

Identification des déterminismes génétiques permettant une élaboration optimale du rendement sous contraintes (climatique, azote et maladies)

Christophe MICHELET¹, Laure DUCHALAIS¹, Bruno CLAUSTRES¹, Delphine HOURCADE², Katia Beauchêne², Stéphane Jezequel², Xavier LACAZE², Olivier ROBERT³, Ellen GOUEMAND³, Denis BEGHIN³

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

² ARVALIS- Institut de Végétal 3 rue Joseph et Marie Hackin, 75116 PARIS

³ BIOPLANTE, Bioplante Florimond Desprez, BP41, 59242 CAPPELLE EN PEVELE

* Coordinateur : Christophe MICHELET, cmichelet@ragt.fr

1. Introduction

L'amélioration de la productivité restera toujours un objectif essentiel de la création et de la sélection variétale. Cependant depuis plusieurs années la forte volatilité des cours des matières premières agricoles, la hausse des prix des intrants (notamment les engrais), la prise en compte des contraintes environnementales dans les pratiques agricoles et l'évolution des conditions climatiques impactent significativement le revenu des exploitations. C'est pourquoi le progrès génétique pour le rendement doit s'accompagner d'une stratégie de sélection qui intègre tous ces facteurs avec pour principal objectif la sélection de variétés de blé adaptées à des situations de contraintes multiples (température, disponibilité en eau, nutrition azotée, pression parasitaire, etc...).

Depuis plusieurs années, le développement des outils de marquage moléculaire a permis de mieux connaître le déterminisme génétique de certains caractères agronomiques et de facteurs de résistances aux maladies et ainsi d'intégrer progressivement l'utilisation de la sélection assistée par marqueurs dans les programmes de sélection pour le développement de variétés de blé tendre présentant des gènes d'intérêt pour l'agriculteur.

Dans un premier temps, des marqueurs moléculaires ont été développés avec succès sur un nombre important de caractères à héritabilité simple comme la résistance aux maladies (gènes de résistance majeur à la rouille brune *Lr*, à la rouille jaune *Yr*, à *Septoria tritici* *Stb* ou au piétin verse *Pch1*), aux virus (gène *Sbm1* de résistance à la mosaïque du blé) et aux insectes (gène *Sm1* de résistance à la cécidomyie orange du blé), certains éléments qualitatifs majeurs (le locus *Ha* contrôlant la dureté du grain), et des composantes contrôlant

la hauteur de la plante (gènes de nanisme *Rht-B1*, *Rht-D1*) ou le cycle de développement de la plante (gène de vernalisation *Vrn-A1*, *Vrn-B1*, *Vrn-D1*). Ces marqueurs sont désormais utilisés dans de nombreux programmes de sélection de par le monde (Landjeva et al 2007, <http://maswheat.ucdavis.edu>).

Pour des caractères plus complexes comme le rendement, des approches statistiques particulières doivent être appliquées pour détecter des composantes génétiques à effet quantitatif (QTL, Quantitative Trait Loci) tout en tenant compte des interactions QTLs x Environnement ainsi que des interactions entre QTLs. Du fait de la ségrégation de nombreuses composantes génétiques au sein d'une population de lignées recombinantes, l'identification de liaisons étroites entre les marqueurs et le ou les QTLs d'intérêt est parfois difficile. Cela nécessite le recours à d'autres stratégies comme la production de lignées quasi iso-géniques obtenues par rétrocroisement du QTL d'intérêt pour affiner le positionnement des marqueurs vis-à-vis de celui-ci tout en limitant les effets du fonds génétique. Par ailleurs, ce type de structure génétique (population de lignées recombinantes) permet uniquement de générer des informations sur la valeur des allèles issues des 2 lignées parentales qu'il faut ensuite valider au sein d'un panel de matériel génétique plus diversifié. D'autre part, la nécessité de développer spécifiquement une population recombinante issue de technologies HD (haplo-diploïdisation) ou SSD (Single Seed Descent) en ségrégation pour le caractère d'intérêt constitue également une contrainte forte de coût et de temps lorsque ce type d'études est initié. Ces difficultés ont pu limiter le développement de marqueurs associés à ces QTLs en comparaison des marqueurs associés à des gènes majeurs.

Pourtant, des QTLs liés au rendement et aux composantes de rendement ont été rapportés dans

de nombreuses études (Quarrie *et al.* 2005, Sun *et al.* 2009; Li *et al.* 2007, Kuchel *et al.* 2007a) et certaines ont été réalisées pour étudier des interactions entre l'élaboration du rendement et l'environnement, en particulier sous des conditions de stress abiotiques (Quarrie *et al.* 2006, Kuchel *et al.* 2007b, Snape *et al.* 2007, Kirigwi *et al.* 2007).

Une autre approche pour la détection de QTLs se développe actuellement : il s'agit des études par génétique d'association qui permettent d'associer des données génotypiques provenant de matériel génétique diversifié issu des programmes de sélection à des données phénotypiques normalement mesurées durant le processus de sélection comme le rendement. L'un des avantages de cette méthode est l'accès à une variabilité allélique plus forte au travers d'un groupe de matériel d'origine génétique diversifiée. Grâce à des méthodes statistiques appropriées, il est alors possible de déterminer l'effet de différents allèles associés à un QTL donné.

L'application de la génétique d'association demande un nombre important de marqueurs pour obtenir une couverture génétique dense et de ce fait nécessite l'accès à des techniques de génotypage à haut débit.

La nature allohexaploïde du génome du blé a limité le développement de marqueurs SNPs à partir de la voie conventionnelle des ESTs (Expressed Sequence Tag) et a compliqué l'analyse génétique du fait des produits d'amplification homéologues (Barker and Edwards, 2009). Malgré ces inconvénients, des progrès ont été réalisés sur l'optimisation de la voie ESTs pour l'obtention des marqueurs SNPs

(<http://wheat.pw.usda.gov/SNP/new/index.shtml>) et de la validation de la plateforme Illumina Golden Gate Assay sur le blé (Akhunov *et al.* 2009). Le développement de marqueurs COS (Conserved Orthologous Set) (Quraishi, 2009) et ISBP (Insertion Site Based Polymorphism) (Paux *et al.* 2008) ainsi que l'exploitation d'une prochaine génération de marqueurs SNPs spécifiques des génomes A, B et D du blé (http://www.cerealsdb.uk.net/search_reads.htm)

contribueront considérablement au développement de cette approche QTL par génétique d'association. L'objectif de ce projet était d'appliquer ces approches nouvelles de génétiques d'association pour l'identification de facteurs génétiques liés au rendement chez le blé tendre et en particulier sous contraintes hydrique, azote et maladies.

2. Matériel et méthodes

2.1 Matériel végétal

Un panel de 206 variétés et lignées a été défini par les 3 partenaires du projet.

Ce matériel végétal se compose du « panel de lignées élites françaises » travaillé par ARVALIS (148 variétés) ainsi que d'un échantillon complémentaire de 58 variétés et lignées en cours de sélection apporté par R2n et BIOPLANTE.

2.2 Phénotypage du matériel végétal

Evaluation phénotypique au champ du panel commun en condition de stress hydrique

Cette évaluation a été conduite sur 2 années (2011 et 2012) sur le site ARVALIS de Gréoux (04) identifié comme particulièrement intéressant pour mener des études sur le comportement variétal face au stress hydrique et thermique. Le sol est de type sablo-argilo limoneux. Ce panel de variétés a été testé sous 2 conditions : non irriguée (SEC) et irriguée (IRR).

En 2011, uniquement 155 variétés ont pu être testées sous la condition irriguée. Cinq irrigations ont été réalisées entre les stades apparition de la dernière feuille (Z37) et grain laiteux (Z75) pour un total de 175 mm d'eau pour la modalité Irriguée. Aucun apport d'eau n'a été effectué pour la modalité SEC.

En 2012, uniquement 131 variétés ont pu être évaluées. Quatre irrigations ont été réalisées entre les stades épi 1 cm (Z30) et grain laiteux-pâteux (Z80) pour un total de 130 mm d'eau pour la modalité IRR. Un apport de 30 mm a été réalisé pour la modalité SEC afin de favoriser un bon établissement du nombre d'épis/m². Le stress hydrique est donc plutôt appliqué pendant la phase du remplissage du grain.

Le programme d'irrigation 2012 est assez différent de celui de 2011 par la quantité (- 40mm en 2012), par la fréquence (4 en 2012 au lieu de 5 en 2011) et par les dates d'apport (80% des apports entre fin montaison à fin épiaison en 2011 contre 80% des apports réalisés entre début épiaison à la fin du remplissage du grain en 2012).

Evaluation phénotypique au champ du panel commun en condition de stress azoté et de faible pression maladies

Le panel de 206 variétés a été évalué sur 2 sites en 2011 (un site R2n à Louville la Chenard = LLC (28-

Eure et Loir) et un site Bioplante à Prémèsques = PRE (59-Nord)) et sur 2 sites également en 2012 (un site R2n à Louville la Chenard = LLC (28-Eure et Loir) et un site Bioplante à Cappelle = CAP (59-Nord)).

Sur chaque site, 2 modalités différentes ont été appliquées :

- **Une Modalité N+F+** correspondant à une conduite à l'optimum technique pour la fertilisation azotée et pour la protection fongicide.
- **Une Modalité N-F-** correspondant à une conduite avec limitation des intrants (azote et fongicide).

Deux répétitions ont été réalisées par essai et par modalité.

Le tableau ci-dessous présente les 2 itinéraires techniques culturaux appliqués sur les 4 sites d'expérimentation en 2011 et 2012 :

Site	Année	Modalité	Précédent	Date de semis	Azote (dose totale)	Fertilisation azotée et protection fongicide
Louville	2011	N+F+	Colza	22/10/2010	200	2 fongicides foliaires (dose 100%) + 2 fongicides anti-fusariose de l'épi (dose 100%)
Louville	2011	N-F-	Colza	22/10/2010	120	2 fongicides foliaires (dose 50%) + 2 fongicides anti fusariose de l'épi (dose 100%)
Prémèsques	2011	N+F+	Pomme de terre	26/10/2010	170	2 fongicides foliaires (dose 100%) + 2 fongicides anti fusariose de l'épi (dose 100%)
Prémèsques	2011	N-F-	Pomme de terre	26/10/2010	120	Pas d'application de fongicides foliaires et anti-fusariose = essais non traité
Louville	2012	N+F+	Colza	18/10/2011	150	2 fongicides foliaires (dose 100%) + 2 fongicides anti fusariose de l'épi (dose 100%)
Louville	2012	N-F-	Colza	18/10/2011	50	2 fongicides foliaires (dose 50%) + 2 fongicides anti fusariose de l'épi (dose 100%)
Cappelle	2012	N+F+	Maïs ensilage	28/10/2011	230	2 fongicides foliaires (dose 100%) + 1 fongicide anti fusariose de l'épi (dose 100%) + régulateur
Cappelle	2012	N-F-	Maïs ensilage	28/10/2011	130	2 fongicides foliaires (dose 50%) + 1 fongicide anti fusariose de l'épi (dose 100%)

La modalité N- avait pour objectif de créer une situation de carence azotée pendant la phase de montaison du blé. Le tableau ci-dessous résume les applications d'azote (kg/ha) sur les 4 sites en 2011 et 2012 :

Site	Année	Modalité	Reliquat sortie hiver	Apport N1	Apport N2	Apport N3	Dose totale
Louville	2011	N+F+	25	50	90	60	200
Louville	2011	N-F-	25	0	60	60	120
Prémèsques	2011	N+F+	62	40	80	50	170
Prémèsques	2011	N-F-	62	0	70	50	120
Louville	2012	N+F+	80	50	60	40	150
Louville	2012	N-F-	80	0	0	50	50
Cappelle	2012	N+F+	25	50	130	50	230
Cappelle	2012	N-F-	25	0	80	50	130

Notations et mesures réalisées sur les essais

Les notations et mesures réalisées sur les essais figurent dans le tableau ci-dessous :

CODE VARIABLE	NOTATION	2011						2012					
		LLC N+F+	LLC N-F-	PRE N+F+	PRE N-F-	GRE IRR	GRE SEC	LLC N+F+	LLC N-F-	CAP N+F+	CAP N-F-	GRE IRR	GRE SEC
D.250	Dose de semis	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
D.255	Dose épaveur 255	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
D.255 Q	Dose épaveur 255 QUANTITE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
D.292	Dose de récolte	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
BFLD	Biomasse foliaire (TMS)	N	N	N	N	Y	Y	N	N	N	N	N	N
BFLD	Biomasse foliaire (TMS)	N	N	N	N	Y	Y	N	N	N	N	N	N
SA	Nb épis/m²	Y	Y	Y	Y	(0)	(0)	Y	Y	Y	Y	Y	Y
TSPS	Nb épis/m² (seuils par esp)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
SSPS	Nb épis/m² (seuils par esp)	N	N	N	N	Y	Y	N	N	N	N	N	N
PSPS	Nb épis/m² (seuils par esp)	N	N	N	N	Y	Y	N	N	N	N	N	N
SNFS	Nb épis/m² (seuils par esp)	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Y
GNSP_C1	Nb grains / épi (Colza 1)	N	N	N	N	(0)	(0)	Y	Y	Y	Y	Y	Y
GNSP_C2	Nb grains / épi (Colza 2)	Y	Y	Y	Y	N	N	N	N	N	N	N	N
GPA_C1	Nb grains/m²	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
GPA_C2	Nb grains/m²	Y	Y	Y	Y	N	N	N	N	N	N	N	N
GPS_C1	Nb grains/m²	Y	Y	Y	Y	(0)	(0)	Y	Y	Y	Y	Y	Y
GPS_C2	Nb grains/m²	Y	Y	Y	Y	N	N	N	N	N	N	N	N
TKW_Plot	PM5 (N, S, P, K)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
TKW_Ear	PM5 (N, S, P, K)	Y	Y	Y	Y	N	N	N	N	N	N	N	N
BT	BT (N, S, P, K)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
PH	PH (N, S, P, K)	N	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
SW	SW (N, S, P, K)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
HA	HA (N, S, P, K)	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	N
GNC	GNC (N, S, P, K)	Y	Y	Y	Y	1 rep	1 rep	Y	Y	Y	Y	Y	Y
SNY	SNY (N, S, P, K)	Y	Y	Y	Y	1 rep	1 rep	Y	Y	Y	Y	Y	Y
GPC	GPC (N, S, P, K)	Y	Y	Y	Y	1 rep	1 rep	Y	Y	Y	Y	Y	Y

2.3 Génotypage du matériel végétal

Différentes plateformes de génotypage ont permis de génotyper l'ensemble du panel étudié.

Nous avons ainsi pu le génotyper avec :

- 1572 marqueurs DArT (Triticarte Pty Ltd)
- 439 marqueurs SSR
- 1 734 SNPs issus du projet Digital
- 35 867 SNP issus de la puce 90K (Illumina)
- 104 727 SNP provenant de la puce TaBW420 développée dans la cadre du projet Breedwheat.

Pour les marqueurs SSR et SNP, le génotypage a été effectué par la plateforme Gentyane basée à l'INRA GDEC.

2.4 Analyse des données

Ajustement des données de phénotypage

Un modèle mixte linéaire a été utilisé pour estimer les effets terrains et pouvoir calculer des moyennes ajustées. Le sous-bloc de précocité et le génotype ont été définis comme fixes et les répétitions ont été définies comme aléatoires dans ce modèle. Les lieux et les modalités ont été traités de façon indépendante.

Analyses d'association

Les marqueurs ayant une Minimum Allele Frequency (MAF) inférieure à 5% ainsi que les marqueurs présentant plus de 20% de données manquantes ou hétérozygotes ont été supprimés des analyses de génétique d'association afin d'éviter les liaisons marqueur/caractère non significatives ou erronées.

Un modèle mixte linéaire (MLM) introduisant une matrice structure et une matrice d'apparentement a été utilisé grâce au package EMMA pour calculer les associations marqueurs – caractères. La date d'épiaison moyenne a également été introduite en

co-variable pour les caractères présentant une forte interaction avec la précocité.

Analyses des interactions génotype x environnement

Données phénotypiques

Une base de données a été constituée par la mise en commun des différents partenaires du projet de données phénotypiques historiques correspondant à 789 environnements au total et une trentaine de variétés inscrites au catalogue français était présente par environnement. Au total 13 197 données individuelles (lieu x année x variété) ont été mises à la disposition d'Arvalis pour l'analyse des interactions QTL x Environnement. Les données ont été acquises entre 2005 et 2010 et concernent uniquement le rendement. La caractérisation environnementale a consisté à calculer pour chaque combinaison lieu x année x variété des indicateurs de stress dérivés des données météorologiques. Les données de base météorologiques utilisées sont issues des stations météorologiques les plus proches possibles des lieux d'essais et correspondent au réseau interne Arvalis de références météorologiques. Les moyennes de ces données brutes journalières (ETP caractérisant l'évapotranspiration, Températures maximales, Températures minimales, Températures moyennes, Rayonnement Global, Précipitations) ont permis de dériver des indicateurs de stress pour des intervalles phénologiques recalculés à partir de l'interface Panoramix (outil modélisateur interne à Arvalis qui permet de recalculer les stades phénologiques d'un cycle de blé tendre à partir des données météorologiques et de sols pour toutes les combinaisons lieu x année x variété considérées). Une fois les stades phénologiques déterminés, nous avons pu pour chacun calculer des indices de stress tels que la somme des températures maximales supérieures à 25°C entre épiaison et grain laiteux. Au total 120 covariables de nature différente ont été calculées pour chaque combinaison lieu x année x variété. Ces informations appelées « covariables environnementales » sont analysées en parallèle avec les données génotypiques et phénotypiques.

Données génotypiques

Les données de génotypage de la puce 90K sont utilisées. Seulement 23 462 marqueurs ont été retenus après filtres sur les données manquantes (taux de données manquantes supérieur à 10%) et sur la MAF (inférieure à 5% des données). Après l'application de ces filtres, les données

génotypiques, phénotypiques et de covariables environnementales pouvaient être analysées en régression factorielle.

Le modèle d'analyse est un modèle linéaire de la forme suivante :

$$Y_{ij} = \mu + cov_{kj} + mk_{mi} + cov_{kj}.mk_{mi} + K + \varepsilon_{ij}$$

Dans cette équation, i correspond au génotype, j à l'environnement considéré, μ à l'intercept ou moyenne générale, cov_{kj} à l'effet principal de la covariable environnementale k qui explique une partie des effets principaux environnement, mk_{mi} correspond au marqueur m caractéristique du génotype i , $cov_{kj}.mk_{mi}$ correspond au terme d'interaction entre la covariable k et le marqueur m .

Ce modèle permet de tester s'il existe des pentes de régression différentes selon les allèles au marqueur considéré vis-à-vis de la covariable environnementale. Si les pentes sont significativement différentes alors la p -value associée au terme décrit ci-dessus devient significative. Enfin, K correspond à la matrice de kinship dérivée de l'apparentement calculé sur la base des marqueurs moléculaires et ε_{ij} à la résiduelle du modèle. En effet, il n'y avait pas de répétition d'un génotype dans un environnement donné. On peut donc tester l'interaction $cov_{kj}.mk_{mi}$ mais les effets d'interaction génotype x environnement sont confondus avec la résiduelle ce qui pose souci pour la représentation des effets des allèles aux marqueurs.

Les analyses ont donc été menées sous ASREML (Gilmour et al., 2009). Le seul effet aléatoire déclaré étant l'effet de la Kinship, tous les autres effets ont été considérés comme fixes.

3. Résultats et discussions

3.1 Evaluation phénotypique du matériel végétal

Résultats des expérimentations IRR / SEC 2011 et 2012 à Gréoux

La figure ci-dessous montre les résultats de rendement obtenu en 2011 et 2012 sur les 2 modalités IRR et SEC. Les rendements moyens obtenus sont très similaires entre les 2 années avec des différentiels d'environ 14 qx/ha malgré des modalités d'irrigations assez différentes entre 2011 et 2012.



Les 2 tableaux ci-dessous montre les facteurs explicatifs de la variation de rendement entre les 2 modalités.

	GREOUX 2011				GREOUX 2012				
	Sec	Irrigué	Diff		Sec	Irrigué	Diff		
Heading time	132	132	0	ns	Heading time	67	69	0	ns
Senescence date	172	172	0	ns	Plant height	479	523	-44	**
Plant height	71	78	-7	***	Nb Ear/m ²	18.0	18.0	-0.1	ns
Biomasse Flo	21.8	22.3	-0.5	*	Nb Spikelet/Ear	2.4	2.4	0.1	ns
N Flowering time	1.6	1.5	0.1	**	Nb aborted spikelet/Ear	15.5	15.7	-0.2	ns
Nb Ear/m ²	495	497	-2	ns	Nb Grain/spikelet	2.7	2.7	0.0	ns
Nb Spikelet/Ear	21.4	20.8	0.6	**	Nb Grain/Ear	42	41	0.4	ns
Nb aborted spikelet/Ear	3.4	3.3	0.1	ns	Nb Grain/m ²	19735	21350	-1615	**
Nb fertile spikelet/Ear	18.0	17.4	0.5	**	TKW	33.8	38.2	-4.4	***
Nb Grain/spikelet	2.3	2.5	-0.3	***	Grain dry matter	6.2	7.8	-1.6	***
Nb Grain/Ear	40	44	-4	***	Straw dry matter	6.8	8.1	-1.3	***
Nb Grain/m ²	19651	21930	-2279	***	Total dry matter	13.0	15.9	-2.9	***
TKW	36.0	39.2	-3.2	***	Harvest Index	0.47	0.49	-0.02	ns
Yield qx/ha	70.7	85.3	-14.6	***	Yield qx/ha	66.1	80.9	-14.8	***
Sp.Weight	73.9	74.3	-0.4	ns	Sp.Weight	76.8	81.2	-4.4	***
Protein content	12.9	10.4	2.5	***	Protein content	14.5	12.9	1.6	***

En 2011, l'écart de rendement s'expliquerait par une diminution du PMG et du nombre de grains/m² suite à une réduction du nombre de grains/épi sans que le nombre d'épis/m² soit pénalisé.

En 2012, une diminution du PMG et du nombre de grains/m² serait aussi responsable de l'écart de rendement Irrigué/Sec. Mais par rapport à 2011, la composante « nombre d'épis/m² » expliquerait en grande partie la réduction du nombre de grains/m² en conservant un nombre de grains/épi stable entre les 2 modalités. On observe également une diminution des teneurs en matières sèches des pailles et du grain entre les 2 modalités en 2012.

Le stress hydrique appliqué en 2011 et 2012 n'affecte donc pas de la même manière l'élaboration de la composante de rendement « nombre de grains/m² ». Le PMG est cependant la composante qui a été systématiquement pénalisée par le stress hydrique.

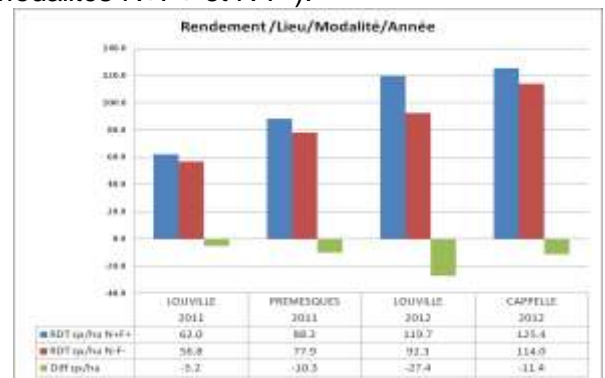
Les corrélations pour le rendement entre les 2 modalités IRR et SEC et les 2 années d'expérimentations sont très faibles :

	SEC_11GRE	IRR_11GRE	SEC_12GRE	IRR_12GRE
SEC_11GRE	1			
IRR_11GRE	0.29197323	1		
SEC_12GRE	0.00886654	0.02357345	1	
IRR_12GRE	0.26286327	0.36964312	0.41187742	1

Ces résultats confirment donc que les stress hydriques appliqués en 2011 et en 2012 n'ont pas affecté le rendement de la même façon.

Résultats des expérimentations 2011 et 2012 en conditions N+F+ et N-F-

La figure ci-dessous présente les résultats rendement obtenus sur les 4 sites en 2011 et 2012 (modalités N+F+ et N-F-).



Essai Louville 2011 : Le rendement moyen N+F+ est égal à 62 qx/ha vs 56.8 qx/ha pour la modalité N-F-. L'écart de rendement entre les 2 modalités est de 5,2 qx/ha pour une réduction de 80u d'azote. Par rapport à une année normale, le potentiel de rendement a été fortement affecté par les conditions sèches du printemps qui ont pénalisé l'absorption de l'azote pendant la phase de montaison expliquant le faible différentiel de rendement N+F-/N-F-.

Essai Prêmesques 2011 : Le rendement moyen N+F+ est égal à 88.2 qx/ha vs 77.9 qx/ha pour la modalité N-F-. L'écart de rendement est de 10.4 qx/ha pour une réduction de 50u d'azote. Le potentiel de rendement N+F+ sur le site de Prêmesques est nettement supérieur à celui de Louville (+26 qx/ha) ce qui dénoterait des conditions d'absorption de l'azote beaucoup plus favorables. L'écart de rendement N+F+/N-F- est 2 fois plus important à Prêmesques (-10.4 qx/ha) qu'à Louville (-5.2 qx/ha). Sur ces 2 sites, la pression des maladies foliaires (septoriose, rouille jaune et rouille brune) était quasi-nulle dans les essais N-F- à cause des conditions particulièrement défavorables du printemps (absence significative de pluies).

Essai Louville 2012 : Le rendement moyen N+F+ est égal à 119.7 qx/ha vs 92,3 qx/ha pour la modalité N-F- soit un écart de rendement de 27.4 qx/ha pour une réduction de 100u d'azote. Ce potentiel de rendement élevé est à relier aux conditions très favorables pour la mise en place des composantes Nb grains/m² et PMG. A noter que malgré une pression maladie très forte, un programme allégé en N-F- a permis de protéger l'essai de manière très satisfaisante (pas d'observation de maladies) et que par conséquent l'écart de rendement observé entre les 2 modalités peut être principalement attribué à la réduction de la fertilisation azotée.

Essai Cappelle 2012 : Le rendement moyen N+F+ est égal à 125,4 qx/ha vs 114 qx/ha en N-F-, soit une différence de 11,4 qx/ha pour une réduction de 100u d'azote. Comme à Louville, des conditions très favorables ont permis l'expression d'un potentiel élevé. Il a été observé la présence de septoriose, de rouille jaune et de rouille brune en N-F-. Par conséquent, l'écart de rendement entre les 2 modalités pourra s'expliquer plus par une interaction entre carence azotée x maladies foliaires.

Les 2 sites de Louville et Cappelle se différencient surtout au niveau du rendement obtenu sous la modalité N-F- avec un plus fort différentiel à Louville (-27 qx/ha vs -11 qx/ha) qui peut s'expliquer par l'absence du 2^{ème} apport d'azote (décision d'impasse suite au reliquat élevé de 80u mais qui n'a certainement pas pu être valorisé à 100%).

Le tableau suivant montre les facteurs explicatifs des écarts de rendement obtenus entre les modalités N+F+ et N-F- en 2011 et 2012 :

	LOUVILLE 2011				PREMESQUES 2011				LOUVILLE 2012				CAPPELE 2012			
	N+F+	N-F-	Dif	Prob	N+F+	N-F-	Dif	Prob	N+F+	N-F-	Dif	Prob	N+F+	N-F-	Dif	Prob
Heading time	131	132	-1	*	133	133	0	ns	142	141	1	ns	149	149	0	ns
Plant height	69	66	2	**	85	85	0	ns	95	90	5	**	86	82	4	**
Yield q/ha	82.0	56.8	5.2	***	88.2	77.9	10.3	***	119.7	92.3	27.4	***	125.3	114.0	11.3	***
Nb Ear/m ²	328	317	11	**	354	342	12	**	602	491	111	***	628	486	142	***
Nb Spikelet/Ear	16.1	15.6	0.5	*	19.2	18.6	0.6	**	16.9	17.6	-0.7	**	17.1	17.0	0.1	ns
Nb Grain/spikelet	2.3	2.4	0.0	*	2.3	2.3	0.0	ns	2.6	2.5	0.2	*	3.0	3.4	-0.4	**
Nb Grain/Ear	45	38	7	***	47	43	4	***	44.4	43.5	0.9	ns	52	57	-6	**
Nb Grain/m ²	14424	11725	2699	***	16332	14462	1870	***	26167	20503	5664	***	32088	27366	4722	***
TKW	43.2	48.6	-5.4	***	54.3	54.2	0.1	ns	46.2	45.5	0.7	*	38.4	41.9	-2.5	**
Sp.Weight	77.6	77.6	0.1	ns	77.4	76.1	1.3	***	75.5	75.5	0.0	ns	77.6	77.4	0.2	ns
Protein content	11.7	11.4	0.3	***	12.7	11.4	1.3	***	12.3	11.2	1.1	***	12.3	11.2	1.1	***

En 2011, l'écart de rendement sur les 2 sites s'explique par une diminution du nombre de grains/m² suite à une réduction combinée du nombre d'épis/m² x nombre de grains/épi. Cependant, le PMG demeure constant entre les 2 modalités à l'inverse de ce qui est observé sur le site de Louville (augmentation du

PMG en N-F- suite à la réduction du nombre de grains/m²).

En 2012, le nombre de grains/m² est la composante qui explique principalement la différence de rendement entre les 2 modalités au travers d'une forte diminution du nombre d'épis/m².

Sur 2 ans, le nombre de grains/m² semble la composante la plus affectée par le stress azoté qui selon les années s'expliquerait plus par une réduction du nombre de grains/épi (2011) ou par une diminution du nombre d'épis/m² (2012). Le PMG interviendrait très peu dans l'explication de l'écart de rendement entre les 2 modalités.

Nous observons également une teneur en protéines beaucoup plus faible sous la modalité N-F- et un effet peu marqué de la diminution de la fertilisation azotée sur le poids spécifique.

Les tableaux suivants montrent les analyses de corrélation entre les 2 années pour chaque modalité. Les corrélations de Pearson sont effectuées sur les moyennes des 2 sites de chaque année. Les corrélations sont très bonnes entre les 2 années pour les 2 modalités concernant l'épiaison et la hauteur qui sont des caractères très héréditaires, mais également pour le nombre d'épillet/épi et le nombre de grains/m². Des corrélations plus faibles mais supérieures à 0.50 sont observées pour le rendement, le nombre d'épis/m² et le nombre de grains/épis ce qui indique une plus forte interaction avec l'environnement.

Annual correlation	MEAN 2011/2012	
	2011 = LLC+PRE	
	2012 = LLC+CAP	
	N+F+	N-F-
Grain Yield	0.52	0.64
Nb Ear/m ²	0.55	0.42
Nb Spikelet/Ear	0.83	0.84
Nb Grain/spikelet		
Nb Grain/Ear	0.58	0.49
Nb Grain/m ²	0.73	0.72
TKW	0.70	0.78
Heading time	0.97	0.97
Plant height	0.83	0.86
Sp.Weight	0.44	0.51
Protein content		

Les corrélations pour le rendement entre les 2 modalités N-F- et N+F+ sont très élevées pour une même année et sont relativement correctes entre les 2 années de phénotypage.

	N-F-_11MOY	N+F+_11MOY	N-F-_12MOY	N+F+_12MOY
N-F-_11MOY	1			
N+F+_11MOY	0.813232116	1		
N-F-_12MOY	0.644920392	0.630389246	1	
N+F+_12MOY	0.507890767	0.516805663	0.757251801	1

3.2 Description génétique du panel

Les différents batches de génotypage ont permis de couvrir de façon homogène le génome. Le nombre de marqueurs par chromosome varie entre 5 534 (4D) et 16 186 (2B). La moyenne est de 10 860 marqueurs et le génome D reste cependant le moins bien représenté.

Une analyse de la diversité génétique à l'aide d'une ACP sur la matrice de Kinship générée avec les marqueurs SNP de la TaBW420 permet de distinguer 4 groupes. La même conclusion aurait été faite sur la base des 439 SSRs, à la différence que les groupes sont plus resserrés.

Le Déséquilibre de liaison (DL) a été calculé avec le package LDHeatMap sous R pour chaque paire de marqueur et par chromosome. L'équation de décroissance du DL de Hill & Weir a été « fittée » sur la courbe et la distance à partir de laquelle le DL décroît en dessous de 0,2 a été déterminée pour chaque chromosome.

chromosome	1A	1B	1D	2A	2B	2D	3A	3B	3D	4A	4B	4D	5A	5B	5D	6A	6B	6D	7A	7B	7D
LD decay < 0.2	0.8	0.2	2.7	0.5	0.8	2.0	0.9	0.2	1.2	1.6	1.1	3.9	0.5	2.6	1.5	0.7	0.4	2.5	0.4	0.7	3.5

Les distances varient entre 0.2 et 3.9 cM indiquant une structure du panel faiblement marquée. Cette information nous permet également de diminuer les intervalles de confiance des pics d'association mis en évidence.

3.3 Analyses QTL

Une multitude de QTLs a été détectée dans l'ensemble des environnements testés. Une large partie d'entre eux est environnement spécifique, c'est-à-dire que leur effet sur le rendement n'est significatif que dans un seul lieu et pour une seule année d'expérimentation. Leur effet est donc trop dépendant de l'environnement pour être utilisé en Sélection Assistée par Marqueur (SAM). Nous nous sommes donc focalisés sur l'identification de QTLs qui se sont exprimés pour les 2 années de phénotypage et les 2 modalités testées pour chacun des stress étudié.

Résultats des expérimentations IRR/SEC 2011 et 2012 à Gréoux

Aucun QTL de rendement n'a pu être identifié pour les 4 environnements testés. Ces résultats peuvent s'expliquer par les très faibles coefficients de corrélation obtenus entre les 4 environnements testés.

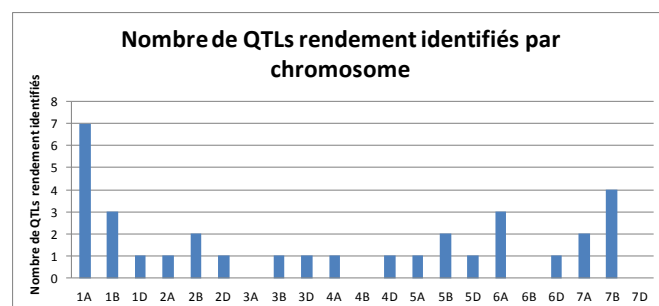
Un seul QTL localisé sur le 5B a pu être identifié pour la modalité SEC pour les 2 années de phénotypage. Il a apporté un gain sur le rendement de +8.3% en 2011 et un gain de +2.4% en 2012. Il

n'est lié à aucun autre caractère et présente donc un intérêt pour une utilisation en SAM à condition d'être en présence d'un stress hydrique puisqu'il ne s'exprime pas pour la modalité irriguée.

Résultats des expérimentations 2011 et 2012 en conditions N+F+ et N-F-

Trente trois QTLs ont été identifiés comme significativement liés au rendement dans l'ensemble des environnements testés pour les 2 conditions N+F+ et N-F-. Ce grand nombre de QTLs communs aux 2 modalités et aux 2 années testées se justifie par les coefficients de corrélations élevés que nous avons eus entre environnements. En moyenne sur les 4 environnements testés, ils permettent d'obtenir un gain sur le rendement variant de +2.7% à +10% entre l'haplotype négatif et l'haplotype positif en fonction du QTL.

Les 33 QTLs rendement identifiés se distribuent sur l'ensemble des 3 génomes comme le montre la figure ci-dessous :



Sept QTLs ont donc pu être identifiés sur le chromosome 1A, 1 sur le 2A, 1 sur le 4A, 1 sur le 5A, 3 sur le 6A, 2 sur le 7A, 3 sur le 1B, 2 sur le 2B, 1 sur le 3B, 2 sur le 5B, 4 sur le 7B, 1 sur le 1D, 1 sur le 2D, 1 sur le 3D, 1 sur le 4D, 1 sur le 5D et 1 sur le 6D. Aucune relation entre le nombre de QTLs identifiés et la saturation en marqueurs des différents chromosomes n'a pu être identifiée.

Les différents haplotypes de ces 33 QTLs sont suffisamment représentés dans l'ensemble du panel étudié à part pour 4 d'entre eux ayant une MAF inférieure à 10%. Pour ces derniers, leurs effets restent donc à valider dans un panel de lignées ou de variétés beaucoup plus large.

Plusieurs associations entre ces QTLs et d'autres caractères ont pu être mises en évidence.

Parmi ces QTLs, 13 sont également liés à la précocité. Bien que nous ayons introduit l'épiaison en co-variable dans nos analyses quand nous avons une forte corrélation entre épiaison et

rendement, plus d'1/3 des QTLs identifiés présente quand même un impact sur la précocité. L'effet de l'haplotype positif pour le rendement engendre une tardification du matériel variant de 1,5 à 2,5 jours. Ceci peut s'expliquer par le fait que les lieux d'essais se trouvaient au Nord de la Loire et que de ce fait, le matériel précoce était moins adapté que le matériel tardif.

Parmi ces 33 QTLs, 12 sont également liés au poids spécifique. Onze sont liés de façon négative et induisent une diminution du PS allant de -0.34 à -0.74 point de PS et 1 seul QTL est lié de façon positive induisant un gain de PS de +0.49 point pour une augmentation moyenne de 4.5% sur le rendement. Ce QTL présente donc un avantage considérable par rapport aux autres pour une utilisation en SAM.

Parmi ces 33 QTLs, 9 sont également liés au PMG. Ces 9 QTLs sont liés à une augmentation du PMG variant de +1,05g à +2,03g. Plus d'un quart des QTLs rendement sont donc liés à une augmentation du PMG. Le PMG semble donc être une composante importante à travailler si nous voulons améliorer le potentiel rendement de nos variétés.

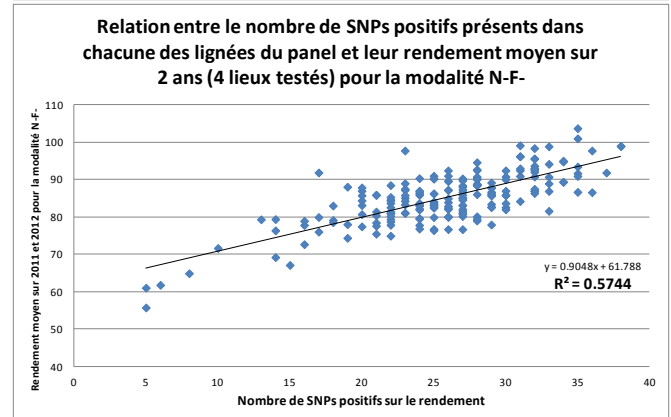
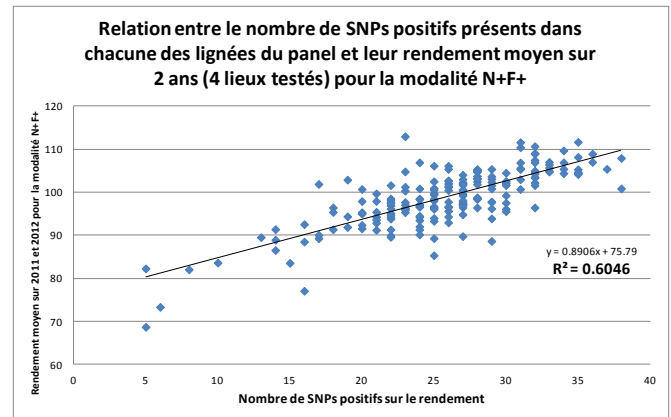
Parmi ces 33 QTLs, 22 sont également liés à la teneur en protéines. Tous ces QTLs sont liés négativement à la teneur en protéines, induisant une baisse variant de -0.13 à -0.74. Deux tiers des QTLs rendement identifiés sont donc liés à une diminution de la teneur en protéines. Une fois encore, il s'avère donc difficile de casser la relation négative entre élaboration du rendement et élaboration de la teneur en protéines. Les QTLs rendement n'ayant pas d'effet sur la teneur en protéines sont donc ceux à privilégier mais ce sont malheureusement ceux qui ont les effets les moins importants sur le rendement.

Parmi ces 33 QTLs, 2 sont également liés à la hauteur et engendrent une légère augmentation de la taille des plantes allant de + 4.0 à + 6.5 cm.

Il est donc extrêmement difficile de trouver un QTL rendement qui n'ait pas d'effet négatif sur un autre caractère. Les QTLs à effet les plus importants sur le rendement sont souvent liés à d'autres caractères de façon négative. Trouver les QTLs à utiliser en SAM s'avère donc comme étant assez complexe.

Finalement, nous avons donc regardé la distribution du rendement en fonction du nombre de SNPs liés de façon positive au rendement présents

dans chacune des lignées pour la modalité N+F+ et pour la modalité N-F-.



Grâce aux deux graphiques ci-dessus, nous pouvons donc constater que le rendement augmente de façon linéaire avec le nombre de SNPs positifs présents.

Pour la modalité N+F+, nous obtenons un coefficient de détermination de 0,60 pour la droite de régression linéaire et nous avons une corrélation de 0.78 entre le rendement moyen sur les 4 environnements testés et le nombre de SNPs positifs dans les différentes lignées.

Pour la modalité N-F-, nous obtenons un coefficient de détermination de 0,57 pour la droite de régression linéaire et nous avons une corrélation de 0.76 entre le rendement moyen sur les 4 environnements testés et le nombre de SNPs positifs dans les différentes lignées.

Nous captions donc avec ces 33 QTLs identifiés une partie très importante de la variabilité génétique du rendement quelle que soit la modalité testée sous une contrainte azotée et nous pourrions davantage nous rapprocher de la valeur moyenne du rendement en construisant un modèle prédictif prenant en compte les effets de chacun de ces 33 QTLs identifiés.

3.4 Analyses des interactions génotype x Environnement

Trois QTL principaux ont été identifiés, un sur le chromosome 4A et deux sur le chromosome 4B, tous les trois en interactions avec les températures échaudantes au cours du remplissage du grain ainsi qu'en interaction avec le stress froid de début de cycle. Les p-values associées au test d'interaction entre les marqueurs et les covariables sont inférieures à 10^{-7} . Pour chaque QTL, un des deux allèles a un effet positif sur le rendement et présente une pente soit croissante soit décroissante vis-à-vis des températures échaudantes. L'autre allèle étant souvent considéré comme ayant un effet nul. Ces QTLs sont donc des gènes en interaction avec l'environnement et peuvent être des gènes centraux dans les voies métaboliques déterminant les caractères principaux comme la précocité et le rendement. L'analyse des interactions QTL x Environnement a donc révélée des QTLs qui n'étaient pas identifiés par ailleurs dans le projet. L'analyse des interactions QTL x Environnement apporte donc une information complémentaire par rapport aux approches traditionnelles de génétique d'association.

identifiés grâce au phénotypage réalisé dans le cadre du projet. Parmi ces QTLs, un a été identifié sur le chromosome 5B et ne s'exprime qu'en condition de stress hydrique alors que trente trois QTLs ont pu être identifiés dans la totalité des environnements testés pour les modalités N+F+ et N-F-. La majorité de ces QTLs sont également liés à d'autres caractères (précocité, PS, PMG, teneur en protéines, hauteur...) ce qui complexifie leur utilisation en SAM qui doit donc se faire avec précaution.

Trois autres QTLs rendement ont pu être identifiés grâce à une base de données phénotypiques constituée par la mise en commun des différents partenaires du projet de données historiques. Grâce à une analyse des interactions Génotype x Environnement, ces QTLs se sont montrés comme étant en étroite interaction avec les températures échaudantes au cours du remplissage du grain et des températures froides en début de cycle.

En vue du nombre important de QTLs rendement identifiés, de leur impact sur d'autres caractères, de leur interaction avec l'environnement et du nombre croissant de marqueurs moléculaires disponibles, l'utilisation de la sélection génomique pourrait donc être une solution pour améliorer la productivité de nos variétés.

4. Conclusions et perspectives

Ce projet a donc permis d'identifier 37 QTLs de rendement. Trente quatre QTLs ont pu être

Références bibliographiques

- Akhunov E., Nicolet E., Dvorak J.** (2009) Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina GoldenGate assay. *TAG* 119:507-519.
- Gilmour A.R., Gogel B.J., Cullis B.R., Thompson R.** (2009) ASReml User Guide Release 3.0 VSN International Ltd, Hemel Hempstead, HP1 1ES, UK, www.vsnl.co.uk.
- Kirigwi F.M., Van Girkel M., Brown-Guerida G., Gill B.S., Paulsen G.M., Fritz A.K.** (2007) Markers associated with a QTL for grain yield in wheat under drought. *Mol. Breeding* 20: 401-413.
- Kuchel H., Williams K.J., Langridge P., Eagles H.A., Jefferies S.P.** (2007a) Genetic dissection of grain yield in bread wheat. I. QTL analysis. *Theor Appl Genet* 115: 1029-1041.
- Kuchel H., Williams K.J., Langridge P. et al.** (2007b) Genetic dissection of grain yield in bread wheat. II. QTL by environment interaction. *Theor Appl Genet* 115: 1015-1027.
- Li S.S., Jia J.Z., Wei X.Y., Zhang X.C., Li L.Z., Chen H.M., Fan Y.D., Sun H.Y., Zhao X.H., Lei T.D., Xu Y.F., Jiang F.S., Wang H.G., Li L.H.** (2007) A intervarietal genetic map and QTL analysis for yield traits in wheat. *Mol Breed* 20:167–178.
- Paux E., Sourdille P., Salse J., Sautenac C., Choulet F., Leroy P., Korol A., Michalak M., Kianian S., Spielmeier W., Lagudah E., Somers D., Kilian A., Alaux M., Vautrin S., Bergès H., Eversole K., Appels R., Safar J., Simkova H., Dolezel J., Bernard M., Feuillet C.** (2008). *Science*: 322: 101-104.
- Quarrie S.A., Steed A., Calestani C., Semikhodskii A., Lebreton C., Chinoy C., Steele N., Pljevljakusic D., Waterman E., Weyen J., Schondelmaier J., Habash D.Z., Farmer P., Saker L., Clarkson D.T., Abugalieva A., Yessimbekova M., Turuspekov Y., Abugalieva S., Tuberosa R., Sanguineti M.C., Hollington P.A., Aragues R., Royo A., Dodig D.** (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.
- Quarrie S.A., Pekic Quarrie S., Radosevic R., Rancic D., Kaminska A., Barnes J.D., Leverington M., Ceoloni C., Dodig D.** (2006) Dissecting a wheat QTL for yield present in a range of environments: from QTL to candidate genes. *Journal of Experimental Botany* 57:2627-2637.
- Quraishi U.M., Abrouk M., Bolot S., Pont C., Throude M., Guilhot N., Confolent C., Praud S., Murigneux A., Charmet G., Salse J.** (2009) Genomics in cereals: from genome-wide conserved orthologous set (COS) sequences to candidate genes for trait dissection. *Funct. Integr. Genomics* 9: 473-484.
- Snape J.W., Foulkes M.J., Simmonds J., Leverington M., Fish L.J., Wang Y., Ciavarella M.** (2007) Dissecting gene x environmental effects on wheat yields via QTL and physiological analysis. *Euphytica* 154:401–408.
- Sun X.Y., Wu K., Zhao Y., Kong F.M., Han G.Z., Jiang H.M., Huang X.J., Li R.J., Wang H.G., Li S.S.** (2009) QTL analysis of kernel shape and weight using recombinant inbred lines in wheat. *Euphytica* 165:615–624.