

Identification de loci candidats associés à la qualité du fruit chez la tomate par une approche de génétique d'association multilocus

Christopher Sauvage
INRA, UR1052, Avignon

Colloque EPGV – 23 au 25 Juin 2014, Evry, France



Les déterminants génétiques de la qualité

Qu'est ce que la qualité ?

- Conservation (gène *Rin*)
- Aspect (forme, taille, ...)
- Organoleptique (Sucres, Vitamine C, ...)

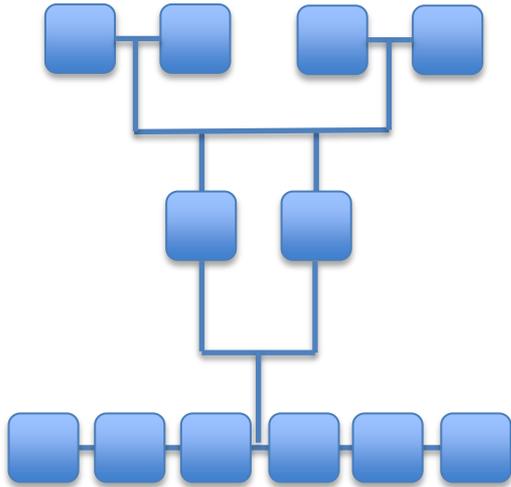
Quelle est l'architecture génétique des caractères ?

- Nombre de gènes
 - Mode(s) d'action
 - Héritabilité
 - Marqueurs liés
-
- Quelles sont les approches adaptées ?

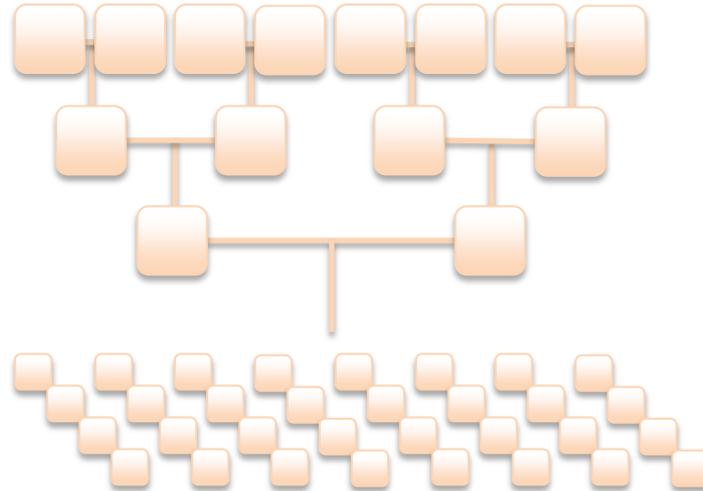


Quelles sont les approches ?

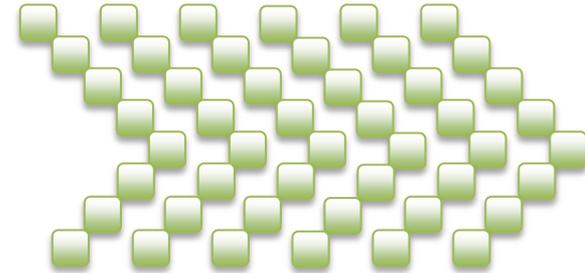
Détection de QTL(s)



Populations MAGIC / NAM



Génétique d'association



Avantages

- Méthodologie rodée
- Diversité des croisements

Inconvénients

- Polymorphisme limité (F_0)
- Intervalle de confiance

- Approche hybride
- Réduction du DL

- Mise en place des croisements
- Difficultés analytiques

- Polymorphisme ancestral
- Modèles adaptés

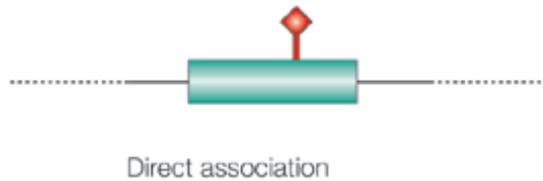
- G x E non considérées
- Structuration de population, DL

GWAS : Intérêt et principes

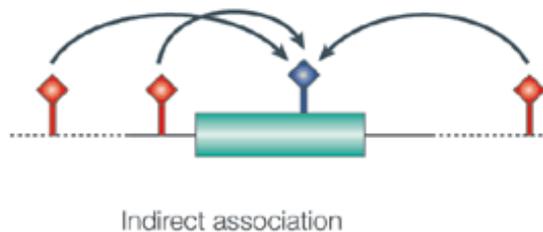
Objectif:

trouver des liens statistiques entre phénotype(s) et génotype

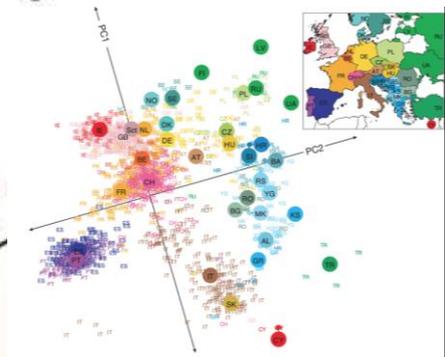
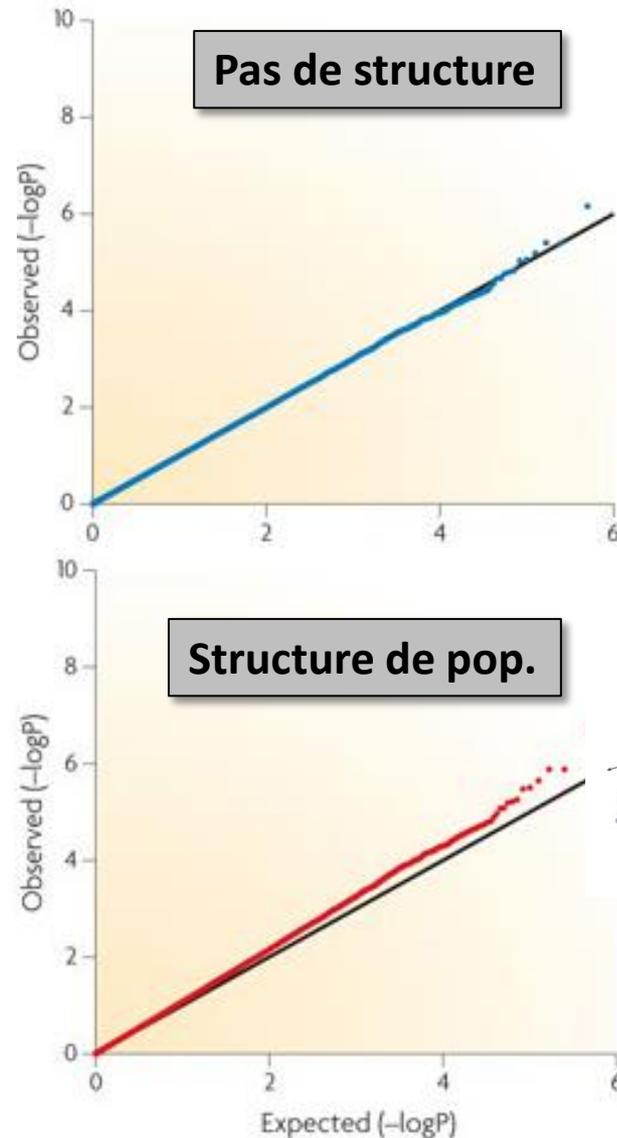
a



b



Nature Reviews | Genetics



Approche méthodologique (3)

MLM

vs

MLMM

$$y = X\beta + Zu + e$$

(Yu et al. 2002)

- Str. de population (Q)
- Parenté (K)
- Génotype (X)
- Phénotype (y)

Implémenté dans Tassel
Rapidité d'exécution

Effet du DL

Surévalue le nombre réel
d'associations

Correction pour tests
multiples

$$y = X\beta + Zu + e$$

(Segura et al. 2012)

- Str. de population (Q)
- Parenté (K)
- Génotype (X)
- Phénotype (y)

Basé sur Emma-X

Rapidité d'exécution

Méthodes de corrections
pour tests multiples

Approche pas à pas (step)

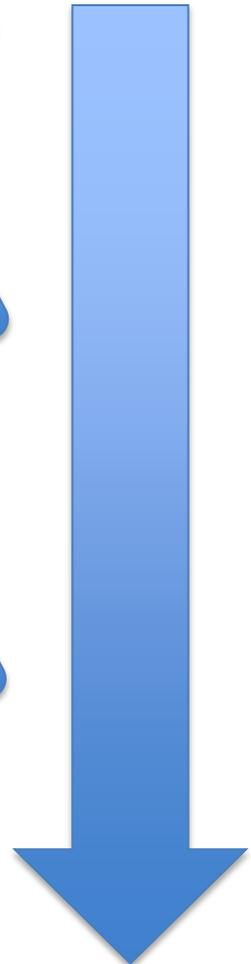
Modélise le DL

Estime h^2_{missing}

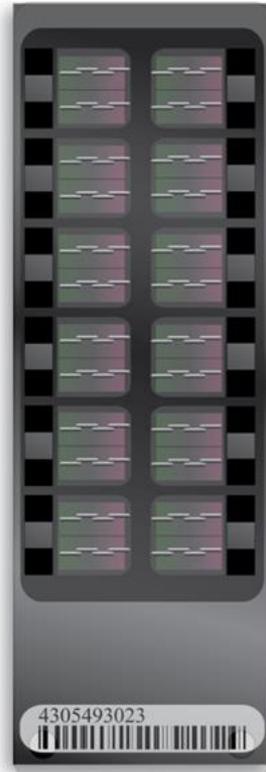
Script R/Python

De précédents travaux de GWA

- Constitution d'une core collection :
20 SSR | 360 Acc : Ranc et al., 2008
- 'GWA' sur le chromosome 2 (FW):
353 SNPs | 90Acc : Ranc et al., 2012
- GWA (8 phénotypes)
192 SNPs | 188 Acc : Xu et al., 2012



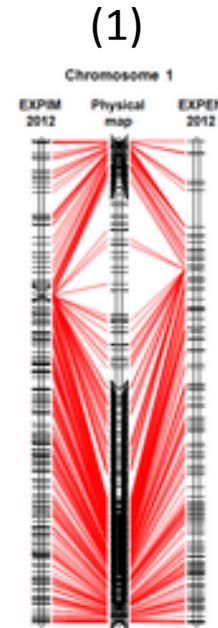
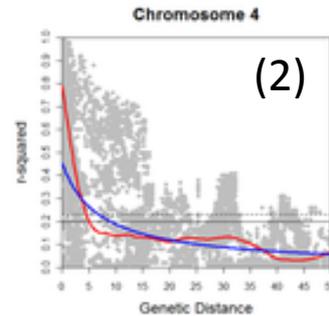
Approche méthodologique (1)



RNAseq
6 accessions
62,576 SNPs
Hamilton et al. (2012)

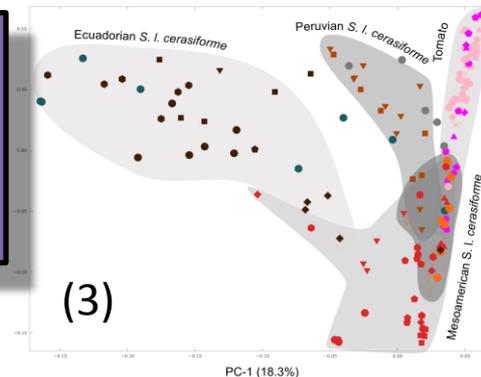
Illumina Infinium SNP array
7720 SNPs
Annotation : ITAG 2.30
Placés sur le génome
> 96% de succès

env. 200 SNP
Issus de
454 + Sanger
(EPGV)



Applications

- (1) Cartographie de liaison : Sim et al., 2012 PlosOne
- (2) Etude du DL : Sim et al., 2012 PlosOne
- (3) Analyse de la diversité: Blanca et al., 2012 PlosOne



Approche méthodologique (2)

- **Phénotypage** (2007 & 2008)
 - GC-MS: métabolites primaires (MPI, Golm)
 - Dosage : ASA, DHA, Sucres
- **Bio-Analyse**
 - Matrice de génotype : Plink!
 - Structure, apparentement : Admixture, Spagedi
 - GWAS : MLMM (Segura et al., 2012, Nat. Gen.)
 - Visualisation : snp.plotter, corrplot [R] packages

Analyse des phénotypes

Mesures

– $N_{\text{phénotypes}} = 76$

Corrections

– Effet année

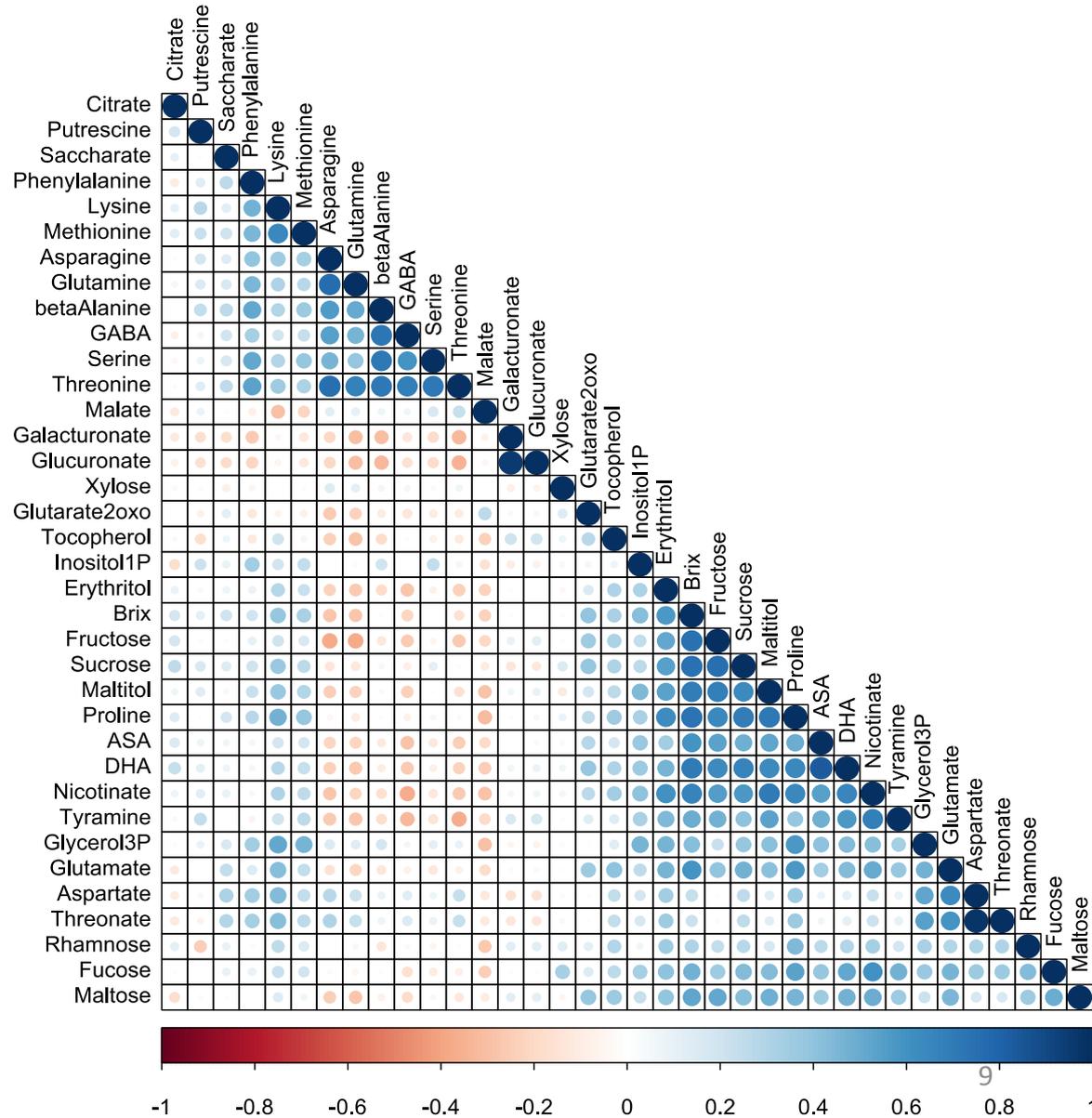
– Distr. Normale

Corrélations

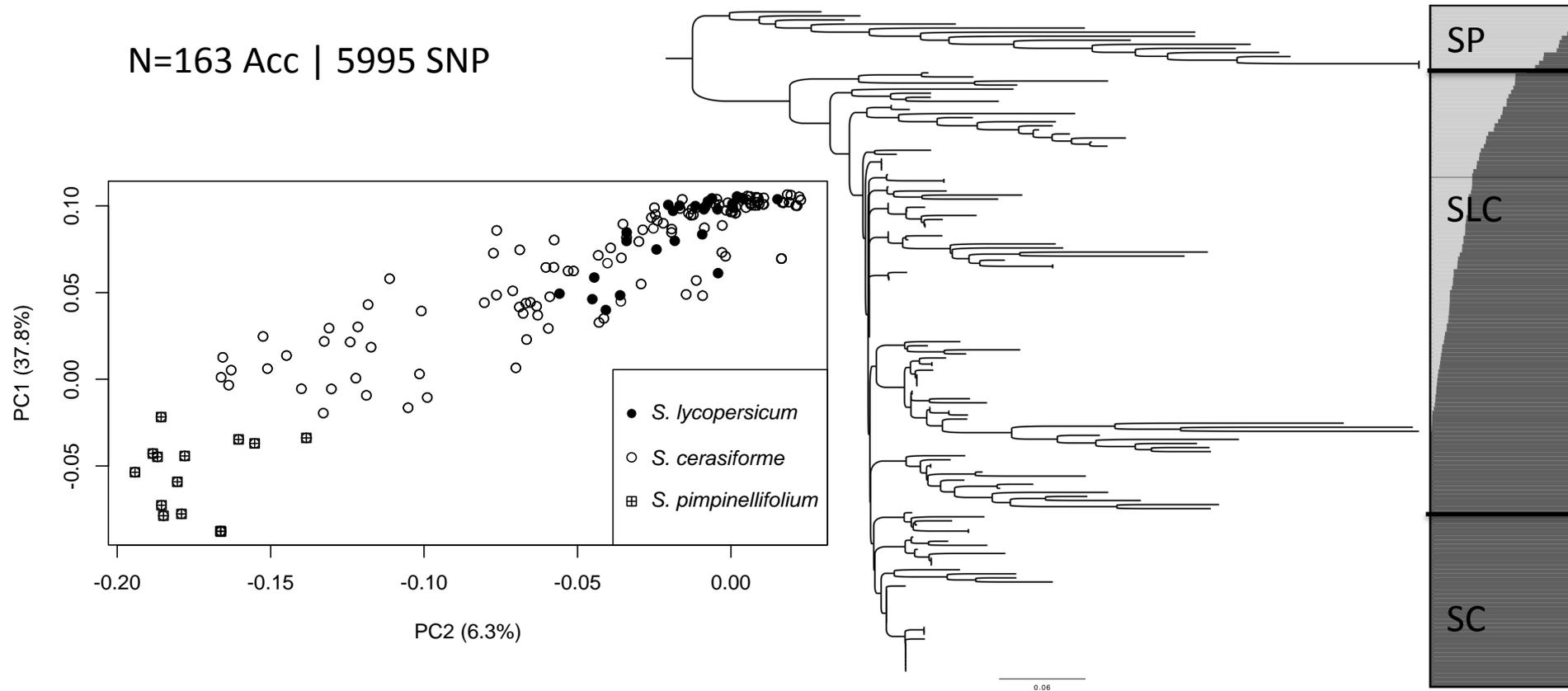
– Entre années

– Entre phénotypes

– $N_{\text{utiles}} = 36$

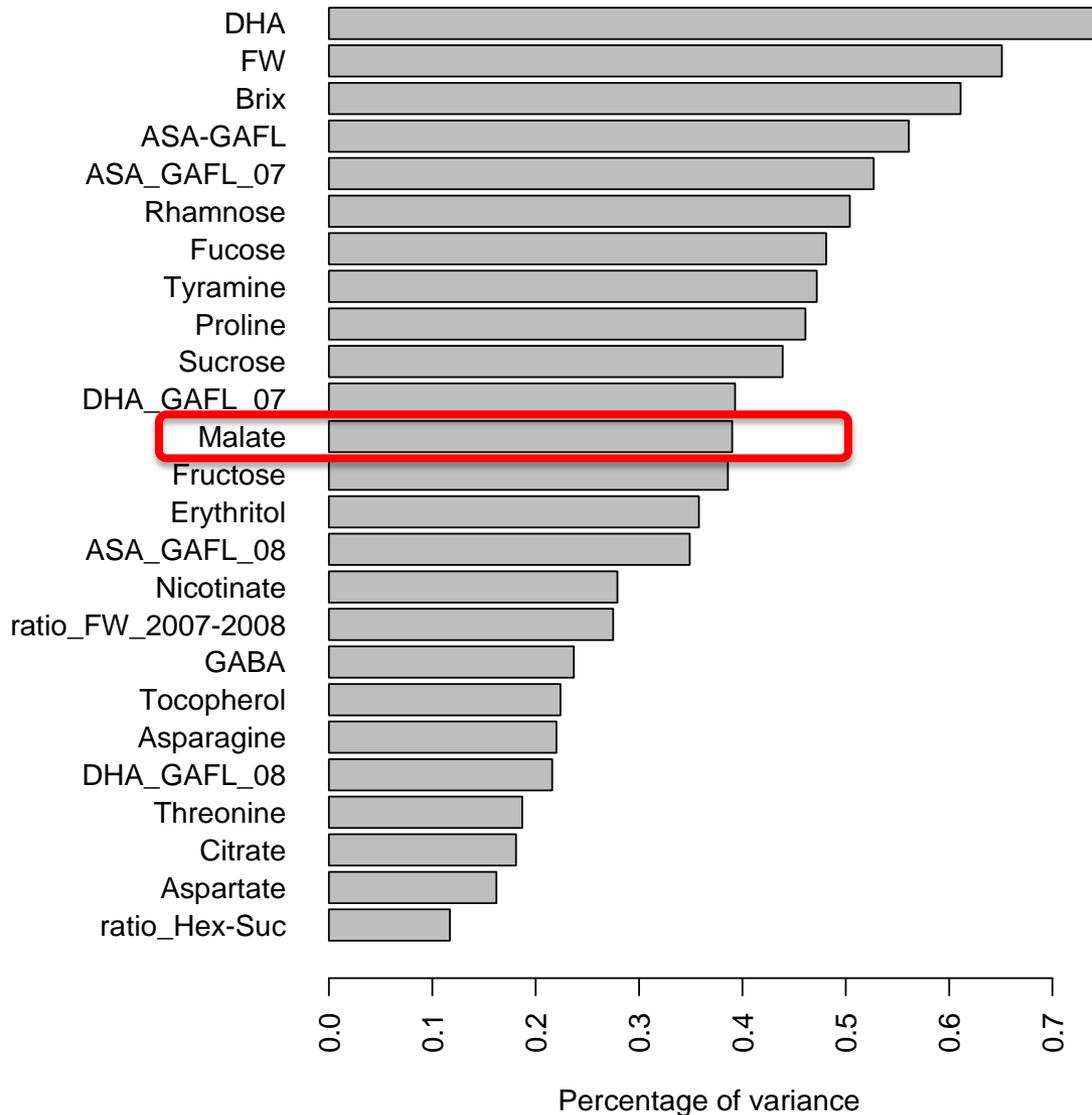


Etude de la structure de population



- ⇒ Forte structure de population ($Q=2$)
- ⇒ Structure liée à la taille du fruit
- ⇒ Apparemment moyen faible ($k=0.07$)

Paramètres estimés par le modèle



■ $0.162 < PVE < 0.743$

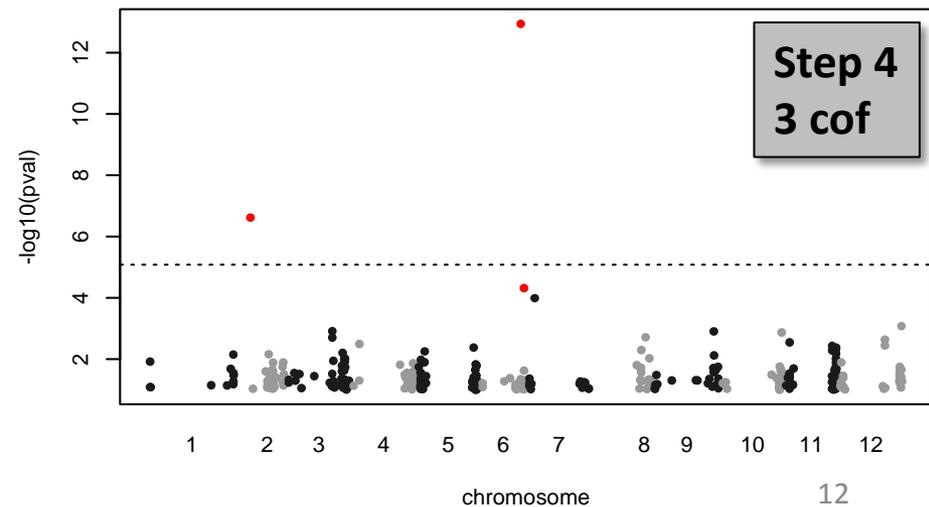
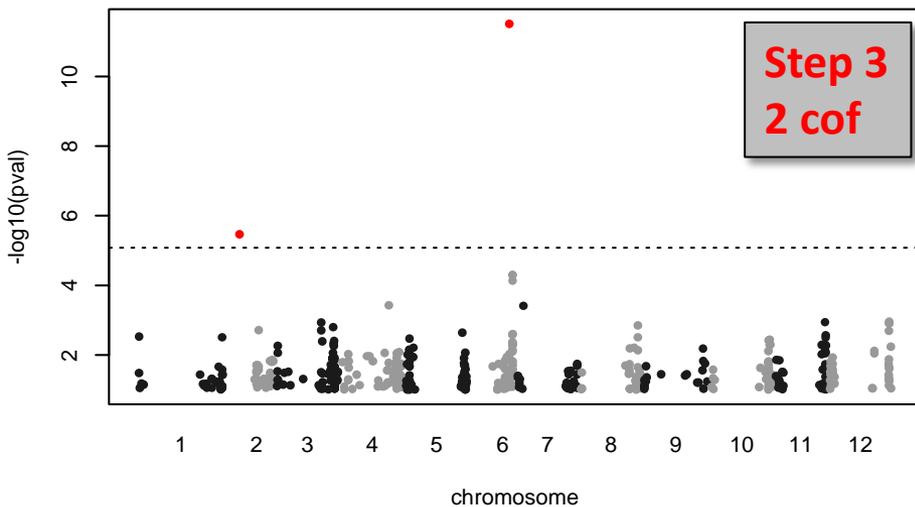
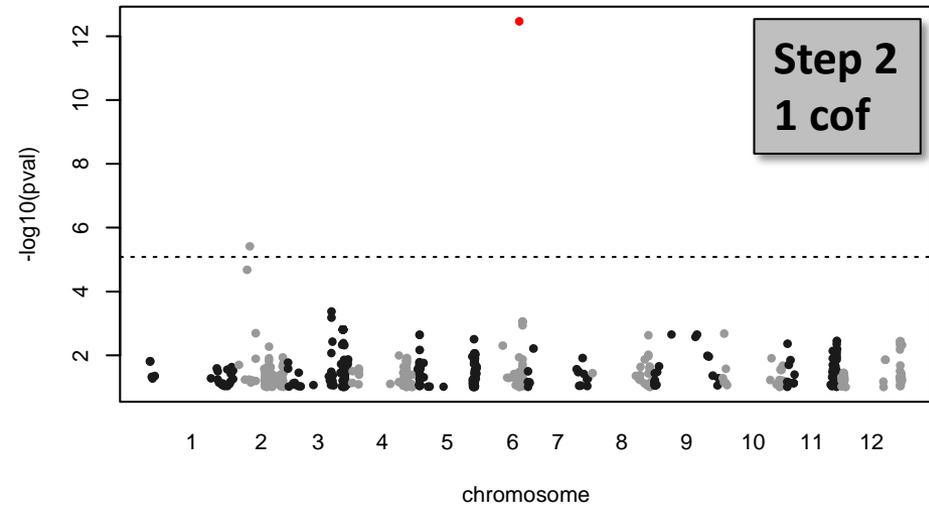
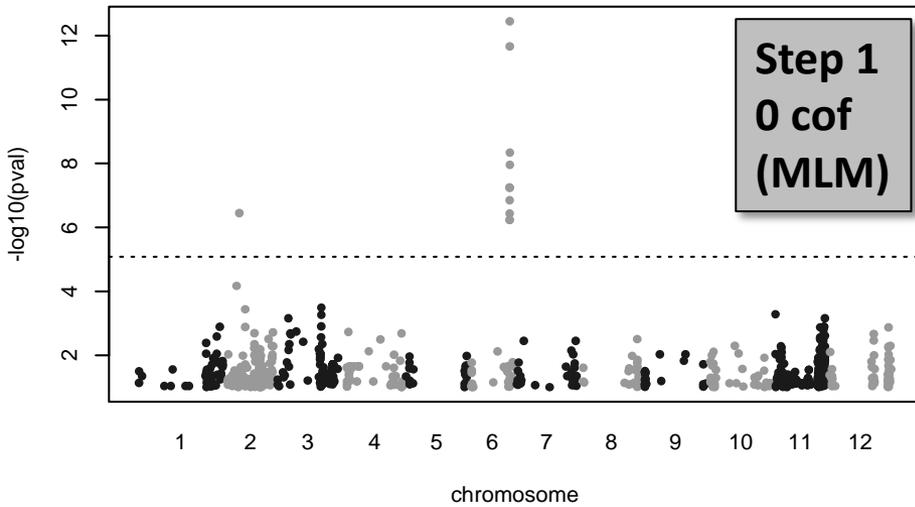
■ $0.12 < h^2 < 0.69$

■ $0.07 < h^2_{\text{missing}} < 0.458$

■ 1 à 9 loci /
phénotype

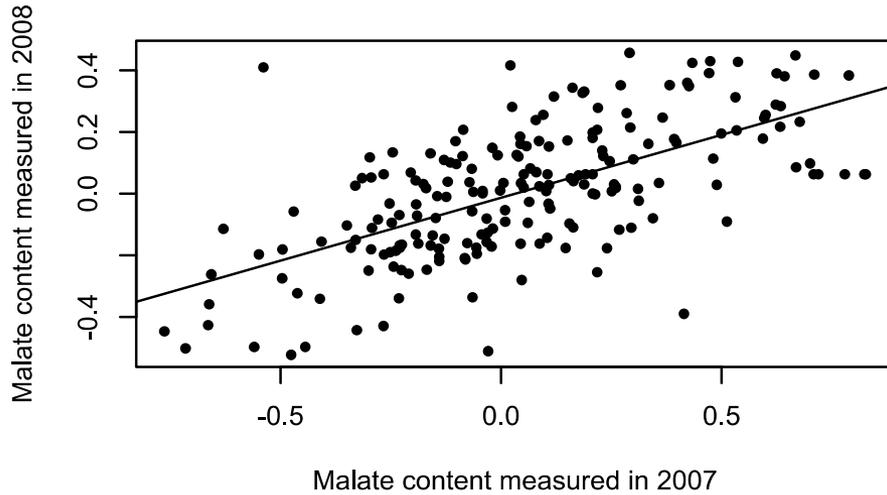
Apport du modèle multilocus

ex: la teneur en malate

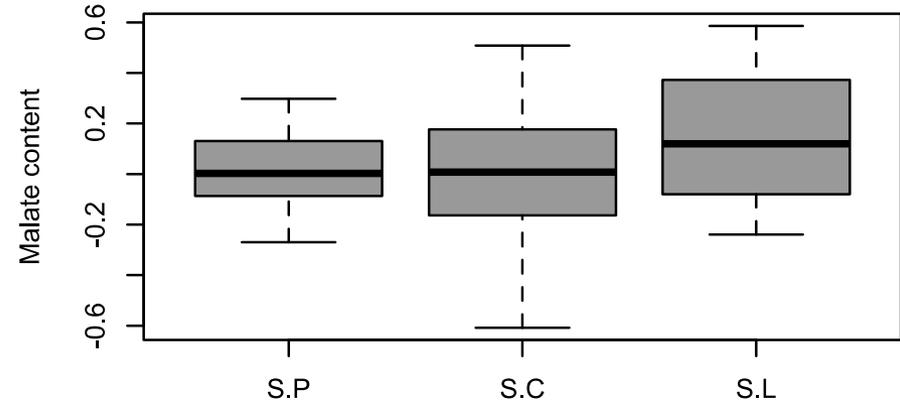


GWA pour la teneur en malate (1)

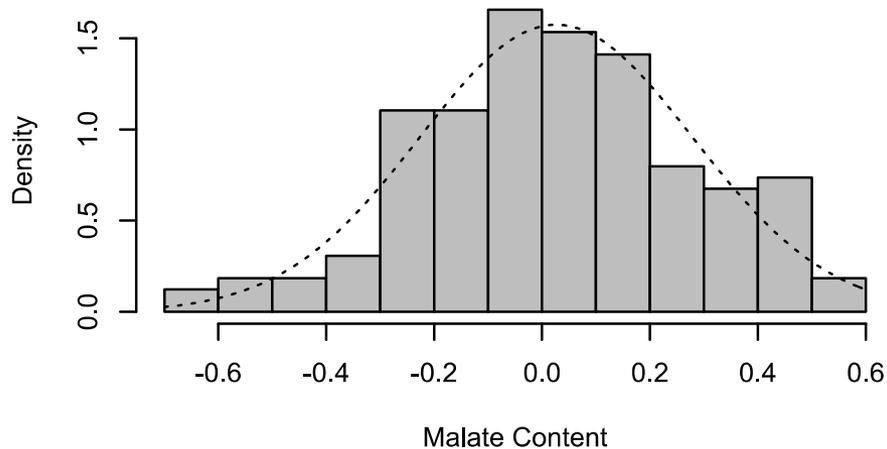
(A) Year effect on the measurement of malate content



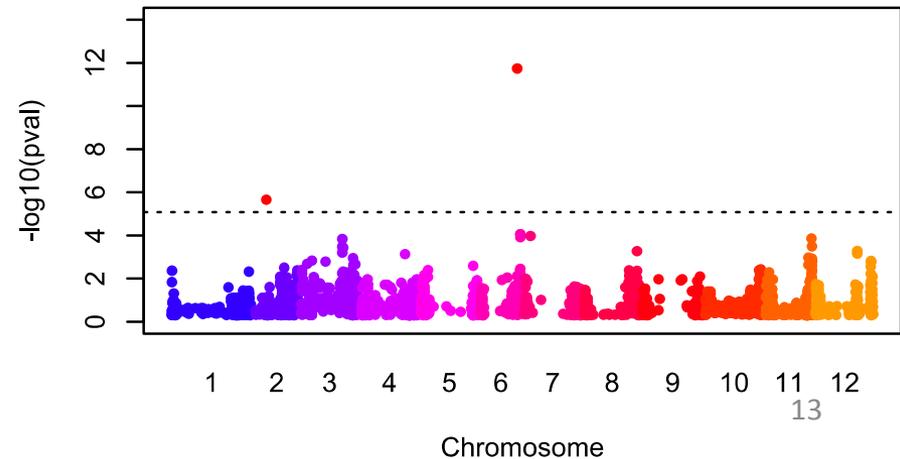
(B) Malate content in each species



(C) Distribution of Malate content in the Collection

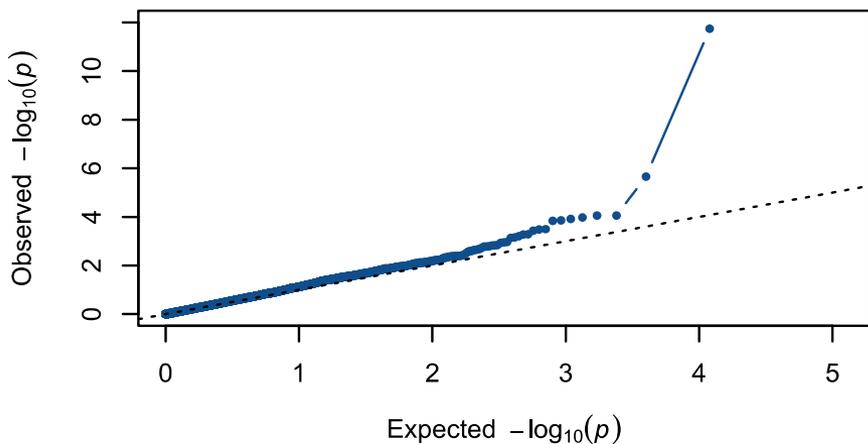


(D) Manhattan plot

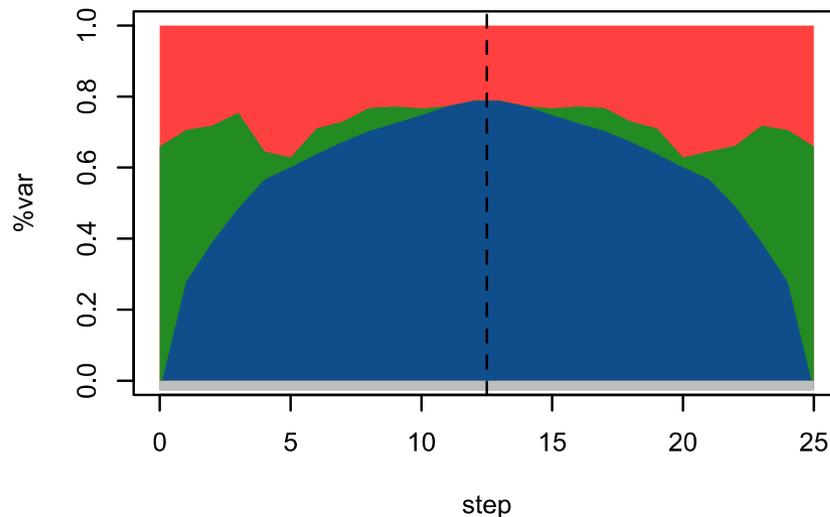


GWA pour la teneur en malate (2)

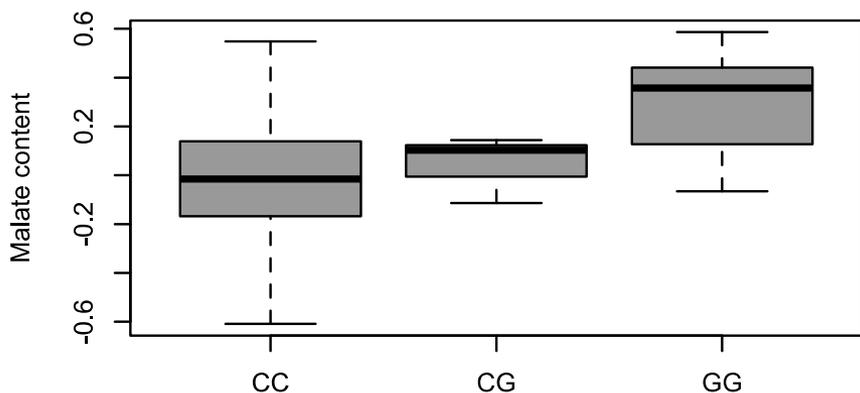
(E) QQplot for Optimal model according to mbonf



(F) Partition of variance at each step of MLMM



(G) Allelic effect for the associated SNP on chromosome 2



(H) Allelic effect for the associated SNP on chromosome 6

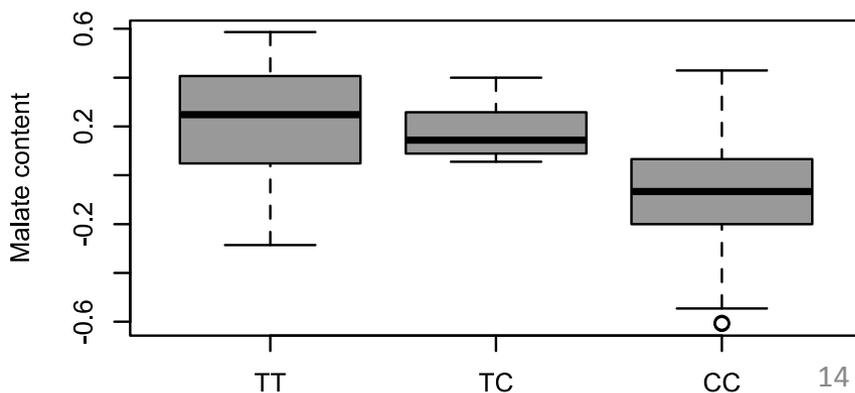
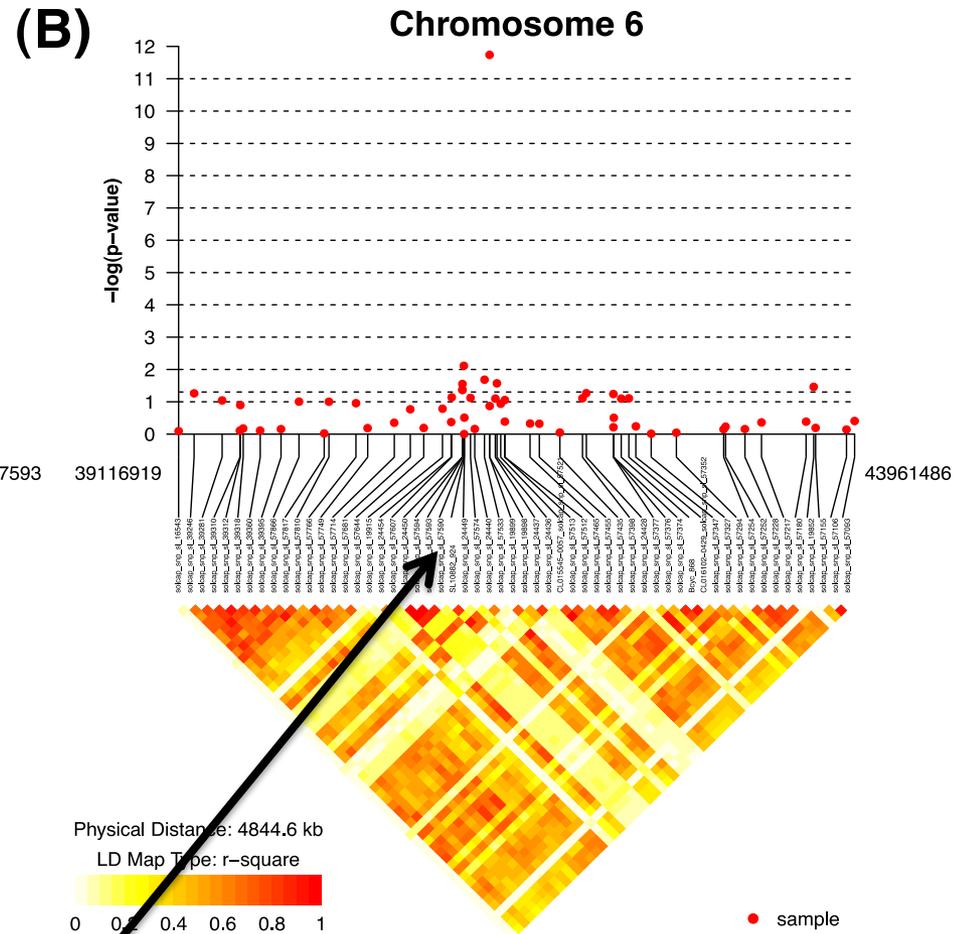
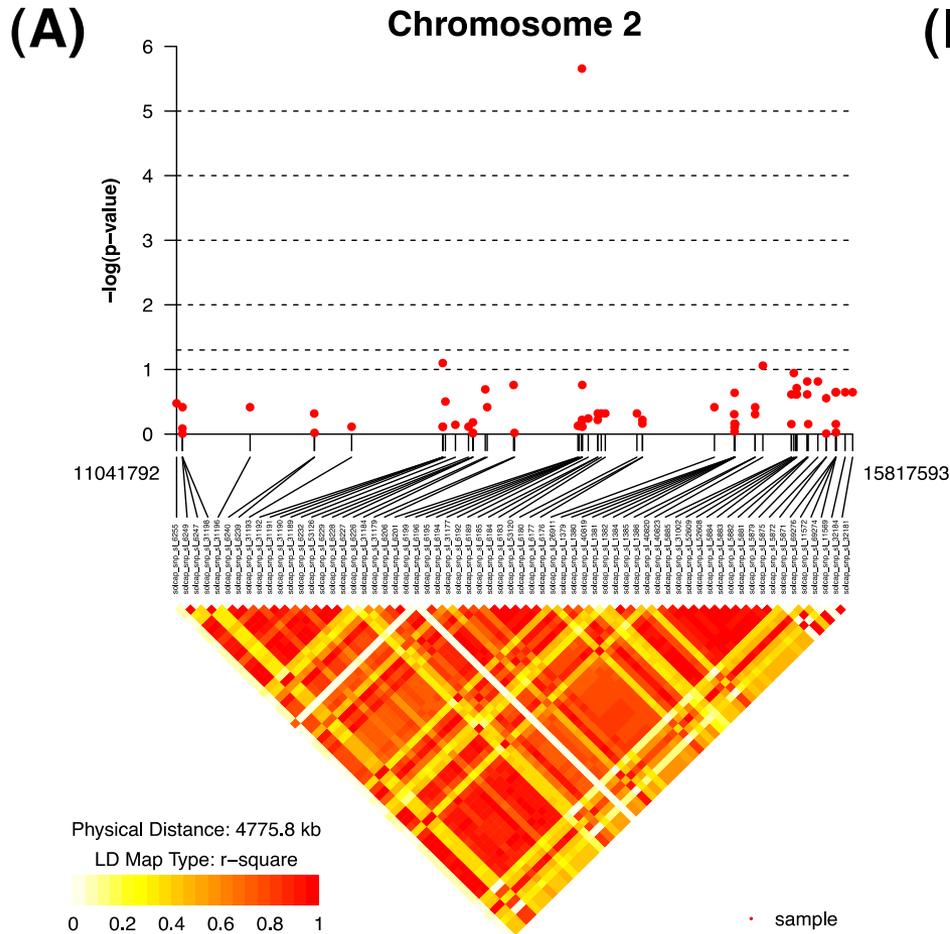


Figure 3 | A focus on the malate level results. (A) Correlation for the malate level over the two years of sampling in the collection of 163 accessions; (B)

A l'échelle locale...



Solyc06g072910.2 : aluminum-activated malate transporter-like !!
=> Approche de screening de la collection de tilling (Evry)

Validation via l'approche MAGIC (1)

12 Genotypes Sampling
8 parental lines + 4 F1

Phenotype

Metabolome

Proteome

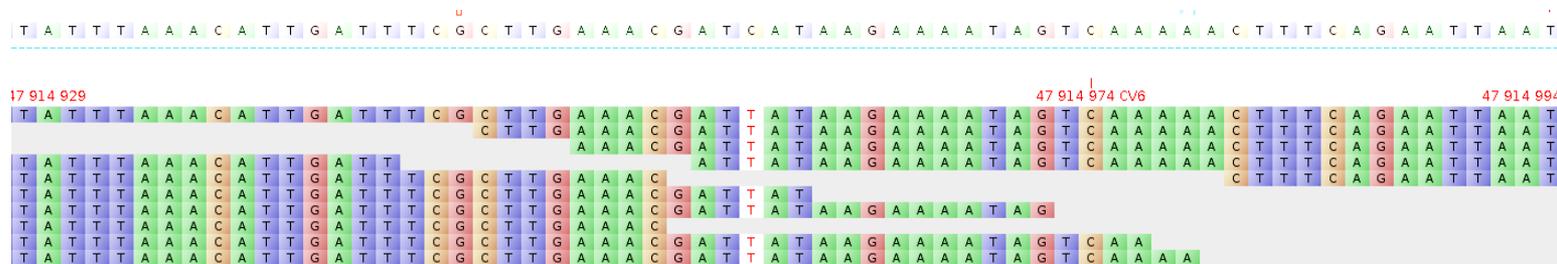
Transcriptome

Genome

8 génomes parentaux reséquencés (GAll, EPGV Projet MagicTomSNP)

→ Depth : 9x-26x (7-17x after cleaning & mapping)

Coverage >90% per line



Détection de SNP (VarScan) :

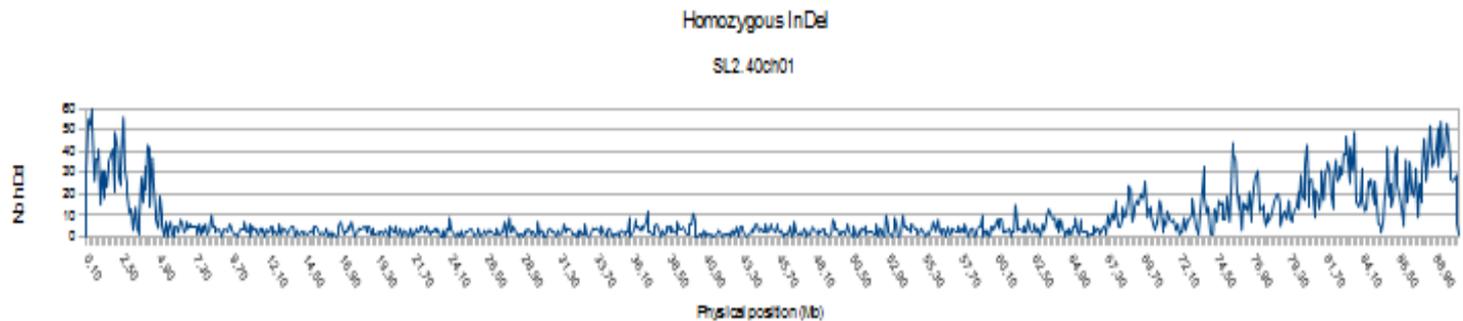
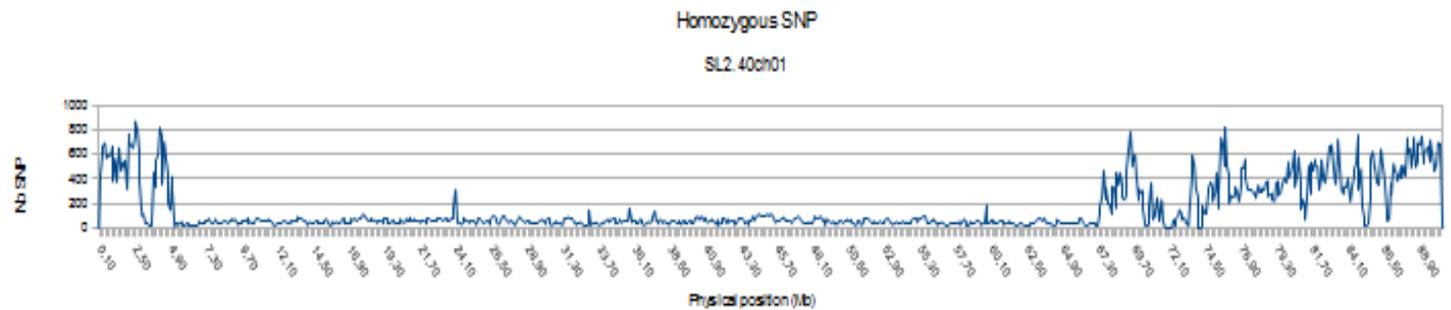
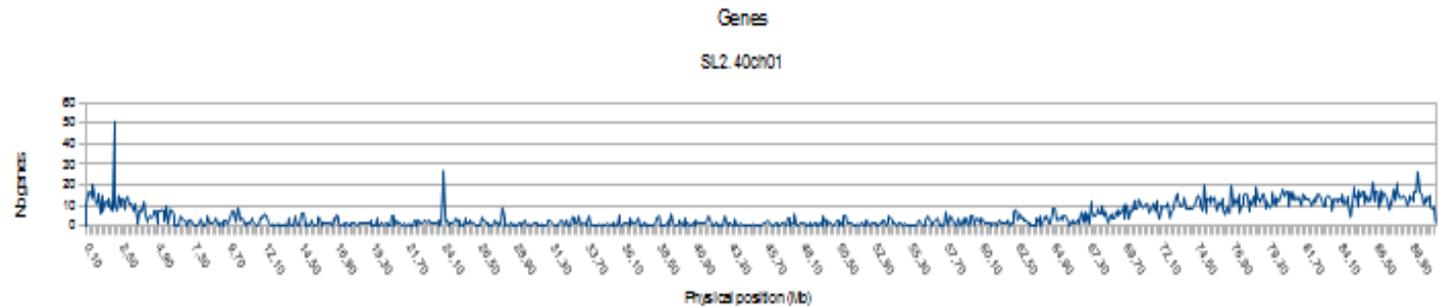
min quality = 30, min coverage = 8, P-value = 1e-3, allelic freq= 0.9

→ 80,000 à 1,500,000 SNP par lignée line

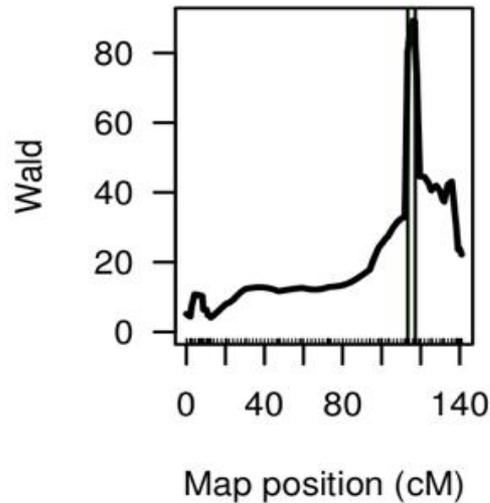
→ > 4.10⁶ SNP

→ Développement d'un panel de 1536 SNP (cartographie de liaison)

Validation via l'approche MAGIC (2)

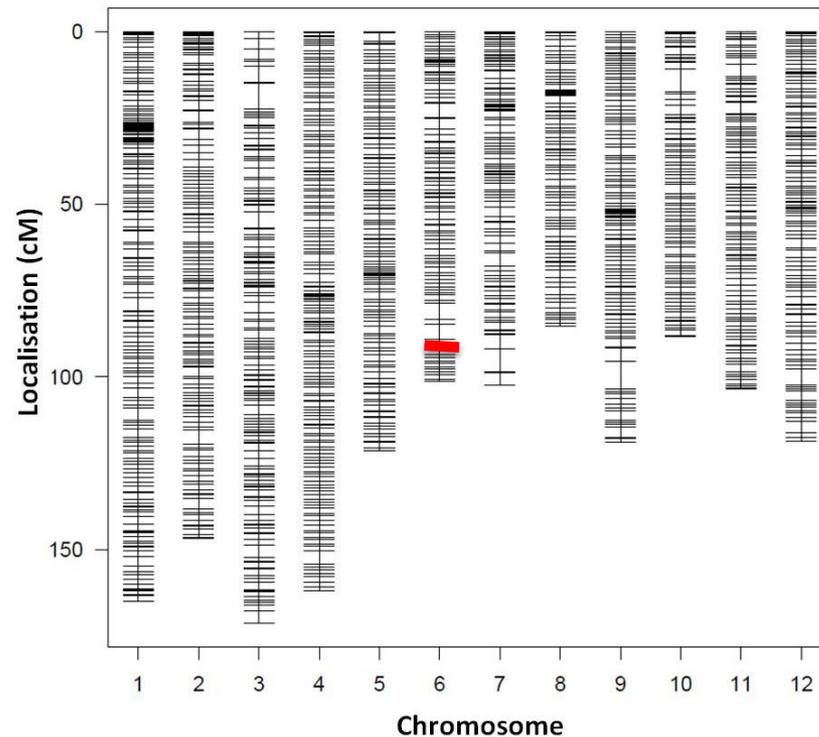
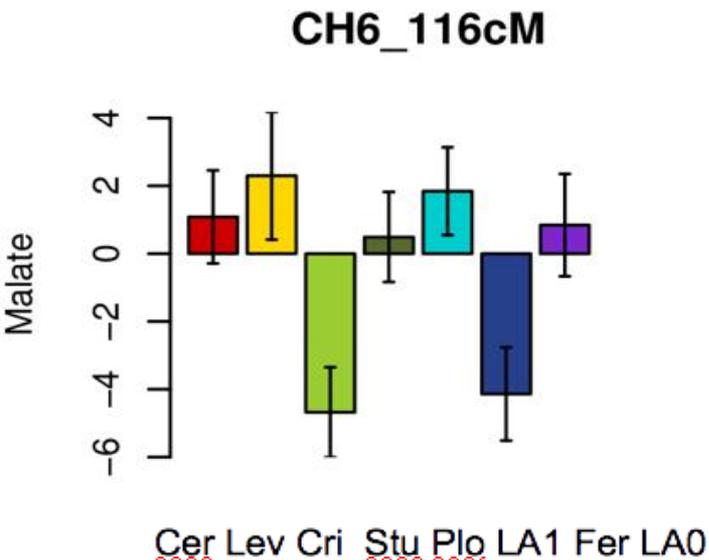


Validation via l'approche MAGIC (3)



QTL lié à la teneur en malate

1 QTL localisé sur le Chr. 6
41.15 cM < 95% CI < 41.60 cM
PVE de 39% (GWAS – 39%!)



A retenir !

Méthologie

- Amélioration des modèles de GWAS
- Effet important de la structure de population
- Qualité primordiale du phénotypage

Résultats

- Architectures génétiques très variables (1 à 9 loci)
- Identification de candidats potentiels 'sérieux'
- Validation par des approches parallèles (MAGIC + GWA)

Perspectives

GWAS

- Validation fonctionnelle de(s) candidat(s) : Tilling, RNAi
- GWA par utilisation des haplotypes
- Utilisation de la puce SNP en SELGEN (?)

MAGIC

- RNAseq : étude transcriptomique (fruit)
- GBS low coverage (1-3x) : cartographie
- Interactions GxE (stress hydrique)

Remerciements:

GWAS

- **INRA Avignon** – G. Bauchet, R. Stevens, M. Causse,
- Y. Carretero, H. Burck & R. Duboscq
- **INRA Orléans** – V. Segura
- **INRA URGI** – D. Steinbach & G. Merceron
- **INRA EPGV** – D. Brunel & MC. Le Paslier
- **MPI Golm** – A.R. Fernie, Z. Nikoloski & P. Thi Do
- **Trait Genetics** – M. Ganal



Financement: FP6 – Projet Européen EuSol

Article associé: Sauvage *et al.* (2014) Plant Phys. On line

Données disponibles sur le GNPis | GNPasso de l'URGI

MAGIC

- **INRA Avignon** – L. Pascual-Banuls, N. Desplat & JP. Bouchet
- **Société Vco**
- **INRA EPGV**



Financement: ANR, Vco, INRA

Email: christopher.sauvage@avignon.inra.fr