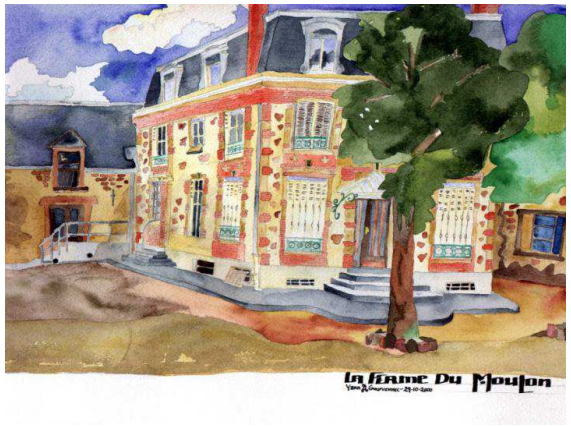
Connected biparental populations in European maize: insights into allelic series and heterosis

L. Moreau, G. Blanc, A. Larièpe, N. Bardol, B. Mangin, S. Jasson, A. Gallais, A. Charcosset

Station de génétique végétale du Moulon,
INRA, UPS, CNRS, INAPG

Multi-parent populations workshop June 12-13, 2013, NIAB

Laurence Moreau



Le Moulon Génétique végétale UMR

INRA

Institut National de la Recherche Agronomique

Research objectives

Genetic diversity characterisation

Determinism of traits of interest:
biomass and grain yield, flowering time, nitrogen use

Optimisation of breeding methods,
Using molecular diversity information

Dilemma for the talk in the context of the meeting …

We have worked (so far) on

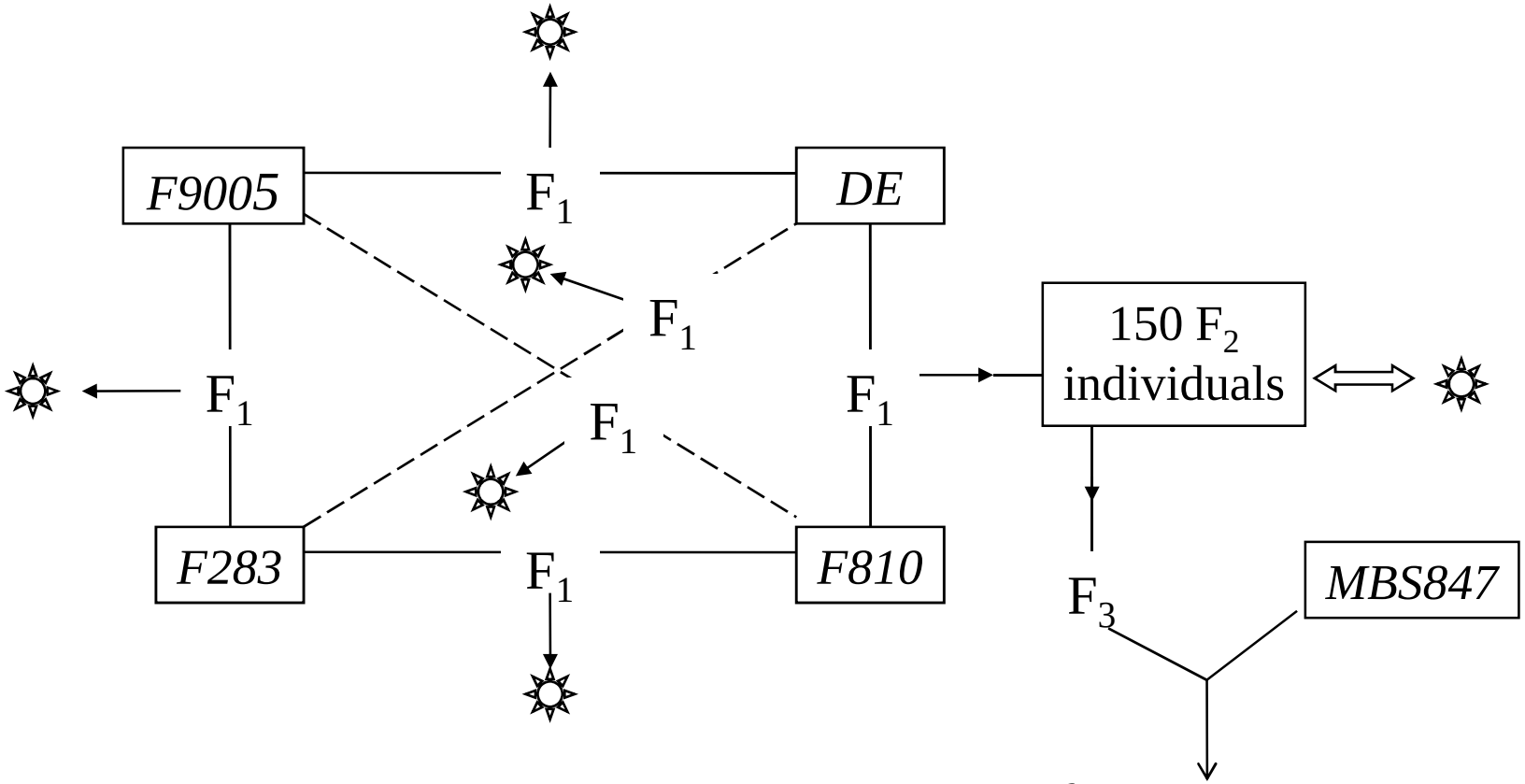
- ✓ a population intermated over several generations, compared to a « classical » version, but starting from a single biparental cross (Huang et al., 2010, Genetics)
- ✓ multiparental designs but with no multi generation intermating …

Outline

1. Use of a multiparental mating design for marker assisted selection in maize
2. Evaluation of parental allele clustering
3. Multiparental QTL mapping insights into heterosis

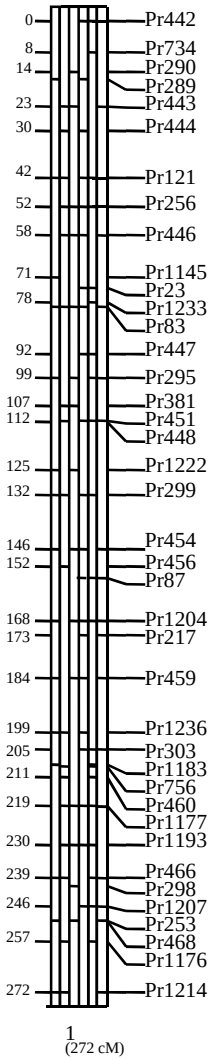
Use of a multiparental mating design for marker assisted selection in maize (Blanc et al., 2006, TAG)

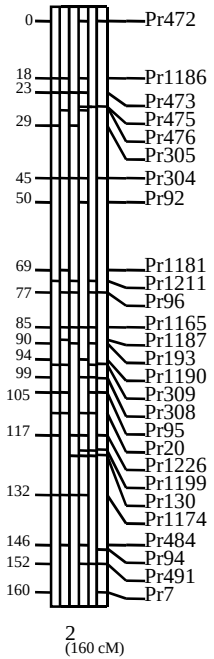
Four parental lines (early flint)

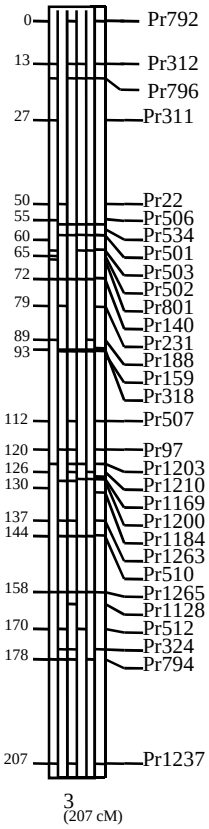


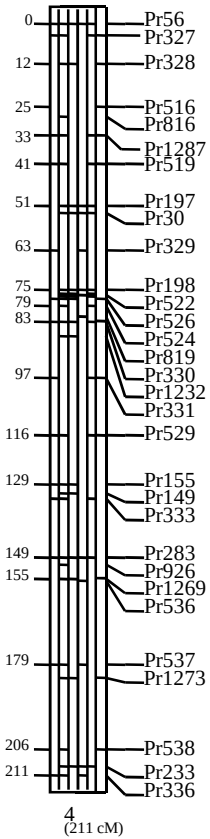
Two years of experimentation of hybrids (10 trials) for grain yield, moisture, flowering time

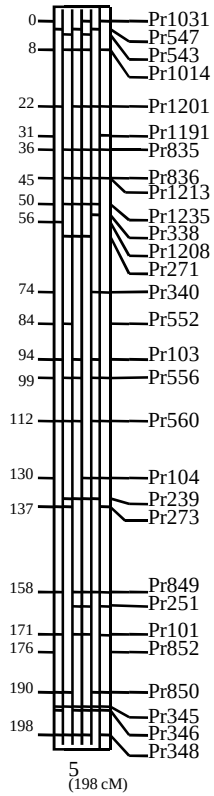
Consensus map











272 microsatellite markers. Polymorphism revealed on agarose metaphor gels

Monopopulation maps : on average 1663 cM, 12,4 cM between two markers

Consensus map : 1794 cM, 7 cM on average between markers

Global QTL analysis (MC QTL software, Jourjon, 2005, Bioinformatics)

Model 1: QTL effects assumed independant across populations

Performances (N x 1)

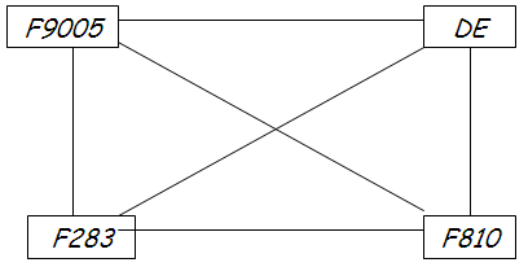
$$\mathbf{Y} = \mathbf{JM} + \mathbf{X_q A_q} + \sum_{c \neq q} \mathbf{X_c A_c} + \boldsymbol{\varepsilon}$$

- $\mathbf{JM}$: Population effects, p -1=5 df
- $\mathbf{X_q A_q}$: QTL effects, p=**6 df**
- $\sum_{c \neq q} \mathbf{X_c A_c}$: Covariates effects
- $\boldsymbol{\varepsilon}$: Random residual

Model 2: allele effects forced to be consistent across populations

(3 df for QTL effect instead of 6)

Rk. Difference between models 1 and 2 makes it possible to test QTL x genetic background epistatic effects, See Jannink and Jansen 2001



Number of QTL detected, average confidence intervals (CI), and percentage of variance explained (R^2) in the different analyses (same global type I risk).

Analyses	Silking Date			Grain Moisture			Grain Yield			Index		
	Nb of QTL	CI	R^2	Nb of QTL	CI	R^2	Nb of QTL	CI	R^2	Nb of QTL	CI	R^2
Single-population *model*	2.2[a] (8[b])	32	29.6	1.8[a] (7[b])	30	28.9	1.3[a] (7[b])	49	25.9	0.8[a] (5[b])	44	13
Multipop dis-connected *model (1)*	8	25 (23[c])	64.3	8	29 (26[c])	52.2	5	(				
Multipop connected *model (2)*	11	28 (16[d])	66.0	13	20 (14[d])	57.9	12	(17[d])			(23[d])	

Joint connected model:
-Increased Nb of QTL
-Reduced CI
-Increased R^2

[a] average number of QTL detected per population
[b] total number of regions detected by all single-population analyses (model(1))
[c] average CI of the QTL also detected in the single-population analyses (model(1))
[d] average CI of the QTL also detected in the multipopulation disconnected analyses (model(2))

(Blanc et al., TAG, 2006)

Global comparison of allelic effects for each QTL (flowering time, days)

QTL	Ch.	Position (in cM)	CI	r^2	Estimated additive effect				Class effect	epistasis with (5%)	Gback
					DE	*F283*	*F9005*	*F810*			
1	1	46	38 - 56	0.06	0.22^{a}	0.03^{a}	0.21^{a}	-0.46^{b}	2	10, 11	
2	1	140	134 - 166	0.06	0.01^{a}	-0.42^{c}	0.06^{a}	0.35^{b}	3	11	
3	2	85	62 - 89	0.07	-0.48^{a}	0.19^{bc}	0.31^{c}	-0.03^{b}	3	8	
4	3	41	33 - 50	0.08	0.27^{a}	-0.51^{b}	-0.19^{b}	0.43^{a}	2	11	
5[ß]	3	150	139 - 188	0.04	-0.07^{a}	0.11^{a}	-0.3^{b}	0.26^{c}	3	-	
6[ß]	4	75	45 - 97	0.02	-0.24^{a}	0.09^{bc}	0.00^{b}	0.15^{c}	3	-	
7[αß]	5	26	10 - 38	0.02	0.06^{a}	-0.29^{b}	0.10^{a}	0.12^{a}	2	-	
8[ß]	6	25	2 - 31	0.04	0.13^{a}	0.08^{a}	-0.37^{b}	0.17^{a}	2	3	
9	7	145	135 - 167	0.04	-0.03^{a}	-0.36^{b}	0.15^{a}	0.25^{c}	3	-	
10	8	58	47 - 65	0.05	-0.23^{a}	-0.03^{a}	0.40^{b}	-0.15^{a}	2	1	
11	10	30	28 - 32	0.18	-0.34^{a}	0.87^{b}	-0.29^{a}	-0.25^{a}	3	1, 2, 4	**

Early flowering alleles

- *Several QTL with three classes of allelic effects (consistent with Buckler et al. 2009)*
- *Contribution of epistasis seems limited*

Yield (t ha^{-1})

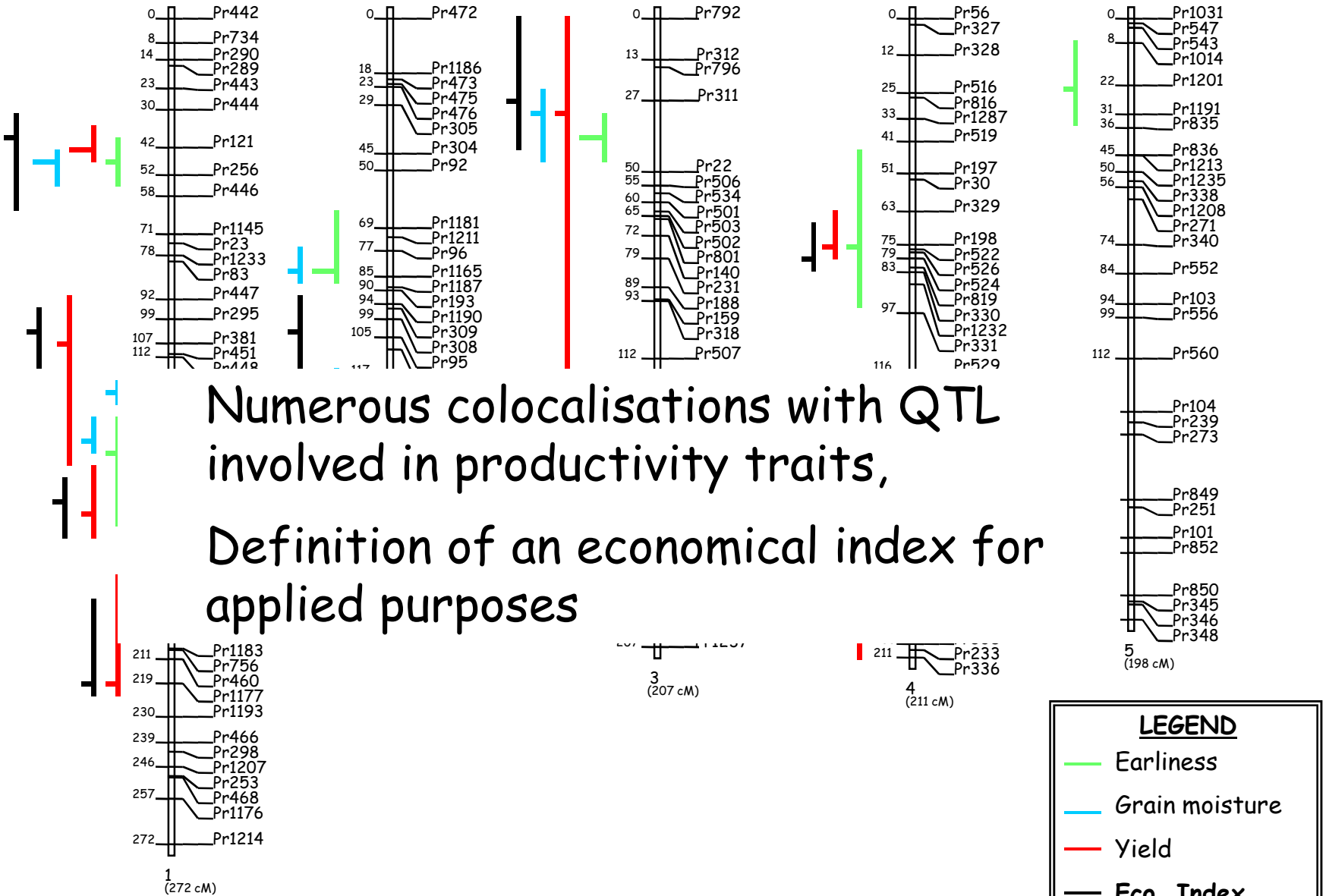
N°	chr	pos	Parental alleles DE	F283	F9005	F810	Nb of class	QTL x QTL	QTL x Backgr.
1	1	44	**0.099** ^{a}	**0.114** ^{a}	-0.017 ^{b}	-0.195 ^{c}	3	3, 11	*
2	1	105	**0.102** ^{a}	-0.086 ^{b}	0.017 ^{c}	-0.033 ^{bc}	3	7, 11	
3	1	160	**0.067** ^{a}	-0.082 ^{b}	-0.083 ^{b}	**0.098** ^{a}	2	1, 7	
4	1	217	**0.049** ^{a}	**0.057** ^{a}	**-0.006** ^{a}	-0.101 ^{b}	2	10, 11, 12	*
5	3	35	**0.039** ^{a}	**0.001** ^{a}	-0.094 ^{b}	**0.055** ^{a}	2	-	
6	4	79	-0.083 ^{a}	0.015 ^{b}	-0.028 ^{ab}	**0.096** ^{c}	3	7, 11, 12	
7	4	164	-0.045 ^{a}	-0.007 ^{a}	**0.103** ^{b}	-0.052 ^{a}	2	2, 3, 6, 10, 11	**
8	6	23	-0.021 ^{a}	**0.094** ^{b}	-0.087 ^{c}	0.014 ^{a}	2	11	
9	7	139	-0.057 ^{a}	-0.057 ^{a}	**0.041** ^{b}	**0.073** ^{b}	2	-	
10	8	33	-0.032 ^{a}	-0.040 ^{a}	**0.073** ^{b}	0.001 ^{a}	2	4, 7	
11	9	75	-0.020 ^{a}	-0.025 ^{a}	-0.054 ^{a}	**0.099** ^{b}	2	1, 2, 4, 6, 7, 8, 12	*
12	10	2	-0.021 ^{a}	**0.088** ^{b}	-0.063 ^{a}	0.003 ^{a}	3	4, 6, 11	*

Most productive alleles

(Blanc et al., TAG, 2006)

Epistasis more important

Colocalisation between QTL detected for flowering time and other traits of interest: grain moisture, yield



Numerous colocalisations with QTL involved in productivity traits,

Definition of an economical index for applied purposes

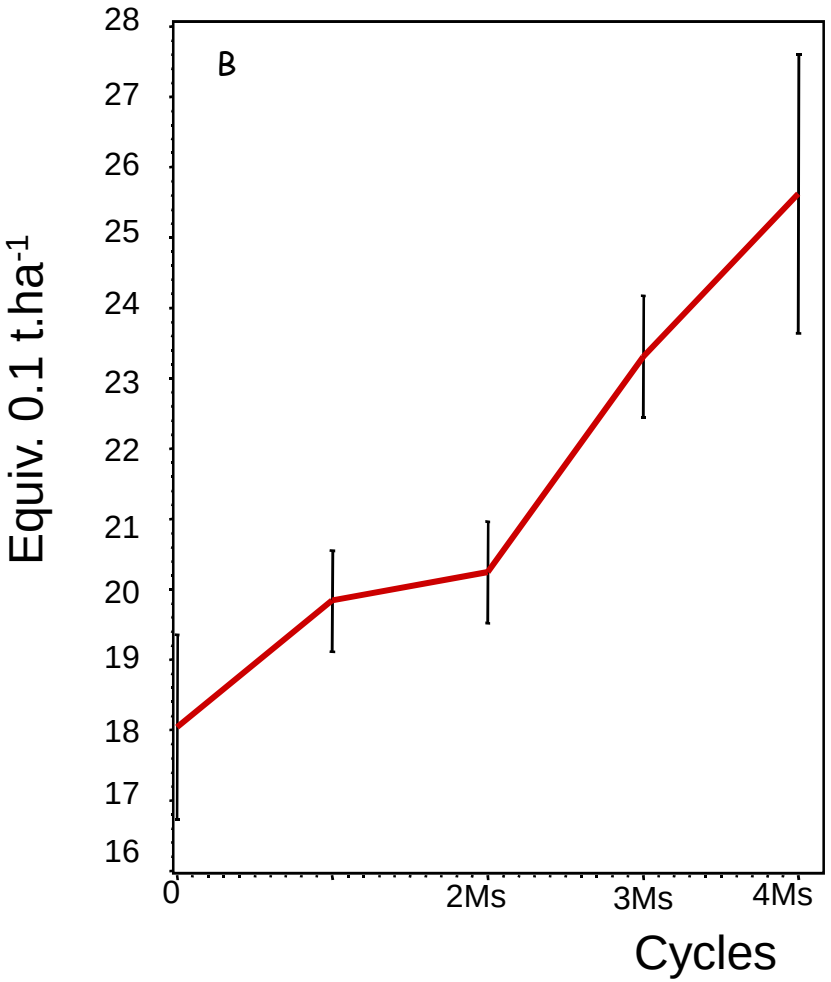
LEGEND

- Earliness
- Grain moisture
- Yield
- **Eco. Index**

Outcome of marker assisted selection program for yield index (evaluation 2004-2005-2007 trials)

- Evolution of expected frequency of favorable alleles from to 0.28 (parents) to 0.80

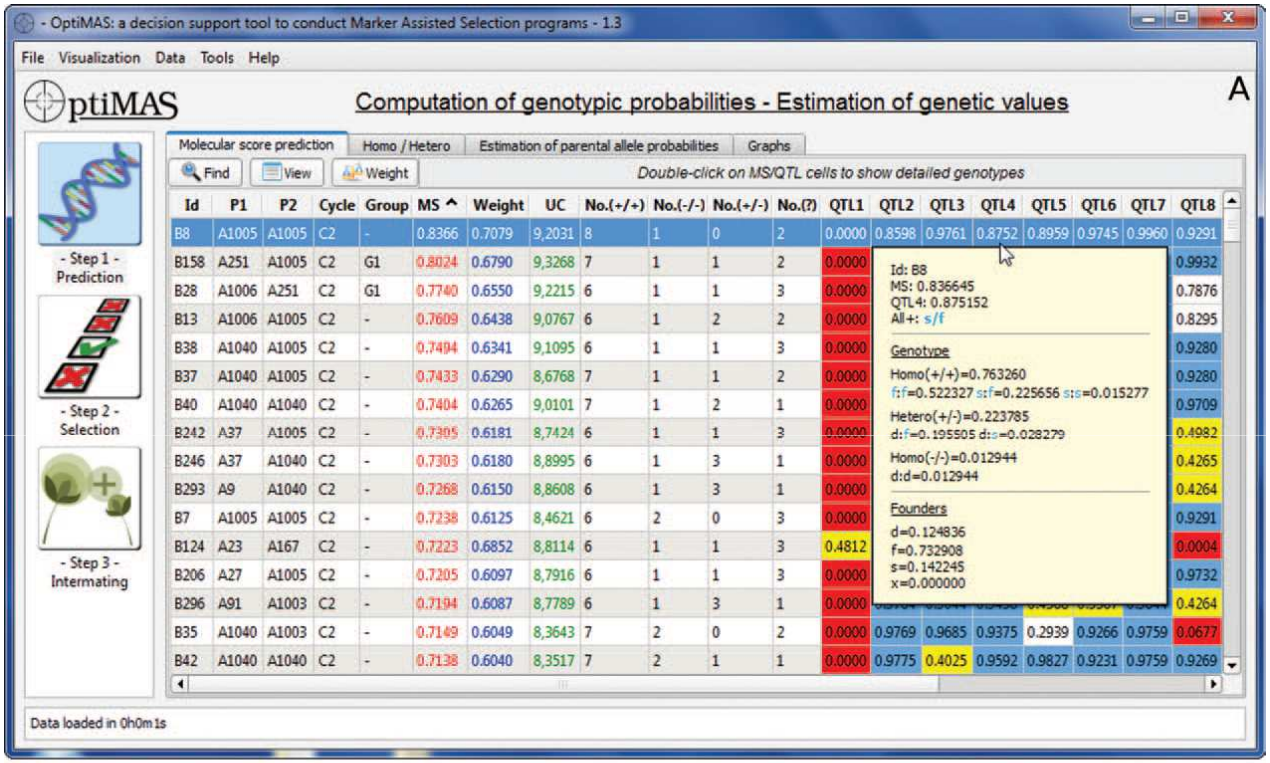
- M cycles efficient: significant genetic gain (50% of expected gain)

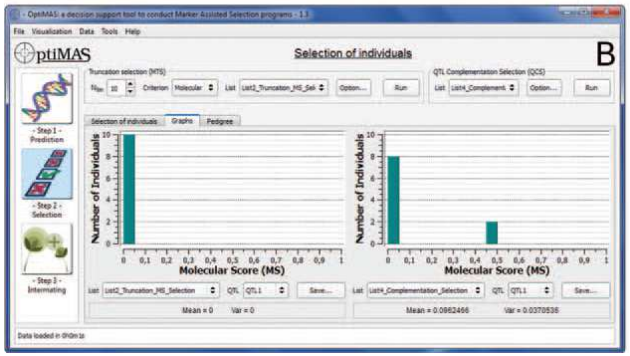


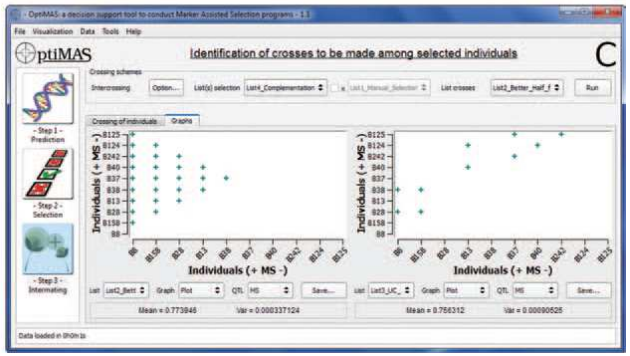
High performance of inbred lines fixed from last cycles confirmed

Multiparental designs have a great potential for marker assisted breeding (assembly of alleles more challenging)

Development of Optimas, an informatic tool to facilitate the allele assembly process following multiparental designs (Valente et al., 2013, J. of Hered.)





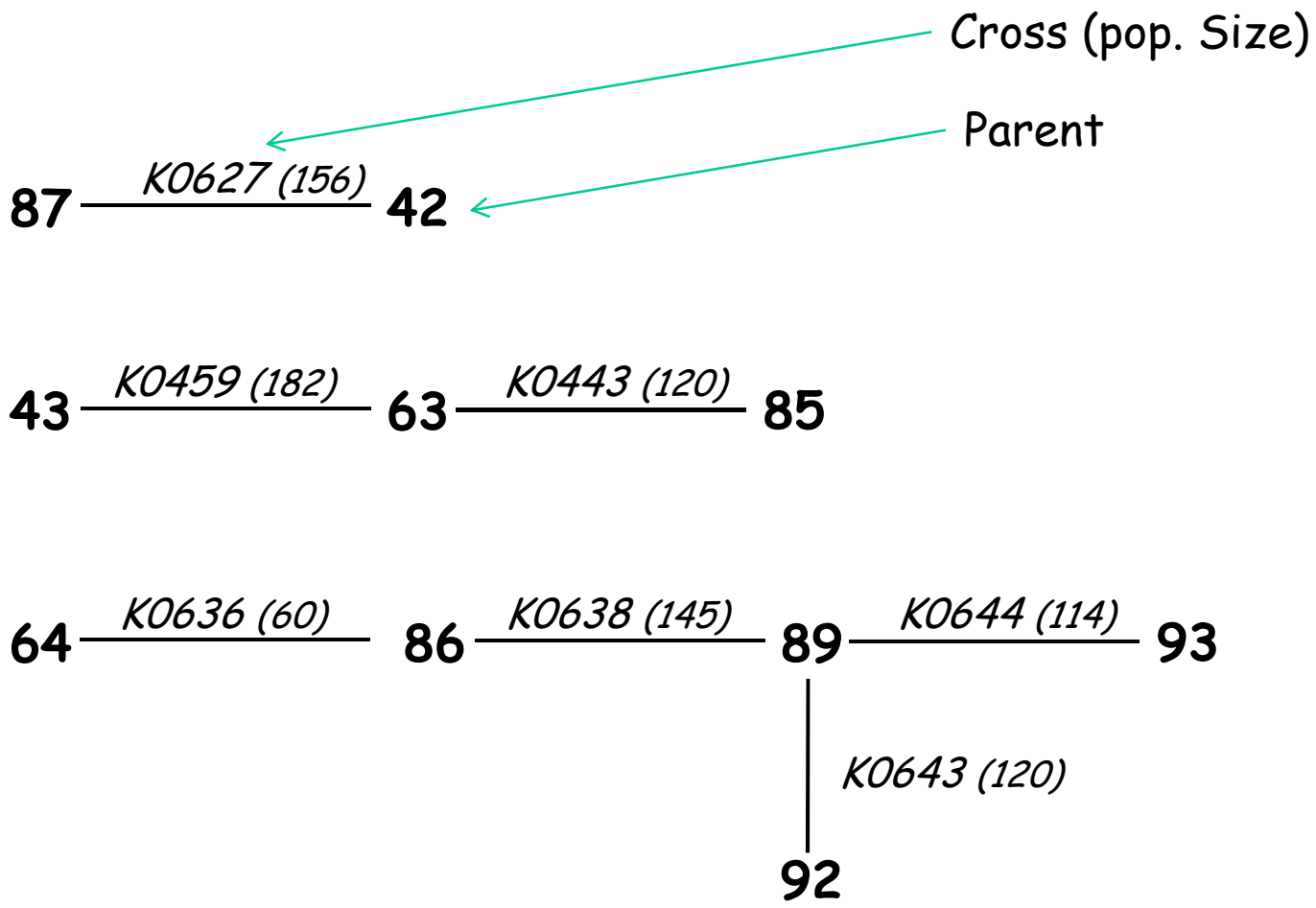


https://www.integratedbreeding.net/

1. Use of a multiparental mating design for marker assisted selection in maize
2. Evaluation of parental allele clustering
3. Multiparental QTL mapping insights into heterosis

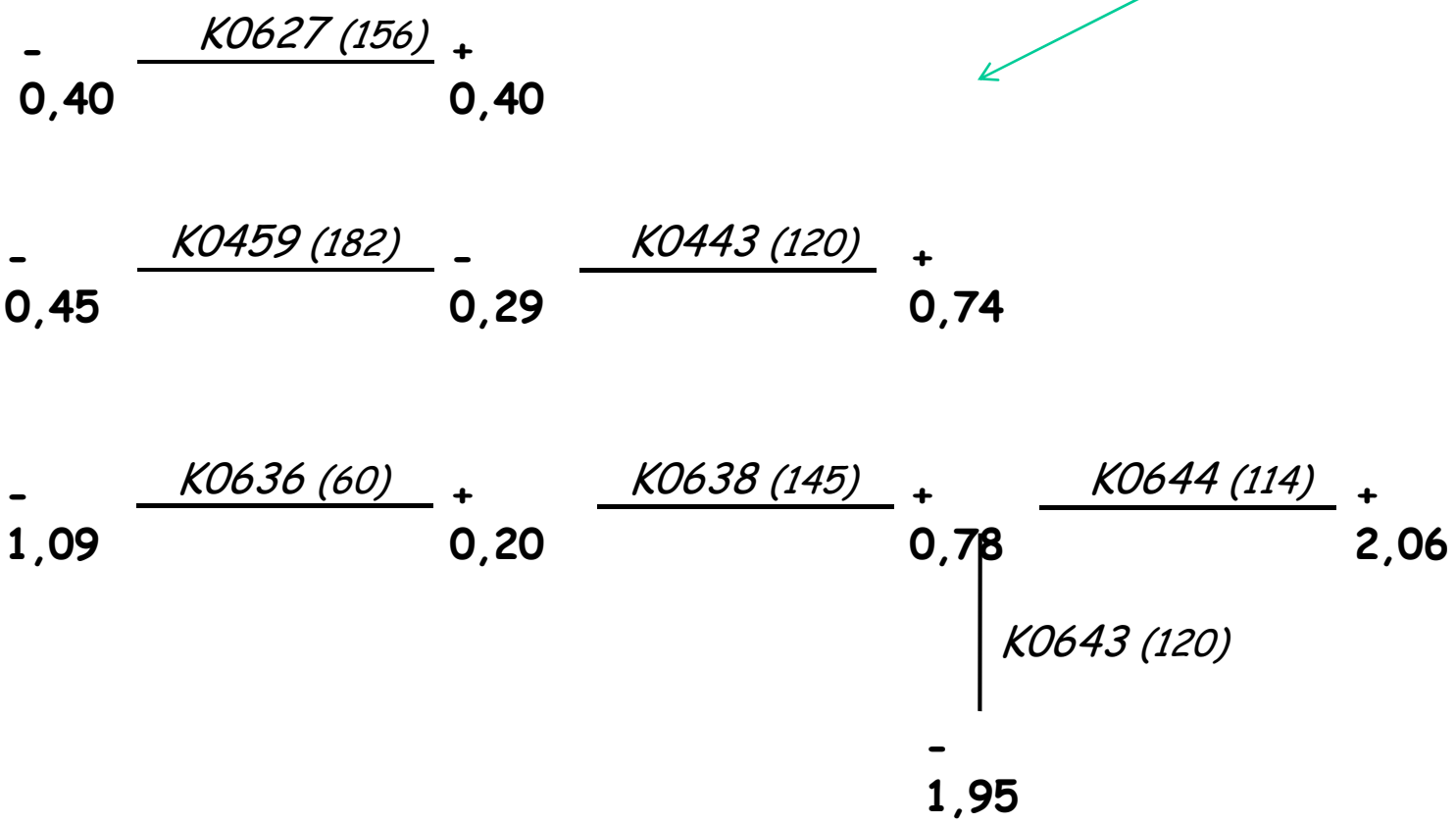
Populations used in applied breeding programs may only be partially connected

Ex: Euralis design (N. Bardol, 2013, TAG)

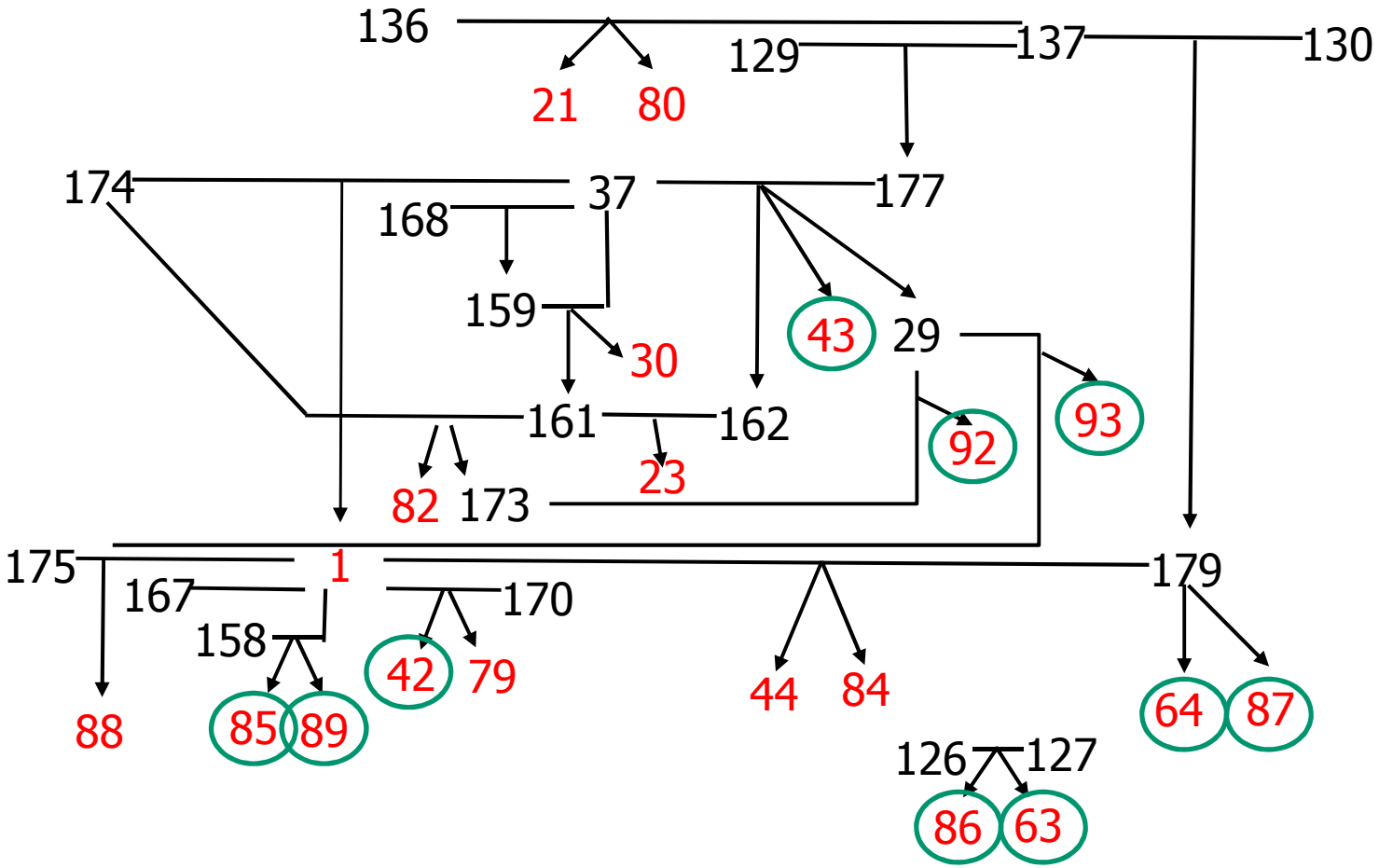


All alleles cannot be compared -> problem to predict genetic values beyond one generation of intermating

Ex. of one QTL analysed with previous models



Parental lines are related and therefore share identical by descent chromosomal segments



-> actual number of alleles is less than number of parents

The linkage disequilibrium – Linkage Analysis concept (LD-LA),

(Meuwissen et Goddard, 2000; Jannink et Wu, 2003)

Thanks to dense genotyping of parents:

- identification of lines carrying same alleles
- reduction in number of parameters to be estimated

May increase connexion and/or increase power of QTL detection

Should we use directly genotypes at individual loci or haplotypes?

Clustering of parental alleles into haplotypes with "Clusthaplo" (Leroux et al. in review)

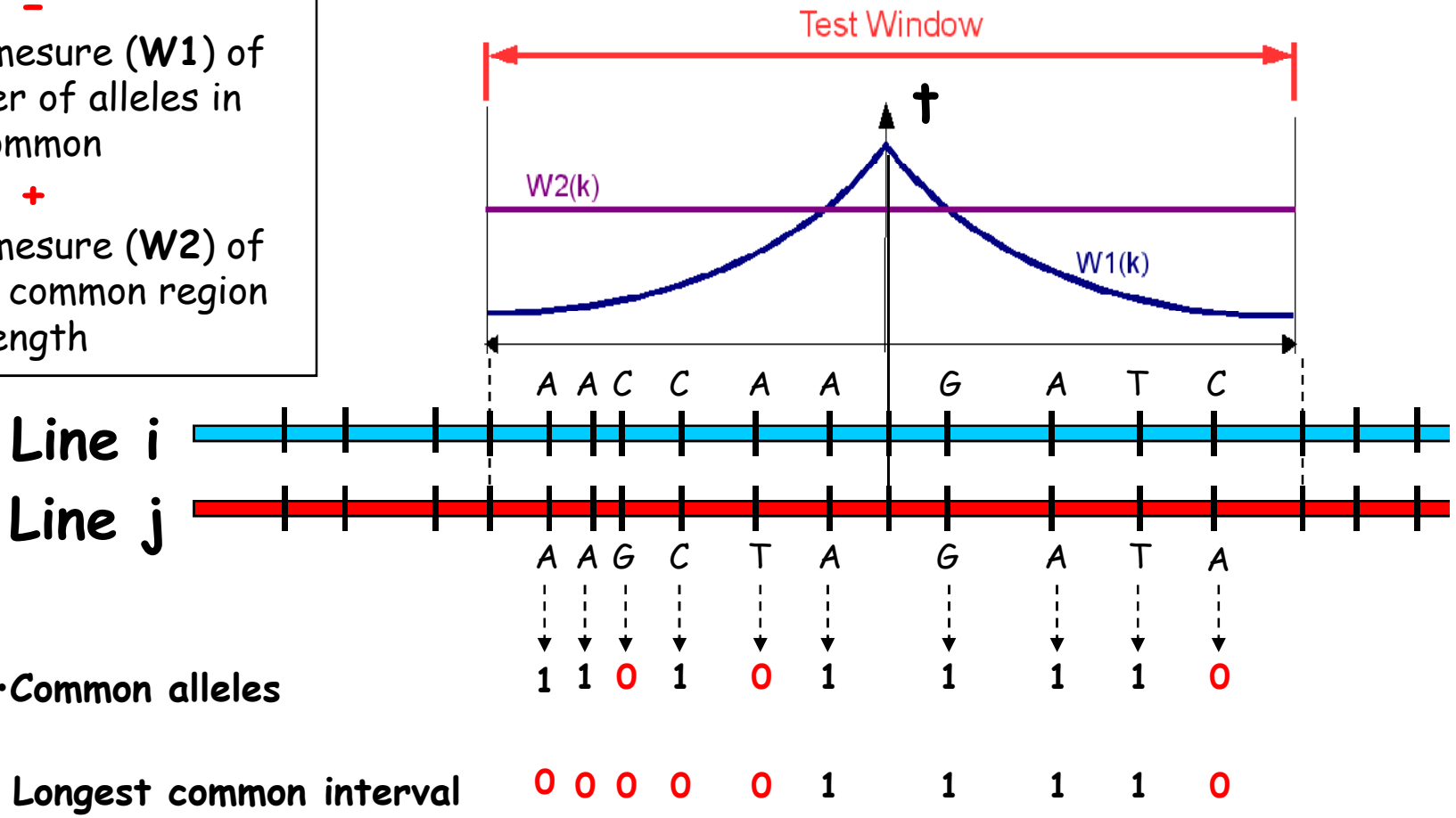
Similarity score

=

Weighted mesure (**W1**) of the number of alleles in common

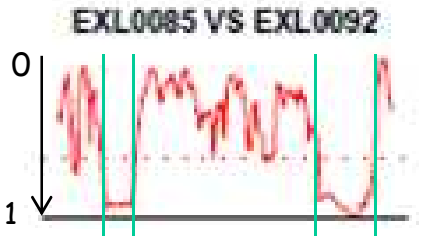
+

Weighted mesure (**W2**) of the longest common region length

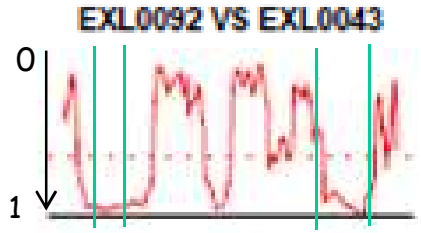


- Common alleles
- Longest common interval

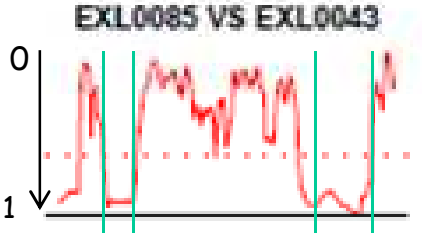
Application to 50KSNP genotyping of parents of Euralis design



Ex. of local similarities between three inbred lines along a chromosome



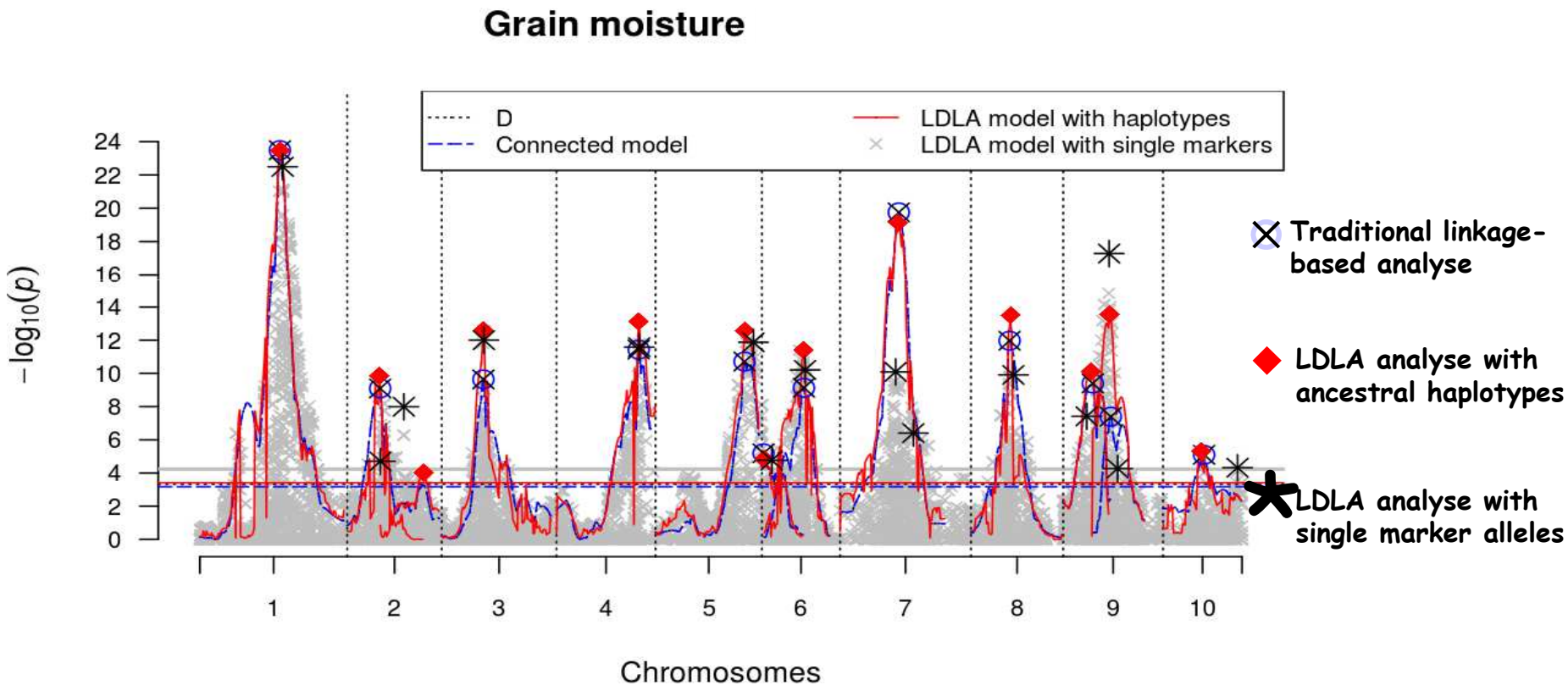
-> Threshold and transity closure rules for clustering



-> three possible definitions for alleles in QTL models

- One allele per parent
- Haplotype grouping
- Single marker grouping

QTL analysis with MCQTL-LD
Coll. BIA Toulouse « Syngenta » « Euralis »



Models	Silking date(days)		Grain moisture(%)	
	Nb QTL	R^2	Nb QTL	R^2
Single population analyses	1^a / 3^b	15.7	2.1^a / 8^b	29.5
(Partially) Connected	5	24.5	12	53.7
LDLA with ancestral alleles	**5**	21.7	**13**	53.2
LDLA with single marker	**7**	38.8	**15**	46.1

Models	Grain yield(q.ha^{-1})		Index	
	Nb QTL	R^2	NbQTL	R^2
Single population analyses	1^a / 7^b	13.0	0.8^a / 6^b	10.0
(Partially) Connected	7	30.8	7	26.3
LDLA with ancestral alleles	**12**	37.5	**11**	30.5
LDLA with single marker	**9**	35.8	**7**	31.7

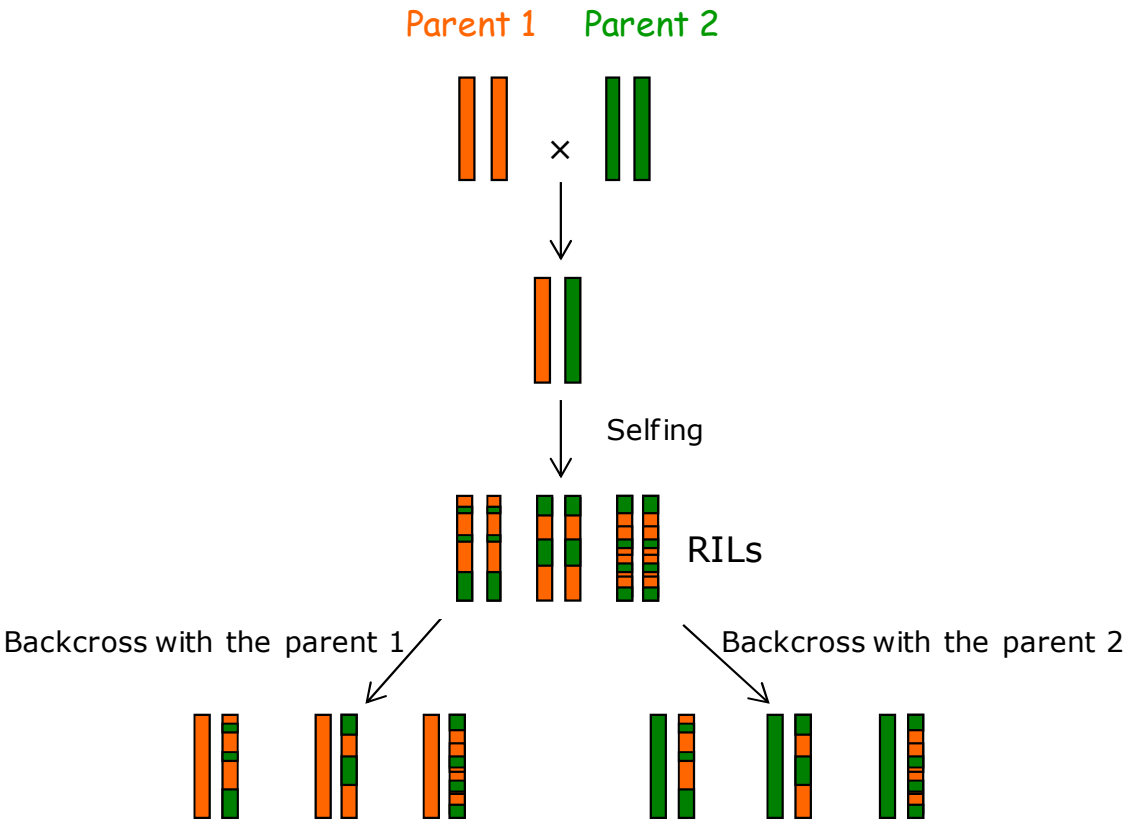
⇒ *More QTL detected on average with LDLA approaches*

⇒ *Differences between models vary among traits suggest different types of allelic effects*

1. Use of a multiparental mating design for marker assisted selection in maize
2. Evaluation of parental allele clustering
3. Multiparental QTL mapping insights into heterosis

QTL approaches of heterosis

How to analyze dominance effects using homozygous lines?

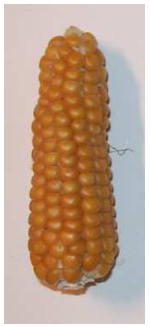


North Carolina design III

Creates homozygous vs. heterozygous segregations

Extension of the NCIII (Larièpe et al., 2012, Genetics)

Three parents

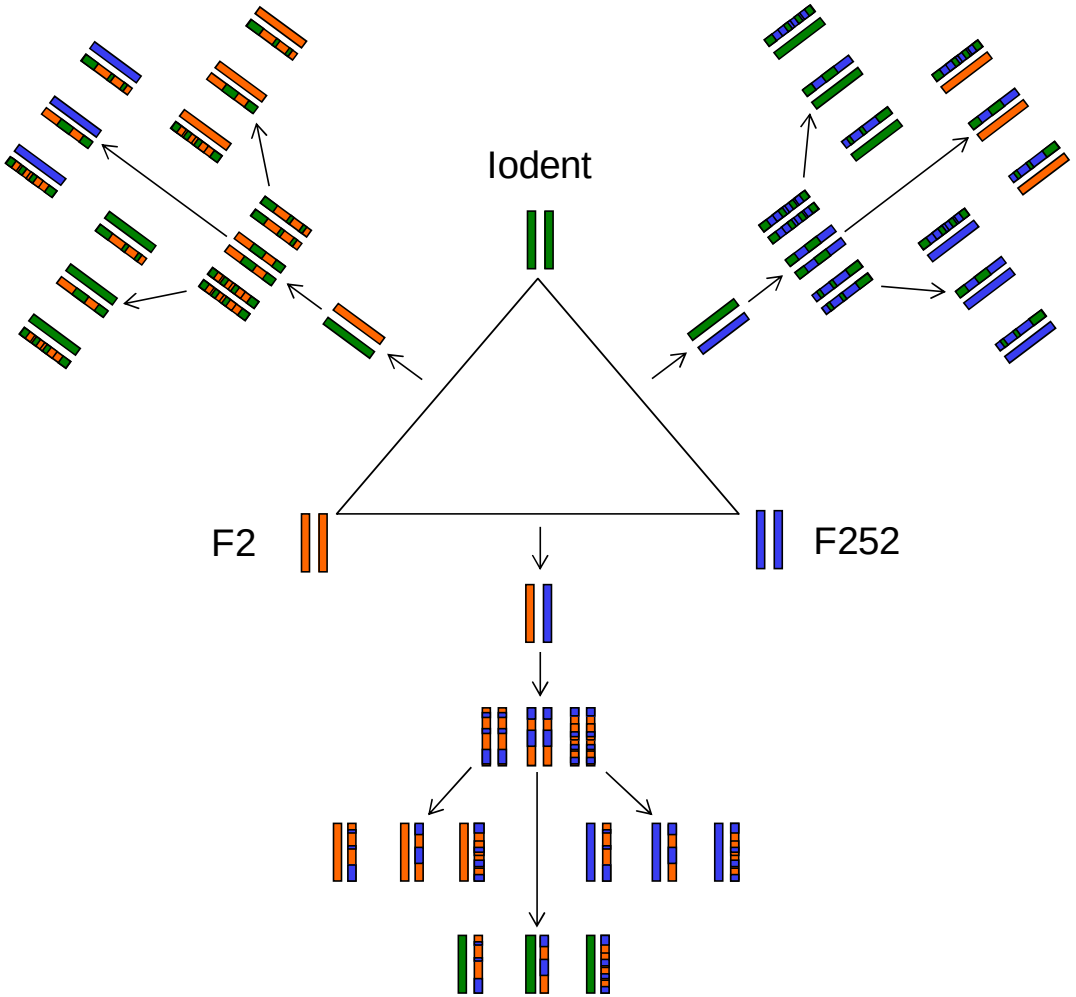


F2: European Flint

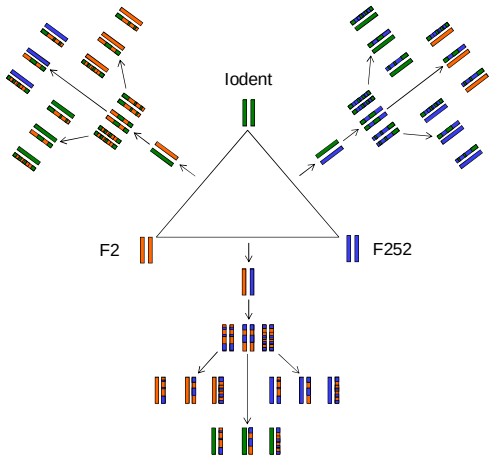
F252: American Dent

Iodent: American Dent

Cross to both related and unrelated parents



9 families of hybrids, including three between unrelated lines (bold)



Population	RILs	× F2	× Io	× F252
F2 × Io (D)	11 / 22	11 / 12	12 / 22	**13 / 23**
F2 × F252 (E)	11 / 33	11 / 13	**12 / 23**	13 / 33
Io × F252 (G)	22 / 33	**12 / 13**	22 / 23	23 / 33

Makes it possible to compare :

✓Homozygous (indirectly)

✓Homozygous / heterozygous

✓Heterozygous (directly)

✓Test for epistatic effects (6 genotypes / 9 comparisons)

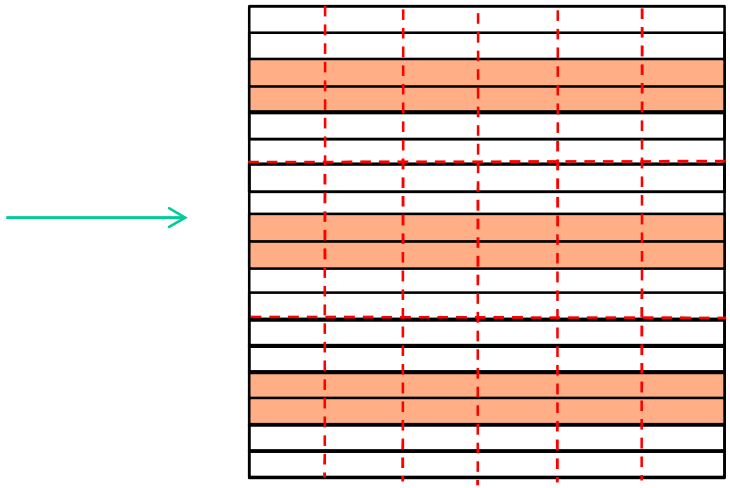
Phenotyping experiment

Materials

- 1278 Hybrids between Rils and parents
- several checks constructed from parental lines (per se, F1 and three way hybrids)

Evaluated in four environments

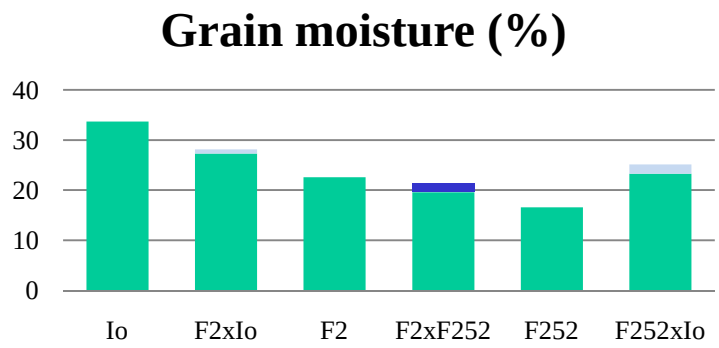
- Single replicates (two rows plots)
- Block design organised to prevent excessive competition between related and unrelated crosses (colored)
- Traits: flowering time (silking date), plant height, grain moisture at harvest, grain yield

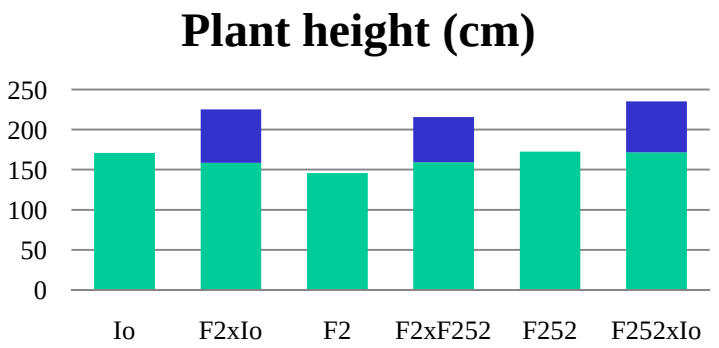


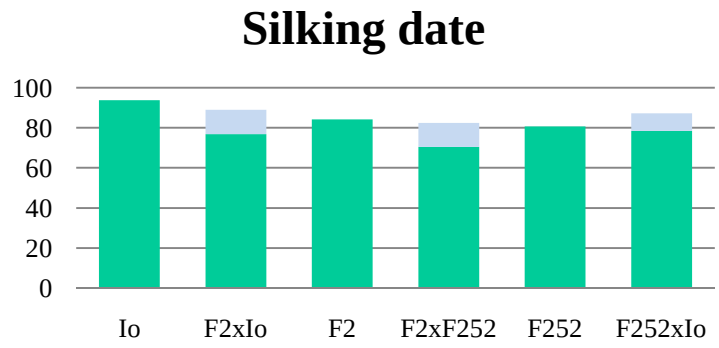
Relative magnitude of GxE limited
-> focus on average of locations

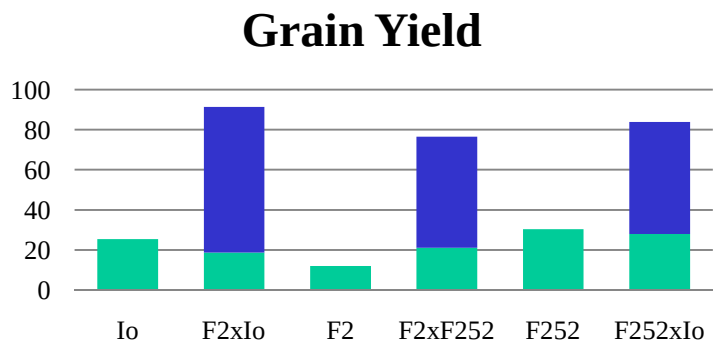
Heterosis

Dark / Light blue: positive / negative deviations of hybrids to average of parents









- ✓ *Dramatic heterosis for yield (probably enhanced by competition)*
- ✓ *Heterosis towards faster flowering*

QTL mapping models

Using McQTL (coop B. Mangin, S. Jasson), multilocus models

- ✓ Simultaneous mapping of the different populations for Z1 and Z2 (without unrelated crosses)
- ✓ global analysis of whole design with additive and dominance effects

$$y = Jm + X_q^* g_q^* + \sum_{c \neq q} X_c^* g_c^* + e,$$

y: performance
Jm: family effect (9 families)
X: genotypes, *g:* effects (3 additive, 3 dominant)

- ✓ Inclusion of epistatic terms:
  with background or beetween detected QTLs,
  additive x additive and terms including dominance

Synthesis of detected effects

Trait	QTL detection on NCIII design				QTL detection on the global design			
	QTL detected for Z_1	R^2 (%)	QTL detected for Z_2	R^2 (%)	No. QTL detected	R^2 (%)	QTL with significant additive effects	QTL with significant dominance effects
Grain moisture	7	36	3	17	13	40	13	2
Silking date	2	12	6	46	12	36	11	7
Plant height	4	26	7	44	15	44	15	8
Grain yield	2	11	9	53	10	34	6	10

- ✓ More QTL detected with global model
- ✓ All QTL detected for yield display dominance effects
- ✓ Moisture fully additive
- ✓ Height and silking intermediate

Visualisation of representative effects

Circles: homozygous Genotypes

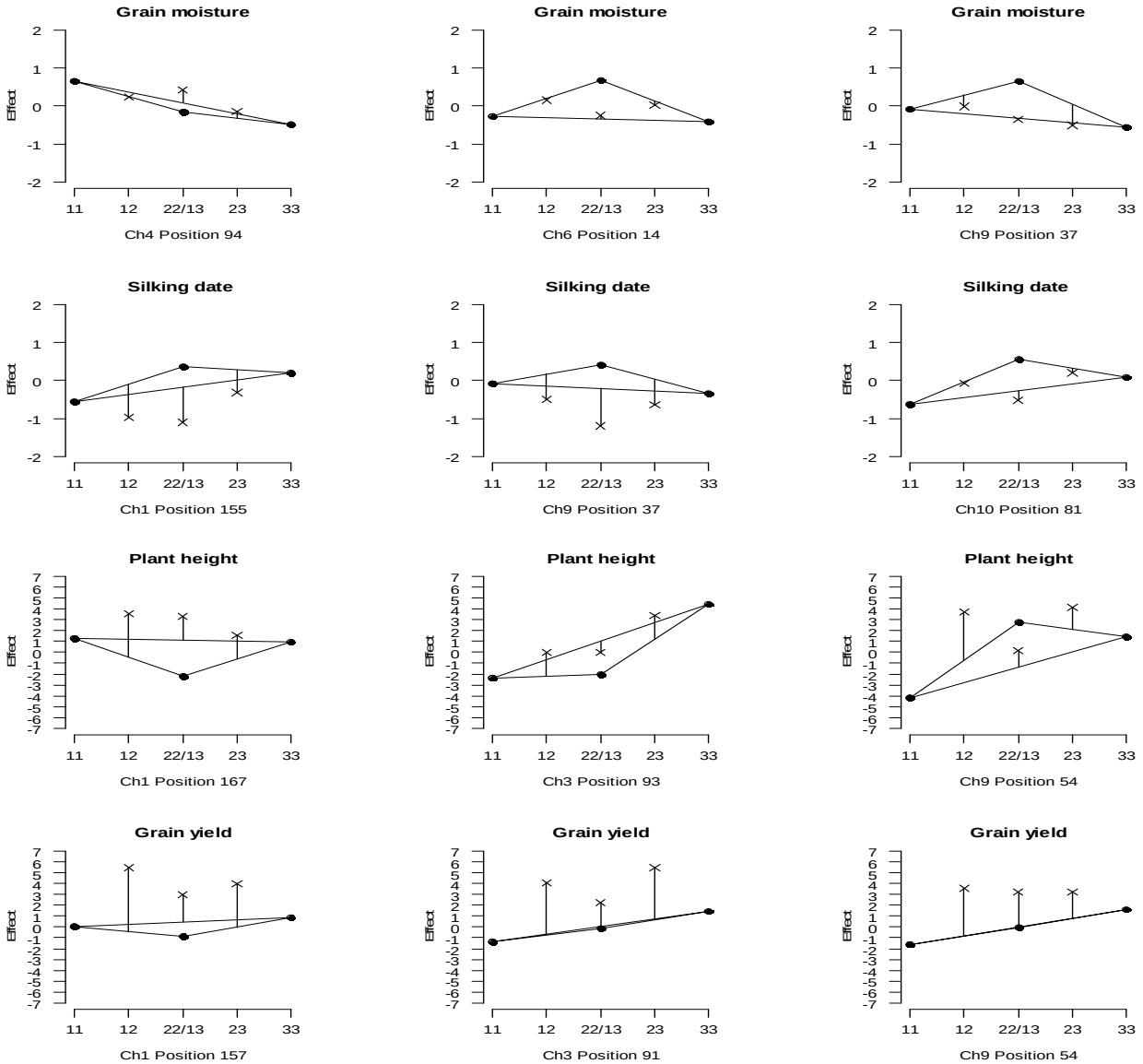
Crosses: heterozygous genotypes

Triangle sides join homozygous genotypes

Vertical lines represent dominance effect

1 indicates the F2 allele, 2 the Io allele, and 3 the F252 allele.

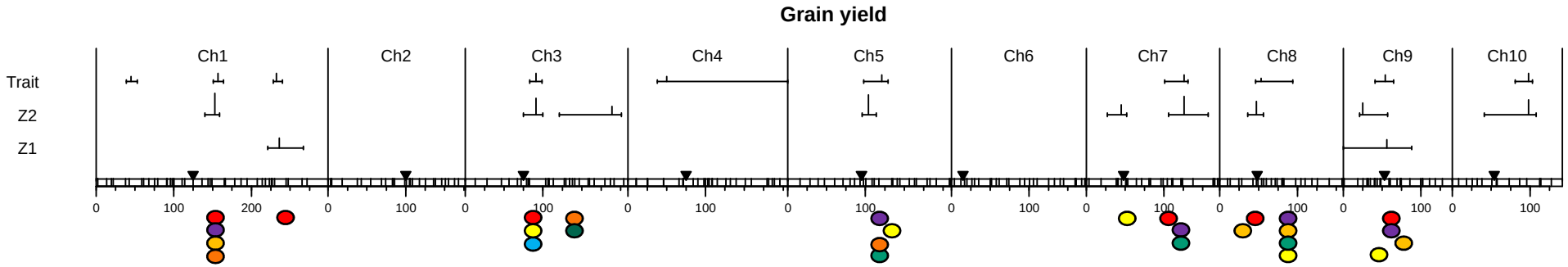
Sum of additive effects contrained to be null



All QTL detected for yield display apparent overdominance

Comparison with yield QTLs reported in published studies adressing heterosis

- Tang et al, 2010 (iF2)
- Frascaroli et al, 2007
- Frascaroli et al, 2007 reanalyzed by Schoen et al, 2010 (Z2)
- Frascaroli et al, 2007 reanalyzed by Schoen et al, 2010 (Z1)
- Stuber et al, 1992 reanalyzed by Schoen et al, 2010 (Z2)
- Stuber et al, 1992 reanalyzed by Schoen et al, 2010 (Z1)
- Lu et al, 2003 reanalyzed by Schoen et al, 2010 (Z2)
- Lu et al, 2003 reanalyzed by Schoen et al, 2010 (Z1)



- ✓ high level of congruency in highly diverse materials contrasts with the expected complexity of heterosis,
- ✓ several close proximities to hypo recombinant centromeric regions suggests a strong effect of linkage (see also McMullen et al., 2009)

Outcomes of Multi Parental Populations in maize

Genetic analyses

See also results on other experiments by Rebaï et al. 1997 , Buckler et al., 2009, Coles et al., 2010)

- Confirms globally high number of QTLs and a large contribution of linkage to detected regions
- Suggests complex allelic series with gradient of effects
- Dominance certainly important for yield, epistasis more limited / difficult to detect

Breeding analyses

- Nice way to increase a list of solid favorable alleles to be assembled in Marker Assisted Selection
- Application in advanced generations requires marker selection, probability computation (Optimas project)

Opportunities to go further

- a priori grouping of alleles based on dense genotyping of parents (see NAM, LDLA concepts) seems promising to gain power and facilitate management of MAS generations
- Genomic prediction / selection

Questions:

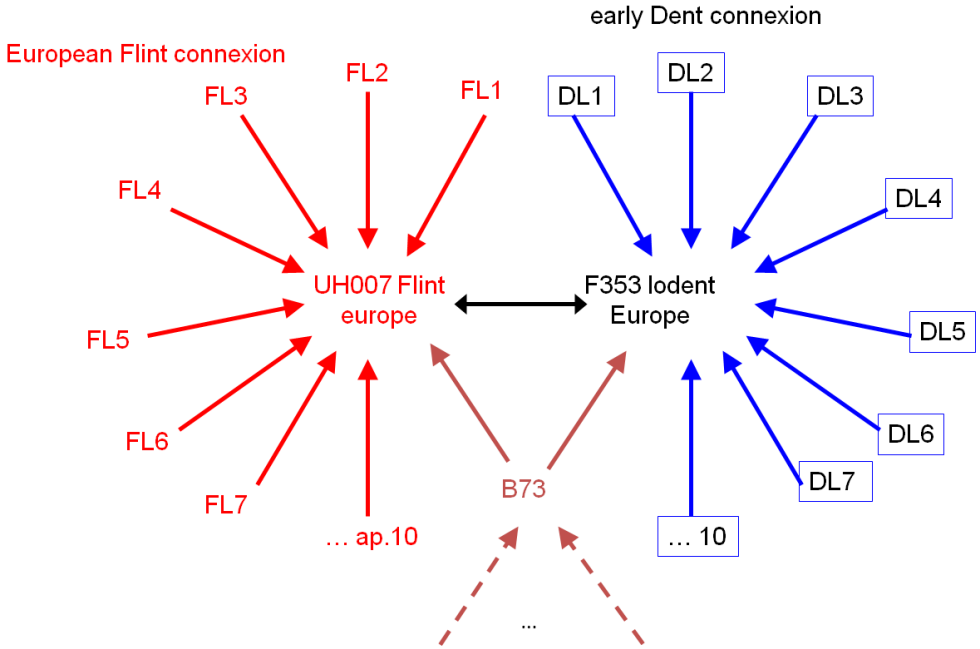
- Nb. of parents and optimisation of design
- Which density of markers needed? (rq. Maize 60 kSNP array just developed, genotyping by sequencing, …)
- Statistical models?

Rk. New early flowering populations developed in cooperative project CornFed,

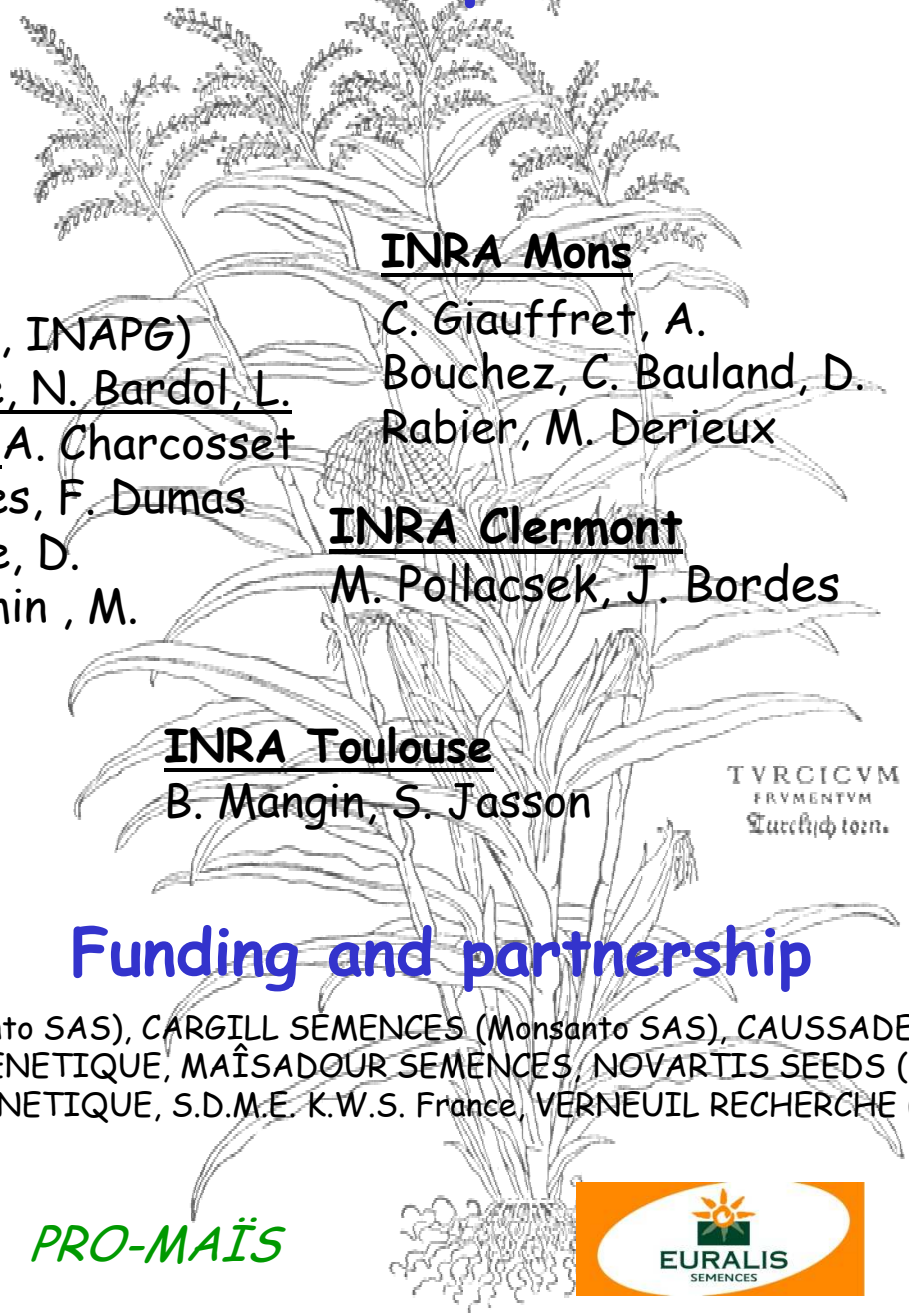
Coop. KWS (M. Ouzunova et al.), Limagrain (P. Flament et al.), University Hohenheim (A. Melchinger et al, TUM (C. Schoen, E. Bauer et al.), Syngenta (N. Ranc et al.),

KWS nursery

29 crosses, mating design constructed as satellites to US NAM, 2600 DH lines genotyped with 50K SNP (TUM), phenotyped for biomass production traits



Participants



SGV Moulon
(INRA, CNRS, UPS, INAPG)
G. Blanc, A. Larièpe, N. Bardol, L. Moreau, A. Gallais, A. Charcosset
D. Madur, V. Combes, F. Dumas
P. Bertin, S. Jouane, D. Coubriche, Ph. Jamin , M. Falque, J. Joets

INRA Mons
C. Giauffret, A. Bouchez, C. Bauland, D. Rabier, M. Derieux

INRA Clermont
M. Pollacsek, J. Bordes

INRA Toulouse
B. Mangin, S. Jasson

Euralis
P. Carolo
Ph. Blanchard
M. Ventolon

Biogemma
P. Dubreuil

Funding and partnership

PROMAIS
ASGROW France (Monsanto SAS), CARGILL SEMENCES (Monsanto SAS), CAUSSADE SEMENCES, CEBECO SEMENCES, EURALIS GENETIQUE, MAÎSADOUR SEMENCES, NOVARTIS SEEDS (Syngenta), LIMAGRAIN GENETICS, PIONEER GENETIQUE, S.D.M.E. K.W.S. France, VERNEUIL RECHERCHE (Limagrain)

Euralis

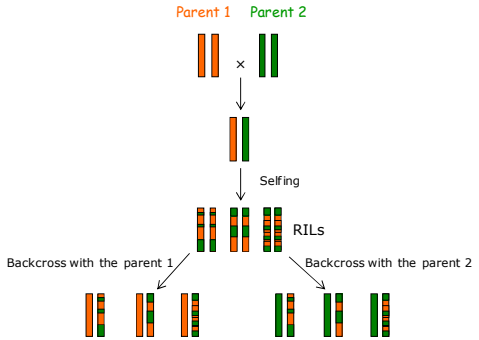
LE MOULON Génétique végétale UMR

PRO-MAÏS

EURALIS SEMENCES

INRA
Institut National de la Recherche Agronomique

QTL analyses with NCIII



1. Separate analysis of the two populations and *a prosteriori* comparison of effects (Stuber et al., 1992)

2. Combination of performances of hybrids with two parents into linear combinations, followed by QTL analysis (in maize: Melchinger et al., 2007*, Frascaroli et al., 2007, Schoen et al., 2010)

Z1 = (RILxP1 + RILxP2)/2
=> linear contrast for additive* effect

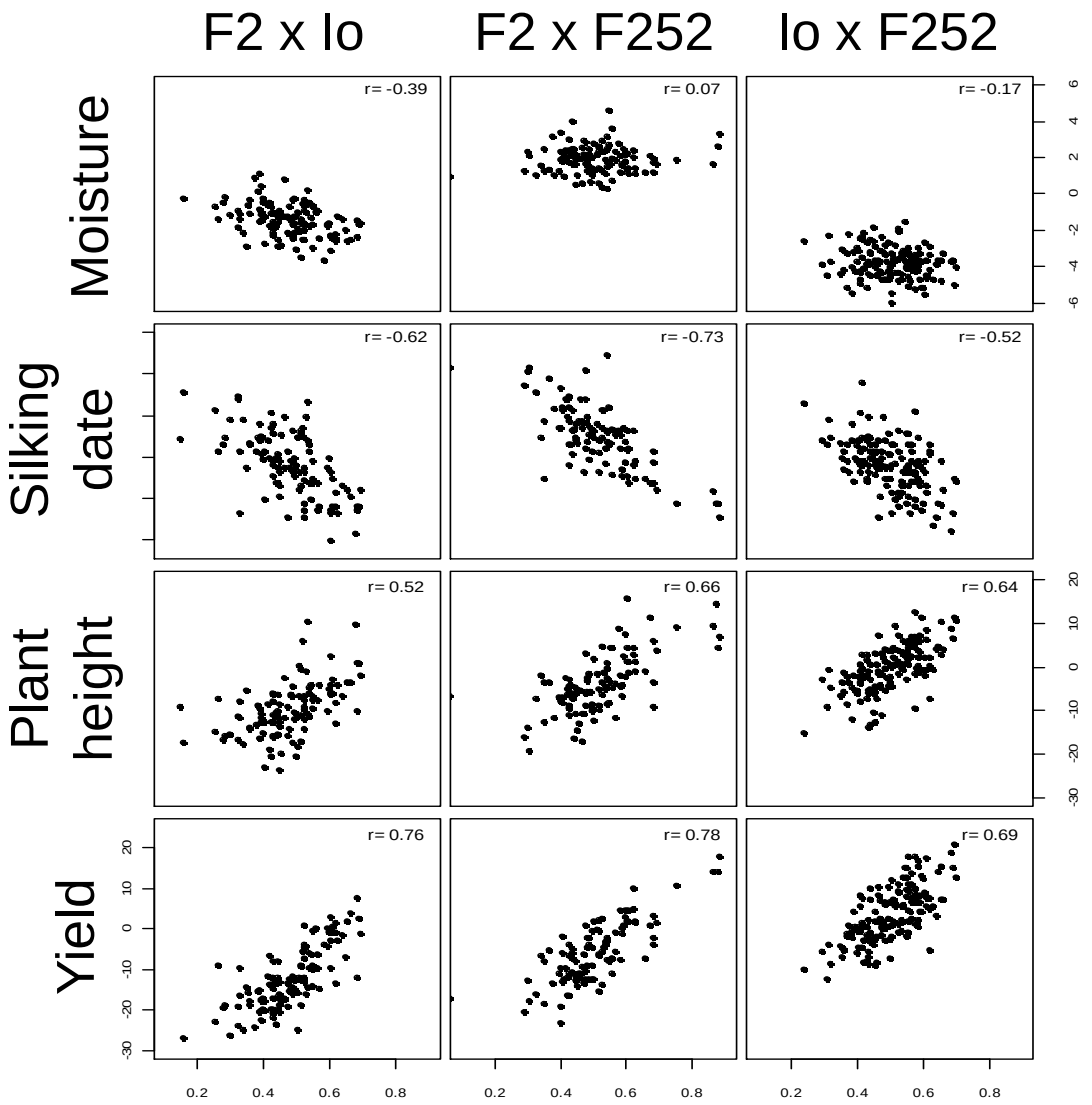
Z2= (RILxP2 – RILxP1)/2
=> linear contrast for dominance* effect

* Augmented by epistatic effects

Global effect of consanguinity

Evaluated through relationship between

- Z2 contrast between hybrids with P1 and P2 (Y axis)
- Distance to Parent1 (X axis), ie. contrast between heterozygosities of hybrids with P1 and P2



Strong magnitude of correlation (but for moisture)

-> polygenic and unidirectional dominance effects (positive for yield and height, negative for flowering time)