

Développement, par génétique d'association, d'un outil moléculaire de prédiction du poids spécifique chez le blé tendre

Jacques BORDES¹, Olivier ROBERT^{*2}, Ellen GOUEMAND², Laure DUCHALAIS³, Valérie LAURENT², Marie Reine PERRETANT¹, François-Xavier OURY¹, Annie LAPIERRE¹, Laëtitia CHEVARIN¹, Denis BEGHIN², Eric STORLIE¹, Bernard ROLLAND⁴, Gilles CHARMET¹

* Coordinateur : Olivier ROBERT¹, olivier.robert@florimond-desprez.fr, Tél. : 03 20 84 94 90

1 - INRA-GDEC - 5 chemin de Beaulieu, 63039 Clermont-Ferrand Cedex 2

2 - BIOPLANTE - Florimond Desprez, BP41, 59242 Cappelle en Pévèle

3 - BIOPLANTE - R2n, 60 rue Léon Beauchamp, 59930 La Chapelle d'Armentières

4 - INRA-APBV - Domaine de la Motte, BP 35327, 35653 Le Rheu cedex

1. Introduction

Le poids spécifique (PS) est un critère économique majeur pour la commercialisation du blé. En effet, les industriels recherchent des poids spécifiques élevés, de façon à avoir de la masse sous un faible volume. Ainsi, les lots apportés à l'intervention doivent répondre à des normes minimales de qualité. De plus, les prix payés par les organismes stockeurs (coopératives, négociants) tiennent compte de ces mêmes normes. Ainsi, le seuil du PS retenu pour les contrats commerciaux est de 76 kg/hl. Le prix d'intervention subit une réfaction de -0.5 €/t lorsque le PS est compris entre 76 et 75 kg/hl et jusqu'à -1.5 €/t kg/hl pour des PS compris entre 74 et 73 kg/hl.

Le PS étant un caractère relativement héritable, il est possible de le sélectionner. Actuellement, les sélectionneurs évaluent le poids spécifique à posteriori par des pesées d'échantillons, mais sans pouvoir orienter efficacement leurs croisements en fonction de ce critère, ni le prendre en compte durant les toutes premières générations de sélection. Le développement dans ce projet, d'un outil moléculaire de prédiction du PS, aidera le sélectionneur dans les choix des croisements et/ou le tri précoce dans les descendances. Cela lui permettra d'obtenir plus efficacement de nouvelles variétés et de fournir aux agriculteurs français, un plus grand choix de variétés avec un poids spécifique amélioré.

La mise au point d'un outil de prédiction nécessite deux étapes : une étape de calibration et une étape de validation. L'étape de calibration consiste à identifier dans des populations, par génétique d'association ("Linkage Disequilibrium Mapping"), des marqueurs moléculaires liés aux caractères étudiés. L'étape de validation consiste à mettre au point des équations de prédiction de sélection génomique pour ces caractères et de les valider dans une nouvelle population, en comparant les résultats de phénotypage observés, à ceux obtenus par la prédiction.

Comme l'analyse QTL, devenue classique depuis l'avènement des marqueurs moléculaires, l'analyse d'association vise à tester la dépendance statistique entre un polymorphisme moléculaire et une variation phénotypique.

En exploitant des populations naturelles d'individus non apparentés, l'accumulation des recombinaisons historiques doit logiquement conduire à une portée très réduite du déséquilibre de liaison (DL), moins de 2 kb chez le maïs (Flint-Garcia *et al* 2003, 2005). Dans ce cas, seuls les polymorphismes très proches du facteur causal de la variation sont liés de façon significative à la variation du caractère. Cette approche a été

appliquée avec succès chez les plantes, en particulier à l'INRA chez le maïs (Thornsberry *et al* 2001) et le blé (Ravel *et al* 2006, Bordes *et al* 2011, Le Gouis *et al* 2012). Des considérations théoriques ont amené à penser que, dans certains cas, la portée du DL pourrait s'étendre à plusieurs cM : une couverture du génome avec quelques milliers de marqueurs, pourrait alors être suffisante. Les paramètres qui peuvent accroître la portée du DL sont, entre autres, l'autogamie et la consanguinité du matériel. La sélection, exercée sur un gène candidat, tendrait également à renforcer la portée du DL local (Rhoné *et al* 2007). Des études expérimentales chez l'orge (Kraakman *et al* 2004) et le blé (Somers *et al* 2007) ont confirmé la faisabilité d'une analyse « genome-wide » avec une densité de marquage de l'ordre du cM. Les marqueurs DArT (Diversity Arrays Technology) obtenus à partir d'une technologie développée par Triticarte Pty Ltd peuvent répondre à cette contrainte. Toutefois, le risque de « spurious associations » demeure, voire est encore augmenté si l'on utilise par exemple un set de variétés européennes ou des lignées de sélectionneurs, qui présentent des taux d'apparentement élevés.

La difficulté dans ce type d'approche est le contrôle des faux positifs causés par l'existence d'une structure des populations ou de différentes parentés entre les individus, se traduisant par des DL significatifs entre loci non liés. Afin de remédier à ce problème, une solution consiste à rechercher le modèle le plus approprié, entre un modèle linéaire prenant en compte la structure (Pritchard *et al* 2000) et un modèle mixte prenant en compte : soit les coefficients de parenté seuls, soit combinés avec la structure, dans les tests d'association (Meuwissen and Goddard 2001, Yu *et al* 2006).

Cependant, avec un nombre important de données (marqueurs, individus et caractères), la méthode MLM nécessite un temps de calcul très lourd. Afin de réduire cette durée, Zhiwu Zhang *et al.* (2010), décrivent deux méthodes, dérivées de l'approche "sire model" utilisée en sélection animale : "compressed MLM method" and "population parameters previously determined" (P3D). Cette dernière, dérivée de la première et qui fait appel aux coefficients de parentés entre individus est maintenant implantée dans le logiciel : TASSEL3.0.66 (<http://www.maizegenetics.net>).

La sélection génomique, proposée par Meuwissen *et al* en 2001, consiste à utiliser de façon globale, c'est-à-dire sans procédure de choix a priori, des marqueurs moléculaires « genomewide », pour prédire la valeur génétique des candidats. Cette valeur génétique est toujours inconnue,

sauf pour des données simulées selon un modèle connu), et la sélection « classique », qui est donc toujours indirecte, porte sur un prédicteur de cette valeur génétique (breeding value), basé sur le phénotype des candidats eux-mêmes ou de leurs apparentés (filles ou petites filles, par exemple, pour estimer la valeur laitière d'un taureau). La réponse à cette sélection « phénotypique » dépend de la corrélation entre la vraie valeur génétique et le phénotype mesuré (régression géno-phénotypique), dont le coefficient est égal à l'héritabilité au sens large du caractère pour la population considérée (Gallais 1990, p55).

Comme l'ensemble des gènes gouvernant les caractères complexes (QTL) n'est pas encore connu, l'utilisation de marqueurs en nombre suffisant devrait permettre de capturer l'information apportée par les QTL. Pour que cette capture soit efficace, il faut que tous les QTL soient en déséquilibre de liaison (DL) significatif avec au moins un marqueur. C'est donc l'étendue et la distribution du DL qui détermine la densité de marqueurs à utiliser. Ainsi, chez le blé, les estimations, bien qu'encore peu nombreuses, donnent des étendues allant de 300 kB à plusieurs cM, selon le matériel considéré (Horvath *et al* 2009, Trebbi *et al* 2011). Il est donc encore difficile de dire quelle densité de marquage sera nécessaire pour avoir des prédictions génomiques précises.

La calibration du modèle de prédiction nécessite d'avoir à la fois des génotypes et des phénotypes (comme succédané de la valeur génétique). Une fois les équations établies et leur qualité éprouvée par validation croisée, la valeur génétique peut être estimée pour des candidats dont on connaît seulement le génotype aux marqueurs. L'utilisation de marqueurs plutôt que des phénotypes pour estimer la valeur génétique présente un intérêt si :

- Le coût du marquage est très inférieur à celui du génotypage
- Le génotypage permet de raccourcir les cycles de sélection

Dans le cas d'une plante annuelle comme le blé, dont le phénotypage pour le rendement est relativement peu cher l'avantage économique paraît moins évident par exemple que pour les bovins laitiers.

Finalement, les principaux objectifs de ce projet étaient multiples :

- Identification de lignées ou de variétés avec un poids spécifique élevé et stable
- Obtention de données moléculaires à partir d'une puce à ADN de 5000 marqueurs DArTs, révélant la diversité génétique du matériel étudié. Ces données pourront être exploitées par exemple pour orienter le choix des géniteurs à croiser.
- Acquisition de la méthode d'analyse "génétique d'association à l'échelle du génome" (contrôle des risques statistiques adaptés au matériel de sélection), qui pourra être réinvestie pour tout autre caractère.
- Identification de marqueurs moléculaires liés au PS, au rendement et à la précocité : outil de prédiction.
- Validation de l'outil de prédiction sur un autre panel de variétés, afin d'en évaluer sa robustesse et sa pertinence.

2. Matériel et méthodes

► Matériel végétal

Le matériel étudié est composé de 2 populations de 376 lignées haploïdes doublées (HD1 et HD2) fournies par

Bioplante et 376 lignées recombinantes (RIL) apportées par l'INRA. Le matériel des deux partenaires a été choisi pour être complémentaire :

- Le taux de recombinaison entre le matériel HD et RIL est différent.
- Le matériel Bioplante (HD) est composé de lignées HD produites pour obtenir des variétés destinées au marché Français voire européen.
- Le matériel végétal (RIL) proposé par l'INRA représente une diversité génétique plus importante.

Le matériel HD est composé de 4 à 5 lignées issues d'environ 80 croisements. Ces lignées auront été prises au hasard. Le matériel RIL est constitué des 376 lignées les plus avancées du programme de sélection INRA (blés rustiques de qualité). Il s'agit de lignées F8 issues du croisement d'une centaine de géniteurs différents et qui ont été évaluées dans le réseau INRA (inter-stations A) entre 2000 et 2010.

► Phénotypage

Les expérimentations ont été réalisées en 2009 (HD1), 2010 (RIL) et 2011 (HD2) dans 3 lieux différents. En 2009, la population HD1 a été implantée à Clermont-Ferrand (63), Cappelle-en-Pévèle (59) et Milly-la-Forêt (91). En 2010, la population RIL a été implantée à Clermont-Ferrand, Cappelle-en-Pévèle et Mons en Chaussée (80). En 2011 la population HD2 a été implantée à Cappelle-en-Pévèle, Milly-la-Forêt et Rennes (35). Le niveau d'intrants correspondait aux pratiques culturales utilisées habituellement dans chaque lieu, pour des objectifs à haut rendement : forte densité de semis, niveau élevé de fertilisation azotée et pesticides appropriés.

Toutes les lignées, cultivées sur des parcelles de 10 m² ont été implantées en blocs aléatoires (1 bloc par lieu). Pour estimer l'hétérogénéité possible du sol, dans chaque bloc, deux variétés Apache et Caphorn ont été répétés 10 fois. Le poids spécifique (PS, kg/hL), le rendement en grains (RDT, Qtx/ha à 0% d'humidité) et la date d'épiaison (Ep, jours à compter du 1er Janvier) ont été déterminés pour chaque parcelle.

► Génotypage

Le génotypage a été réalisé par Triticarte Pty Ltd, à partir d'une puce à ADN de 5000 marqueurs DArT, testés sur les 3 populations. Avec les 2700 qui étaient polymorphes, une carte génétique a été réalisée à partir d'une carte consensus, construite avec le logiciel MetaQTL (Veyrieras *et al* 2007), à partir de données publiées. Les marqueurs non cartographiés dans la carte consensus ont été placés à côté du marqueur DArT avec lequel le DL était le plus fort. Ce DL a été calculé avec R.

Nous avons considéré les allèles rares (<2,5% du nombre total d'allèles) comme données manquantes.

► Déséquilibre de liaison

Le DL entre les loci (DArTs cartographiés), représenté par le R² a été calculé avec TASSEL3.0.66. Pour chaque population, le DL a été calculé séparément pour les loci non liés et ceux portés par le même chromosome. Au sein d'une région chromosomique, si plusieurs marqueurs étaient signalés au même locus de la carte de référence, un seul était utilisé pour l'analyse. Le seuil de signification par paire a été calculé pour 1000 permutations. Le seuil p < 0.001 a été retenu. La valeur

critique du R² (environ 0.20) a été déterminée en utilisant la valeur du quantile à 95% de la distribution de la racine carrée des R² des loci non liés (méthode décrite par Brescghello et Sorrells, 2006).

La décroissance du DL sur la distance génétique a été représentée par un histogramme pour les R² significatifs en fonction des distances.

► Associations génétiques

Afin d'éviter les risques de type I (détection de fausses associations) et de type II (non détection de vrais associations), différentes méthodes statistiques d'analyse de données, pouvant prendre en compte la structure des populations et la parenté entre individus, ont été développées. Une première étape de recherche du modèle le mieux adapté, a été réalisée à partir de l'épiaison, caractère très héritable. Pour chaque population deux matrices de structure (Q) ont été déterminées, une à l'aide du logiciel "Structure" (Q) et une deuxième à partir des coordonnées des axes d'une analyse en composante principale (ACP) calculée avec R⁺. Les coefficients de parenté ou "Kinship" (K) entre les individus, ont été calculés avec TASSEL3.0.66.

Les modèles testés sur l'épiaison ont été : GLM (sans correction de structure de la population), GLM-Q et GLM-ACP (avec matrice de structure et matrice des coordonnées des ACP pour correction de la structure) et MLM-K (avec Kinship pour correction des différentes parentés), MLM-Q-K et MLM-ACP (correction à la fois de la structure et des différentes parentés).

La recherche du meilleur ajustement a été faite à partir de l'appréciation graphique de la concordance entre une distribution observée et un modèle théorique représentée par le diagramme *probabilités observées-probabilités cumulées* (Kang et al 2008).

Le modèle le mieux adapté a ensuite été utilisé pour la recherche d'associations avec les autres caractères.

► Construction d'un outil de prédiction par les marqueurs moléculaires (sélection génomique)

Les méthodes suivantes ont été comparées :

- 1 - pedigree-BLUP : il s'agit du BLUP « modèle animal » ou la matrice de parenté, habituellement reconstruite à partir des pédigrées, est estimée à l'aide des marqueurs (library pedigree de R)
- 2 - RR-BLUP : ridge regression BLUP (Piepho 2009) : un estimateur BLUP des effets de chaque marqueur est estimée par une méthode de régression avec pénalisation pour éviter le surparamétrage liée au problème $n \gg p$ (plus de marqueurs que d'observations)
- 3 - Deux méthodes bayésiennes : Bayesian Ridge Regression (BRR) et Bayesian LASSO : les deux méthodes diffèrent par les distributions des effets dans les prior. Ces deux méthodes ont été mises en œuvre avec la library BRR de R (Perez et al 2010)
- 4 - Une méthode non paramétrique apparentée aux « support vector machine » (SVM), la random Hilbert Kernel Space (RKHS)
- 5 - Une méthode basée sur l'échantillonnage parmi des classifications semi-supervisées : la random forest regression.

Toutes les méthodes ont été mises en œuvre sous R. Les scripts des fonctions ont été échangés entre les partenaires.

3. Résultats et discussions

► Phénotypage

Les Tableaux 1 à 3 indiquent, pour chaque population, les résultats moyens, l'étendue et les corrélations entre les lieux, des caractères rendement (RDT), poids spécifique (PS) et épiaison (Ep).

Lieu-Caractère	Moyenne	Etendue	Ca-RDT	Clit-RDT	Mi-RDT	Ca-PS	Clit-PS	Mi-PS	Ca-Ep	Clit-Ep
Ca-RDT	97.6	72.8 - 114.4	1.00							
Clit-RDT	88.6	58.5 - 110.8	0.34	1.00						
Mi-RDT	96.2	64.9 - 112.7	0.62	0.08	1.00					
Ca-PS	78.2	72.5 - 83.9	-0.04	0.20	-0.24	1.00				
Clit-PS	77.5	70.5 - 83.0	-0.07	0.36	-0.21	0.75	1.00			
Mi-PS	76.2	69.9 - 82.1	-0.15	0.18	-0.17	0.67	0.65	1.00		
Ca-Ep	150.6	143 - 158	0.25	-0.20	0.37	-0.22	-0.20	-0.25	1.00	
Clit-Ep	145.2	136 - 153	0.27	-0.22	0.39	-0.20	-0.24	-0.23	0.90	1.00
Mi-Ep	140.5	133-149	0.27	-0.21	0.35	-0.19	-0.22	-0.24	0.92	0.92

En rouge : corrélation significative à $p < 0.05$. Ca : Capelle, Clit : Cleemont, Mi : Milly

Tableau 1 Population HD1 : Moyenne, étendue et corrélation entre les lieux pour rendement (RDT), poids spécifique (PS) et épiaison (Ep)

Lieu-Caractère	Moyenne	Etendue	Ca-RDT	Clit-RDT	Mo-RDT	Ca-PS	Clit-PS	Mo-PS	Clit-Ep
Ca-RDT	102.7	66.3 - 130.9	1.00						
Clit-RDT	92.1	61.2 - 112.8	0.23	1.00					
Mo-RDT	86.1	59.4 - 100.9	0.36	0.26	1.00				
Ca-PS	74.3	66 - 80.6	0.34	-0.05	-0.16	1.00			
Clit-PS	71.2	61 - 79.2	-0.04	0.06	-0.10	0.36	1.00		
Mo-PS	75.9	69.7 - 81.2	0.07	-0.06	-0.12	0.65	0.62	1.00	
Clit-Ep	137.7	128 - 147	0.21	0.14	0.31	-0.14	0.02	-0.02	1.00
Mi-Ep	150.4	142 - 159	0.19	0.06	0.17	-0.15	-0.13	-0.17	0.76

En rouge : corrélation significative à $p < 0.05$. Ca : Capelle, Clit : Cleemont, Mo : Mons

Tableau 2 Population RIL : Moyenne, étendue et corrélation entre les lieux pour rendement (RDT), poids spécifique (PS) et épiaison (Ep)

Lieu-Caractère	Moyenne	Etendue	Ca-RDT	Mi-RDT	Re RDT	Ca-PS	Mi-PS	Re-PS	Ca-Ep	Re-Ep
Ca-RDT	106.2	78.9 - 124.8	1.00							
Mi-RDT	97.5	73.9 - 118.4	0.27	1.00						
Re-RDT	107.8	56.8 - 134.0	0.31	0.22	1.00					
Ca-PS	79.4	73.4 - 83.2	0.16	-0.25	-0.08	1.00				
Mi-PS	80.0	73.6 - 88.0	-0.23	-0.01	-0.02	0.33	1.00			
Re-PS	75.8	68.8 - 81.0	0.07	-0.28	0.19	0.75	0.38	1.00		
Ca-Ep	137.7	127 - 151	0.51	0.05	0.02	0.32	-0.31	0.21	1.00	
Re-Ep	128.5	114 - 140	0.50	0.09	-0.02	0.33	-0.17	0.23	0.87	1.00
Mi-Ep	132.1	122 - 141	0.45	0.06	-0.06	0.33	-0.16	0.22	0.88	0.93

En rouge : corrélation significative à $p < 0.05$. Ca : Capelle, Mi : Milly, Re : Rennes

Tableau 3 Population HD2 : Moyenne, étendue et corrélation entre les lieux pour rendement (RDT), poids spécifique (PS) et épiaison (Ep)

Le rendement moyen, relativement fort pour l'ensemble des lieux, indique un matériel végétal de très bon niveau agronomique. Le niveau, sensiblement équivalent entre HD1 et RIL est légèrement supérieur pour HD2. Excepté entre Clermont et Milly pour HD1, les corrélations entre les lieux sont significatives. Elles sont en général assez faibles (0.22 à 0.62), indiquant la présence d'une interaction génotype*milieu. L'étendue, importante à très importante, proche de 80 Qtx pour HD2 sur le lieu Rennes, montre une forte variabilité des 3 populations pour ce caractère.

Le poids spécifique semble relativement plus faible pour la population RIL. Il y a une relativement meilleure corrélation entre les lieux pour la population HD1. Quelle que soit la population considérée, l'étendue est égale ou supérieure à 10. Elle peut aller jusqu'à 18 pour la population RIL à Clermont.

L'étendue pour l'épiaison de la population HD2 est la plus importante. La corrélation entre les lieux est très bonne pour HD1 et HD2, elle est relativement plus faible pour RIL.

► Génotypage et Déséquilibre de liaison

Finalement 2575 (HD1), 1436 (HD2) et 2172 (RIL) DArTs ont été utilisés pour tester les associations, 1154 sont communs aux 3 populations (Tableau 5).

L'étendue du déséquilibre de liaison (DL) est représenté Figure 1. Il est sensiblement équivalent pour les 3 populations avec toutefois une étendue relativement plus faible pour HD2.

	HD1	HD2	RIL	Communs
	2575	1436	2172	
HD1 - HD2	X	X		1435
HD1 - RIL	X		X	1951
HD2 - RIL		X	X	1235
HD1 - HD2 - RL	X	X	X	1154

Tableau 4 : Nombre de DArTs polymorphes et communs entre les populations

Le nombre d'étapes de recombinaisons réduit à 2 dans le processus d'obtention des lignées HD par rapport aux lignées RIL en 8^{ème} année d'autofécondation, ne change pas l'étendue du DL.

Population	Paires	DL $p < 0.001$	DL $p < 0.001$ et $R^2 > 0.2$
HD1	25617	5.16	0.16
HD2	22347	5.00	0.13
RIL	24043	7.35	0.10

Tableau 5 : Paires de locus inter-chromosomiques en déséquilibre de liaison

Le nombre de paires de locus inter-chromosomique en DL, significatif à $p < 0.001$, légèrement supérieur pour la population RIL, va dans le sens opposé à ce qui est attendu du fait du plus grand nombre de recombinaison subi par les lignées autofécondées. Si l'on ne retient que les DL significatifs à $p < 0.001$, mais avec un $R^2 > 0.2$, le pourcentage très faible est équivalent pour les 3 populations. Ceci indique que les 3 populations n'ont pas une structure très marquée.

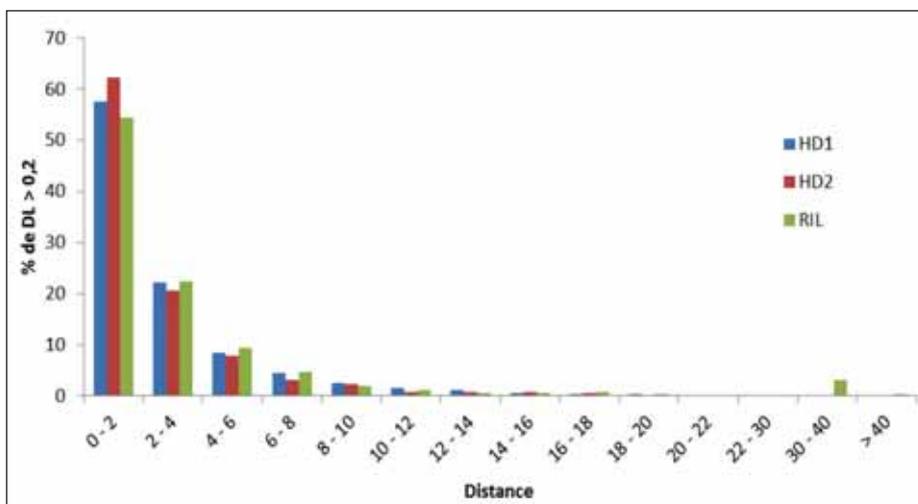


Figure 1 : Fréquence des DL en fonction de la distance génétique



Figure 3 : QTLs communs à 2 ou 3 populations pour Ep, PS et RDT

Le seuil de probabilité retenu pour déclarer une association significative a été de $p < 0.01$.

Le bilan global de tous les marqueurs associés au poids spécifique, au rendement et à l'épiaison, dans chaque lieu

et pour la moyenne des 3 lieux, est indiqué Annexes 1, 2 et 3. Puisque le DL entre 2 marqueurs liés peut être significatif jusqu'à 10 cM (Figure 1) nous avons considérés que les marqueurs significatifs, séparés par une distance voisine de 10 cM, faisaient partie du même QTL (Figure 3).

Pour le poids spécifique, 143 marqueurs répartis sur l'ensemble du génome, étaient associés. On peut observer 8 QTLs communs aux 3 populations (HD1-HD2-RIL) et 9 communs à 2 populations. Pour le rendement, 186 marqueurs étaient impliqués. 5 QTLs sont communs aux 3 populations et 18 sont communs à 2. Pour l'épiaison, 126 marqueurs étaient impliqués. 3 QTLs sont communs aux 3 populations et 16 sont communs à 2. Malgré un dispositif expérimental relativement simple et à partir de populations différentes, on observe une bonne répétabilité des résultats. Celle-ci pourrait être améliorée, soit en augmentant le nombre de lieux ou de répétitions, soit en ayant une meilleure couverture du génome en augmentant le nombre de marqueurs. La deuxième solution pourrait être réalisée à un coût moindre.

4 QTLs colocalisent entre les 3 caractères, dont 3 sur le groupe 6. 11 colocalisent entre le PS et le RDT. Avec ce matériel de type "élite", ce résultat est conforme à ce que l'on peut attendre. Le poids spécifique n'est pas corrélé au rendement (Tableaux 1, 2 et 3), mais étant devenu un critère important de sélection, il est tout à fait envisageable que l'on soit en présence d'un phénomène de "linkage drag". La sélection au cours du temps s'opérant sur un haplotype comportant 2 gènes, un impliqué au niveau du RDT et l'autre sur le PS.

En conclusion, l'expérimentation conduite a permis l'identification de lignées ayant un poids spécifique élevé.

Des marqueurs DArTs polymorphes à partir d'une puce à ADN de 5000 marqueurs ont été cartographiés : 2575 (HD1), 1436 (HD2) et 2172 (RIL) révèlent ainsi une diversité génétique du matériel étudié. Ces données pourront donc être exploitées, principalement au niveau de la matrice de kinship, pour orienter le choix des géniteurs à croiser.

Génotypes	Croisements de Prebreeding	FDHD42
FDHD42		1.00
FDHD38	a	0.79
HDAssSR141	b	0.63
FDHD61		0.40
HDAssSR126		0.25
HDAssSR41		0.25
HDAssSR40		0.25
HDAssSR24		0.15
HDAssSR80		0.10
HDAssSR172		0.09
CF08084		0.08
EM08221		0.07
HDAssSR21		0.05
HDAssSR25		0.04
FDHD74	d	0.04
HDAssSR143		0.04

Coefficient de parenté : 1 = parenté complète et 0 = absence de parenté.
Exemples de croisements possible avec FDHD42 : a et b) FDHD38 et HDAssSR141 = entre lignées proches, c) HDAssSR141 avec une lignée plus éloignée et d) avec une lignée sans aucune parenté

Tableau 6 : Matrice de parenté entre FDHD42 et quelques lignées de la population HD2 et exemple de croisements

Une méthodologie d'application de la "génétique d'association à l'échelle du génome" avec recherche du "meilleur" modèle, adapté au matériel de sélection, pour le contrôle des risques statistiques, a été décrite et expérimentée. Des marqueurs moléculaires, liés au PS, au rendement et à la précocité ont été identifiés ils peuvent servir d'outil de prédiction.

► Prédiction génomique

Sur le panel d'entraînement 2010 (INRA), le comportement des différentes méthodes varie d'un caractère à l'autre (donc probablement selon l'architecture inconnue des QTL), mais aussi selon les lieux, comme illustré Figure 4 :

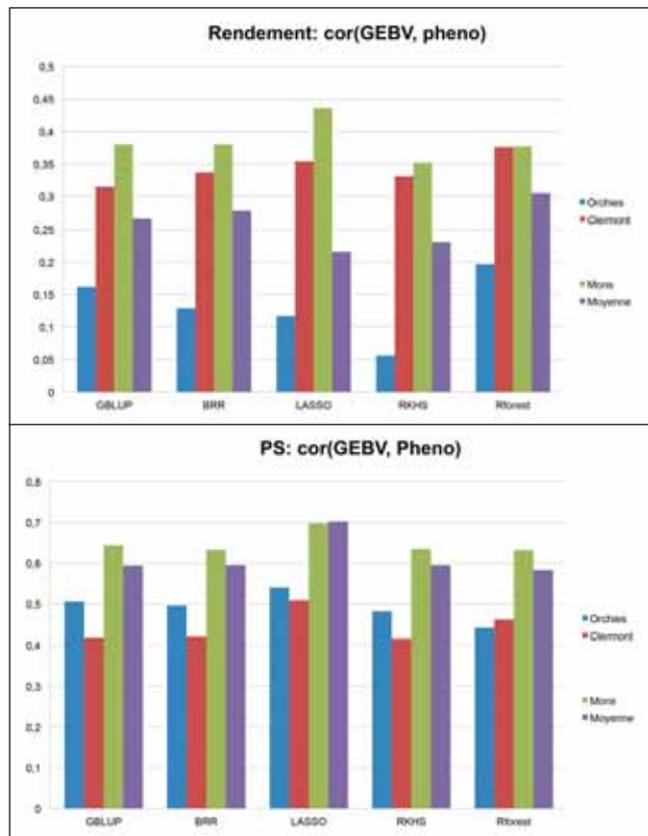


Figure 4 : Efficacité (corrélations entre GEBV et phénotype) par validation croisée pour 5 méthodes statistiques dans les 3 lieux et la moyenne, pour 2 caractères cibles.

Pour le rendement, c'est la méthode LASSO qui semble se détacher dans certains lieux, mais pas pour la moyenne. Et pour le poids spécifique, c'est encore la méthode LASSO, mais cette fois avec une plus grande stabilité entre lieux.

Mais l'intérêt d'une prédiction génomique est également, et sans doute surtout, de permettre de prédire la GEBV de lignées autres que celles du panel d'entraînement.

Pour estimer cette portabilité des prédictions, un troisième panel de 382 lignées (342 Bioplante et 40 INRA) a été implanté en 2011 dans 3 lieux Capelle (59), Milly (77) et Rennes (35). Les panels 2009, 2010 ou 2009+2010 ont été utilisés comme populations d'entraînement et les équations ont servi à prédire les GEBV du panel 2011. Les résultats sont présentés en Figure 5 pour les 3 caractères cibles. (Les lieux n'étant pas les mêmes, on n'a considéré que la moyenne des 3 lieux). Seuls les marqueurs communs (1098) peuvent être utilisés.

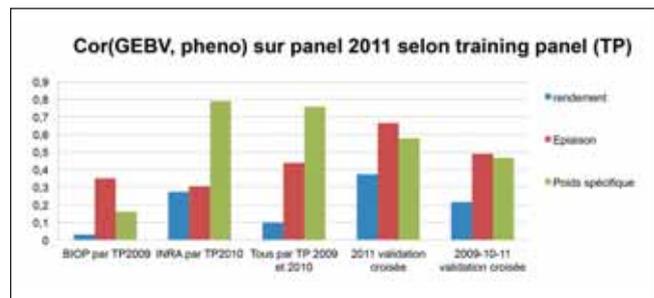


Figure 5

On constate que la qualité de prédiction est du même ordre de grandeur que celle obtenue au sein d'un même panel, donc toujours mauvaise pour le rendement, mais sans être vraiment pire. Par contre, la qualité de prédiction pour le caractère cible du projet, à savoir le poids spécifique, est parfois très bonne, que ce soit au sein du matériel INRA (0.790, NB : 0.802 avec LASSO) ou même pour la prédiction globale de l'ensemble du panel 2011 (Bioplante + INRA) en utilisant les 2 panels 2009 et 2010 comme population d'entraînement ($r=0.758$).

La Figure 6 illustre la qualité de la prédiction dans le cas le plus favorable (INRA2011 par INRA2010).

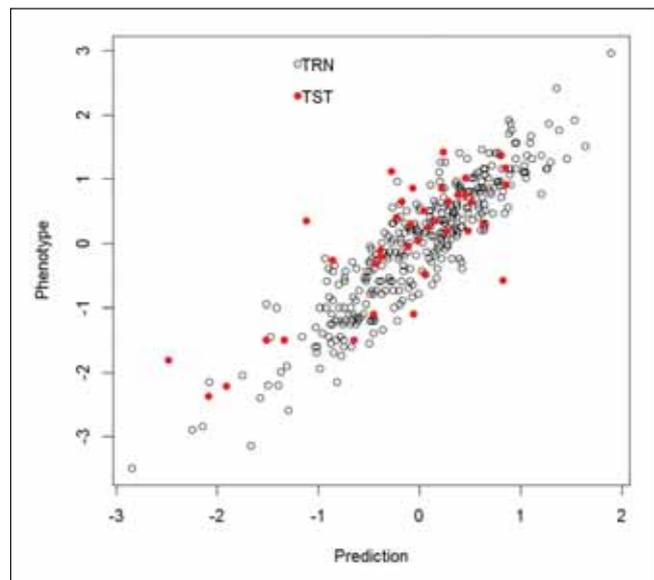


Figure 6

Pour le rendement, on peut avancer l'hypothèse d'une couverture insuffisante du génome, mais aussi celle d'une taille insuffisante de la population de référence, qui ne permet pas aux méthodes, même si elles sont théoriquement prévues pour, de bien répondre au problème de surparamétrisation.

Il est intéressant de regarder, dans les méthodes qui estiment un effet par marqueurs (ridge regression et méthodes bayésienne), quels sont les marqueurs ayant les plus forts effets : retrouve-t-on des marqueurs significativement associés ? Nous avons pris pour exemple la prédiction du poids spécifiques par le panel 2010 ($r=0.790$), et reporté dans le Tableau des associations les 50 marqueurs présentant les plus forts effets génomiques. Sur les 50 marqueurs ayant les plus forts poids dans la prédiction GEBV, 12 correspondent à des associations significatives. Les 2 approches, de philosophie différente

(en SG, les poids peuvent être répartis sur de nombreux marqueurs en DL), donnent des résultats cohérents.

En conclusion, nous pensons que les méthodes de prédiction génomique constituent une alternative intéressante à l'évaluation des phénotypes, surtout quand celle-ci est longue et/ou coûteuse. Néanmoins, des recherches sont encore nécessaires pour améliorer la qualité des prédictions et leur portabilité entre différents germplasmés. Pour cela, la taille de la population de référence est certainement un élément critique, et une

collaboration entre recherche publique et obtenteurs serait sans doute un bon moyen de constituer une population de grande taille, au bénéfice de tous.

3. Remerciements

Nous remercions Laurence Moreau et Alain Charcosset pour leurs conseils au niveau de la génétique d'association et Jacques Le Gouis pour son aide dans la construction de la carte consensus.

Références bibliographiques

- Bordes J., Ravel C., Le Gouis J., Charmet G., Balfourier F. (2011) Use of global wheat core collection for association analysis of flour and dough quality traits. *Journal of Cereal Science* 54: 137-147
- Bresegghello F., Sorrells M.E. (2006) Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics* 172:1165-1177
- Gallais A. (1990) Théorie de la sélection en amélioration des plantes. Masson Ed, Paris, Milan, Barcelone, Mexico.
- Horvath A., Didier A., Koenig J., Exbrayat F., Charmet G., Balfourier F. (2009) Analysis of diversity and linkage disequilibrium along chromosome 3B of bread wheat. *Theor. Appl. Genet.* 119:1523-1537
- Kraakman A.T.W., Niks R.E., Van den Berg P.M.M.M., Stam P., Van Eeuwijk A. (2004) Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. *Genetics* 168:435-446
- Le Gouis J., Bordes J., Ravel C., Heumez E., Faure S., Praud S., Galic N., Remoué C., Balfourier F., Allard V., Rousset M. (2012) Genome wide association analysis to identify chromosomal regions determining components of earliness in wheat. *Theor. Appl. Genet.* 124: 597-611
- Meuwissen T.H.E., Hayes B., Goddard M.E. (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157:1819-1829
- Pérez P., De los Campos G., Crossa J., Gianola D. (2010) Genomic-enabled prediction based on molecular markers and pedigree using the Bayesian Linear Regression Package in R. *The Plant Genome* 3:106-116
- Piepho H.P. (2009) Ridge regression and extensions for genomewide selection in maize. *Crop Sci* 49:1165-1175
- Pritchard J.K., Stephens M., Donnelly P. (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959
- Ravel C., Praud S., Murigneux A., Linossier L., Dardevet M., Balfourier F., Dufour P.H., Brunel D. and Charmet G. (2006) Identification of *Glu-B1-1* as a candidate gene for the quantity of high-molecular-weight glutenin in bread wheat (*Triticum aestivum* L.) by means of an association study. *Theor. Appl. Genet.* 112: 738-743
- Rhoné B., Raquin A.L., Goldringer I. (2007) Strong linkage disequilibrium near the selected *Yr17* resistance gene in a wheat experimental population. *Theor. Appl. Genet.* 114: 787-802.
- Somers, D.J., Banks T., DePauw R., Fox S., Clarke J., Pozniak C. and McCartney C. (2007) Genome-wide linkage disequilibrium analysis in bread wheat and durum wheat. *Genome* 50:557-567
- Thornsberry J.M., Goodman M.M., Doebley J., Kresovich S., Nielson D. et al (2001) Dwarf8 polymorphisms associate with variation in flowering time. *Nat. Genet.* 28: 286-289.
- Trebbi D., Maccaferri M., De Heer P. et al (2011) High-throughput SNP discovery and genotyping in durum wheat (*Triticum durum*). *Theor. Appl. Genet.* 123: 555-569
- Veyrieras J.B., Goffinet B., Charcosset A. (2007) MetaQTL: a package of new computational methods for the meta-analysis of QTL mapping experiments. *BMC Bioinformatics* 8: 1-16
- Yu J., Pressoir G., Briggs W.H., Vroh Bi I., Yamasaki M., Doebley J., McMullen M.D., Gaut B.S., Nielsen D.M., Holland J.B., Kresovich S., Buckler E.S. (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet* 38: 203-208.
- Zhiwu Zhang, <http://www.nature.com/ng/journal/v42/n4/full/ng.546.html> - a1 Elhan Ersoz, Chao-Qiang Lai, Rory J. Todhunter, Hemant K. Tiwari, Michael A. Gore, Peter J. Bradbury, Jianming Yu, Donna K. Arnett, Jose M. Ordovas & Edward S. Buckler (2010) Mixed linear model approach adapted for genome-wide association studies, *Nature Genetics* 42: 355-360

Annexe 1 - Marqueurs associés au poids spécifique pour les 3 populations sur l'ensemble des lieux et $r^2 < 0.01$

Chromosome	Marqueur	Distance	HD1-Ca	HD1-Clt	HD1-Mi	HD1-Moy	HD2-Ca	HD2-Mi	HD2-Re	HD2-Moy	RIL-Ca	RIL-Clt	RIL-Mo	RIL-Moy
1A	wPt-73490	0.45					2.25							
	wPt-665724	16.90					2.20							
	wPt-4765	57.42					2.11							
1B	wPt-1911	21.89		2.17										
	wPt-7480	21.89	1.85	2.86		2.55								
	wPt-730718	100.00					2.49			2.20				
1D	wPt-3738	32.07		2.43		2.01								
	wPt-1387	39.51		3.07		2.41								
	wPt-7946	40.00		3.34		2.43								
	wPt-664824	44.48									2.48			2.00
	wPt-0487	44.60								2.28				
	wPt-5320	52.07					1.80					2.01		
	wPt-5503	52.10										2.54		
	wPt-734229	52.70					2.03							
	rPt-4471	53.59										2.27		
	wPt-3743	70.80	2.56	1.97	2.00	2.83								
	wPt-666719	70.87	2.57	1.98		2.73								
wPt-733835	71.35	2.18												
wPt-667287	72.16	2.13	2.17		2.31								1.94	
2A	wPt-733012	17.45											2.12	
	wPt-743211	39.78			2.25									
2B	wPt-4664	37.46									2.82			
	wPt-2425	91.04									3.36			
	wPt-732040	99.13		3.09										
2D	wPt-734106	0.12			2.74									
	wPt-731336	128.81										3.41		
3A	wPt-745076	16.86										3.79		
	wPt-0714	17.07										3.01		2.18
	wPt-6942	40.00									2.87			
	wPt-0286	57.26					1.99							
	wPt-3816	59.17					2.00							
	wPt-1562	60.83					2.06							
	wPt-4725	82.08	2.01											
	wPt-741749	85.00		2.73										
	wPt-4128	94.91									3.33			
	rPt-9901	95.37									3.38			
wPt-1888	106.82							2.78						
3B	wPt-10006	20.88									2.52		2.06	2.32
	wPt-7504	23.84										3.24	4.07	4.37
	wPt-667891	24.00										3.26	2.75	4.36
	wPt-741485	26.72										3.07		2.25
	wPt-10192	29.55											3.15	2.17
	wPt-742172	31.62									2.60			
	wPt-5432	37.78											3.63	
	wPt-733477	44.09									2.79			
	wPt-666139	46.24										2.03		2.49
	wPt-742225	47.12									3.92			
	wPt-742060	47.12									2.96			
	wPt-8915	52.27										2.17		1.94
	wPt-11295	58.58	1.91											
	wPt-5939	58.95					1.96							
	wPt-4364	68.70						1.87						
wPt-73110	69.59						1.91							

Annexe 1 - (suite)

Chromosome	Marqueur	Distance	HD1-Ca	HD1-Ct	HD1-Mi	HD1-Moy	HD2-Ca	HD2-Mi	HD2-Re	HD2-Moy	RIL-Ca	RIL-Ct	RIL-Mo	RIL-Moy
3D	wPt-731116	36.75	1.90			1.91								
	wPt-731378	38.18			1.94									
	wPt-667098	38.23	1.99			2.08								
	wPt-732812	42.15		2.07										
	wPt-671603	42.64							1.89					
	wPt-4045	42.64	1.88											
	wPt-744794	42.82							2.28					
	wPt-742687	42.84								1.97				
	wPt-741119	42.84								2.19				
	wPt-7535	42.97							1.89					
	wPt-669412	43.79							2.32					
wPt-9389	44.46							2.32						
wPt-9662	50.85							2.28						
4A	wPt-3515	47.28									3.17			
	wPt-744256	71.37		2.83										
	wPt-4620	71.49		2.18										
	wPt-4680	71.49		2.25										
	wPt-5172	71.69			1.97									
	wPt-669203	71.96										2.16		
	wPt-2057	72.25									3.20			
	wPt-0447	0.94			1.81									
	rPt-0238	72.34			1.88									
wPt-730913	72.44			2.40										
4B	wPt-4607	47.86									2.46			
	wPt-1046	47.90									2.47			
4D	wPt-0710	20.00	2.34			2.20								
5A	wPt-1712	47.19			1.94									
	wPt-9452	48.01	1.89											
	wPt-9748	61.50			3.93	2.72								
	wPt-3924	79.42												1.90
	wPt-0907	79.42												2.33
	rPt-6495	164.72					2.03	2.03						
	wPt-4203	166.40					3.03							
	wPt-6462	166.43					2.67							
wPt-4351	168.81					2.82								
5B	wPt-8157	64.09								2.20				
	wPt-4281	64.48								1.87				
	wPt-6680	68.00					2.24							
	wPt-0935	107.23			1.99									
6A	wPt-666927	2.64		2.23		1.87								
	wPt-9832	3.99												2.21
	wPt-667740	8.82									2.64			
	wPt-0168	18.20								1.87				
	wPt-1060	19.30											1.95	
	wPt-2206	87.23									2.72			1.98
	wPt-672030	94.77							1.86					
	rPt6189	95.00							2.14					
	wPt-7548	96.00							2.03					
	wPt-731524	98.00	2.53											
wPt-731002	98.00	2.05												
6B	wPt-664276	20.69					2.61							
	wPt-666793	20.92			2.14									
	wPt-1325	31.25										2.63		
	wPt-3207	34.34									2.41			
	wPt-2810	35.00					2.47							
	wPt-1264	36.03			2.35									
	wPt-3045	36.06			1.90									
wPt-7636	0.88			2.16										

Annexe 1 - (suite)

Chromosome	Marqueur	Distance	HD1-Ca	HD1-Clt	HD1-Mi	HD1-Moy	HD2-Ca	HD2-Mi	HD2-Re	HD2-Moy	RIL-Ca	RIL-Clt	RIL-Mo	RIL-Moy
6D	wPt-733130	36.48					2.90		3.36					
	wPt-664682	37.56					2.85		3.37					
	wPt-731887	37.65					2.89		3.37					
7A	wPt-9255	15.80					1.86							
	wPt-6668	16.42										2.30		
	wPt-9207	17.92						2.85						
	wPt-4835	19.68						2.42						
	wPt-6447	20.42						2.25						
	wPt-3002	68.50		1.83										
	wPt-5476	69.20												
	wPt-5524	91.82		2.22	2.68	2.50								
	wPt-671471	92.00		2.17	2.03	2.35								
	wPt-1706	92.00			1.93									
	wPt-671548	93.16		2.49	2.92	2.83								
	wPt-0639	96.58	2.86		3.40	3.01								
	wPt-0494	109.30							1.91					
	wPt-1023	112.00						3.05						
wPt-1259	113.00						3.71		2.37					
7B	wPt-0963	62.52								1.92				
	wPt-5892	114.26	3.29			2.07								
	wPt-7887	115.93	3.40			2.21								
	wPt-5975	117.00	3.30			2.08								
	wPt-8615	119.00					2.21							
7D	wPt-743140	34.60											1.97	
	wPt-743384	34.60											1.96	
	wPt-744124	34.60											1.96	
	wPt-744866	34.60											1.96	
	wPt-745121	34.60											1.96	
	wPt-731493	75.00									2.49			
	wPt-2256	126.26							2.86					
	wPt-4315	140.20							2.53					
	wPt-4441	144.03											3.09	

Annexe 2 - Marqueurs associés au Rendement pour les 3 populations sur l'ensemble des lieux et $r^2 < 0.01$

Chromosome	Marqueur	Distance	HD1-Ca	HD1-Ct	HD1-Mi	HD1-Moy	HD2-Ca	HD2-Mi	HD2-Re	HD2-Moy	RIL-Ca	RIL-Ct	RIL-Mo	RIL-Moy
1A	wPt-0580	48.10							4.12					
	wPt-734218	55.40							3.37	1.93				
	wPt-732546	55.50							3.05					
	wPt-669499	90.11											2.27	
	wPt-5577	106.75									3.78			2.70
	wPt-7049	107.60										2.74		
	wPt-734074	108.44								2.07				
	wPt-5160	108.62								2.16				3.24
	wPt-6005	108.98								1.65				2.62
	wPt-7117	109.05								1.94				3.25
wPt-732616	109.13								1.82				3.14	
wPt-744613	109.44								2.22					
1B	wPt-2988	4.63			2.06									
	wPt-568	5.60			1.95									
	wPt-2395	6.35			1.97									
	wPt-668076	6.35			1.97									
	wPt-1637	24.70				2.06								
	wPt-5313	28.63					2.87							
	wPt-9973	30.20											2.84	
	wPt-3451	35.92			2.81					1.92				
	wPt-3579	46.10										2.98		
	lPt-7980	57.61											4.05	
	wPt-1103	61.66										2.21		
	wPt-4532	71.37							2.11					
wPt-2861	91.24				1.95									
wPt-730718	100.00					2.05								
1D	wPt-1387	39.51		3.56										
	wPt-7946	41.00		2.88										
	wPt-1145	70.82				1.85								
	wPt-732556	76.73					3.33				2.13		3.00	
	wPt-6059	79.62				1.92	2.77						3.10	
2A	wPt-4711	113.60		3.87		2.20								
	wPt-5738	13.19								1.82				
	wPt-4197	15.75												
	wPt-5839	16.71								2.36				
	wPt-8132	17.36									3.14			
	lPt-0228	17.40									3.03			
	wPt-9712	17.64								2.49				
	wPt-8242	17.77								2.81				
	wPt-7175	18.11								2.70				
	wPt-742886	18.59								2.00				
	wPt-6431	18.65								2.06				
	wPt-9104	19.00								2.32				
	wPt-2309	19.13								2.27				
	wPt-3976	20.07								1.94				
	wPt-7626	27.24								2.26				
2B	wPt-9320	46.31										2.03		
	wPt-9793	117.57					1.96							
	wPt-9274	8.95			1.94									
	wPt-744022	11.82										2.67		
	lPt-4802	12.50			1.83									
	wPt-5672	53.02		2.62										
	wPt-4125	53.53		2.20										
	wPt-0408	53.74										3.45		
	wPt-8548	64.04										2.92		
	wPt-6174	64.07										3.08		
	wPt-2397	90.84					2.87			2.38				
	wPt-732386	90.84					3.00			2.57				
	wPt-0471	99.15							2.38	2.55				
	wPt-733725	99.72							3.73					
	wPt-2261	102.53											2.15	
wPt-665550	105.00					2.99			1.88					
wPt-744181	105.00					4.58			3.27					

Annexe 2 - (suite)

Chromosome	Marqueur	Distance	HD1-Ca	HD1-Clt	HD1-Mi	HD1-Moy	HD2-Ca	HD2-Mi	HD2-Re	HD2-Moy	RIL-Ca	RIL-Clt	RIL-Mo	RIL-Moy
2D	wPt-73169	14.21					2.14							
	wPt-666223	18.12									2.80			
	wPt-666518	79.13							1.86					
	wPt-668239	80.33									4.88			
	wPt-732704	94.24					2.92							
3A	wPt-741190	7.07									2.56			
	wPt-0286	57.26		2.76										
	wPt-3816	59.17		2.81										
	wPt-562	60.83		2.84										
	wPt-2127	62.61					3.08							
	wPt-9761	97.37										3.25		
	wPt-743015	106.31										3.38		
3B	wPt-1888	106.82					1.94	2.24		2.70				
	wPt-9066	23.50						2.05						
	wPt-741465	26.72											3.50	
	wPt-741887	27.00						2.50						
	wPt-3260	30.20										2.65		
	wPt-2510	48.30								2.37				
	wPt-741685	0.23						2.17						
	wPt-0446	67.45											4.40	
	wPt-2904	88.91			2.39	1.9								
	wPt-2491	116.08										2.13		
	wPt-3760	126.39										2.96		
	wPt-7514	126.60	2.12			2.24						2.96		
	wPt-0492	127.41	2.51			3.13								
	wPt-0438	128.95							1.88					
wPt-8983	128.99							1.88						
wPt-0280	129.42							1.95						
3D	wPt-3480	42.64			1.96									
	wPt-4991	42.64			1.95									
	wPt-6107	42.64			1.85									
	wPt-731402	42.64	2.76		2.66									
	wPt-9525	42.64	1.90											
	wPt-4276	43.48			1.98									
	wPt-8655	43.50			1.80									
4A	wPt-730913	72.44			2.86									
	wPt-6502	0.52			2.05									
	wPt-9675	72.52		2.11										
4D	wPt-6966	75.71			2.16									
	wPt-731627	28.95		3.90		2.12								
5A	wPt-3924	79.42												
	wPt-742997	83.00										2.50		
	wPt-743708	83.00										2.09		
	IPt-6495	164.72						2.94						
	wPt-4203	166.40					1.90							
	wPt-6462	166.43					1.82							
5B	wPt-4351	168.81					1.91							
	wPt-1302	37.96											2.57	
	wPt-9666	44.55				1.92			2.23					
	wPt-1432	53.39									3.37			
	wPt-5737	61.43									2.89			
	IPt-7755	75.11	2.03											
	wPt-3661	92.74	2.14											
	wPt-7238	99.37										2.15		
	wPt-4577	107.83										2.60	2.87	4.30
	wPt-6191	111.61										2.59	2.87	4.15
	wPt-2707	115.61										2.13	2.64	3.77
wPt-665750	137.07											2.10		
wPt-730009	137.12											2.31		

Annexe 2 - (suite)

Chromosome	Marqueur	Distance	HD1-Ca	HD1-Ct	HD1-Mi	HD1-Moy	HD2-Ca	HD2-Mi	HD2-Re	HD2-Moy	RIL-Ca	RIL-Ct	RIL-Mo	RIL-Moy
6A	wPt-9858	0.00					2.58							
	wPt-6904	8.99	2.66											
	wPt-3524	14.12	1.87											
	tPt-0877	14.76	2.66											
	tPt-2833	14.76	2.99											
	wPt-666574	14.76	1.92											
	wPt-671855	14.76	2.93			2.01								
	wPt-731842	14.76	3.23			2.14								
	wPt-731834	14.76	2.02											
	wPt-8266	16.35					1.91							
	wPt-732062	16.83				2.32	2.21							
	wPt-0168	18.15				2.15								
	wPt-667405	19.17										2.19		
	wPt-730460	20.06												2.00
	wPt-9132	20.34				2.01								
	wPt-733151	25.28						1.84						
wPt-4229	9141										2.81			
wPt-734140	9168												2.04	
wPt-668773	92.17			1.96										
wPt-731786	92.48								2.11					
6B	wPt-7642	20.91										2.37		
	wPt-741515	21.03	2.12			2.06			2.52					
	rPt6847	22.45	1.81											
	wPt-1241	23.40	2.10			2.38			2.66					
	wPt-5256	24.11	2.11			1.98			2.34					
	wPt-7935	34.73	2.97											
	wPt-2810	34.80							2.33					
	wPt-0419	34.80							2.74					
wPt-4924	44.32							3.21						
wPt-741804	44.35							3.40						
6D	wPt-7394	12.46	3.79											2.62
	wPt-3350	105.65											2.42	3.05
	wPt-731816	106.57												2.36
	wPt-1314	108.23											2.24	2.52
7A	wPt-6417	18.85					2.17							
	wPt-4637	62.80		2.26										
	wPt-6083	63.35		2.07										
	wPt-2100	63.61		2.02										
	wPt-9555	66.58						1.83						
	wPt-4796	69.48		2.07										
	wPt-6620	100.10												2.38
	wPt-731500	105.85										3.07		
wPt-4831	106.70												3.19	
wPt-7947	109.18						1.85							
7B	wPt-664517	20.53	3.47		1.87	3.31								
	wPt-665112	20.53	3.47		1.87	3.31								
	wPt-671801	22.91	3.71		1.90	3.53								
	wPt-0963	62.52					2.40			1.79				
	wPt-0318	63.00					2.35							
	wPt-9467	70.00					2.47			2.36				
	wPt-5975	117.00								1.89				
	wPt-8615	118.00								2.02	3.16			
	wPt-9992	123.50							2.24					
	wPt-4673	128.00												2.21
	wPt-8989	128.00												2.07
wPt-3939	135.05	4.67			3.20									
wPt-665293	139.26	2.22												
wPt-1066	142.36	1.99												
7D	wPt-742840	34.60										2.66		
	wPt-669512	125.70										2.22		
	wPt-664290	126.06										2.07		
	wPt-2258	100										2.14		
	wPt-664391	127.20										2.20		
wPt-4315	140.20						1.95							

Annexe 3 - Marqueurs associés à l'épiaison pour les 3 populations sur l'ensemble des lieux et $r^2 < 0.01$

Chromosome	Marqueur	Distance	HD1-Ca	HD1-Ct	HD1-Mi	HD1-Moy	HD2-Ca	HD2-Mi	HD2-Re	HD2-Moy	RIL-Ca	RIL-Ct	RIL-Mo	RIL-Moy
1A	wP1-0527	34.30	2.32	2.19		2.12								
1B	wP1-741297	5.53			2.45								2.81	
	wP1-0308	13.98												
	wP1-4325	25.05					2.63	2.63						
	wP1-0170	32.62												
	wP1-4655	32.70					2.27							
1D	wP1-668027	36.29			2.69									
	wP1-0705	44.50										3.09		2.05
1D	wP1-664824	44.48												
	wP1-0487	44.60					2.16		2.09					
	wP1-8854	61.94	2.36											
2A	wP1-0100	-27.77		2.06	3.25		1.82							
	wP1-4197	15.75										2.38		
	wP1-668033	17.80				2.13								
	wP1-5245	17.84											2.20	
	wP1-1657	37.80					1.98							
2B	wP1-6361	59.39										3.56	3.29	3.82
	wP1-8058	9.98			2.69									
	wP1-8916	11.37			2.01									
	wP1-744022	11.82	2.47	2.98	4.03	2.70								
	wP1-8398	22.00			3.88									
	wP1-8349	36.29	2.31	2.28	2.64	2.39								
	wP1-4664	37.46	2.04	1.98	2.29	2.00								
	wP1-6932	41.83					1.79							
	wP1-1646	83.97											2.38	2.79
	wP1-6643	89.57								1.85				
	wP1-2397	90.84								2.19				
	wP1-732386	90.84								2.16				
	wP1-7004	92.32					3.26	2.23	2.38	1.99				
	wP1-8916	92.74										3.81	2.27	3.21
	wP1-0471	99.15							2.13					
wP1-665550	0.95						2.62	2.33	3.14					
wP1-744181	0.94								2.80					
wP1-6780	16.16										2.53			
2D	wP1-733227	66.10			3.06									
	wP1-667584	69.44			2.99									
	wP1-731148	69.61			2.89									
	wP1-0298	71.15			2.22									
	wP1-666436	94.24		1.88										
	wP1-3812	94.24					2.84	4.88	2.52					
	wP1-732704	94.24							2.17					
	wP1-731665	94.24							1.88					
	wP1-730934	94.24							2.11					
	wP1-729821	94.24							2.11					
	wP1-671742	94.24		1.98										
3A	IP1-0519	0.42										2.56		
	wP1-745076	16.86										3.29	2.91	3.75
	wP1-2127	62.61							2.12					
3B	wP1-1162	10.55	3.14	2.65	3.58	3.15								
	wP1-2644	31.02			1.91									
	wP1-0302	31.04							2.28	1.97				
	wP1-1191	46.60					2.19		2.05					
	wP1-743395	48.61												1.98
	wP1-4209	66.89										2.20		
wP1-0367	18.36										2.13	2.68	2.98	
3D	wP1-4045	42.64						2.03						
4A	wP1-5857	48.88			2.43									
	wP1-6997	49.71												
	wP1-5124	50.00	2.40	3.53	3.16	3.17								
	wP1-5374	72.08			2.03									
	wP1-4828	72.15						1.90						
wP1-9883	72.24					2.40								
4B	wP1-6669	0.76								1.96				
	wP1-4243	64.00											2.27	
5A	wP1-5588	47.20											2.15	

Annexe 3 - (suite)

Chromosome	Marqueur	Distance	HD1-Ca	HD1-Ct	HD1-Mi	HD1-Moy	HD2-Ca	HD2-Mi	HD2-Re	HD2-Moy	RIL-Ca	RIL-Ct	RIL-Mo	RIL-Moy
5B	wPt-1784	59.40	3.60	2.26	3.19	3.00								
	wPt-667015	109.10			2.08									
	wPt-0927	109.53			1.87									
	wPt-9858	0.00								2.14				
	wPt-666927	2.64	1.96											
	wPt-669271	3.35	1.96											
	wPt-1742	3.57	1.89											
	wPt-731250	15.40					1.97							
	wPt-8266	16.40								1.83				
	wPt-9690	16.54					2.45	3.09	2.32					
	wPt-2573	18.35					1.83	2.20	2.12					
	wPt-730631	18.40						1.91						
	wPt-9132	20.34		2.43	2.26	2.26								
	wPt-5696	9147												2.00
wPt-732370	1017									1.96				
6B	tPt-4887	20.55								2.56				
	wPt-666793	20.92	2.22		2.07									
	wPt-729979	22.62	1.99											
	wPt-743974	25.01					2.39							
	tPt-3506	25.54								2.40				
	wPt-8721	33.16						1.98						
	wPt-9784	33.39	2.66		2.34	2.73								
	wPt-3733	33.50	2.63		2.46	2.56								
	wPt-9952	35.96		2.08	2.40									
	wPt-1264	36.03					2.18	2.14						
	wPt-3045	36.06					2.18	2.14						
	wPt-8493	40.83			1.95									
	wPt-743099	41.06					3.79	2.56	2.75					
	7A	wPt-6417	18.85								1.90			
wPt-1080		60.03	1.85											
wPt-1706		92.00						1.85						
wPt-3425		96.03	1.93											
wPt-533		143.08					2.14							
wPt-2565		5.45											2.05	
wPt-8211		19.87	2.46		2.88	2.16								
wPt-6498		64.96							2.79					
wPt-9467		70.00								2.50				
wPt-9925		96.55							1.88	1.97				
wPt-742351		97.00			2.20									
wPt-731357		97.00			1.87									
wPt-3093		126.86		1.98	2.03									
wPt-4743		142.44					2.01							
wPt-800258	100					1.93								
7D	wPt-733859	31.25											2.15	
	wPt-744148	32.07		2.67										
	wPt-5765	33.49					2.29							
	wPt-6833	33.55					3.34							
	tPt-6015	33.55					3.15							
	rPt7959	33.55					3.27							
	tPt-2440	33.55					3.46							
	wPt-744784	34.05		2.47									2.14	
	wPt-663820	34.75											2.04	
	wPt-663918	34.75											2.08	
	wPt-733104	34.75											2.34	
	wPt-743698	34.75											2.13	
	wPt-743749	34.75											2.15	
	wPt-744889	34.75											2.24	
	wPt-745067	37.99											1.96	
	wPt-731193	75.00										4.02	4.23	4.41
wPt-800147	128.00	4.17	2.66	4.89	3.23									
wPt-666097	141.07	2.20			2.13									