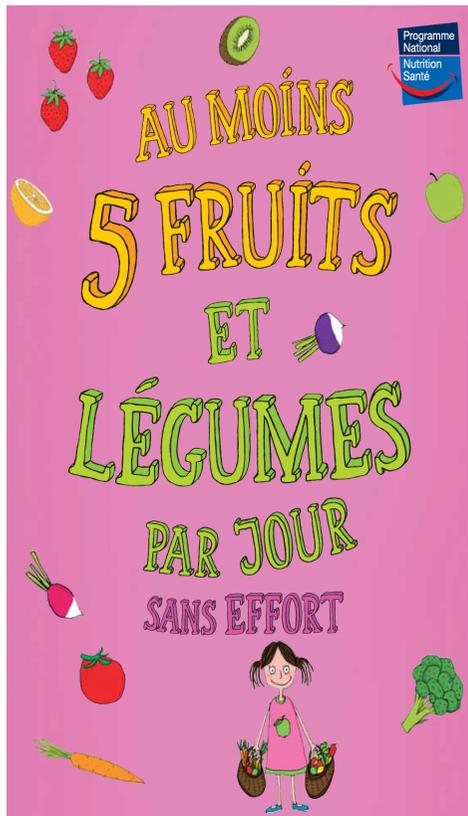


# Identification de locus liés à la qualité du fruit par généétique d'association chez la tomate – projet VTP

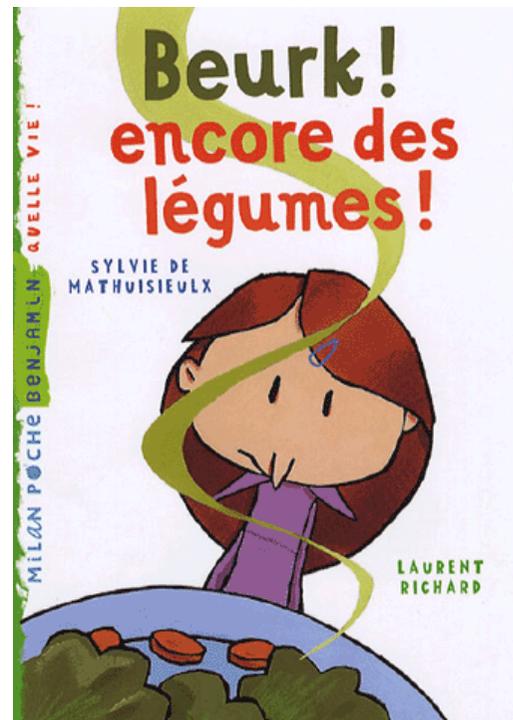


Nicolas Ranc

# La Qualité chez la tomate



Recommandé par le **Programme National Nutrition Santé**



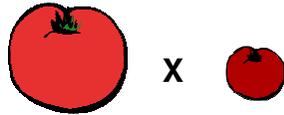
Ne peut se développer sans prendre en compte la **qualité organoleptique et nutritionnelle** des produits

**La tomate = espèce modèle pour les plantes à fruits charnus  
Génome en cours de séquençage**

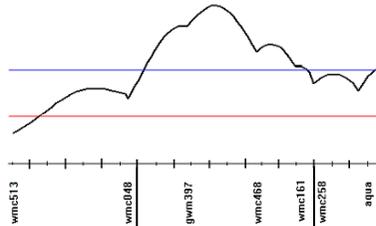
# Identification des bases moléculaires

## Linkage Mapping

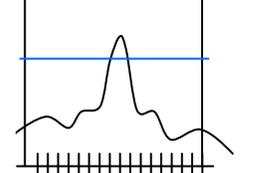
Construction de populations  
F2, BC, HD



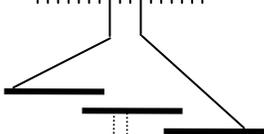
Détection de QTL



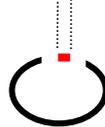
Cartographie fine



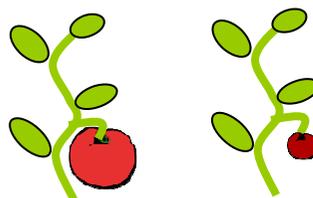
Cartographie physique



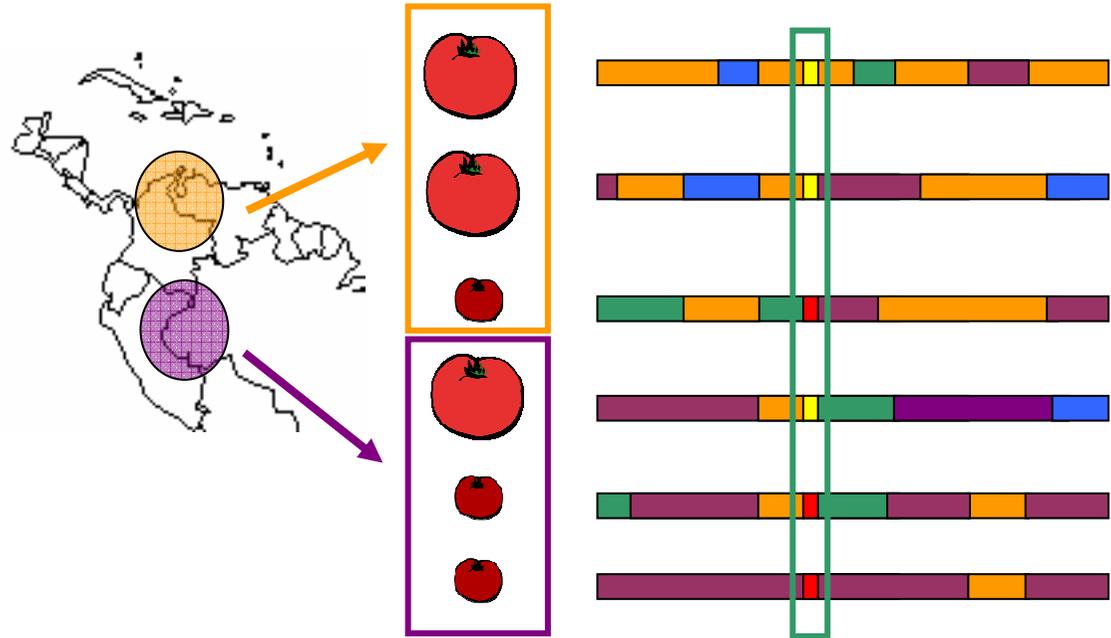
Clonage



Validation fonctionnelle



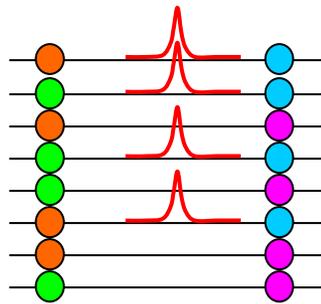
## Association Mapping



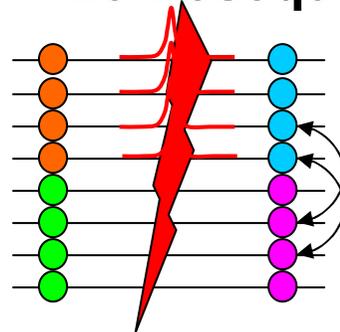
- 🍅 Approche récente pour la **découverte d'allèles** d'intérêt et la **validation de gènes** candidats.
- 🍅 Développée pour la **génétique humaine** et adaptée aux **plantes cultivées allogames**
- 🍅 **Qu'en est il des plantes cultivées préférentiellement autogames comme la tomate?**

# Introduction

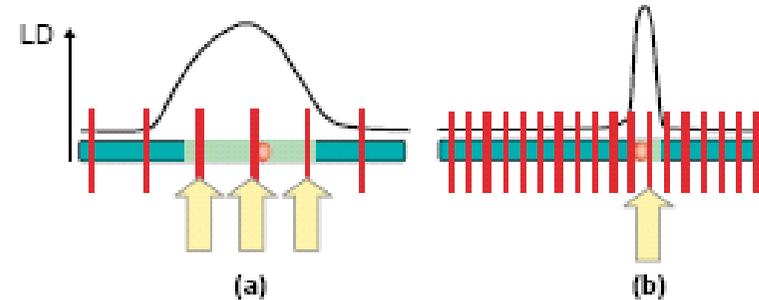
## Le Déséquilibre de Liaison



DL = 0



DL = max



©Current Opinion In Plant Biology

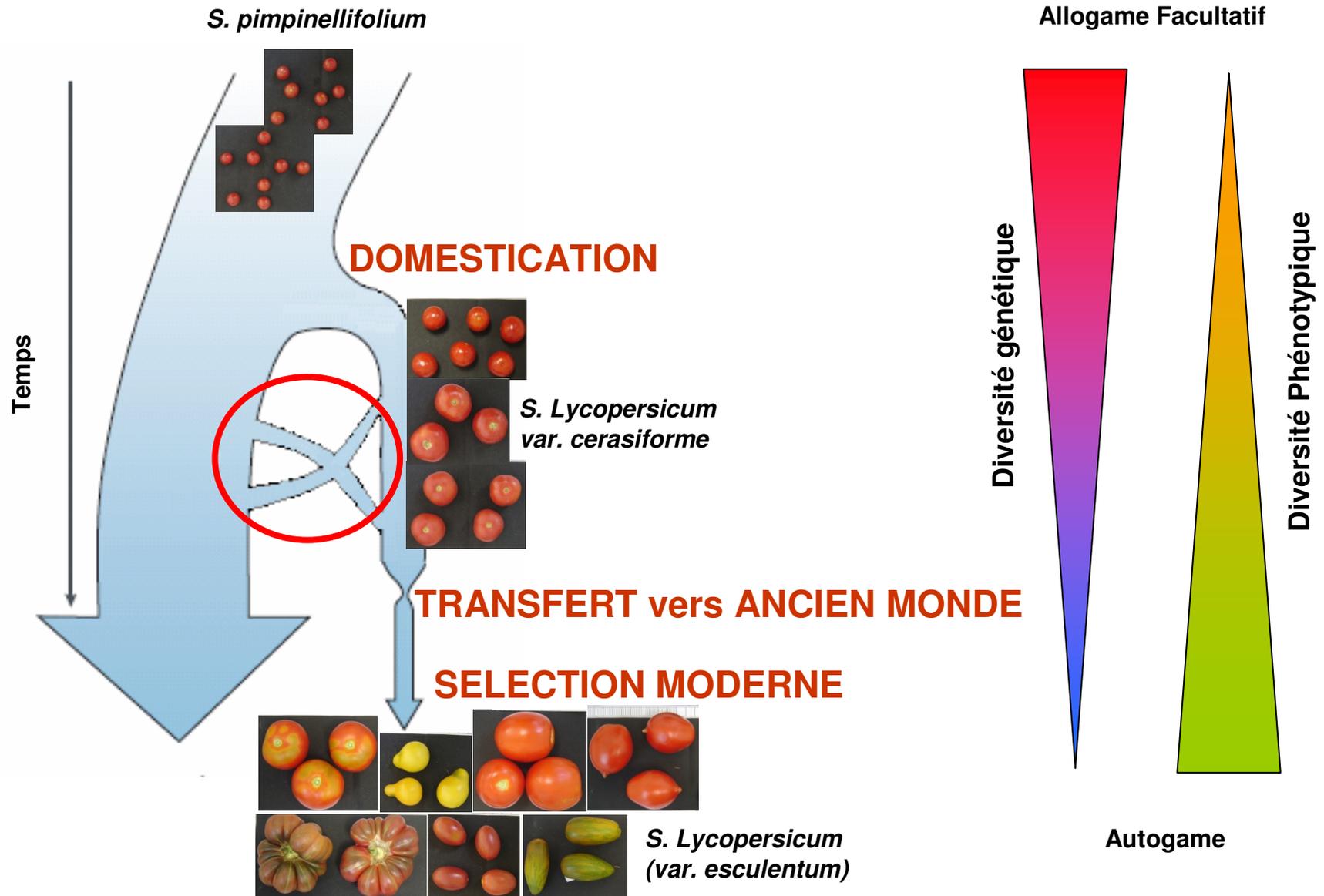
- Faible diversité moléculaire
- Autofécondation → Fort DL attendu

DL	Fort	Faible
Resolution	faible	élevée
Nb de marqueurs requis	faible	élevé
Approche de génétique d'association	"whole-genome scan"	gène candidat

Les collections de tomates ne semblent pas adaptées:



# Histoire évolutive de la tomate



# Methodologie

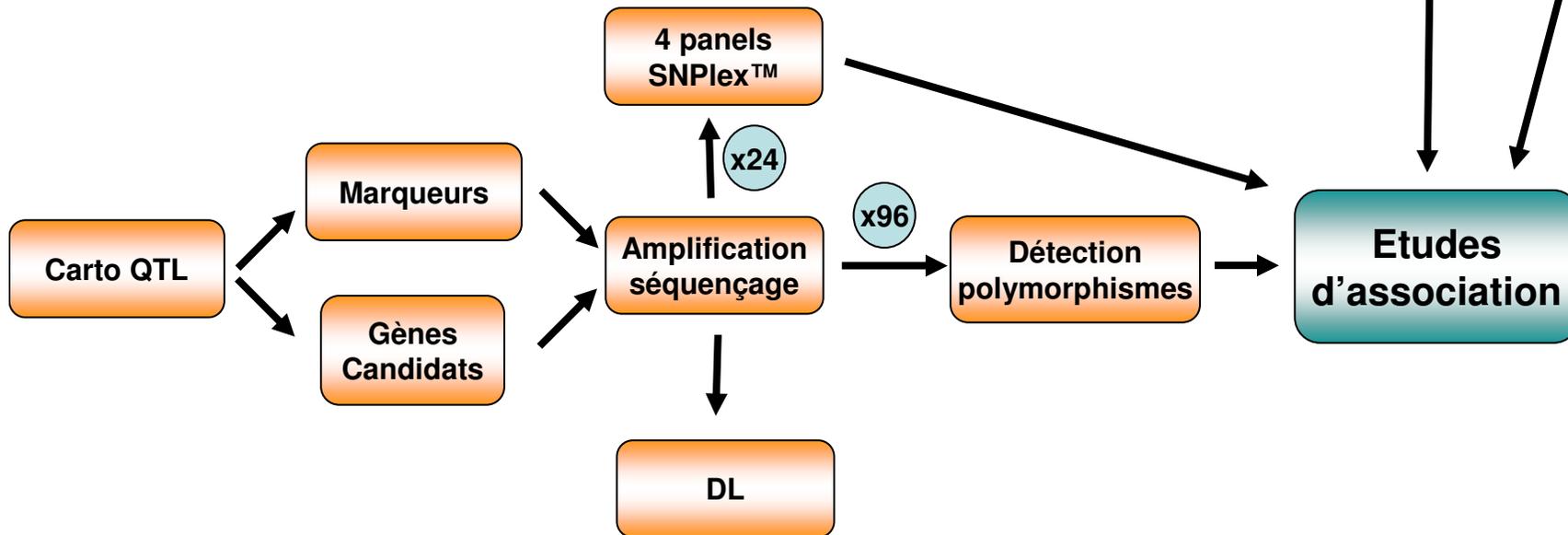
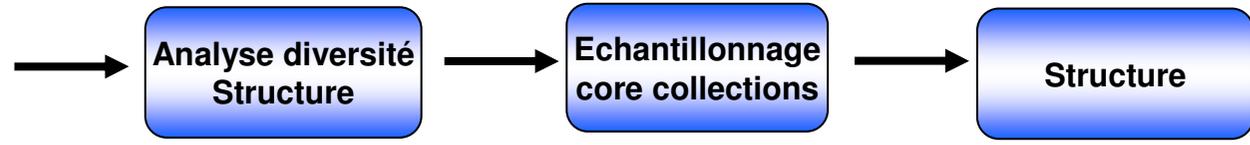
360 accessions

130 *Solanum lycopersicum*

144 *S. l. cerasiforme*

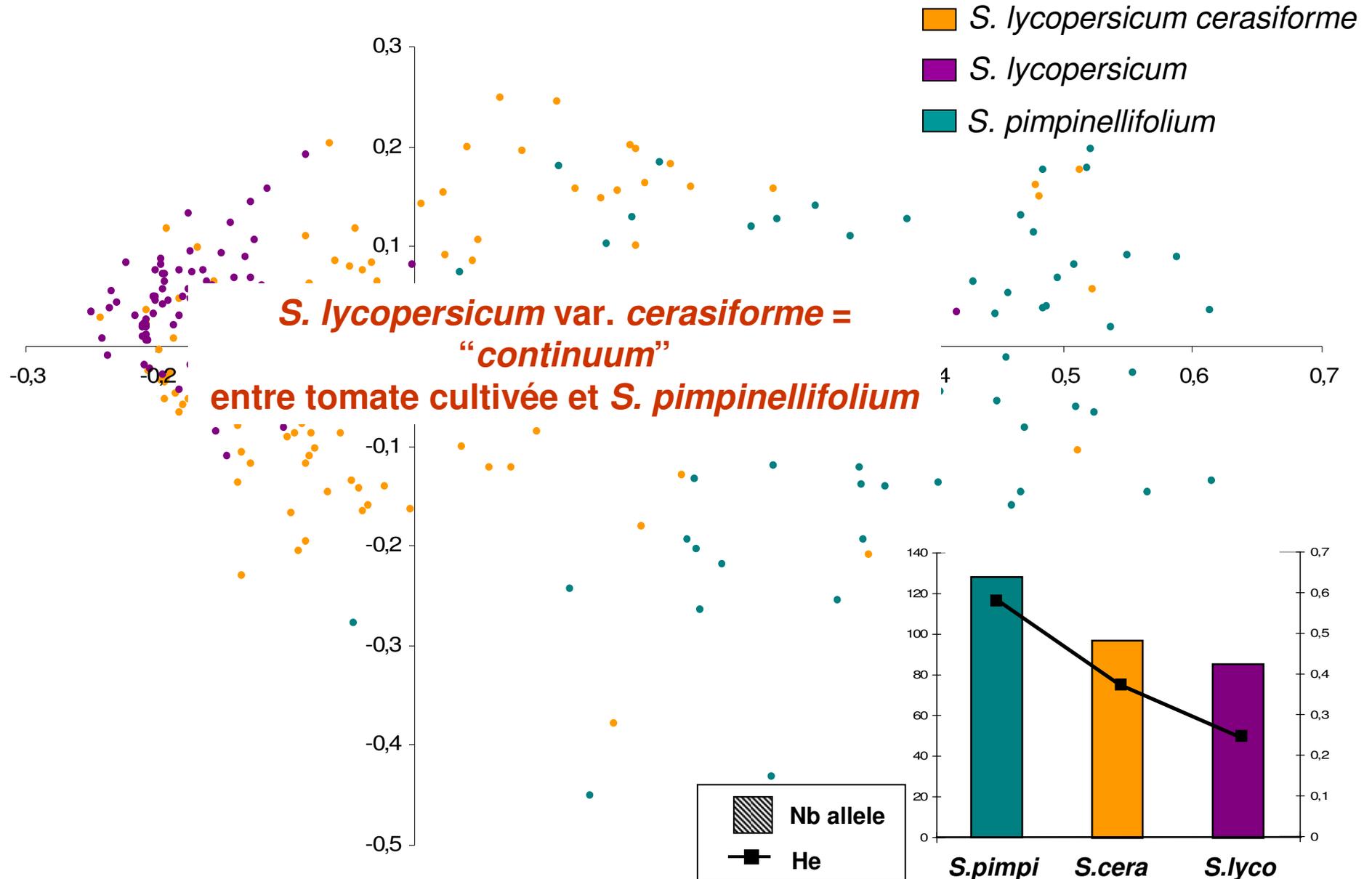
66 *S. Pimpinellifolium*

20 sauvages éloignées



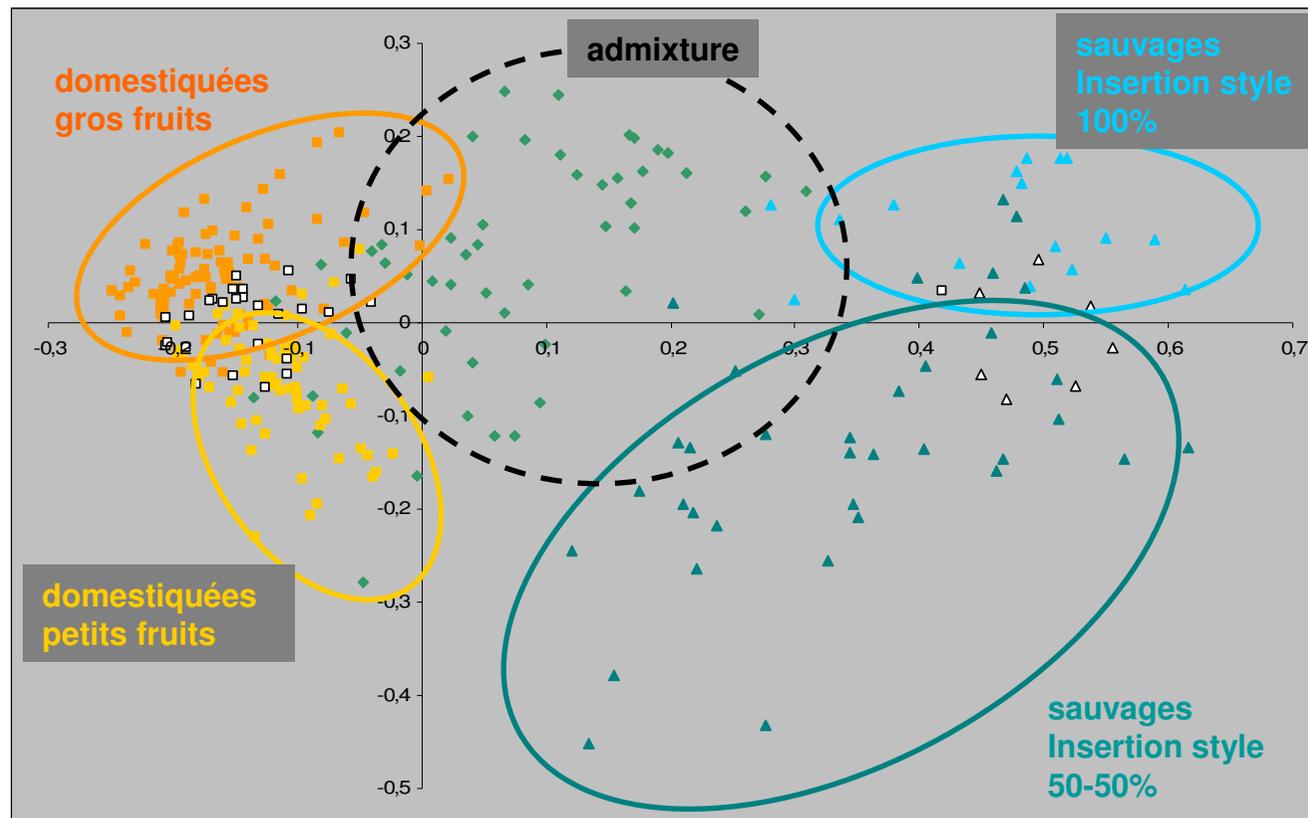
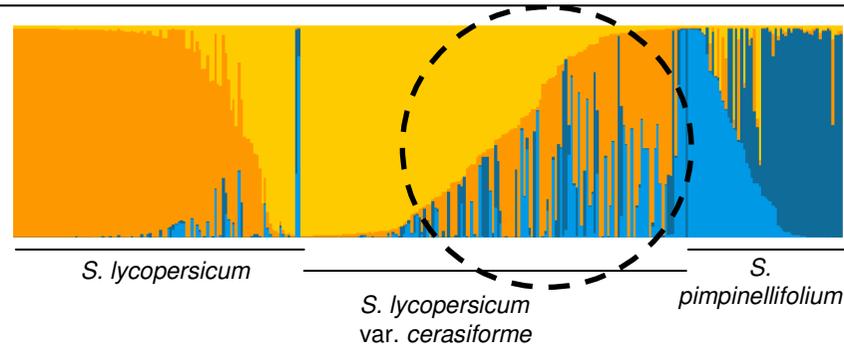
# Analyse de diversité

Analyse en Coordonnées Principales sur les accessions à fruits colorés



# Analyse de diversité

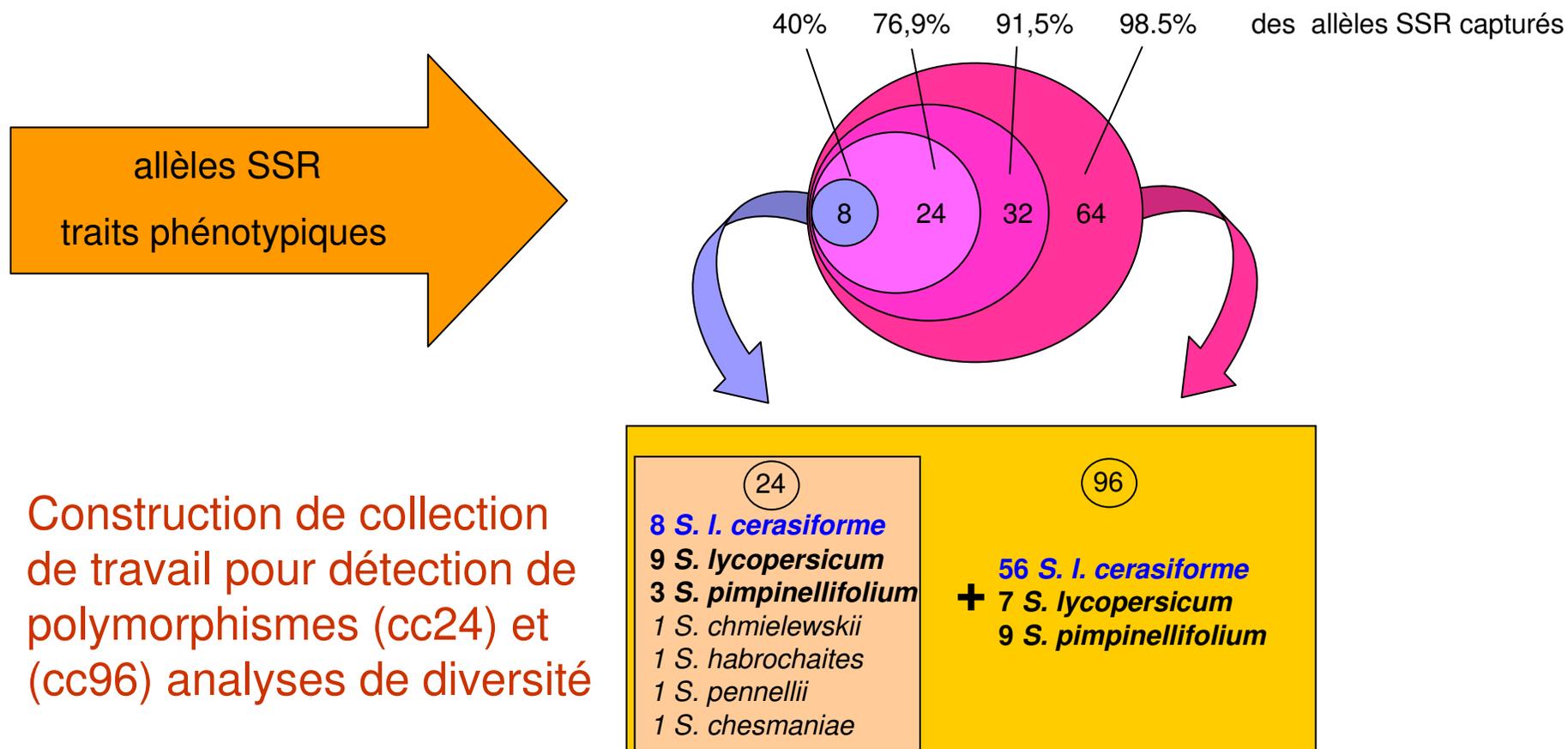
Structure2.0  
Hyp K = 4 groupes



*S. lycopersicum* var. *cerasiforme* semble être un échantillon d'intérêt pour les études d'association

# Construction de Core collections

Construction de core collections emboîtées de *S. lycopersicum* var. *cerasiforme* avec un algorithme de maximisation de la diversité allélique



# CGIS

**Faible polymorphisme moléculaire**  
(1 SNP/4Kb décrit)

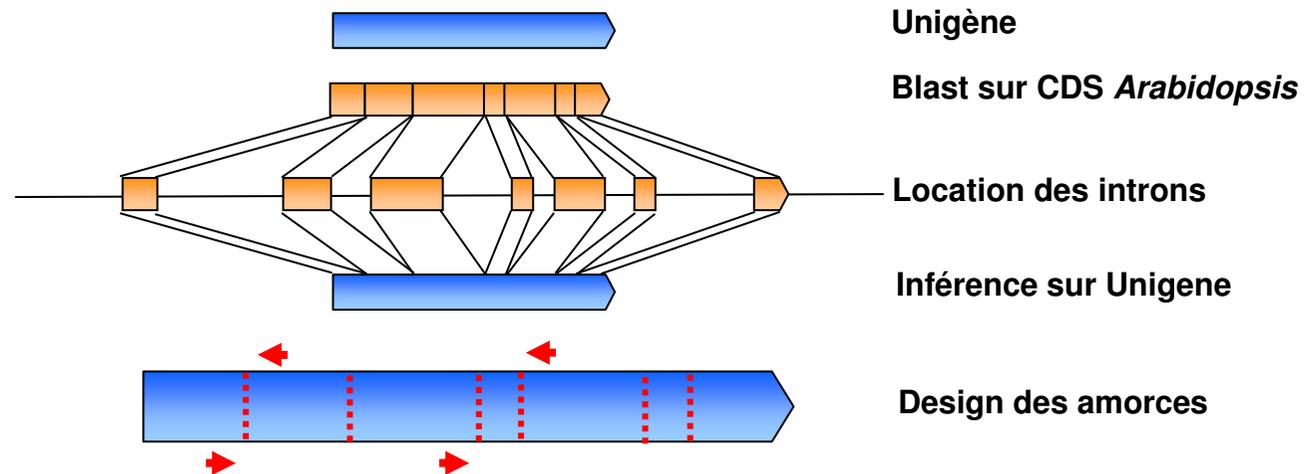
**Séquençage focalisé sur régions non codantes** : Introns & régions intergéniques

Définition des fragments à séquencer :

Séquences Intergéniques → Utilisation Primer3

Gènes candidats positionnels → Utilisation CGIS

Marqueurs COS (Unigènes) → Utilisation CGIS

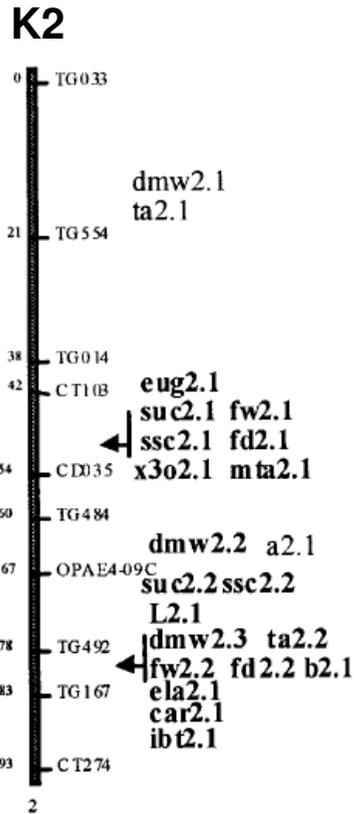


CGIS: an information system used for designing primers of candidate genes using *Arabidopsis* whole genome and *Solanaceae* EST databases, Bres C., Bouchet J.P. Nucci L., Caromel B., Lefebvre V. , 2005 , Triennial Conference of the EAPR

# Séquençage allélique

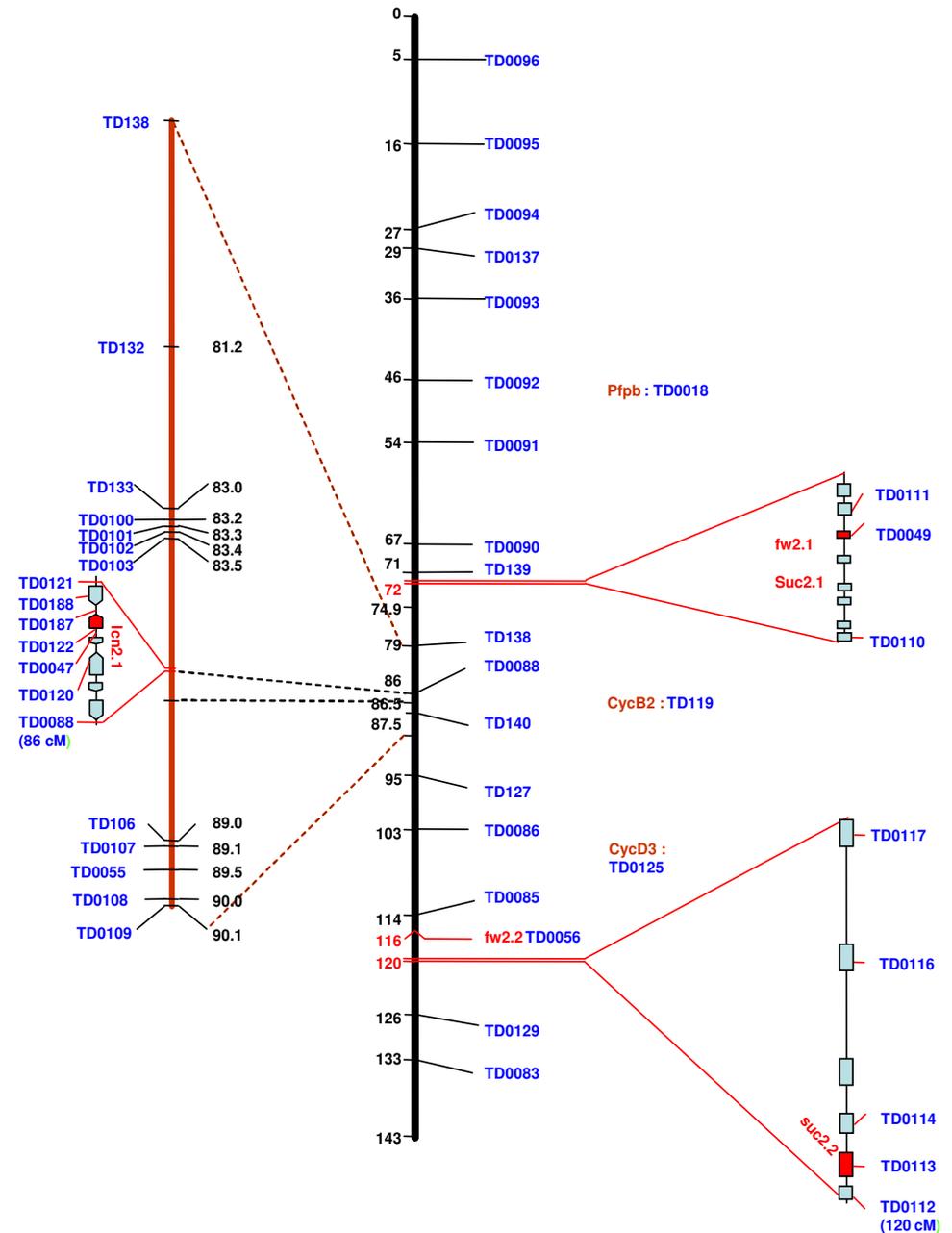
Cartographie  
**QTL qualité**  
 dans  
 population  
*esculentum*  
 X  
*cerasiforme*

**K2 =**  
 Cluster de  
 QTL



Définition de marqueurs sur **3 échelles**  
 différentes sur **K2**:

- 1- **Chromosome** (1 mk / 8 cM)
- 2 - **Carto fine** (1 mk / cM)
- 3 - **BAC**



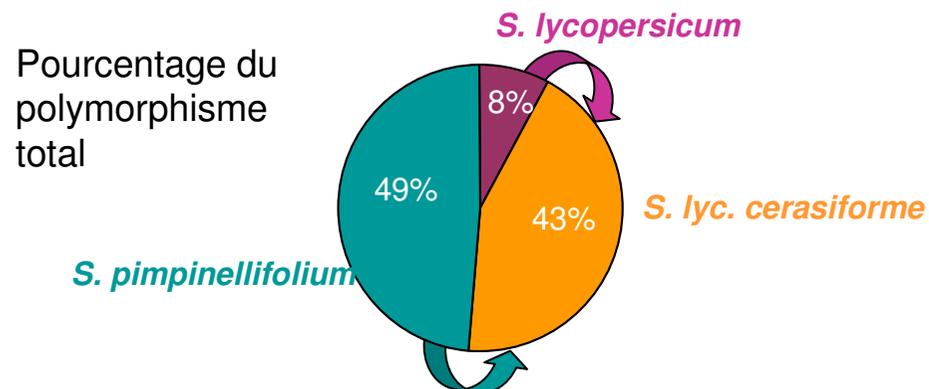
# Identification de SNP

130 fragments séquencés

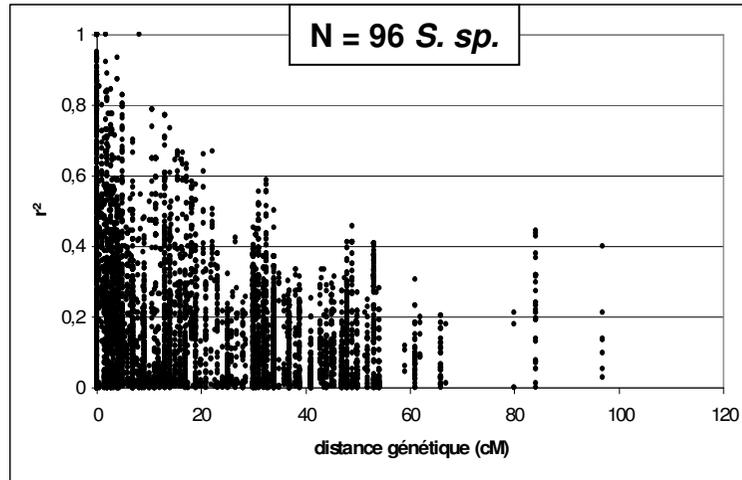
- 91 gènes candidats fonctionnels (qualité du fruit) sur 24 individus (tous chr)
- 62 gènes candidats et fragments "neutres" sur 96 individus (que chr 2)

Fréq. moy. de SNP sur 20 acc. fruits colorés	Longueur fragments polymorphes (kb)	Nb SNPs*	Fréquence SNP dans introns	Fréquence SNP dans exons
Total (+ <i>S. pimpinifolia</i> )	130 (70 kb)	395	1 / 120 bp	1 / 360 bp
<i>S. Lycopersicum</i> (9)		78	1 / 620 bp	1 / 1800 bp
<i>S.l. cerasiforme</i> (8)		328	1 / 130 bp	1 / 380 bp

\*singletons exclus

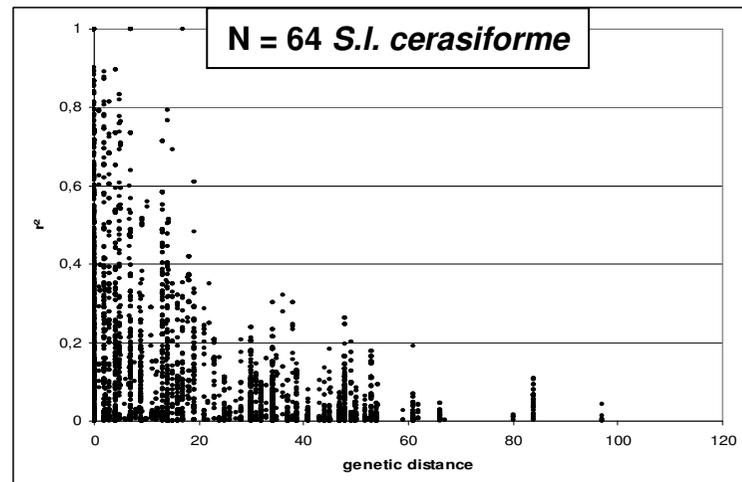


# Déséquilibre de liaison sur le chromosome 2



Echantillon structuré

Effet de la structure  
génétique sur l'étendue  
du DL



Echantillon non structuré

Prise en compte de la structure pour éviter de détecter de fausses associations

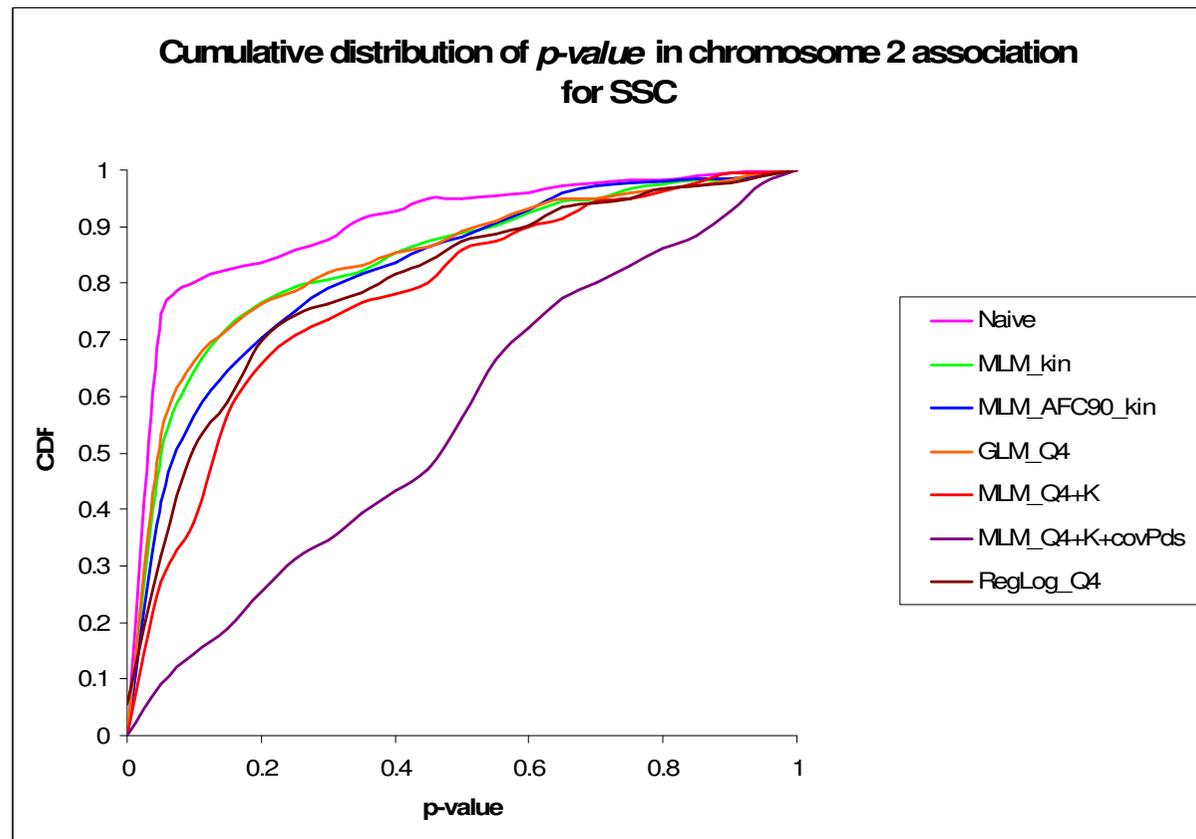
# Analyse d'association – différents modèles

Génétique d'association ~ détection de QTL sur populations naturelles.

**Modèle logistique** (Pritchard, 2001)

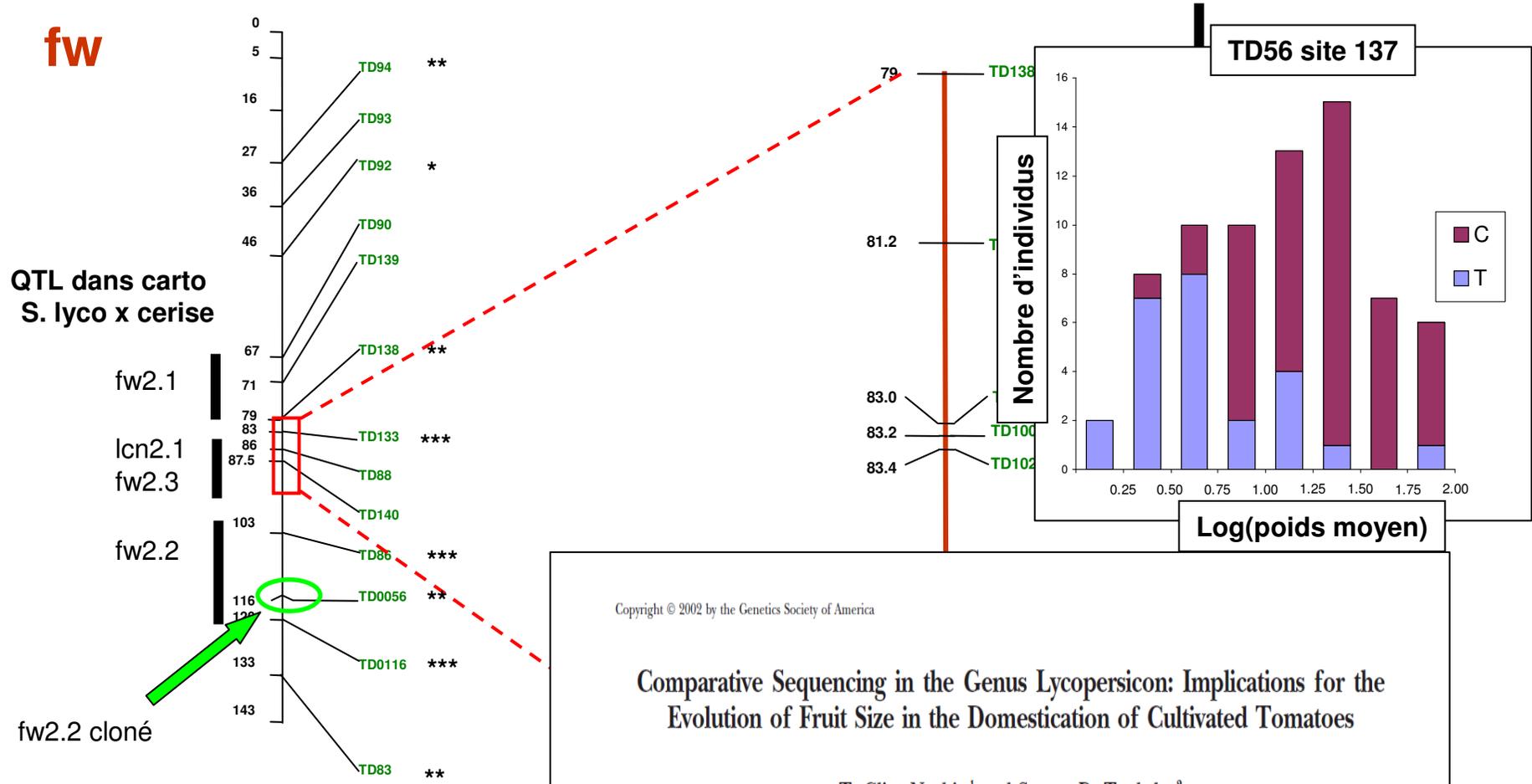
**Modèle Linéaire Généralisé** prenant en compte la structure génétique : **Q**

**Modèle Linéaire Mixte** prenant en compte la structure et l'apparentement entre individus : **Q + K** ou **AFC + K** (Yu et al. 2007; Zao et al. 2007)



# Analyse d'association sur chromosome 2

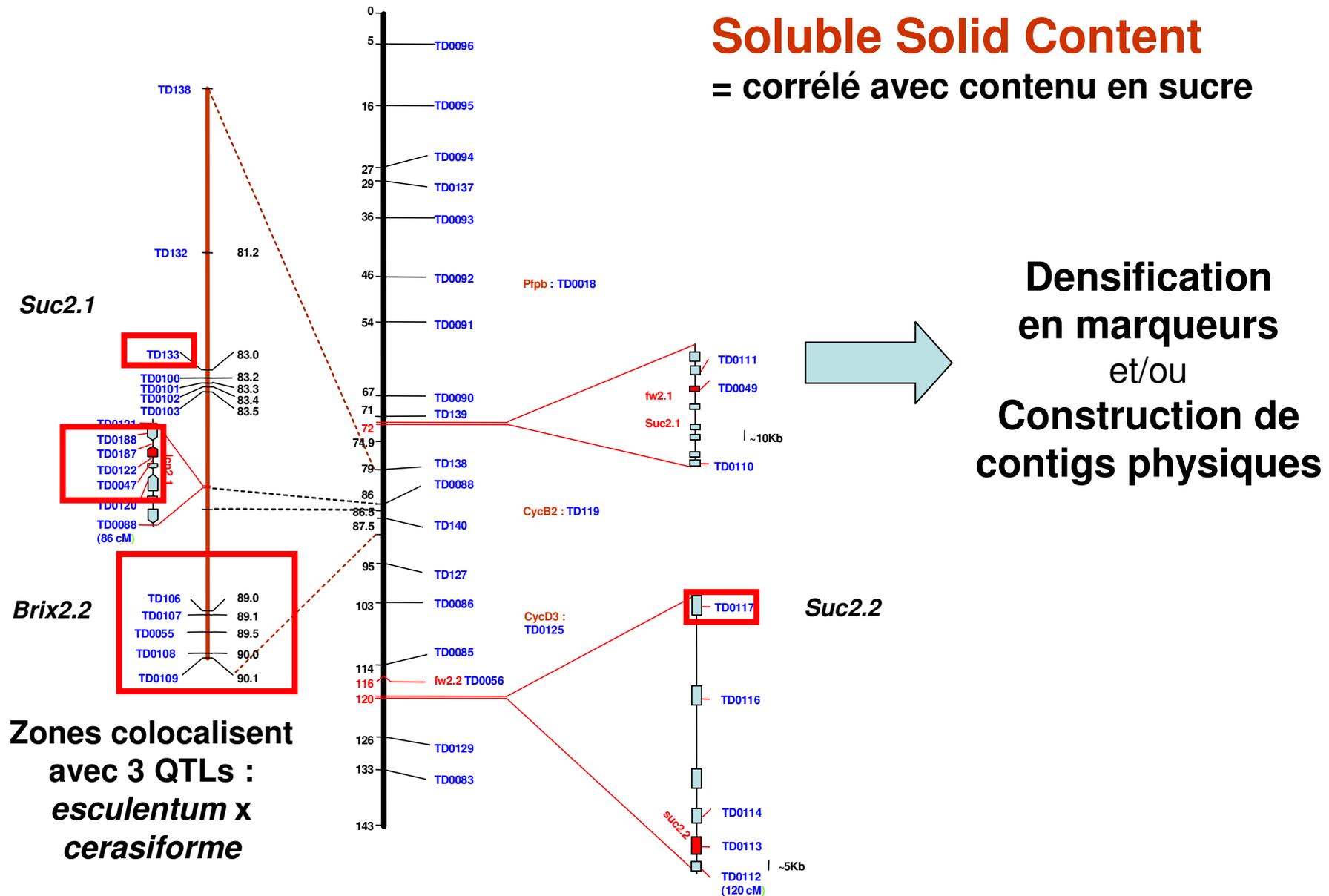
GLM sur 92 accessions, phenotypes = log(poids)



\* <0,01 ; \*\* <0,005 ; \*\*\* <0,001

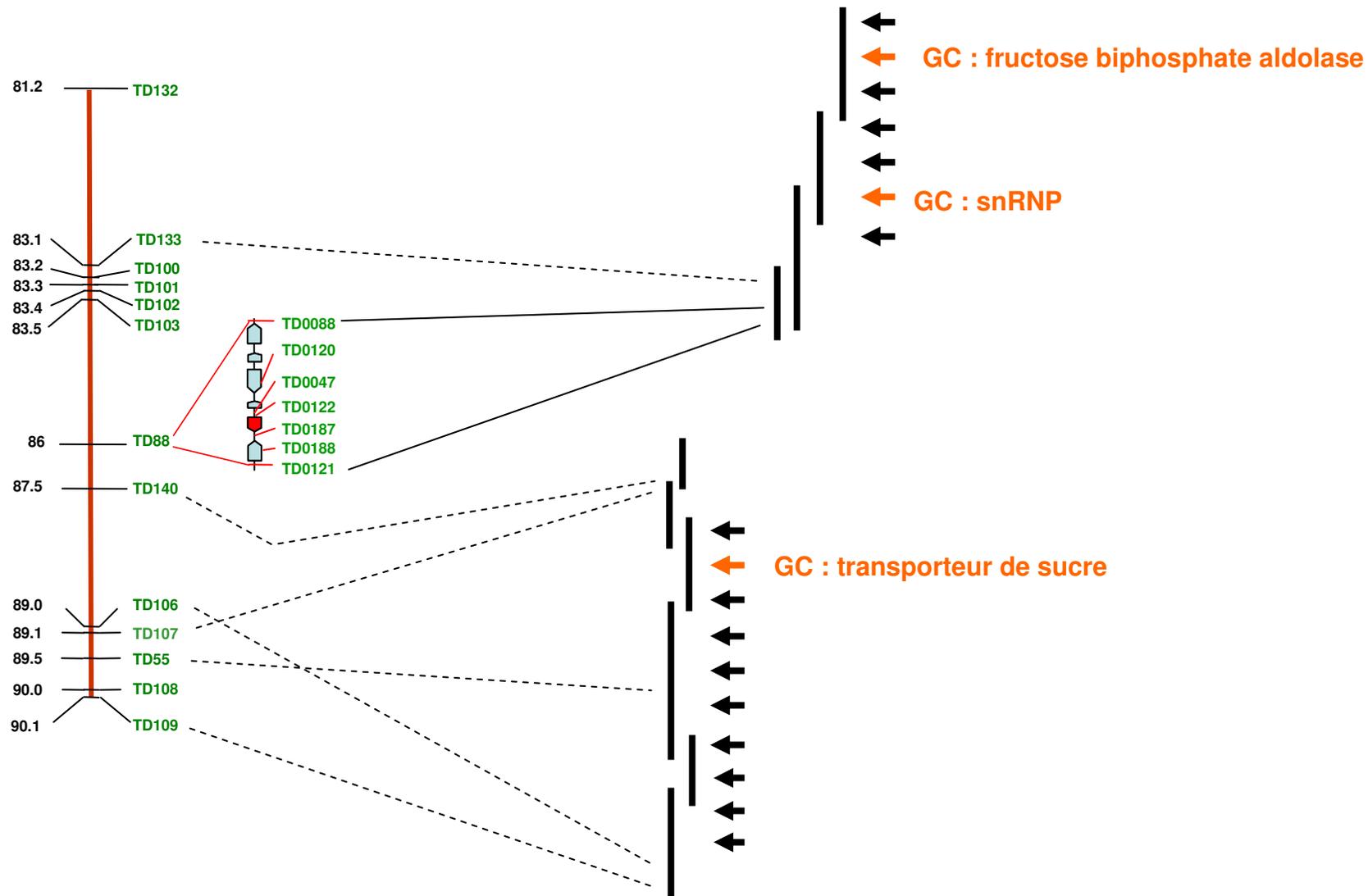
# Analyse d'association sur chromosome 2

MLM : sur 92 accessions avec log(poids) en covariable



# Analyse d'association sur chromosome 2

Grâce à l'avancée du séquençage du génome de la tomate

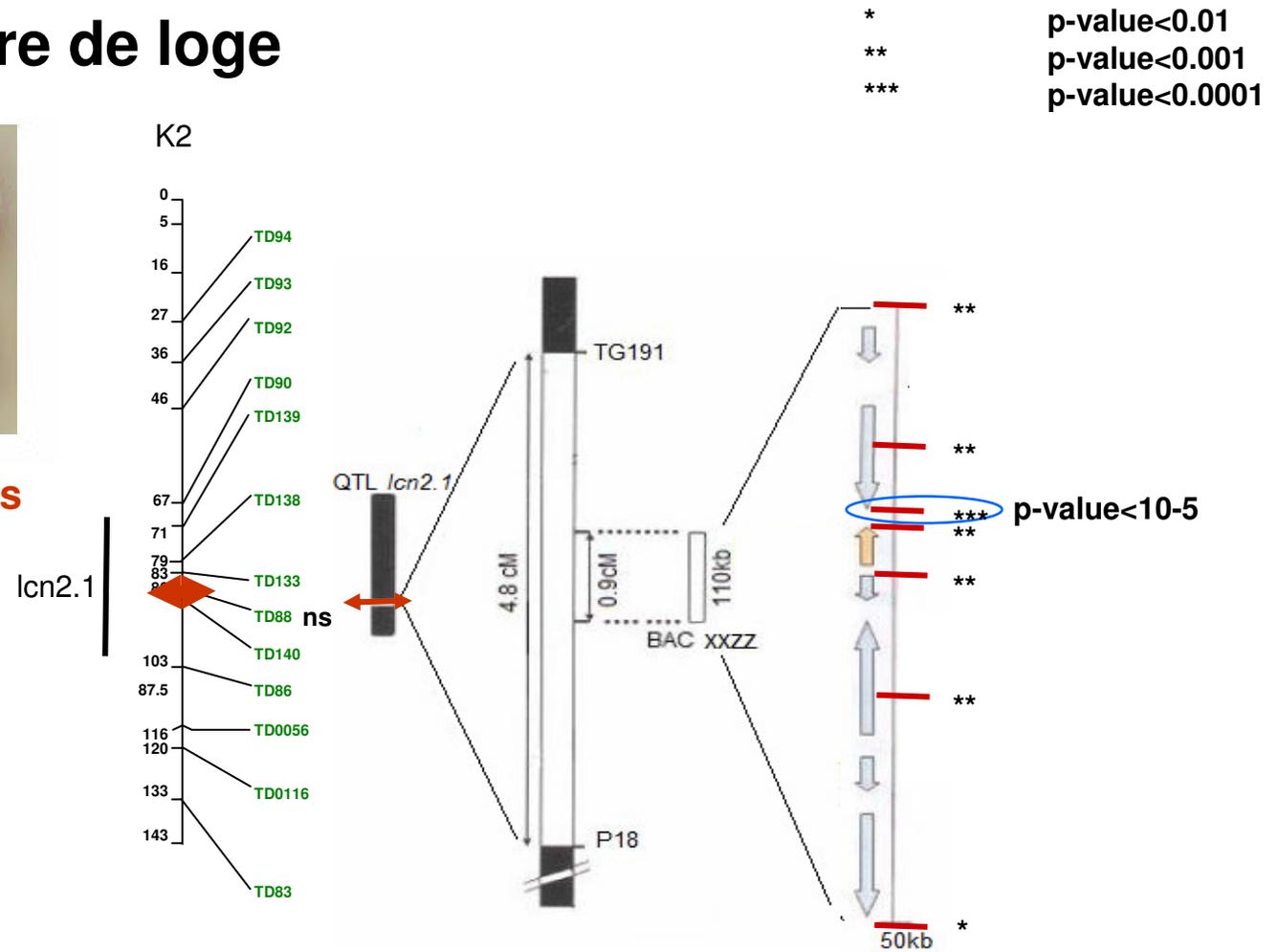


# LD mapping pour un QTL cloné – validation de l'approche

## QTL pour nombre de loge



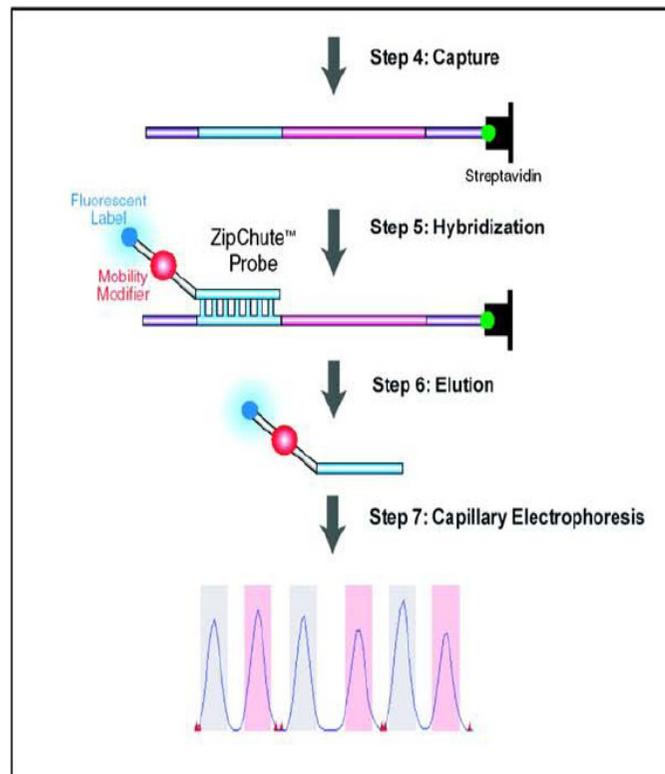
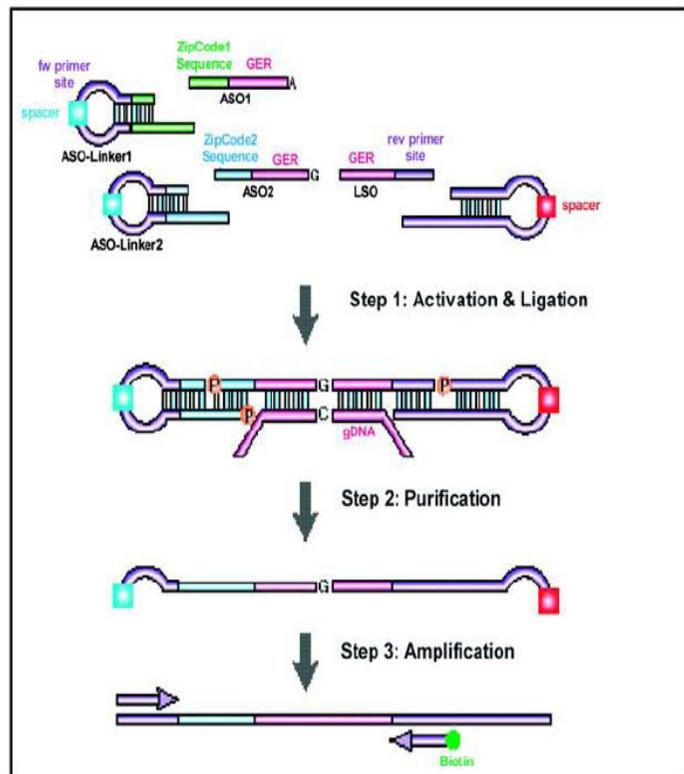
Cervil      Levovil  
**lcn2.1 : 2 vs 4 carpelles**



Clonage positionnel de lcn2.1

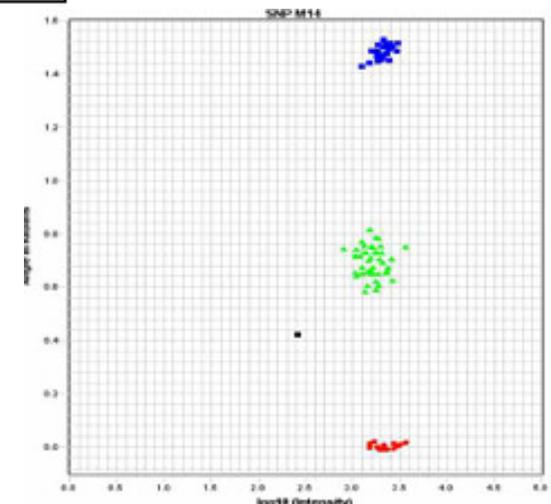
**Validation** du locus candidat (1800 bp), **réduction de la région** à deux SNPs non codants  
 Séquençage x96 = génotypage SNP + détection de trace de sélection

# Génotypage SNPlex™



**4 panels dessinés = 192 SNPs**  
**génotypés sur 360 accessions** (prévision de 300  
supplémentaires : cc Zamir)

**SNPs répartis sur génome et GC**  
**2 en cours d'analyse**  
**2 en commande**



# GnpSNP

Quick search:



Welcome nranc

- Home
- GBrowse

- Projects
- Thematics
- Taxons
- Lines
- Batches
- Sequences
- SNPs
- SubSNPs
- Contacts

- User guide
- FAQ
- GnpSNP news
- Release notes

- About
- Data submission
- GnpIS
- GnpArray
- GnpGenome
- GnpMap
- GnpSeq
- SIREgal

## Genotype card



### Information

Display  results per page  
 23 items found, displaying 1 to 10  
 <<< 1 2 3 >>>

### Results

[\[Help\]](#)

The batch [TD001](#) contains 12 markers.

Positions		60	409	423	481	543	557	570	734	742	745	756	777
Ref. Sequence	<a href="#">Heinz 1706</a>	C	A	T	IN	T	C	C	T	T	C	T	C
Line	<a href="#">LA1840</a>	G	T	C	TT	C	T	A	A	A	G	A	A
Line	<a href="#">LA1464</a>	C	A	T	--	T	A	T	G	G	T	G	G
Line	<a href="#">LA716</a>	C	T	T	TT	C	A	A	G	A	G	A	G
Line	<a href="#">LA1401</a>	C	T	T	--	T	A	T	G	G	T	G	G
Line	<a href="#">Wva 700</a>	C	A	T	--	T	A	T	G	G	T	G	G
Line	<a href="#">LA0147</a>	C	A	T	--	T	A	T	G	G	T	G	G
Line	<a href="#">TinyTim</a>	C	A	T	--	T	A	T	G	G	T	G	G
Line	<a href="#">Vir 1011</a>	C	A	T	--	T	A	T	G	G	T	G	G
Line	<a href="#">Orange cocktail</a>	C	T	T	--	T	A	T	G	G	T	G	G
Line	<a href="#">Cervil</a>	C	A	T	--	T	A	T	G	G	T	G	G

**Soumission**  
**dans la base**  
**de données**  
**GnpSNP**

# Merci à



**INRA Avignon:**

**Mathilde Causse  
Stéphane Muñoz  
Sophie Rolland  
Hélène Burck  
Yolande Carretero  
Esther Pelpoir**

**EPGV-UR1279**

**INRA Evry:**



**Dominique Brunel  
Marie-Christine Le Paslier  
Aurélie Bérard  
Aurélie Chauveau  
Rémi Bounon**



**UMR DIVERSITE ET ADAPTATION DES PLANTES CULTIVEES  
UMR DIVERSITE ET ADAPTATION DES PLANTES CULTIVEES**

**INRA Montpellier:**

**Sylvain Santoni  
Jacques David**



**INRA Versailles:**

**Nacer Mohelibi  
Fabienne Granier**



**eusol**

