

Introgression massive de diversité par l'utilisation de population de pré-breeding : quelle diversité dans les zones d'intérêt ?

Déterminisme génétique du rendement semoulier et de ses composantes

Variabilité allélique dans le(s) QTL(s) identifié(s)

Frédéric Compan

Martin Ecartot

Aline Rocher

Morgane Ardisson

Etienne Paux

Gilles Charmet

Dimitri Sanchez

Ingrid Vilmus

Jacques David

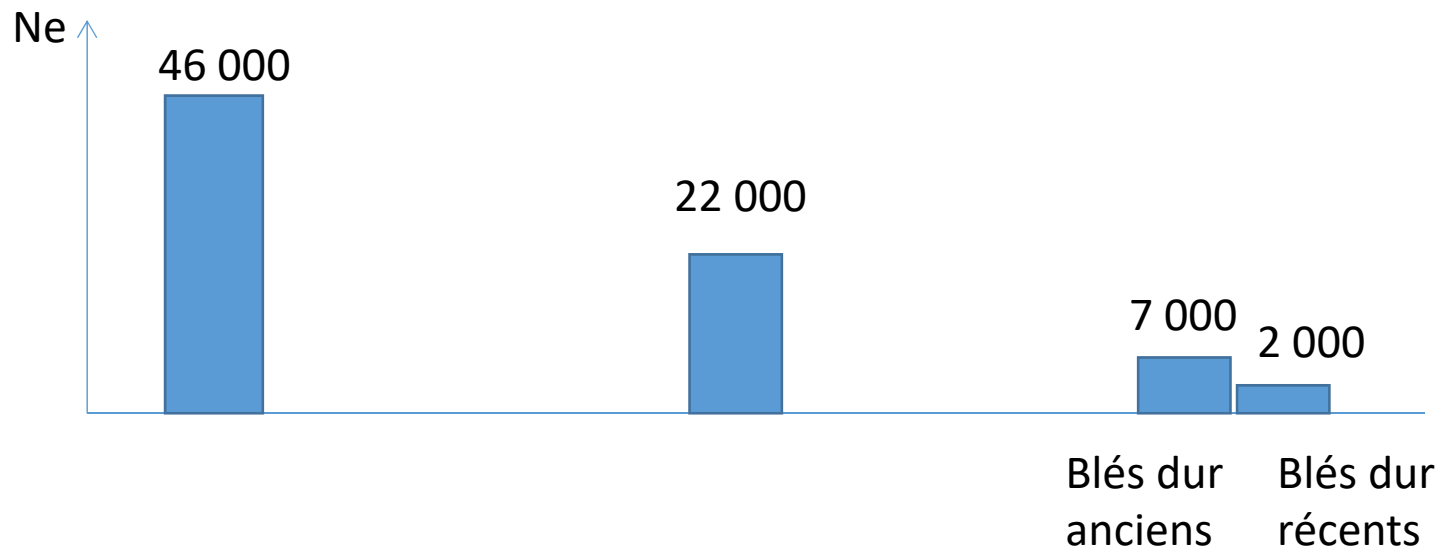
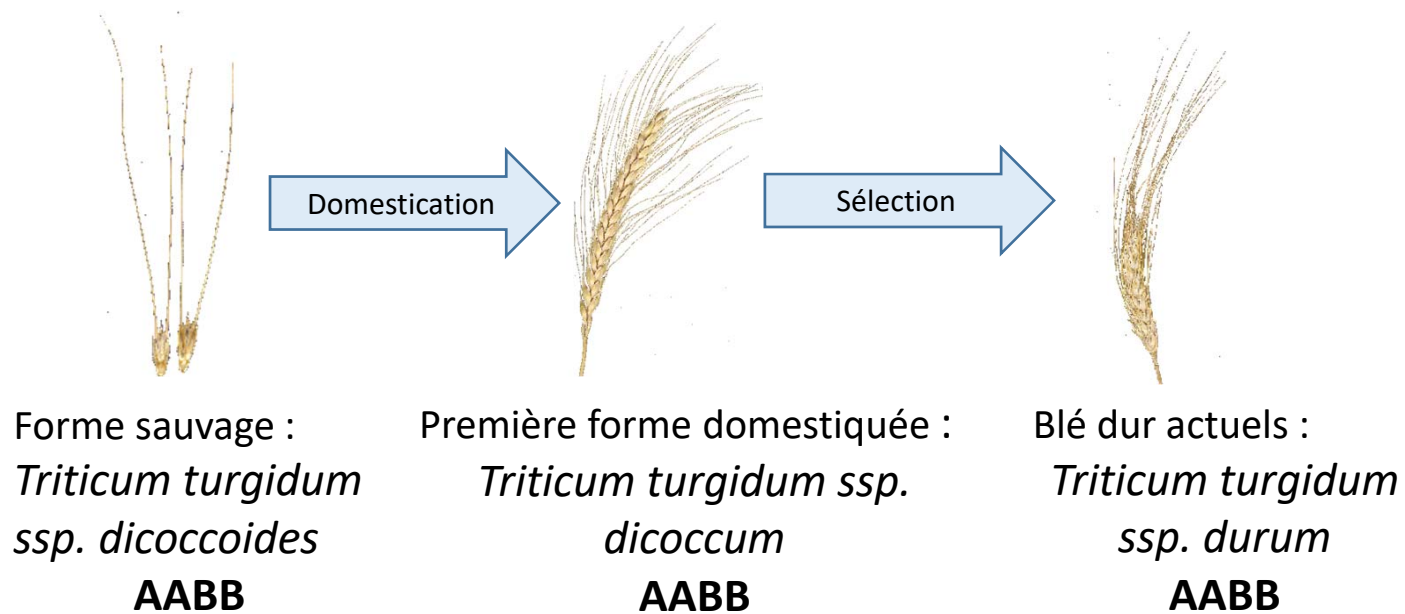
Pierre Roumet

Muriel Tavaud-Pirra



© Wikimedia Commons

Perte de diversité génétique au cours de l'histoire du blé dur



Thuillet, 2003
Haudry, 2008

Réintroduire de la diversité chez le blé dur élite

De manière classique : introgression ciblée par série de rétrocroisements

- ✓ Nécessite de connaître les caractères/ régions à introgresser
- ✓ Apporte des allèles favorables localement, sur ces régions ciblées

Idée : **Introduire massivement** de la diversité par l'utilisation de populations de prébreeding

- ✓ Apporter une large diversité non ciblée, sans connaissance *a priori*
- ✓ Favoriser les recombinaisons
- ✓ Sélectionner pour augmenter progressivement la valeur agronomique de la population



Constitution d'une population de *T. turgidum* à base génétique large

Constitution de la population EPO : Evolutionary Prebreeding Population

1. Apporter de la diversité génétique



Core collection de *T. turgidum*
(plusieurs sous-espèces)

2. Favoriser le brassage génétique



Pool initial : population de *T. t. ssp durum*
dans laquelle ségrège un gène de stérilité

Identification de la stérilité male dans la population

Plante hermaphrodite



Glumes fermées



Stigmate en cours de disparition

Autofécondation (97%)

Plante male stérile



Glumes entrouvertes



Stigmate plumeux

Allofécondation



Gestion du taux d'allofécondation

Constitution de la population EPO :

1. Apporter de la diversité génétique



Des individus de plusieurs sous-espèces (*T. turgidum*)

2. Favoriser le brassage génétique



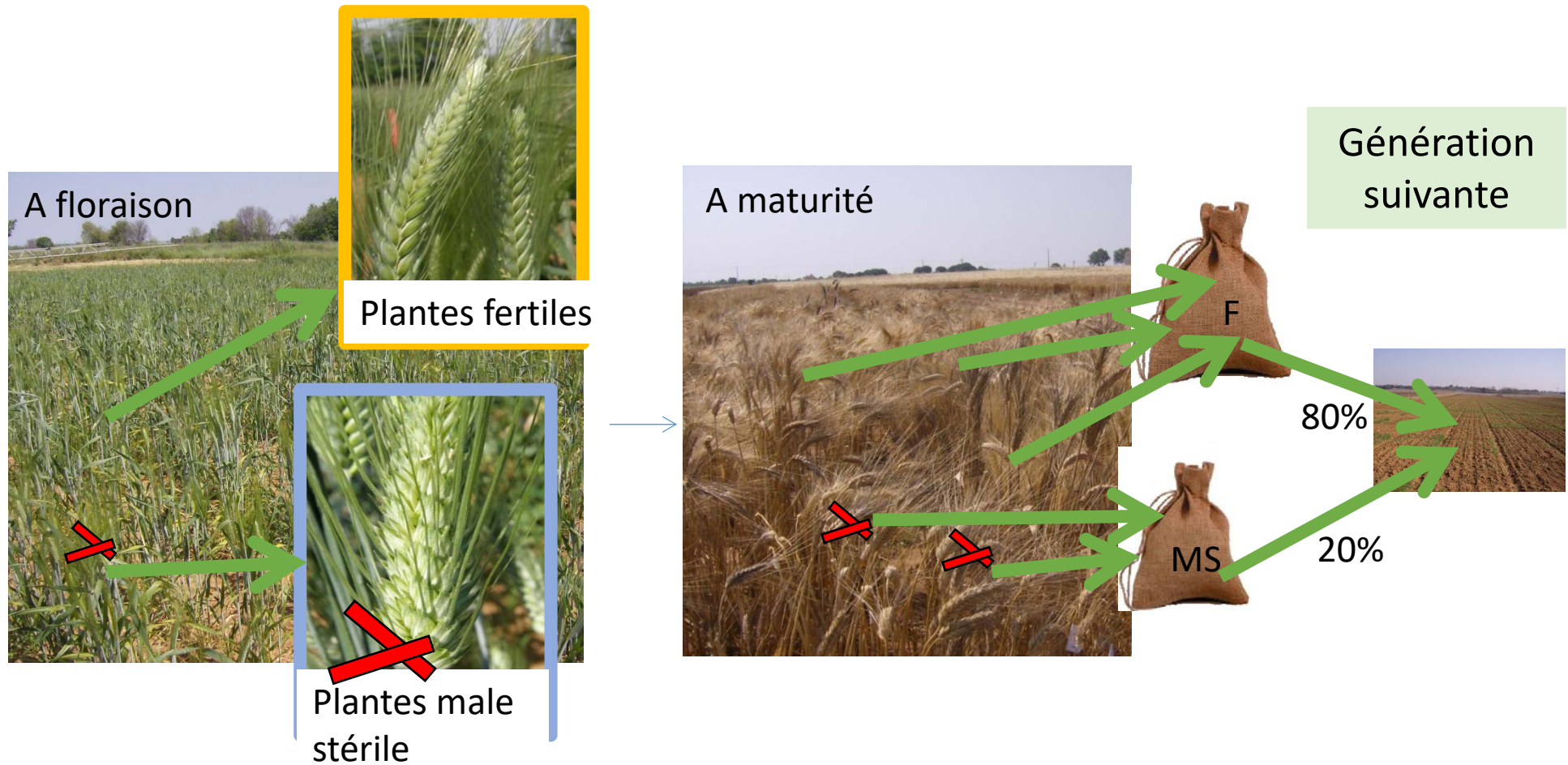
Pool initial : population composite dans laquelle ségrége un gène de stérilité

2 années de « pré-brassage »

Population EPO
(1997)

Population initiale

La stérilité male : un outil de gestion du taux d'allogamie

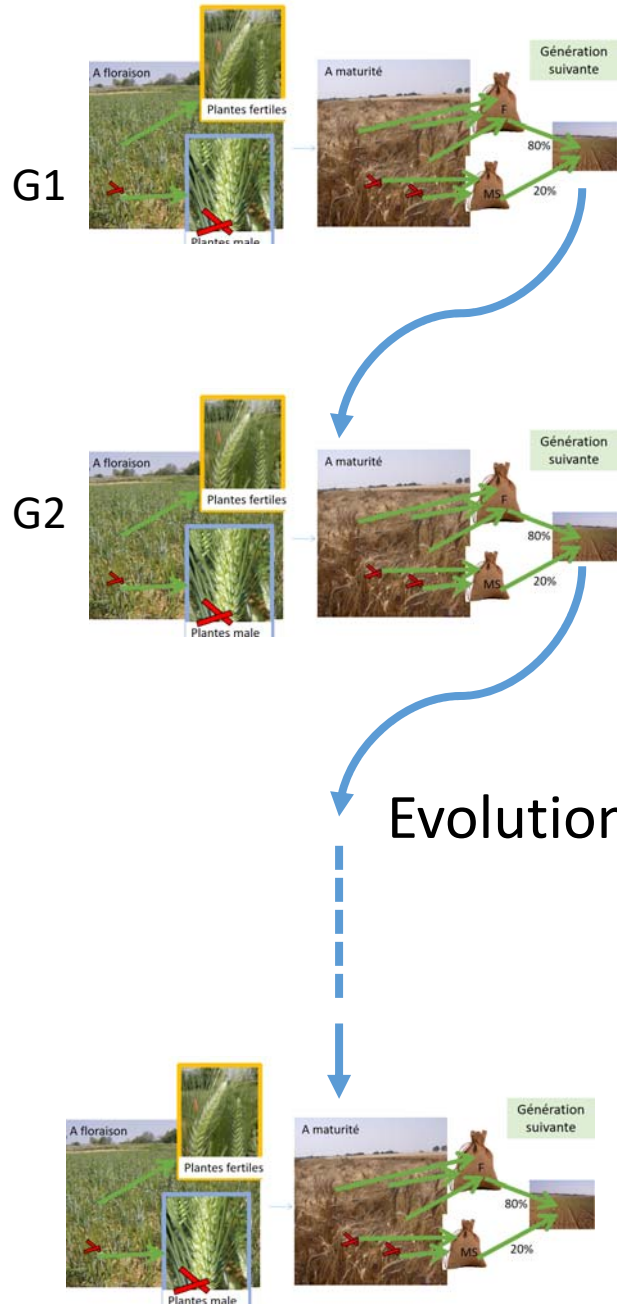


1. Identification des plantes stériles

2. Récolte des graines dans des sacs différents

3. Semis d'un mélange de grains issus des plantes F et MS

Evolution de la population de pré-breeding



Sélection pour un idéotype « cultivable »
dans les systèmes de production actuels

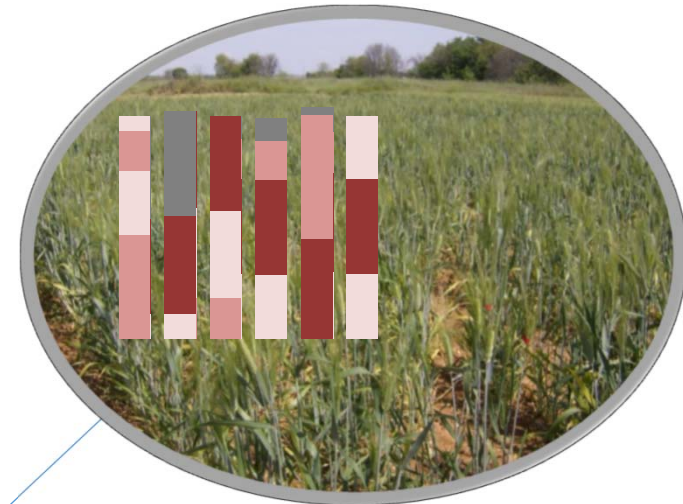
Contre sélection des plantes trop sensibles
ou très peu productives

Evolution durant 15 générations

Création de combinaisons alléliques nouvelles



Sous espèces sauvages et primitives : nombreuses recombinaisons
Diversité importante



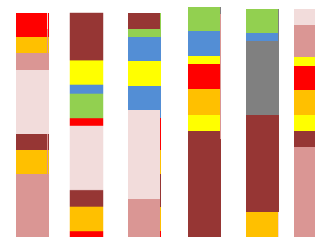
Population de *T. turgidum* ssp *durum* :
Diversité plus faible

G0

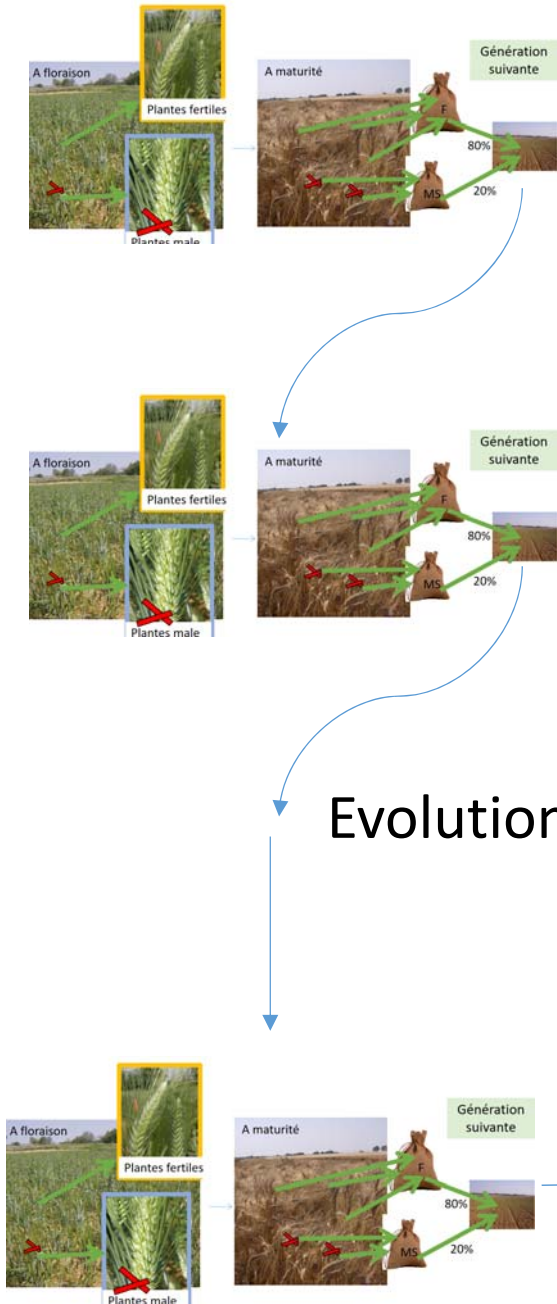
Recombinaisons efficaces favorisées par l'allofécondation

Structure complexe du DL
Combinaisons alléliques nouvelles

G15



Evolution de la population de pré-breeding



Sélection pour un idéotype
« cultivable » dans les
systèmes classiques

Contre sélection des plantes
trop sensibles ou très peu
productives

Evolution durant 15 générations

En 2010 (génération 15) :

Echantillonnage et fixation de 450
lignées par SSD : **lignées EPO**

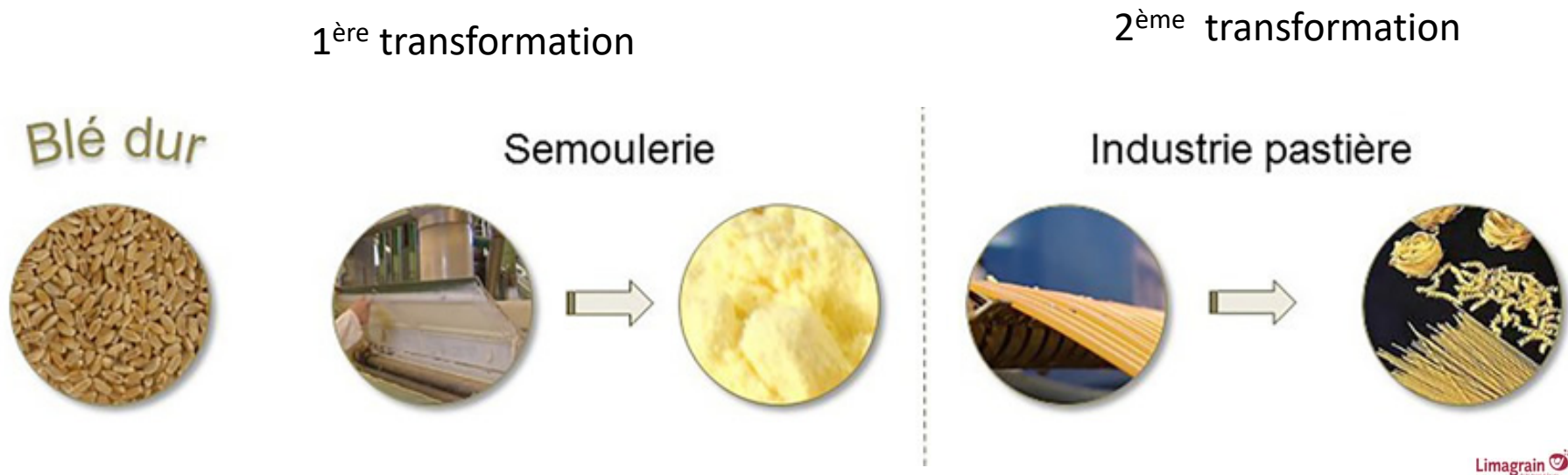
Question générale :

Quel est l'intérêt de cette population de prébreeding pour introduire de la diversité génétique notamment aux QTLs d'intérêt?



Etude des régions impliquées dans le déterminisme génétique du rendement semoulier

Importance du rendement semoulier



Rendement semoulier (SY) : proportion de semoule obtenue à partir d'un lot de grains

$$SY = \text{Poids de la semoule extraite} / \text{poids des grains avant mouture}$$

Dépend de nombreux facteurs impliqués dans :

- ✓ La composition de l'endosperme : notamment teneur en protéine -> mitadin
- ✓ L'adhérence enveloppe/amande
- ✓ Le rapport amande/enveloppe : importance de la forme du grain -> PMG, morphologie des grains

Caractère difficile à mesurer directement (lots de graines > 1kg)

Objectifs de l'étude:

- Définir une méthode d'estimation indirecte du rendement semoulier
- Estimer la variation et l'héritabilité du rendement semoulier et des caractères corrélés dans les lignées EPO
- Etudier le déterminisme génétique du rendement semoulier
- Comparer au déterminisme génétique des caractères corrélés
- Explorer la diversité génétique dans la(les) zone(s) d'intérêt

Matériel végétal et phénotypage

181 lignées EPO fixées issues de la population EPO choisies parmi les 450 lignées, en ayant éliminé les lignées soeurs

Expérimentations à la station expérimentale de Mauguio en 2016 et 2017
Dispositif en 2 blocs complets randomisés



Phénotypage :

- Mesures manuelles : PMG, précocité d'épiaison
- Spectroscopie proche infra-rouge des grains
- Mesures de la morphométrie des grains avec l'optomachine

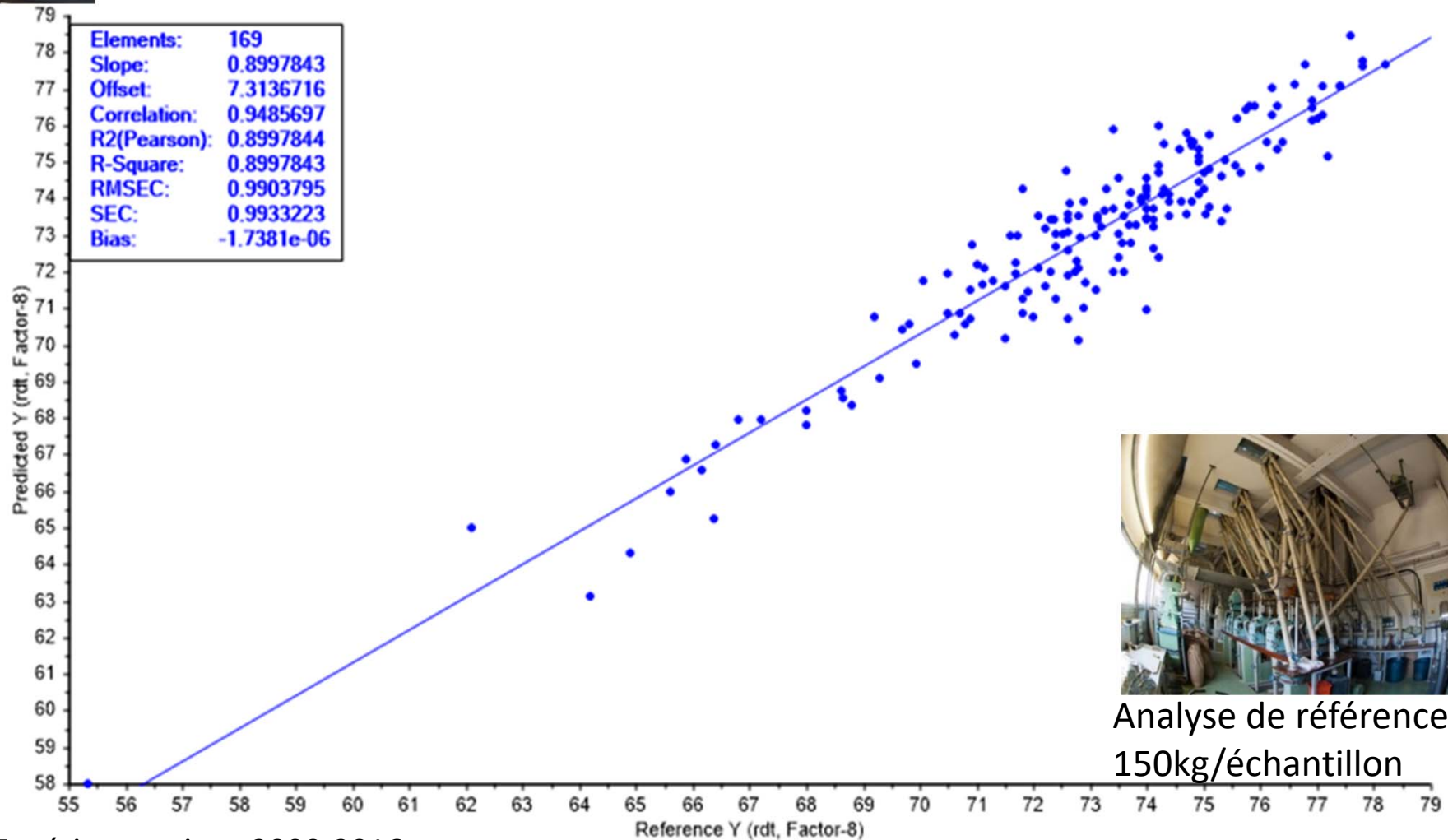


Evaluation du rendement semoulier par la spectroscopie (Nirs)



Analyse Nirs :
200g/échantillon

Equation de calibration de bonne qualité : $R^2=0.9$
Prédiction des valeurs de rendement semoulier par
spectroscopie proche infra-rouge



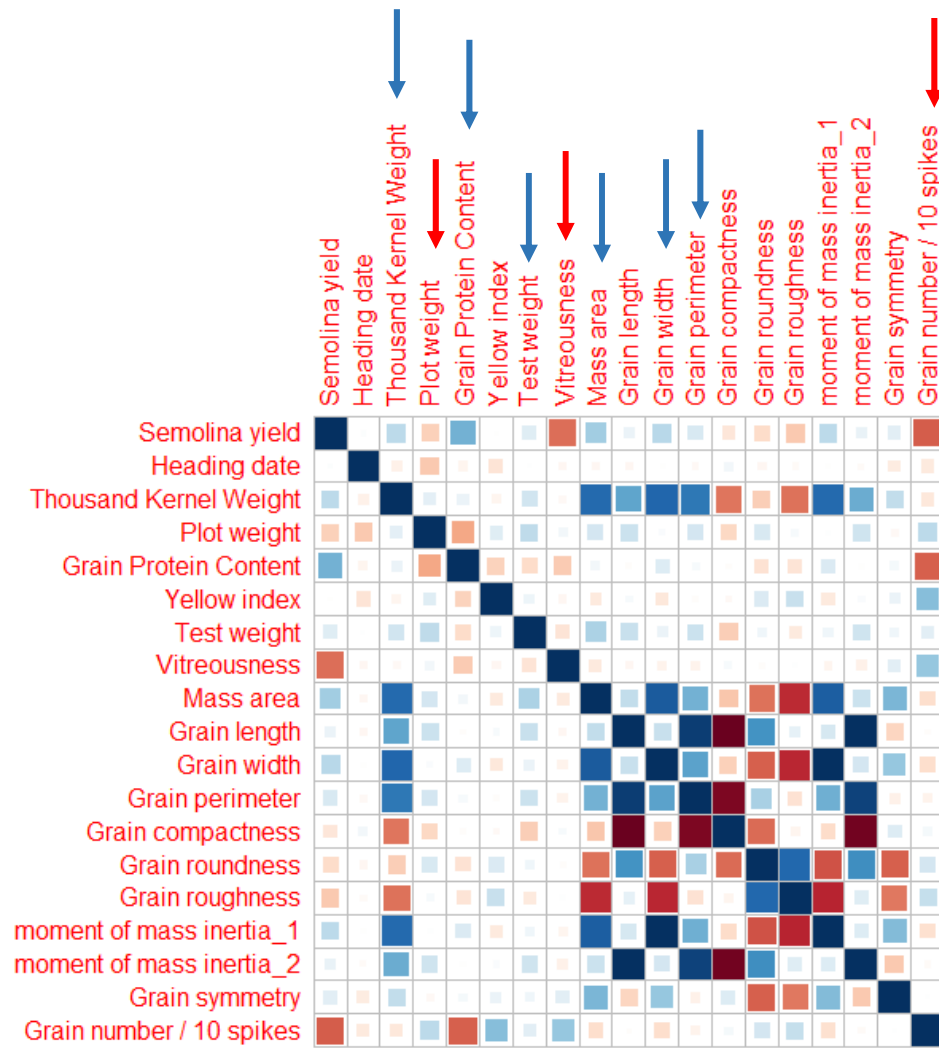
Analyse de référence:
150kg/échantillon

Variation phénotypique des caractères mesurés

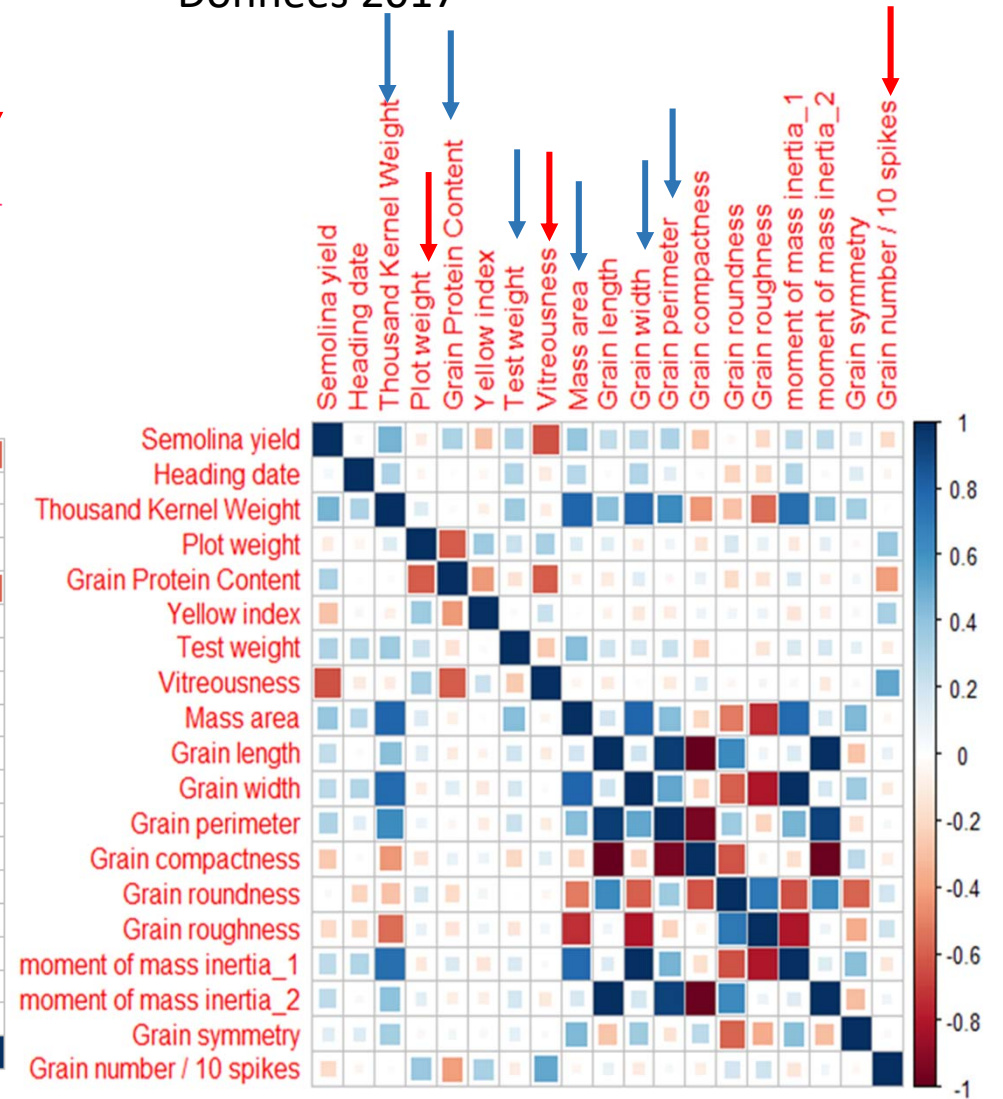
Variable	2016			2017		
	Moyenne	Variance	Héritabilité	Moyenne	Variance	Héritabilité
Rendement semoulier	72.32	9.31	0.89	68.86	2.93	0.89
Nombre de grains/10 epis	120.98	55.47	0.63	122.97	30.38	0.92
Poids de 1000 grains (PMG)	47.99	27.45	0.81	50.34	39.45	0.88
Rendement de la parcelle	2990.50	375564.94	NA	3018.10	313841.15	0.56
Poids spécifique	84.10	4.73	0.38	79.72	4.06	0.66
Masse surfacique	2.55E-03	2.61E-08	0.89	0.00	3.64E-08	0.78
Teneur en protéines des grains	12.26	1.55	0.65	13.03	1.44	0.65
Date de précocité	114.94	8.82	NA	110.02	10.01	0.96
Indice de jaune	47.76	8.97	NA	41.20	6.35	0.67
Mitadin	0.86	15.47	NA	50.15	351.78	0.67
Circularité des grains	1.32	2.40E-03	0.91	1.34	2.66E-03	0.92
Compacité des grains	0.01	4.52E-07	0.91	0.01	4.34E-07	0.92
Largeur des grains	3.40	0.02	0.91	3.50	0.03	0.79
Longueur des grains	7.77	0.15	0.94	8.27	0.18	0.93
Périmètre des grains	21.87	0.83	0.94	23.08	0.96	0.87
Rugosité des grains	0.28	8.33E-05	0.83	0.28	7.57E-05	0.75
Symétrie des grains	0.90	7.47E-05	0.93	0.90	8.52E-05	0.91
Moment d'inertie d'ordre 1	69.69	39.97	0.90	72.79	53.57	0.77
Moment d'inertie d'ordre 2	346.52	1292.43	0.93	388.77	1599.20	0.93

Des corrélations phénotypiques cohérentes pour les 2 années

Données 2016



Données 2017





Données génotypiques : puce 420 K, BreedWheat

En collaboration dans le cadre du méta-programme Selgen CropDL

- 420 K SNPs définis sur le blé tendre dont 280 K sont publics.
- Filtres pour la qualité du génotypage : 168 725 SNPs
- 134 566 polymorphes sur les lignées EPOs
- $MAF > 0.05$

Positionnés par blast sur la carte physique de Zavitan

- 97 477 SNPs pour analyse 2017, 97 507 SNPs pour 2016 (en fonction des données manquantes sur les phénotypes)

Génétique d'association

Association génotype/phénotype en prenant en compte l'apparentement entre lignées

Utilisation du logiciel GEMMA, en déclarant le modèle suivant :

$$Y = \mu + X\beta + Zu + e$$

avec Y = vecteur des BLUEs pour chaque année (estimés après correction spatiale avec le package Spat),

μ = moyenne générale du caractère étudié pour l'ensemble du panel

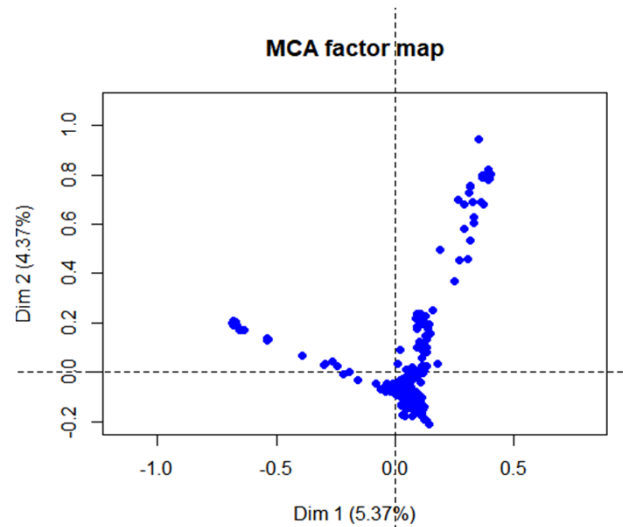
β = effets de substitution des marqueurs

u = effet de l'apparentement des lignées EPO

Z et X sont des matrices de design

e = variance résiduelle

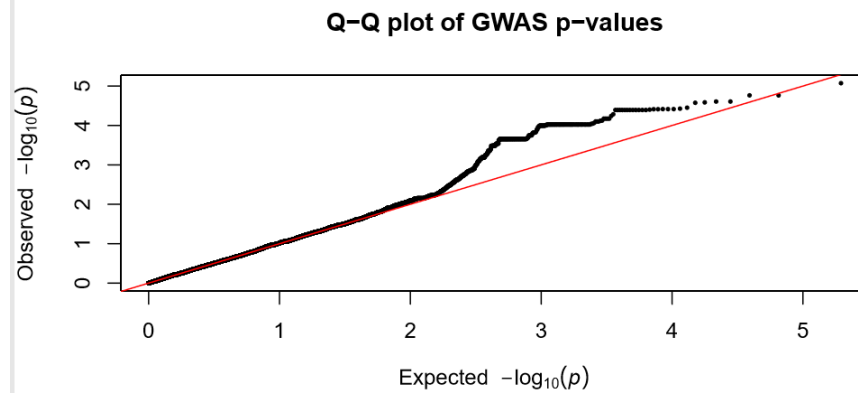
Justification du modèle statistique utilisé:



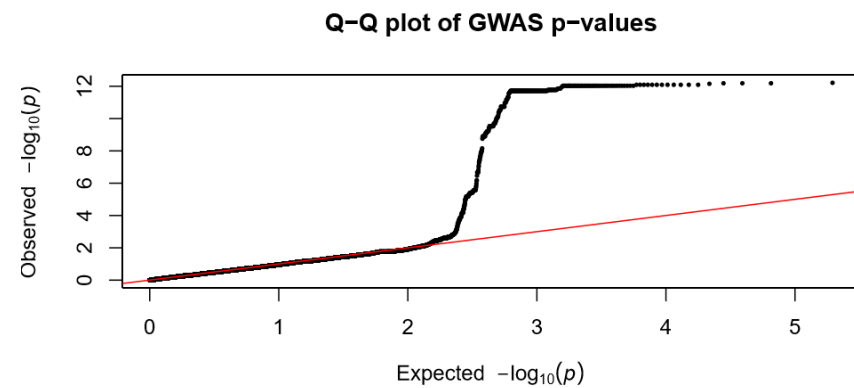
Légère structure observée

Prise en compte de l'apparement

Qqplot : peu de faux positifs observés



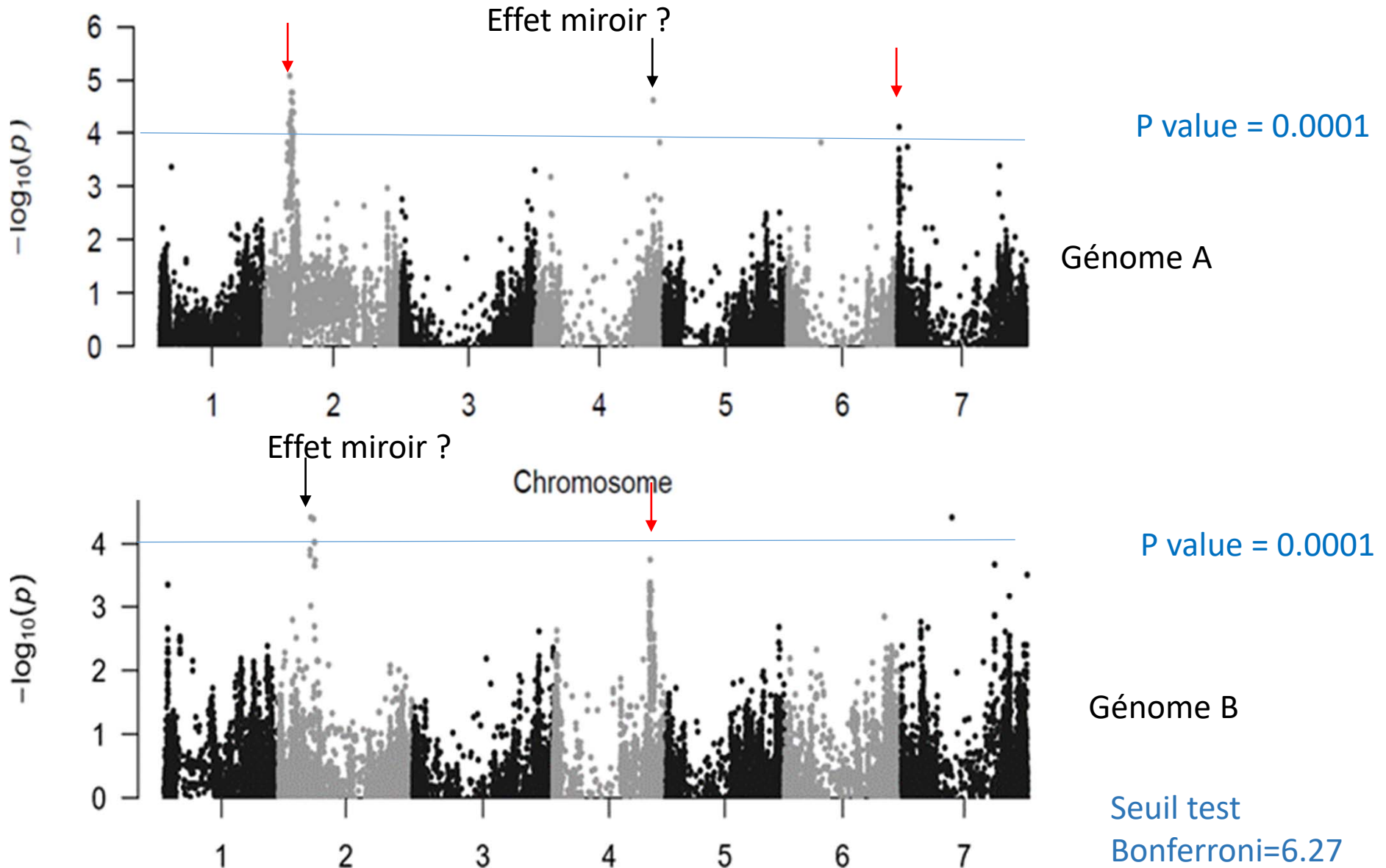
Rendement semoulier



Poids de 1000 grains

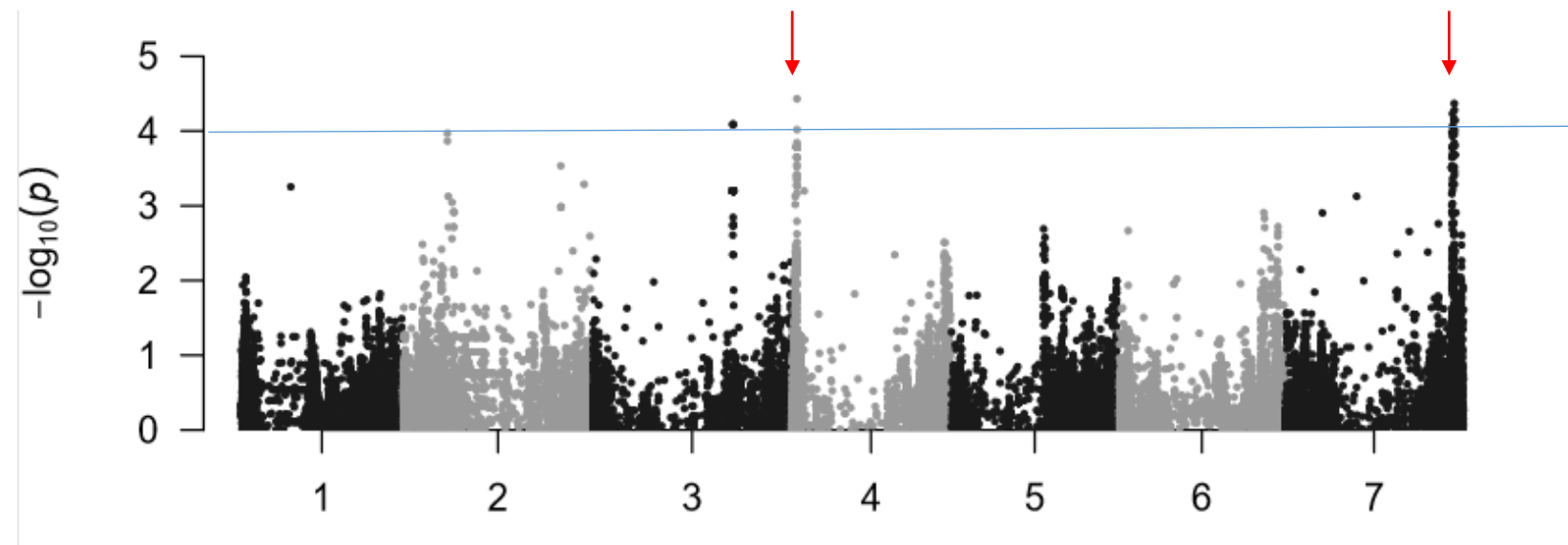
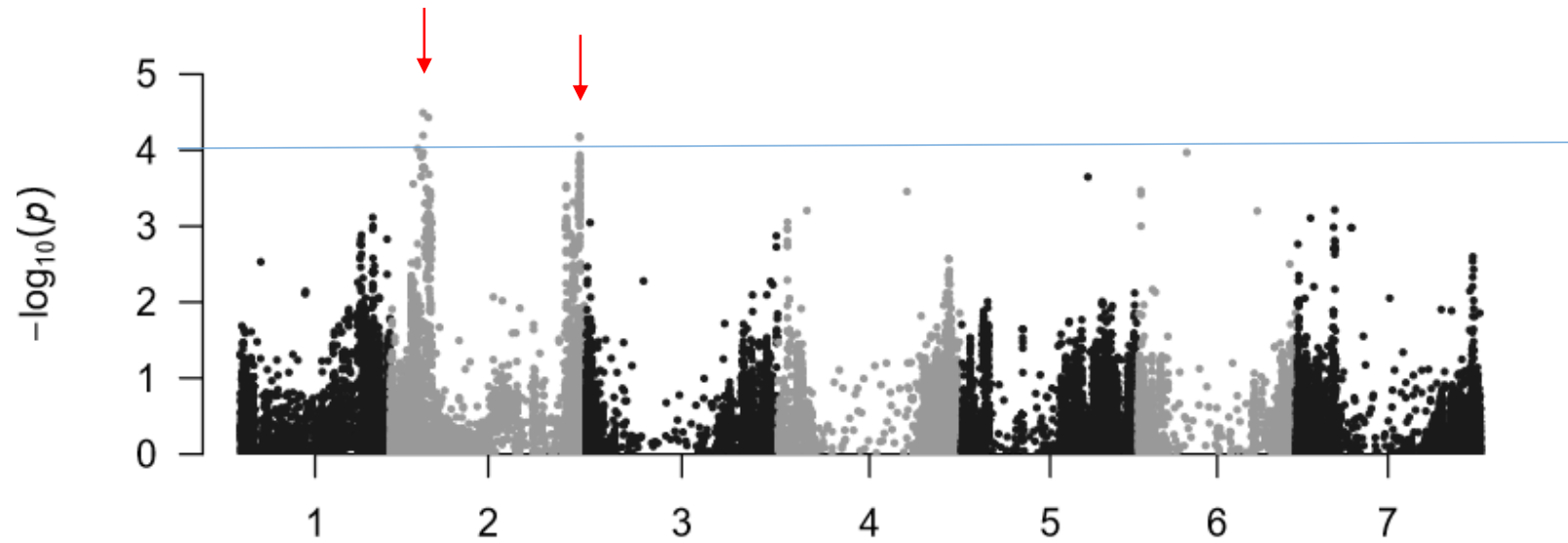
Déterminisme génétique du rendement semoulier

Données 2017



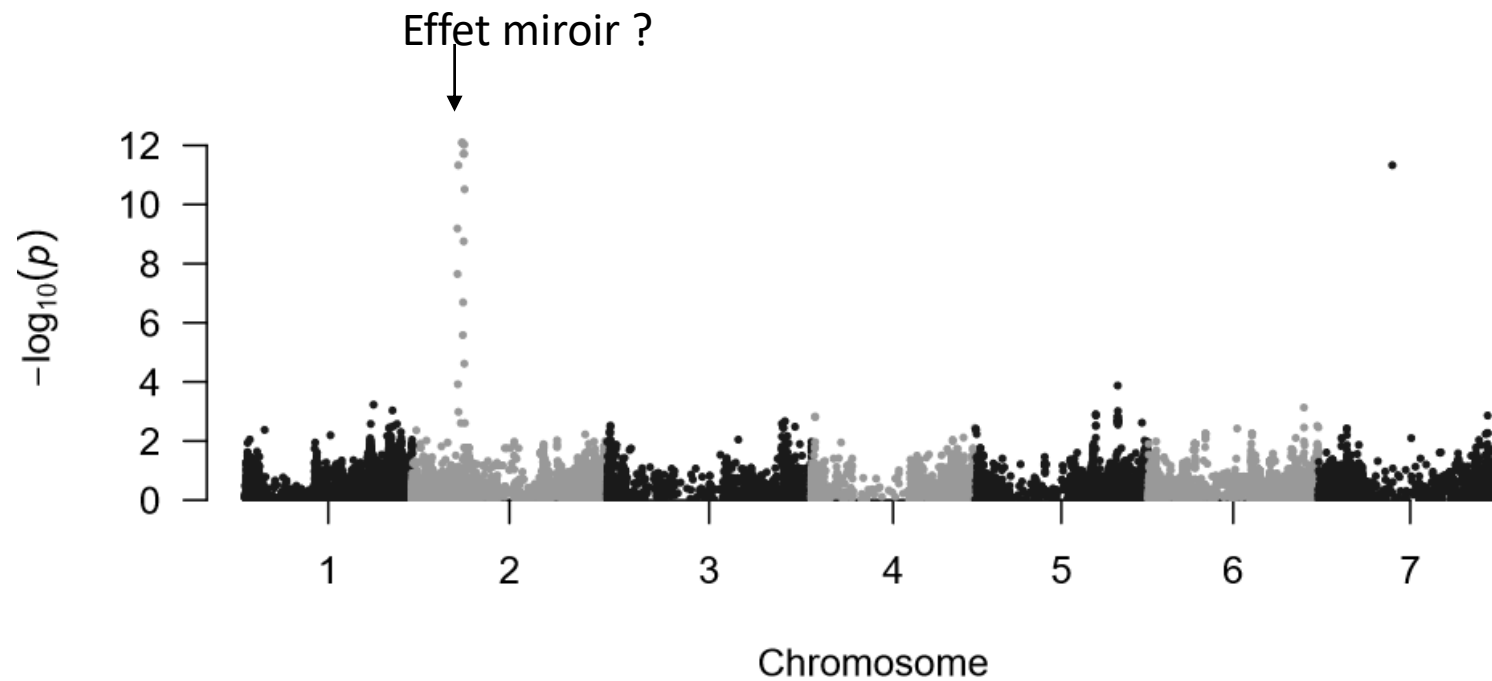
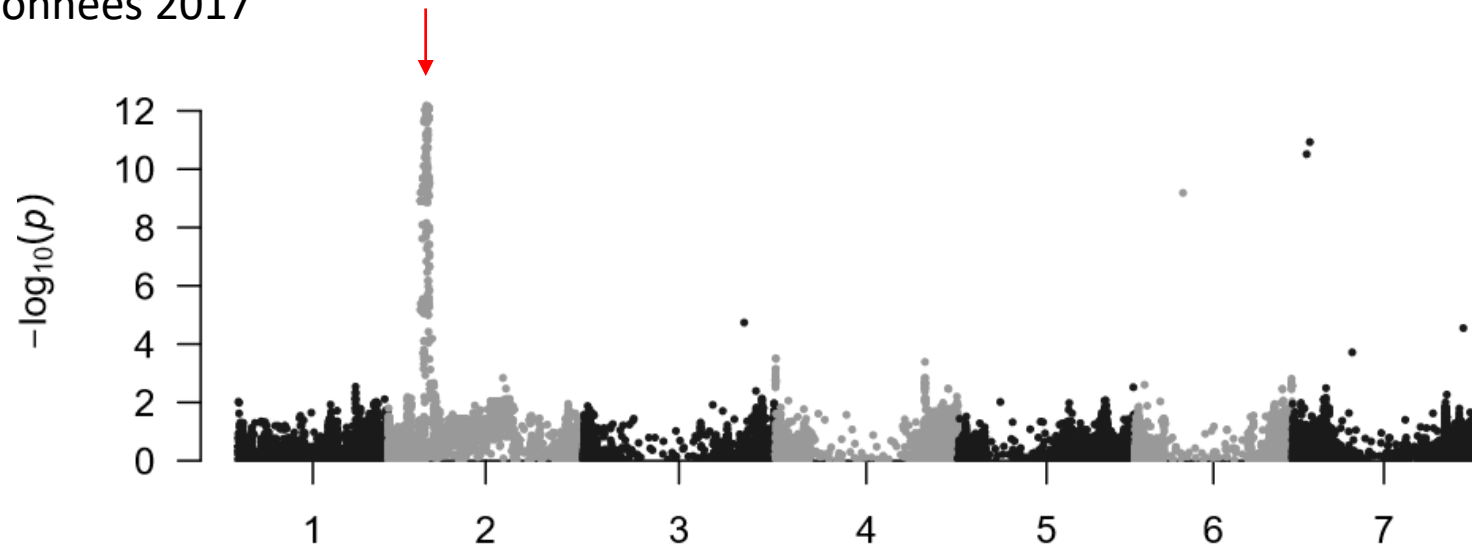
Déterminisme génétique du rendement semoulier

Données 2016



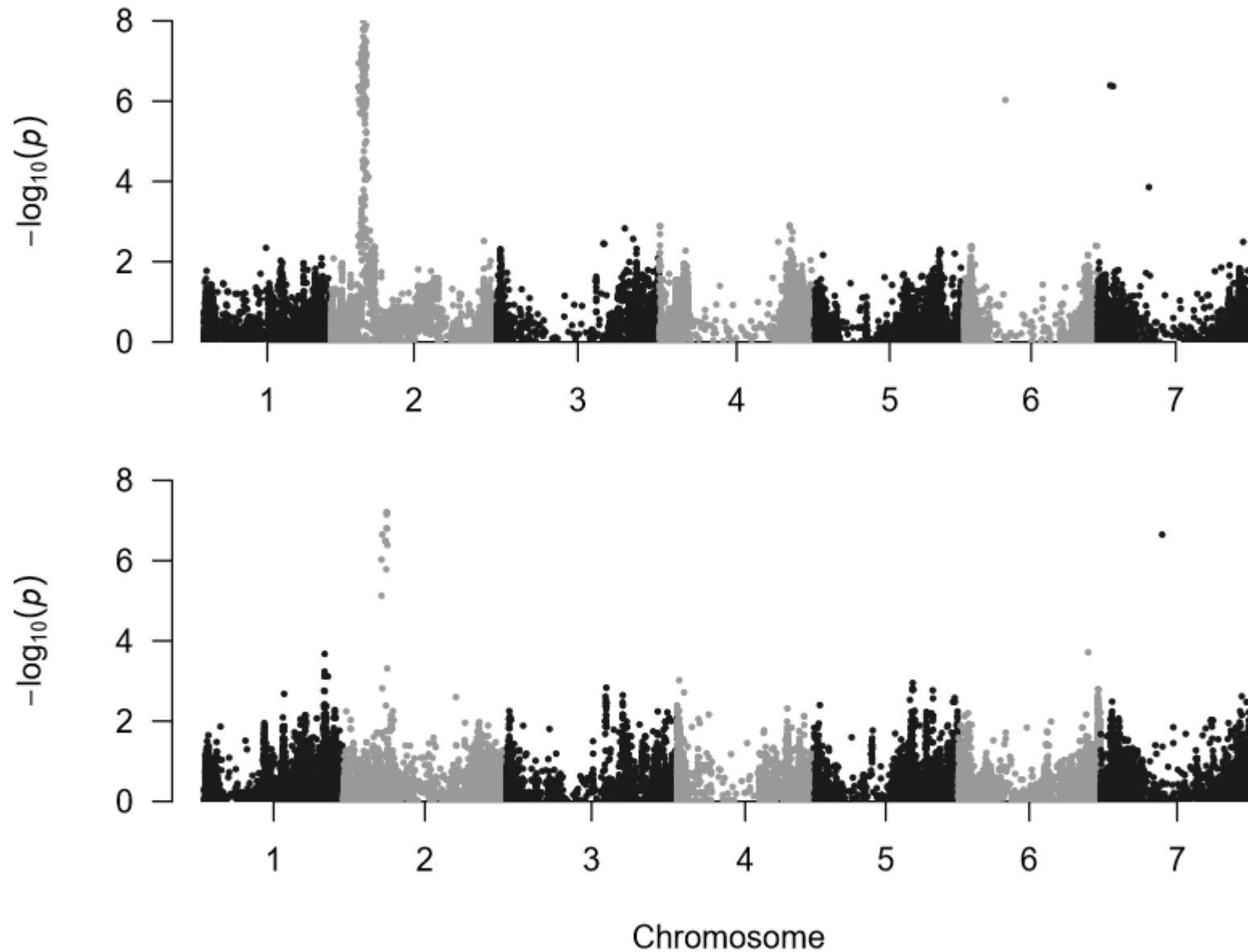
Déterminisme génétique du poids de 1000 grains (PMG)

Données 2017



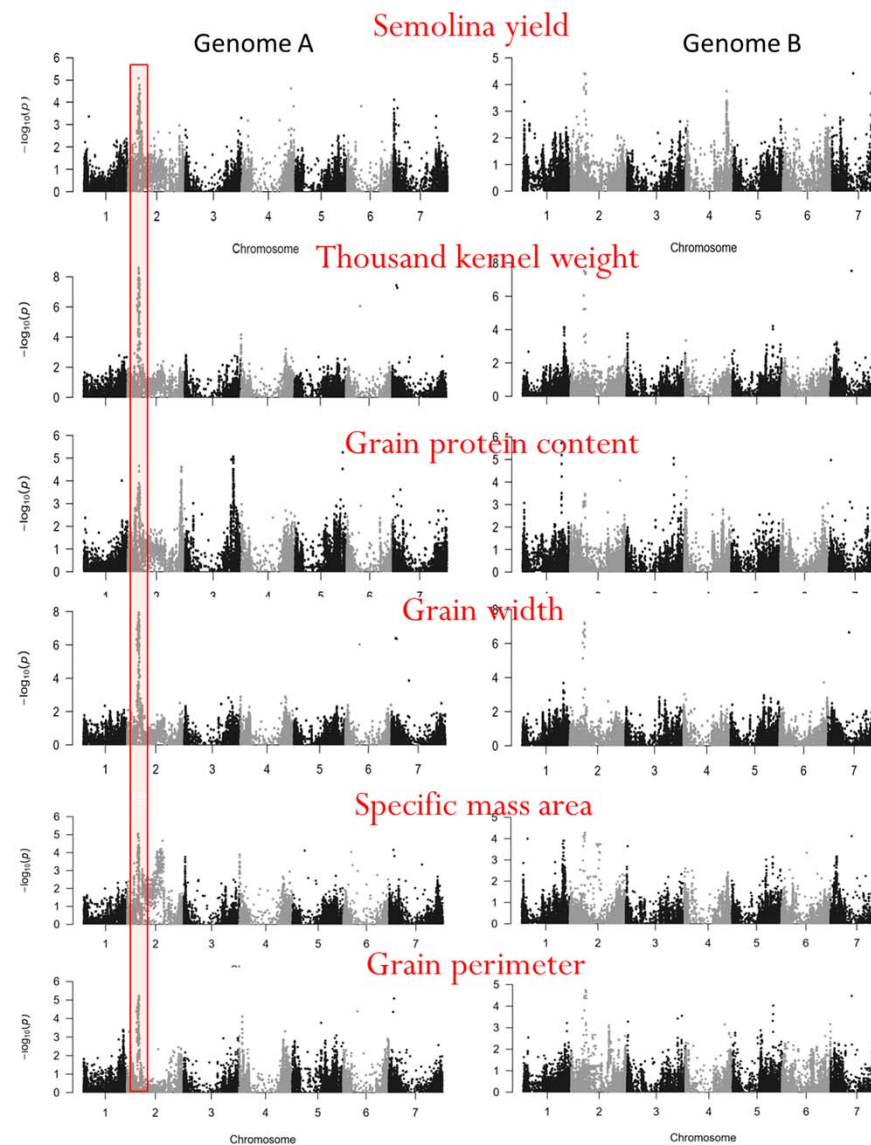
Déterminisme génétique de la largeur du grain

Données 2017



Déterminisme génétique : rendement semoulier et caractères corrélés

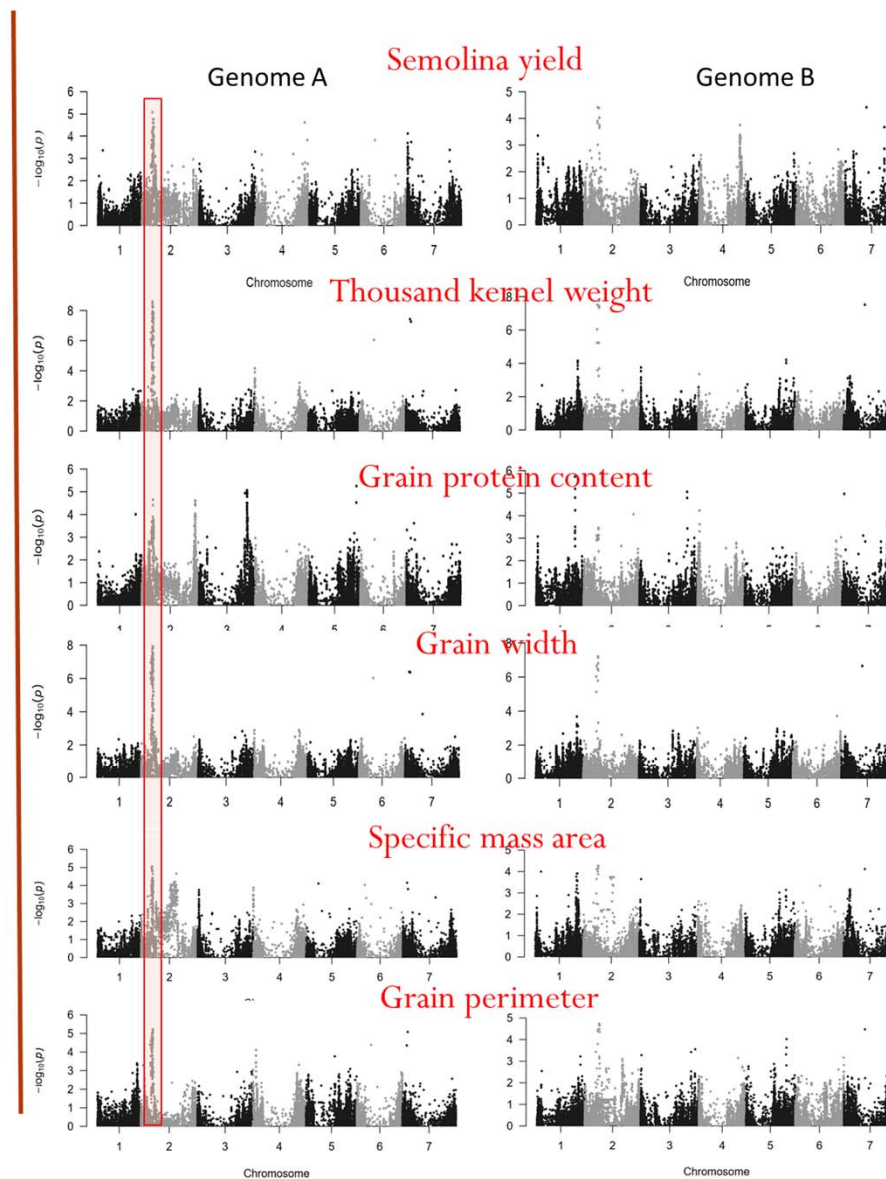
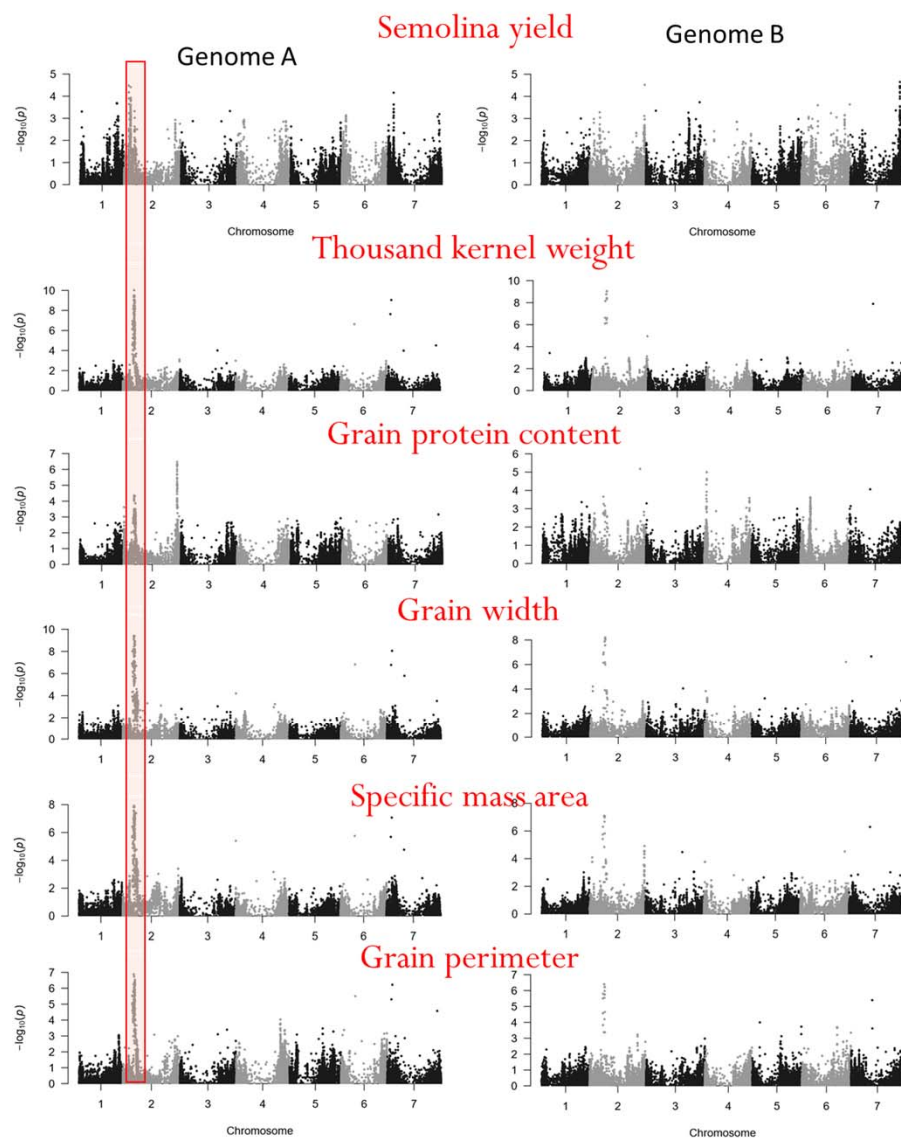
2017



Déterminisme génétique : rendement semoulier et caractères corrélés

2016

2017



Déterminisme génétique du rendement semoulier:

- ✓ Plusieurs QTLs identifiés : 2A, 7A et 4B (2017), 2A, 4B et 7B (2016)
- ✓ Un QTL sur le chromosome 2A observé en 2016 et 2017
- ✓ Un QTL est retrouvé très significativement dans cette région sur les caractères corrélés : PMG, largeur de grains, périmètre des grains, masse surfacique, teneur en protéines

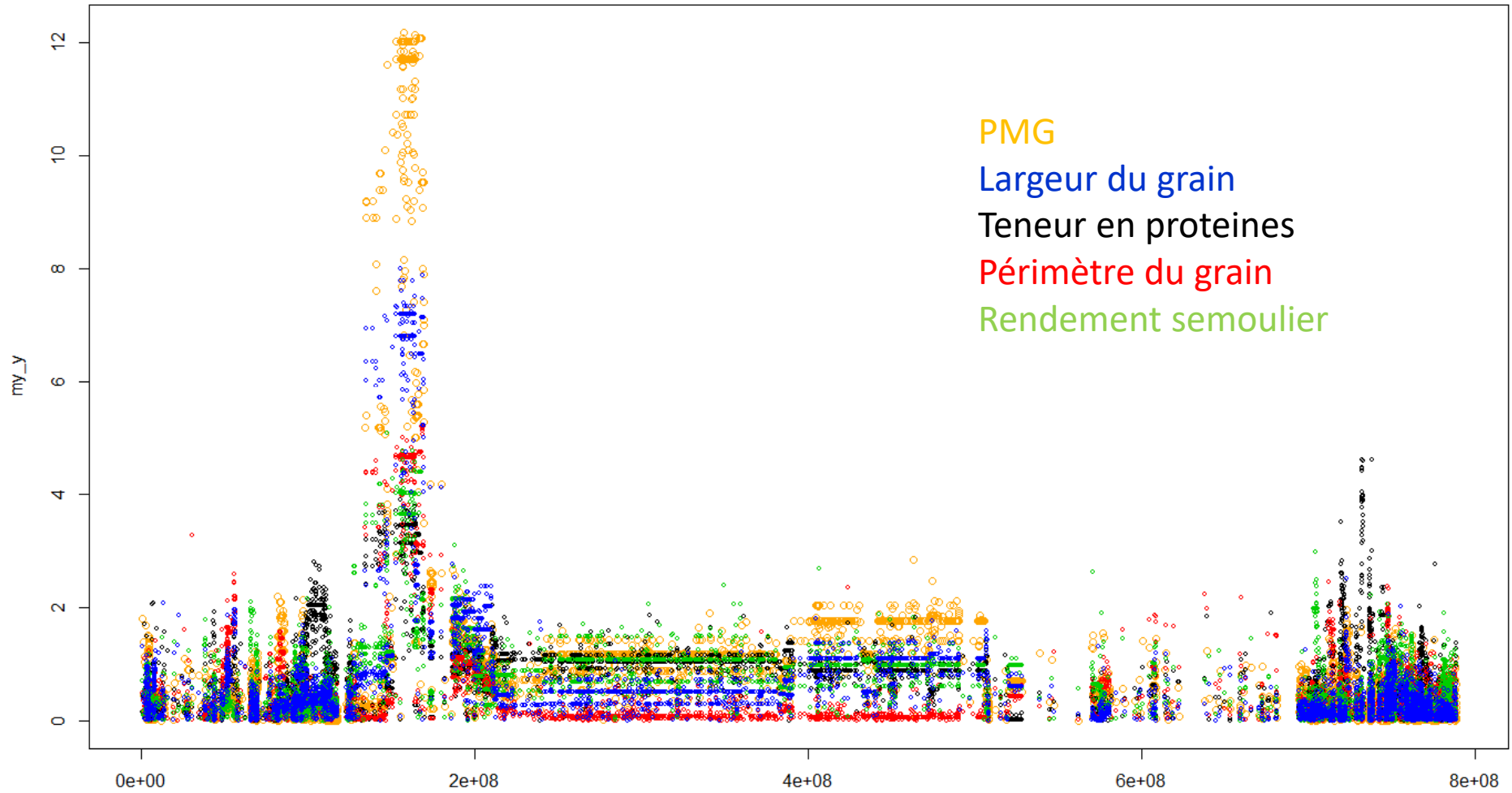


Zoom sur la region du chromosome 2A

- Est ce le même QTL pour les différents caractères ?
- Quelle diversité allélique existe dans cette region ?

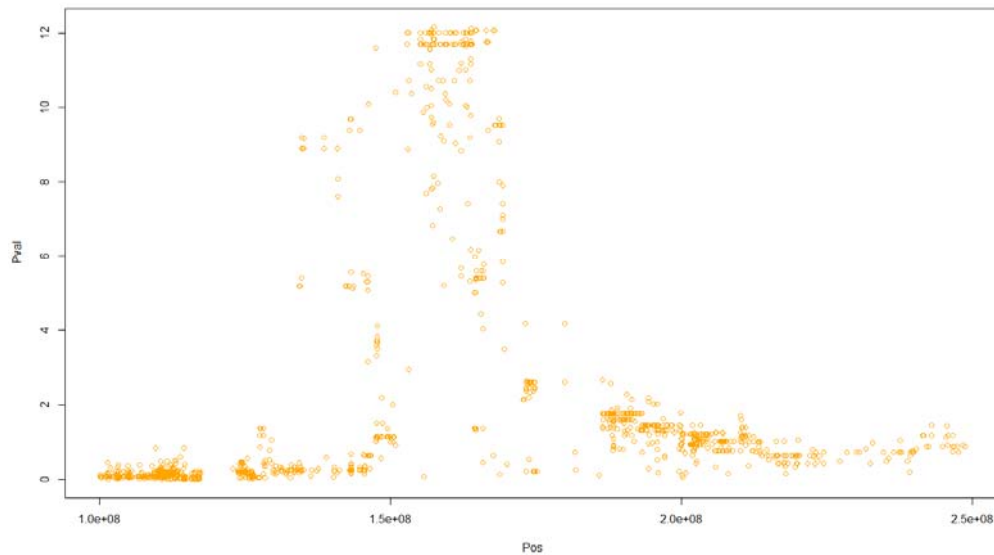
Manhattan plot sur le chromosome 2A

$-\log(\text{Pvalue})$



Position sur le chromosome 2A

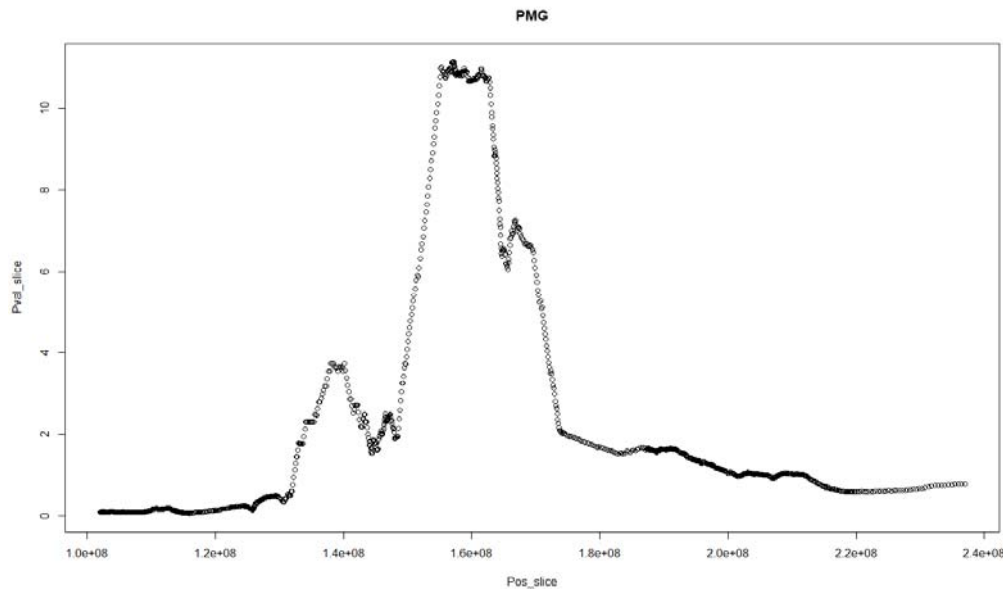
Manhattan plot sur la région d'intérêt : 100Mb-250 Mb



Poids de 1000 grains

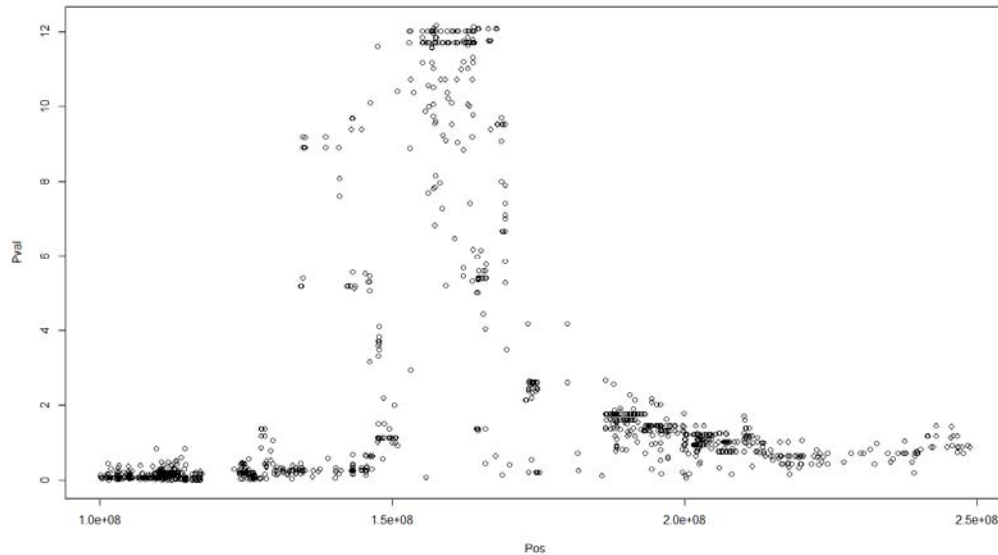
3043 marqueurs

Valeurs observées



Fenetre glissante pour
50 marqueurs

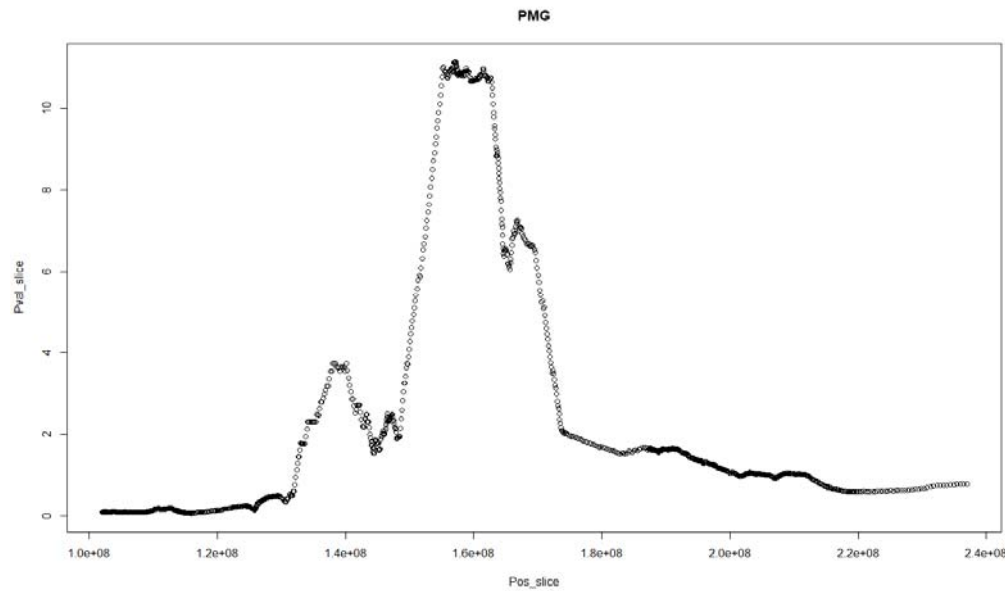
Manhattan plot sur la région d'intérêt : 100Mb-250 Mb



Poids de 1000 grains

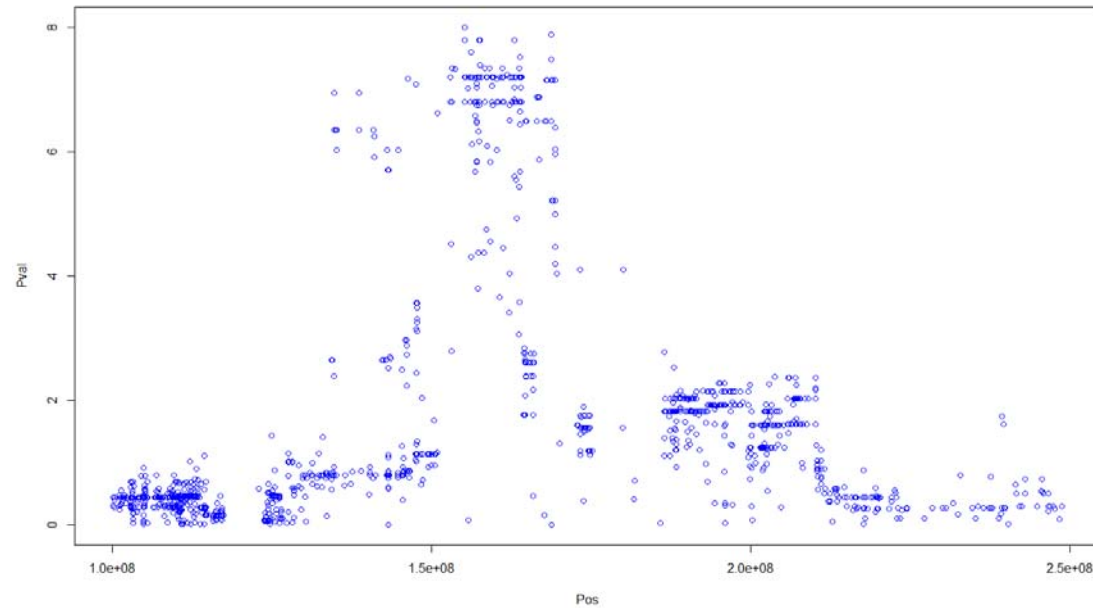
3043 marqueurs

Pvalues observées



Fenetre glissante pour
50 marqueurs

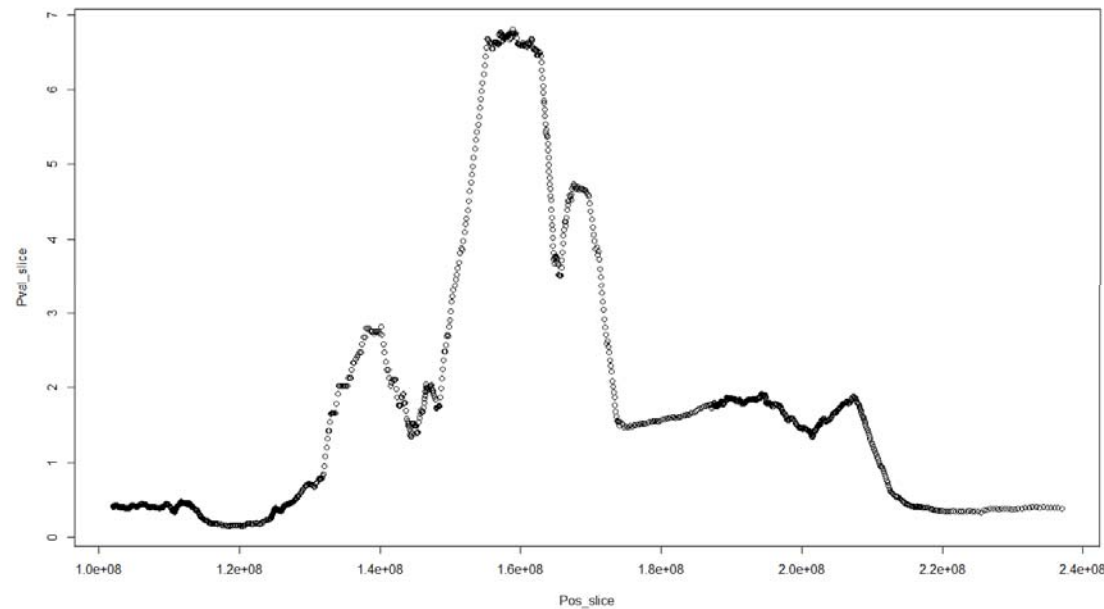
Manhattan plot sur la région d'intérêt : 100Mb-250 Mb



Largeur des grains

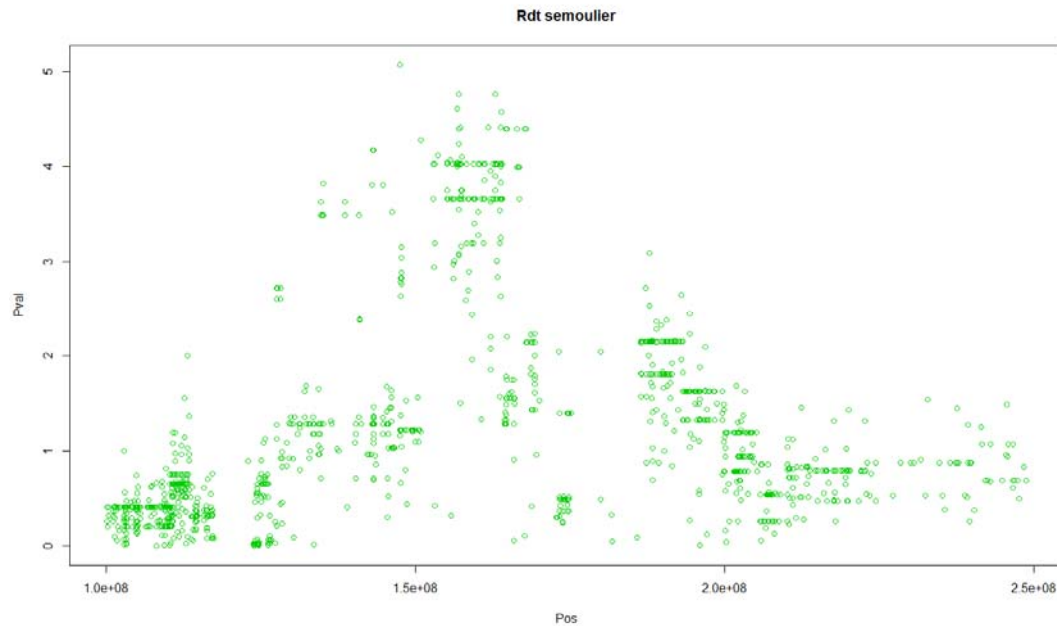
3043 marqueurs

Valeurs observées



Fenetre glissante pour
50 marqueurs

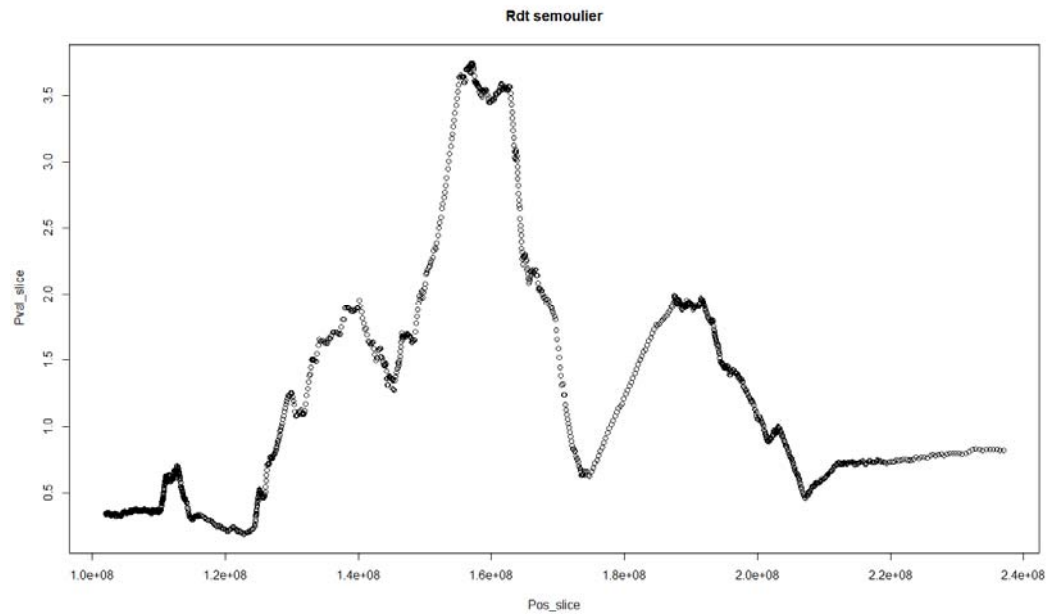
Manhattan plot sur la région d'intérêt : 100Mb-250 Mb



Rendement semoulier

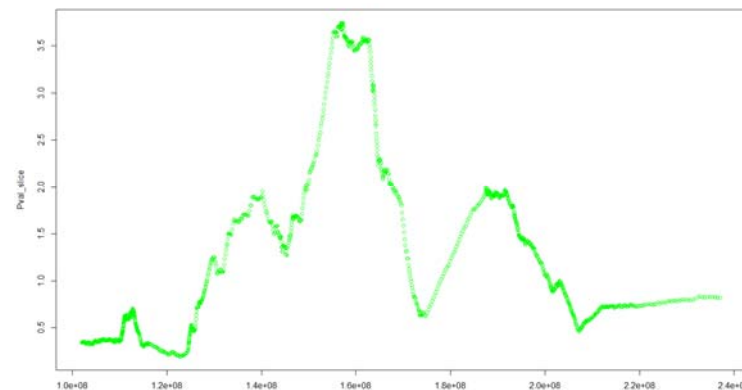
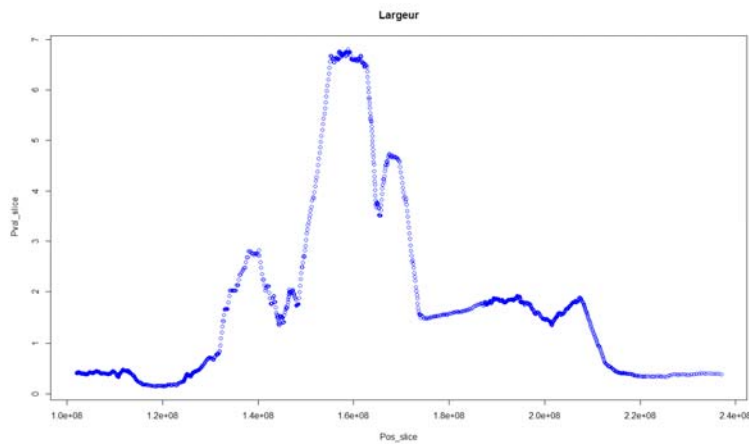
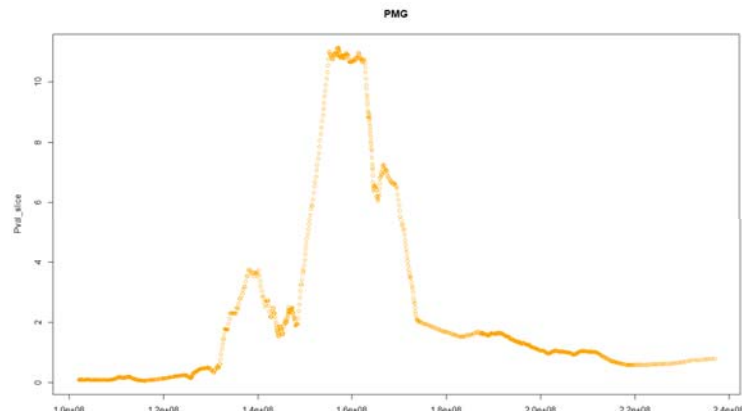
3043 marqueurs

Valeurs observées



Fenetre glissante pour
50 marqueurs

Manhattan plot sur la région d'intérêt : 100Mb-250 Mb



➤ Une même région QTL pour les caractères corrélés

➤ Paramètres de ce QTL:

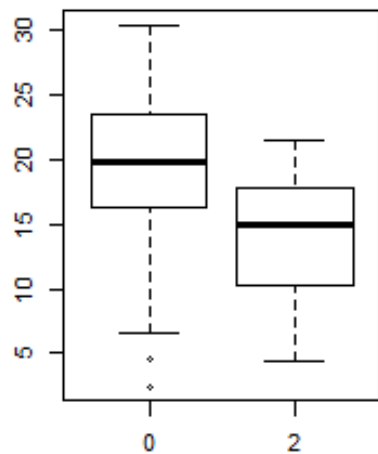
Caractère	Position pic	Effet de substitution	R ²
PMG	157508711	-4.04	0.18
Largueur du grain	155151476	0.10	0.15
Rendement semoulier	147452132	-0.76	0.09

➤ Diversité allélique de cette region ?

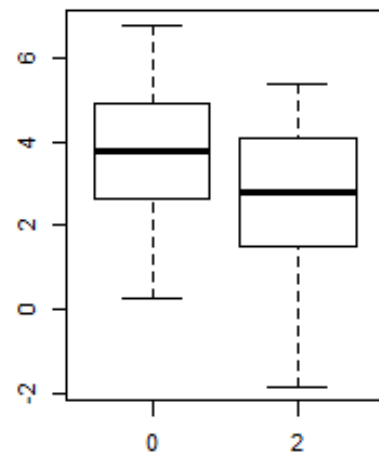
Variations phénotypiques au pic du QTL de PMG

Marqueur
cfn1520694

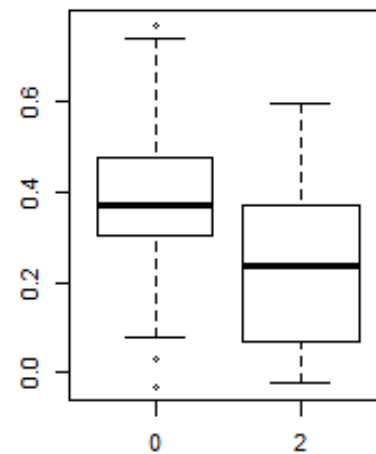
Poids 1000 grains



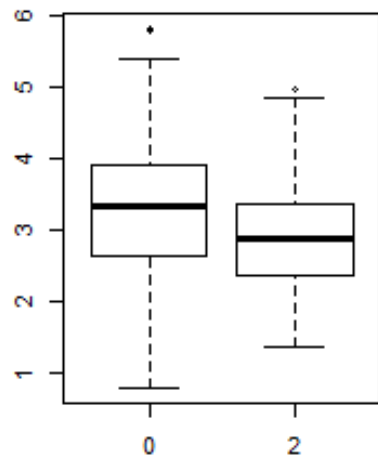
Rendement semoulier



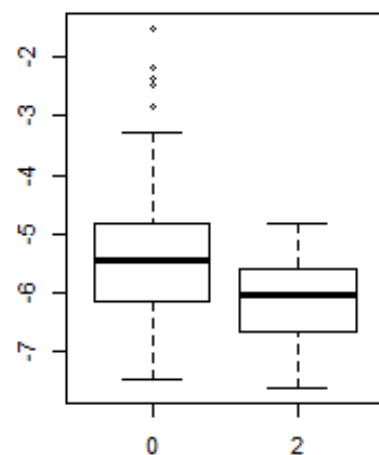
Largeur du grain



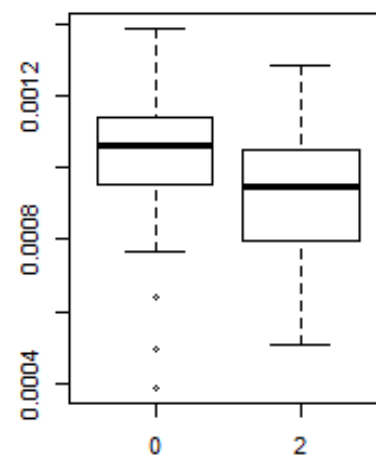
Périmètre du grain



Teneur en protéines



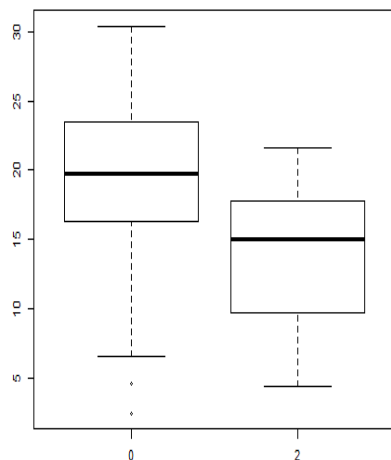
Masse surfacique



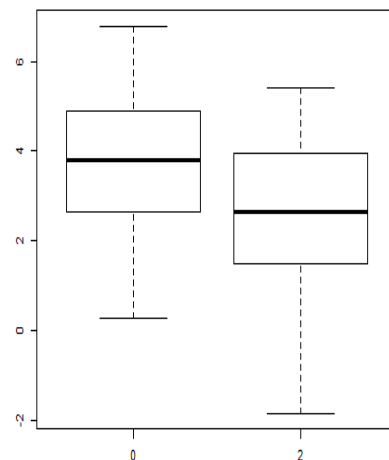
Variations phénotypiques au pic du QTL de Rendement semoulier

Marqueur
cfn0591489

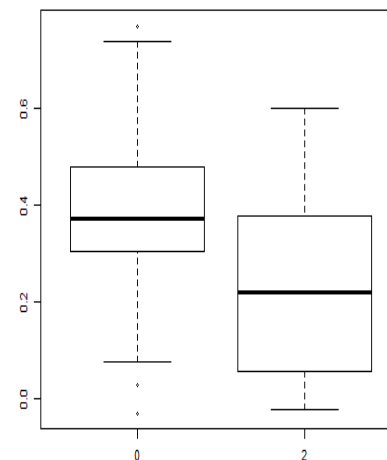
Poids 1000 grains



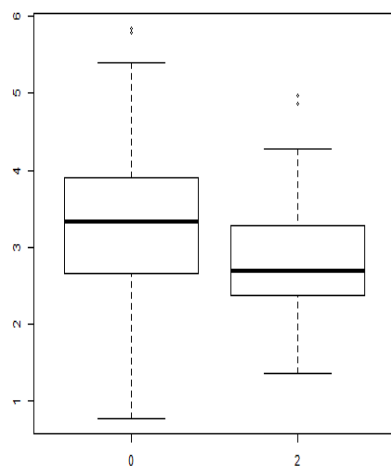
Rendement semoulier



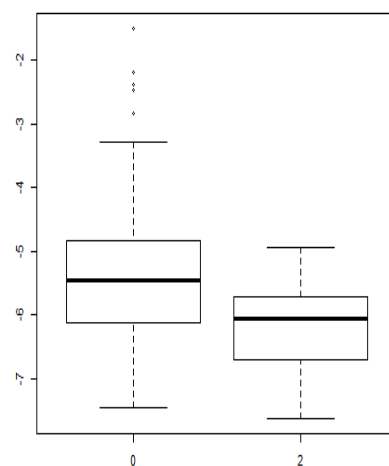
Largeur du grain



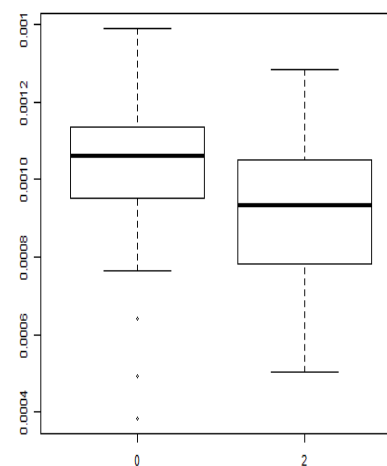
Périmètre du grain



Teneur en protéines



Masse surfacique



Allèles observés pour les marqueurs aux pics, dans les différents groupes de lignées analysées

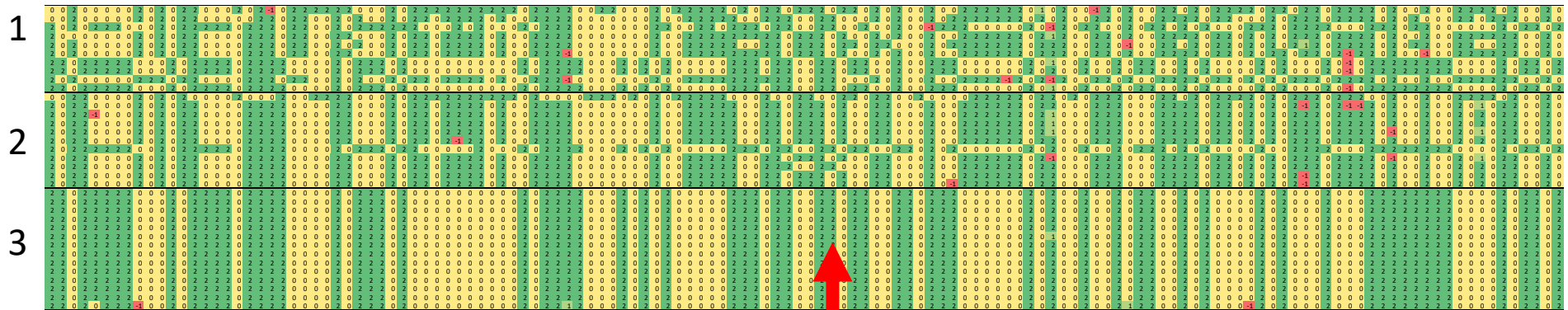
Groupes	Effectifs	Allèles + ↓		Allèles + ↓	
		Cfn1520694 A1A1	Cfn1520694 A2A2	Cfn0591489 A1A1	Cfn0591489 A2A2
<i>T. t. dicoccoides</i>	10	4	6	9	1
<i>T. t. dicoccum</i>	11	1	10	10	0
<i>T. t. durum</i>	14	14	0	14	0
Lignées inscrites	108	101	7	108	0
EPO	425	322	101	322	90

Dans les lignées EPO : **une diversité non présente** dans le blé dur

➤ Identification de QTLs non observés chez les dispositifs de blé dur

Haplotypes observés dans la zone cible (-log(Pval)>10)

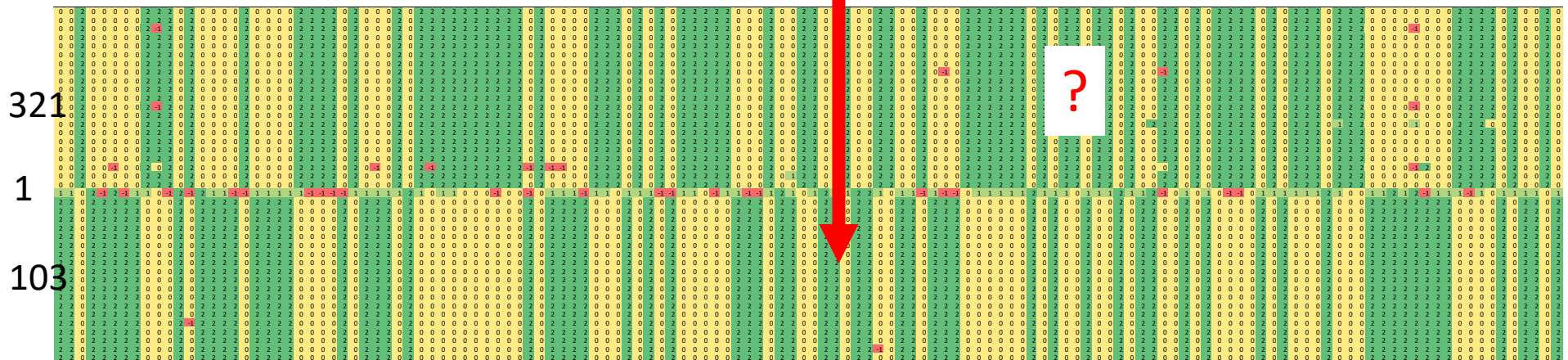
1. *T. t. dicoccoides*
2. *T. t. dicoccum*
3. *T. t. durum*



Lignées EPO

1 haplotype de type durum

Effectifs



DL très fort : r^2 moyen = 0.99 sur l'ensemble des 137 marqueurs (21 Mb)

Déterminisme génétique : rendement semoulier et caractères corrélés

Pertinence de l'estimation du rendement semoulier par Nirs

Nombreux caractères de morphométrie et de composition de la graine corrélés au rendement semoulier

Déterminisme du rendement semoulier :

- quelques régions impliquées dont la région proximale du 2B observée sur les 2 années

Pic conservé dans la région du chromosome 2A pour différents caractères

- QTL de PMG, de largeur du grain, de périmètre du grain, de masse surfacique, de teneur en protéines
- Allèles polymorphes chez *T. dicoccoides* et *T. dicoccum* mais **allèle favorable déjà fixé chez les *T. durum* élite**

Et maintenant ? Prochaines analyses et perspectives

- ✓ Vérifier par DL la position cartographique des marqueurs : effet miroir 2A et 2B
- ✓ Analyse plus fine de la région du 2A :
 - Réseaux d'haplotypes présents dans les lignées EPO, et du DL dans cette zone
 - analyse de la diversité de cette région dans les lignées inscrites au catalogue
 - Etude des gènes annotés présents dans cette région chez Zavitan
- ✓ Etude pluriannuelle : 2016 et 2017