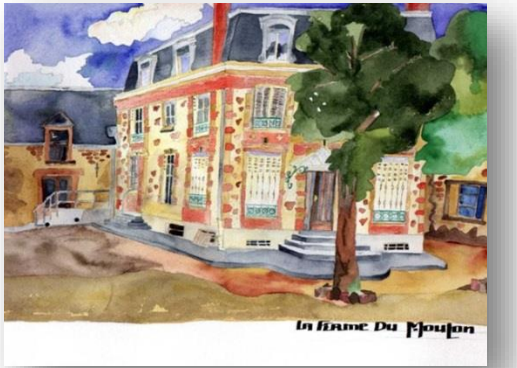
15 YEARS OF GWAS IN EUROPEAN CORN, WHAT HAVE WE LEARNED?

KWS meeting

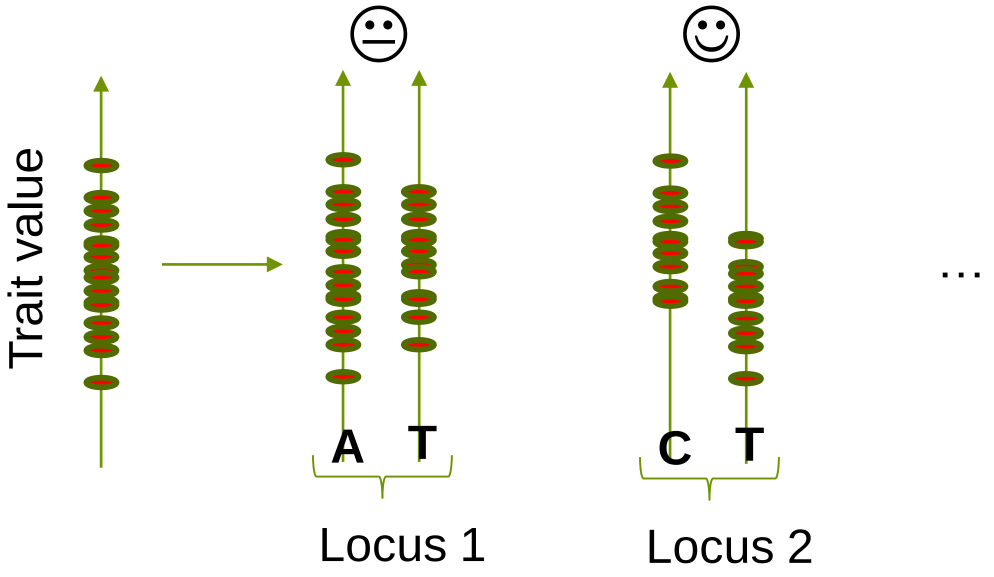
February 14th, 2018



Alain Charcosset, Stéphane Nicolas, J. Fievet, Camille Clipet, Maud Tenaillon, Brigitte Gouesnard, François Tardieu, Claude Welcker, Cyril Bauland, Carine Palaffre, Laurence Moreau

What is Association Genetics?

Quantitative Trait Loci (QTL) mapping in diverse populations where relatedness is limited or does not follow a regular structure (use of existing populations)



Potential applications in breeding: identification of alleles and diagnostic markers

Association Genetics mostly started in plants with the seminal paper of Thornsberry et al. 2001, Buckler's group

letter

© 2001 Nature Publishing Group http://genetics.nature.com

Dwarf8 polymorphisms associate with variation in flowering time

Jeffry M. Thornsberry[1], Major M. Goodman[2], John Doebley[3], Stephen Kresovich[4], Dahlia Nielsen[5], & Edward S. Buckler IV[1,6]

Rk., there had been some preliminary attempts

Euphytica **84:** 145–154, 1995. 145
© 1995 *Kluwer Academic Publishers. Printed in the Netherlands.*

Relationship between genetic markers and morphological traits in a maize inbred lines collection

A. Bar-Hen[1], A. Charcosset[2], M. Bourgoin[3] & J. Guiard[1]
[1] *G.E.V.E.S. La Minière 78280 Guyancourt cedex, France;* [2] *I.N.R.A. Station de Génétique Végétale, Ferme du Moulon, 91190 Gif sur Yvette, France;* [3] *G.E.V.E.S. Le Magneraud, B.P. 52, 17700 Surgères, France*

Received 20 December 1993; accepted 28 April 1995

Why did we become comfortable with the idea of conducting such analyses?

- ✓ New models of population structure yielding covariates to avoid confounding effects (STRUCTURE, Pritchard, 2001)
- ✓ Extended later to relatedness using mixed models (Yu et al. 2006)

First analyses conducted with Candidate Genes, then genome wide with advances in genotyping technologies
-> Genome Wide Association (GWAS)

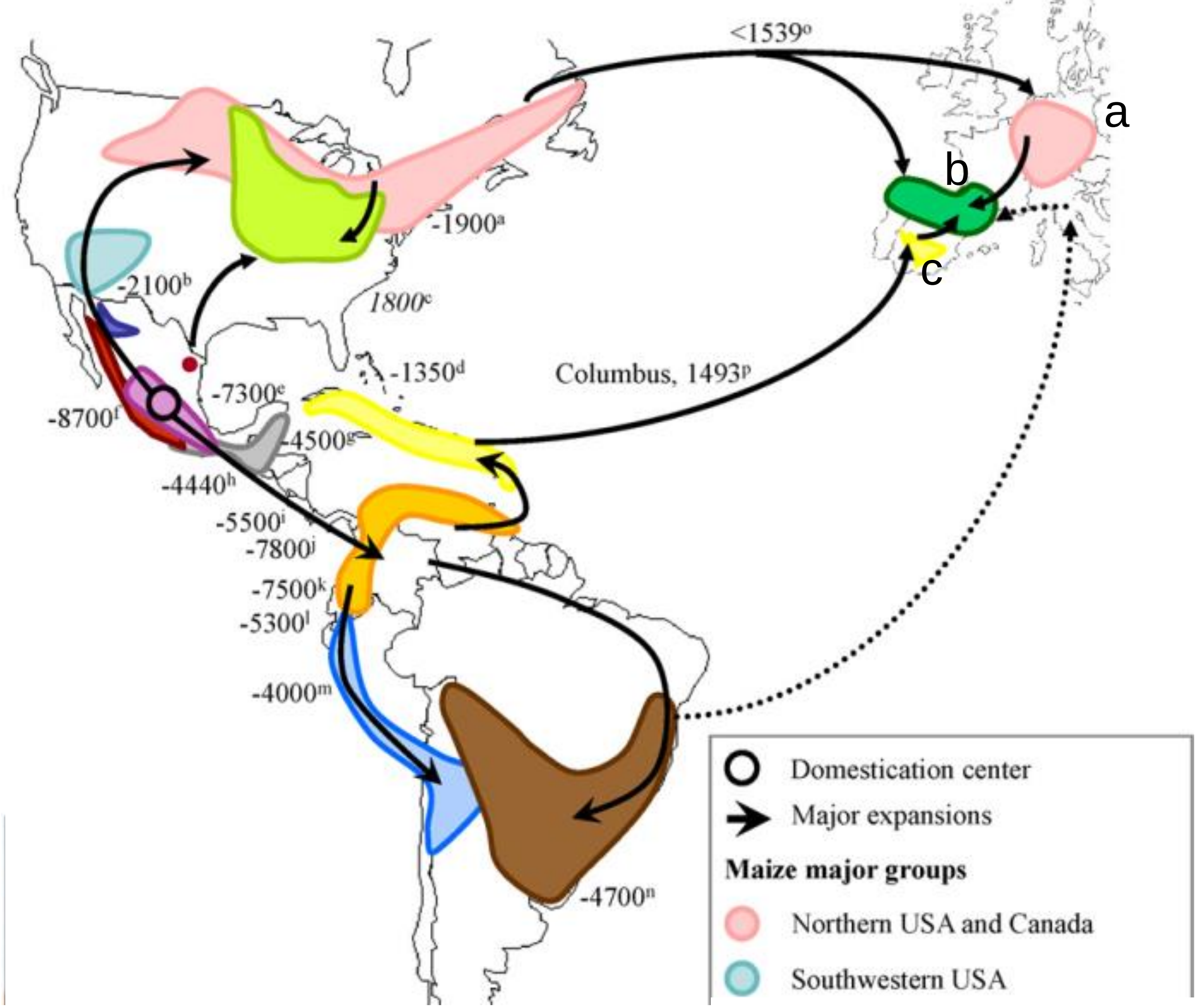
Talk outline

1. Flowering time and global adaptation in a diverse maize panel evaluated per se
2. Environmental adaptation in Flint and Dent hybrid panels, identification of adaptive regions / alleles
3. Factors affecting power of associations
4. New panels in progress

Talk outline

1. Flowering time and global adaptation in a diverse maize panel evaluated per se
2. Environmental adaptation in Flint and Dent hybrid panels, identification of adaptive regions / alleles
3. Factors affecting power of associations
4. New panels in progress

Marker and history based scheme for Maize migration / introduction into Europe (West European perspective)



Tenaillon et Charcosset, 2011

a

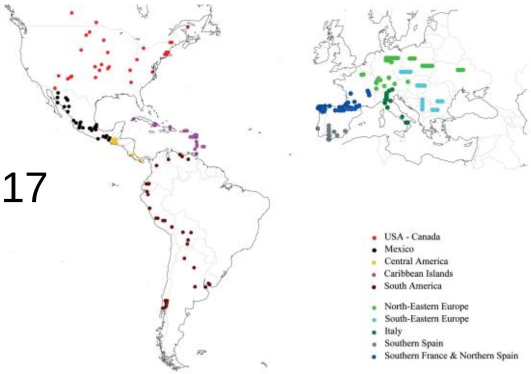
b

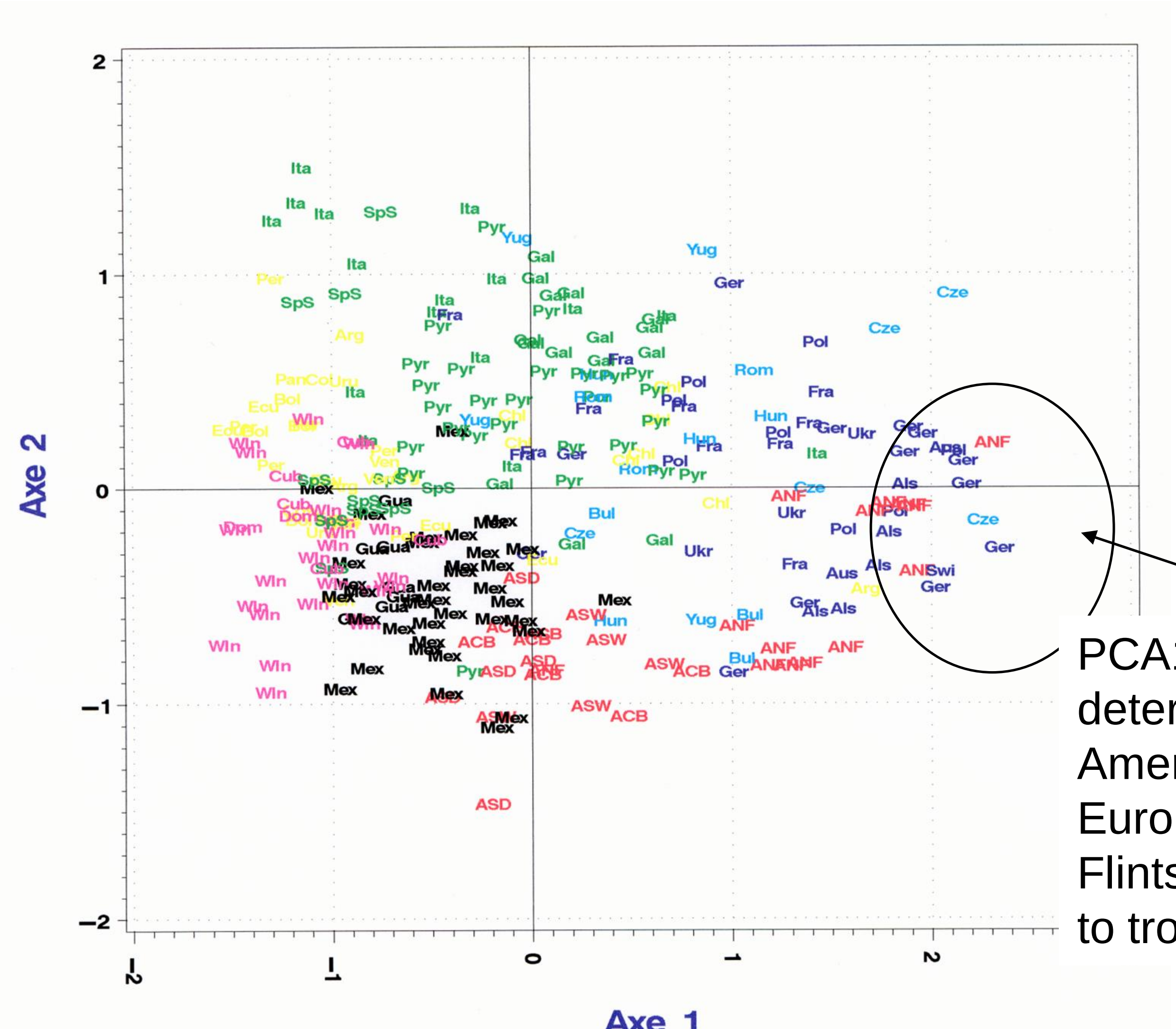
c

Traditional European population varieties (OPVs) essentially flint (vitreous endosperm)

Organization of diversity using molecular markers (PPS, diversité corné, Resgen88, diversité Zea)

here 24 SSR, coop. CIMMYT, Dubreuil et al., 2006, see also Dubreuil et al., 1998, Rebourg et al., 2003, Brandenbourg et al., 2017





America:

- ⬛ Mexico - Guatemala
- 🟥 USA - Canada - Chili
- 🟪 Caribbean
- 🟨 Southern America

Europe:

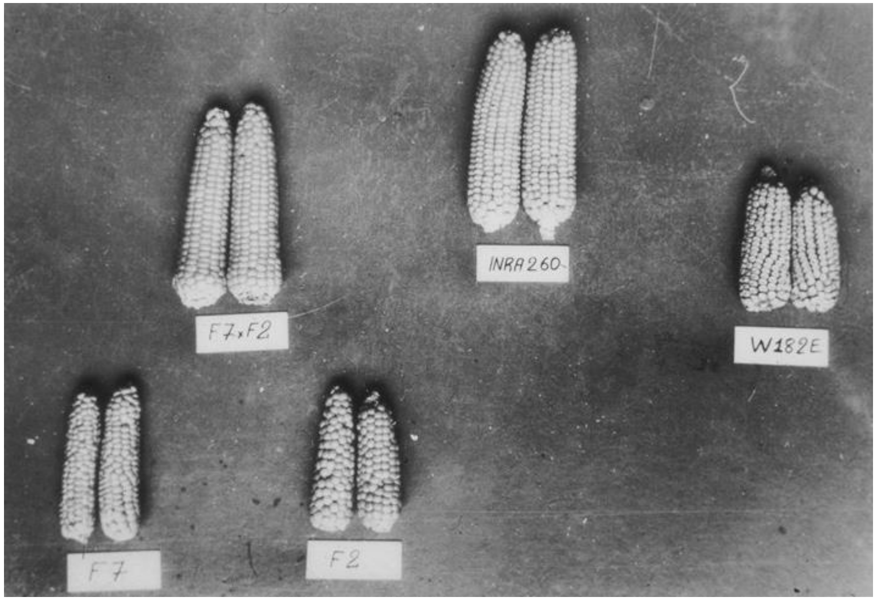
- 🟩 Spain – Pyr - Italy
- ▨ France– North-Eastern
- 🟦 Balkans

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

Development of modern breeding with maize hybrids in Northern Europe

Ex. of France: Flint x Dent Hybrids developed by INRA after WW2 (late 1950s) then private companies

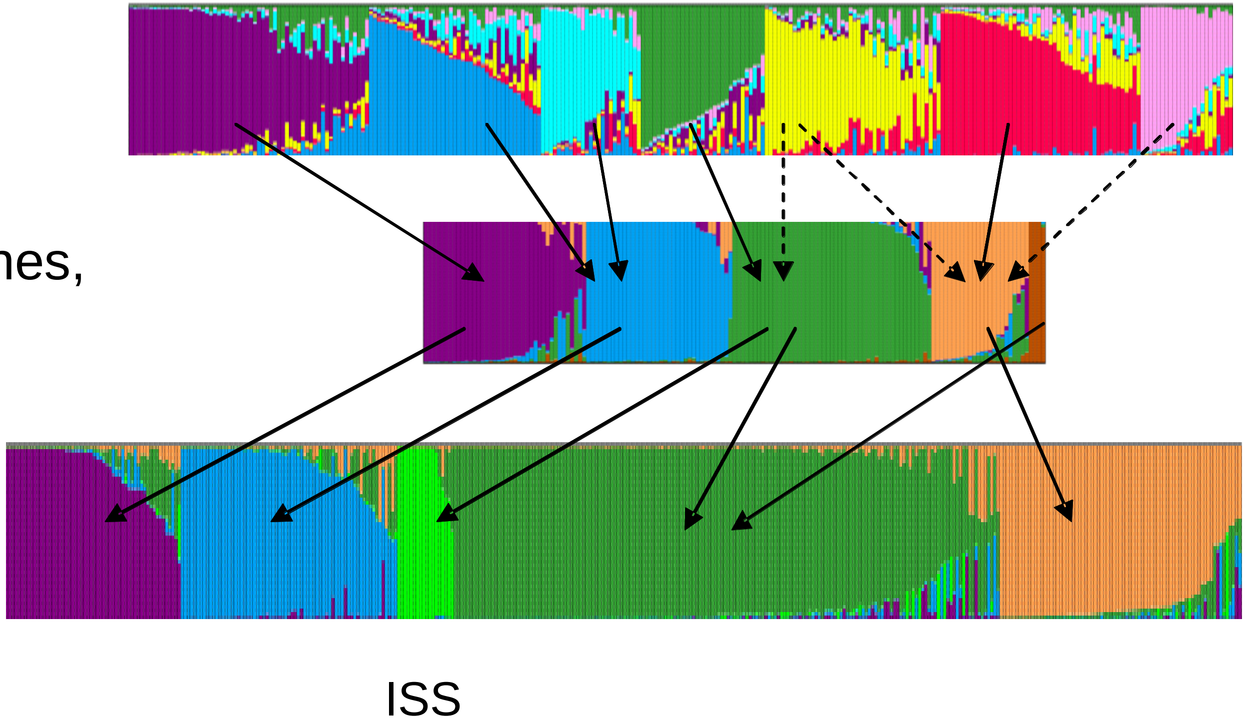
Original inbred lines developed from European Flint traditional population varieties
-> brought adaptation features (cold spring, wet autumns) complementary to productivity of Dents from Northern America



Association genetics with European materials

(Camus-Kulandaivelu et al., 2006, Genetics)

- 🟪 Northern Flint (Eur.+Am.)
- 🟦 Pyrenees-Galicia Flint (Eur.)
- 🟦 Italian Flint (Eur.)
- 🟩 Corn Belt Dent
- 🟨 Mexico (highland)
- 🟥 Caribbean (lowland)
- ◻ Andean



275 populations OPVs (landraces)

153 first cycle inbred lines, from selfing in OPVs

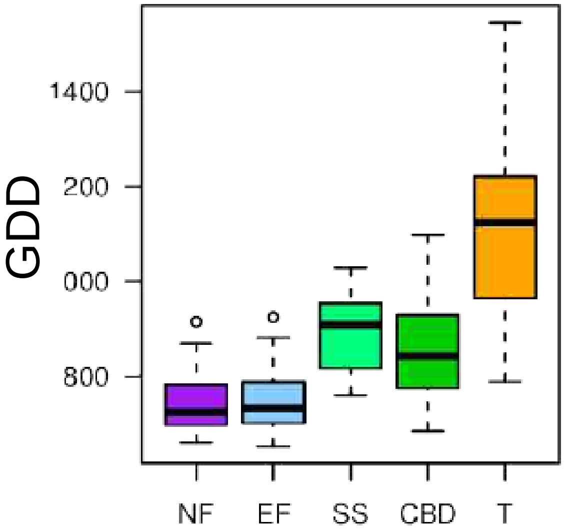
+222 inbred lines from more advanced generations (pedigree breeding, rec. selection)

-> First "Generic" panel of 375 diverse lines tracing to European and American origins

Evaluation for phenology and plant architecture « per se » in 9 environments (2002-2005)

Including one at KWS
(Bouchet et al. 2017, Heredity)

Variation within and among groups for flowering time (GDD)

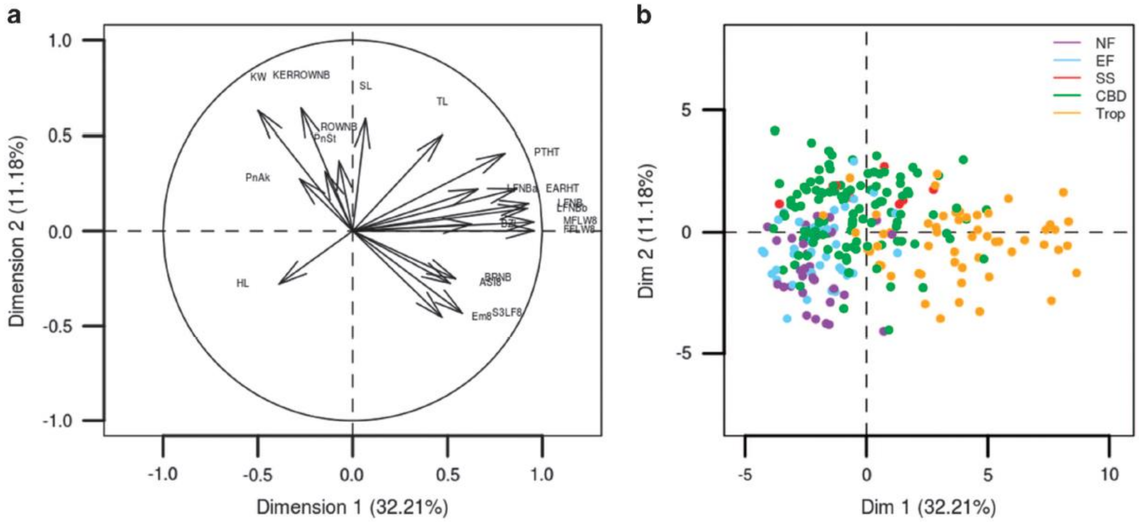


Multitrait:

Axis 1: flowering time driven

Axis 2: yield component driven (kernel number)

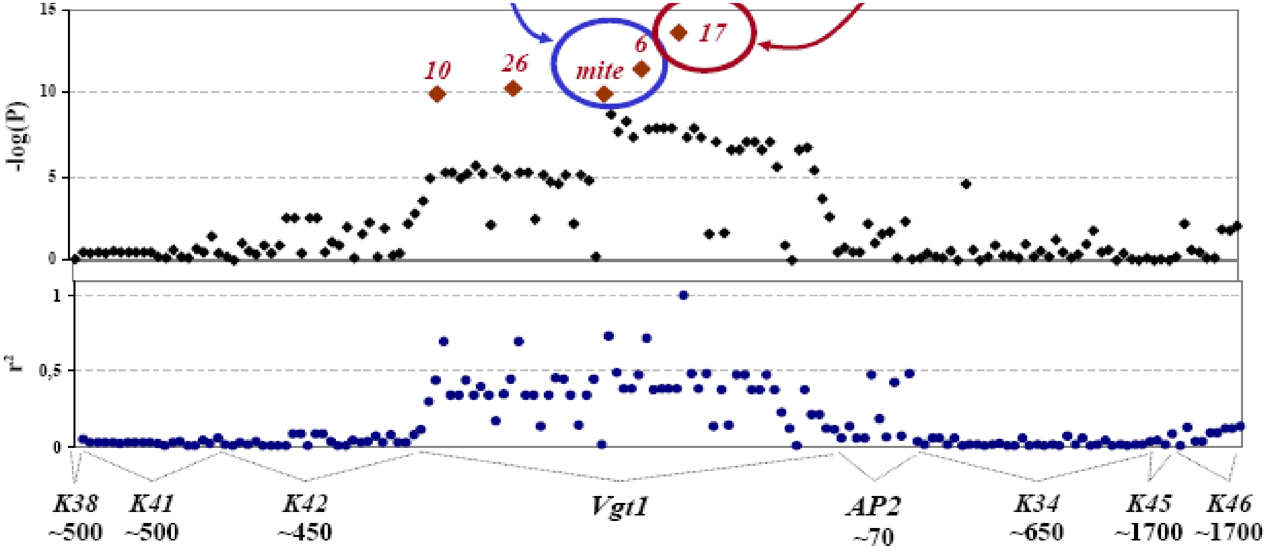
Axes3+: tillering, husk leaves, other



Association genetics in specific regions

✓ *D8* region (Camus-Kulandaivelu et al. 2006, Genetics)

✓ Confirmation of major role of *Vgt1* (cloned by Salvi et al., 2005): very strong associations with FT (> 4% of variance) (Ducrocq et al., 2008)

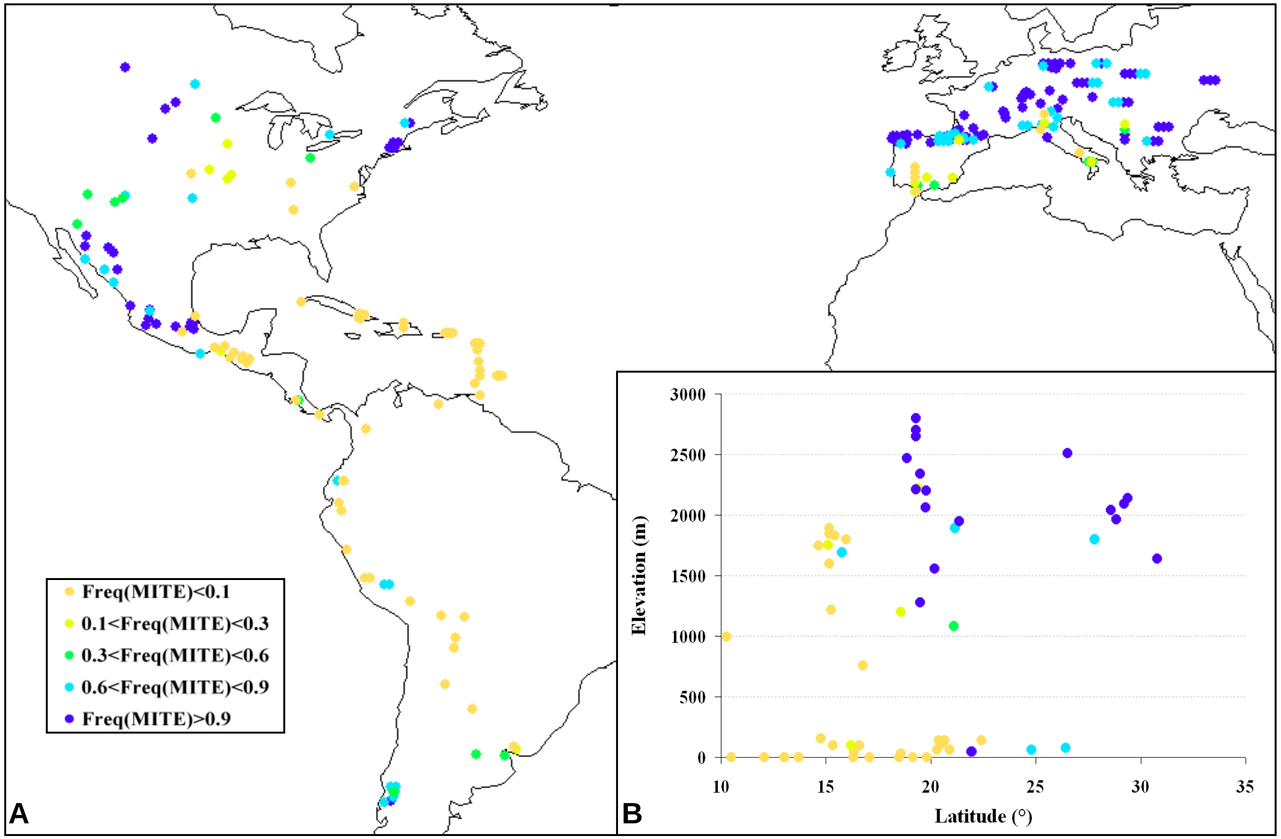


LD with SNP17

Distance to *Vgt1* (kb)

Vgt1: major role in worldwide climatic adaptation

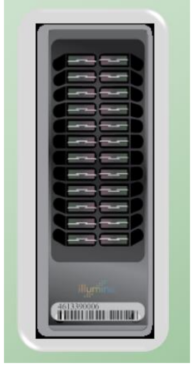
Mite insertion (early allele) frequency in populations



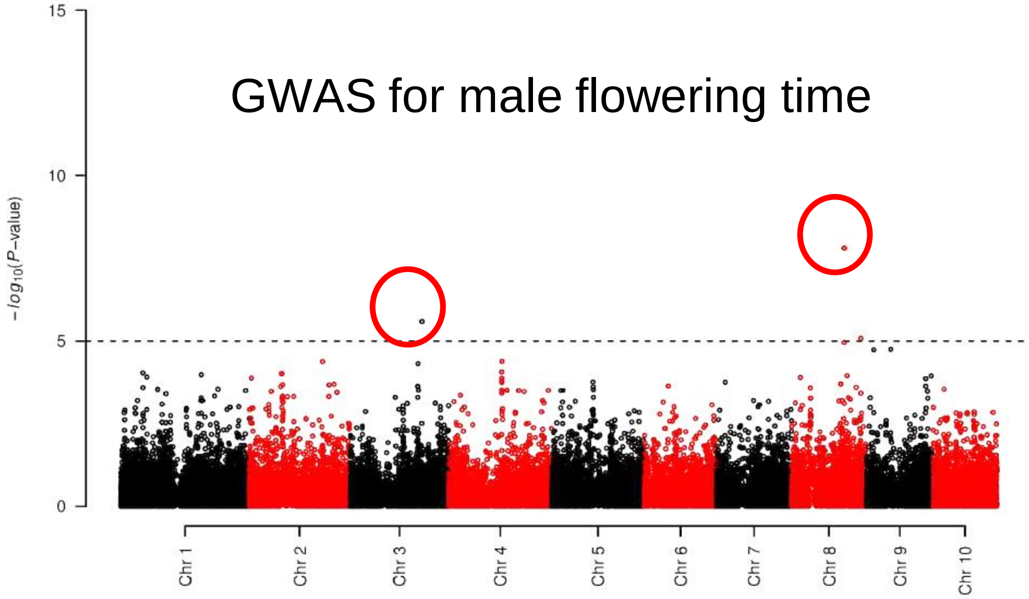
Adaptation to cool temperate climates but also differenciation of tropical « highland » vs. « lowland »

Towards genomewide studies with 50k SNP array (Ganal et al., 2011 Plos One)

Bouchet et al. 2013, Plos One; 2017 Heredity)



- Highlights *Vgt3* region on Chr 3
- Strong association for *Zcn8* on chromosome 8, most likely explains *Vgt2*



GWAS for male flowering time

But LD analysis calls for densification in most regions (improved now with 600k and GBS)

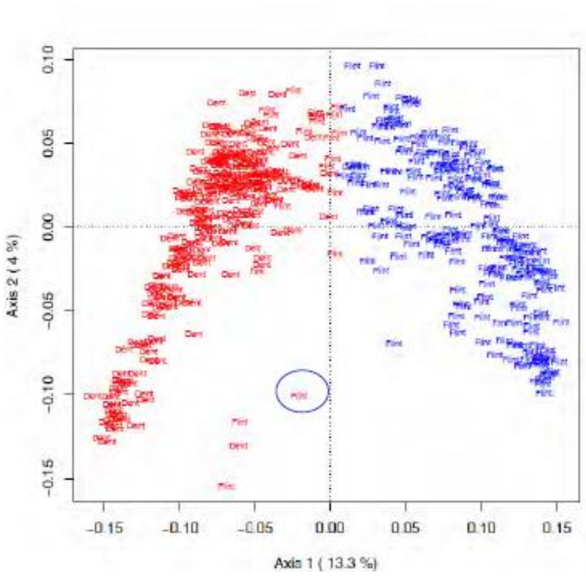
Talk outline

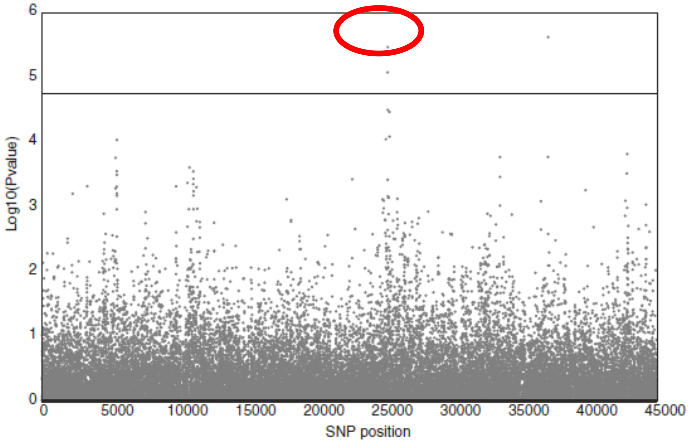
1. Flowering time and global adaptation in a diverse maize panel evaluated per se
2. Environmental adaptation in Flint and Dent hybrid panels, identification of adaptive regions / alleles
3. Factors affecting power of associations
4. New panels in progress

Towards Dent and Flint specific panels within CornFed project

Previous panel adapted to traits of interest for lines per se, but hybrid yield evaluation difficult (heterogeneous phenology and diversity of genetic origins)
-> cooperation with Univ Hohenheim, CSIC and CRAG to define Flint and Dent panels of approx. 300 lines each, Evaluation: per se and hybrids with a « tester » from the opposite group (Rincent et al. 2014, Theor Appl Genet)

Ex. Early vigor and photosynthetic activity in cold conditions in Flints (Revilla et al., 2016, BMC Plant Biology)





Further focus on Dents within Drops-Amaizing projects (F. Tardieu)

Millet et al. 20016, Plant Physiol

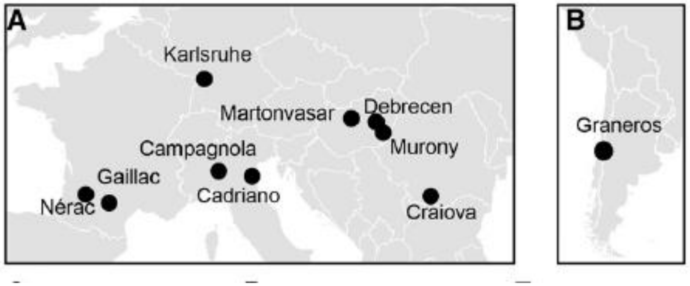
Further evolution of CornFed dent panel to narrow down flowering time window
-> 250 lines, crossed to flint tester

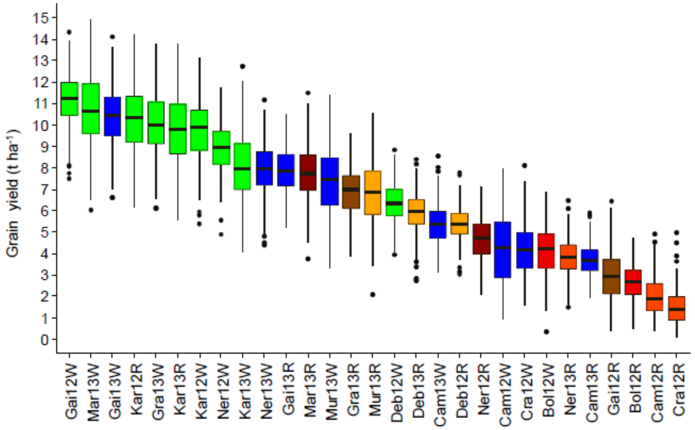
29 environmental conditions (site x year x management) with accurate « envirotyping » to infer water status, T, …

Genotyping (S. Negro et al., in prep):

- 50K array,
- 600 K array
- GBS

Statistical analysis with WUR (F. van Eeuwijk, W. Kruijer)



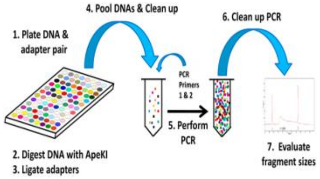


Grain yield variation within and among conditions

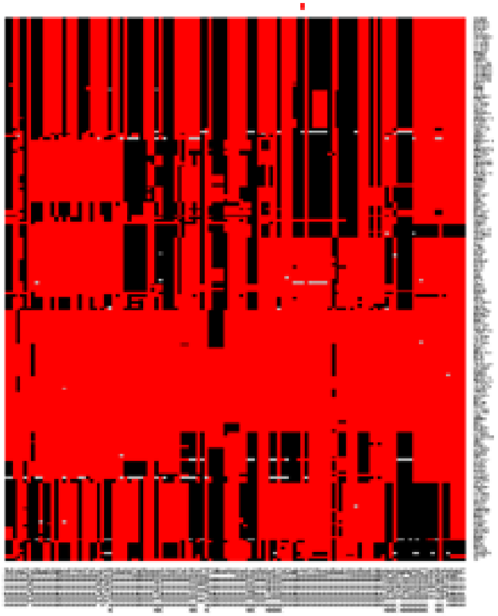
High density SNP genotyping

50K Illumina Infinium HD
(Ganal et al., 2011)
~ 50 € / sample
~ 50 000 SNP

600K Affymetrix Axiom
(Unterseer, 2014)
~ 200€ / sample
~ 600 000 SNPs



Genotyping By Sequencin
(Glaubitz et al., 2014)
~ 33€ / sample
~ 600 000 SNPs
(Imputation needed)



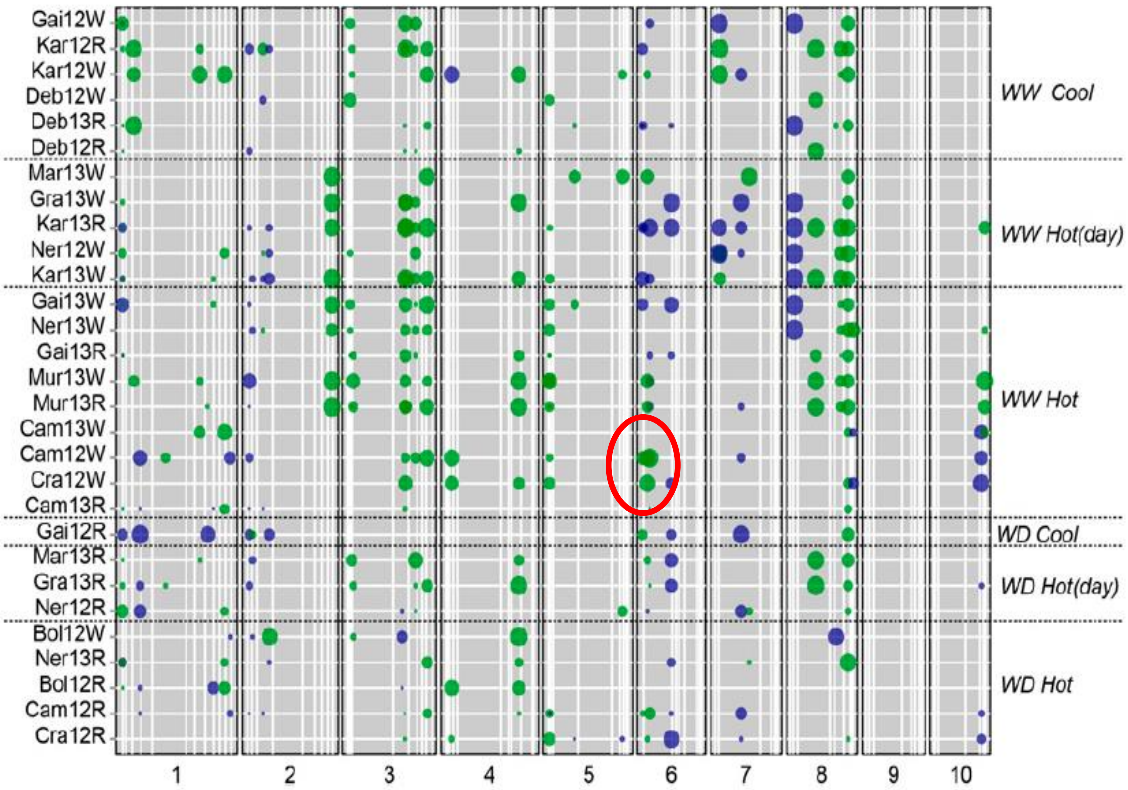
Clustering = average Kinship= IBS Admixture > 0.7

Ex. 211 markers (columns)
in 100 kb on Chr1
rows: 250 lines dent
DROPS-Amaizing panel

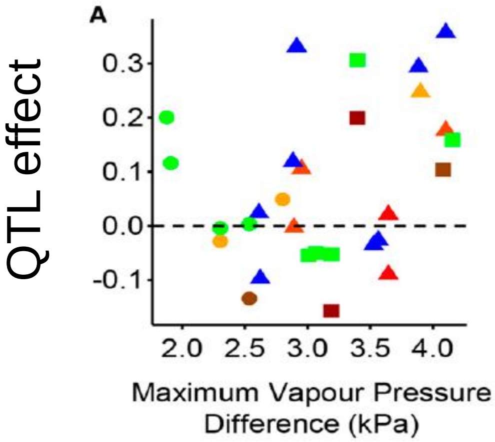
More than 1 M markers in total (heavy work for managing data)

Summary of associations for grain yield

Reveals several regions with important effects, Including again *Vgt3* (although flowering time variation was narrow)



One QTL with strong effects in hot well watered conditions on Chr 6, in a region with strong structural variation

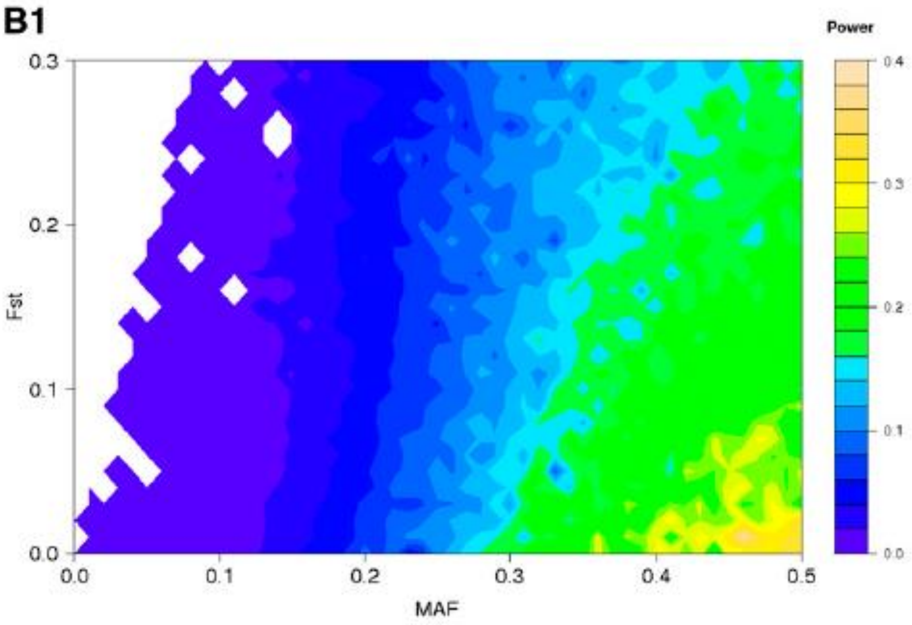


Talk outline

1. Flowering time and global adaptation in a diverse maize panel evaluated per se
2. Environmental adaptation in Flint and Dent hybrid panels, identification of adaptive regions / alleles
3. Factors affecting power of associations
4. New panels in progress

Power of association tests and possible statistical improvement (Rincent et al., 2014, Genetics)

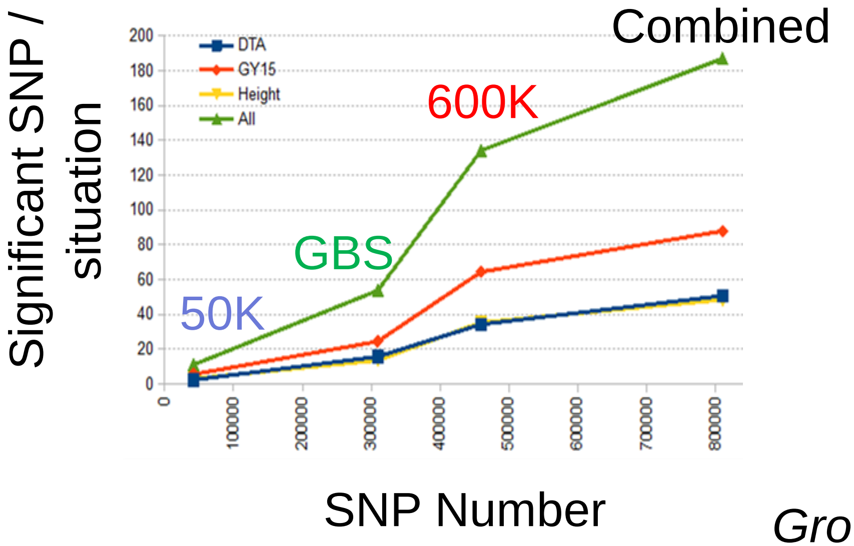
- ✓ For a given QTL effect power decreases with unbalance between allelic frequencies and/or strong correlation to population structure



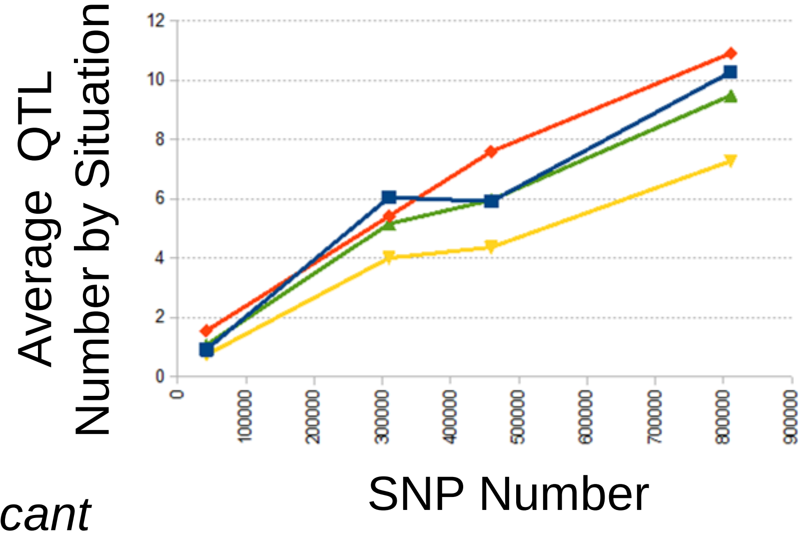
- ✓ Power can be increased by approx. 40% by discarding the chromosome being scanned from kinship estimation. Avoids QTL the effect to be included in the model both as fixed and random effect.

Effect of SNP density and genotyping technologies on QTL detection (Negro et al., in prep)

~ 62 significant SNPs / situation



~ 9.5 QTLs / situations



Group significant SNPs according to LD

1. Average number of QTL detected by situation (year x treatment x location) increased linearly with SNP density whatever the trait
2. GBS and 600K were complementary to detect QTL
3. QTL detection by GBS (~90$/QTL) was cheaper than 600K (430$/QTL) and 50K (462$/QTL)

Talk outline

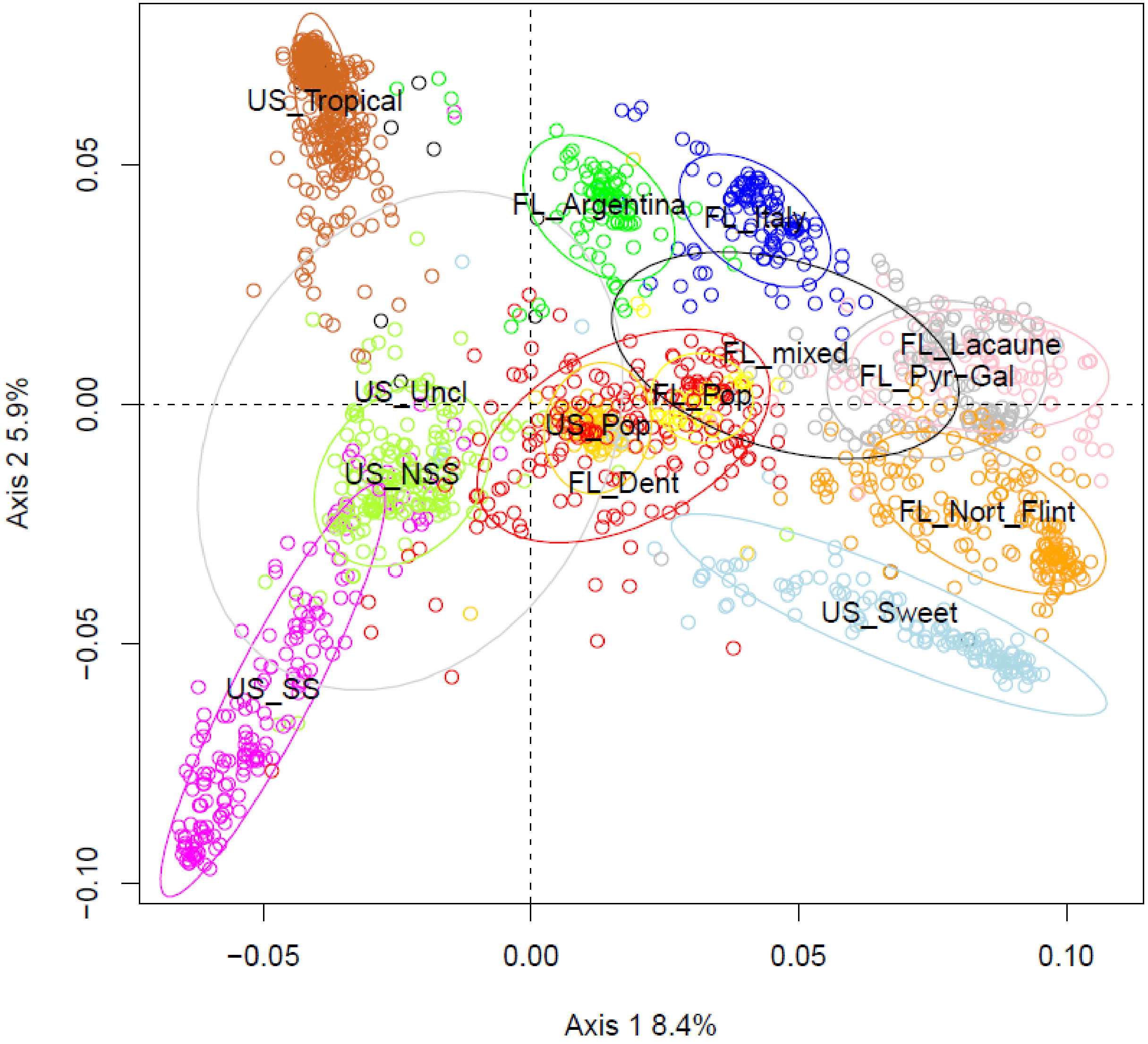
1. Flowering time and global adaptation in a diverse maize panel evaluated per se
2. Environmental adaptation in Flint and Dent hybrid panels, identification of adaptive regions / alleles
3. Factors affecting power of associations
4. New panels in progress

aMaizing

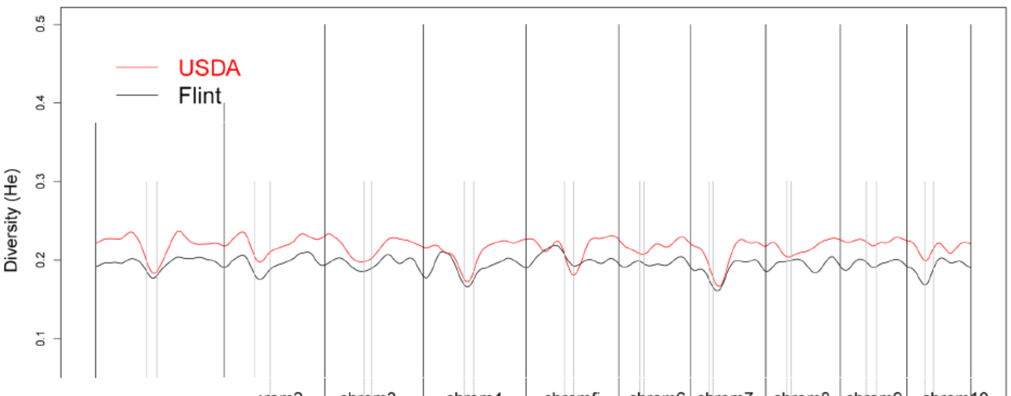
Expanding flint panels: Genotyping by sequencing, coop. Buckler group (Gouesnard et al., 2017)

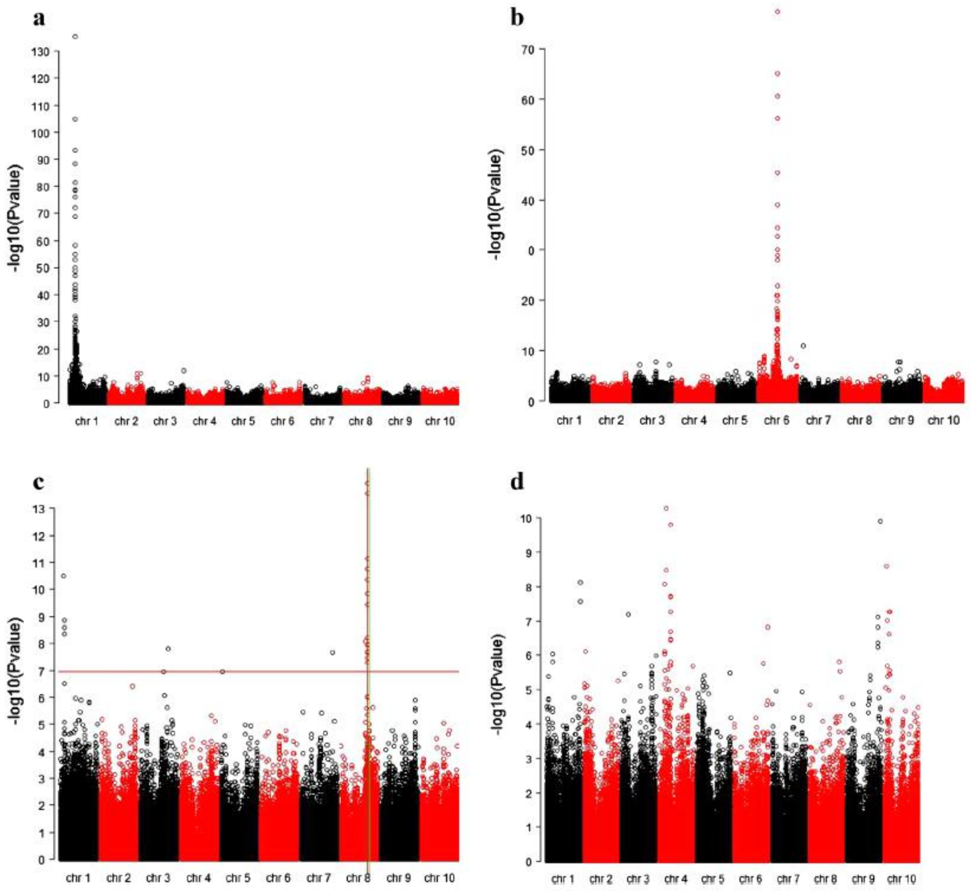
1200 flint inbred lines from European collections / breeding programs (FL_) analyzed at 514.000 SNPs and compared with USDA collection (US_, Romai et al., 2013)

-> confirms the originality of European Flint inbreds, with several subgroups (Italy)



Diversity patterns along genome in both collections





GWAS (Flint collection)

a cob color,

b color of the *top* of the kernel,

c for the growing degree days to anthesis. *Vertical lines* for *Vgt1*(*green*), and *Vgt2* (*red*). *Horizontal lines* for Bonferroni test limit;

d number of rows per ear

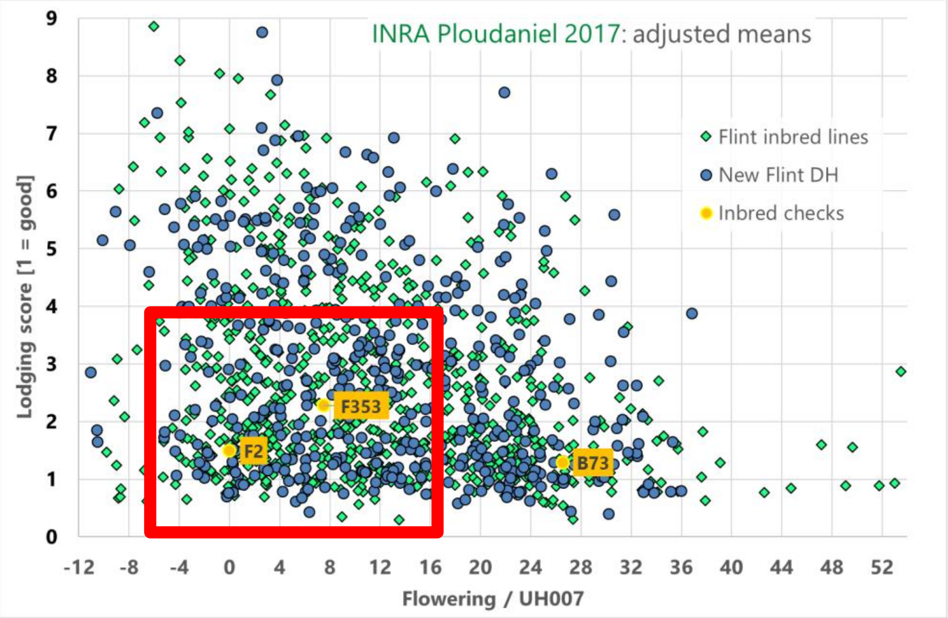
Expanding flint panels: development of “structure breaking” inbred lines

✓ 600 representative inbred lines selected to optimize representation of flint diversity

✓ 300 F1 crosses selected to maximize recombination between flint groups -> 514 DH

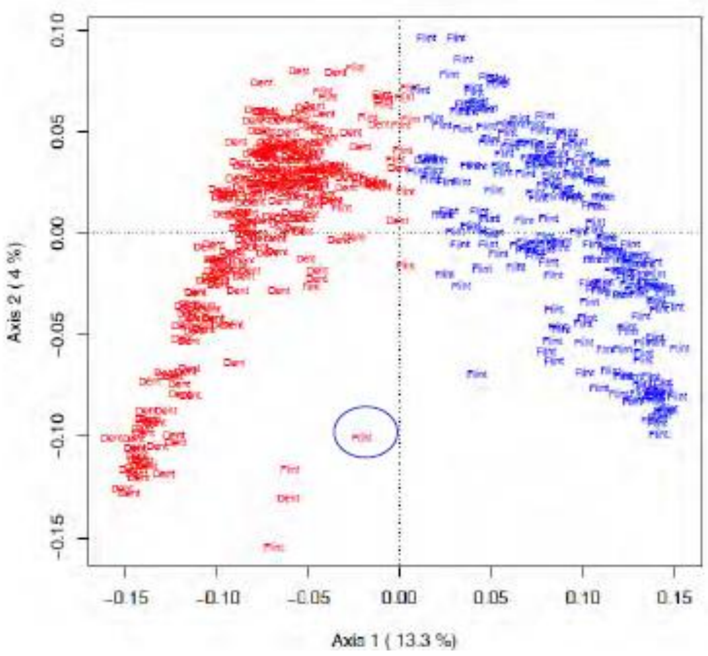
-> all evaluated per se (flowering, lodging, diseases)
-> GWAS to be conducted

-> selection of lines to be included in the new hybrid flint panel (trials 2018, 2019)

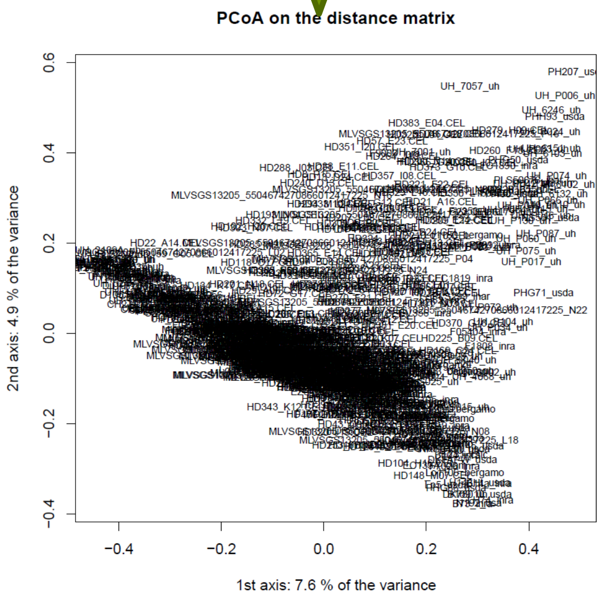


Further breaking population structure

Objectives: evaluate (i) main determinants of divergence and complementarity between Flints and Dents, (ii) possible gene fluxes, (iii) suitable genomic selection models



573 New DH Flint x Dent lines (coord. Cyril Bauland, SMH). Genotyped with 15k SNP Limagrain Affymetrix



Analysis: Delphine Madur

- Per se evaluation of new inbreds and their parents (>1000), two years (PhD Simon Rio)
- 343 hybrids according incomplete factorial between 400 FxD lines (5 trials 2016 - 2017, coord. expe Cyril Bauland and Julie Fievet,)

General conclusions

- ✓ Inbred lines created at different periods since then constitute highly valuable materials for evolutionary and quantitative genetics studies
- ✓ Association genetics reveals important genomic regions: *Vgt1, Vgt2 (Zcn8), Vgt3* for flowering time, Chr 4 for cold tolerance, Chr 6 for heat stress
- ✓ GWAS efficient to find QTL with the strongest contribution to the variation of a given trait x environment combination. Multi environment analyses help understanding adaptive effects
- ✓ Increasing marker densities for GWAS (DROPS and AMAIZING projects): 600 k array, Genotyping by sequencing -> increases number of detected QTLs
- ✓ Limits in power better understood (frequency, pedigree structure …)

Rk. Same experiments can be used to identify genetic resources of interest and calibrate genomic prediction equations

Ongoing work, Perspectives

- ✓ Validation of associations with introgression populations, epistasis
- ✓ Further increase and optimize panels for GWAS (F x D, Flint)
- ✓ Integrate results through appropriate GWAS meta-analysis
- ✓ Implement hybrid specific designs (no tester)
- ✓ Further densify genotyping including imputation
- ✓ New reference genome F7, Ep1 by TUM (available at maize GDB), F2 and other ongoing within project "Amaizing"

  -> opportunities for structural variation analysis, new marker development, …

Rk. Linkage mapping (CornFed EU NAM; Promaïs SAMMCR, VALRG) bring complementary information for rare alleles

Acknowledgements

GQE Le Moulon
(INRA, CNRS, UPS, INAPG)
S. Bouchet, R. Rincent, S. Negro,
S. Nicolas, C. Clipet, J. Fievet, L. Moreau, A. Charcosset
D. Madur, V. Combes, F. Dumas
P. Bertin, S. Jouane, D. Coubriche,
C. Montheil, Ph. Jamin , M. Falque,

INRA Lepse Montpellier:
E. Millet, F. Tardieu, C. Welcker

INRA Saint Martin de Hinx
J. Laborde, C. Palaffre and colleagues

INRA AGAP Montpellier
B. Gouesnard

Cornfed Partners: CSIC, CIAM, EPGV, INRA Mons TUM, Univ. Hohenheim, IPK Gattersleben, Syngenta, Limagrain,KWS, Biogemma

Amaizing Partners: Caussade Semences, Euralis, KWS, Limagrain, Maïsadour, RAGT, Syngenta

Funding