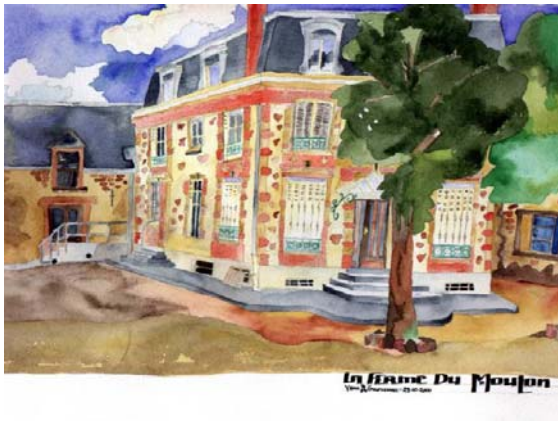


QTL detection and marker assisted selection in multiparental designs: A case study in temperate maize

Guylaine Blanc, Alain Charcosset, André Gallais,
Laurence Moreau

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



Promais



Institut National de la Recherche Agronomique

Possible factors for lack of progress with M cycles (whereas progress went on with phenotypic selection)

- ✓ GxE interactions (environmental conditions of QTL detection different from those of final evaluation)
- ✓ Very strong efficiency of first (C) cycle: high heritability thanks to number of trials, large size of the population selection 17/300 -> fixation of key alleles
- ✓ Evolution of genetic background (suggests epistasis):
 - minor QTL get « silent » in next generation
 - QTL « silent » in first cycle generate variation in next cycle

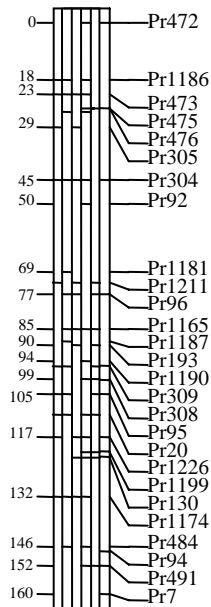
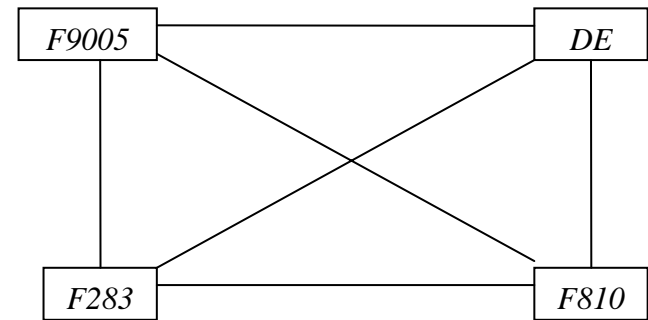
Towards mutiparental designs

Rationale:

- more parents -> more favorable alleles in total
- assembling more challenging -> greater opportunity for MAS
- closer to « usual » breeding practice of many parents and crosses, each with limited population size

Implementation in temperate maize (Blanc et al., 2006)

4 parental early flint lines
6 F2 populations,
150 individuals each,
hybrid evaluation with
dent tester for grain yield,
moisture, flowering time
10 trials



Construction of a synthetic map:
272 markers in total,
(approx. 120-150 per population)

Global QTL analysis (MC QTL software)

Model 1: QTL effects assumed independant across populations

$$\begin{array}{l} \text{Performances} \\ (N \times 1) \end{array} \quad \mathbf{Y} = \mathbf{JM} + \mathbf{X}_q \mathbf{A}_q + \sum_{c \neq q} \mathbf{X}_c \mathbf{A}_c + \boldsymbol{\varepsilon}$$

Random residual

Population effects
(N*p) x (p*1)
p - 1 = 5 df

QTL effects
(N x 2p) x (2p x 1)
p = 6 df

Covariates effects
c x p = cx 6 df

Model 2: allele effects consistent across populations

(3 df for QTL effect instead of 6)

Results

Analyses	Grain Yield		
	Nb of QTL	CI (cM)	R ²
Single-population <i>model (0)</i>	1.3 ^a (7 ^b)	49	25.9
Multipop dis-connected <i>model (1)</i>	5	40 (21 ^c)	18.9
Multipop connected <i>model (2)</i>	12	52 (17 ^d)	46.9



Comparison of results for a same global type I risk

^aaverage number of QTL detected per population

^b number of different regions detected by single-population

Joint connected model:

-Increased Nb of QTL

-Reduced CI

-Increased R²

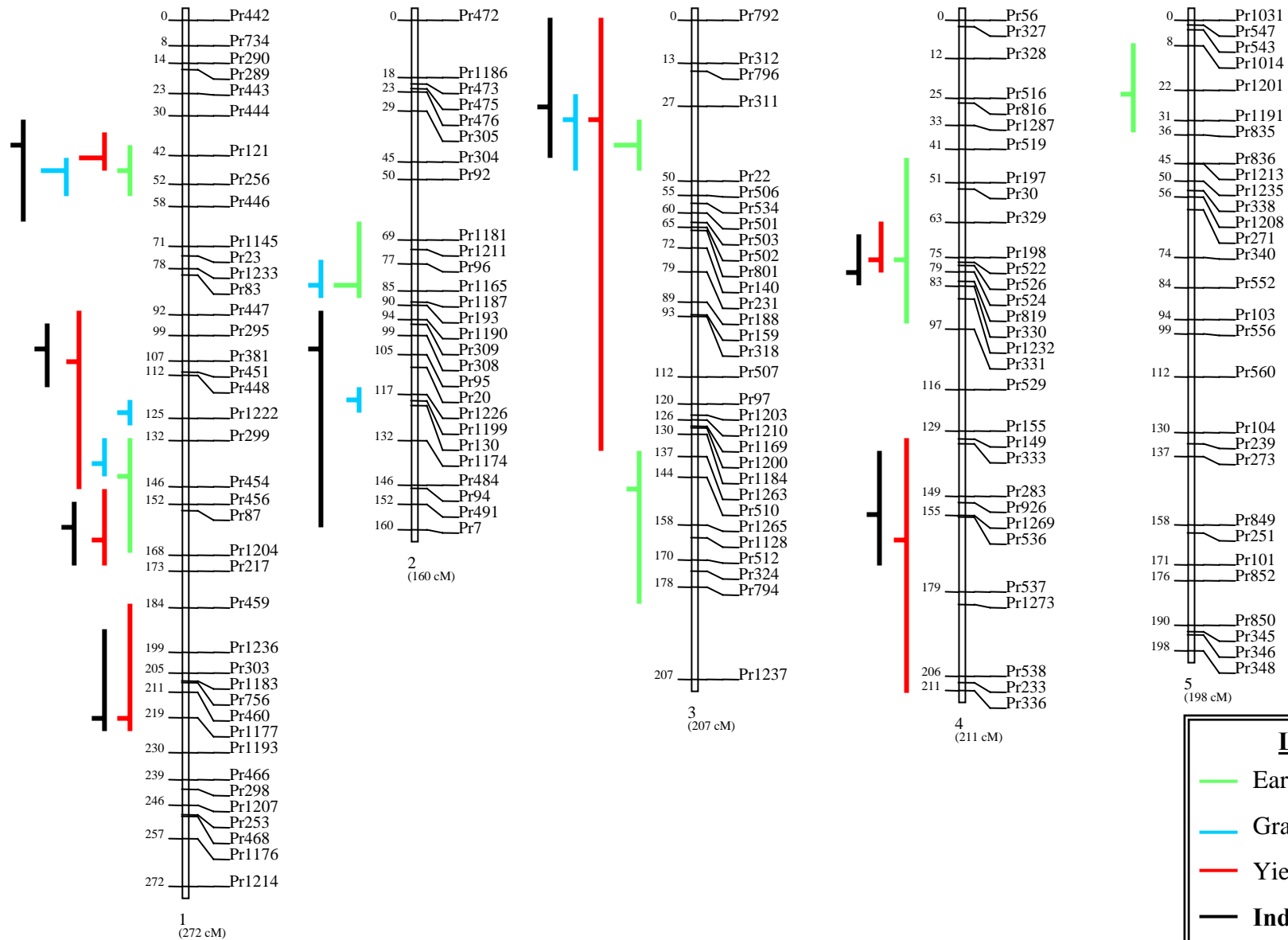
tion

detected in the multipopulation disconnected analyses

Genetic effects and identification of favorable alleles: Grain yield (t ha⁻¹)

N°	chr	pos	Parental alleles				Nb of class	QTL x QTL	QTL x Backgr.
			DE	F283	F9005	F810			
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	*

Numerous colocalisations among traits (generally antagonistic)



LEGEND

- Earliness
- Grain moisture
- Yield
- Index

(Blanc et al., 2006, TAG)

Selection of individuals and determination of couples for intercrossing

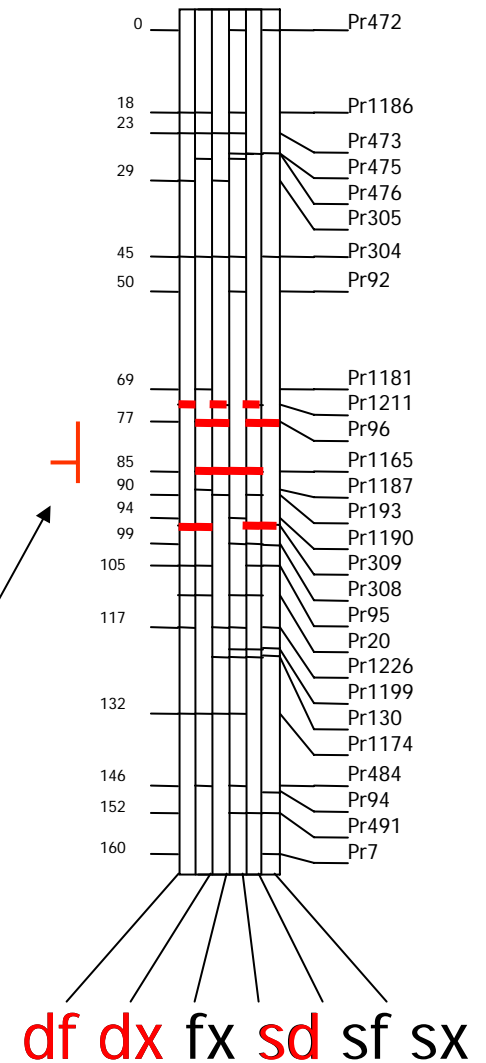
Two selection programs:

- increase grain yield while keeping moisture constant (index),
- early flowering time

For each program 30 individuals selected at each generation (out of 600 and 300, resp.) based on expected number of favorable QTL alleles

Selection of markers to infer the probability of allele over generations, Taking into account allelic relationships at markers

Ex. flowering time QTL, chr. 2, favorable allele = **d**



Calendar of the selection

2001-2001: initial evaluation

2002: first cycle M, (P for yield)

2003: second cycle MM

2004: third cycle MMM.

Evaluation of genetic gains 2004 et 2005

Testcross values of:

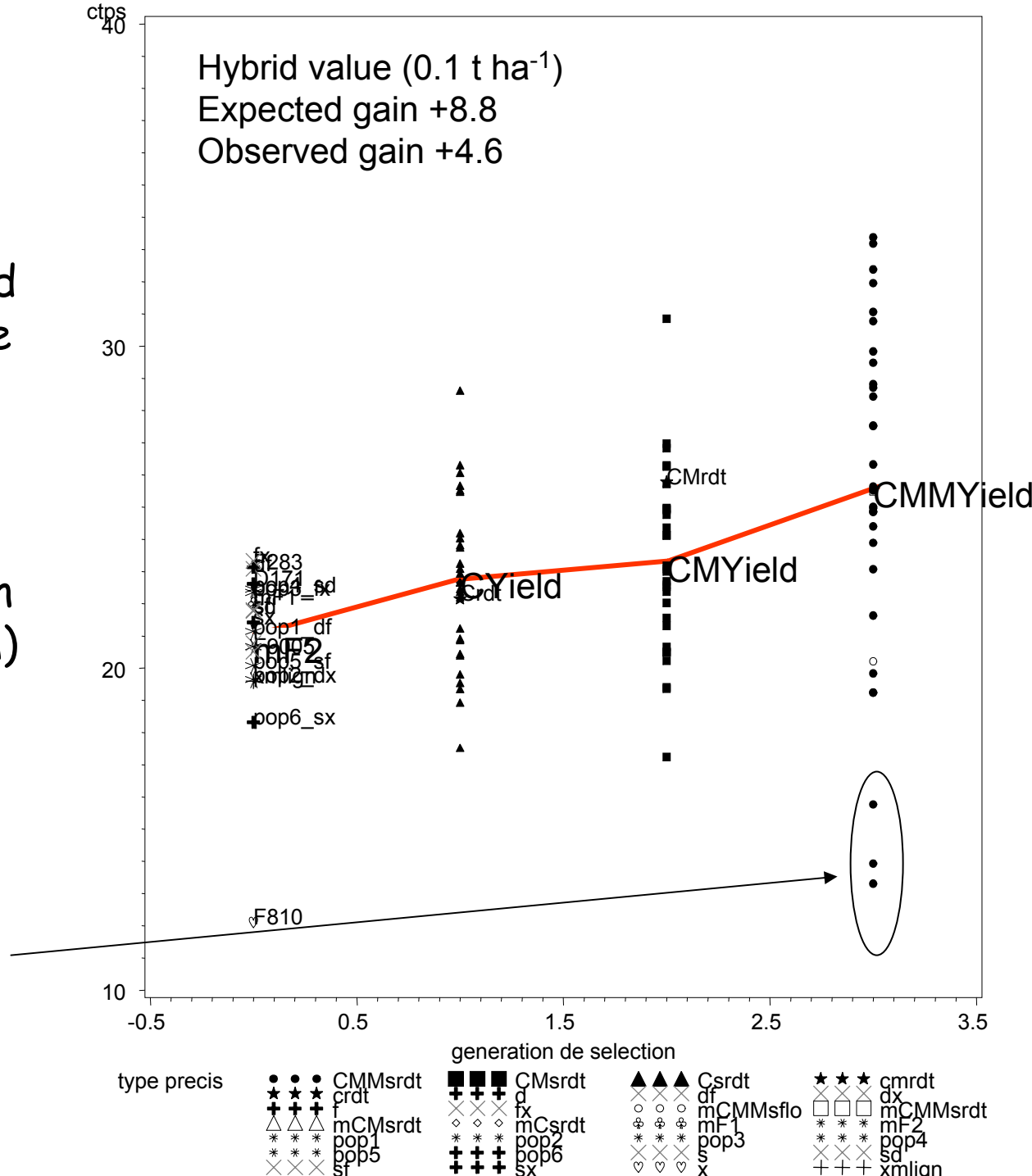
- 4 parental lines
- 6 F1 hybrids
- 6 initial F₂ populations as pools
- Selected genotypes (20-30) at each cycle M, MM, MMM*
- Populations resulting from the intercrosses M and MM*, as pools

*2005 only, 8 trials each year.

Yield index (2004-2005 trials)

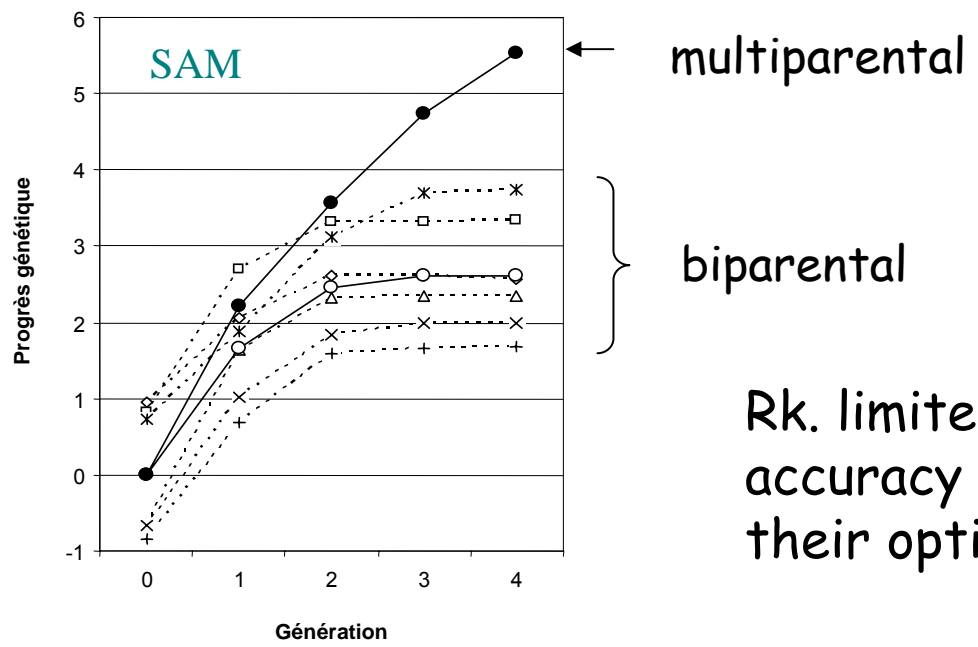
- Evolution of expected frequency of favorable alleles from 0.28 (parents) to 0.80
- M cycles efficient: significant genetic gain (50% of expected gain)

Individuals selected to prevent loss of the single favorable allele brought by F810



Genetic gain with MAS over several cycles, interpreted as the impossibility to assemble the key favorable alleles in a single generation, due to different parental origins.

Simulations show an advantage compared to several parallel biparental programs (same total means) (Blanc et al., in press),



Rk. limited contribution of higher accuracy of QTL detection, if both at their optimal type I risk for MAS

Conclusions

Multiparental approaches seem promising

Limits however in the number of parents that can be managed with previous approach (assuming one allele per parent)

calls for further investigations to broaden the diversity addressed

- design optimisation,
- Consider QTL effect as random,
- A priori grouping of alleles (eg. Jansen et al., 2003)



Acknowledgments

B. Mangin and colleagues for statistical aspects (MCQTL),

M. Merlino, V. Combes, F. Dumas, D. Madur for genotyping

INRA experimental stations

Lusignan: X. Charrier

Mons: C. Bauland, D. Rabier

Moulon: P. Bertin, P. Jamin, D. Coubriche, S. Jouane

The private companies members of Promais:

ADVANTA, CAUSSADE SEMENCES, EURALIS
GENETIQUE, LIMAGRAIN GENETICS, MAISADOUR,
MONSANTO, NORDSAAT FRANCE, PIONEER, R2N,
S.D.M.E. / KWS France, SYNGENTA.

Agriobtentions

French ministry of Agriculture and the French ministry of
research

Selection of individuals and determination of couples for intercrossing

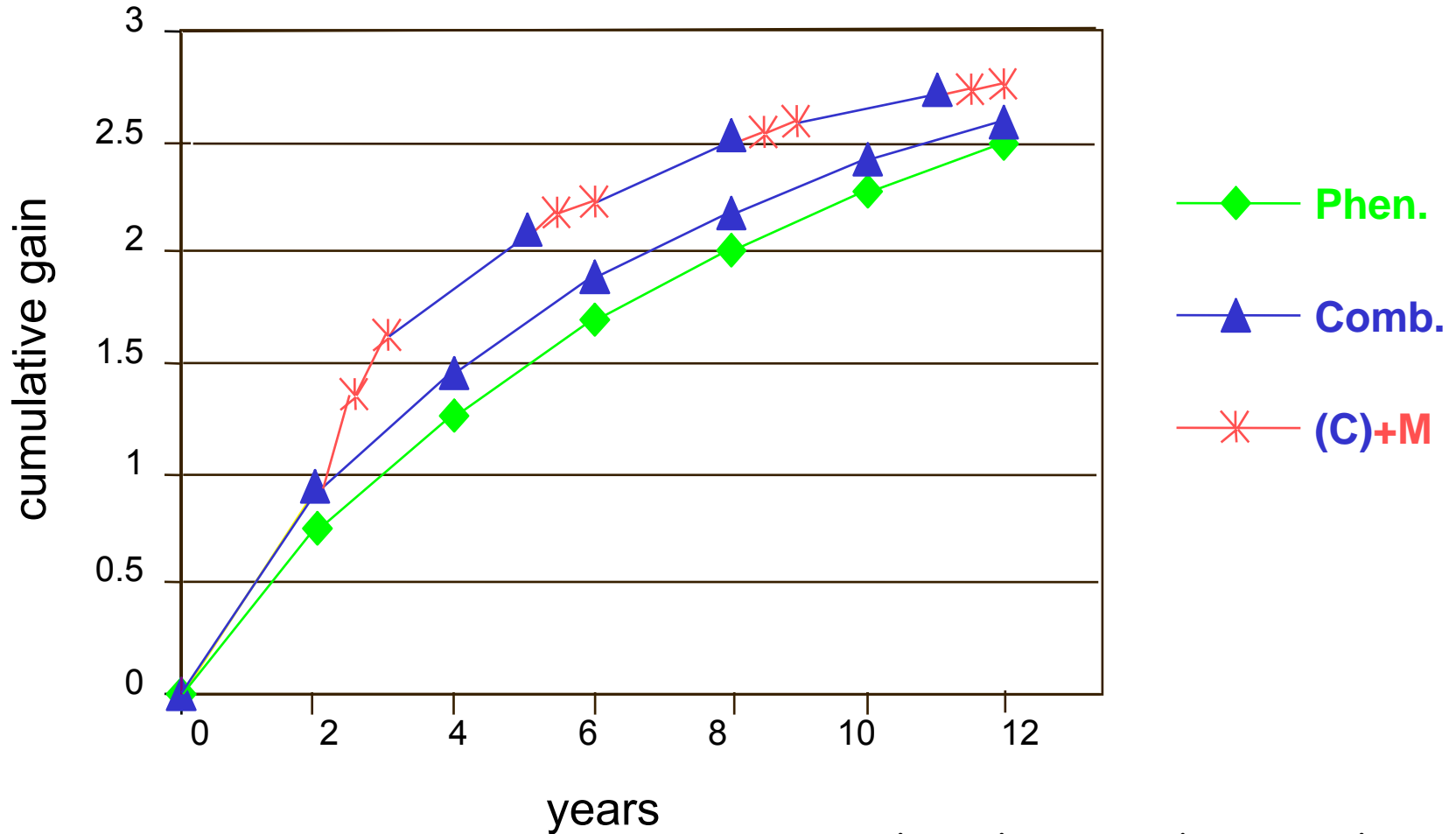
Choice among couples based on an utility criterion to favor couples that will produce the best individuals as possible in the next generation

$$U_{\text{couple}} = \text{Mean of the parental values of the couple} + i \sigma_{\text{couple}}$$

Variance = only heterozygotes genotypes at QTL contribute to variance in the next generation.

Empirical criterion, way to give more importance to top individuals. See more on the topic in Bernardo, Moreau and Charcosset (Crop Science, 2006)

Combination of « combined selection » and “marker only” selection cycles: Results of stochastic simulations (20 QTL, 200 individuals, $h^2=0.5$)



Determination of couples for intercrossing

Choice among couples based on an utility criterion to favor couples that will produce the best individuals as possible in the next generation

$$U_{\text{couple}} = \text{Mean of the parental values of the couple} + i \sigma_{\text{couple}}$$

Variance = only heterozygotes genotypes at QTL contribute to variance in the next generation.

Empirical approach, based on a strong assumption of normality of the molecular score, do not optimize the probability of obtaining a superior genotype in the whole population...

QTL detection models

Model 2 Joint analysis. QTL effects nested within populations (full model, disconnected model)

$$\begin{array}{l}
 \text{Performances} \\
 (N \times 1)
 \end{array}
 \mathbf{Y} = \mathbf{JM} + \mathbf{X}_q \mathbf{A}_q + \sum_{c \neq q} \mathbf{X}_c \mathbf{A}_c + \boldsymbol{\varepsilon}$$

Random residual

Population effects
(N*p) x (p*1)
p-1=5 df

QTL effects
(N x 2p) x (2p x 1)
p=6 df

Covariates effects
c x p = c x 6 df

Model 3 Joint analysis using connections between populations (reduced model, connected model)

$$\mathbf{Y} = \mathbf{JM} + \mathbf{X}_q^* \mathbf{A}_q^* + \sum_{c \neq q} \mathbf{X}_c^* \mathbf{A}_c^* + \boldsymbol{\varepsilon}'$$

Random residual

QTL effects
(N x k) x (k x 1)
k-1=3 df

Covariates effects
c x (k-1) df = c x 3 df

k=number of parental lines=4

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 (1)	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 disconnected model (2)	8	25 (23 ^c)	64.3	8	29 (26 ^c)	52.2	5	32 (17 ^d)	40.9	10	33 (23 ^d)	55.0
Multipop connected model (3)	11	28 (16 ^d)	66.0	13	20 (14 ^d)	57.9	12	32 (17 ^d)	40.9	10	33 (23 ^d)	55.0

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

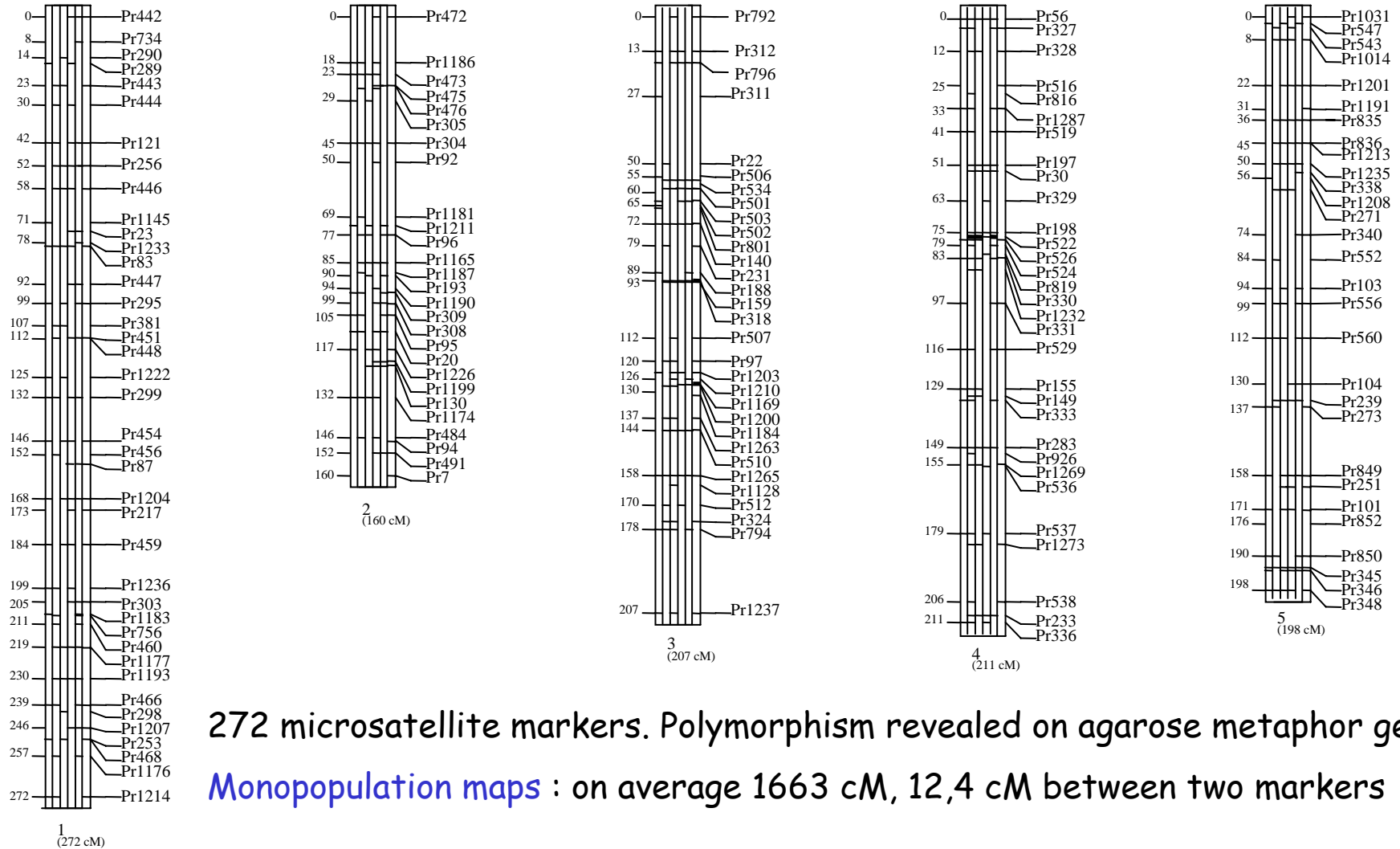
^aaverage number of QTL detected per population

^b number of different regions detected by single-population analyses (model(1))

^caverage 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))

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

Elements of interpretation for lack of genetic gain observed with MAS in biparental scheme

- High heritability and high selection intensity for the first C cycle \rightarrow fixation of key alleles
- effects at minor QTL still variable decreased in the selected background, suggests epistasis
- \rightarrow new QTL have been involved in P gain in next generations

Combination of marker information at detected QTL and phenotypic performance (Lande and Thompson, 1990, Genetics)

◆ Molecular score (M)

$$G = M = \sum_q a_q \theta_q$$

a_q additive effect of parent A allele at QTL q

θ_q Expected number of parent A allele at QTL q (inferred from close markers)

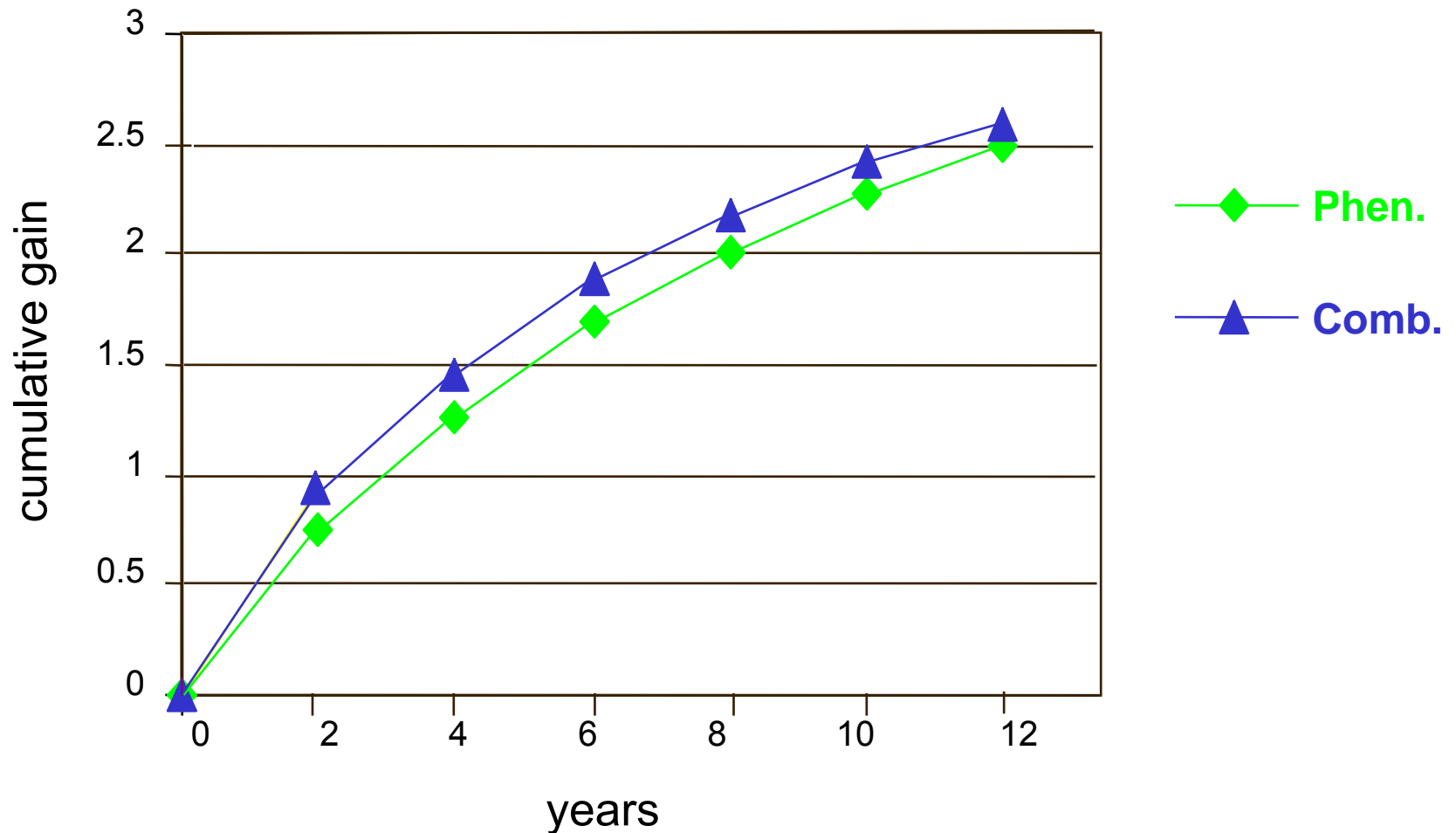
◆ Best prediction of genetic value (C) by combining (M) and phenotype (P)

$$G = b_p P + b_m M$$

b_p and b_m depend on h^2 and R^2_p
if $h^2 \nearrow$, $b_m \searrow$ et $b_p \nearrow$

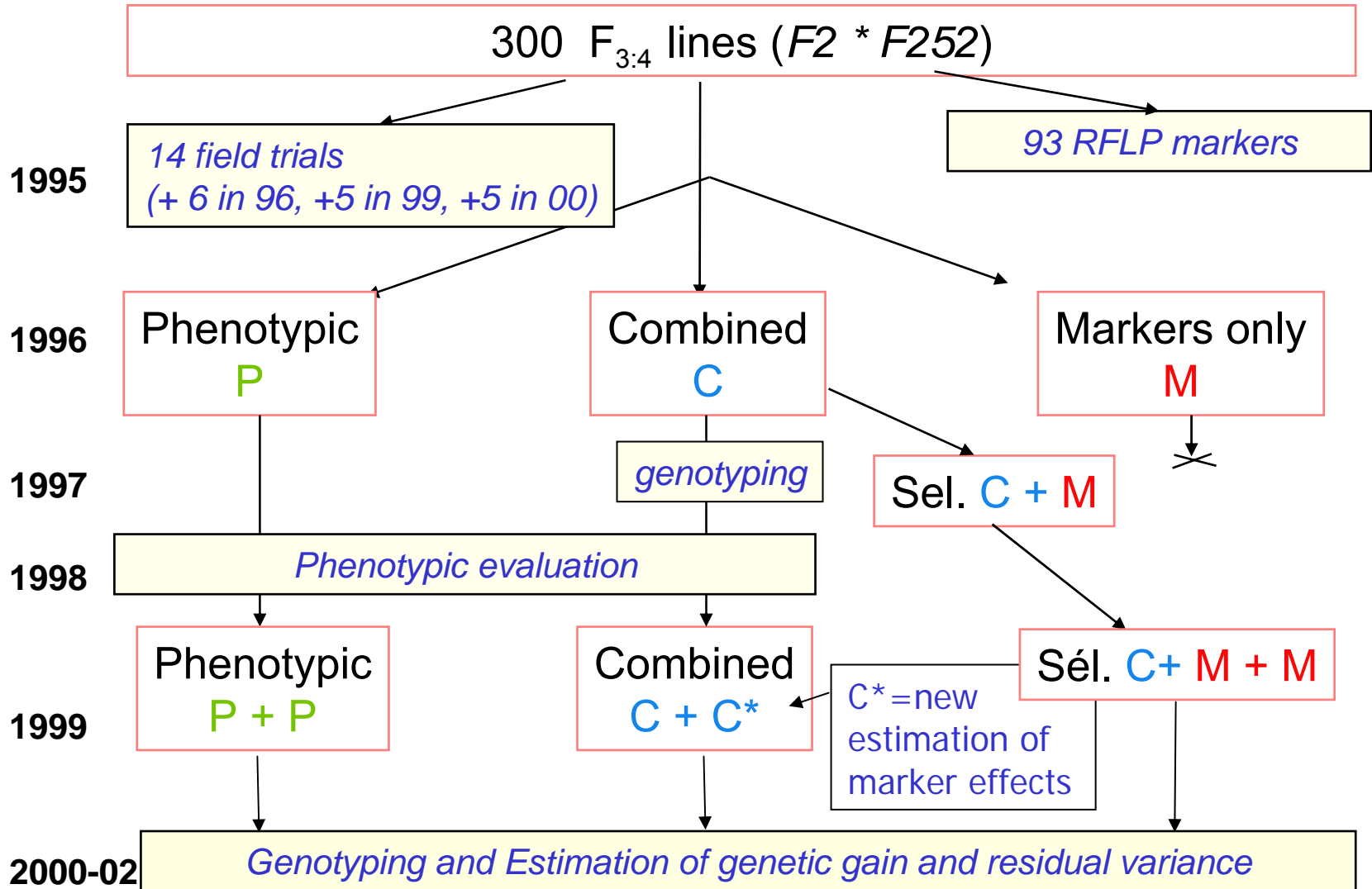
Rk. : sequential selection on M and then P also possible, (Xie and Xu, 1998, Heredity)

Combination of « combined selection » and “marker only” selection cycles: Results of stochastic simulations (20 QTL, N=200, $h^2=0.5$)



Hospital et al.,1997, Theor Appl Genet

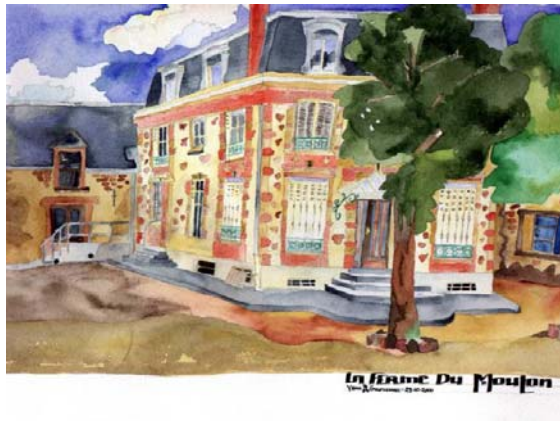
Marker-Assisted Selection in a biparental population (Promaïs program)



Bridging genomics and genetic diversity: Evolution of population structure in temperate maize and consequences for association genetics

Alain Charcosset et al.

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



Institut National de la Recherche Agronomique

Calendar of the experiments

2001-2001: initial evaluation

2002: first cycle M

2003: second cycle MM

2004: third cycle MMM. Evaluation of M and MM

2005: Evaluation of M, MM and MMM

Evaluation of genetic gains

Testcross values of:

4 parental lines

6 F1 hybrids

6 initial F2 populations as pools

Selected genotypes at each cycle M, MM, (MMM)

(25-30 selected genotypes per cycle)

2(1) Population(s) resulting from the intercrosses as pools

2 (1) year(s) of experiment, 8 trials each year.

Calendar

2001-2001: initial evaluation

2002: first cycle M (flowering time), P (yield)

2003: second cycle MM

2004: third cycle MMM.

Calendar of the experiments

2001-2001: initial evaluation

2002: first cycle M (flowering time), P (yield)

2003: second cycle MM

2004: third cycle MMM. First evaluation of M and MM

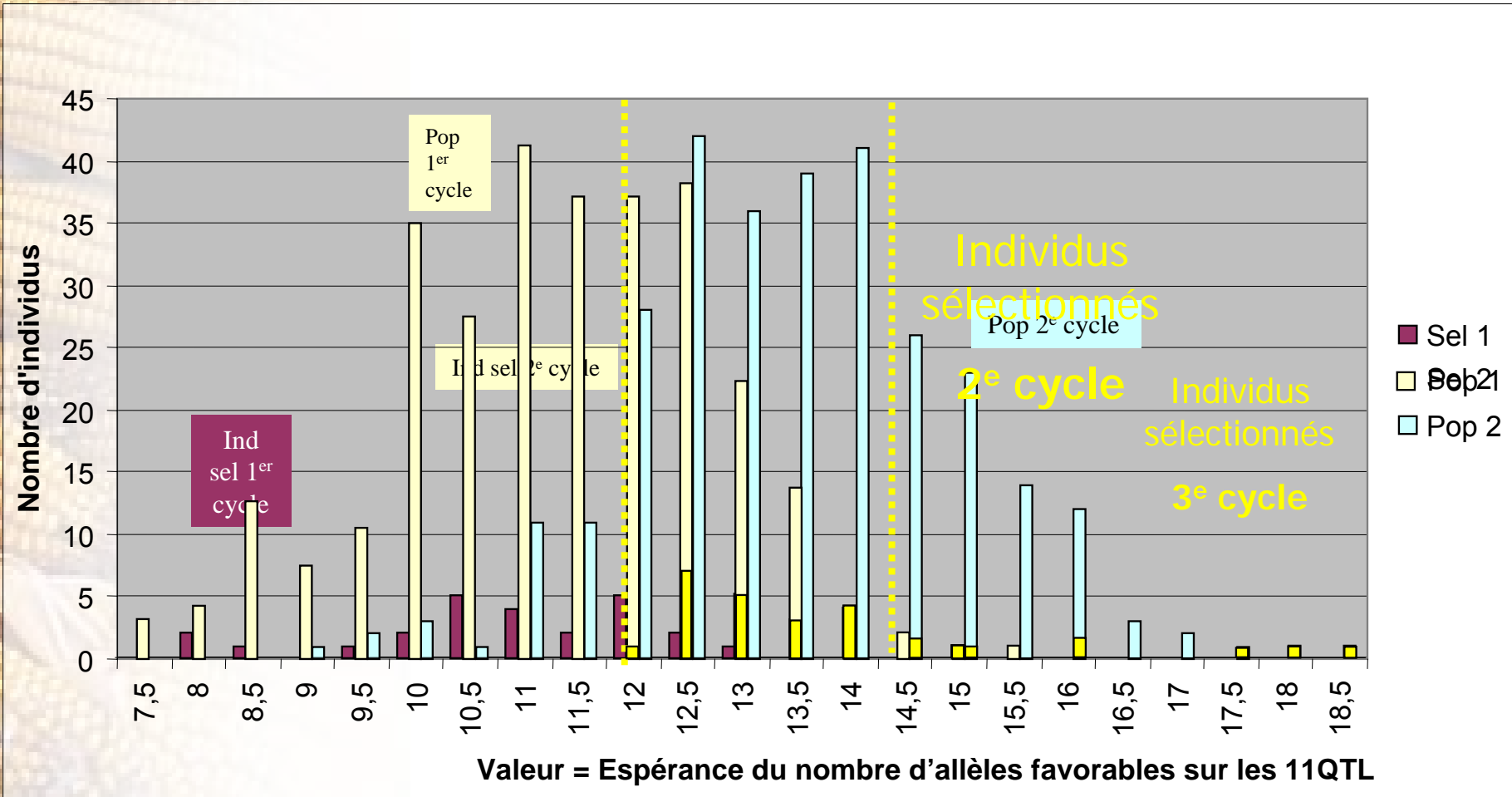
2005: Evaluation of M, MM and MMM

Towards multiparental designs

Rationale for moving to multiparental designs:

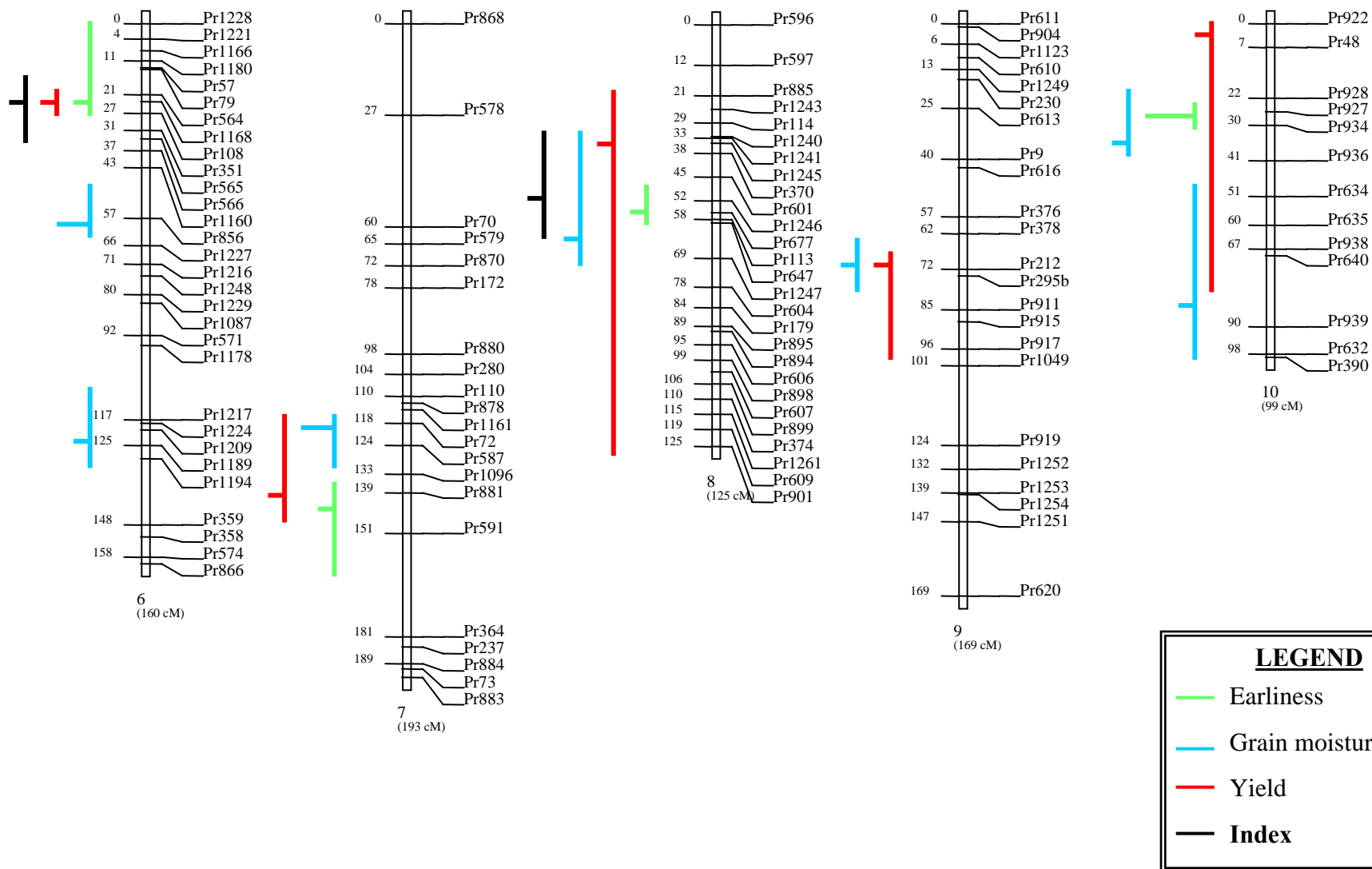
- more parents -> more favorable alleles in total
- assembling more challenging -> greater opportunity for MAS
- closer to « usual » breeding practice of many parents and crosses, each with limited population size

Valeur des individus sélectionnés pour la floraison



Augmentation de la valeur des individus

QTL detection with connected multipopulation model (MCQTL, genomewide $\alpha=10\%$)



(Blanc et al., 2006, TAG)

