes into the five classes varied between batches, as summarised in Table 6.

Background	Poly	Mono	OTH	NMH	BT	OTV
Expert	12254	9401	6334	4427	2266	375
Bastide	12687	8975	6323	4513	2336	395
Robigus	12380	7584	6519	6450	1808	402

Table 6: classification of Axiom 35k probes based upon standard analysis

As an example of the genotyping QC, two Robigus parental samples (Robigus\_1 call rate=99.641; 0.085% het; Robigus\_2 call rate=96.443; 1.38% het) were included: the contrast between the two, and consequences for genotyping, are shown in Table 7.

Robigus_1	Robigus_2	Instances	Action
AA	AA	9014	Retain probe
AA	AB	112	Check segregation in progenies; generally revert to AA
AA	BB	7	Discard
AA	No call	291	Check segregation in progenies; generally revert to AA
AB	AA	3	Check segregation in progenies; revert to AA
AB	AB	2	Discard
AB	BB	0	Not applicable
AB	No call	1	Discard
BB	BB	9396	Retain probe
BB	AA	2	Discard
BB	AB	124	Check segregation in progenies; generally revert to BB
BB	No call	268	Check segregation in progenies; generally revert to BB
No call	AA	9	Check segregation in progenies; generally revert to AA
No call	AB	0	Not applicable
No call	BB	4	Check segregation in progenies; generally revert to BB
No call	No call	1	Discard

Table 7: Conflicting allele calls between parental replicates

Upon inspection, most of the genotype calls revert back to the better Robigus\_1 sample. Most of the probes discarded were in the NMH or OTV classes.

Figure 7 illustrates a typical PolyHiRes genotyping result from the Axiom software.



Figure 8: allele calling output from Axiom software

The distribution of mapped / unmapped markers across chromosomes is shown in Table 8 for all three genotyping exercises.

Chromosomea	Bastide	Expert	Robigus
1A	403	403	431
1B	538	531	614
1D	244	209	222
2A	535	543	592
2B	682	623	758
2D	159	177	175
3A	415	396	730
3B	512	494	256
3D	54	49	51
4A	271	208	281
4B	254	261	290
4D	57	46	54
5A	510	483	550
5B	627	611	564
5D	136	140	250
6A	161	165	179
6B	468	450	511
6D	109	103	114
7A	314	329	339
7B	330	330	324
7D	57	52	53
7A/7Db	22	21	20
Other	10644	10449	11864
Total c (proportion) A	2609 (0.381)	2527 (0.383)	3102 (0.422)
Total c (proportion) B	3411 (0.499)	3300 (0.500)	3317 (0.452)
Total c (proportion) D	816 (0.119)	776 (0.118)	919 (0.125)

Table 8: Distribution of markers across chromosomes

<sup>a</sup> Based upon Avalon/Cadenza and Savannah/Rialto crosses, according to Axiom outputs; <sup>b</sup> according to Avalon/Cadenza; <sup>c</sup> unambiguous mapped markers for each genome (proportion of all unambiguous mapped markers in brackets)

#### 'Expert' material

For the 'Expert' subset of material, 17134 markers were analysed across 112 genotypes, including parental samples and replicates. Figure 9 shows graphical genotypes as a global view (Fig 9a) and in greater detail (Fig 9b). This indicates that there seem to be regions of the genome under positive selection pressure (a relatively high proportion of progenies carrying donor alleles or heterozygous) and whilst other regions are under apparent negative selection pressure (generally recurrent parent genotype). For example, expanding in on some of the 1D markers shows contrasting patterns of inheritance (Figure 10, positive selection; Figure 11, negative selection).



Figure 9: Graphical genotypes for 'Expert' material

Graphical genotypes for 112 materials, with genotypes shown as rows and markers shown as columns. Parental lines are shown at the top: three replicates of 'Expert' (above first horizontal line), followed by NIAB SHW parents (between horizontal lines) the progenies (beneath second horizontal line). Genotypes are indicated as green (same SNP allele as 'Expert'), red (homozygous for alternative SNP allele), blue (heterozygous) and white (missing data). Schematics are shown for (a) all markers and (b) those mapped to chromosome 1D.

The FLAPJACK 'Pedigree Verification' tool counts matches to the presumed parents for all genotypes across all markers. Figure 12 shows it has identified outcrossed lines in the Expert / SHW-029 progenies, with the remainder carrying 70-90% of the parental Expert genome (mean 80.2%).



Figure 12: Parental contributions to Expert / SHW-029 progenies

Parental contributions calculated using the proportion of 'Expert' alleles relative to the number of Expert + SHW-029 alleles, across all polymorphic markers for the cross, using the FLAPJACK 'Pedigree Verification' feature. Probable outcrossed lines are highlighted.

#### 'Bastide' material

For the 'Bastide' subset of material, 17503 markers were analysed across 121 genotypes, including parental samples and replicates (Figure 13).

			T																																				
			1			1		-																		100		-			-	_			_	_	_		
3_Expert_1	A	C	C	C	G	T	T	A	ŧτ	A	C	C	G	C	G	Т	T	A	T	A	G	G	T	F (	A	A	T	Т	T	A	A	C	T	A	G	T	T	A	A
4_Expert_2	A	C	C	C	G	T	Т	A	T	A	C	C	G	C	G	T	T	A	T	A	GI	G 1	TT	T (	A	A	T	Т	T	A	A	C	Т	A	G	Т	Т	A	A
Expert_3	A	C	C	C	G	T	T	A	T	A	C	C	G	C	G	Т	T	A	T	A	GI	G	T	Г	CA	A	Т	Т	T	A	A	C	T	A	G	T	T	A	A
Expert 4	A	C	C	C	G	T	T	A	T	A	C	C	G	C	G	T	T	A	Т	A	GI	G	T	T C	CA	A	Т	Т	Т	A	A	C	Т	A	G	T C	T	A	A
NIAB_SHW 027	Ag	С	C	T	C	T.		G	T.	6.4%	A	C		C	G	Te	C	G	C	A	A	G I	-	C .	G		C	C	Te	Alla	A	T			9	C		G	C
NIAB SHW 029	G	T	T	T	G	Т	T	A	Т	C	A	C	A	A	G	C	C	G	C	A	100	A	T (	0 1	G	T	C	Т	T	A	A	C	C	T	C	C	Т	G	A
NIAB SHW 036	G	T	T	T	C	C	T	A	C	C	A	G	A	A	G	C	C	G	C	A		A	T	0 1	T G	T	C	Т	Т	A	A	C	C	T	C	C	т	G	A
NIAB SHW 043	G	T	C	T	C	C	C	A	T	C	A	C	A	A	T	C	C	G	C	A	A	A 1	CO	6 7	T G	Ť	C	T	T	A	A	C	C	T	C	C	T	G	A
NIAB SHW 051	G	Ť	T	T	č	C	Ť		C	C	A	G	A	A	G	C	C	G	C	A		A	T	6 5	T G	Ť	C	Ť	T	A	A	C	C	Ť	ĉ	C	τI	G	A
NTAR SHN 070	G	T	C	T	č	T	0	A	C	č	A	G	A	A	G	č	č	G	C	A	A	A 1	C (	6 7	T G	Ť	C	Ť	ET .	A	A	T	č	T	č	č	τÌ	G.	A
2 NTAR SHW 073	G	Ť	T.	T	c	Ť	C	4	Т	C.	A	C	A	A	T	č	č	G	č	A	Δ ]	Δ 1	c i	2 1	T G	Ť	C.	Ť	Ť	Δ	A		Č.	Ť	C	č	т	-	A
NTAB SHN 074	G	Ť	C	Ť	č	Ť	č		Т	Δ	A	č	G	C	G	T	T	Ğ	č	2	A		T		- 6	Ť	č	C.	Ť	A	A	T	č	Ť	č	T	Ť	A	4
S NTLB CHW 077	G	Ť	C	T	č	÷	G		-	-	2	C.	G	č	G	Ť	-	G	č	2	2		÷ i	2	6	÷	č	č	Ŧ	A	A	÷	č	÷	č	÷.	÷ i	2	-
12 NTAB SHW 078	G	Ť	č	Ť	č	÷	č		T	A	A	č	G	č	G	T	T	G	č		2	č i	Ť l	č (	6	T	C.	č	Ť	A	A	÷.	č	T	õ	Ť	T	2	6
A NTER GUY ORE	G	÷	č	÷	č	÷	č		Ť		-	č	G	č	G	T	+	G	č	2	2		÷ i	2		Ť	č	č	÷	2	2	÷	č	T	č	C.	÷ i	2	6
THINK SEAL	G	+	Ť	÷.	č	C	č			6	2	G	4	A	T	C	C	G	č	21	2	A .	0		r G	÷	č	T	C	2	2	C	T	-	G	T	4	2	2
	0	+	+	-	č		0		č	č	2	6	2	12	-	č	č	0	č	21	2		č			+	č	ċ	č	2	2	č	+	2	0	T	- 1	2	2
_18015_3575	0	0	0	-	0					-	0	0	0	A	-	T	T	0	U.	2	2 1		T				T	T	U	~	2	č	-	~	0	T	-	2	1
5_1ND15_3549	A	0	0	2	0	4	-		-	2	5	5	0	0	0	+	-	2	+	2	G		÷ .			10	+	+	-	-	2	0	+	~	0	-	+	2	0
5_INDIS_3559	-	00	0	6	0	4	4		4	A	5	6	0	0	6	-	-	A	+	~	6	0	÷			10	4	+	4	A	~	č	+	A	0	1	-	~	2
S_INDIS_3562	2	0	0	2	0		-			10	5	6	6	5	0	+	-	2	+	2	6	9	+ -			12	-	+	-	2	2	5	+	2	0	+	-	2	0
9_INDIS_3573	-	5	~	-	0					-	6	0		~	0			~	-	2	G	2		1	- <u>^</u>	~	-	+		~	4	0	4	~	0	-	1	2	2
76091_01		-0	10	10	0					-		č			0	6		0	0	2			-	9	0 0	-	0	+		2	2	2	+	2	0	+	- 1	2	2
0_76097_UT	G	1	-		0						A	6	A	A	6	6	5	6	6	~	200. /	A	+ +		1. 0		6	4		~	1	5	4	4	2	1	1	2	1
0_76115_UT	4	G	C T	4	G		1			A	6	6	6	6	6	-	-	2	-	A .	G	9	1				-	+		A	4	C	-	2	G	-		2	0
0_76125_UT	G				G					C	A	C	A	A	6	10	6	6	G	A.	2.94	A.	1		[ G		G			A	10	C	C.		G	G.	1.1	6	0
_75981_UT	A	C	C	C	G			1		A	C	G	G	C	G	1		A	1	A	G	6	1	1 0	A	A	1	1	1	A	A	C	1	A	G	1	I I	6	2
0_76163_UT	A	C	C	C	G			1		A	C	C	G	C	G	1	1	A	T	AL.	G	G	T.		CA	2.1	-	T	1	A	A	C	1	A	G	Т	T	<u>A</u> .	A
_76069_UT	G	T.	Π.	1	G		1	- P	1	C	A	C	A	A	G	C	C	G	C	A.	10	Α.	T	6	G	T	C	T	T.	A	A	C	C	Т.	C	C	13	G	A
_76165_UT	A	C	C	C	G	I	1	A	1	A	C	C	G	C	G	T	T.	A	T	A	GI	G	1		C A	A	1	T	T.	A	A	C	T	A	G	T	T.	A	A
_76071_UT	A	C	C	C	G	1	I	A	1	A	C	C	G	C	G	Т	T	A	T	A.	GI	G	T	r (	A	A	Т	T	Т	A	A	C	T	A	G	T	T .	A	A
5_7€087_UT	G	T	T	1	G	1	1	A	I	20	: 2.0	C	10	20	G	C	100	2 a	1 C	A	G	6	T 🛃	C '	c ? .	1.1		T	Т	A	A	C	T	A	G	T	Τ.	A.	A
1_65829	G	I	Л.	1.	G	1		A	1	A	C	C	G	C	G	Т	T	A	Т	A.	G	G	I		A	A	UT.	T.	Τ.	A	A	C	Т	A	G	T	Ι.	A	A
3_65961	G	d	10	1.0	5		T	A	10		C	C	20	C	100	C	T	A	1	A	G	G	T i		c A		T	T	T	A	A	C	C	A	G	10	T.	A.	A
1_65843	G	T	T	T	G	T	1	A	T	C	A	C	A	A	G	C	C	G	C	A	0	A I	T (	0 1	T G	T	C	Т	T	A	A	C	T	A	G	T	T.	A.	A
11_65847	G	T	T	T	C	C	C	G	C	C	A	G	A	A	T	C	C	G	C	A	A	A	C	0 1	T G	T	C	C	C	A	A	C	T	A	G	T	T	A	A
1_65851	G	T	T	T	G	Т	T	A	T	C	A	C	A	A		C	C	G	C	A	101	A T	6	0 1	T G	A	T	Т	Т	A	A	C	T	A	G	Т	T	A	A
1_65859	G	T	T	T	G		T	A	T	A		C		C			T	A	T	A	GI	G	T .	Г	A	A	T	Т	T	A	A	C	Т	A	G	T	T	A	A
1_65903	G	T	T	T	G		T	A	T	C	A		A	A	T.	C	C			A	20	AT	c (	0 1	G	A	T	T	T	A	A	C	Т	A	G	T	T	A	A
11 65935	G	T	T	T	G	T	T	A	T	A	C	C	G	C	G	T	T	A	T	A	GI	G	T	T (	CA	A	T	Т	T	A	A	C	T	A	G	Т	Т	A	A
19 65827	G	T	T	T	C	C	C	0	C	C	A	G	A	A	T	C	C	G	C	A	A	A I	CO		r G	T	C	C	C	A	A	C	T	A	G	T	T	A	A
00033 11	G	T	T	T	G	T	Т		т	A		0	6	100	0	0	0	6	T		0.4		-	2 2			T	-	-	1		-	-		0	T	T		

Figure 10: Example of positive selection for SHW alleles on chromosome 1D

Screenshot showing detail of genotypes across chromosome 1D. Boxed region shows inheritance of SHW alleles at 4 SNPs which seem to have been actively selected for in Expert / SHW-029 progenies.

	h:	_																													
	1 1		-				-	22	22	20		2	-			-							_	_	_	_					
M23 Expert 1	AT	C	C	C	T	C	C	T .	A A	NE /	AA	A	A 4	A	A	A	A	C	C	C	C	C	C	C	C	C	C	C	C	6 6	G
M24 Expert 2	A T	C	C	C	Т	C	C	T	A A		A J	Α,	A A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	C	C	GG	G
N23_Expert_3	AT	C	C	C	Т	C	C	T	A A	57	A A	A . /	A A	A	A		A	C	C	C	C	C	C	C	C	C	C	C	C	GC	G
N24 Expert 4	AT	C	C	C	Т	C	C	T	A A	1	A A	4 /	A A	A	A	A	A	TC	C	C	C	C	C	C	C	C	C	C	C	GO	G
CO9_NIAB_SHW_027	GT	C	C	C	C	Т	C	T	6 A		A J	A /	A C	9	A	A	A	C		C		C D	C	C	C	C	T C	C	C G	G A	A 0
M11_NIAB_SHW_029	GT	-	C	C	C	T	C	T	G A	1	C 1	1		C	G	G	G	C	T	A	G	G	C	T	C	T	C	C	C	AA	A
M19_NIAB_SHW_036	GT	Т	C	C	C	Т	C	C A	0 4		A	T P	. 0	C	A	10	G	C	C	A	G	G	C	T	C	C	C	C	C	AA	1 0
002 NIAB SHW 043	GT	Т	C	C	C	T	C	T .	A A	1 (	0 1	T 👌	8 0	C	G	G	G	C	T	A	G	G	C	Т	C	Т	C	A	C	A A	A
A08_NIAB_SHW_051	GT	T	C	C	C	T	C	0	0 4	1	A	1		C	A	G	G	C	C	A	G	G	C	T.	C	C	C	C	C	AA	A
KOS_NIAB_SHW_070	GT	T	C	C	C	T	C	T	A A	1	A	T A	0 0	C	G	G	G	C	C	C	G	G	C	Т	C	C	C	C	C	AA	A
2A12_NIAB_SHW_073	GT	T	C	C	C	T	C	Т	F	1	A	F A	. 0	C	G	G	G	C	C	C	G	G	C	T	C	C	C	C	C	A A	A
2E12_NIAB_SHW_076	GJ	T	C	C	C	T	C	T	0 4	1	A	T I	A A	C	A	G	A	C	C	A	C	G	C	T	C	C	C	C	C	A A	G
2G12_NIAB_SHW_077	GT	T	C	C	C	T	C	T	6 A	1	A	F /	A A	C	A	G	A	C	C	A	C	G	C	T	C	C	T C	C	C	A A	G
2I12_NIAB_SHW_078	GT	T	C	C	C	T	C	T	A A	1	A	T ,	A.	C	A	G	A	C	C	A	C	G	C	C	C	C	T c	C	C	AA	
2C14_NIAB_SHW_085	GT	Т	C	C	Т	T	C	T	A A	1	A 1	Γ.,	A A	C	A	G	A	C	C	A	C	G	C	C	C	C	T C	C	C	G A	G
B11_INDYS_3581	A T	C	C	C	C	T	C	T	G A	1	A A	A ./	A 4	A	A	A	A	T	C	C	C	C	A c	C	C	C	C	C	C	GC	G
D11_INDYS_3595	AT	C	C	C	Т	C	C	Τ.	A A		A A	4 1	A A	A	A	A	A		C	C	C	C	C	C	C	C	C	C	C	GC	G
H09_INDYS_3549	GT	C	C	C	C	T	C	T	0 4		A J	4	A A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	C	C	GC	; G
L09_INDY5_3559	A T	C	С	C	Т	C	C	T	A A		A A	4 /	A 4	A	A	A	A		C	C	C	C	C	C	C	C	C	C	C	GG	G
N09_INDYS_3562	AT	C	C	C	T	C	C	T	A A	1	A /	A /	A A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	C	C	GC	G
P09_INDYS_3573	AT	C	C	C	Т	C	C	T	A A	1	A A	A 9	A 4	A	A	A	A	C	C	C	C	C	C	C	C	C	C	C	C	GC	G
B10_76091_UT	GT	C	C	C	Т	C	С	T	A A	1	A A	4	A 4	A	A	A	A	C	C	C	C	C	C	C	C	C	C	C	C	GG	G
D10_76097_UT	GT	C	C	C	Т	C	C	T	A A		A 4	A /	A 4	A	A	A	A	C	C	C	C	C	C	C	C	C	C	C	C	GG	G
F10_76115_UT	GT	C	C	C	Т	C	C	T	A A		A A	A /	A A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	C	C	GC	G
H10_76125_UT	A T	C	C	C	Т	C	C	T	A A		C		A 4	C	G	G	G	C	C	A	G	G	C	T	C		C	C	C	G A	A
J08_75981_UT	AT	C	C	C	Т	C	C	T	A A		A /	A 1	A A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	C	C	GG	G
J10_76163_UT	GT	C	C	C	T	C	C	T )	A A		A /	A 🥖	A A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	C	C	GC	G
L08_76069_UT	A T	C	C	C	T	C	C	T	A A		6		A 4	C	G	G	G	C	C	_	G	G	C	T	C		C	C	C	GA	A
L10_76165_UT	GT	C	C	C	T	C	C	Τ.	A A		A /	A ()	A A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	C	C	GC	G
N08_76071_UT	AT	C	C	C	T	C	C	T	A A	1	A A	A /	A A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	C	C	GG	G
P08_76087_UT	GT	C	C	C	T	C	C	T	A A	1	A	A /	A /	A	A	A	A	C	C	C	C	C	C	C	C	C	C	C	C	GG	G
2A21_65829	AT	C	C	C	T	C	C	1	A A	1	A	A /	A A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	C	C	6 0	G
2A23_65961	AT	C	-	C	T	_	C	1	A A	1	A A	A /	A A	A	A	A	20	10	C	C	C	C	C	C	C	C	C	C	C	GG	G
2C21_65843	AT	C	C	C	T	C	C	1	AA		A	A /	A /	A	A	A	A	C	C	C	C	C	C	C	C	C	C	C	C	GG	G
2E21_65847	AT	C	C	C	1	C	C	1	A A	1	A /	A /	A A	A	A	A	A	T	C	C	C	C	C	C	C	C	C	C	C	GO	G
2G21_65851	GT	C	C	C		C	C	1	A A		A /	A /	A A	A	A	A	A	C	C	C	C	C	C	C	C	C	C	C	C	GG	G

Figure 11: Example of negative selection against SHW alleles on chromosome 1D

Screenshot showing detail of genotypes across chromosome 1D. Boxed region shows inheritance of SHW alleles at SNPs which seem to have been actively selected against in Expert / SHW-029 progenies.

95



#### Figure 13: Graphical genotypes of Bastide material

Graphical genotypes for 121 materials, with genotypes shown as rows and 17503 markers shown as columns. Parental lines are shown at the top: three replicates of 'Bastide' (above first horizontal line), followed by NIAB SHW parents (between horizontal lines) and progenies (beneath second horizontal line). Genotypes are indicated as green (same SNP allele as 'Bastide'), red (homozygous for alternative SNP allele), blue (heterozygous) and white (missing data).

#### Robigus material

In the Robigus background, 19222 markers were analysed across 301 genotypes, including parental samples and replicates (Figure 14).





Graphical genotypes for 301 materials, with genotypes shown as rows and 19222 markers shown as columns. Parental lines are shown at the top: two replicates of 'Robigus' (above first horizontal line), followed by NIAB SHW parents (between horizontal lines) and progenies (beneath second horizontal line). Genotypes are indicated as green (same SNP allele as 'Robigus'), red (homozygous for alternative SNP allele), blue (heterozygous) and white (missing data).

#### Selected lines

The graphical genotypes of 31 of the selected lines, in a Robigus background, are shown in Figure 15.



Figure 15: Graphical genotypes of selected lines

Graphical genotypes for Robigus selections, with genotypes shown as rows and 19222 markers shown as columns. Parental lines are shown at the top: two replicates of 'Robigus' (above first horizontal line), followed by NIAB SHW parents (between horizontal lines) and selected progenies (beneath second horizontal line). Genotypes are indicated as green (same SNP allele as 'Robigus'), red (homozygous for alternative SNP allele), blue (heterozygous) and white (missing data).

#### Combined genotyping approach

Bringing together all three genetic backgrounds resulted in a common marker set after Axiom filtering of 13,166 markers. After cleaning on the basis of missing values, minor allele frequency and heterozoygosity, this gave a final marker set of 7998 markers. Around 200 markers were dropped because they gave different allele calls for single samples, depending on which genotyping batch they were included in (Figure S9).

Five lines have more than 10% missing values, and twelve lines have more than 10% heterozygous calls. Two lines were clearly 'Bastide' selfs.

Multidimensional scaling analysis uncovered clear structure in the samples, with three pools of highly related material representing the three genetic backgrounds, and a fourth cluster of SHW donor lines (Figures 16 and 17).



Figure 16: Multidimensional scaling analysis: axes 1 and 2



Figure 17: Multidimensional scaling analysis: axes 1 and 3

#### QTL analysis

The Cereals DB genetic map locations were used as a basis for GWAS analyses. Manhattan plots are shown for ADJ\_YLD (Figure 18) and AWN (Figure 19).

For other traits, GWAS results are shown in Supplementary Data as follows:

PROT (Figure S10), SW and HD (Figures S11-S12); PH and LODG (Figures S13-S14); YR and BR (Figures S15-S16); GLAU (Figure S17).



Figure 18: GWAS for adjusted yield



Figure 19: GWAS for presence/absence of awns

#### 4. Discussion

The project has been successful in introducing valuable new genetic diversity into French breeding programmes. Early genotyping results suggest that the lines tested in yield trials each inherited approximately 20% of the SHW donor genome, compared to the 25% expected by chance in BC1-derived material (Figure 12).

#### Hybrid necrosis and population development

Hybrid necrosis (HN) can be a feature of wide crossing programmes in wheat (Zeven, 1981). Parental lines appear to be perfectly healthy, but their F1 seedlings develop patches of dead leaf tissue which spread as the plants grow, and can lead to affected F1 plants producing no seed. HN is caused by two complementary dominant genes (*Ne1* and *Ne2*), which have been mapped to approximately 10 cM intervals on chromosome arms 5BL and 2BS, respectively (Chu et al., 2006).

NIAB ran a large pre-breeding programme to explore SHWs developed by CIMMYT. Many of the F1s from crosses between the UK variety 'Xi19' and SHWs showed high HN levels; some had such extreme symptoms that entire lineages did not progress beyond the F1 generation (Howell et al., 2017). Subsequent crosses between these SHWs and elite varieties showed that F1s with most UK winter wheats displayed high, often lethal, HN levels, with the notable exception of the variety 'Robigus'. This is the main reason 'Robigus' was chosen as a key parent for NIAB's germplasm development within WISP, the UK public-sector wheat pre-breeding programme (Howell et al., 2017).

For this reason, a pre-project HN screen was used within this project. The breeder partners proposed a panel of thirty elite French varieties which were crossed against a selection of CIMMYT SHWs known to give rise to F1s displaying necrosis. The resulting F1s were scored for HN symptoms and two varieties, Expert and Bastide, were selected on the basis of their relatively low scores.

As crossing work for the project progressed, it became clear that whilst HN levels had been lower in Expert and Bastide than most of the other varieties screened, in practice this was still hampering population development: many F1 and BC1 plants were lost to hybrid necrosis, especially during and after vernalisation. For this reason, additional material was generated in the 'Robigus' background, as WISP crossing had shown that this material was free from HN symptoms. However, this delayed the population development such that BC1F2 material was distributed to the breeders across 3 years, in order to meet the targets outlined in the project proposal.

#### Early generation selection

Typical commercial wheat breeding programmes follow a pedigree system, with hundreds of thousands of plants from around 1000 crosses screened at the F2 generation, progressing to tens of thousands of ear rows (F3), thousands of ear-row families (F4) and yield testing and purification beginning at the F5 generation (1000 lines), at which point most lines appear relatively uniform.

Early generation selection in pre-breeding material from wide crosses such as within INDYS requires a different approach. There is much greater segregation for a much wider range of phenotypes than is typical for commercial programmes - characters which commercial breeders rarely encounter, such as black or red chaff, tenacious glumes and extreme height differences. Often this requires a modification of the pedigree system, focusing on single plant progenies instead of ear-row families.

Because there is so much variation it is also difficult to balance the requirement for high selection pressure, to manage numbers carried forwards in the programme, with the wish to maintain the diversity so painstakingly captured in the pre-breeding lines. Rather than expecting yield levels in excess of elite varieties, it is more realistic to hope to recover pre-breeding lines which capture genuine novel diversity but yield at around 95-100% of the recurrent parent level.

The planned yield trials of 2016-17 were deferred by one year due to the unusually poor weather encountered across key wheat growing regions during spring and summer 2016 (Figure 20).



Figure 20: Cool temperatures and high rainfall across France, early grainfill 2016

Deviations from the 20-year mean (1996-2015) are shown for (a) temperature and (b) rainfall for the period May 20 - June 20, 2016 (data from Meteo France, redrawn from Arvalis report).

The resulting low yields and high incidence of ear disease, especially *Michrodochium nivale*, led to the 2015-16 multiplication plots producing grain of insufficient quantity and quality to plant into yield trials. Weather conditions had improved by the autumn planting period, so there was no lasting damage to the programme into subsequent years. By repeating the stage 2 multi-location testing in 2016-17, additional selection and observation was possible.

#### Yield performance

As a whole, the INDYS trials appeared to match the CTPS trial series relatively well (Figure 7). Despite the regular performance of the controls, the yield levels of INDYS lines were generally lower in all trials (Figure 6, Table 5) than would be typical for elite breeding lines, reflecting the diversity of the material tested and the selection criteria being used. A high proportion of INDYS lines tested are in the late maturing Robigus background. This probably explains the particularly low yields in the southernmost CS trial: later maturing types rarely fulfil their potential in CTPS South trials, as high temperatures, rapid crop senescence and terminal drought often coincide with their mid- to late-grainfill period, whereas earlier varieties pass through these stages sooner and thus escape the abiotic stresses. The lowest yield recorded came from a line deliberately selected to retain negative traits from the synthetics.

Some high yielding INDYS lines were apparent across all five trials; the average yield levels of the 36 INDYS selections was only 5 qx/ha behind the control average (Table 5), with this yield deficit ranging from 1-10 qx/ha across locations. This deficit increases in the more southern trials, reflecting the high proportion of 'Robigus' lines within the set of 36 selections.

Each trial included INDYS lines which yielded more than the highest control variety. However, INDYS lines were not replicated within sites whilst control varieties were highly replicated, so additional trials are required to determine their true yield potential. Replication across sites at least gave some indication of heritability and yield stability.

### Other traits

Heritabilities for specific weight were lower than expected (Table 4); this probably reflects the impact of terminal drought on the grainfill of late varieties in the CS trial. In addition, specific weight was measured in some trials using the on-combine weighing system, which can be adversely affected by how well the grains thresh out, but for other trials it was measured on grain samples which had passed over a seed cleaner. The three parental backgrounds, Expert, Bastide and Robigus, have only moderate specific weights, so the presence of some good progeny lines indicates that SHW are bringing positive variation for this trait.

For other agronomic and quality traits, variation was more or less as expected. For example, heading date was much earlier in the southern CS location than the other locations (Figure S3, Table S3). Lodging was very sporadic, with some tall weak varieties susceptible in each trial; the only trial to show appreciable differential lodging was the 2017 SUR nursery (Figure S6).

Managing disease levels is often a problem in pre-breeding material, which often makes use of unadapted exotic parents. Within elite breeding programmes, there is usually a strong focus during early generations on selection for disease resistance. For example, F2-F4 breeding nurseries are generally left untreated, and include 'spreader' varieties (susceptible to key foliar diseases) at regular intervals, into which fungal inoculum is introduced to build an artificially high disease pressure. This simplifies the selection of resistant progenies. In contrast, early generation pre-breeding nurseries are often treated with fungicides, to ensure that useful novel diversity for other performance traits is not masked by generally high disease levels.

This was exaggerated for INDYS: as well as the influence of the exotic SHW parents, a large number of INDYS lines were grown in the very susceptible 'Robigus' background to evade hybrid necrosis problems. This led to the collection of a lot of disease scores, especially for yellow rust and brown rust. Whilst some data was collected for other diseases, a wider testing network is required to build up a fuller picture. In this way, some novel resistances might be detected; ideally this should happen on residue seed from earlier generations, in order to capture as much diversity as possible.

Observations for the binary DUS traits presence/absence of awns and ear glaucosity were useful in checking that trials and nurseries had been drilled correctly. Both the presence of awns and the glaucous appearance of ears are generally recessive characters. The very high heritability of the AWN trait (Table 4) indicates that trials were planted and scored in the correct sequence.

### Genotyping

The recent development of large SNP arrays in wheat has considerably reduced the price per genotyping data point. The Axiom 35K Wheat Breeders Array used here was designed around inter-varietal SNPs and is commonly used by the UK wheat genetics and breeding community. However, there is inevitably ascertainment bias in any such array and it is clear from our genotyping results that markers which map to the Dgenome chromosomes are under-represented (11-12% of the total mapped markers; Table 8). Other products such as the Axiom Wheat 820k HD Genotyping Array have been developed using exome capture of a wider selection of source materials, often including wild relatives and land races. These may be more suitable for genotyping pre-breeding material, but due to high costs are not routinely used. It is somewhat ironic that prebreeding projects which target greater D-genome diversity, such as those based on SHWs, are still hampered by an inherent lack of D-genome diversity in design of highthroughput genotyping platforms.

The underlying structure within the INDYS genotypes, exemplified in Figures 16-17, clearly shows the four clusters of related material: Robigus, Bastide and Expert derivatives, plus NIAB SHWs. It is likely that the diversity within the NIAB SHWs is being underestimated. Most of these SHWs are based on the AB-genomes of Hoh-501, a German winter durum type bred at University of Hohenheim, but with different D-genomes from diverse *Ae. tauschii* accessions (Jones et al., 2013). As we have seen, the Axiom 35k array is biased towards polymorphisms on the A and especially B genomes and thus will be underestimating the NIAB SHW pool of D-genome variation.

The genotyping results reported here are a first sift of the variation within the population. A more detailed analysis will be carried out to determine the best subset of markers, carry out thorough quality control of markers and samples, thoroughly check against pedigree records and fieldbooks, and make a better attempt at assigning markers to chromosomes and

ordering them into the correct sequence. This will be essential for more detailed trait dissection work and will certainly continue after the end of the INDYS project.

However, these primary results do give an interesting snapshot of the diversity captured. In contrast to most published studies, this work has been carried out on 'live' pre-breeding selections rather than on unselected experimental populations. We would therefore expect to see patterns of selection in the progenies genotyped: for markers where a high proportion of progenies carry alleles inherited from the SHW donor parent, this suggests that these markers are linked to genes where the donor allele confers an advantage which breeders have selected for. The global overviews suggest that this is the case in all three backgrounds: Figures 9, 13 and 14 show a high proportion of red colouration (indicating SHW alleles) running down vertical columns, indicating markers where high proportions of progenies have inherited SHW alleles. Closer scrutiny shows examples of apparent selection both for SHW alleles (Figure 10) and against SHW alleles (Figure 11) on chromosome 1D in Expert / NIAB SHW-029 material. Selection patterns such as these, which are coincident with significant phenotype differences for key traits, should prove to be interesting areas for future trait dissection projects.

#### QTL mapping

DNA was extracted from single BC1F4 plants, whereas most of the phenotypes were carried out on BC1F5 bulks (each tracing back to single BC1F3 plants). There will inevitably be discontinuities between genotype and phenotype due to sampling and genetic drift: for example, if a common progenitor was heterozygous but contrasting alleles have been inherited by the genotyped and phenotyped samples. These are more likely to result in increased experimental noise, and thus false negatives, than false positives. Any hits identified in the preliminary GWAS reported here are likely to be genuine, especially as they are based on BLUE trait data derived from multiple locations. Here, we largely concentrate on hits with log10(p) values above a very stringent threshold of 6.0.

Scans will be repeated for each location/year that data has been collected for, in order to build up a more robust picture of the number, significance and stability of QTLs detected. It is reassuring to see that the AWN trait gives a massive QTL peak on chromosome 5A (Figure 19): this is almost definitely the *Tipped (B1)* locus on 5AL, also detected in other studies (Mackay et al., 2014; Yoshioka et al., 2017). Whilst this is not a novel result, it acts as a good sense-check that the phenotyping, and genotyping and GWAS are correct.

The most significant GWAS hit for yield was from an unmapped SNP (Figure 18). Because this trait was based on multi-location BLUE data, it is also an indication of yield stability across contrasting environments, and will certainly be investigated further. Other hits for yield were less significant, but still with  $\log_{10}(p)$  values >4.0, for example on chromosomes 1D, 2B, 2D, 3B and 5D (Figure 18). These may become more significant when individual locations are analysed. All five of these regions have been previously associated with QTLs for yield and yield components (for example Assanga et al, 2017; Azadi et al, 2014; Li et al, 2007).

Significant GWAS hits for grain protein were detected on chromosomes 3A and 5B (Figure S10). A recent study, also based on SHW-derived materials, also detected grain protein QTLs on these chromosomes (Nedelkou et al., 2017). A significant hit was found on chromosome 5A for specific weight (Figure S11). Further work is required to follow this up, it will be interesting to see if this is in a similar position to a QTL for grain size on 5A, which is known to influence pericarp cell size and thus grain shape and size (Brinton et al., 2017), which can both influence specific weight.

Three hits are found for heading date (Figure S12). The most significant (chromosome 2B) is likely to be the major flowering time gene *Ppd-B1*, which affects flowering time through the interaction with daylength through a copy number polymorphism (Diaz et al., 2012). This variant is very uncommon in European wheat germplasm; if this is the causal gene then it demonstrates that SHWs can also being in useful variation from sources other than the D-genome.

Similarly, a secondary hit on 2D is likely to be its homoeologue *Ppd-D1* (Beales et al., 2007). A third hit on chromosome 4A may be a QTL for heading date detected in a recent Canadian study (Fowler et al., 2016).

Plant height is generally under the major control of the GAinsensitive 'semi-dwarfing' genes *Rht-B1* and *Rht-D1* on chromosomes 4B and 4D, respectively (Börner et al, 1997). It is likely that the GWAS hit for PH on chromosome 4D (Figure S13) is *Rht-D1b*, carried by Expert and Bastide. The highly significant hit on an unmapped marker may be *Rht-B1b*, carried by Robigus. As diagnostic markers exist for both loci, this should be relatively simple to resolve; adding them as co-factors should enable minor QTLs to be detected. It is likely that the coincident significant hits for lodging (Figure S14) are also due to these plant height effects; again, correcting for *Rht* status should allow the detection of minor QTLs. Other significant hits for LODG are found on 7B, which may be related to another height QTL (Cadalen et al., 1998).

Several significant hits are detected for yellow rust, on chromosomes 1B, 3A, 3B and 4B (Figure S15). Hundreds of YR resistance QTLs have been mapped across most of the genome: for example, a recent association study of global spring wheat varieties detected 130 significant SNPS, representing 70 unique QTLs (Muleta et al., 2017). Possible candidates for GWAS hits here include the recently cloned gene *Yr15* (1BS; Klymiuk et al 2018), the durum-derived resistance *Yr24/26*, introduced through SHWs (McIntosh and Lagudah, 2000) or a host of QTLs on 1BL; field resistance QTLs on 3B and seedling resistance QTLs on 4A (Muleta et al., 2017).

#### Next steps

Following the trial results of the 2017-2018 campaign, the breeders decided to make a selection of 36 lines with a 3-fold goal: a) characterize thoroughly the best lines for yield and diseases in order to incorporate them in the most efficient way into their respective breeding programmes; b) have a subset of the lines created in the project representing good agronomic value as well as diversity, to incorporate them in further studies for traits like NUE or drought tolerance, for which SHW-derived varieties are known to carry favourable QTLs; c) use them to create populations to dissect the QTLs identified here and in the subsequent analyses.

The thirty-six lines selected have already been planted in 2018-19 replicated yield trials and disease observation nurseries with each breeder; quality analyses (proteins, alveograph, SDS test) will be performed on the harvested grains. Maintenance of pure stocks is with Secobra and KWS Momont. The last sixty-four new BC1F4 selections, tracing back to the BC1F2 first grown in 2016-17, are in 2018-19 yield trials at Secobra, and disease nurseries with other breeders.

## 5. Supplementary Figures





Figure S2: Specific weight, 2017-18









Figure S4. Plant height, 2016-17 and 2017-18



Figure S5: maturity scores from SC trial, 2017-18

100







Figure S7: Yellow rust

Yellow rust observations from successive cohorts: a) material in BC1F5 trials 2018, tested in nurseries 2017 and 2018; b) material in BC1F4 observations 2018



Figure S8: Brown rust

Brown rust observations from successive cohorts: a) material in BC1F5 trials 2018, tested in nurseries 2017 and 2018; b) material in BC1F4 observations 2018



Figure S9: Genotyping artefacts caused by miscalling of alleles across batches

Four elite lines (77563UT, 77586 UT, 77588 UT and 77592 UT) were included within all three genotyping exercises.

Axiom software changed the allele calling (left panel) resulting in the appearance of related triplets in dendrograms (right panel).



Figure S10: GWAS for grain protein content



Figure S12: GWAS for heading date



Figure S14: GWAS for lodging

Figure S11: GWAS for specific weight



Figure S13: GWAS for plant height







Figure S16: GWAS for Brown Rust

## 6. Supplementary Tables

	SC (78)	KM (28)	LD (59)	SUR (60)	4-loc avge
INDYS Min	10.5	9.7	10.4	11.8	10.9
INDYS Max	16.6	16.4	15.5	15.2	15.2
INDYS 36 Min	11	10.2	11.1	12.1	11.4
INDYS 36 Max	12.6	12.9	13.2	14.1	12.9
Std Min	11.4	10.7	11.3	12.4	11.7
Std Max	12.9	13.1	12.8	13.4	12.7
INDYS Avge	12.3	11.5	12.3	13.3	12.3
INDYS 36 Avge	11.9	11.1	12	12.8	12.0
Std Avge	12.1	11.4	12.2	12.9	12.1

Table S1: Grain protein content

	SC (78)	CS (82)	KM (28)	LD (59)	SUR (60)	5-loc avge
INDYS Min	72.2	63.7	59.8	66.1	68.7	68.1
INDYS Max	82.5	76.7	77.5	80.6	82.4	79.6
INDYS 36 Min	77.5	65.9	66.2	73.6	73.2	72.0
INDYS 36 Max	82.2	74.4	77.5	79.4	81.1	79.1
Std Min	80.4	71.1	70.6	77.8	77.9	75.7
Std Max	84.8	78	79.5	82.6	85.3	83.0
INDYS Avge	79.3	70.6	70.9	76.1	77.2	74.9
INDYS 36 Avge	79.9	71	72.1	76.9	78	75.9
Std Avge	82.4	74.3	75.7	80.2	81.4	79.4

Table S2: Specific weight

	SC (78)	CS (82)	KM (28)	LD (59)	SUR (60)	5-loc avge
INDYS Min	127	120	127	134	131	127.8
INDYS Max	148	143	145	153	152	148.2
INDYS 36 Min	134	120	131	140	137	120.0
INDYS 36 Max	148	143	145	152	152	143.0
Std Min	135	122	132	143	138	122.0
Std Max	143	132	139	146	147	147.0
INDYS Avge	142	133	138	148	145	141.1
INDYS 36 Avge	142	132	138	148	145	140.9
Std Avge	139	125	136	144	142	137.3

Table S3: Heading date from 2017-18 trials

	SC (78)	KM (28)	LD (59)	SUR (60)	5-loc avge
INDYS Min	70	80	71	75	74.0
INDYS Max	115	130	125	115	121.3
INDYS 36 Min	75	85	81	80	80.3
INDYS 36 Max	95	100	99	100	98.5
Std Min	75	80	84	75	78.5
Std Max	90	100	94	90	93.5
INDYS Avge	86	92	91	91	89.9
INDYS 36 Avge	85	91	90	90	89.1
Std Avge	82	87	88	83	84.9

Table S4: Plant height measured in 2017-18 trials

## Références bibliographiques

Araus JL and Cairns JE (2014). Field high-throughput phenotyping: the new crop breeding frontier. Trends PI Sci 19: 52-61

Assanga S.O et al. (2017) Mapping of quantitative trait loci for grain yield and its components in a US popular winter wheat TAM 111 using 90K SNPs. PLoS One 12:e0189669

Azadi et al (2015) QTL Mapping of Yield and Yield Components under Normal and Salt-stress Conditions in Bread Wheat (Triticum aestivum L.). PI Mol Biol Rep 6:2030-2040

**Beales J et al.** (2007). A pseudo-response regulator is misexpressed in the photoperiod insensitive Ppd-D1a mutant of wheat (Triticum aestivum L.). Theor Appl Genet 115:721-733.

#### Beddington, 2009. Available online at

https://webarchive.nationalarchives.gov.uk/20121206120858/http://www.bis.gov.uk/assets/goscience/docs/p/perfect-storm-paper.pdf

**Börner A** *et al.* (1997) Comparative molecular mapping of GA insensitive Rht loci on chromosomes 4B and 4D of wheat (Triticum aestivum L.). Theor Appl Genet 95: 1133-1137.

**Brinton J et al.** (2017). Increased pericarp cell length underlies a major quantitative trait locus for grain weight in hexaploid wheat. New Phytol 215: 1026-103

**Cadalen T** *et al.* (1998). Molecular markers linked to genes affecting plant height in wheat using a doubled-haploid population. Theor Appl Genet 96:933-940

Chu C.G., Faris J.D., Friesen T.L., Xu S.S. (2006). Molecular mapping of hybrid necrosis genes Ne1 and Ne2 in hexaploid wheat using microsatellite markers. Theor Appl Genet 112:1374-1381.

**Diaz A et al.** (2012). Copy number variation affecting the Photoperiod-B1 and Vernalisation-A1 genes is associated with altered flowering time in wheat (Triticum aestivum). PLoS One 7, e33234

**Feldman M.** 2001. Origin of cultivated wheat. In: Bonjean AP, Angus WJ (eds). The world wheat book: a history of wheat breeding. (pp. 3-56). Lavoisier Publishing

Fowler DB et al. (2016) Quantitative Trait Loci Associated with Phenological Development, Low-Temperature Tolerance, Grain Quality, and Agronomic Characters in Wheat (Triticum aestivum L.). PLoS One doi:10.1371/journal.pone.0152185

Fulton TM, Chunwongse J and Tanksley SD (1995) Microprep protocol for extraction of DNA from tomato and other herbaceous plants. Plant Mol Biol Rep 13: 207-209.

van Ginkel M.and F Ogbonnaya (2007). Novel genetic diversity from synthetic wheats in breeding cultivars for changing production conditions. Field Crops Research 104: 86-94.

Howell PM, FJ Leigh, AR Bentley and AJ Greenland (2017). Pre-breeding in UK cereal crops - plant science into practice! Aspects of Applied Biology 134: 191-198

Janninck J-L, Lorenz AJ, Iwata H (2010). Genomic selection in plant breeding: from theory to practice Brief Funct Genom 9: 166-177

**Jones H et al.** (2013). Strategy for exploiting exotic germplasm using genetic, morphological, and environmental diversity: the Aegilops tauschii Coss. example. Theor Appl Genet 126:1793-808 Kalinowska K, Chamas S, Unkel K, Demidov D, Lermontova I, Dresselhaus T, Kumlehn J, Dunemann F, Houben A (2018). State-of-the-art and novel developments of in vivo haploid technologies. Theor Appl Genet https://doi.org/10.1007/s00122-018-3261-9

**Klymiuk V et al.** (2018). Cloning of the wheat Yr15 resistance gene sheds light on the plant tandem kinase-pseudokinase family. Nature Comm 9: 3735

Li S et al. (2007). An intervarietal genetic map and QTL analysis for yield traits in wheat. Mol Breeding 20:167-178

Lin, C.-S., and Poushinsky, G. (1985). A modified augmented design (type 2) for rectangular plots. Can. J. Plant Sci. 65:743-149.

Mackay IJ et al. (2014). An Eight-Parent Multiparent Advanced Generation Inter-Cross Population for Winter-Sown Wheat: Creation, Properties, and Validation. G3 4: 1603-1610

McIntosh RA and Lagudah ES (2000). Cytogenetical studies in wheat. XVIII. Gene Yr24 for resistance to stripe rust. Plant Breed. 119: 81-83

Mexico, D.F et al. (2001). Registration of ten synthetic wheat and six bread wheat germplasms resistant to karnal bunt. Crop Science 41: 1652 - 1653

Milne I, Shaw P, Stephen G, Bayer M, Cardle L, Thomas WTB, Flavell AJ and Marshall D. 2010. Flapjack - graphical genotype visualization. Bioinformatics 26:3133-3134.

**Moore G** (2015). Strategic pre-breeding for wheat improvement. Nature Plants 1:15018 doi:10.1038/nplants.2015.18

Mujeeb-Kazi, A., and G.P. Hettel, eds. 1995. Utilizing Wild Grass Biodiversity in Wheat Improvement: 15 Years of Wide Cross Research at CIMMYT. CIMMYT Research Report No. 2.

**Muleta KT** *et al.* (2017). Loci associated with resistance to stripe rust (Puccinia striiformis f. sp. tritici) in a core collection of spring wheat (Triticum aestivum). PLoS One 12: e0179087

**Nedelkou I-P** et al. (2017). Exotic QTL improve grain quality in the tri-parental wheat population SW84. . PLoSOne https://doi.org/10.1371/journal.pone.0179851

United Nations 2017. World Population Prospects. Available online at https://population.un.org/wpp/

**Wang S et al.** (2014) Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. Plant Biotechnol. J. 12:787-796.

Watson A et al. (2018). Speed breeding is a powerful tool to accelerate crop research and breeding. Nature Plants 4:23-29

Yoshioka M et al. (2018). Three dominant awnless genes in common wheat: Fine mapping, interaction and contribution to diversity in awn shape and length. PLoSOne https://doi.org/10.1371/journal.pone.0176148

**Zeven AC** (1981). Eighth supplementary list of wheat varieties classified according to their genotype for hybrid necrosis. Euphytica 30:521-539

## Synthèse des programmes de recherche 2014



# Fonds de soutien à l'Obtention Végétale

Retrouvez l'ensemble des informations sur les programmes de recherche sur le site : www.fsov.org

#### 



Section Semences de Céréales à paille et Protéagineux 44, rue du Louvre - 75001 PARIS Tél. : 0142 33 85 05 - E-mail : fsov@gnis.fr www.fsov.org



