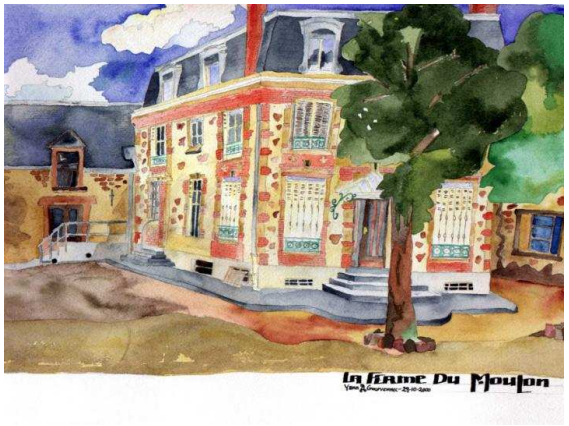
Genetic architecture of yield and related traits in European maize: insights into the effects of linkage and allelic series. Consequences for marker assisted selection

Alain Charcosset, Guylaine Blanc, Sébastien Ducrocq, YunFeng Huan, André Gallais, Laurence Moreau

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



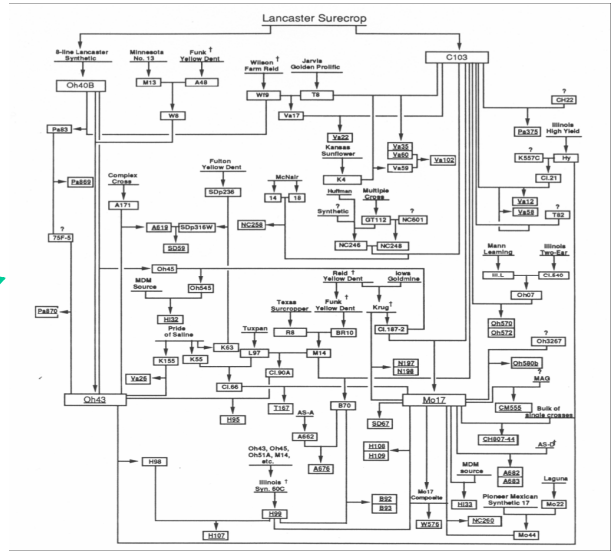
LE MOULON Génétique végétale UMR

INRA
Institut National de la Recherche Agronomique

Elements on breeding of annual plants

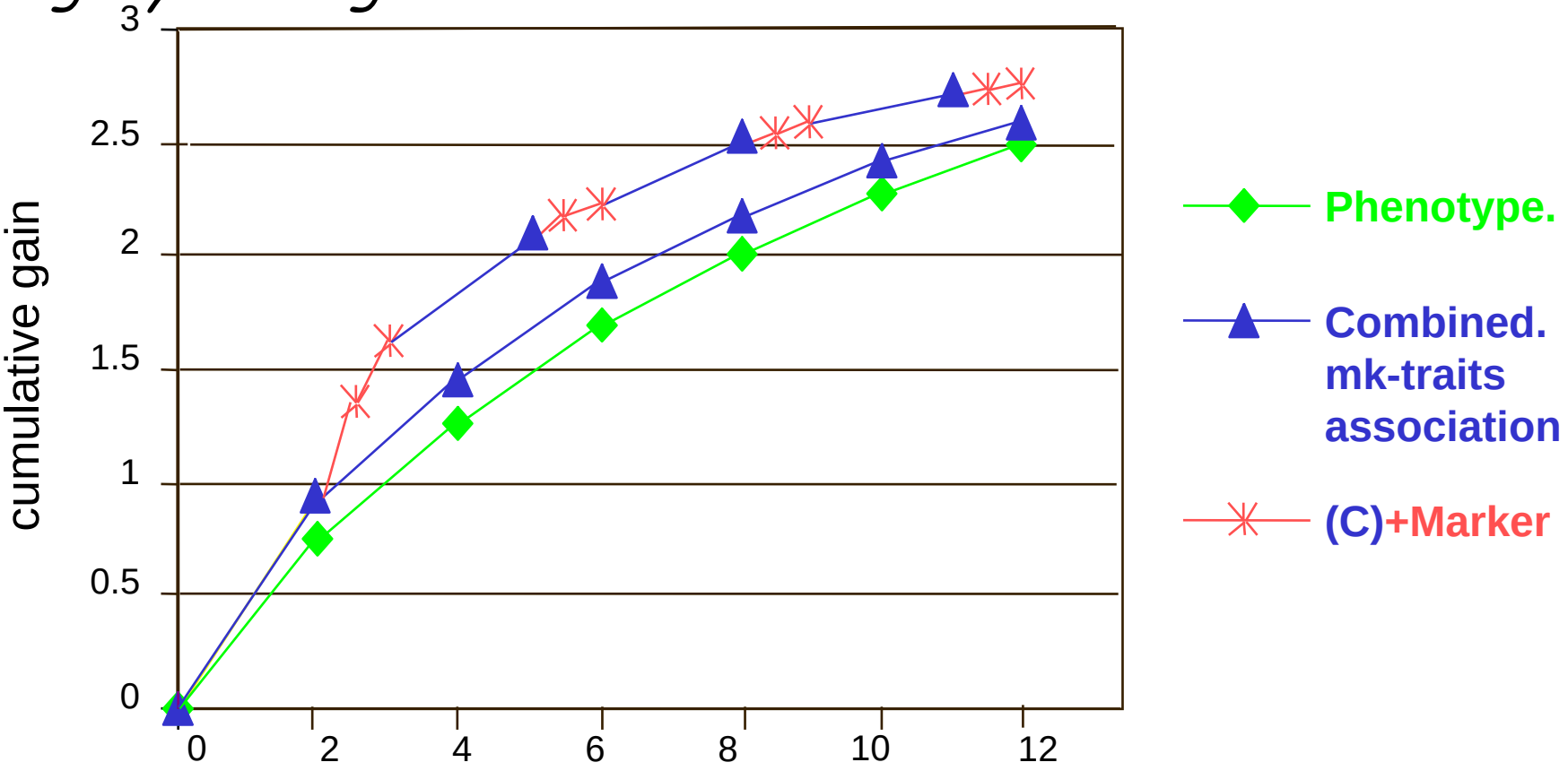
- Many independant breeding programs (companies)
- Development of inbred lines: varieties or hybrid parents
- Within each program: many crosses (P1xP2) leading to populations of limited size

- Phenotypic evaluation during or after fixation of new inbred lines
- Heavy progeny testing, generally multilocal (GxE)
- Data analysis generally not managed globally (no BLUP), integration of results mostly through breeder's expertise
- De facto recurrent selection through intercrossing of superior inbred lines or elite families



Main expectations for Marker Assisted Selection for complex traits:

a strategy to accelerate genetic gain thanks to fast selection cycles based on markers only rather than progeny testing



Simulations (20 QTL, 200 inds, h^2=0.5)
Hospital et al.,1997, Theor Appl Genet

Talk outline

1.QTL mapping in mutiparental connected populations and application for MAS

2.Fine mapping of specific QTL based on linkage and Linkage Disequilibrium

3.Global fine mapping using Advanced Intercross Lines

Model: maize varieties adapted to Northern Europe, presently hybrids between lines of European and North American origins

1. QTL mapping in mutiparental connected populations and application for MAS

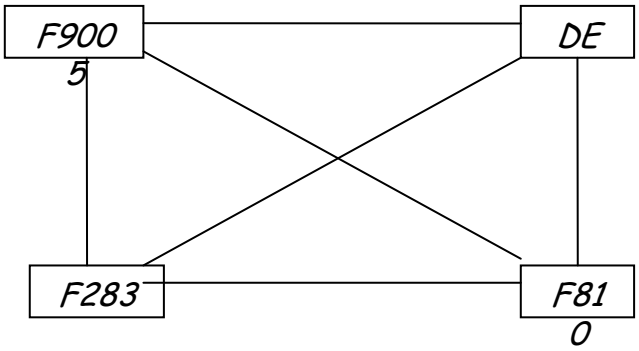
QTL mapping in annual plants started mostly with biparental populations analyzed in independent studies
-> informative
-> (generally) efficient for MAS

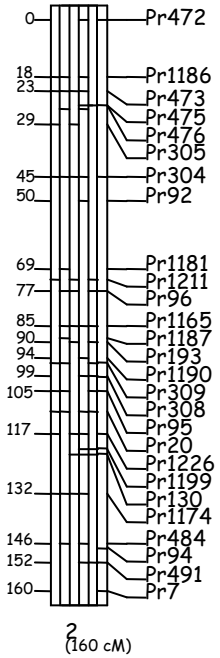
<u>But :</u>
•To reach enough power, size of each population beyond usual breeding practice
•No (direct) possibility to integrate results from different populations

-> interest for more global designs

multiparental QTL mapping in temperate maize (Blanc et al., 2006, TAG)

4 parental early flint lines
6 F_2 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, Jourjon, 2005, Bioinformatics)

Model 1: QTL effects assumed independant across populations

Performances (N x 1)

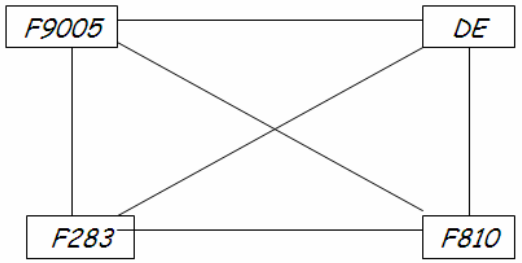
$$\mathbf{Y} = \mathbf{JM} + \mathbf{X_q A_q} + \sum_{c \neq q} \mathbf{X_c A_c} + \boldsymbol{\varepsilon}$$

- $\mathbf{JM}$: Population effects, (N*p) x(p*1), p -1=5 df
- $\mathbf{X_q A_q}$: QTL effects, (N x 2p) x (2p x 1), p=**6 df**
- $\sum_{c \neq q} \mathbf{X_c A_c}$: Covariates effects, c x p= cx 6 df
- $\boldsymbol{\varepsilon}$: 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



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	
5ß	3	150	139 - 188	0.04	-0.07^{a}	0.11^{a}	-0.3^{b}	0.26^{c}	3	-	
6ß	4	75	45 - 97	0.02	-0.24^{a}	0.09^{bc}	0.00^{b}	0.15^{c}	3	-	
7αß	5	26	10 - 38	0.02	0.06^{a}	-0.29^{b}	0.10^{a}	0.12^{a}	2	-	
8ß	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
Contribution of epistasis seems limited

Yield (t ha^{-1})

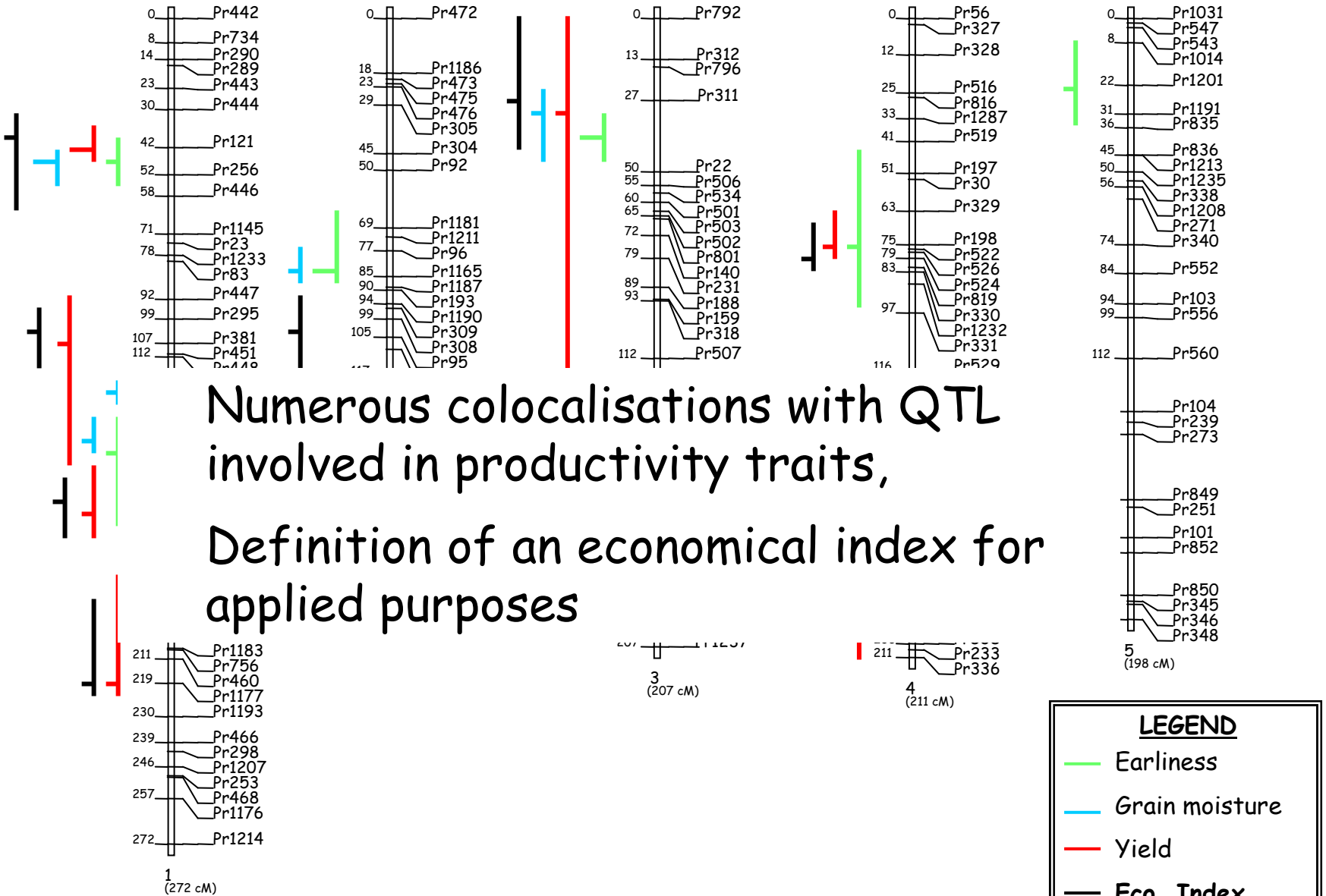
N°	chr	pos	Parental alleles: DE	F283	F9005	F810	Nb of class	QTL x QTL	QTL x Backgr.
1	1	44	**0.099** ^{a}	**0.114** ^{a}	-0.017 ^{b}	-0.195 ^{c}	3	3, 11	*
2	1	105	**0.102** ^{a}	-0.086 ^{b}	0.017 ^{c}	-0.033 ^{bc}	3	7, 11	
3	1	160	**0.067** ^{a}	-0.082 ^{b}	-0.083 ^{b}	**0.098** ^{a}	2	1, 7	
4	1	217	**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

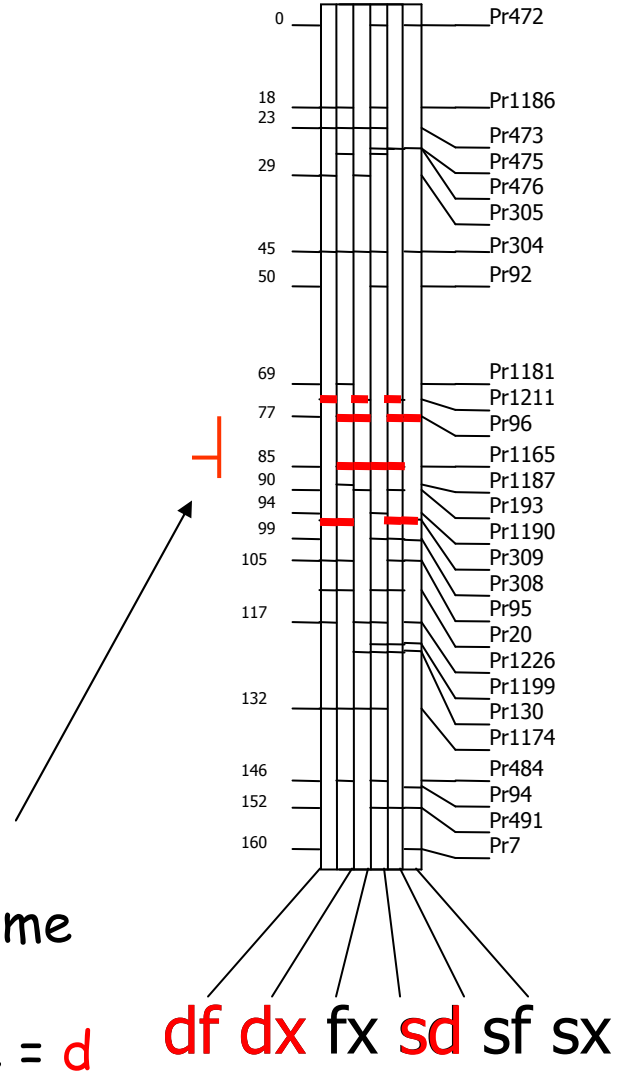
LEGEND

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

Marker based selection of individuals

Two programs: (i) yield index and (ii) flowering time 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 -> approx. 35 markers per cycle

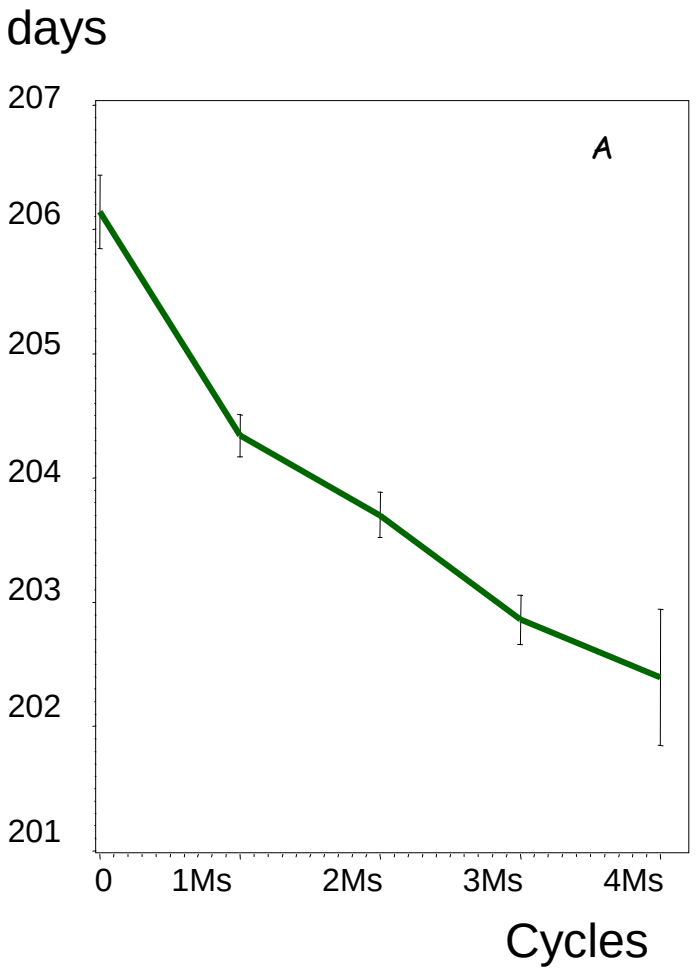


Ex. flowering time QTL, chr. 2, favorable allele = d

Flowering time program
(evaluation: 2004-2005-2007 trials)

- Evolution of expected frequency of favorable alleles from to 0.36 (parents) to 0.86

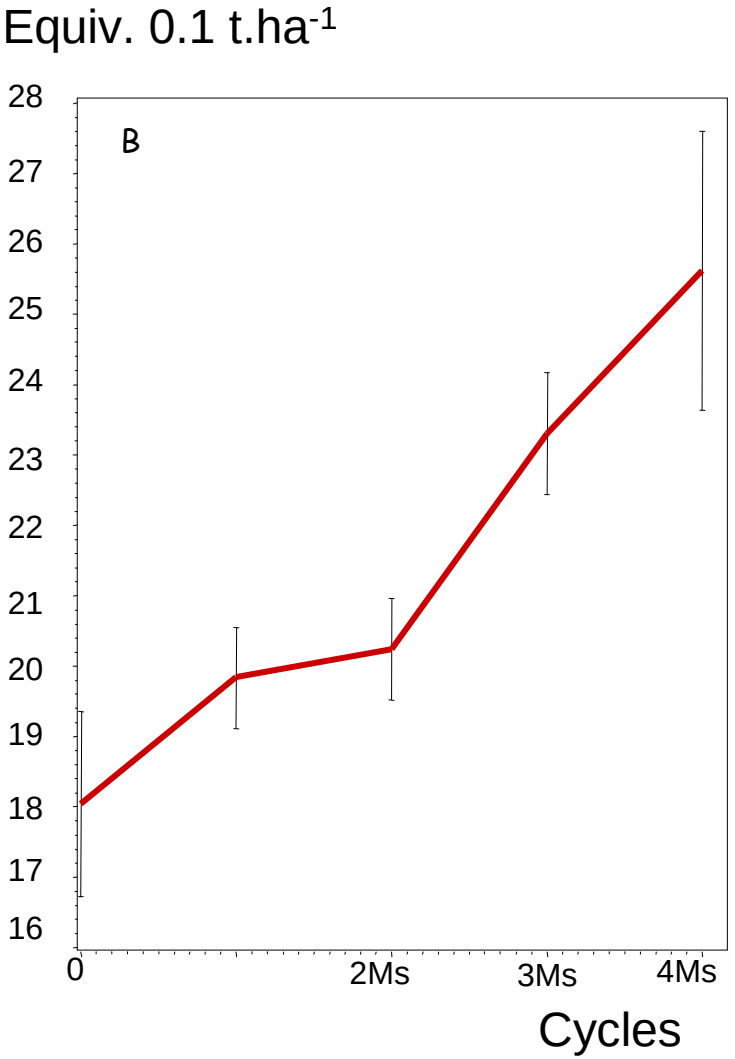
- M cycles efficient: significant genetic gain (70% of expected gain)



Yield index program
(evaluation 2004-2005-2007 trials)

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

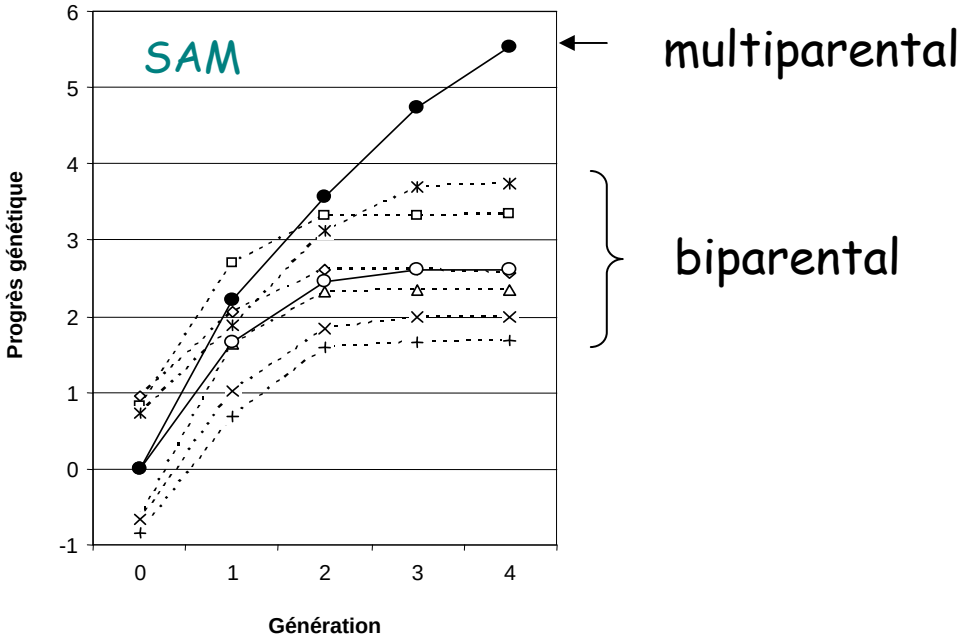
- M cycles efficient: significant genetic gain (50% of expected gain)



High performance of inbred lines fixed from last cycles confirmed

To which extent multiparental QTL mapping is beneficial for marker assisted selection?

Simulations show an advantage compared to several parallel biparental programs (same total means) (Blanc et al., 2008),



Faster recycling of best individuals obtained from different crosses

Efficient but opportunities to go further

- a priori grouping of alleles based on dense genotyping of parents : gain in power and faciliated management of MAS generations
- Genomic selection

Raises however questions:

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

Better understanding of determinism should be beneficial to optimize strategies

- Number of QTL and effect of linkage?
- local LD and allelic series at QTL?

2. Brief elements on Fine mapping of specific QTL based on linkage and Linkage Disequilibrium

Development of near isogenic materials in QTL regions:

-residual heterozygosity in F_5 (Heterozygous Inbred Families)

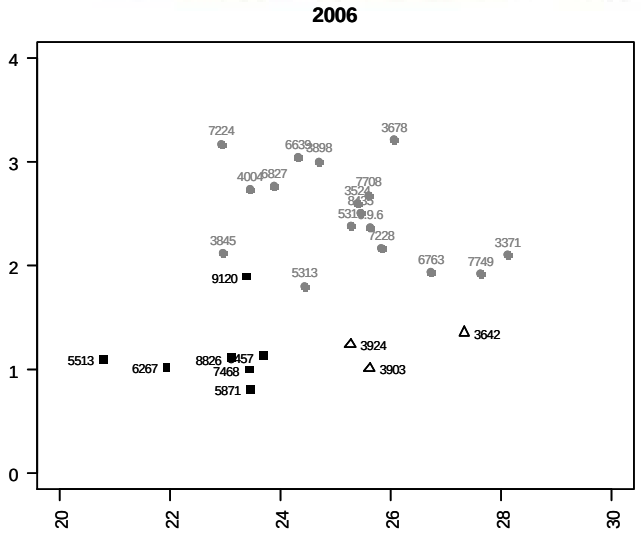
- introgressions in a "recipient" line

...

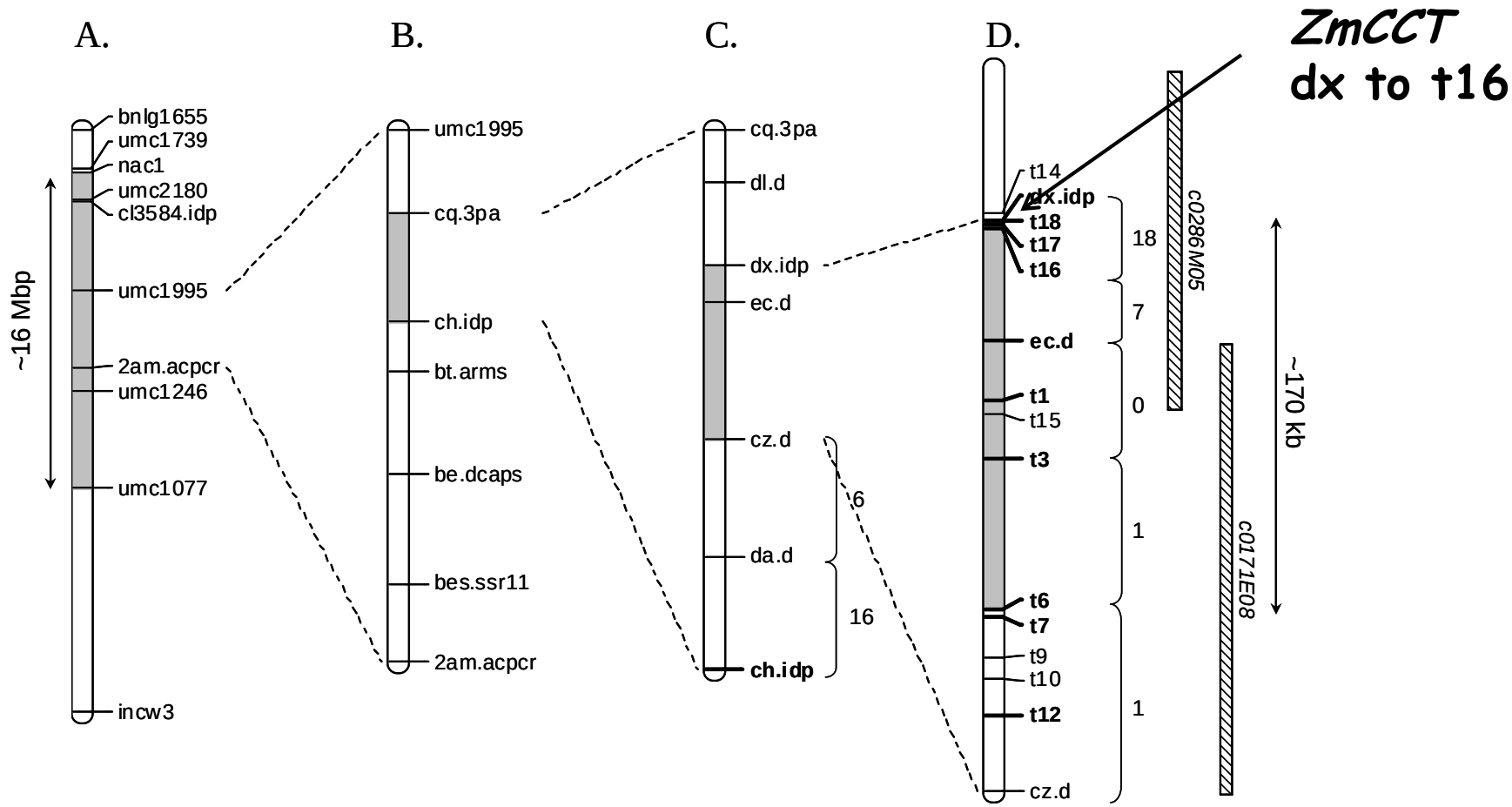
Ex. Linkage based fine mapping of major flowering time QTL on Chr. 10 (Ducrocq et al. 2009, Genetics, coop C. Giauffret et al. INRA Mons)

Screening of recombinants within 8000 plants (HIF within F331 x F286) RILs

Scoring of individual plants within selfed families:
genotype at QTL determined by combining mean and variance (up to 4 environments)



Development of markers: exploitation of synteny, sequence of BACs ….

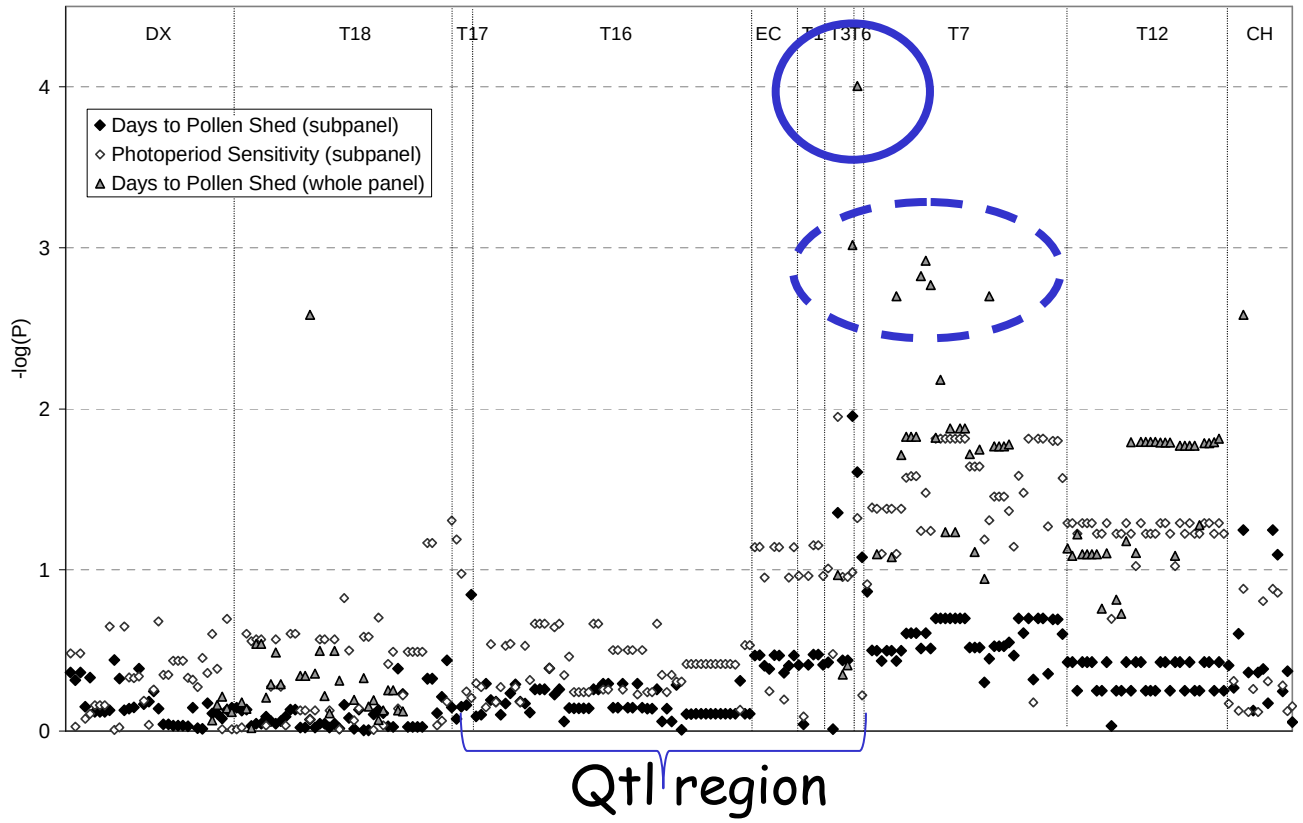


Highly unbalanced recombination

-> 170 kb interval, in the vicinity of *ZmCCT*, homologue to rice flowering time gene *HD7* (Xue et al., 2008) but excluding it (rk. Preliminary expression studies of ZmCCT between parents not conclusive)

Association Genetics (375 diverse inbred lines panel)

Region characterized very high LD (three main haplotypes considering QTL region and ZmCCT)



divergent effects: late flint allele (consistent Fv283 effect in Blanc et al. 2006), a most common intermediate allele (three other flint lines) and an earlier allele

Validation of association genetics effects:

late allele: clearly validated by « reverse fine mapping » using 1700 plants within new nearly isogenic materials (introgressions),

early allele: to be investigated in a similar way

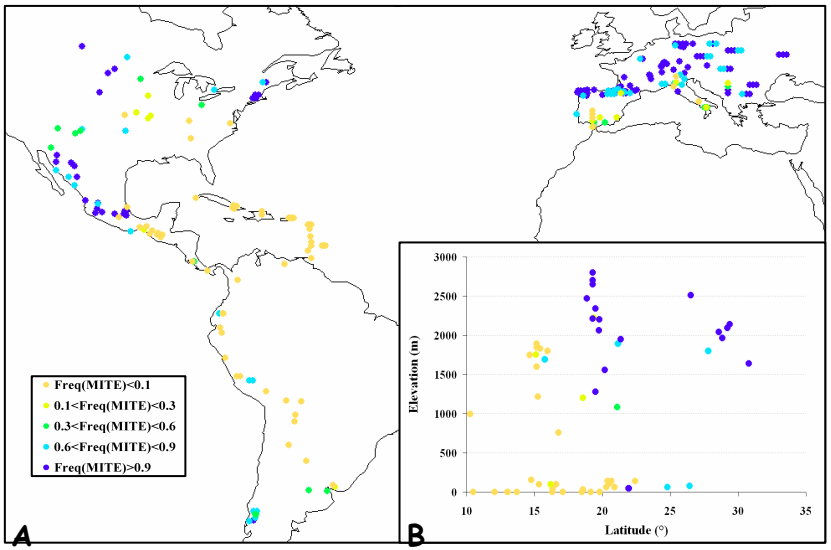
Similar case study for vgt1 (Salvi et al. 2007, Ducrocq et al., 2008)

-Causal factor 70 kb upstream of gene

-Lower Linkage Disequilibrium

-Stronger association genetics tests due to more balanced frequencies within groups

- consistency of allele distribution with ecogeographical conditions



Mite insertion (early allele) frequency

Possible to finely characterize QTL, Combination of linkage and LD beneficial

-> biological information, eg. role of distant regulatory factors
-> diagnostic markers expected to be robust in different populations
-> association genetics can suggest new alleles of interest

But long and risky

- Effect may depend on genetic background
- One QTL may become several during linkage fine mapping
- Scoring of unmabiguous phenotypes (mendelization) not always possible

Possibility to anticipate and accelerate fine mapping on a genome wide scale?

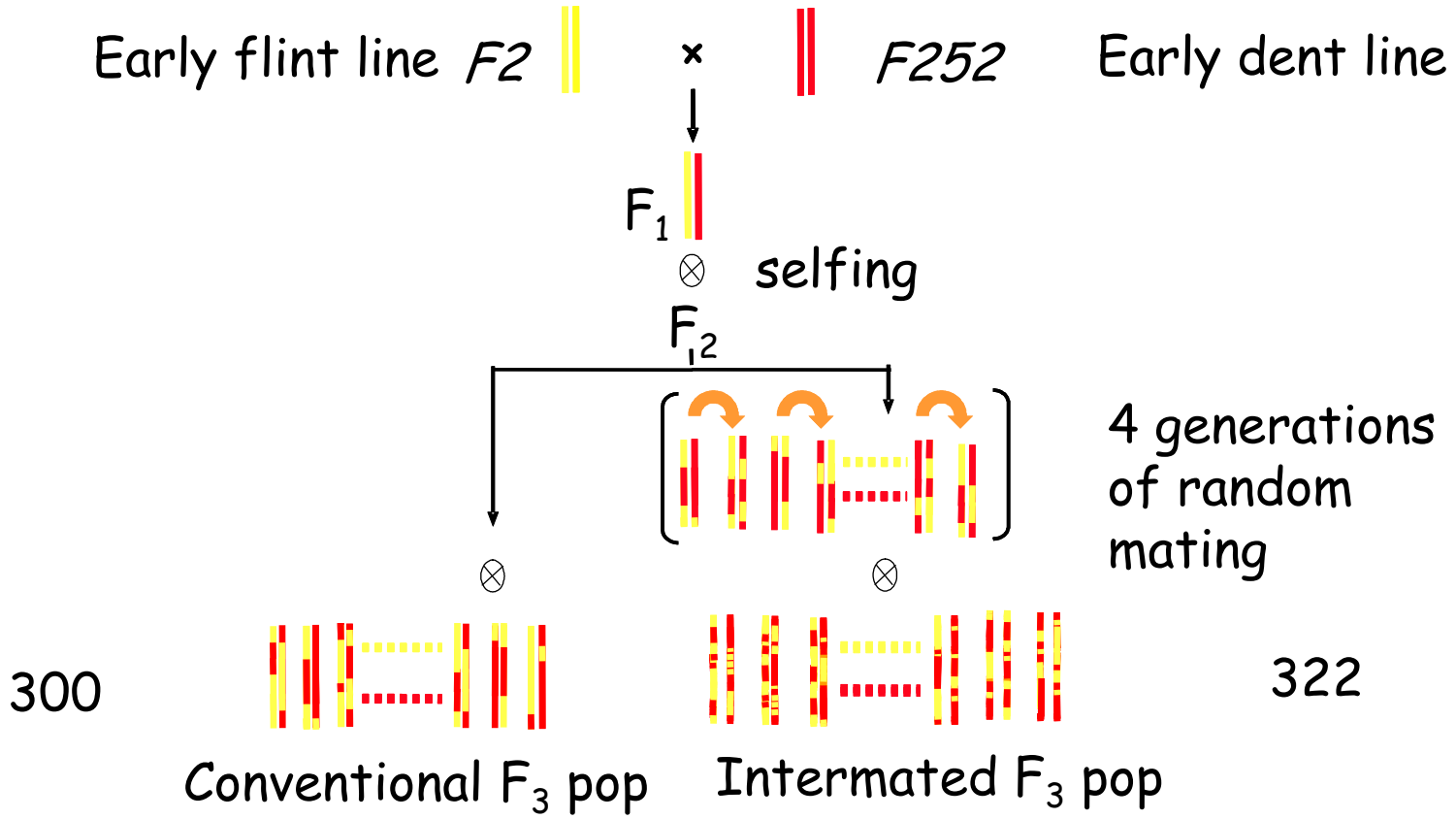
3. Global fine mapping using Advanced Intercross Lines (AIL)

Principle (Darvarsi and Soller, 1995):
add generations of random intercrossing when developing segregating populations to increase the number of recombination events between linked loci.

-> improved accuracy for ordering markers on a map.
-> improved accuracy of QTL detection (provided a higher marker density is used)
- -smaller confidence intervals
- -possibility to detect linked QTL previously masked by coupling or repulsion

Experimental evaluation of the interest of AIL by comparison with a conventional population with similar size, same parental lines and evaluated in the same environments

Plant Material



$F_{3:4}$ families crossed with tester *MBS847* and hybrid progenies evaluated in 10 field trials (5 in 1999, 5 in 2000) for Dry Grain Yield (DGY), Grain Moisture at harvest (GM) and Silking Date (SD)

Phenotypic evaluation

		Parental lines F2	Parental lines F252	Conventional population	Intermated population
DGY (qx/ha)	mean	94.6	93.9	91.6	90.9
	σ^2_G			21.2 *(17.5-26.2)*	12.0 *(9.7-15.2)*
	H^2			0.81*(0.77-0.84)*	0.73 *(0.68-0.77)*
GM (%)	mean	33.2	29.6	32.1	31.8
	σ^2_G			0.74 *(0.61-0.91)*	0.75 *(0.63-0.92)*
	H^2			0.82 *(0.79-0.85)*	0.83 *(0.80-0.85)*
SD (day)	mean	207.3	210.3	210.0	209.1
	σ^2_G			1.34 *(1.11-1.64)*	1.78 *(1.49-2.16)*
	H^2			0.84 *(0.81-0.86)*	0.86 *(0.84-0.88)*

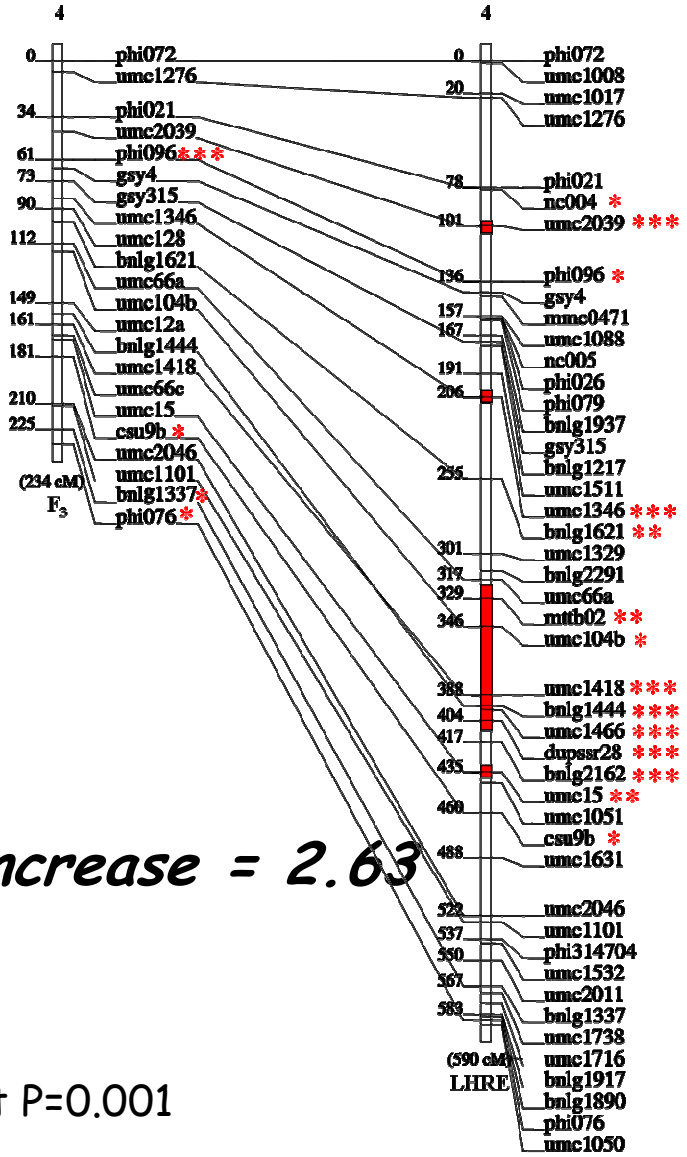
Population means lower than average parental value for DGY
=> epistasis with positive interaction between alleles from same parents

Genetic variance for DGY lower in intermated pop. (factor almost 2)
=> recombination events broke some clusters of QTL in coupling phase.

Genetic Maps

Conventional F_3:

- 199 markers
- 2115.7cM
- 11.2cM between 2 markers



Intermated F_3:

- 358 marqueurs
- 5568.3 IcM
- 16.2 IcM between 2 markers

174 common markers

Rate of map length increase = 2.63

▬ : allelic frequency ≠ 0.5 at P=0.001

*, **, ***: segregation distorsion P=0.05, 0.01, 0.001

Elements on QTL detection methods

PlabQTL (Utz et al., 2003) using Composite Interval Mapping approach (Zheng, 1994)

Lod threshold obtained from 1000 permutation test for a genomewide level risk of 10%

Consistent with increased nb of recombination events,
LOD threshold F3 = 2.7 < LOD threshold iIF3 = 3.12

Confidence Intervals: 2-LOD units fall

QTL for IF3 and their CI projected onto the F3 map to compare positions

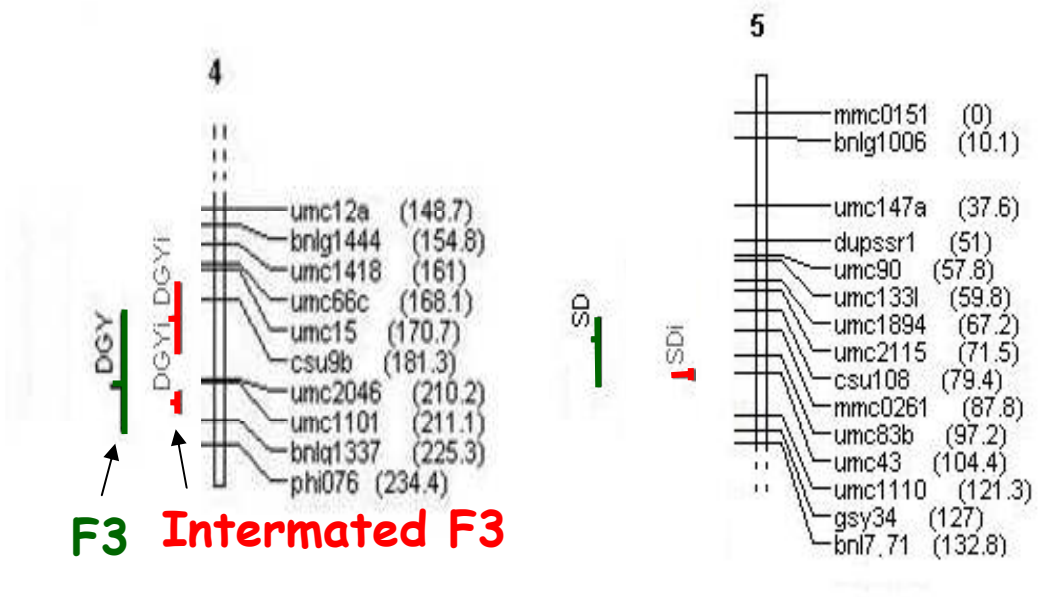
Cross-validation in the two populations, to correct estimation of QTL effects for selection bias (Beavis, 1994)

QTL Results

QTL detection	Conventional F3			Intermated iIF3		
Whole data set analyses	DGY	GM	SD	DGY	GM	SD
Nb QTL	9	9	12	11	5	5
R^2_P	48.4	45.5	45.7	41.5	20.6	26
Nb of QTL in both pop over the total	7 over 30			8 over 21		
Average CI of QTL in both pop.	CI=32.9			CI=14.3		

Results confirm the gain of precision of Intermated F3 / F3. Factor = 2.31, consistent with map expansion factor

Exemples of QTL detected in the two populations



After projection onto F3 map,

- One F3 QTL for DGY split into two QTL in coupling phase in iF3
- More accurate QTL position estimates

QTL Results

QTL detection	Conventional F3			Intermated iF3		
Whole data set analyses	DGY	GM	SD	DGY	GM	SD
Nb QTL	9	9	12	11	5	5
R^2_P	48.4	45.5	45.7	41.5	20.6	26
Nb of QTL also detected in other pop	7 over 30			8 over 21		
Average CI of QTL	CI=32.9			CI=14.3		
Cross-validation analyses						

However, fewer QTL detected in iF3 and poor level of consistency between the QTL detected in the two pops.

Power issue?
Based on F3 av. QTL effect R^2=5% power about 60%, one would expect 60% of the F3 QTL to be also detected in iF3.

QTL detection	Conventional F3			Intermated iIF3		
Whole data set analyses	DGY	GM	SD	DGY	GM	SD
Nb QTL	9	9	12	11	5	5
R^2_P	48.4	45.5	45.7	41.5	20.6	26
Nb of QTL in both pop over the total	7 over 30			8 over 21		
Average CI of QTL in both pop.	CI=32.9			CI=14.3		
Cross Validation Analyses (1/5 for validation)						
Detection set: Average nb QTL	9.0	6.0	8.3	5.7	3.8	5.9
Detection set R^2_P	54.0	42.3	41.0	30.6	21.4	33.9
Validation set R^2P	39.5	31.1	20.1	12.2	6.1	17.2
Selection bias DS/VD	1.4	1.4	2.0	2.5	3.5	2.0
Estimated average individual QTL R^2	4.4	5.2	2.4	2.2	1.6	2.9

For DGY and GM cross-validation gives a different picture of trait architecture in iF3/F3. *Suggests QTL with smaller individual effects in iF3 than F3.*

-> Actual number of QTL certainly higher than expected from F3 results

Conclusions

Multiparental designs (see also Buckler et al., 2009)
-> Offers opportunities to accelerate genetic gain by MAS, thanks to a faster assembling of favorable alleles contributed by different parents
-> suggest allelic series with gradient of effects

Nature of molecular variation and physical size of these alleles?, Mode of action?
-Probably confounding effect of linkage (see AIL)
-Can be resolved to local haplotypes using fine mapping scales (see Chr. 10), LD magnitude highly variable
-Role of non coding regions (should not be discarded in LD scans and related approaches)

Perspectives and questions

Strong interest for combining LD and linkage analyses

New possibilities for dense genotyping of parents (and population of possible parents)
-> « rough » estimation of effects to facilitate identification of promising crosses to form new breeding populations
-> Lower density genotyping of progeny within these populations and imputation (eg. Nested association mapping, Yu et al., 2008) -> effects to be used for selection

Genomic selection and/or « Post-QTL » marker assisted selection?
-> is there room for individualized QTL to facilitate allele tracking during MAS process?

Acknowledgments

B. Mangin and colleagues for statistical aspects (MCQTL),
J.B Veyrieras for the simulations
E. Bouty for preliminary QTL detections

Lab: M. Merlino, V. Combes, F. Dumas, D. Madur, C.L. Ky, for genotyping

INRA experimental stations
Lusignan: X. Charrier
Mons: D. Rabier
Moulon: P. Bertin, P. Jamin, D. Coubriche, S. Jouane, C Bauland

The private companies members of Promaïs:
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
Genoplante


Thank you

Conclusions

- This study confirms the interest of AIL for improving QTL detection accuracy

- However, for complex traits such as DGY and usual population sizes, recombination events split clusters of QTL into pieces, each of small individual effect that become hardly detectable.

- Interest of AIL for Marker Assisted Selection?
  QTL effects more accurate in AIL but fewer QTL detectable.

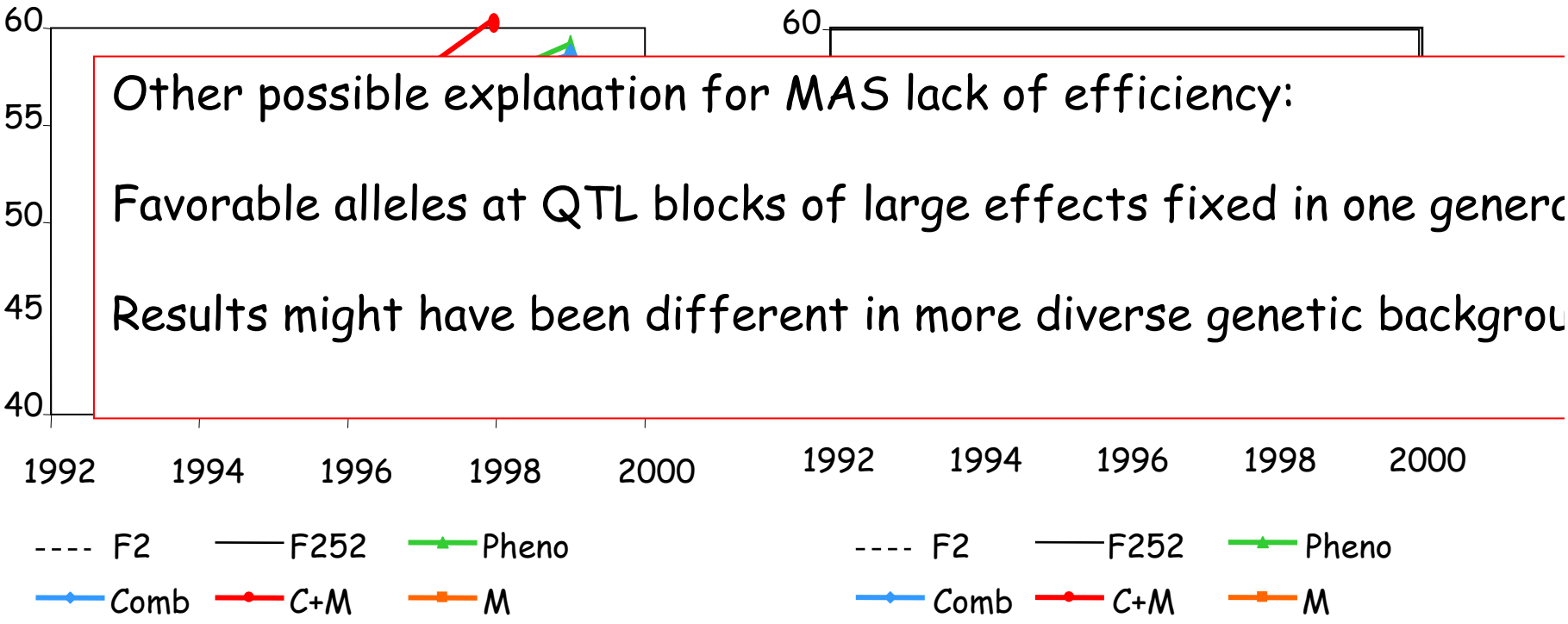
- Our results highlight the fact that:
  - ✓number of QTL involved in trait variation certainly high (at least 35 for DGY, 50 for GM…).
  - ✓numerous situations on linkage and repulsion.
  - ✓this might reduce long-term efficiency of MAS based on QTl detected in intial population.
    -> genomic selection more appropriate
  - ✓might partly explain the lack of efficiency of MAS in

Already shown by Schön et al., 2004

Genetic gains in MAS project carried out in the F3 population

Expected gain based on predicted values (index. 0.1 t ha^{-1})

Observed genetic gain (index. 0.1 t ha^{-1})



Comb= MAS with marker effects estimated at each generation
C+M =MAS with marker effects estimated from intial QTL detection

Outline

1. Organization of genetic diversity of traditional Open Pollinated Varieties from Europe (with respect to that of America)
2. Quantitative Trait Loci involved in flowering time variation: linkage mapping in multiparental designs, Meta-Analysis of published studies, linkage based fine mapping and association genetics
3. Will genotyping assisted selection make it possible to accelerate or maintain genetic gain (per cost unit) for complex traits?

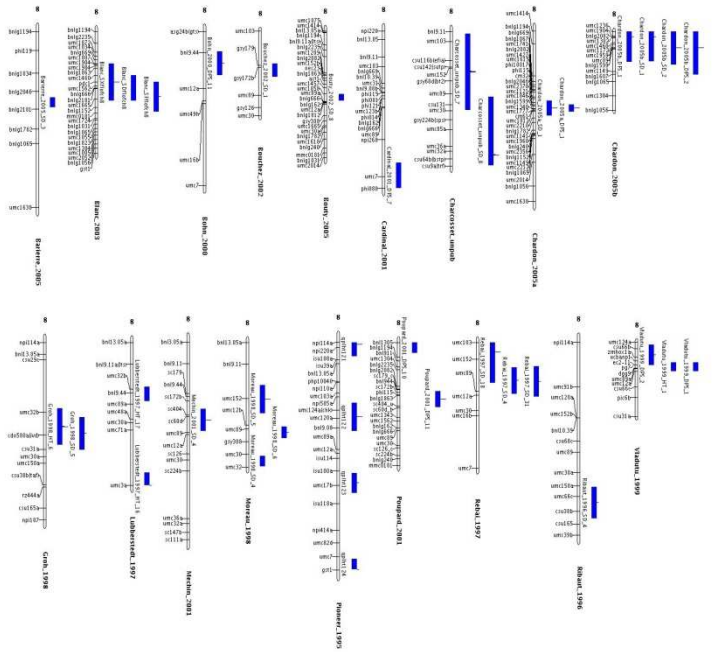
Conclusions and Perspectives

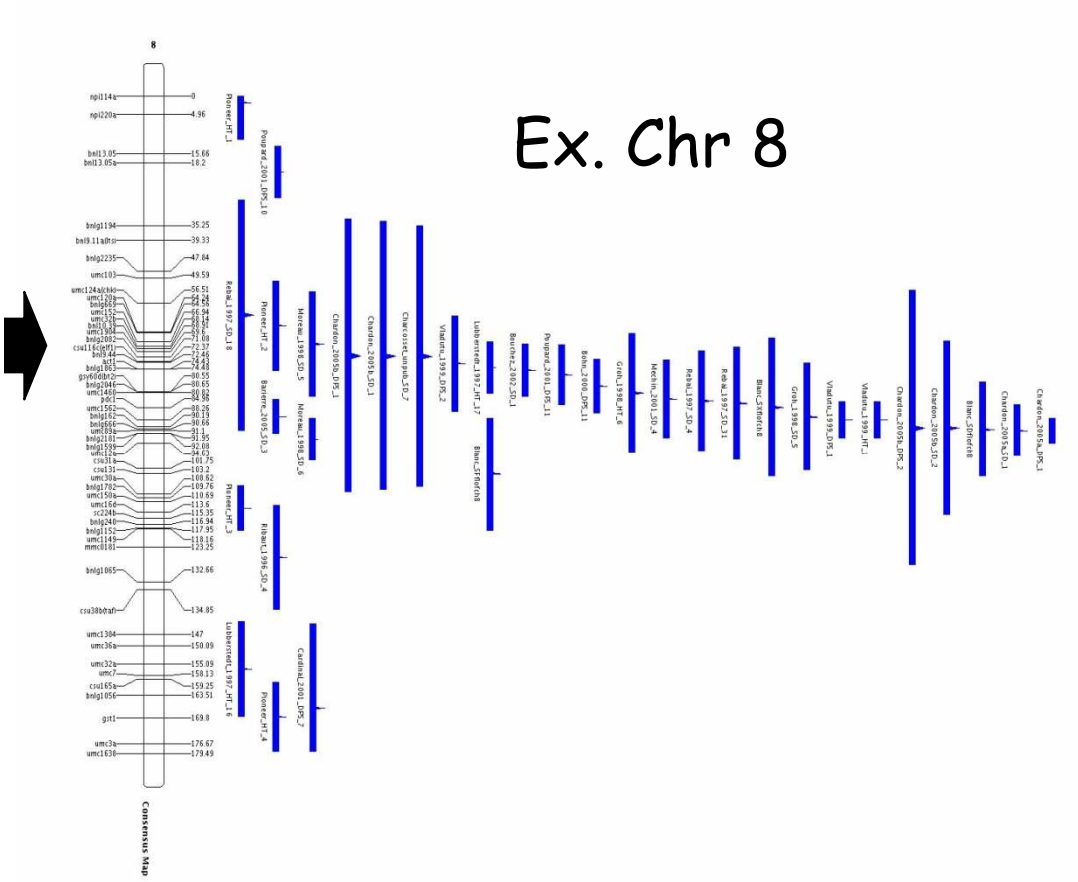
1. Organization of genetic diversity of traditional Open Pollinated Varieties from Europe (with respect to that of America)

2. Quantitative Trait Loci involved in flowering time variation

Second approach: a posteriori Integration of existing QTL mapping results using Meta-analysis

Step1. Map synthesis and QTL projection for litterature results (Veyrieras et al. 2007, BMC Bioinformatics)



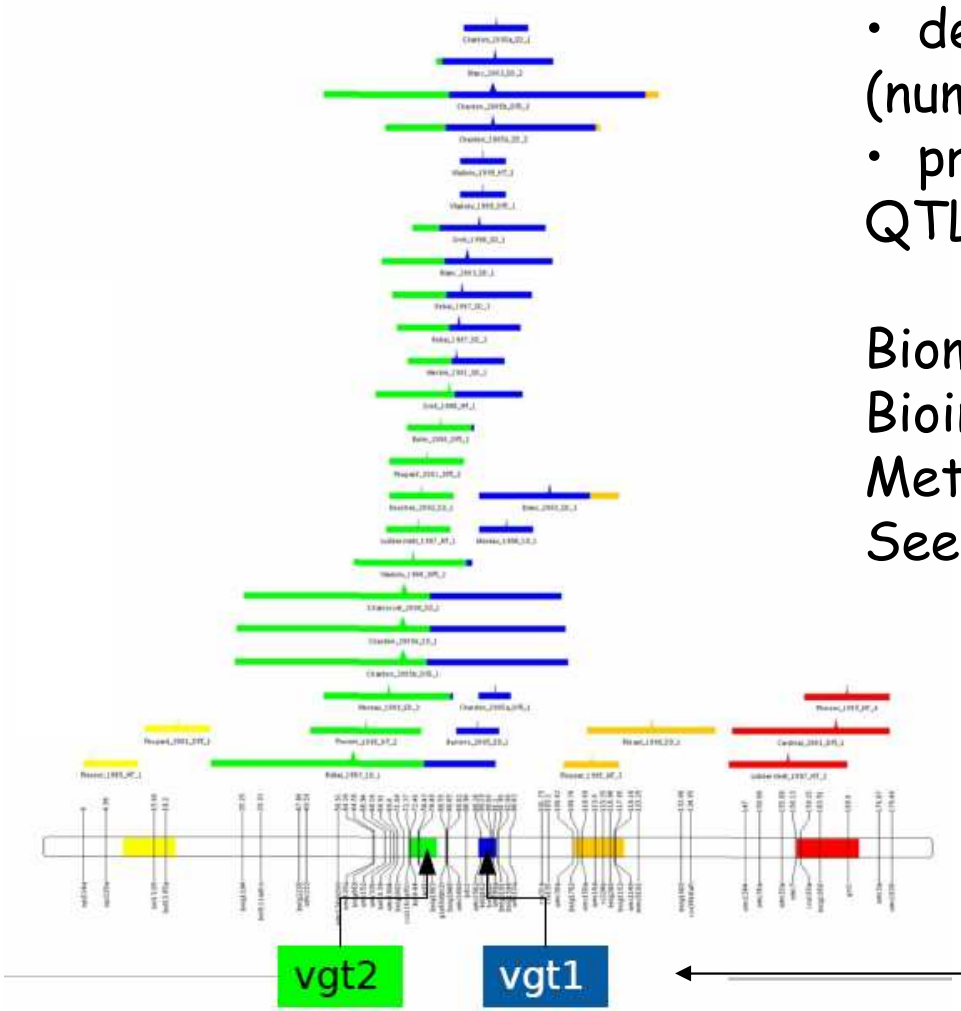


Ex. Chr 8

Algorithms for :

- optimized synthesis of maps
- Projection of QTL

Step 2: clustering of QTL



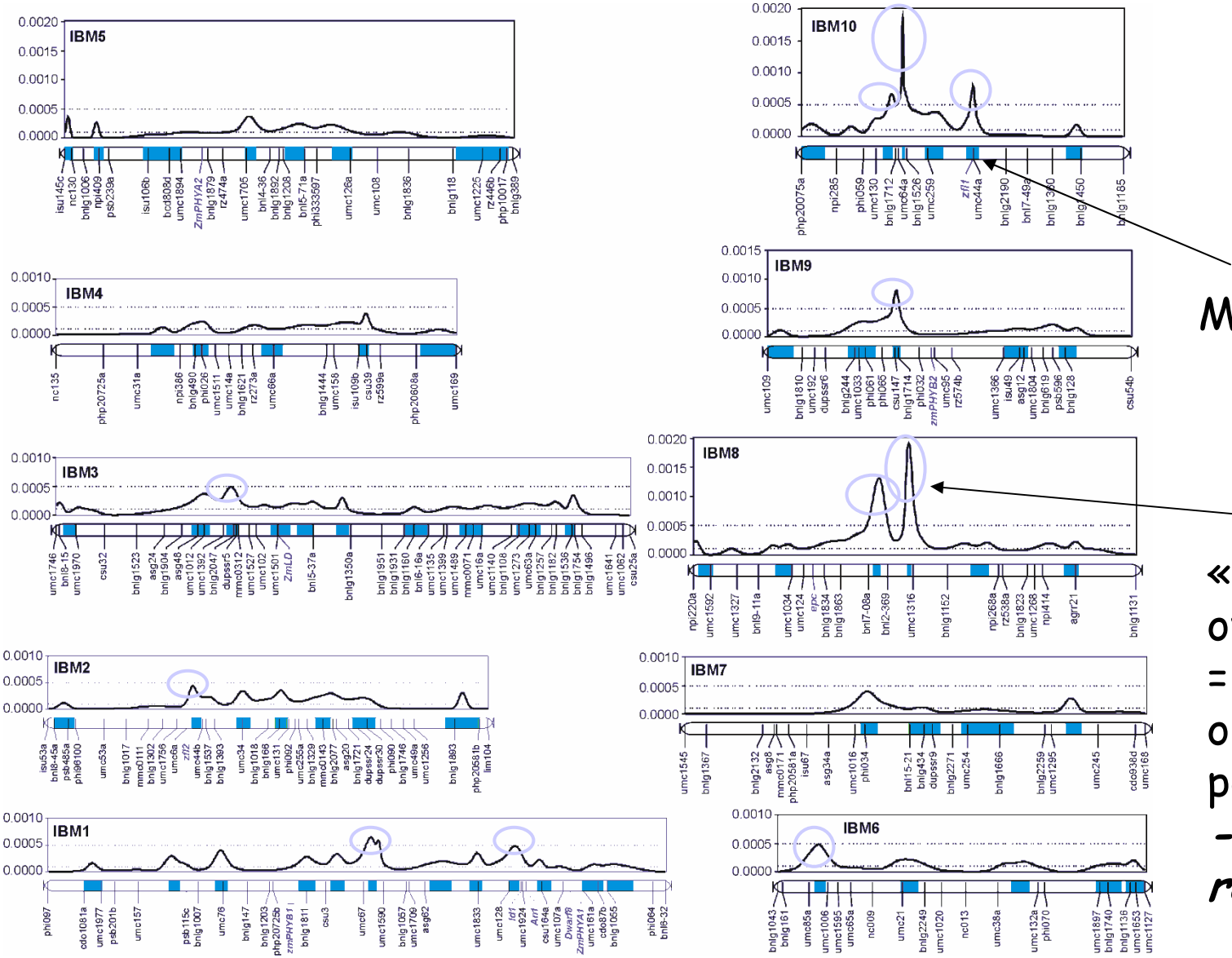
Algorithms for :

- determination of synthetic model (number of underlying QTLs): here 5
- probabilistic assignation of initial QTL to metaQTL

Biomercator (Arcade et al., 2004, Bioinformatics, coll. J. Joets)
MetaQTL (Veyrieras et al., 2007)
See poster 216

Increased mapping precision of meta QTL / most precise initial QTL,
validated for vgt1 thanks to positional cloning (Salvi et al., 2000, 2007)

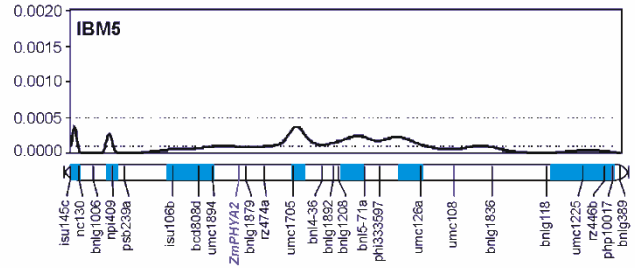
A global view of FT control in maize: 317 initial -> 67 metaQTL (Chardon et al., 2004, Genetics)

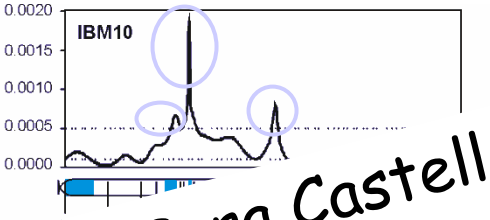


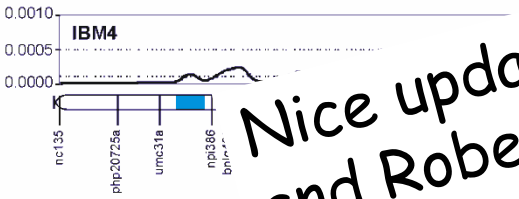
MétaQTL

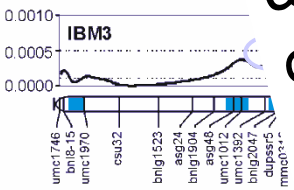
« empirical overview » = p(QTL in one population) *-> major regions*

A global view of FT control in maize:
317 initial -> 67 metaQTL (Chardon et al., 2004, Genetics)



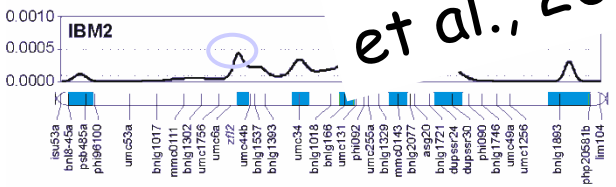


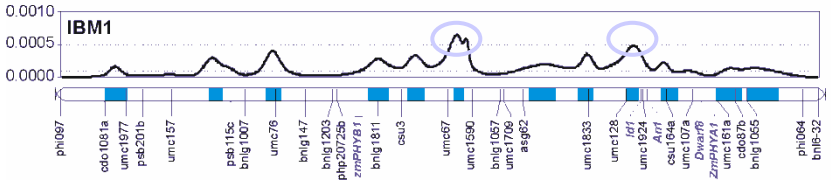


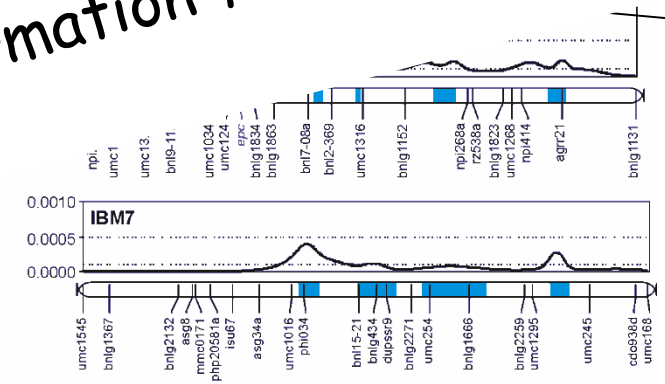


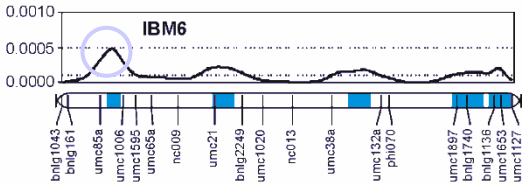
Nice update by Silvio Salvi, Sara Castelletti and Roberto Tuberosa in Maydica issue just dedicated to Ron Philips

Including information from NAM (Buckler et al., 2009)









QTL

« empirical overview » = p(QTL in one population)

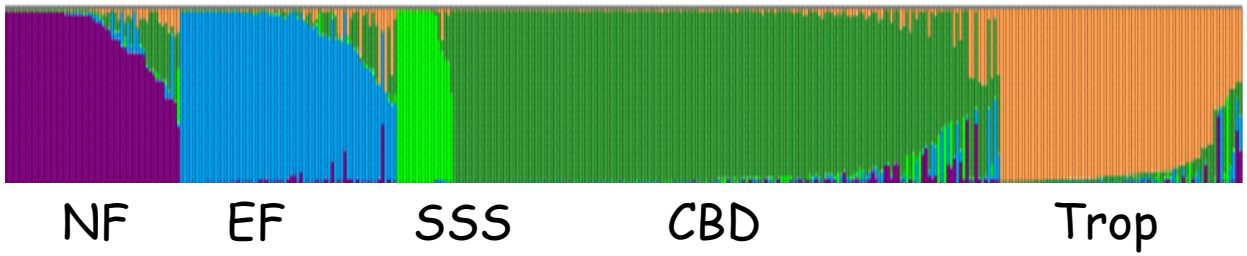
-> major regions

Towards identification of causal polymorphisms: association genetics and linkage based fine mapping

Definition of a panel of 375 representative inbred lines complementing US panel (Camus-Kulandaivelu et al., 2006), see poster 177

Evaluation of phenotypic traits (NIRS for kernel traits)

Analysis of global population structure using neutral SSR markers



↳ *Effect on trait variation?*

Evaluation of phenotypic traits and relationship with population structure

Trait	h2	R2*
Male flowering (GDD)	**0.97**	**0.51**
Vitreousness (% endosperm)	0.92	0.31
1000 Kernel weight (g)	0.77	0.27
Kernel oil (%)	0.86	*0.06*
Kernel protein (%)	0.88	*0.04*
…		

*

$$T_j = a_o + \Sigma_1^{k-1} a_i g_{ij} + e_j$$

Group effects for FT and vitreousness
-> Population structure (and kinship) need to be taken into account in association tests (Yu et al., 2006)

Trait	NF	EF	CBD	SS	Trop.
Male flowering (GDD)	**932**	950	1034	1065	**1201**

Panel used so far for investigating genetics of kernel traits and flowering time, first based on candidate genes

kernel traits

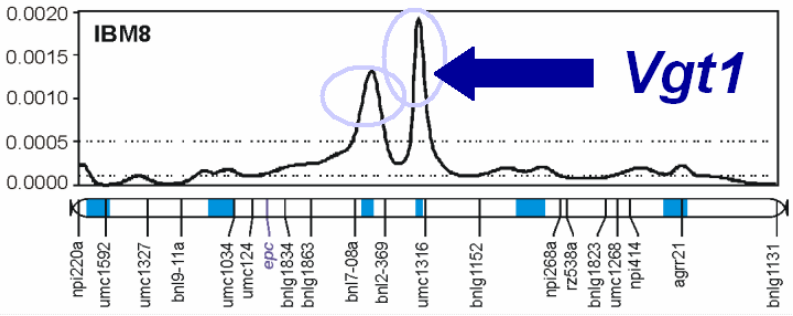
- Candidate genes (Sh2, Bt2, …)
- *epistatic interaction between transcription factor O2 and its target CyPPdk* (Manicacci et al., 2009, Plant Physiol.)

Flowering time

- Candidate genes, *sometimes disappointing so far (eg. ZmHd3a, Chr 6)*
- In depth analysis of 300 kb region between tb1 and D8 (Camus-Kulandaivelu et al., 2006, 2008, Genetics)
- *fine mapped regions: Vgt1 and Chr 10 major QTL*

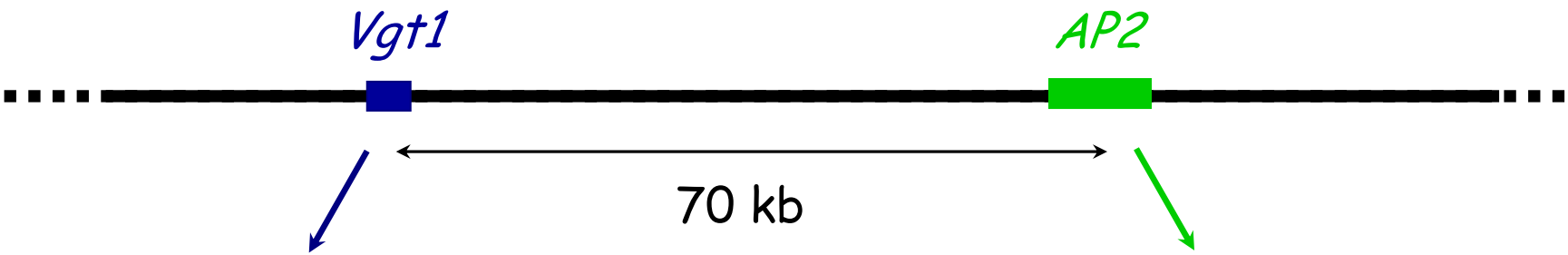
Case study 1: Vgt1 and flowering time

(Ducrocq et al. Genetics, 2008)



Region repetively detected as QTL (see meta-analysis, Chardon et al.. 2004)

Positional cloning by (Salvi *et al.*, 2007, PNAS)



non coding 2.7 kb region

Association with FT (95 lines)

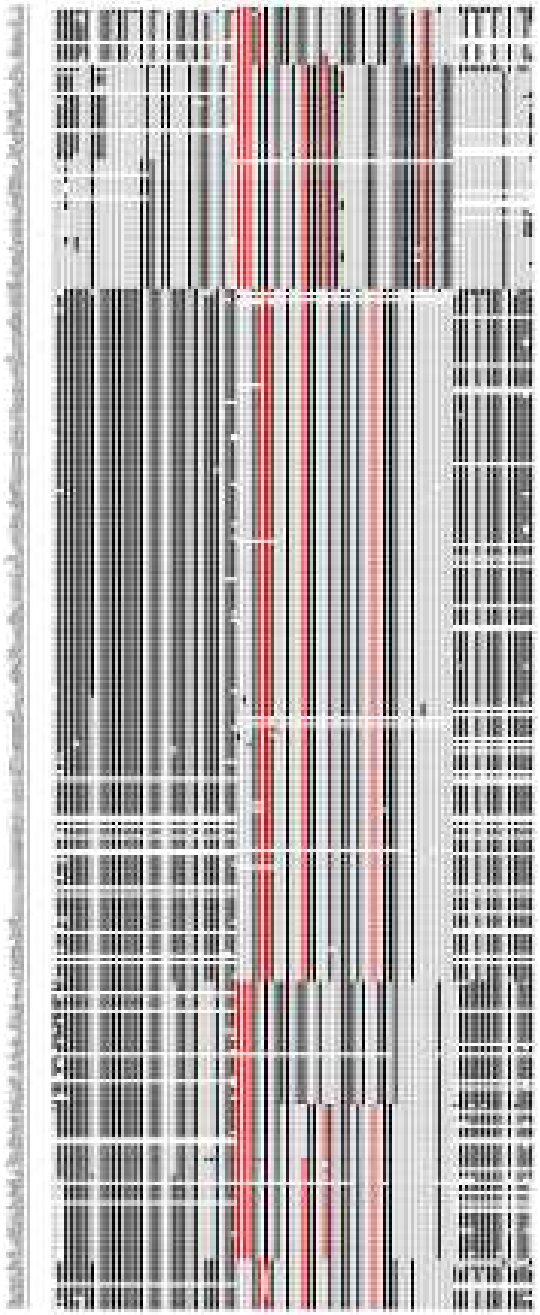
Regulatory element

Overexpression / Repression

Gene inhibits FT

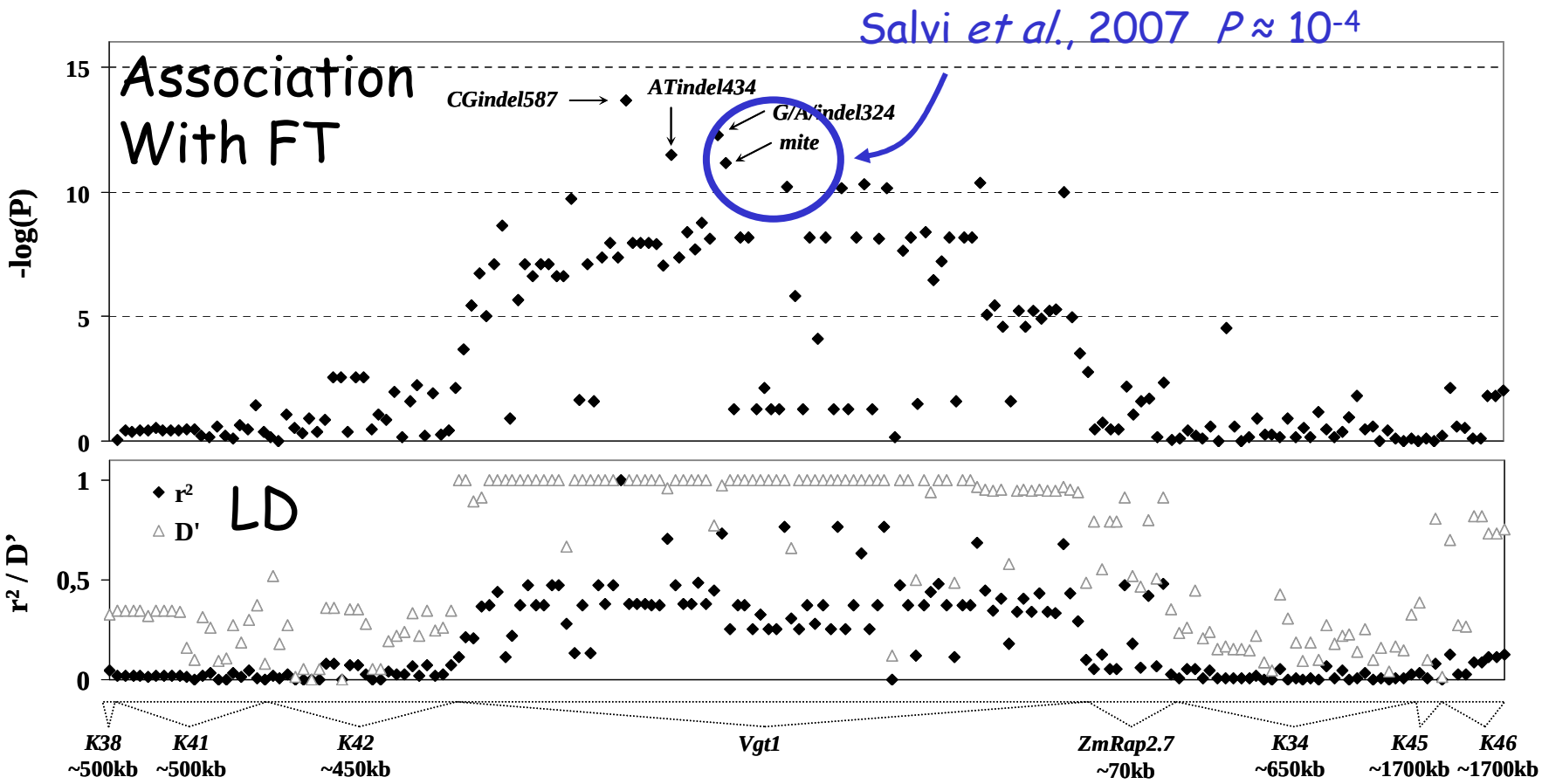
Polymorphism investigation:
sequencing amplicons for the whole 375 inbred lines collection (S. Ducrocq, D. Madur, coll. CNG, KWS)

- here vgt1 (approx. 2.5 kbp) -> suggests ancestral haplotypes
- Ap2,
- more distant regions as controls



Investigation of Vgt1 – AP2 region in 375 inbred panel

(S. Ducrocq, et al., 2008)



Segments analyzed and distance to vgt1 (kb)

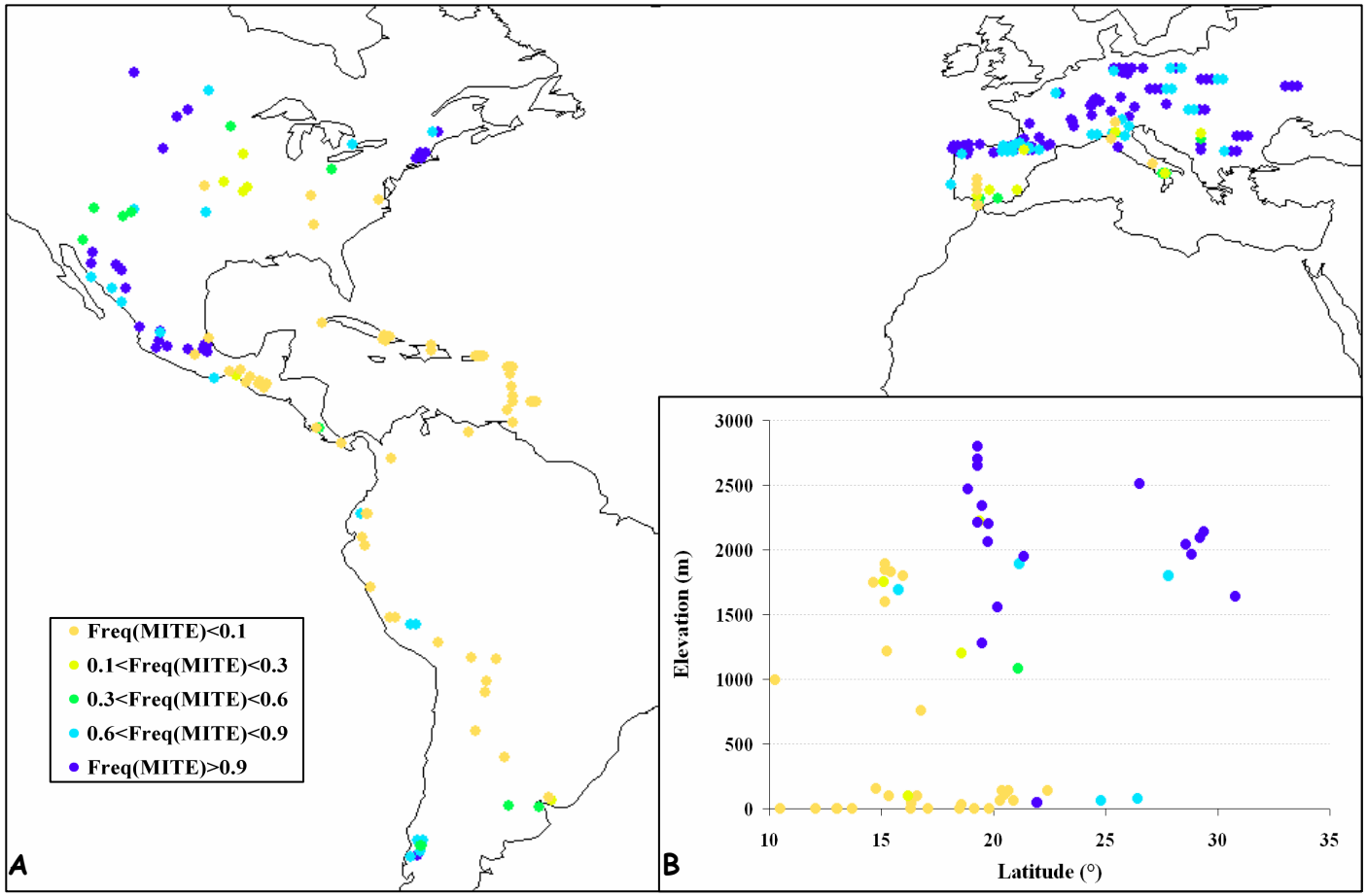
Very high significance levels -> *power due to population size and balanced alellic frequencies within groups (good representation of early flowering materials)*

Strong effect of the region (100 GDD, $R^2 \geq 4\%$ after accounting for populations structure), probably provides an « upper » limit for what can be observed for FT in this panel

GCindel1587 vs. mite: recombination with high frequency in « Iodent » dent family, effects deserve further comparison in diverse materials. Mite is at worse an excellent proxy.

Major role of vgt1 in climatic adaptation

Mite insertion (early allele) frequency



Adaptation to cool temperate climates but also differenciation of tropical « highland » vs. « lowland »

Absence in Andes consistent with migration (Vigouroux et al. 2008)

3. Will Marker / Omics assisted selection make it possible to accelerate or maintain genetic gain (per cost unit) for complex traits?

Prediction of genetic value based on 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* alleles at QTL *q* (inferred from close markers)

- Combining (*M*) and phenotype (*P*), if available

$$G = b_p P + b_m M$$

- opportunities to monitor the assembling of alleles by crossing complementary individuals

Yield (in fact economic index)
All four parents are needed to create the ideotype assembling all superior alleles

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	*

(Blanc et al., TAG, 2006)

Conclusions

Northern Flint divergence is striking (see Doebley 1986), After extremely fast expansion in North America, rapid expansion in Europe and contribution to new genetic pools
-> Calls for the development of specific resources connected to American designs (KBBE project, Promaïs)

Genotyping offer promises to accelerate breeding process, particularly when broad diversity is addressed
-> important in present context (growing food demand and environmental issues)

Advances in technologies (sequencing, genotyping) offer unprecedented opportunities in MAS:

- Linkage and LD based fine deciphering of allelic series to have reliable markers at QTL and avoid heavy probability computation through generations
- Genomic selection ss. (forget about QTL, calibrate a prediction formula, as for NIRS …)

Optimum is probably a combination of both

To be addressed synergistically with more fundamental issues

- QTLs: true complexity of the architecture of quantitative traits and the role of linkage (*see poster 183*)?
- Alleles at QTLs: number?, nature of molecular variation and physical size?, Mode of action?
- …

Participants

SGV Moulon
(INRA, CNRS, UPS, INAPG)
P. Dubreuil, C. Rebourg, F. Chardon, S. Ducrocq
G. Blanc, L. Moreau, A. Gallais
D. Manicacci, M. Tenaillon, M. Fourmann, S. Barraud
L. Camus-Kulandaivelu,
J.-B. Veyrieras, A. Larièpe, C. Mir, M. Truntzler, S. Bouchet
D. Madur, V. Combes, F. Dumas
P. Bertin, S. Jouane, D. Coubriche, Ph. Jamin
A. Charcosset, M. Falque, J. Joets

INRA Montpellier, St Martin
B. Gouesnard
M. Dupin
J. Laborde

INRA Mons
C. Giauffret, A. Bouchez, C. Bauland

KWS
M. Ouzunova
T. Presterl
M. Kloiber-Maitz

CNG
D. Brunel
C. Dupuits,
H. Mc Kahn

Biogemma
A. Murigneux
J.-P. Martinant
L. Consoli,
P. Dubreuil

CIMMYT
M. Warburton
D. Hoisington

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)

Genomics, association genetics: Génoplante + Gabi-Génoplante

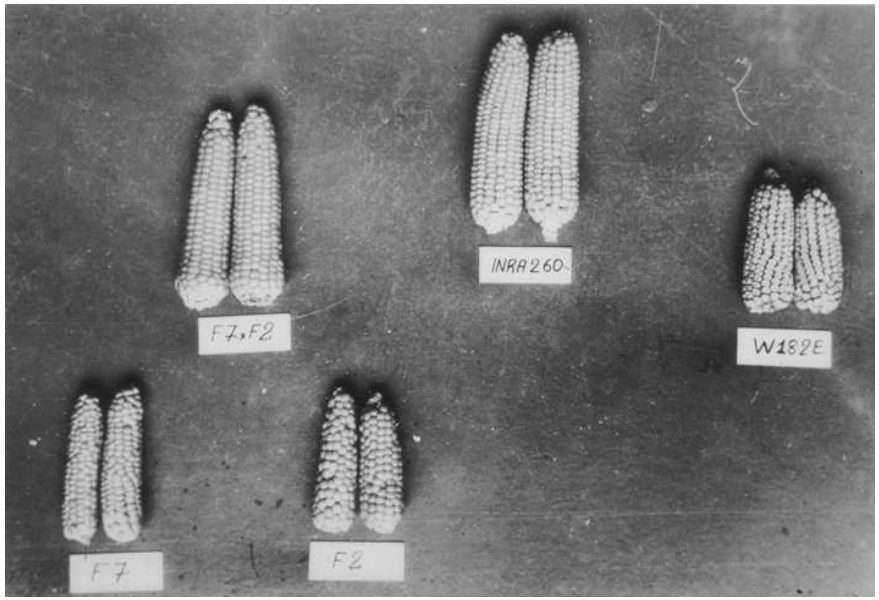
LE MOULON Génétique végétale UMR
INRA Institut National de la Recherche Agronomique
CNRS
PRO-MAÏS
GENOPLANTE
CIMMYT

Maize introduced and cultivated in Europe, rapidly after Discovery, including temperate cool regions (Bock 1539, Fuchs 1542),

Traditional varieties essentially flint (vitreous endosperm)

European Flint basis of hybrid maize breeding in northern Europe by crossing with dent origins

Still remain at least a major source of adaptation



Maize domestication from teosinte (Matsuoka et al., 2002) in Mexico, followed by southwards and northwards migration

Varieties thermic requierement for flowering:

650 GDD

700 GDD

800 GDD

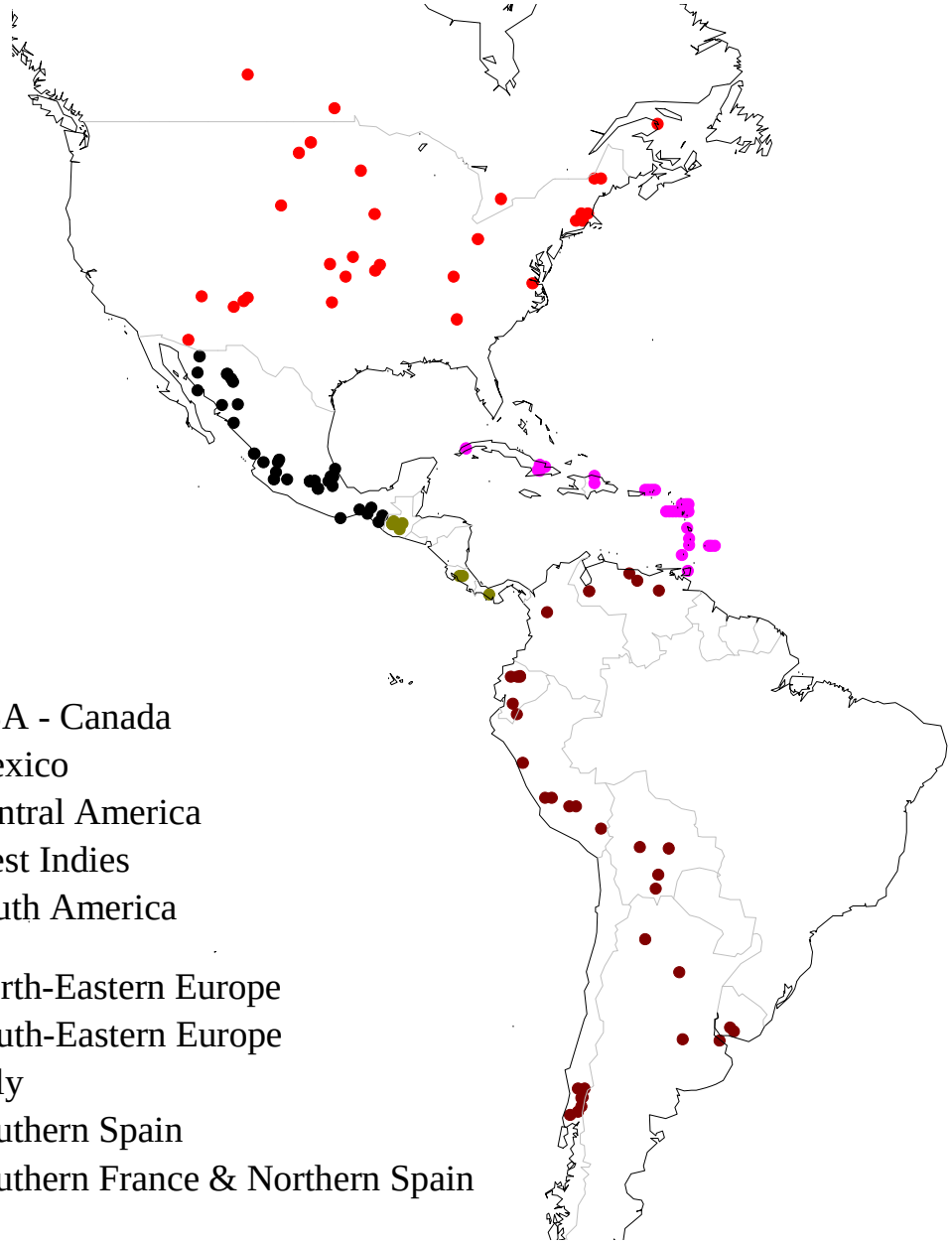
1000 GDD short days

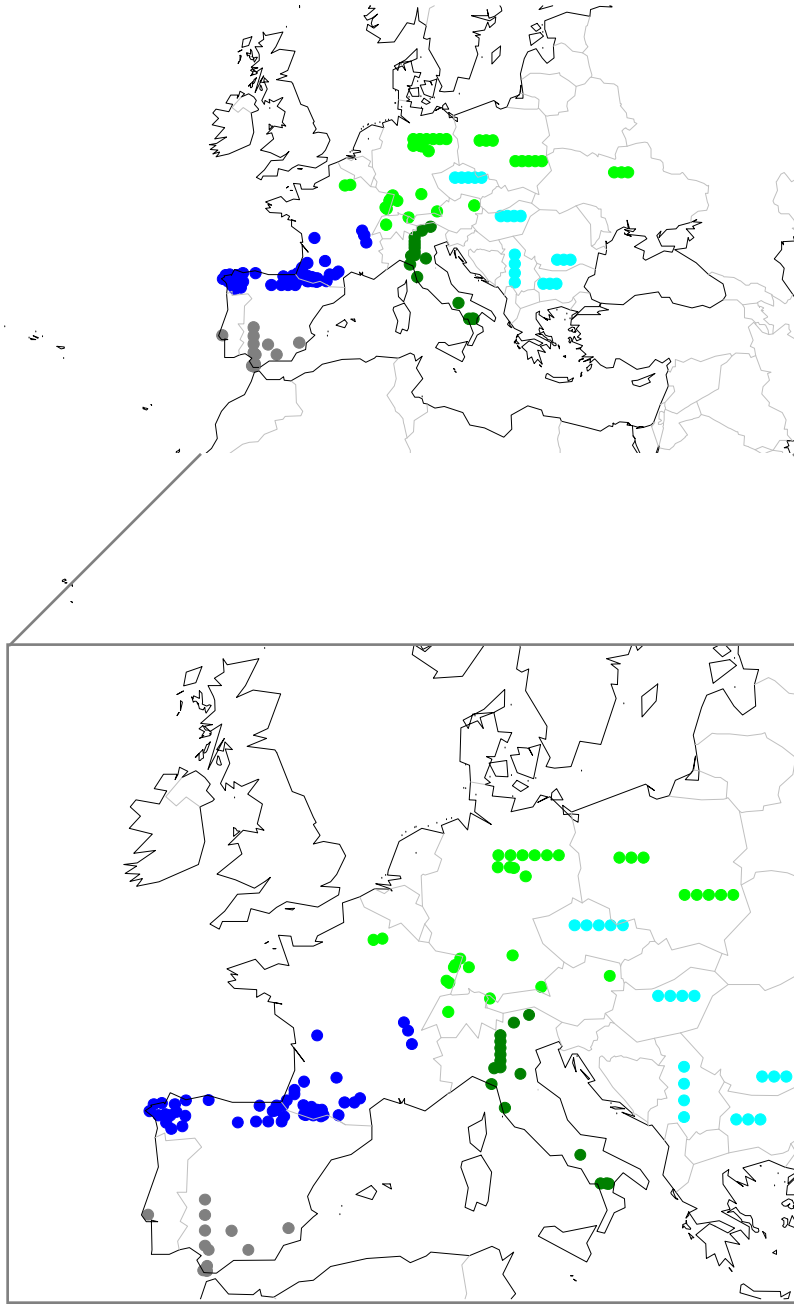
1300 GDD long days

Context

- Replacement of traditional open pollinated varieties of maize by hybrids in many European regions following WW2
- Preservation of this material thanks to collection and conservation efforts (2895 pops for EU). *Utilization limited by lack of information on the structure of diversity within the collections.*
- Investigation of the potential of marker analysis to clarify the phenomena that shaped the diversity of European traditional maize varieties, to better utilize genetic resources collected before the switch to hybrids

Sampling accessions
131 European, 145 American OPVs





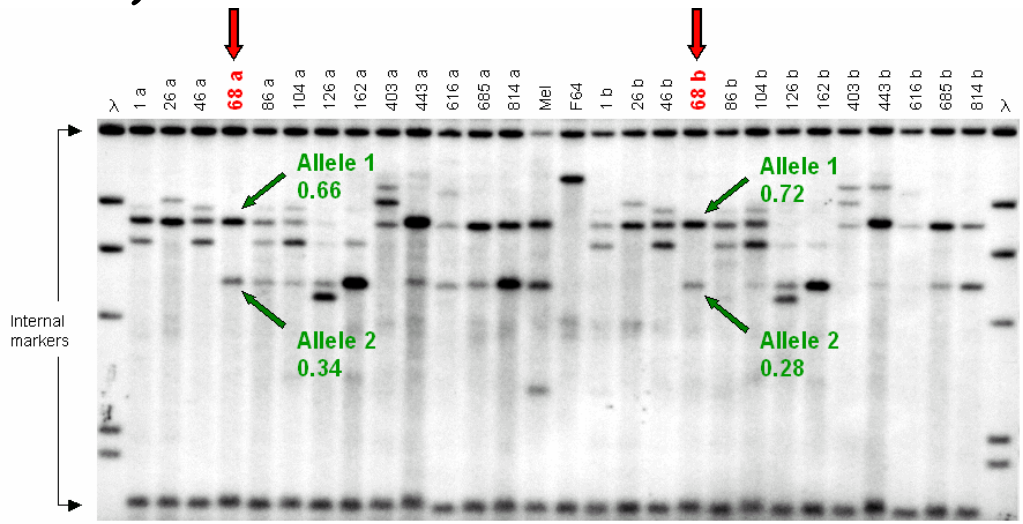
- USA - Canada
- Mexico
- Central America
- West Indies
- South America

- North-Eastern Europe
- South-Eastern Europe
- Italy
- Southern Spain
- Southern France & Northern Spain

Neutral markers based evaluation of diversity

Augmented scale of analyses thanks to DNA pools (2x15 inds) and direct estimation of frequency based on signal intensities

a) 29 RFLP loci



b) 24 SSR loci (coll. CIMMYT), deconvolution

c) SNPs, …

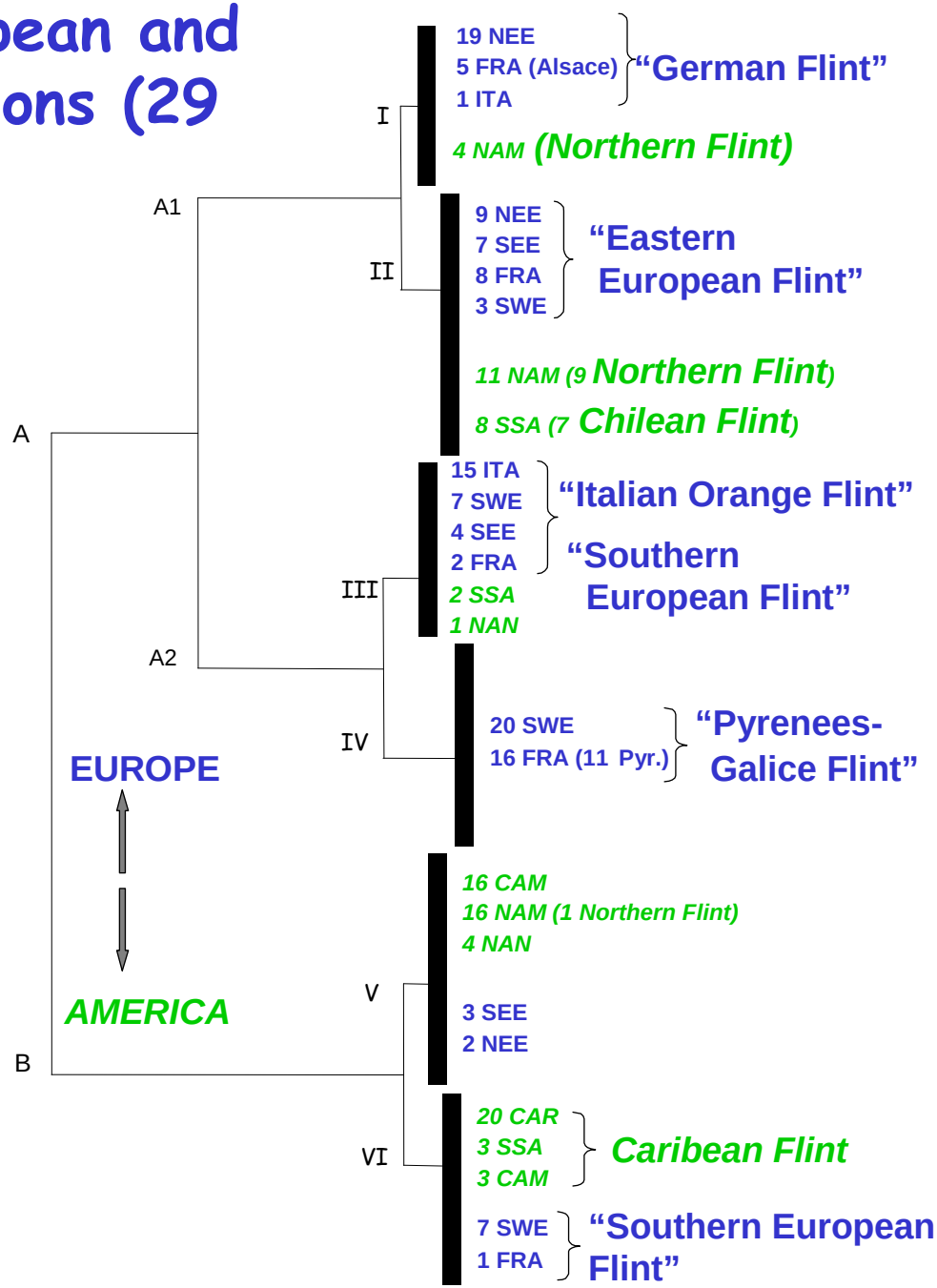
Joint Analysis of 131 European and 89 American maize Populations (29 RFLP), (Rebourg et al., 2003)

European diversity is included in the American one (22 spec alleles vs. 102)

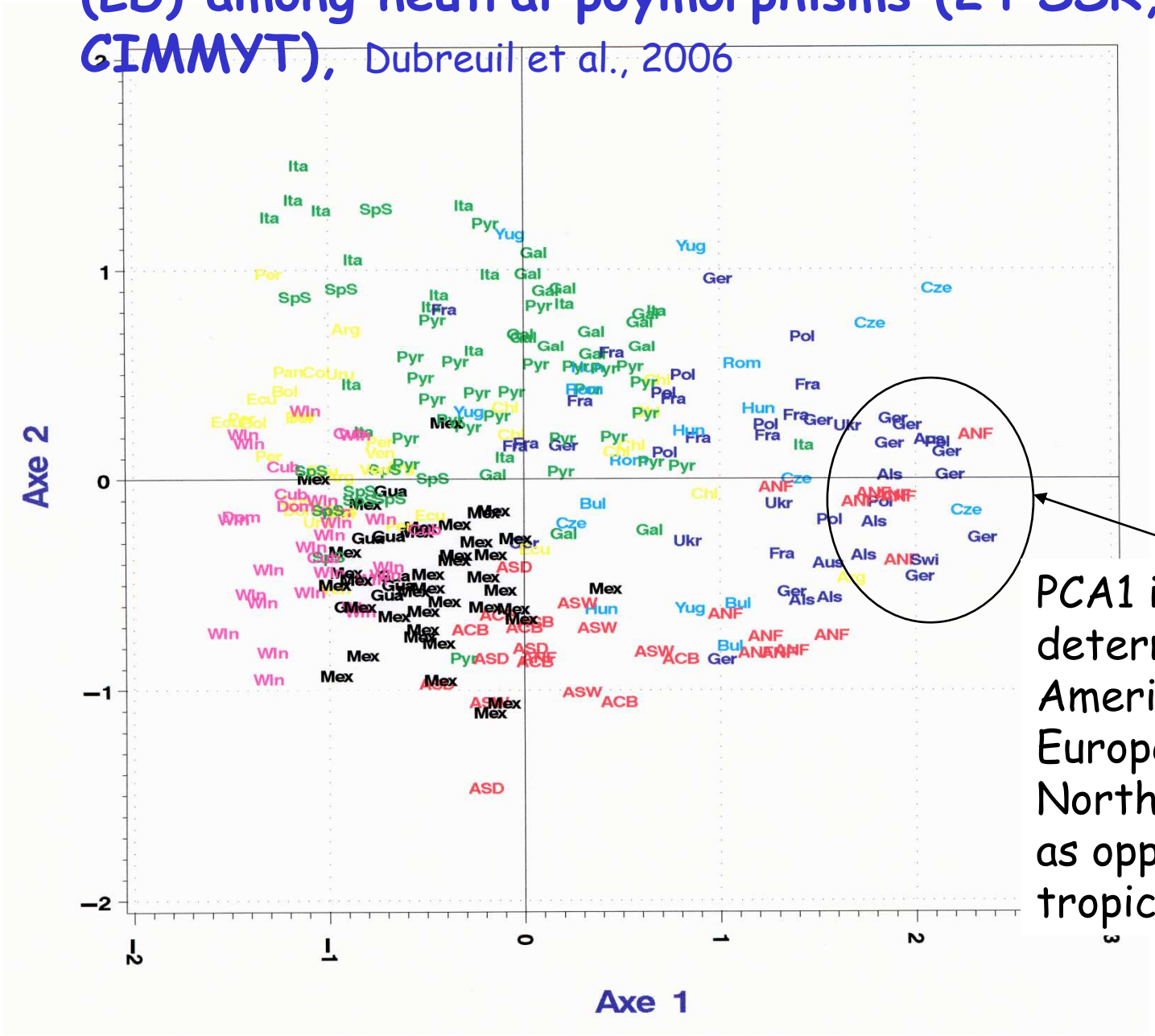
Similarities between :

- Southern Spain and Caribbean material (consistent with 1st introductions by Colombus)
- Northern Europe and American Northern Flints (and also material from southern Chili)
- Italian and South American material

Original group in the Pyrenees



Principal Component Analysis of Linkage Disequilibrium (LD) among neutral poymorphisms (24 SSR, coop. CIMMYT), Dubreuil et al., 2006



America:

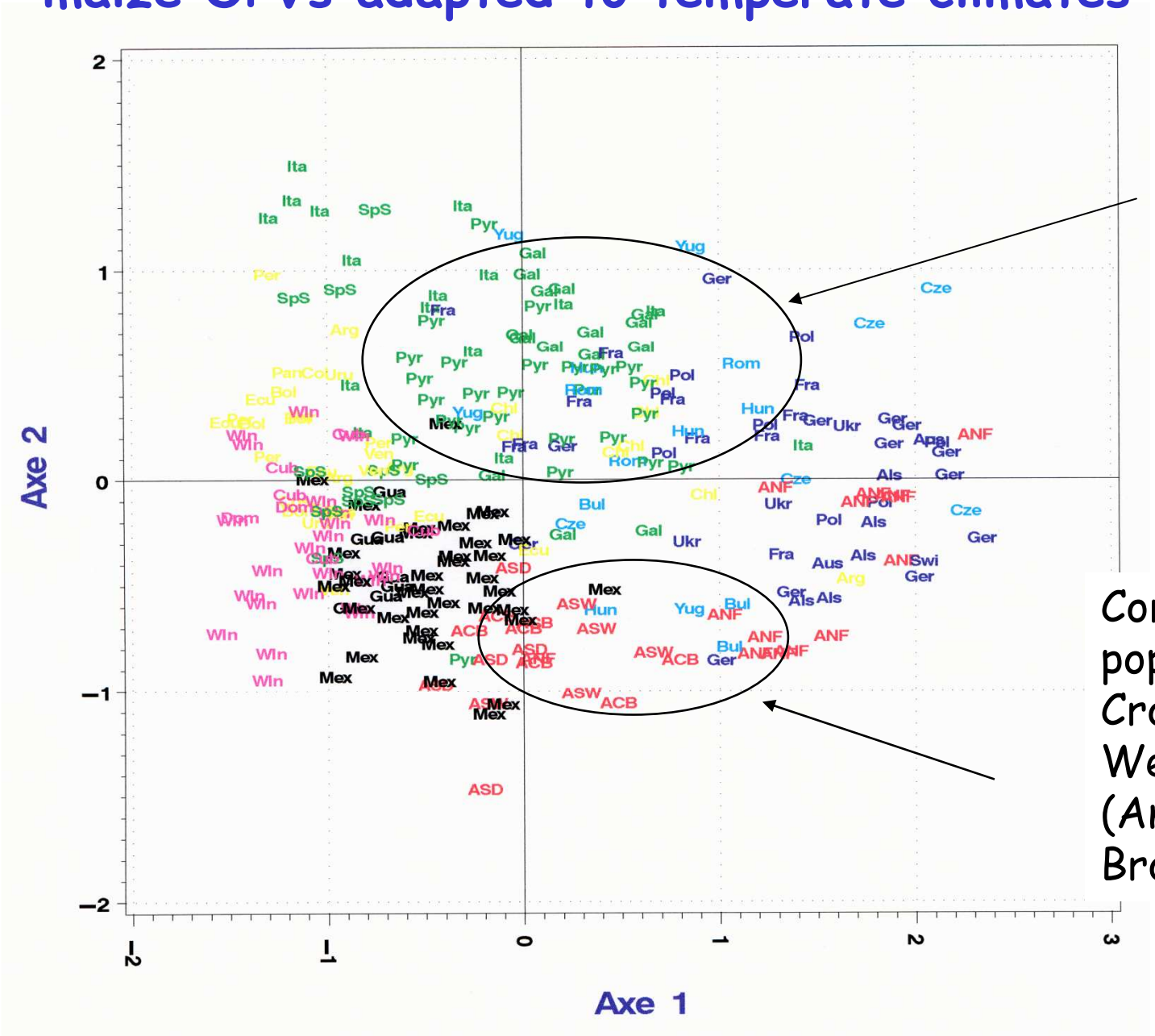
- Mexico - Guatemala
- USA - Canada - Chili
- Caribbean
- Southern America

Europe:

- Spain – Pyr - Italy
- France– North-Eastern
- Balkans

PCA1 is determined by American and European Northern Flints, as opposed to tropical material

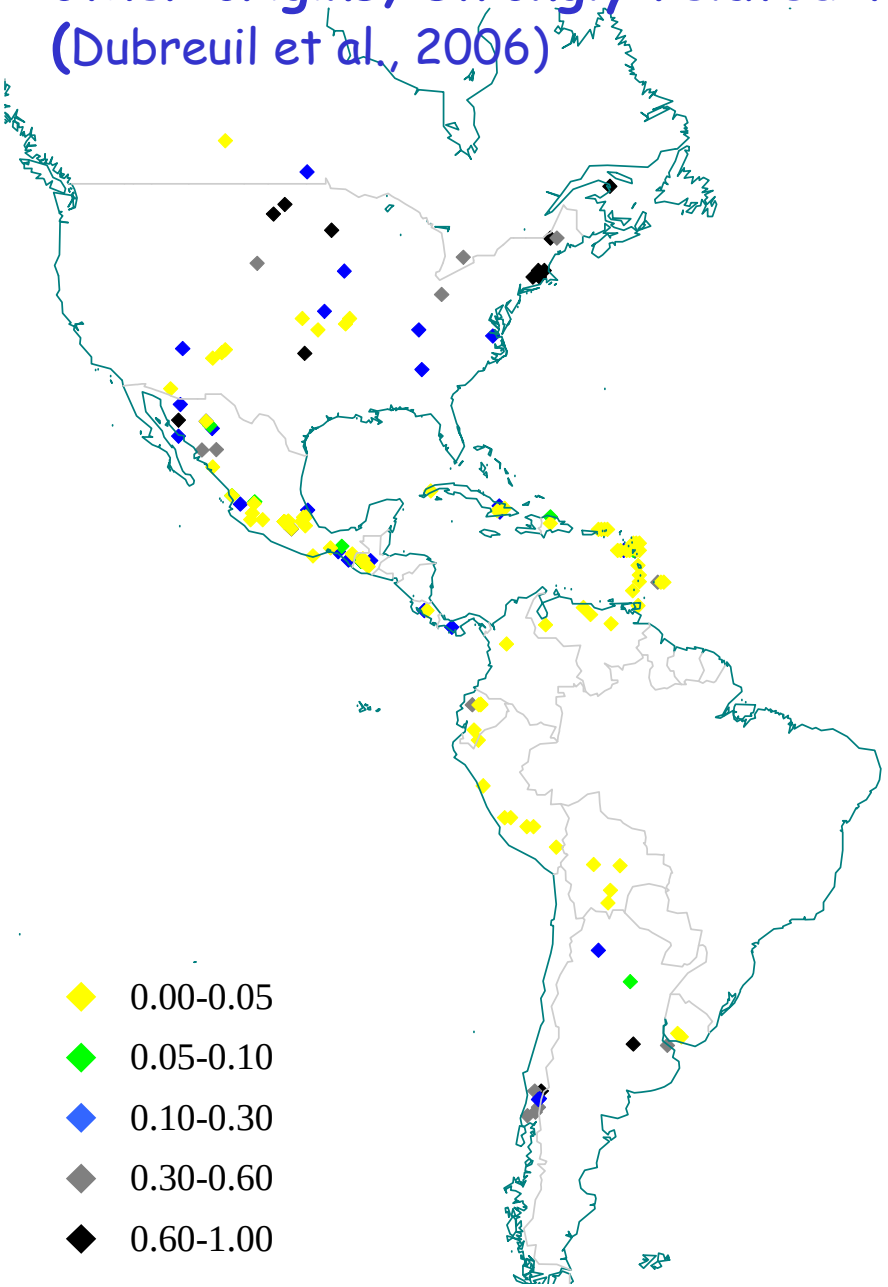
Contribution of Northern flint to the development of new maize OPVs adapted to temperate climates

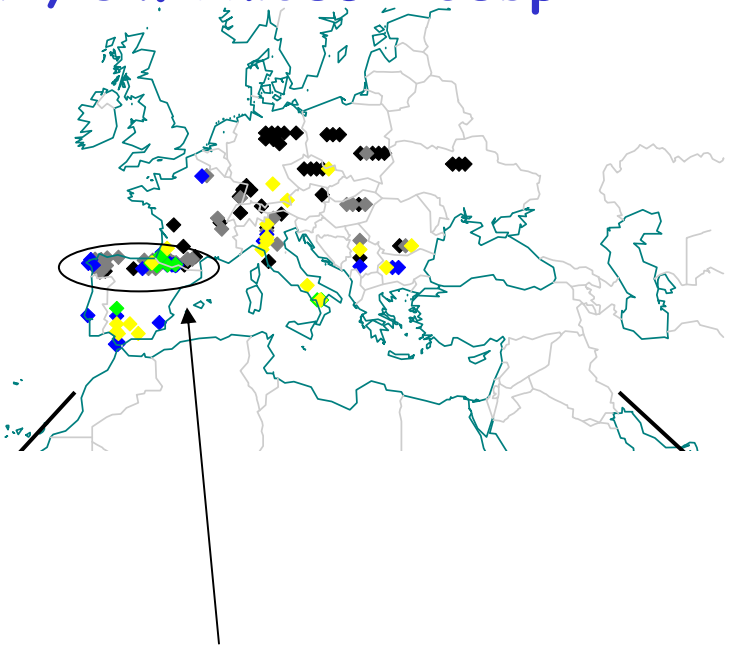


Europe specific OPVs complementary to Corn belt dent for northern Europe hybrids

Corn Belt Dent populations: NF Crossed to Southern Western Dent 1800- (Anderson and Brown, 1952)

Identification of « diagnostic » alleles of Northern Flint (NF) vs. other origins, strongly related to PCA1, ex. Phi085-238bp (Dubreuil et al., 2006)



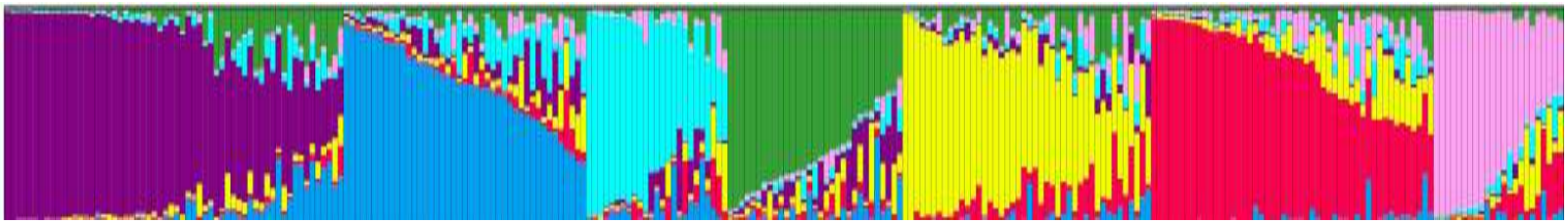


Suggests an European specific hybridization between NF and tropical materials

Modeling of population structure with STRUCTURE software (Bayesian approach, Pritchard et al., 2000)

-determination of ancestral groups underlying population structure, with limited within group LD
-quantitative evaluation of genome proportion inherited from each ancestral group (admixture)
-Also can be viewed as quantitative clustering

-> concludes here to seven groups (Camus-Kulandaivelu et al., 2006)



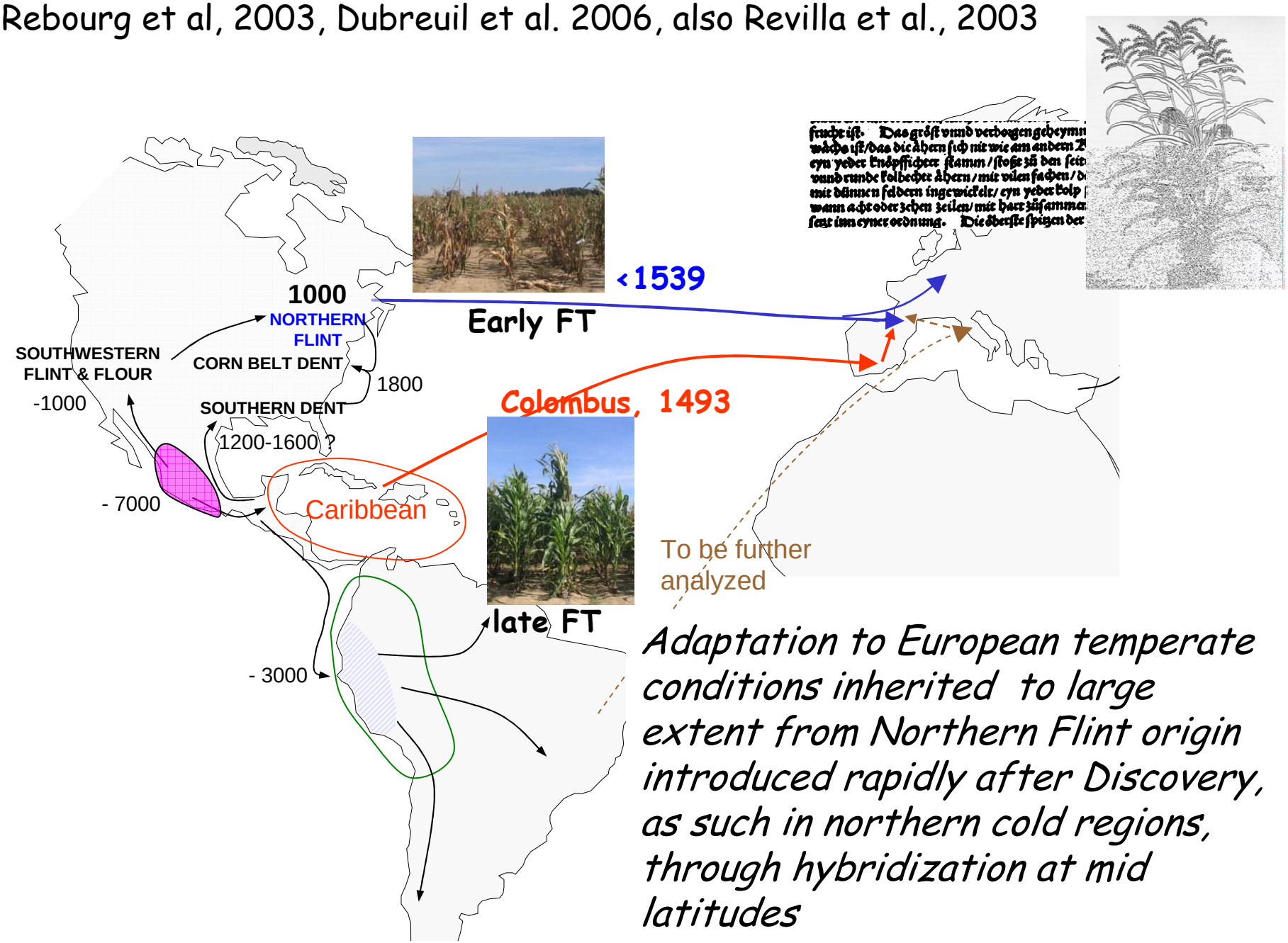
Each population = column

- Northern Flint (Eur.+Am.)
- Pyrenees-Galicia Flint (Eur.)
- Italian Flint (Eur.)
- Corn Belt Dent
- Mexico (highland)
- Caribbean (lowland)
- Andean

-> Historical hybridization with Northern Flint at origin of European and Corn-Belt materials not detected (due to recombination events accumulation) , possibly reinforced by "correlated frequencies" option

Historical interpretation, coll. M. Chastanet (CNRS)

Rebourg et al, 2003, Dubreuil et al. 2006, also Revilla et al., 2003



Adaptation to European temperate conditions inherited to large extent from Northern Flint origin introduced rapidly after Discovery, as such in northern cold regions, through hybridization at mid latitudes

Heterotic groups either preexist hybrid breeding (Nothern Europe) or have been shaped by it (Corn Belt)

AIL have been developed in several species, mainly

-mice: Iraqi et al., 2000, Hernandez-Valladares et al., 2004…

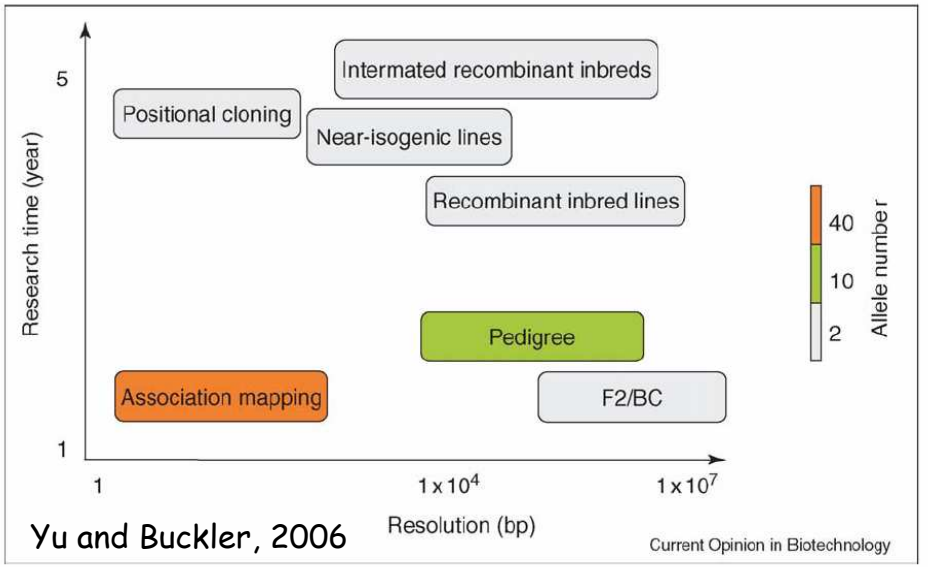
-maize: "Intermated B73 Mo17" (IBM) population, Reference population in the maize community for mapping projects,(Lee et al. 2002) but not adapted to Northern European conditions.

Aim of this study:

Evaluate interest of AIL for QTL detection using a conventional population and a AIL with similar size, both derived from the same parental lines and evaluated in the same environments.

QTL detection by linkage analysis in plants

- Historically, QTL detected using biparental populations such as F_2, BC or RIL
- Even if powerful and robust,
  - Low resolution due to low number of recombination events between linked loci (CI often >20 cM)
  - Narrow genetic variability explored
    - Detected QTL depends on the cross, not usable for MAS in other crosses



Yu and Buckler, 2006

- To go further we developed several approaches
  - Intermated lines
  - QTL detection in connected populations and application in MAS
  - (Meta-analysis)
  - Association genetics

QTL detection in intermated lines

"Advanced Intermated Line" (AIL) first proposed by Darvarsi and Soller, 1995.

Basic principle, add generations of random intercrossing when developing segregating populations to increase the number of recombination events between linked loci.

This leads to:

- ✓ improved accuracy for ordering markers on a map.
- ✓ improved accuracy of QTL detection (provided a higher marker density is used)
  -smaller confidence intervals
  -possibility to detect linked QTL previously masked by coupling or repulsion

QTL detection in intermated lines

AIL have been developed in several species, mainly

-mice: Iraqi et al., 2000, Hernandez-Valladares et al., 2004…
-maize: "Intermated B73 Mo17" (IBM) population,
Reference population in the maize community for mapping projects,(Lee et al. 2002) but not adapted to Northern European conditions.

Several fine mapping studies have been published in maize

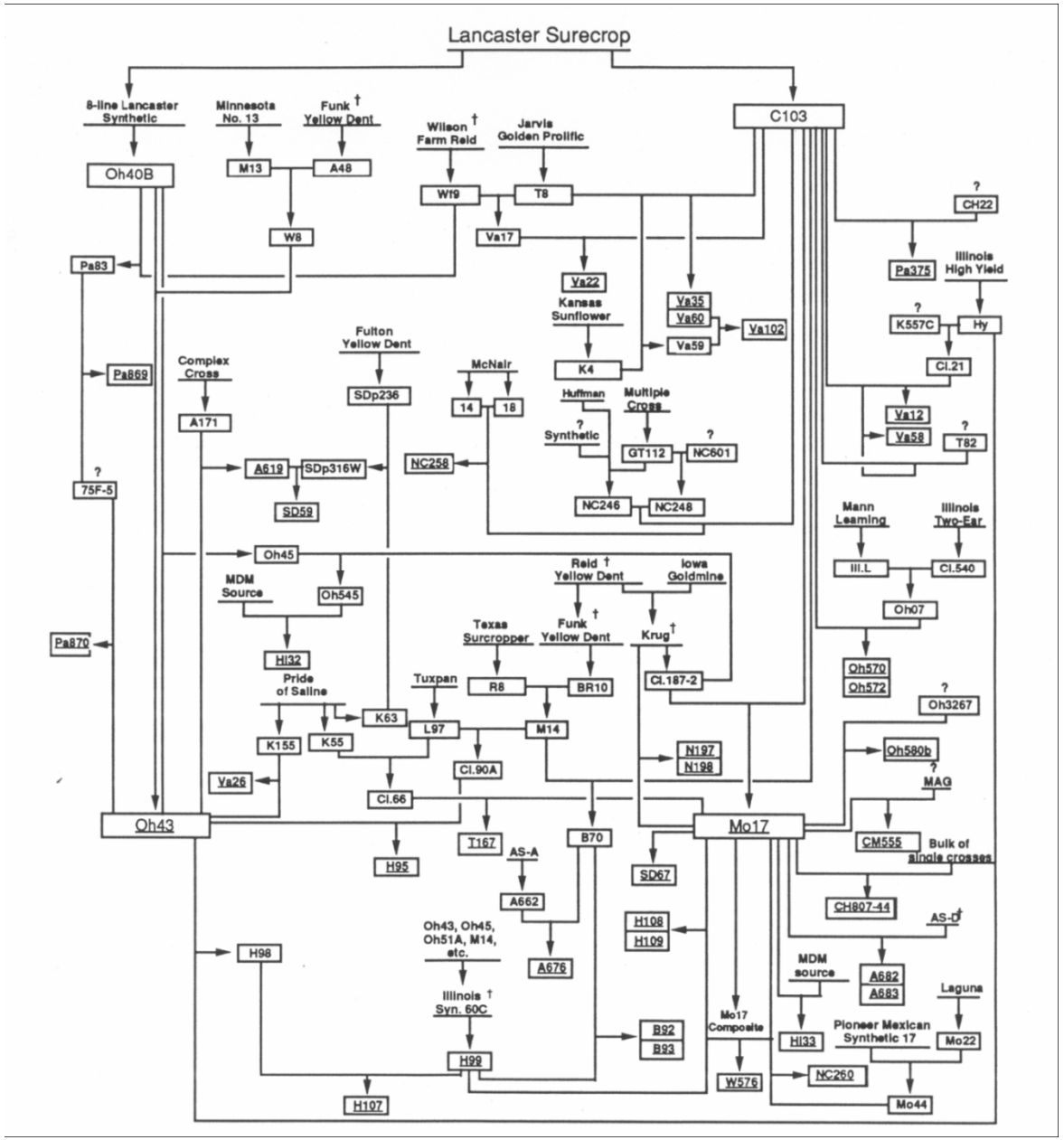
Balint-Kurti et al. 2007, 2008 compared QTL results obtained in IBM population and a conventional populations derived from the same maize inbred lines (B73 and Mo17).

In some cases, QTL CI was highly reduced in IBM (factor of 50).

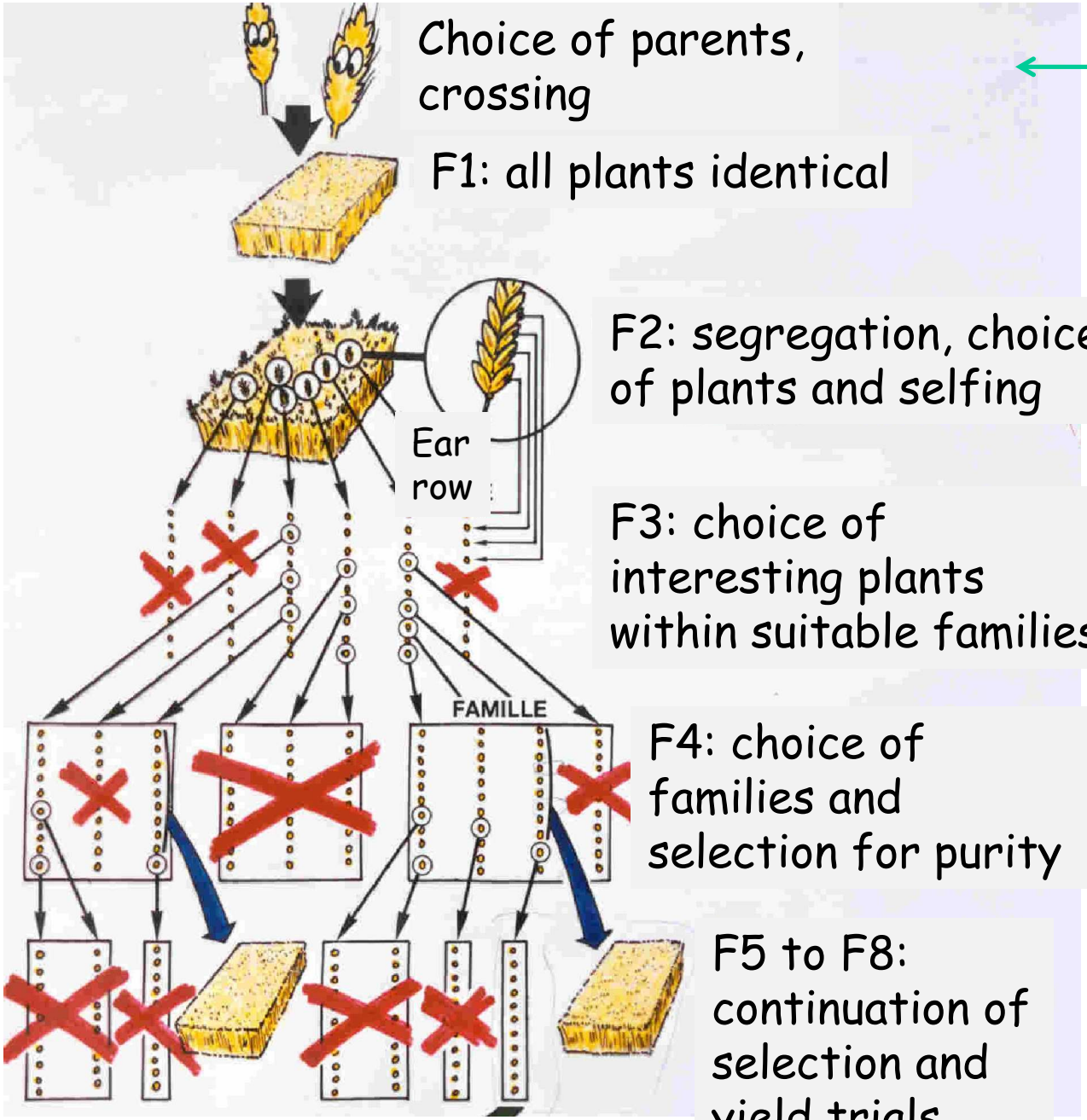
However population sizes and phenotypic evaluation were different, preventing from a possible direct comparison between QTL results

A typical program

- hybrids between inbred lines from complementary genetic groups
- reciprocal selection of these groups
- within group selection:
- within each group:
- Inbred line development followed by hybrid testing (with « tester »)



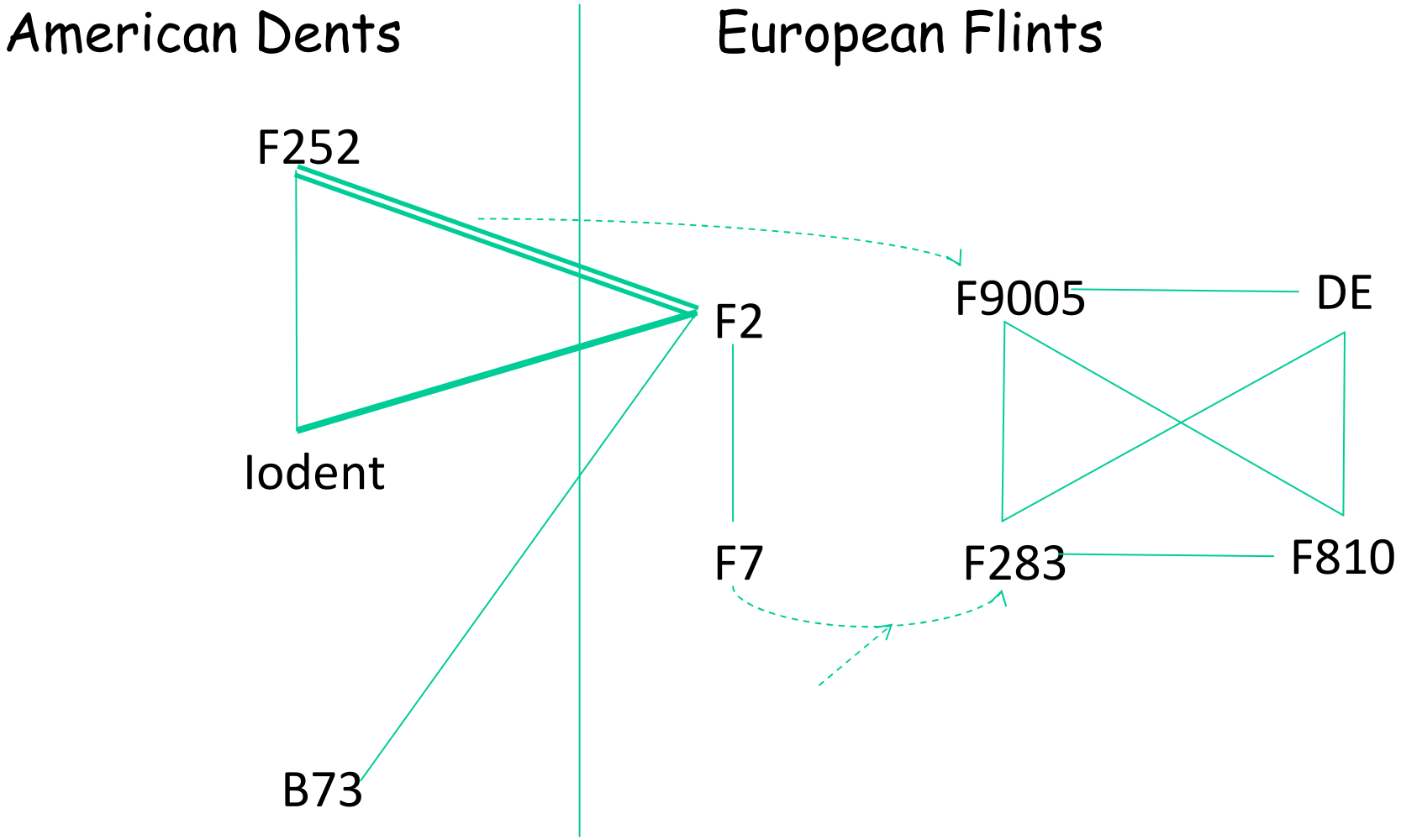
Coors and Tracy



Source : GNIS

Multilocal testing

QTL mapping populations
Early interest for multiparental designs



Research objectives

Genetic diversity characterisation

Determinism of traits of interest

Model: maize varieties adapted to Northern Europe, presently hybrids

Optimisation of breeding methods in plants,
Using emerging technologies

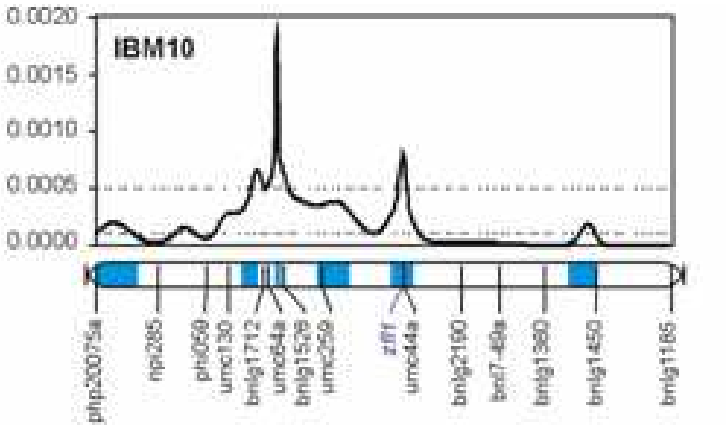
Highly unbalanced recombination

-> 170 kb interval, in the vicinity of *ZmCCT*, homologue to rice flowering time gene *HD7* (Xue et al., 2008) but excluding it (rk. Preliminary expression studies of ZmCCT between parents not conclusive)

Fine mapping of one specific region (flowering

More global approach based on advanced intercross line

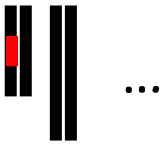
Linkage based fine mapping of major flowering time QTL on Chr. 10 (Ducrocq et al. 2009, Genetics, coop C. Giauffret et al. INRA Mons)



Main region in Blanc et al., region also repetitively detected as QTL in diverse studies (see meta-analysis, Chardon et al.. 2004 + recent results)

R^2=41% for flowering time in long days in F_5 RILs from Fv331 (highland tropical) x F286 (temperate) aiming at cold tolerance

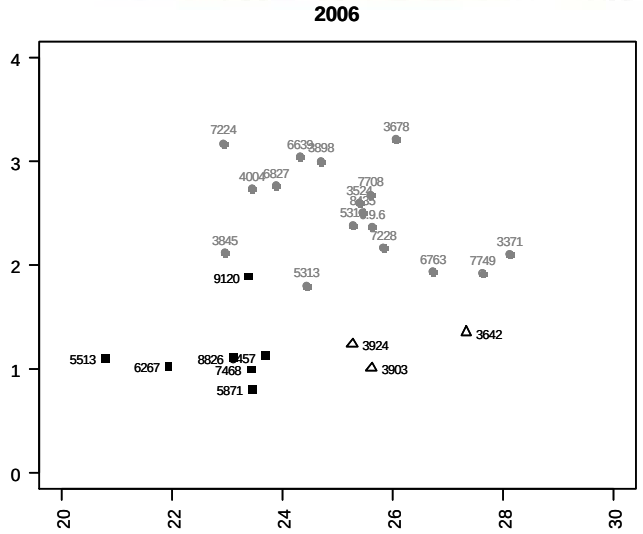
Development of near isogenic materials using residual heterozygosity in F_5 (HIF)



...

Screening of recombinants within 8000 plants

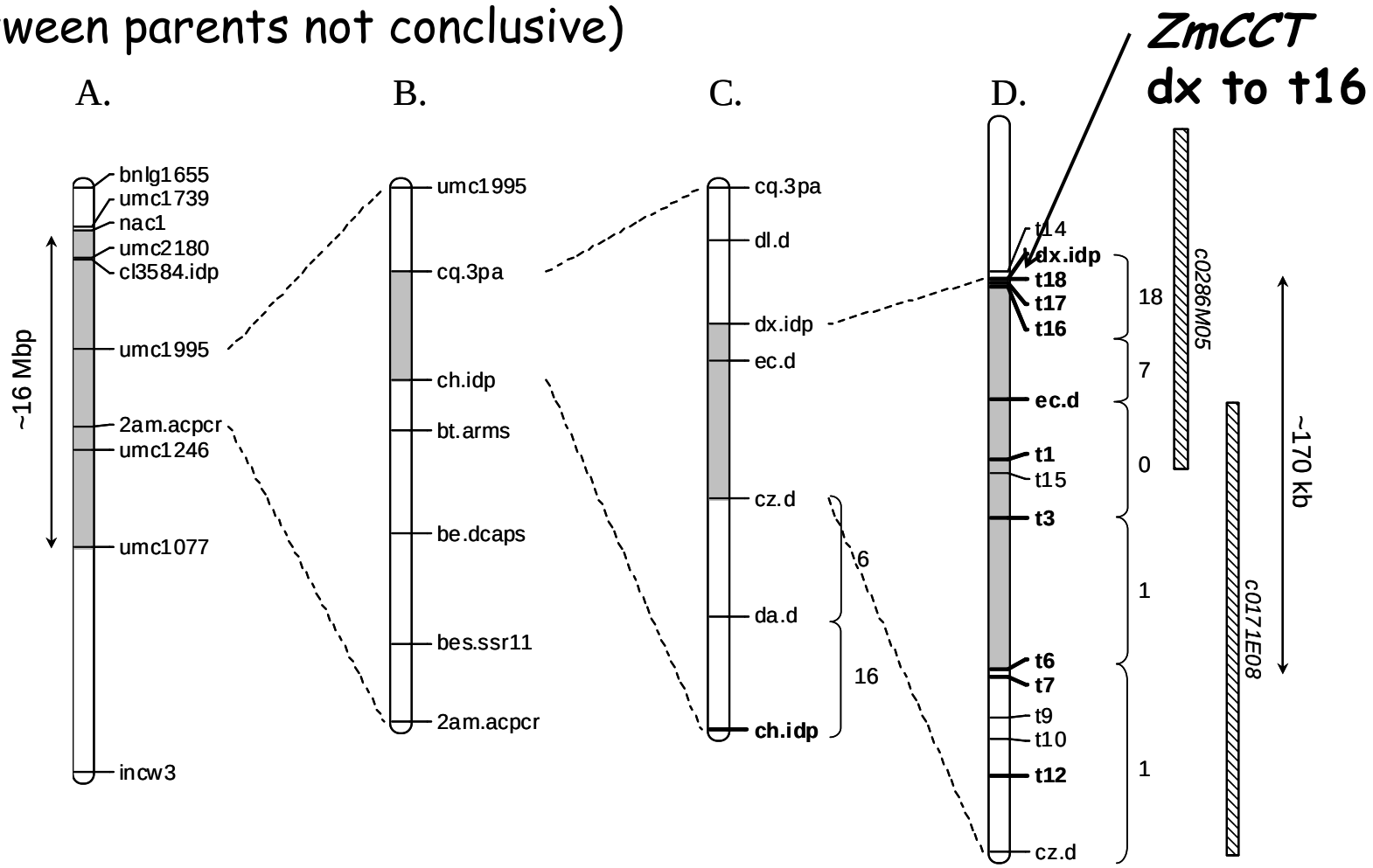
Scoring of individual plants within selfed families:
genotype at QTL determined by combining mean and variance (up to 4 environments)



Development of markers: exploitation of synteny, sequence of BAC c0286M05 (was publicly available), BAC c0171E08 (was not available, sequenced within program)

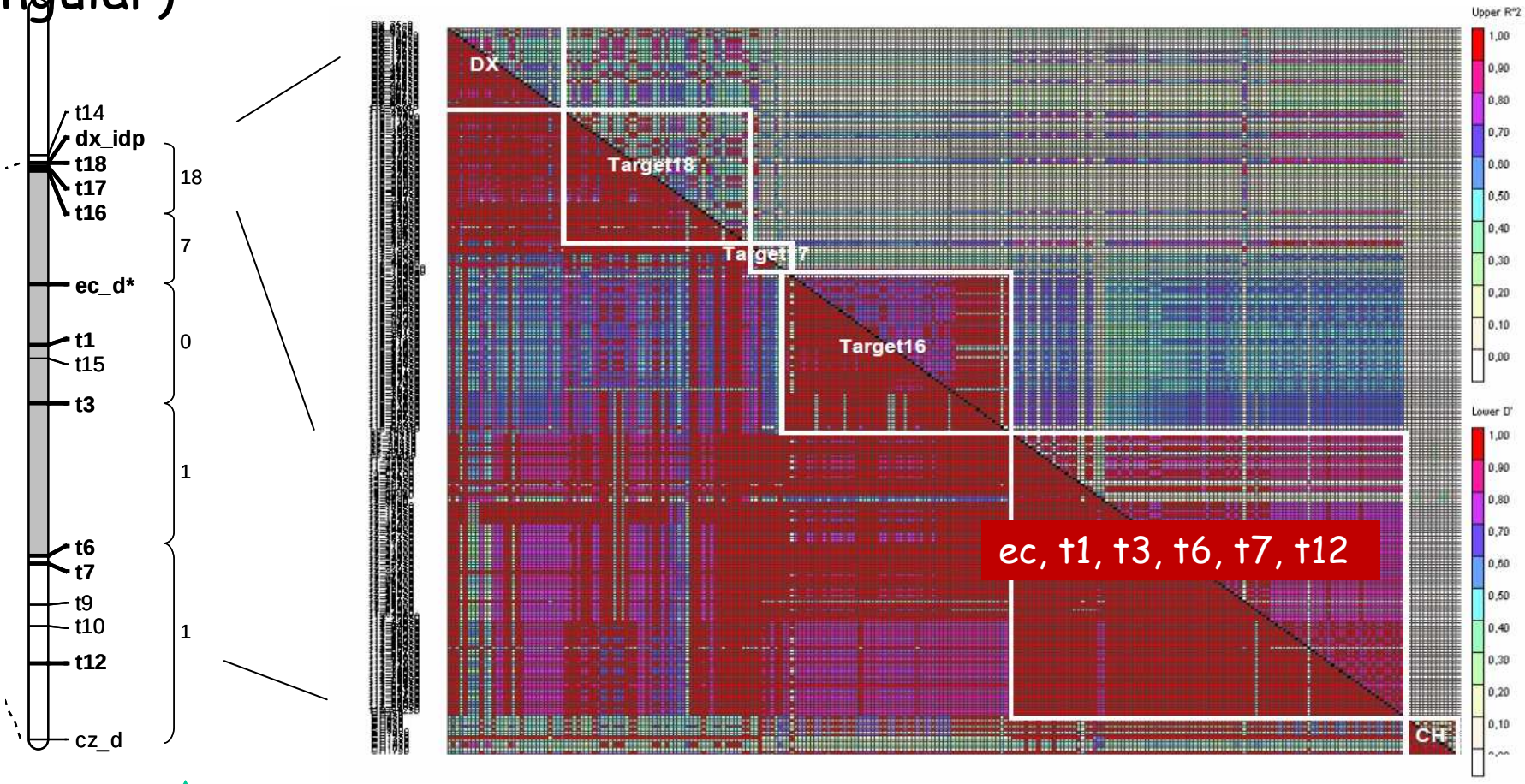
Highly unbalanced recombination

-> 170 kb interval, in the vicinity of *ZmCCT*, homologue to rice flowering time gene *HD7* (Xue et al., 2008) but excluding it (rk. Preliminary expression studies of ZmCCT between parents not conclusive)



Association Genetics (375 diverse inbred lines panel)

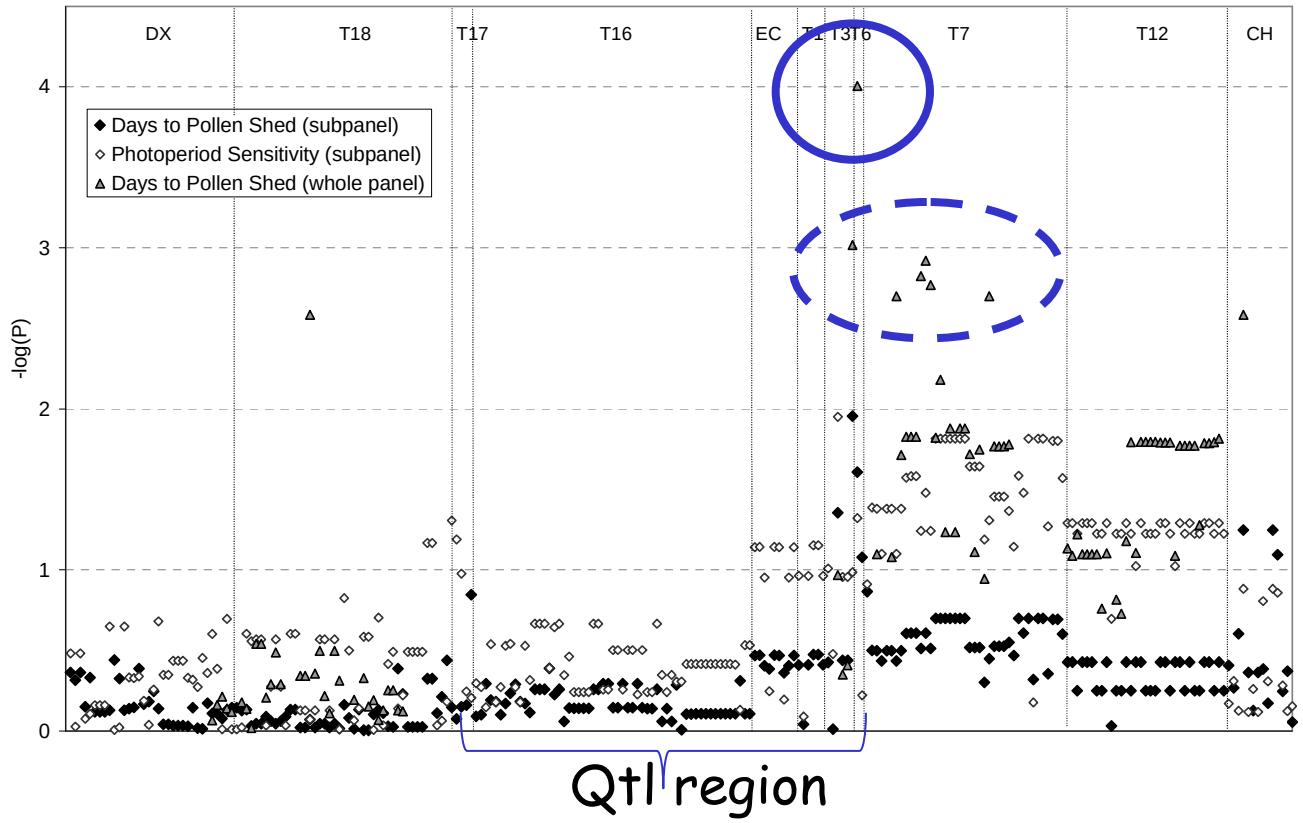
Region characterized very high LD (two haplotypes in QTL segment, one less frequent segregating in Flints only, three haplotypes considering also ZmCCT, rq. Fv331 appears as singular)



Recomb inds (/8000)

DL in panel: R^2 (upper) and D' (Lower)

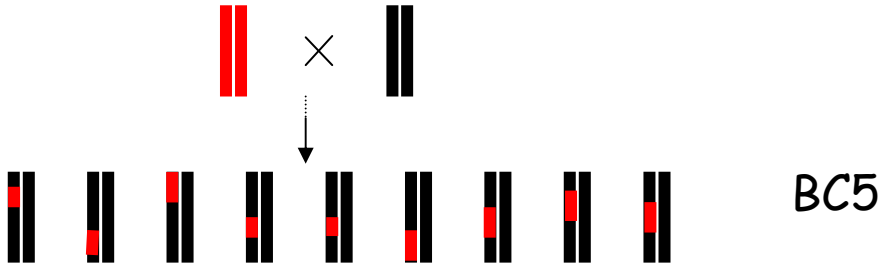
Association Genetics (375 inbred panel) tests reach their maximum for T6 (three alleles polymorphism) and close SNPs



Suggests three haplotypes with divergent effects, including a late flint allele (consistent Fv283 effect in Blanc et al. 2006), a most common intermediate allele (three other flint lines) and an early dent allele

« Reverse fine mapping » to check associations

Ongoing work by S. Bouchet



Effect of late flint allele supported by introgressions of Fv283 into Fv2

-> delays male flowering time (+4.2 days, $p<10^{-4}$) and more female flowering time (+8.8 days)

Fine mapping (1700 plants screened in 2009) confirms that the QTL is located in the same region

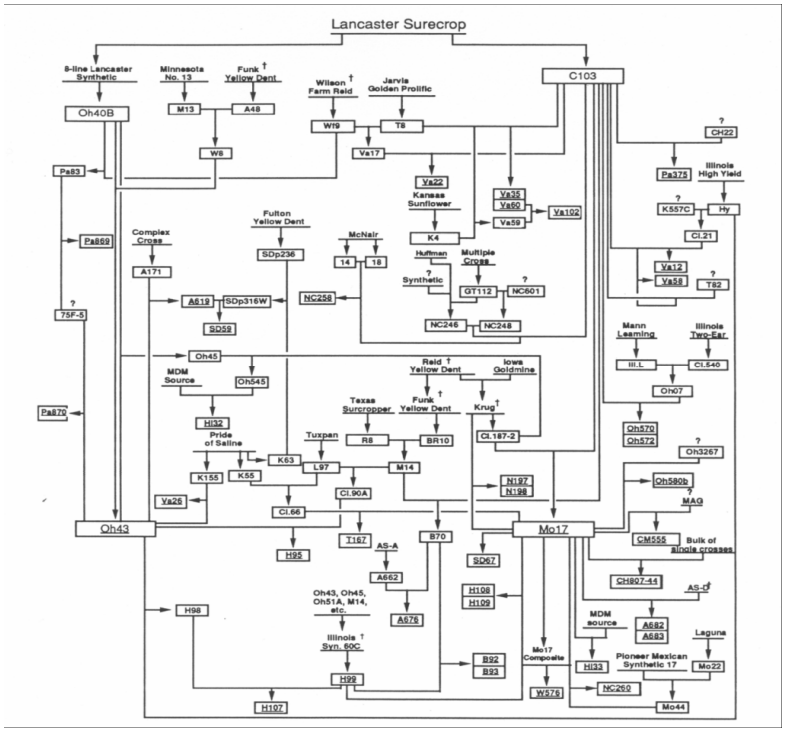
1.QTL mapping in mutiparentzl connected populations and application for MAS

2. fine linkage and LD mapping, an exemple

3.Complexity … AIL

Elements on conventional breeding of annual plants

- Many independant breeding programs (companies)
- Development of inbred lines: varieties or hybrid parents
- Within each program: many crosses (P1xP2) leading to populations of limited size
- Phenotypic evaluation during or after fixation of new inbred lines
- Heavy progeny testing
- Data analysis generally not managed globally (no BLUP), integration of results mostly through breeder's expertise
- De facto recurrent selection through intercrossing of superior inbred lines or elite families



A typical pedigree:
« Lancaster » lines (Tracy and Coors)

Developed in several species

-mice: Iraqi et al., 2000, Hernandez-Valladares et al., 2004…

-maize: "Intermated B73 Mo17" (IBM) population,

Reference population in the maize community for mapping projects,(Lee et al. 2002) but not adapted to Northern European conditions.