

FsoV



Sélection génomique pour prédire rendement et qualité brassicole en orge d'hiver six rangs

Gilles Charmet/ Amélie Genty



Deux types d'orge



- 2-row vs. 6-row
- 1 gene/30.000 genes



Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene

Takao Komatsuda[†], Mohammad Pourkheirandish^{†*}, Congfen He^{*}, Perumal Azhaguvel^{*}, Hiroyuki Kanamori⁵, Dragan Perovic[¶], Nils Stein[¶], Andreas Graner[¶], Thomas Wicker[¶], Akemi Tagiri^{*}, Udda Lundqvist^{**}, Tatsuhito Fujimura[‡], Makoto Matsuoka^{††}, Takashi Matsumoto^{*}, and Masahiro Yano^{*}

^{*}National Institute of Agrobiological Sciences, Tsukuba 305-8602, Japan; [†]Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba 305-8572, Japan; ⁵Institute of Society for Techno-Innovation of Agriculture, Forestry, and Fisheries, Tsukuba 305-0854, Japan; [¶]Leibniz Institute of Plant Genetics and Crop Plant Research, D-06466 Gatersleben, Germany; [¶]Institute of Plant Biology, University of Zürich, CH-8008 Zürich, Switzerland; ^{**}Svalöf Weibull AB, SE-268 81 Svalöv, Sweden; and ^{††}Bioscience and Biotechnology Center, Nagoya University, Nagoya 464-8601, Japan



GENOMALT: CONTEXTE

La France est le 2e exportateur d'orge brassicole dans le monde

1,97 million d'hectares d'orge en 2021, soit 21 % de la surface céréalière sur 120 000 exploitations. **10,4 M de tonnes** produits, dont 5,7 M exportées. Les orges de brasserie **représentent plus de 4 millions de tonnes** soit plus d'un tiers du total

La culture de l'orge d'hiver 6 rangs brassicoles est une **particularité française qui est grandement responsable de la compétitivité** de la filière Orge-Malt-Bière

La plupart des schémas de sélection variétale 6RH utilise la production d'haploïdes doublés (HD) comme méthode d'accélération

Sources : SSP, FranceAgriMer, Malteurs de France

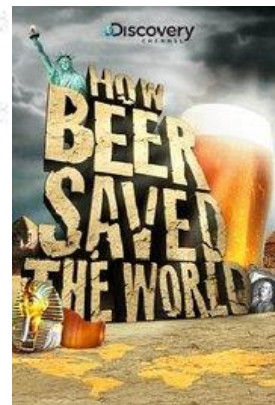
Au niveau mondial, 4° céréale 132 Mt/an



30 %

1 %

70 %



GENOMALT: CONTEXTE

De l'orge au malt...

La malterie française compte 14 unités de production regroupées en 3 groupes qui sont les 3 premiers mondiaux.

En 2020/2021, 1,6 million de tonnes d'orges de brasserie ont été transformées en **1,4 million de tonnes de malt**. Depuis 1967, la malterie française est le 1er exportateur mondial de malt. La France exporte 80 % de sa production de malt soit 1,2 million de tonnes par an.

Sources : FranceAgriMer, Malteurs de France

Du malt à la bière...

Des nombreuses brasseries ont été créées en France ces dernières années. Aujourd'hui, avec **plus de 2 300 brasseries** la France est le 1er pays européen en nombre de brasseries. La brasserie française emploie 7 900 personnes.

La consommation française de bière est de 22 millions d'hectolitres en 2020, soit une consommation de **33 litres par habitant et par an**. Cette consommation place les Français dans les plus faibles consommateurs de bière européens, juste avant l'Italie.

Sources : Brasseurs de France, Brewers of Europ

17/05/2022



Les critères de qualité brassicole (maltière)

Qualité brassicole de l'orge:

- Calibrage: taille des grains pour avoir une bonne teneur en amidon = % grains > 2.5mm
- Protéines: suffisamment pour "nourrir" les levures. mais pas trop (filtration): 9.5 à 11.5%
- Bglucanes (fibres solubles): colmatage des filtres. trouble de la bière (mais recherchées pour nutrition humaine)

Qualité du malt:

- Taux d'extrait: quantité de matière solubilisée au brassage
- Friabilité
- Pouvoir diastasique: activité des amylases dont beta essentiellement

Les tests de micromaltage sont long/couteux/exigeants en grains

Intérêt des prédictions génomique en (pre)selection



GENOMALT: Matériel et méthodes

MATERIALS

259 breeding lines (DH) from SECOBRA

315 breeding lines from RAGT

105 « founder lines ». i.e. registered varieties freely available under UPOV agreement

PHENOTYPING: 2 harvest Years 2018-2019

SECOBRA: 3 locations Thoiry. Auffay. Warmeriville. 2018: 1872 plots; 2019: 1327 plots

RAGT: 2 locations Cupperly and Premesques. 2018 & 2019: 904 plots

Each compagny evaluated its own material + part of founder lines (88 in common)

Variables available in most locations: Yield (dt/ha). protein content (%). thousand grains weight (g). test weight (Kg/hl). Calibration (%). heading (days from january 1st). plant height (cm)

Malting traits evaluated from a single location by breeder & by year (no replicate)

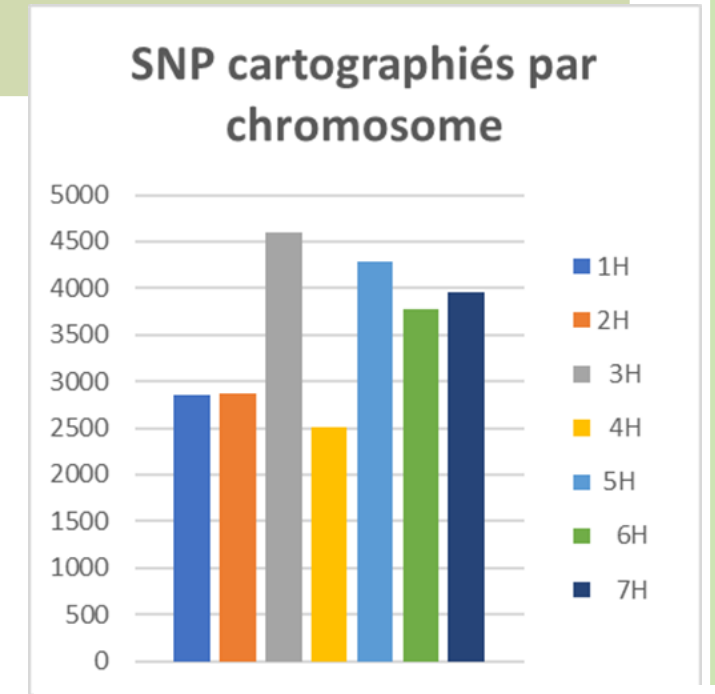
Later in winter: malt-related traits: Friability. extract. viscosity & B-Glucan



GENOMALT: Matériel et méthodes

Genotyping Infinium HD iSelect (JHI) 50K

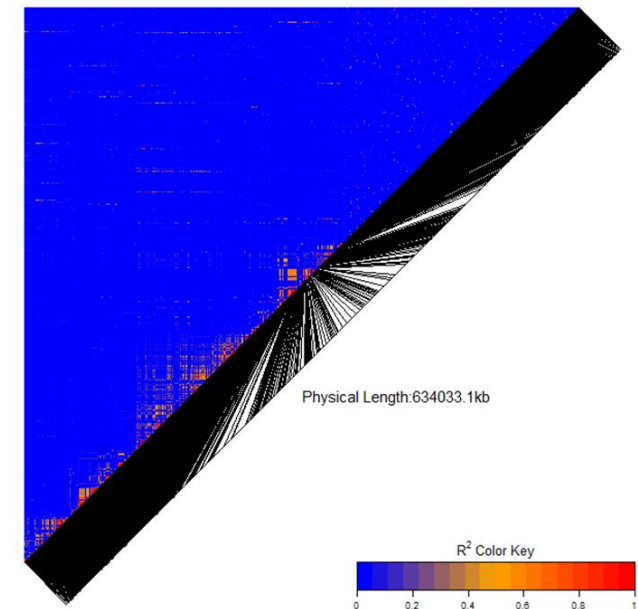
- SECOBRA 44 040 SNP
- RAGT 40 705 SNP (QC)
- In COMMON 38 366 SNP
- Discarding SNP with NA>20% 38 223 SNP
- Discarding SNP with Het > 0.05 36 612 SNP
- Discarding SNP with MAF <0.01 24 945 SNP
- Keeping SNP mapped on 1H-7H 24 101 SNP



Data pre-processing

- Homogeneising codes A/A to A:A. C/C to C:C etc
- Recoding minor allele -1. heterozygotes 0. major allele +1
- Computing relationship matrices (GMALT.K)
- Quality control by checking LD heamaps

Pairwise LD - Chromosome 7H



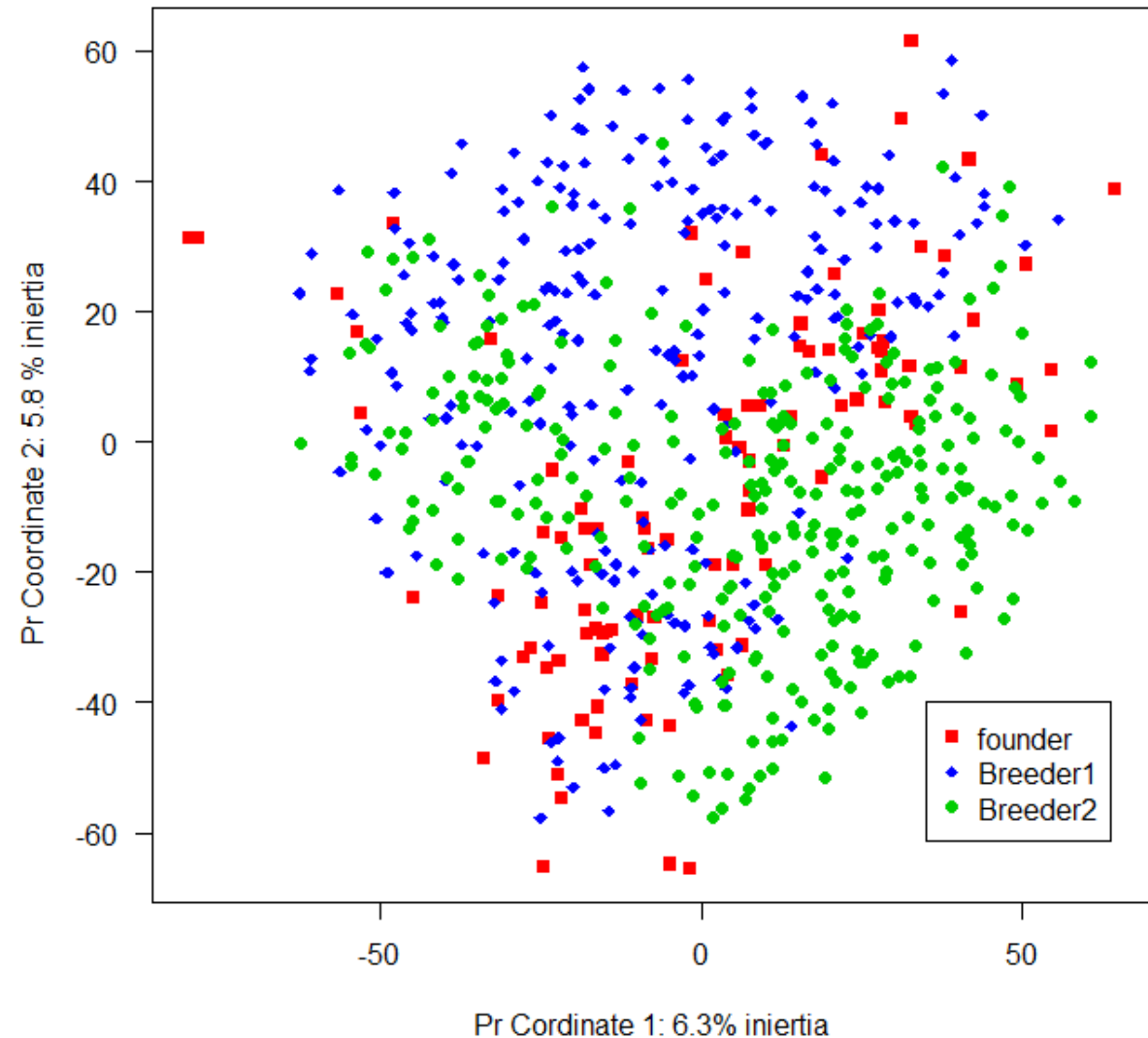
Graphical display of global K matrix

Nuage globalement sphérique: pas de structure clairement tranchée.

Conséquence du germplasm limité à R6H, serait différent si R2vsR6 ou Hiver vs printemps

Large chevauchement des deux sélectionneurs, avec variétés au milieu, mais début de divergence génétique $F_{st} = 0.03$ (vs 0.09 à 0.15 aux USA)

CMD-SCALE of GENOMALT lines from 24 101 SNP



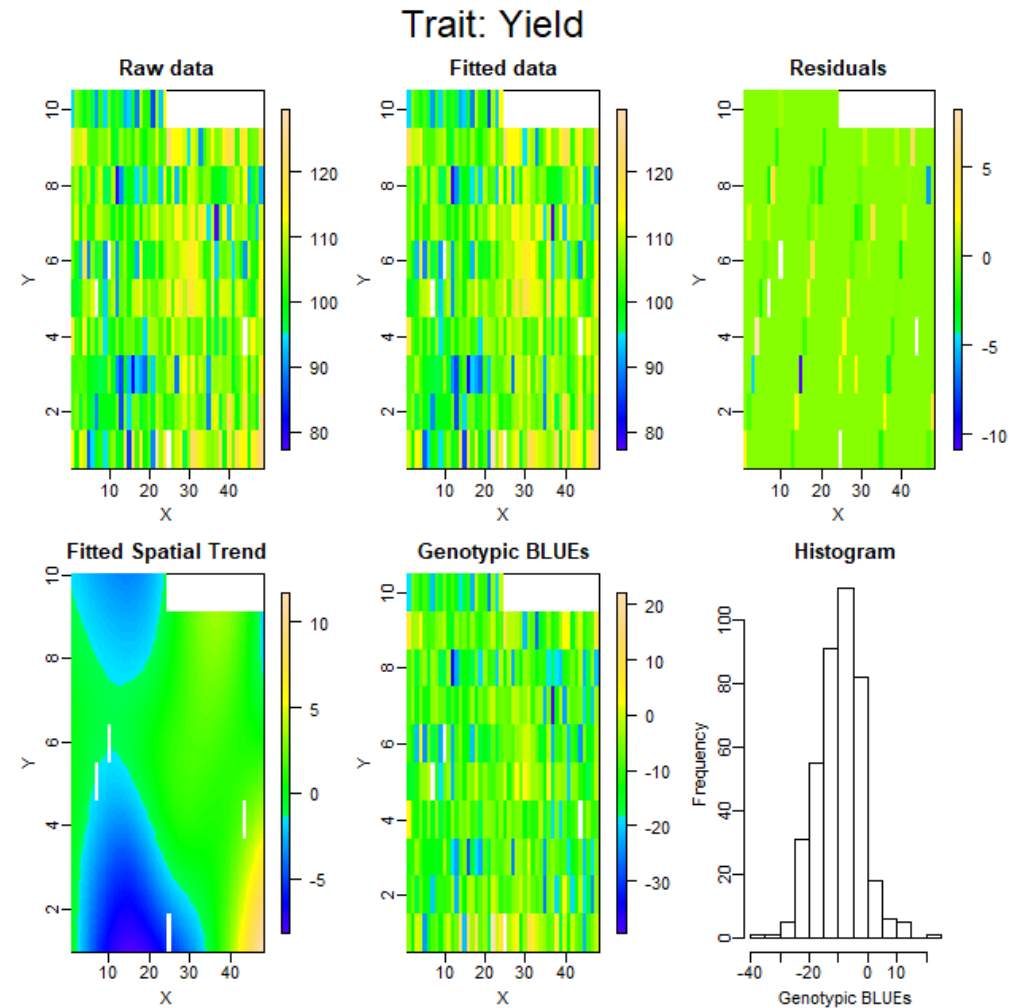
ANALYSES OF FIELD DATA 2-YEARS 2018-2019

Site:Block ***: Spatial adjustment: site by site

```
# e.g. Yield_RAGT2019_PRE.M0=  
SpATS(response="Yield".spatial=~SAP(  
X.Y.nseg=c(10.20)).genotype="GENO".  
data=RAGT2019_Pre.  
control=list(tolerance=1e-03))
```

Then assemble spatially adjusted plot
data into a dataframe: GMALTSPAT

Not applicable to malting traits (no
replicate)



Combined analysis of field Data 2018+2019

Merge the spatially adjusted Dataset:

- `GMALT2YEARSPAT=rbind(GMALT2018SPAT. GMALT2019SPAT)`
- 6 Traits: Yield. Prot. TestW. TGW. Cal. Head)

Global ANOVA on spatially corrected Data e.g.

- `Yield2YEAR.LM=lm(Yield~Year + Year:Site + GENO+ GENO*Year + GENO*Site . data=GMALT2YEAR.dtf)`
- `anova(Yield2YEAR.LM)`

ANOVA on raw data for malt-related traits

- `Friability2YEAR.LM=lm(Friability~Env+ GENO+ GENO*Env. data=GMALT2YEAR.dtf) (Env = year*geno)`
- `Anova (Friability2Year.LM)`



Combined analysis of Data 2018+2019

F (Res=1148 df)	Yield	Prot	TestW	TGW	Cal	Head	Friabi	Extract	Visco	BGglu
Year (1) 2	1282	274	909	95	118	220				
GENO (698)	2.97	4.78	4.55	7.64	7.74	2.77	30	11	20.5	27.8
Site/Env (4)	99	98		716	545	1280	218	67	270	281
Year*Site (4)	845	149	546	309	76	632				
Y*GENO (664)	1.23	1.14	1.28	1.28	1.29	1.39				
Site*GENO (1243)	1.12	1.10	0.95	1.09	1.54	0.98	2.8	1.9	4.1	3.9

GENO MAIN EFFECT >> GxE



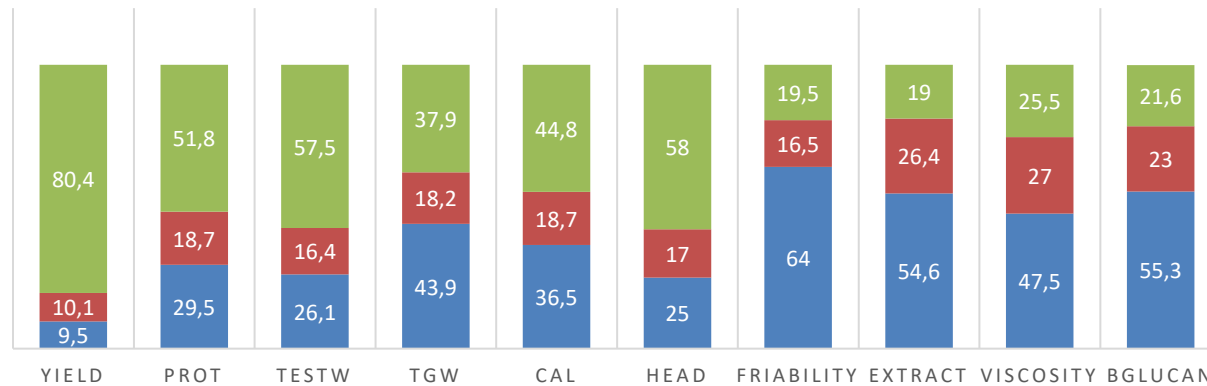
work on ADJUSTED MEANS



Combined analysis of Data 2018+2019

% SUM OF SQUARES

■ GENOTYPE ■ GxE ■ Environment



**Genotype SS very high for malt related traits
Same for variance components (see
Frontiers)
Broad sense heritabilities ranging from
0.60 (yield) to 0.89 (friability)**

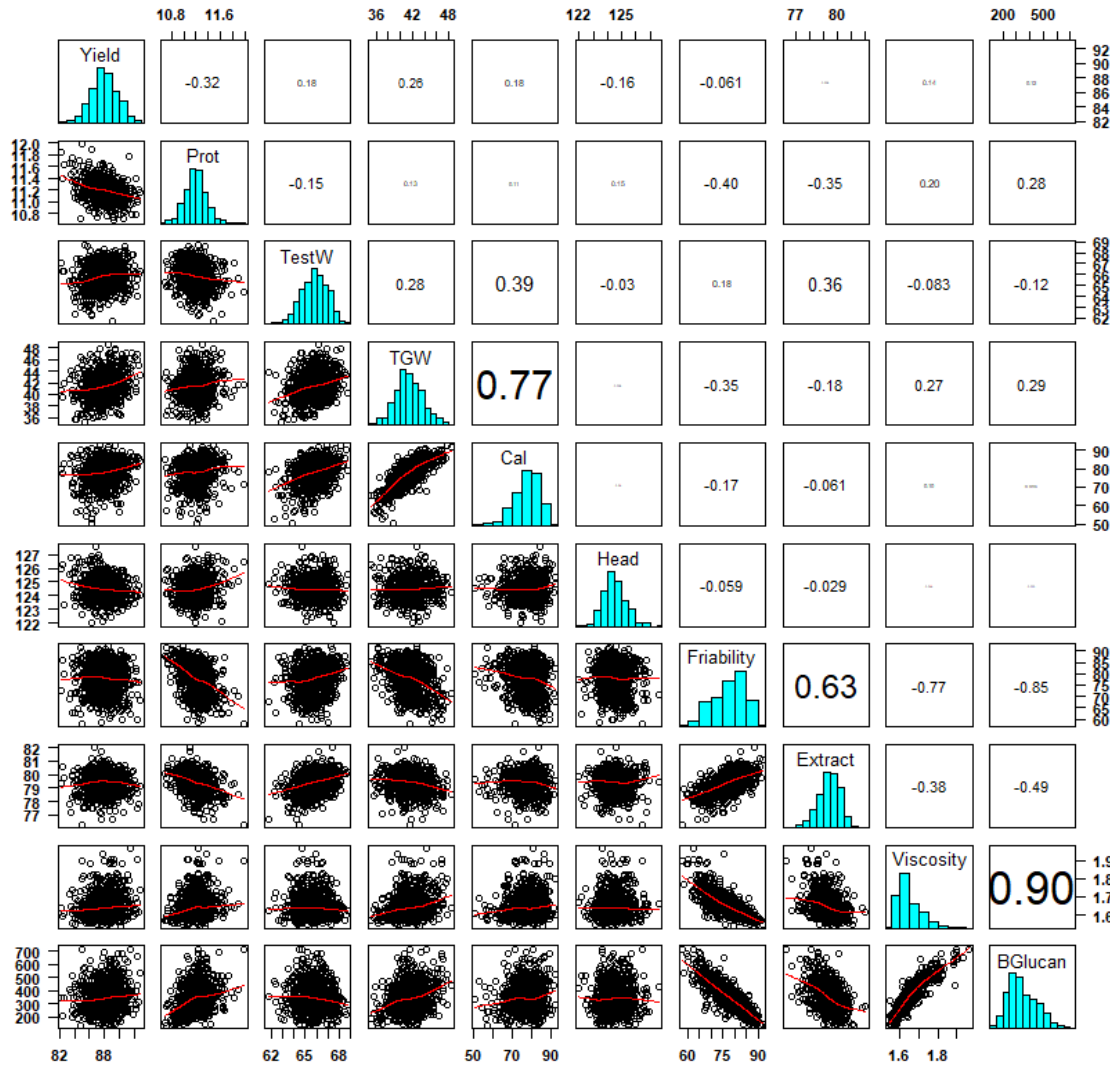
Estimation of main GENO effects with model:

- $\text{Yield2YEAR.LM} = \text{lm}(\text{Yield} \sim \text{Year} + \text{Year:Site} + \text{GENO}, \text{data} = \text{GMALT2YEARS PAT})$
- $\text{Yield2YEAR.ADJ} = \text{Yield2YEAR.LM}\$coefficients[3:712]$
- Add Grand Mean μ to get relevant value
- Look at « genetic » correlations

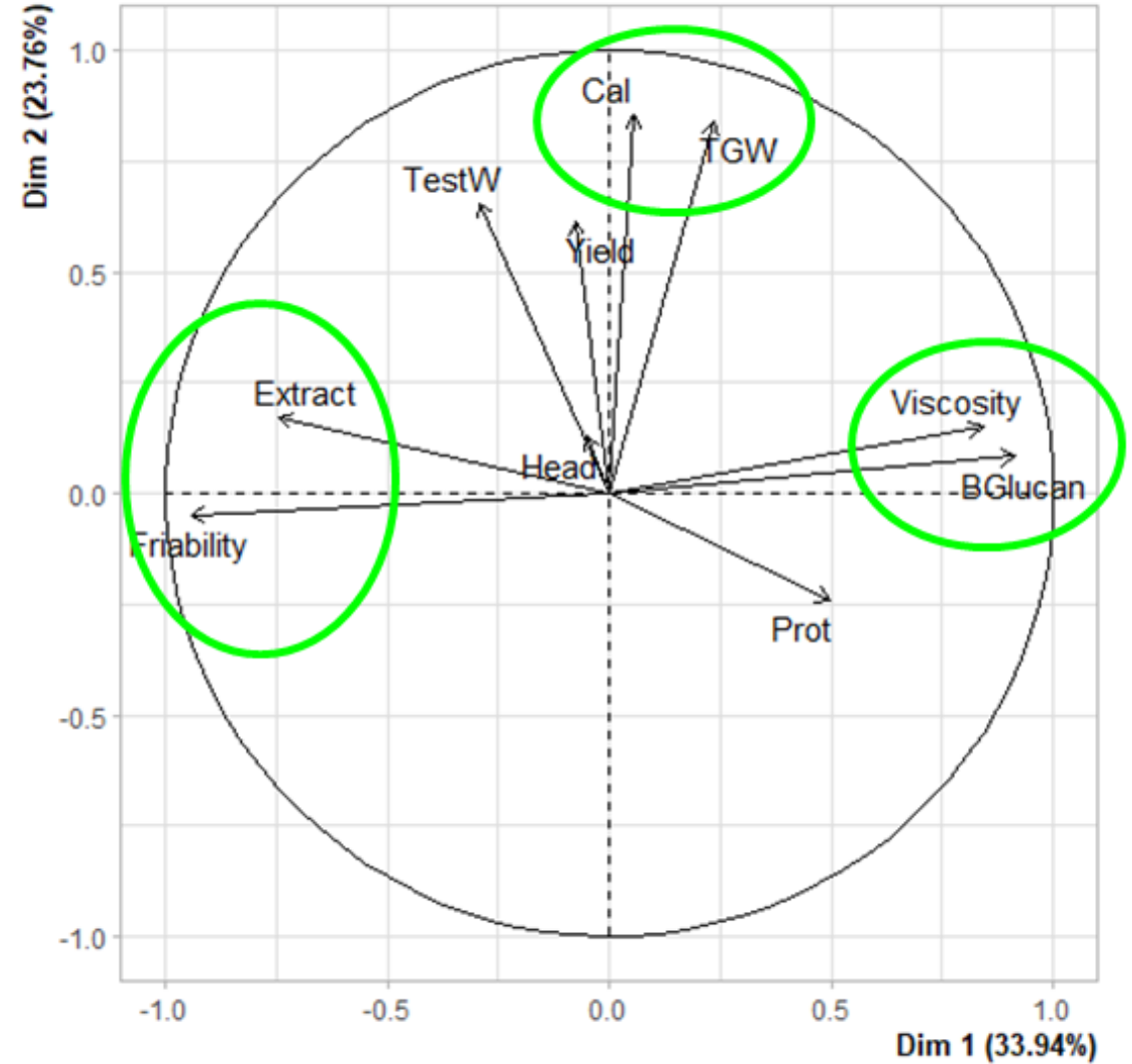


Genotypic correlations: favourable or neutral

correlation of genotypic adjusted values



PCA graph of variables



GENOMIC PREDICTIONS

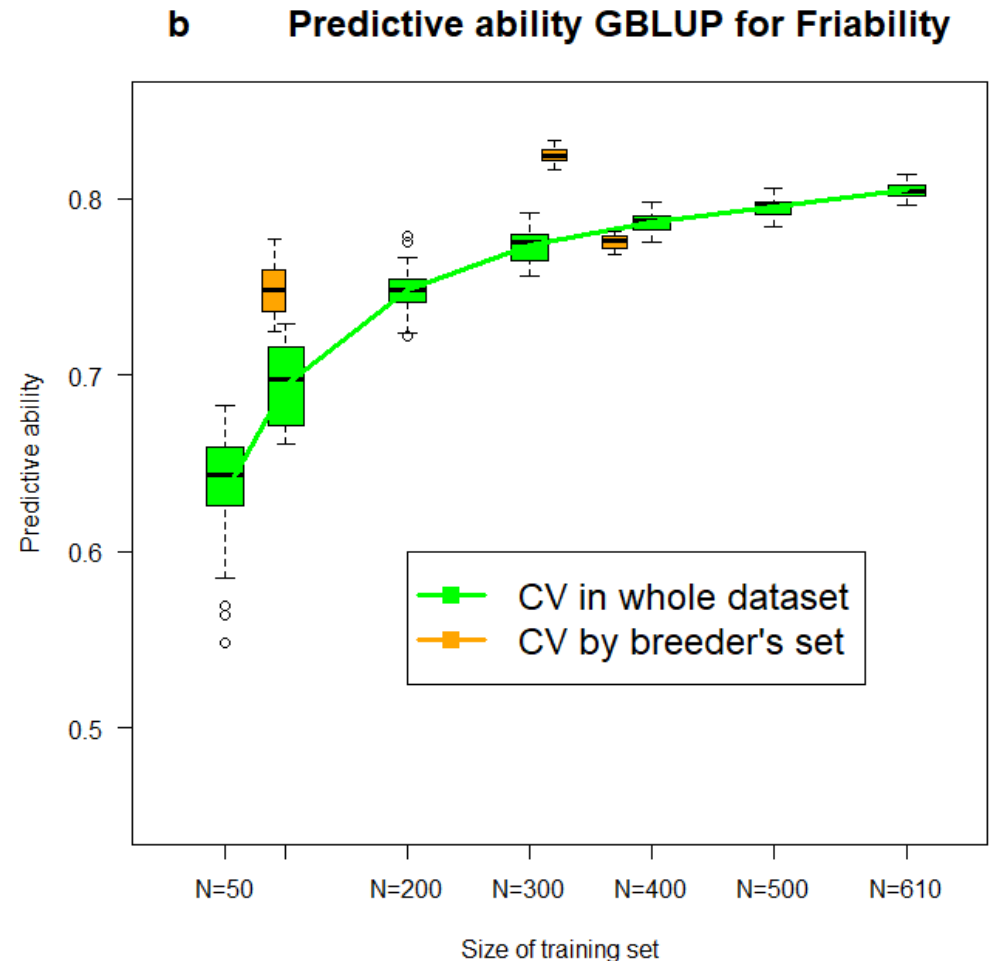
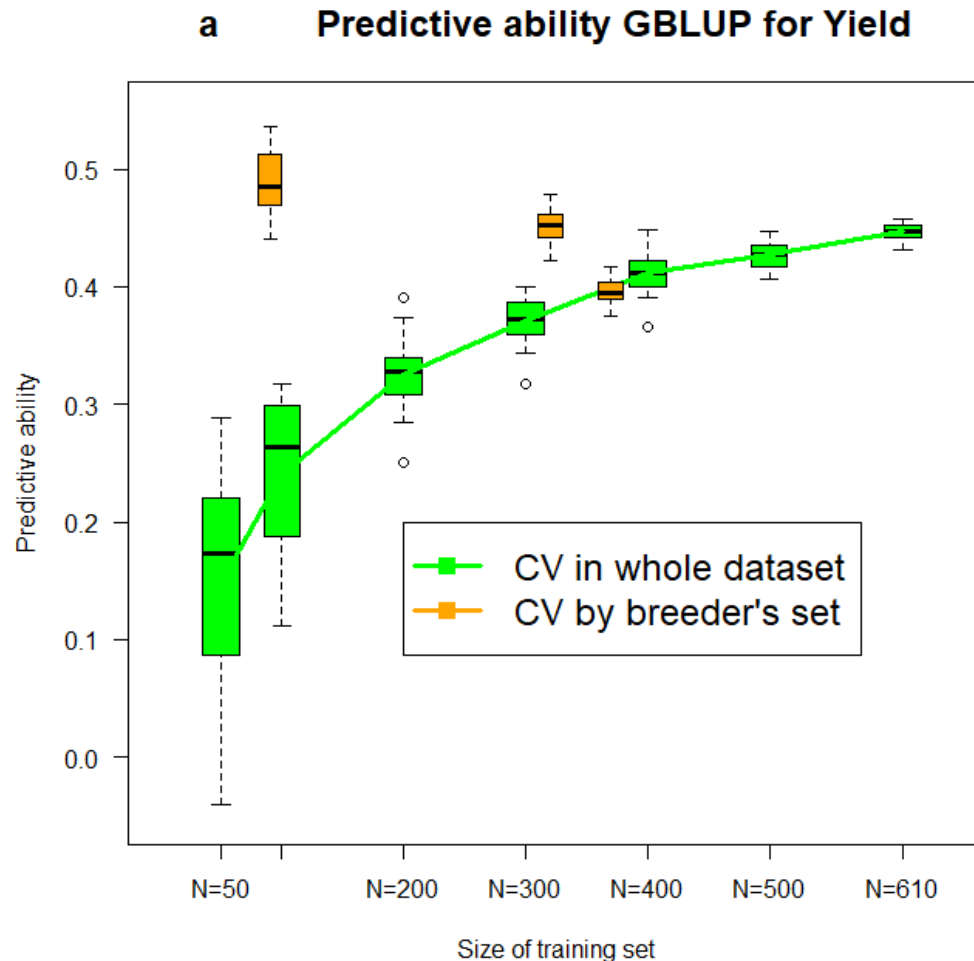
GBLUP cross validation

TRAIT	randomCV N = 679)	BRE1+FO.CV (N = 364)	BRE2+FO.CV (N=420)	FounderCV (N = 105)
Yield	0.556 / 0.012	0.530 / 0.026	0.431 / 0.023	0.490 / 0.056
Prot	0.515 / 0.014	0.645 / 0.016	0.215 / 0.032	0.379 / 0.064
TGW	0.692 / 0.010	0.763 / 0.014	0.540 / 0.022	0.585 / 0.048
TestW	0.661 / 0.014	0.722 / 0.016	0.578 / 0.016	0.658 / 0.050
Cal	0.714 / 0.012	0.697 / 0.016	0.598 / 0.020	0.350 / 0.072
Head	0.655 / 0.019	0.632 / 0.032	0.676 / 0.034	0.313 / 0.104
Friability	0.814 / 0.006	0.823 / 0.009	0.782 / 0.014	0.745 / 0.032
Extract	0.696 / 0.008	0.766 / 0.027	0.654 / 0.014	0.785 / 0.028
Viscosity	0.698 / 0.011	0.743 / 0.011	0.651 / 0.020	0.706 / 0.036
BGlucan	0.762 / 0.010	0.796 / 0.011	0.725 / 0.017	0.740 / 0.022

High to very high predictive abilities, particularly for malt-related traits



Effect of size of training set

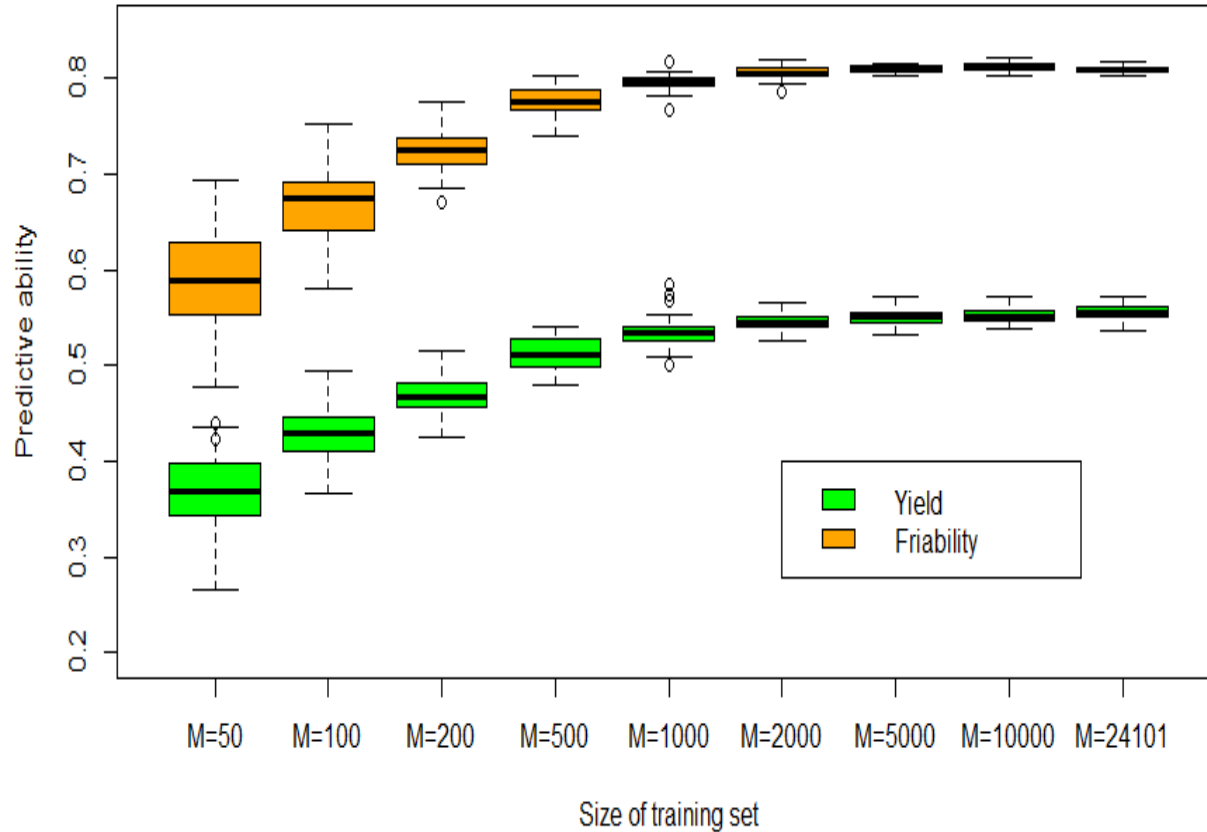


**Predictive abilities of breeder's subsets > random subsets (particularly varieties)
Genetic diversity of higher quality of phenotypes?**

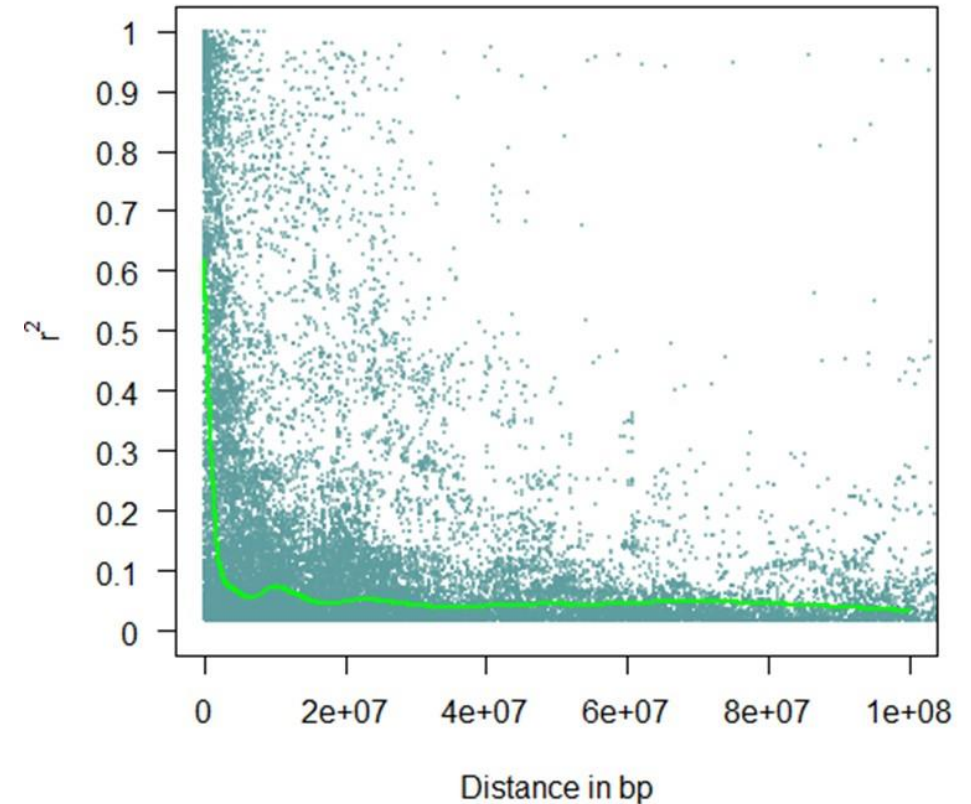


Effect of number of markers

Effect of Marker number on predictive ability



LD decay on chromosome 1H



**Optimum marker number 1000/2000; may be explained by LD extent:
LD >0.4 up to 2 Mb (genome size 4,250 Mb)**



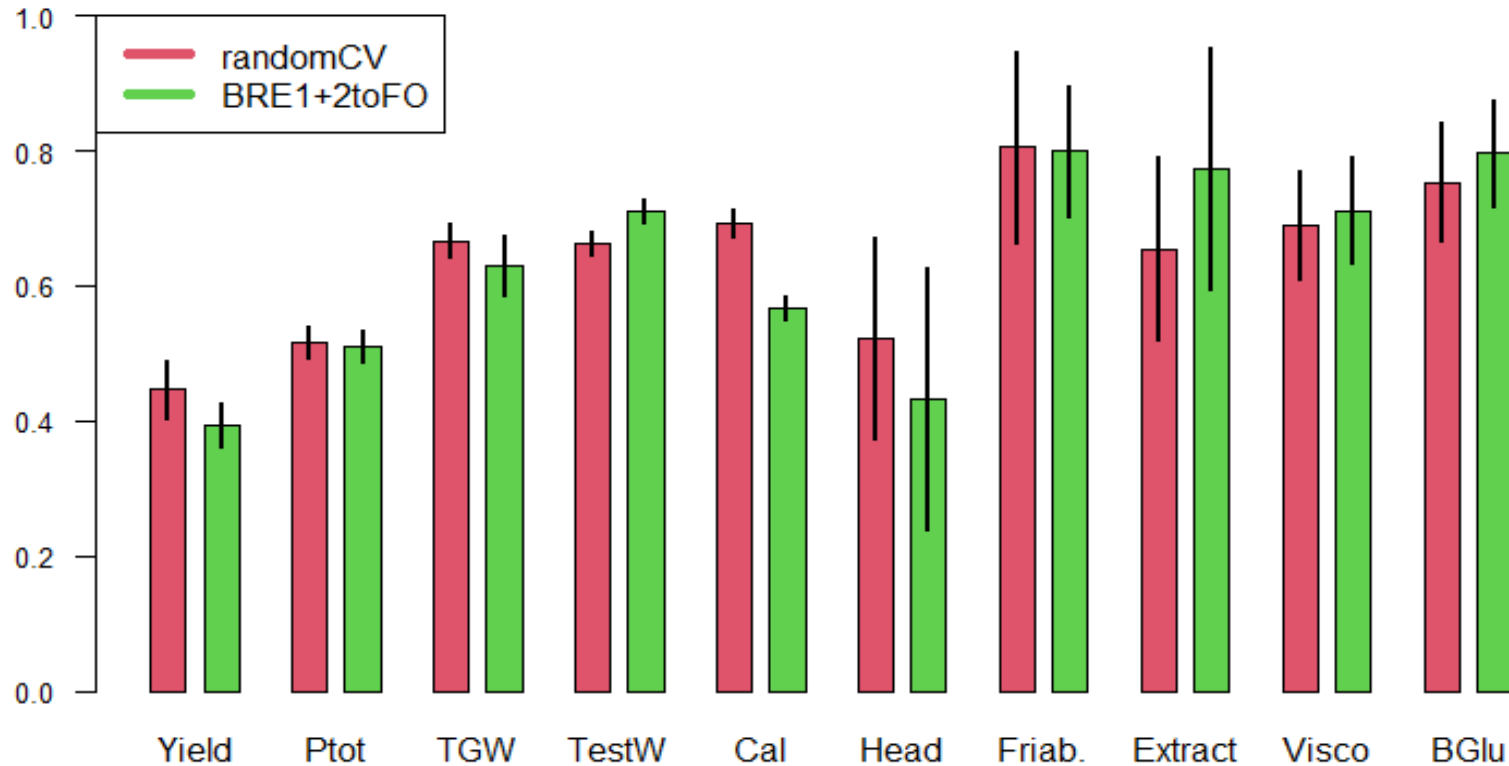
GBLUP cross population validation

TRAIT	BRE1+2 to FO	BRE1 to FO	BRE2 to FO
Yield	0.394 / 0.075	0.254 / 0.101	0.248 / 0.107
Prot	0.509 / 0.097	0.575 / 0.086	0.250 / 0.095
TGW	0.630 / 0.071	0.595 / 0.064	0.547 / 0.096
TestW	0.710 / 0.049	0.562 / 0.083	0.556 / 0.066
Cal	0.568 / 0.068	0.529 / 0.071	0.472 / 0.101
Head	0.432 / 0.090	0.408 / 0.095	0.332 / 0.111
Friability	0.799 / 0.040	0.718 / 0.052	0.677 / 0.050
Extract	0.773 / 0.039	0.653 / 0.058	0.726 / 0.040
Viscosity	0.712 / 0.044	0.644 / 0.061	0.572 / 0.072
BGlucan	0.796 / 0.040	0.745 / 0.048	0.706 / 0.060

High PA with independent validation set, no clear advantage of merging breeder's data



GBLUP random CV vs cross pop validation



Independent validation nearly as predictive as random CV



Predictive abilities of other genomic models

TRAIT	h ²	h	GBLUP	Bayes Cpi	LASSO	EGBLUP
Yield	0.551	0.742	0.446 / 0.022	0.443 / 0.026	0.338 / 0.030	0.463 / 0.016
Prot	0.613	0.783	0.517 / 0.016	0.514 / 0.013	0.482 / 0.020	0.513 / 0.012
TGW	0.837	0.915	0.667 / 0.012	0.669 / 0.015	0.627 / 0.011	0.671 / 0.010
TestW	0.775	0.880	0.662 / 0.012	0.666 / 0.012	0.647 / 0.017	0.677 / 0.009
Cal	0.853	0.923	0.693 / 0.013	0.690 / 0.013	0.639 / 0.021	0.705 / 0.013
Head	0.652	0.807	0.522 / 0.022	0.519 / 0.018	0.511 / 0.019	0.518 / 0.017
Friability	0.895	0.946	0.805 / 0.009	0.806 / 0.006 0.814	0.789 / 0.008	0.805 / 0.011
Extract	0.753	0.868	0.654 / 0.009	0.658 / 0.009	0.650 / 0.010	0.669 / 0.009
Viscosity	0.769	0.876	0.690 / 0.011	0.697 / 0.007	0.657 / 0.015	0.700 / 0.009
BGlucan	0.851	0.922	0.753 / 0.009	0.754 / 0.009	0.722 / 0.015	0.761 / 0.009

Few differences, LASSO underperformed, EGBLUP slightly > for low h² traits



Take-home message

- Reliable phenotypic data, despite small and unbalanced design
- Most genetic correlations are favourable to breeder's goals
- Few genetic divergences between breeder's germplasms
- High GBLUP predictive abilities of malt-related traits, either with random CV or independent validation
- (including major QTL detected by GWAS as fixed term does not improve PA)
- No clear advantage of merging breeders' data to improve PA. Small training set of 105 varieties give high PA
- Few differences along prediction models, EGBLUP slightly better
- **RESULTS HIGHLY ENCOURAGING FOR ROUTINE USE OF GS**



Acknowledgements

SECOBRA

Amélie Genty

Pierre Pin

Guillain Cresté

Pierre-Marie Leroux

INRAE

Gilles Charmet

Sophie Bouchet

IFBM

Marc Schmitt

Catherine Colin

Sophie Schwebel

RAGT 2N

Bruno Claustres

Nathalie Leroy

Christopher Burt

Régis Dupont

Mathieu Tison

Gaetan Touzy





merci

