Dissecting GxE and QTLxE of maize yield in contrasting scenarios of light, heat and water deficit

Emilie J. Millet, Willem Kruijer, Santiago Alvarez Prado, Llorenç Cabrera-Bosquet, Aude Coupel-Ledru, Sébastien Lacube, Yves Gibon, Sylvain Prigent, Claude Welcker, Alain Charcosset, et al.

To cite this version:

Emilie J. Millet, Willem Kruijer, Santiago Alvarez Prado, Llorenç Cabrera-Bosquet, Aude Coupel-Ledru, et al.. Dissecting GxE and QTLxE of maize yield in contrasting scenarios of light, heat and water deficit. XXIVth Eucarpia Maize and Sorghum Conference, TUM School of Life Sciences Weihenstephan, Oct 2019, Freising, Germany. hal-04154261

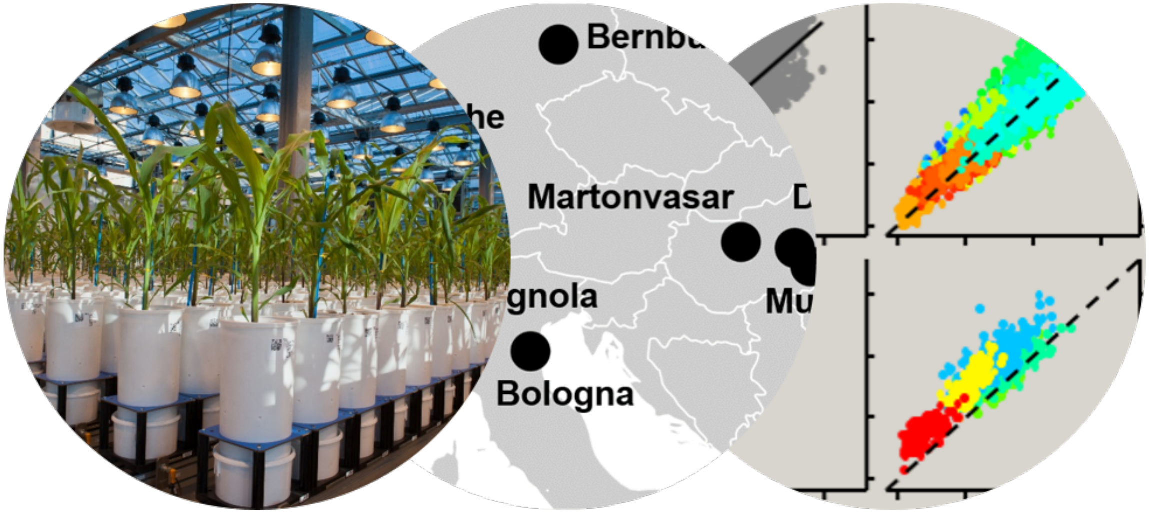
HAL Id: hal-04154261
https://hal.inrae.fr/hal-04154261v1

Submitted on 6 Jul 2023

Dissecting G×E and QTL×E of maize yield in contrasting scenarios of light, heat and water deficit.

Millet E, Kruijer W, Alvarez Prado S, Cabrera-Bosquet L, Coupel-Ledru A, Lacube S, Gibon Y, Prigent S, Welcker C, Charcosset A, van Eeuwijk F, Tardieu F

XXIVth EUCARPIA Maize and Sorghum Conference, 8th October 2019, Munich



WAGENINGEN
UNIVERSITY & RESEARCH

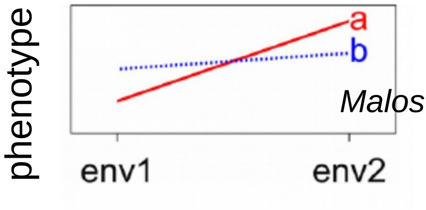
European
Plant Phenotyping
Network
EPPN 2020

DROPS

amaizing

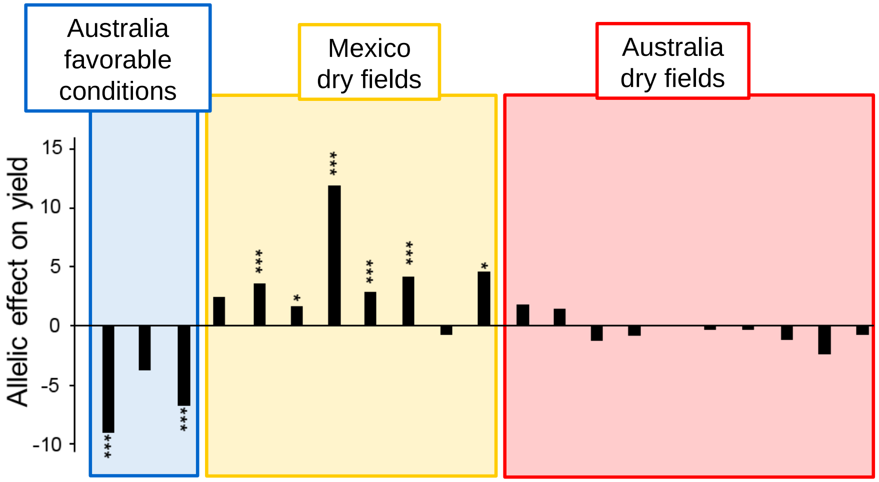
Background

Phenotypic response can differ for different varieties = G × E



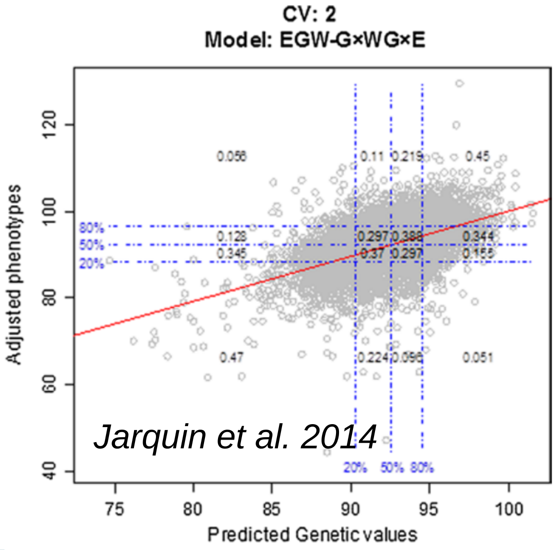
Malosetti et al. 2013

Detection of QTL will depend on the environment
⇒ environment dependent effects

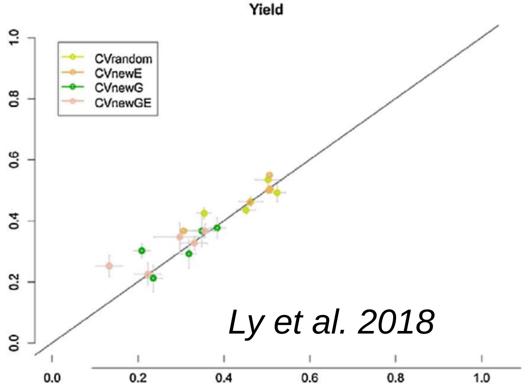


redrawn from
Bonneau et al 2013
ACPFG Austraila

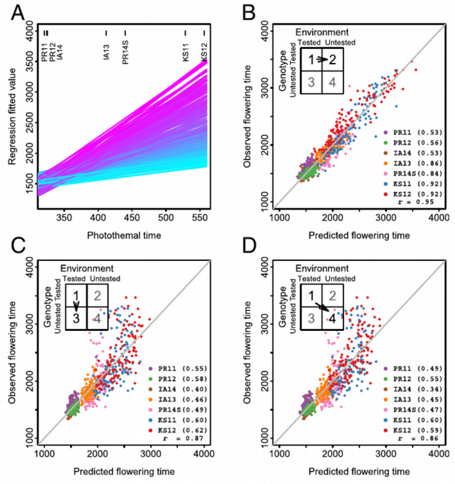
Genomic prediction can assess the performances of thousands of genotypes but faces the difficulty G×E
⇒ Few studies dissect the G×E for predicting grain yield



Jarquin et al. 2014



Ly et al. 2018



Li et al. 2018

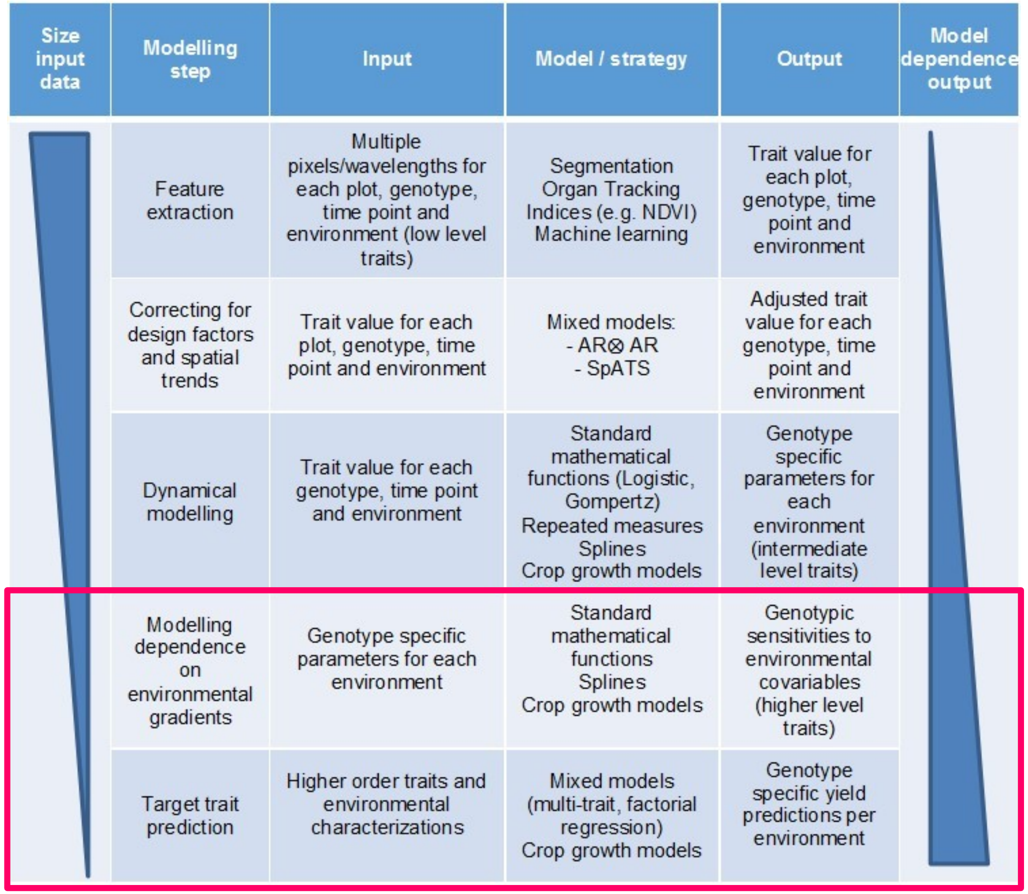
Background

New tools and methods for dissecting the G×E: the «ics»

- Phenomic (new phenotyping)
  → field network, high throughput platform
- Genomic (High density genotyping)
- Environmental characterization (envirotyping, enviromic (Resende et al. 2019 bioRxiv))
  → sensor network, Geographic Information Systems
- Statistics (new class of models)

➔ New modelling strategies to incorporate new phenotyping information in different classes of genotype-to-phenotype (G2P) models

➔ From large numbers of basic and raw secondary phenotype data → relatively few genotype specific parameters with high predictive power



Size input data	Modelling step	Input	Model / strategy	Output	Model dependence output
	Feature extraction	Multiple pixels/wavelengths for each plot, genotype, time point and environment (low level traits)	Segmentation Organ Tracking Indices (e.g. NDVI) Machine learning	Trait value for each plot, genotype, time point and environment	
	Correcting for design factors and spatial trends	Trait value for each plot, genotype, time point and environment	Mixed models: - AR⊗AR - SpATS	Adjusted trait value for each genotype, time point and environment	
	Dynamical modelling	Trait value for each genotype, time point and environment	Standard mathematical functions (Logistic, Gompertz) Repeated measures Splines Crop growth models	Genotype specific parameters for each environment (intermediate level traits)	
	Modelling dependence on environmental gradients	Genotype specific parameters for each environment	Standard mathematical functions Splines Crop growth models	Genotypic sensitivities to environmental covariables (higher level traits)	
	Target trait prediction	Higher order traits and environmental characterizations	Mixed models (multi-trait, factorial regression) Crop growth models	Genotype specific yield predictions per environment	

van Eeuwijk et al. 2019 Plant Science

Background

Novel approaches for dissection of genetic architecture of yield under drought and heat that combines:

- environmental characterization (sensor networks)
- statistical modelling
- genome-wide association
- phenomic tools
- genomic prediction

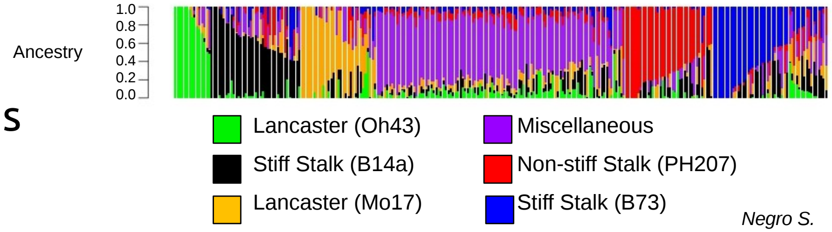
Weclker C.

Tardieu F.

van Eeuwijk F.

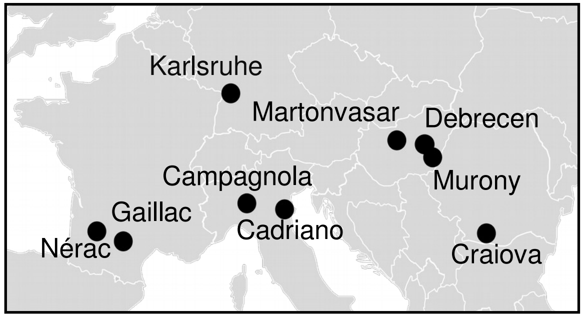
Charcosset A.

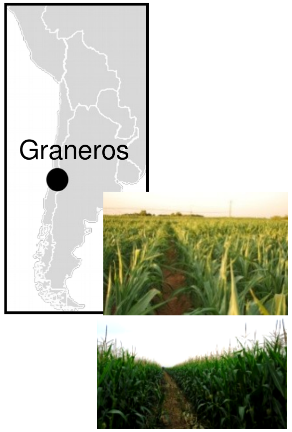
The DROPS panel ~ 250 maize hybrids
(Dent diversity group × flint tester)



Negro S.

Climatic transect over Europe + Chile :
10 Sites, 2 Years (2012, 2013)
diverse drought scenarios : 2 water regimes
→ 25 experiments (Site x Year x Regime)

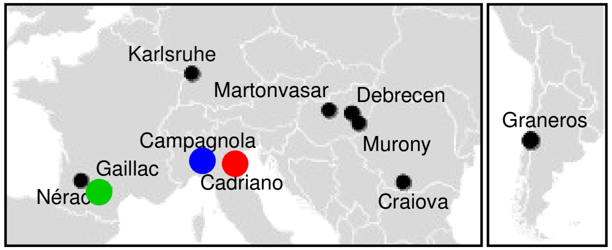




High throughput phenotyping platform (Phenoarch):
3 experimental series
2 water regimes
→ 8 experiments

Large variation of grain yield linked to environment variability ?

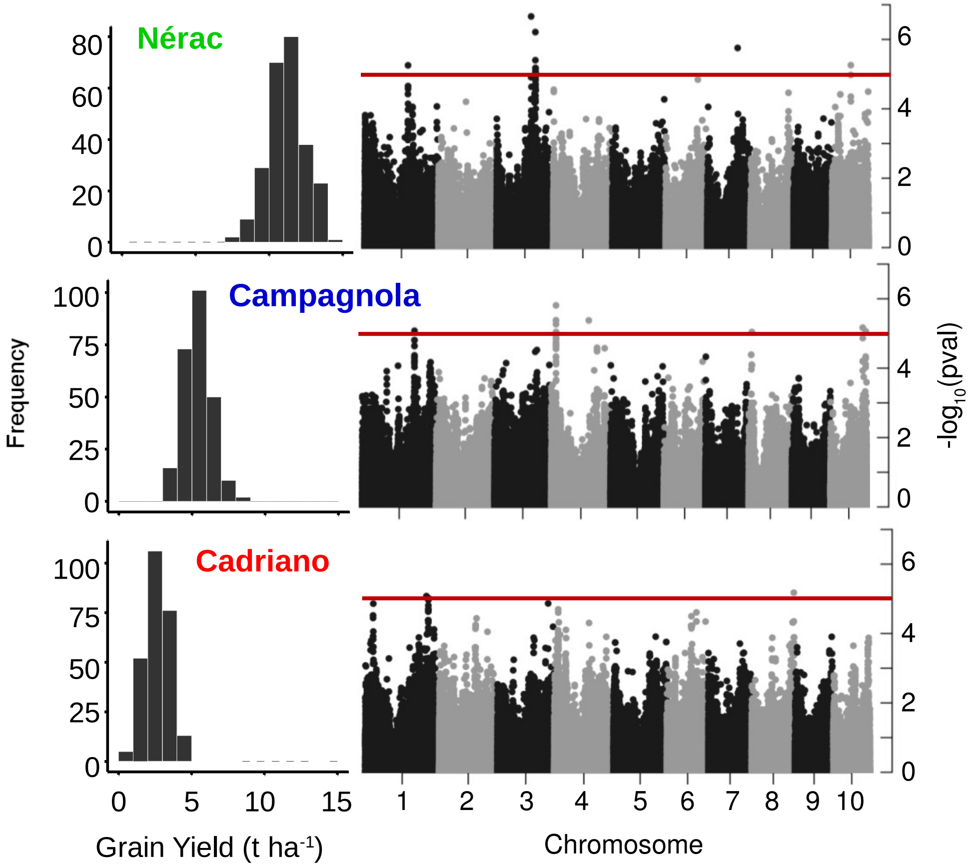
Large distribution of grain yield within and between experiments with a large G×E



467 significant SNPs for grain yield => 48 QTLs

QTL are significants in 7 experiments (from 2 to 23)

=> Can we use the environmental classification to explain the QTL×E ?



WW - Hot(day)

WW - Cool

WD - Hot

Weclker C.

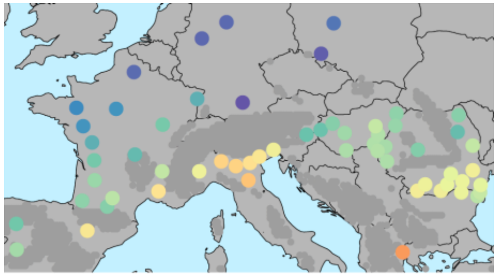
Kruijer W.

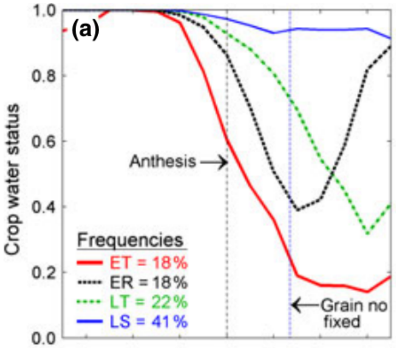
Coupel-Ledru A.

Nicolas S.

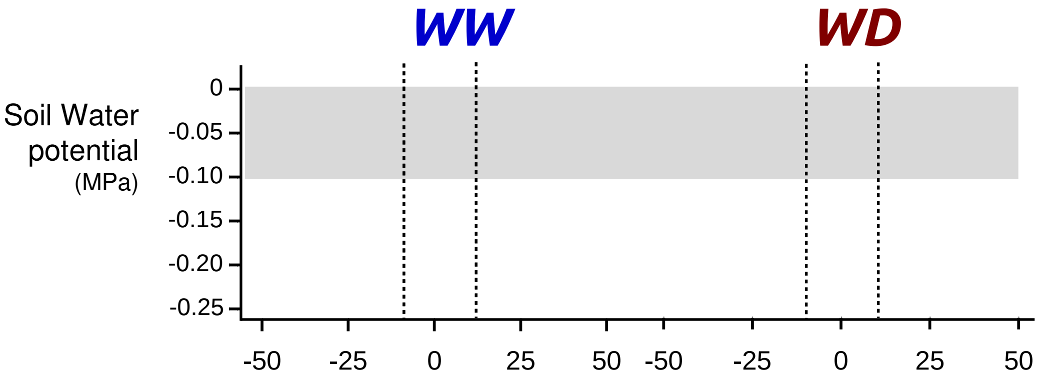
Environmental classification over Europe

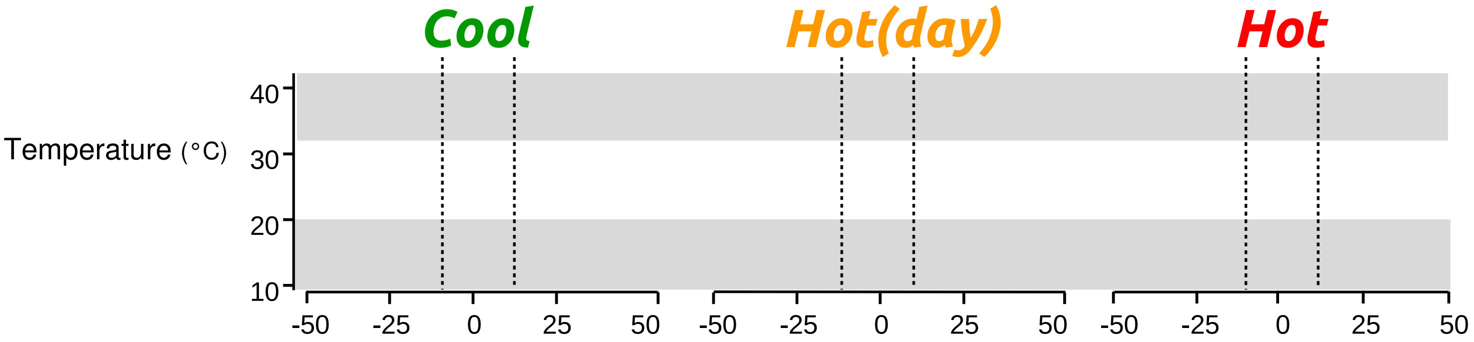
Which scenarios in Europe ?
Simulation of drought scenarios (APSIM) : 55 sites x 35 years
⇒Typical scenarios of drought in Europe





Harrison et al. 2014

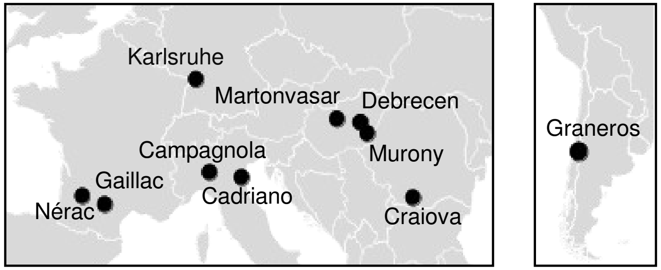


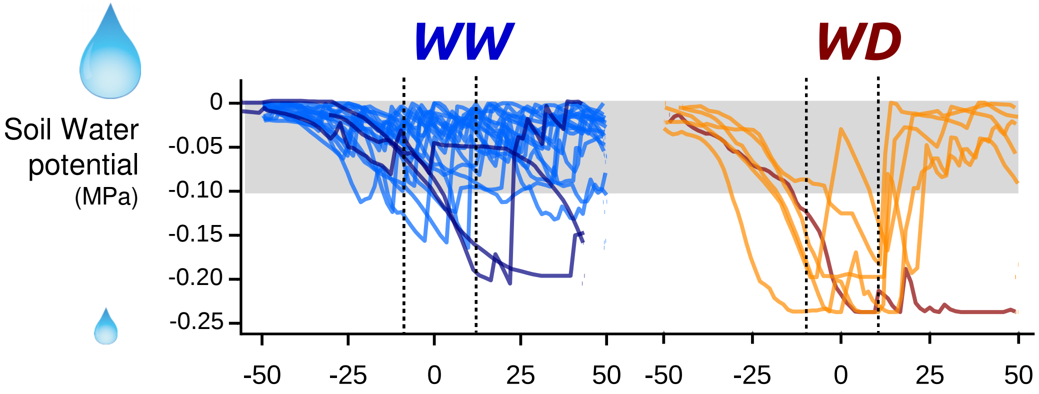


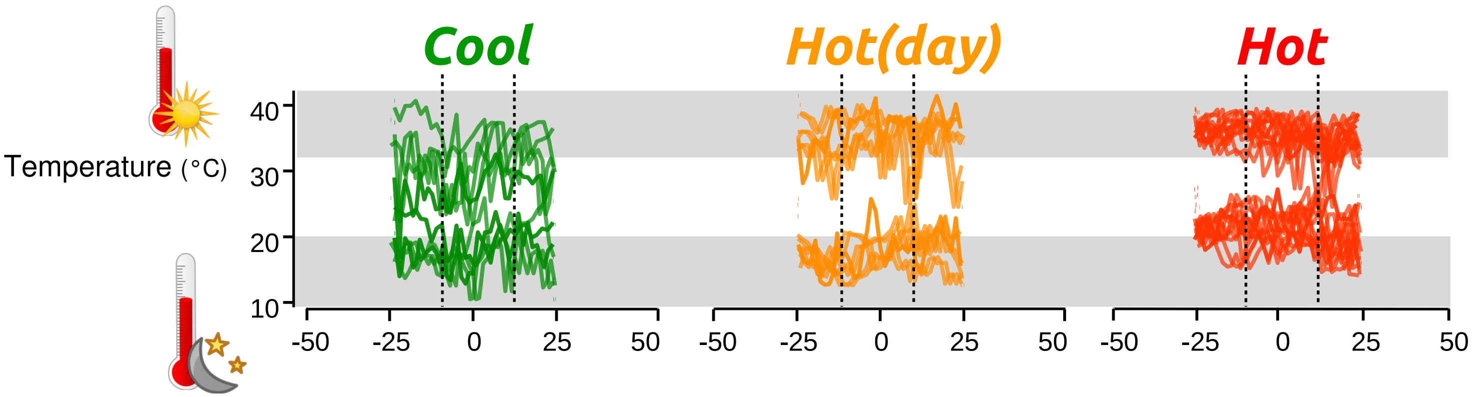
Environmental classification over Europe

Which scenarios in Europe ?
Simulation of drought scenarios (APSIM) : 55 sites x 35 years
⇒Typical scenarios of drought in Europe

Clustering the experiments in these scenarios



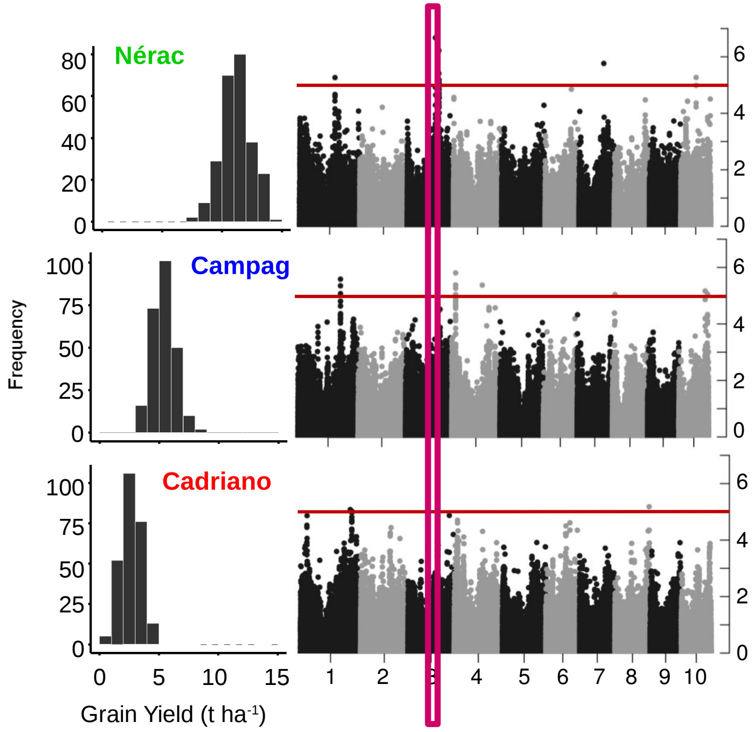




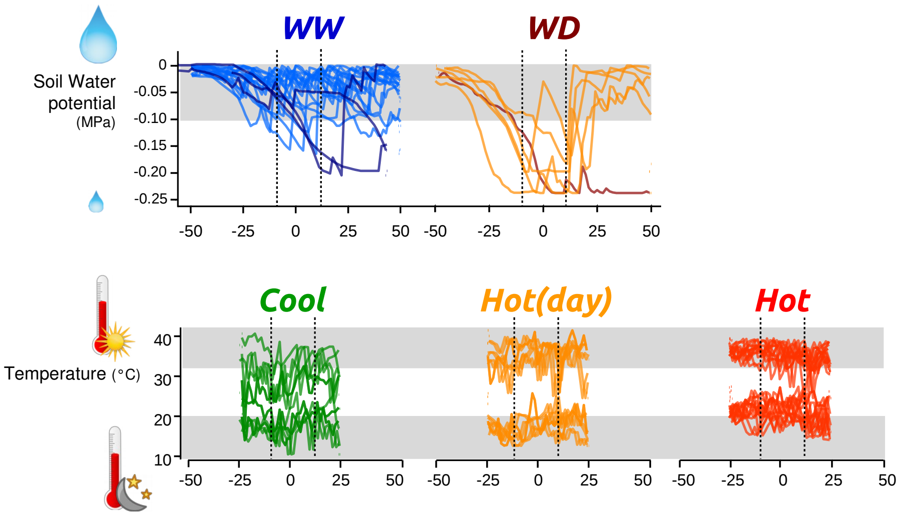
Millet et al. 2016 Plant Physio.

Environmental classification over Europe to explain the allelic effects

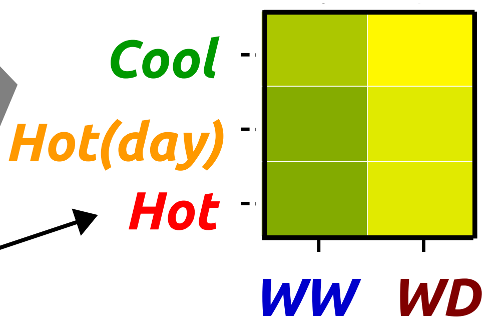
Allelic effect of one QTL

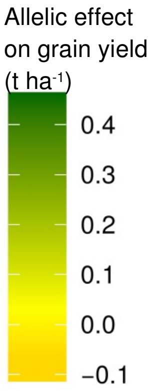


Clusters of experiments



Allelic effect of one QTL, grouped per scenario





Genes of heat tolerance

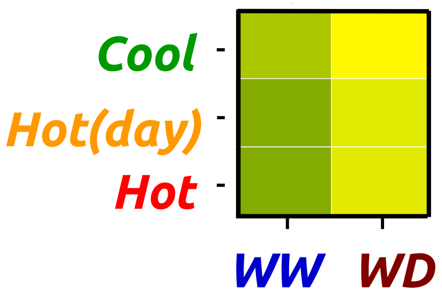
Genes of drought tolerance

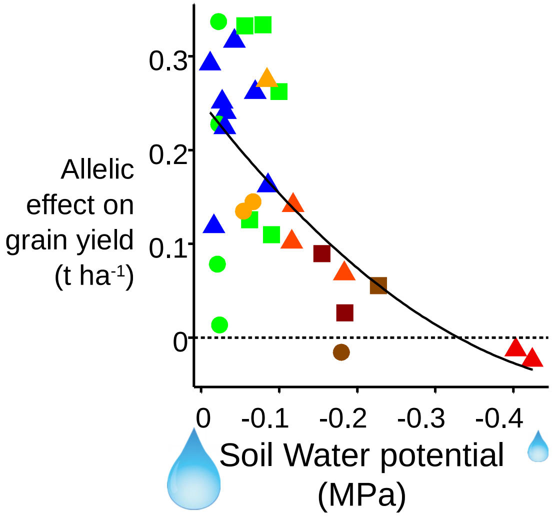
WAGENINGEN
UNIVERSITY & RESEARCH

Millet et al. 2016 Plant Physio.

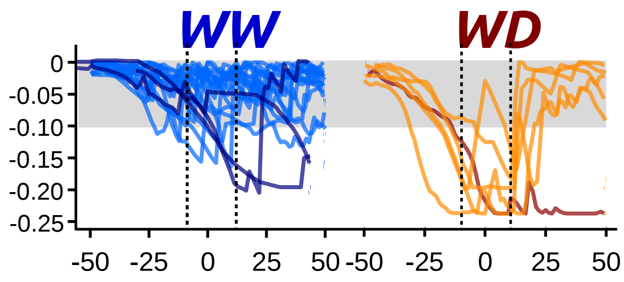
Genomic region displayed scenario-dependent effect

Classification of the QTLs using the QTL×Scenario and QTL×Co-variable





(environmental variable average around flowering time)

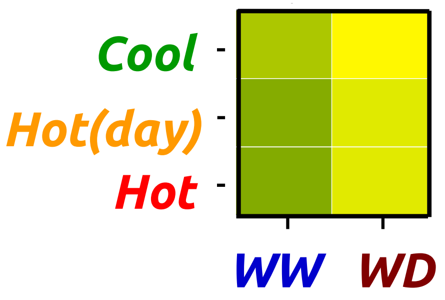


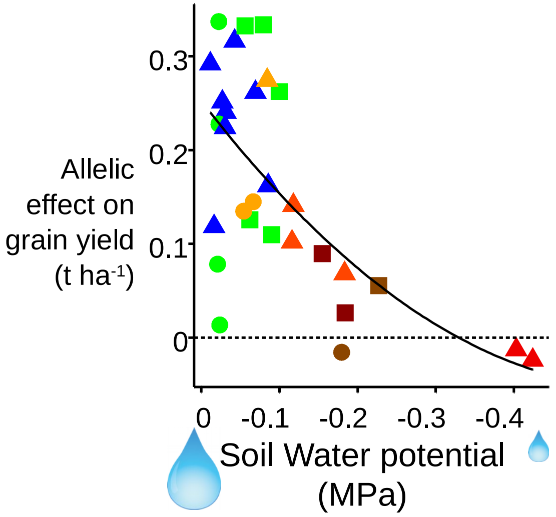
- ➔ 25 QTLs with larger allelic effect in favourable conditions
- ➔ 8 QTLs with high effects in *Hot* and null effect in *Cool*
- ➔ 12 QTLs with high effects in *WD* and no or lesser effect in WW
- ➔ 3 QTLs with no clear pattern of allelic effect

Millet et al. 2016 Plant Physio.

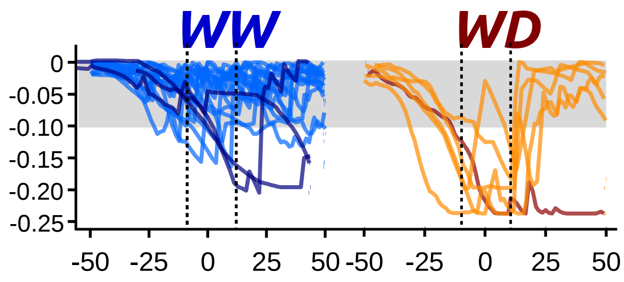
Genomic region displayed scenario-dependent effect

Classification of the QTLs using the QTL×Scenario and QTL×Co-variable





(environmental variable average around flowering time)

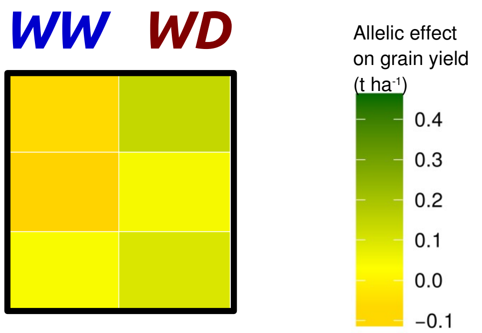


- ➔ 25 QTLs with larger allelic effect in favourable conditions
- ➔ 8 QTLs with high effects in *Hot* and null effect in *Cool*
- ➔ 12 QTLs with high effects in *WD* and no or lesser effect in WW
- ➔ 3 QTLs with no clear pattern of allelic effect

Targeting the genes in the interval based on the response to environment:

- Positive in WD, negative in WW
- QTL increases ABA in leaves (Giuliani et al., 2005)
- Gene involved in ABA biosynthesis

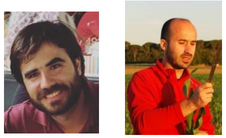
→ high ABA can cause reduction in growth in WW while maintaining growth in WD



Millet et al. 2016 Plant Physio.

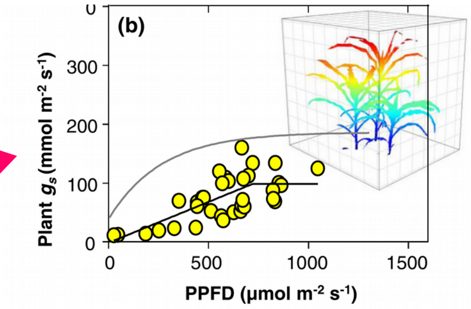
Genomic region displayed scenario-dependent effect: in platform too

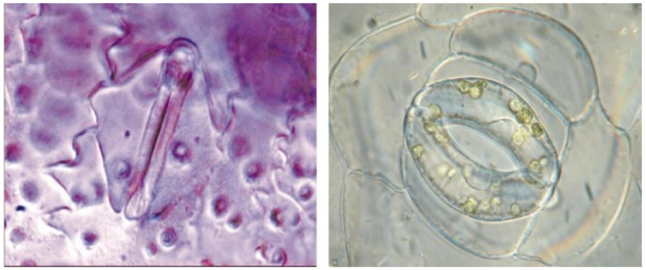
Stomatal conductance from water flux in the DROPS panel (inversing the Penman Monteith equation)



Alvarez Prado S.
Cabrera-Bosquet L.

$$ET_{ref} = \frac{s\times R_n + \rho\times C_p\times g_a\times VPD}{\lambda\times\left(s+\gamma\times\left(1+\frac{g_a}{g_s}\right)\right)}$$





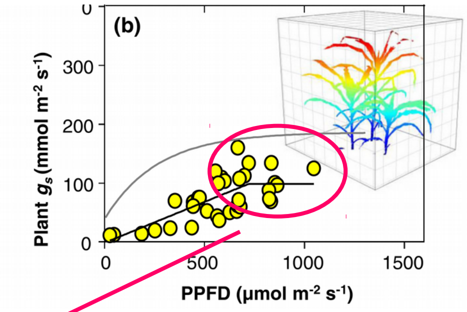
Hetherington & Woodward, 2003

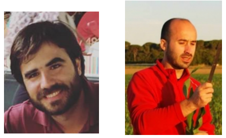
Alvarez Prado et al. 2018 PCE

Genomic region displayed scenario-dependent effect: in platform too

Stomatal conductance from water flux in 250 genotypes (inversing the Penman Monteith equation)

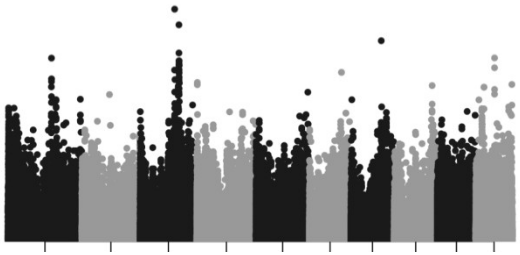
$$ET_{ref} = \frac{s\times R_n + \rho\times C_p\times g_a\times VPD}{\lambda\times\left(s+\gamma\times\left(1+\frac{g_a}{g_s}\right)\right)}$$





Alvarez Prado S.
Cabrera-Bosquet L.

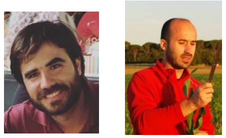
GWAS in each experiments on g_{smax} = 16 QTLs with large QTL×E



Alvarez Prado et al. 2018 PCE

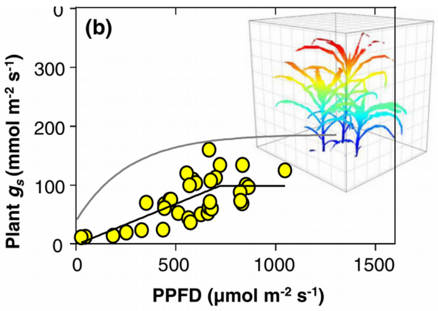
Genomic region displayed scenario-dependent effect: in platform too

Stomatal conductance from water flux in 250 genotypes (inversing the Penman Monteith equation)



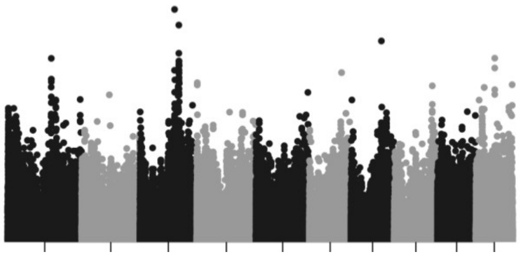
Alvarez Prado S.
Cabrera-Bosquet L.

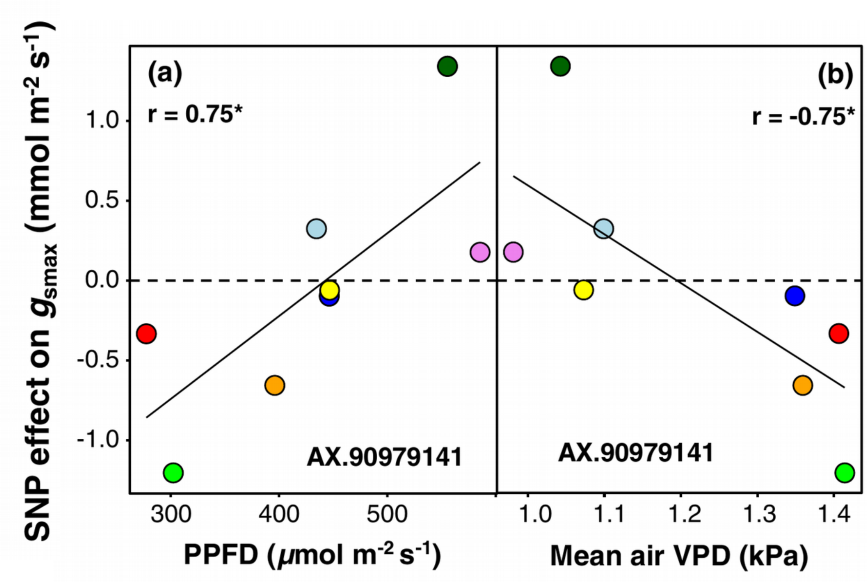
$$ET_{ref} = \frac{s \times R_n + \rho \times C_p \times g_a \times VPD}{\lambda \times \left(s + \gamma \times \left(1 + \frac{g_a}{g_s}\right)\right)}$$



GWAS in each experiment on g_{smax} = 16 QTLs with large QTL×E

Using environmental variables to explain the QTL×E:





Vapour pressure deficit and light with opposite directions:

Genes of ABA transcription factors + two ethylene related proteins

→ involved in the drought-induced processes of stomatal closure

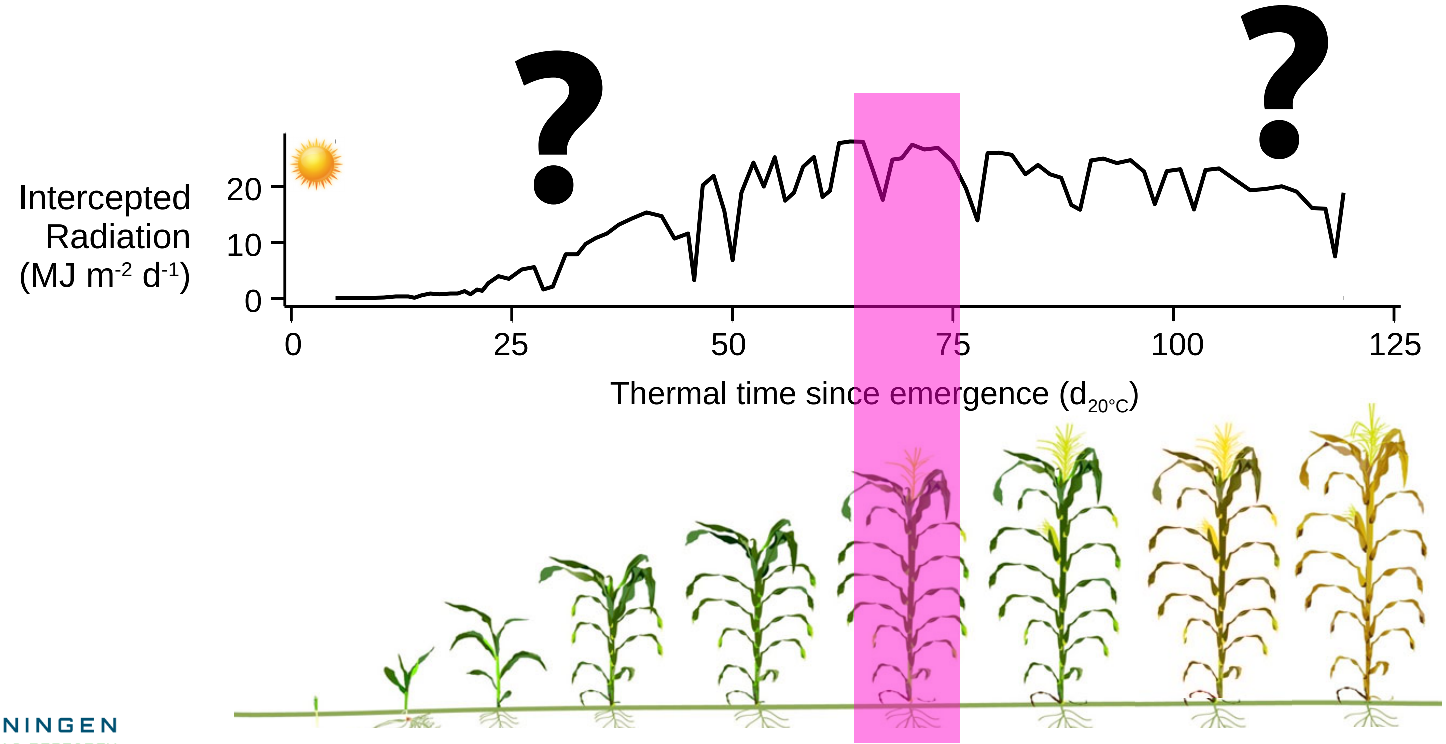
Alvarez Prado et al. 2018 PCE

Dissecting QTL×E: a summary

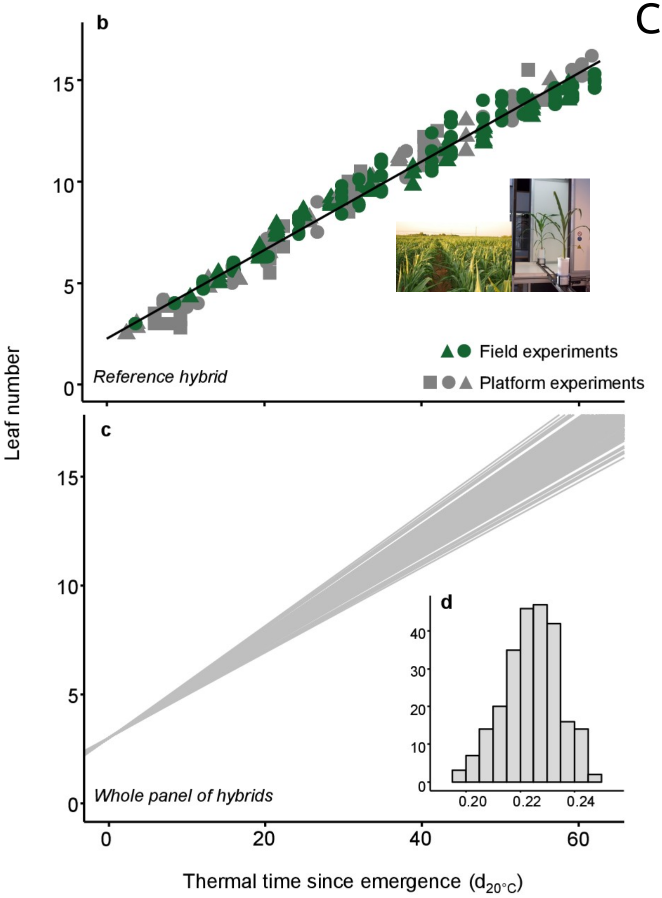
Using sensor network only (no phenomic) we could classify QTLs and identify candidate genes linked with response of GY to change in environment

Also applicable in platforms for secondary traits that can also display G×E (environmental variability also in platform)

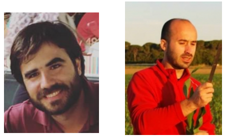
In field, we used variables around flowering time BUT the sensitivity of yield to environmental conditions largely varies with the timing of stressing periods linked to the plant phenology
=> simulation the progression of phenological stages



Dividing the crop cycle into phenological phases

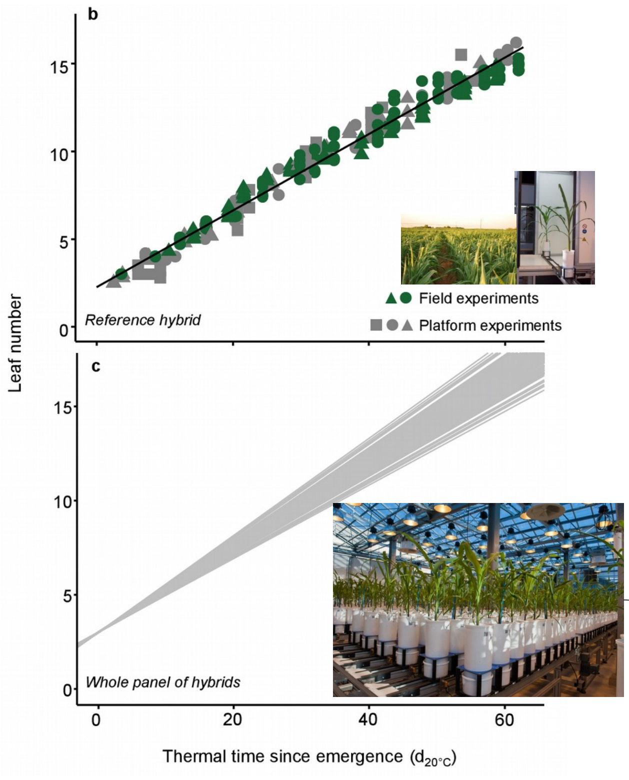


Common progression of stages in field and in platform (Phenoarch)



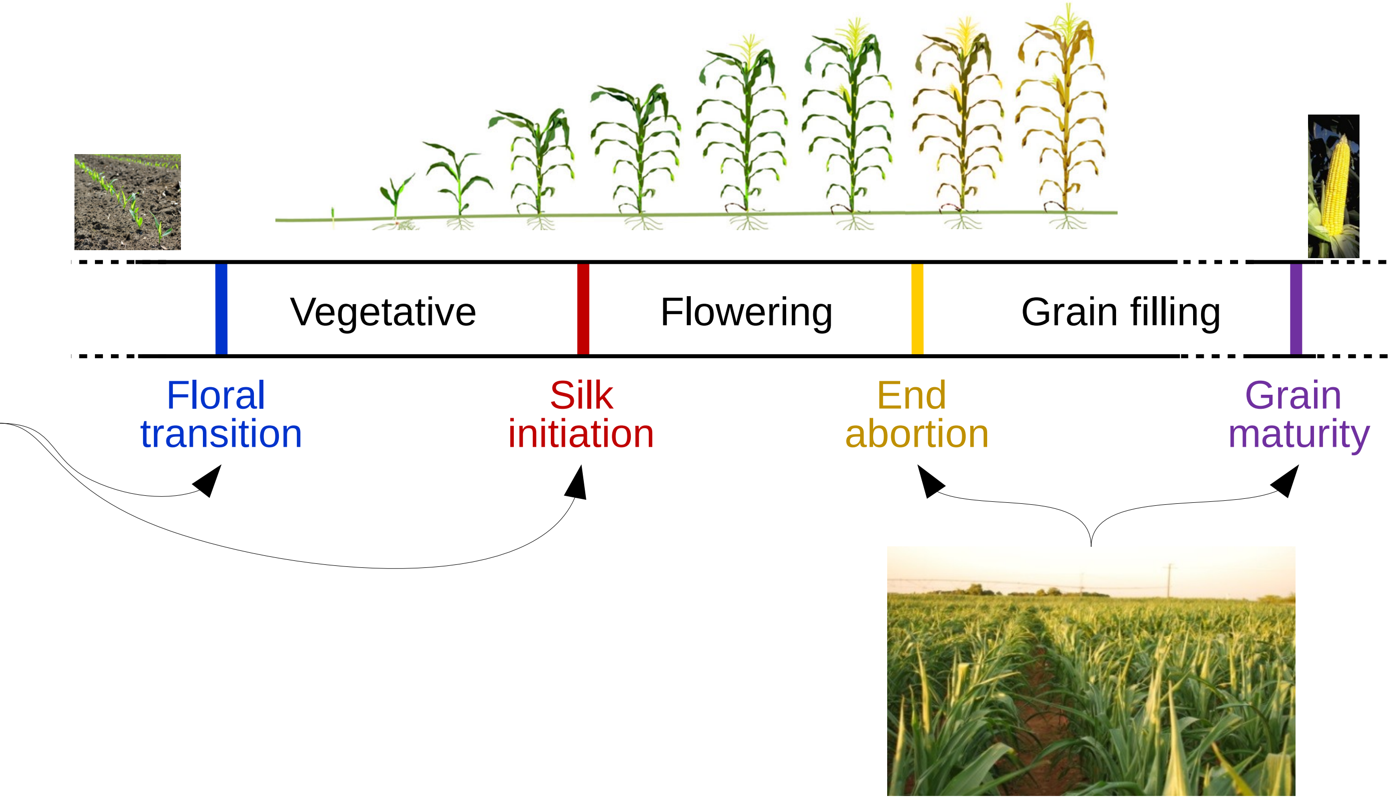
Alvarez Prado S.
Cabrera-Bosquet L.

Dividing the crop cycle into phenological phases



Common progression of stages in field and in platform

Combined with stages measured in field = 3 phases



Vegetative

Flowering

Grain filling

Floral transition

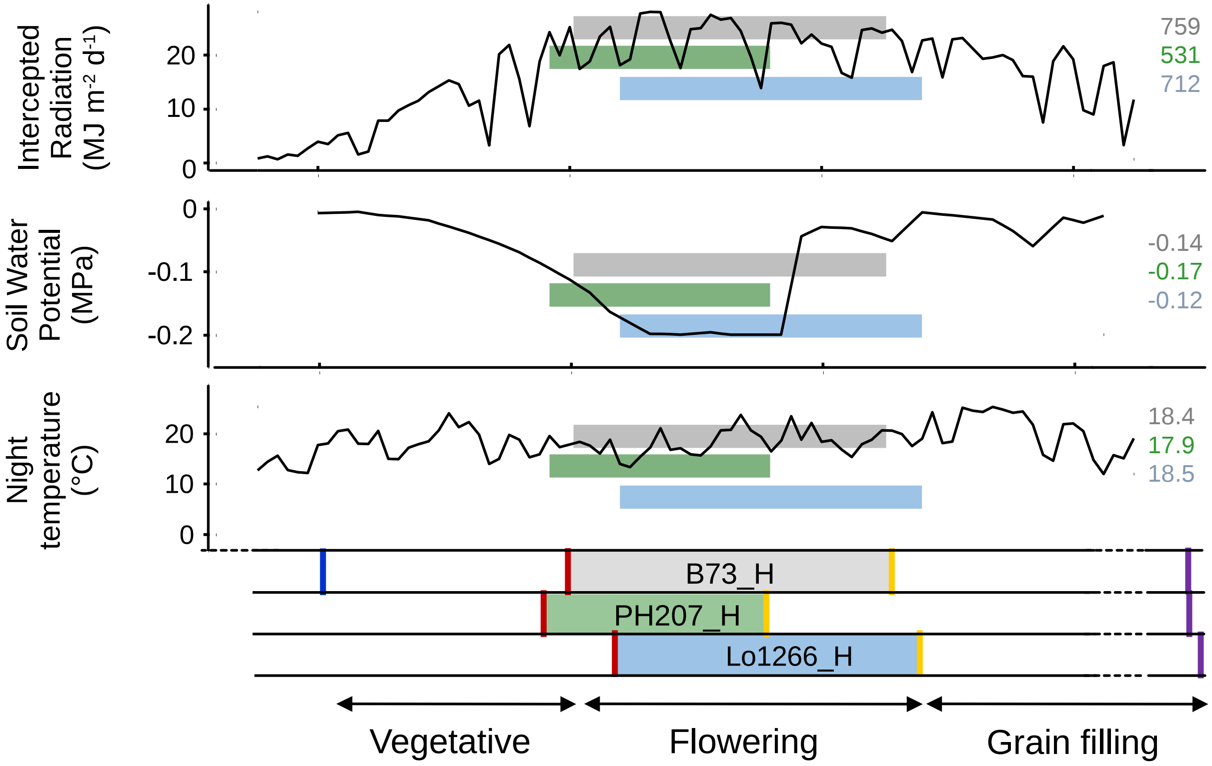
Silk initiation

End abortion

Grain maturity

Different duration and position of the periods due to the differences in hybrids phenology

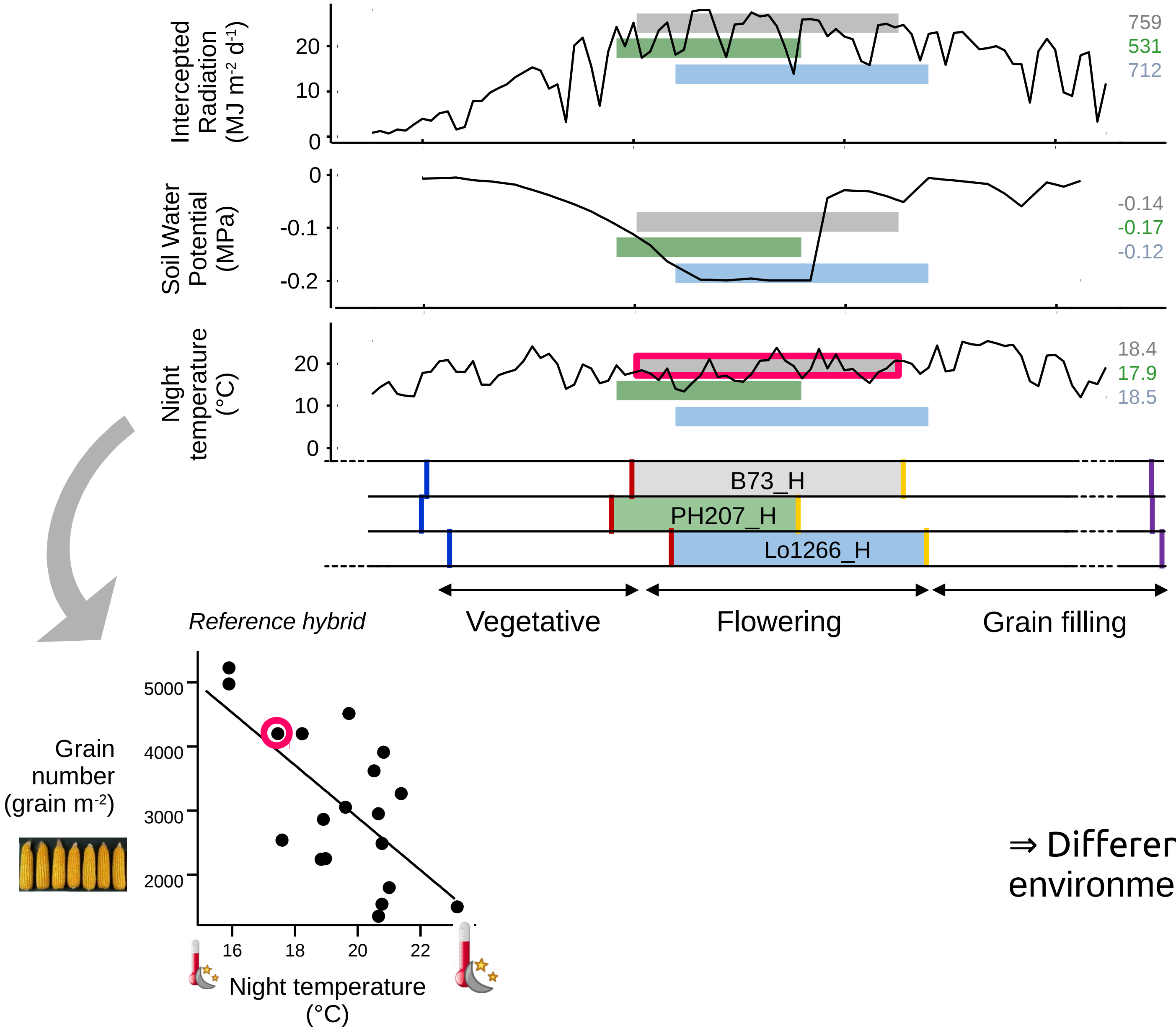
Different phenologies = different conditions sensed by the plants a the same phenological stage



Millet et al. 2019 Nature Genetics

Different duration and position of the periods due to the differences in hybrids phenology

Different phenologies = different conditions sensed by the plants a the same phenological stage

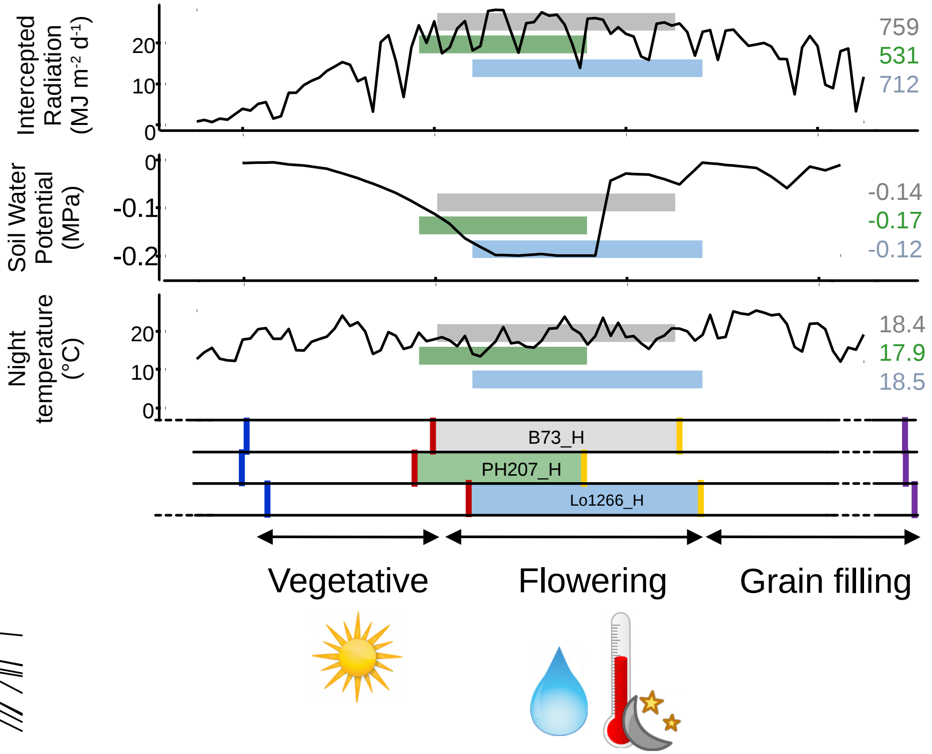


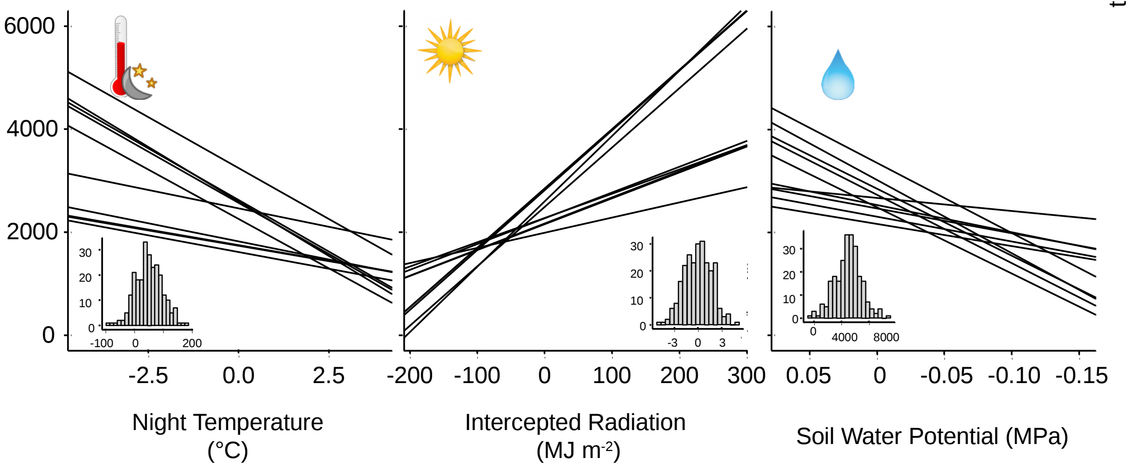
⇒ Different sensitivities to environmental conditions

Millet et al. 2019 Nature Genetics

Dissection of grain number

$$Grain\ Number_{ij} = \mu + e_j + g_i + \sum_{l=1}^{s} \beta_{l,i}\, z_{l,\,j,i}$$





Millet et al. 2019 Nature Genetics

Establishing the model of grain yield dissection

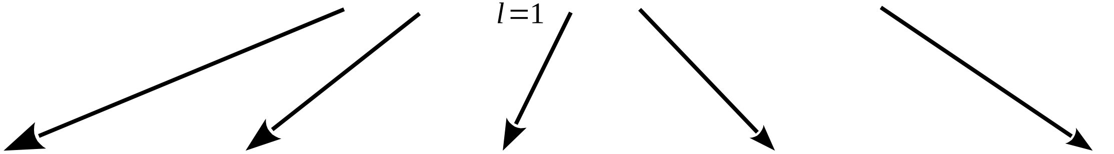
Grain yield dissection into components and responses to environment (few parameters)

$$\mathit{Grain\,Yield}_{ij} = \mathit{Grain\,Number}_{ij} \times \mathit{Grain\,Weight}_{i}$$

$$\mathit{Grain\,Yield}_{ij} = \left(\mu + e_{j} + g_{i} + \sum_{l=1}^{s} \beta_{l,i} z_{l,j,i}\right) \times \mathit{Grain\,Weight}_{i}$$

van Eeuwijk F.

Kruijer W.



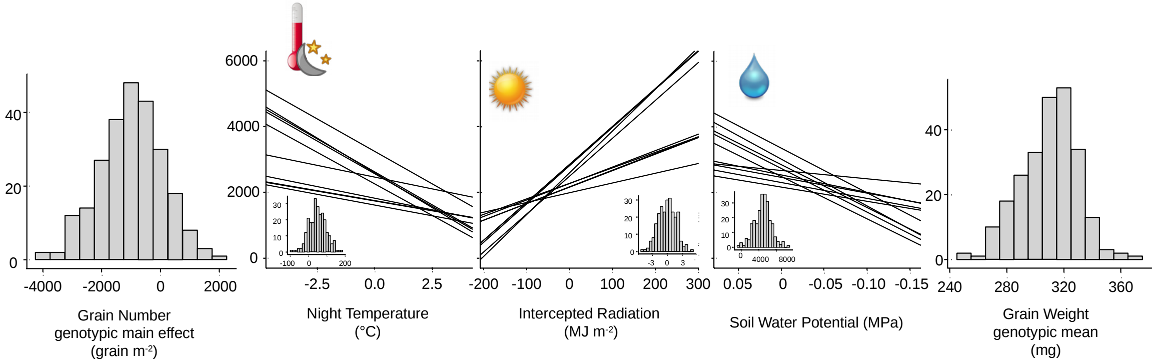
environmental main effect

genotypic main effect

genotypic sensitivities

environmental covariables

genotypic grain weight



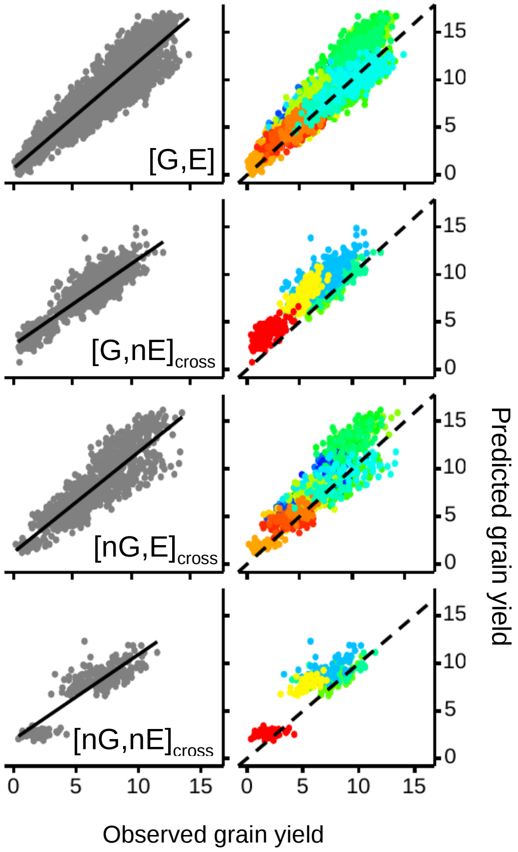
⇒ Can be predicted using marker profile and/or environmental indices

Whole genome regression with 50K SNPs and BayesR (Moser et al. 2015)

Prediction of grain yield in new experiments for new hybrids

Cross-validation and test of the model by sampling hybrids and experiments:
80% training – 20% test

	G (200)	nG (46)
E (20)	**Training** Estimation of model parameters	**Test set** Prediction of genotypic parameters using marker profile
nE (5)	**Test set** Prediction of main environmental effect using sensor data	**Test set** Prediction of genotypic and environmental parameters using marker profile and sensor data



Correlation coef per env:

0.60 – 0.91

0.43 –0 .85

0.21 – 0.71

0.20 – 0.74

⇒ A simple model allows a good estimation and prediction of grain yield

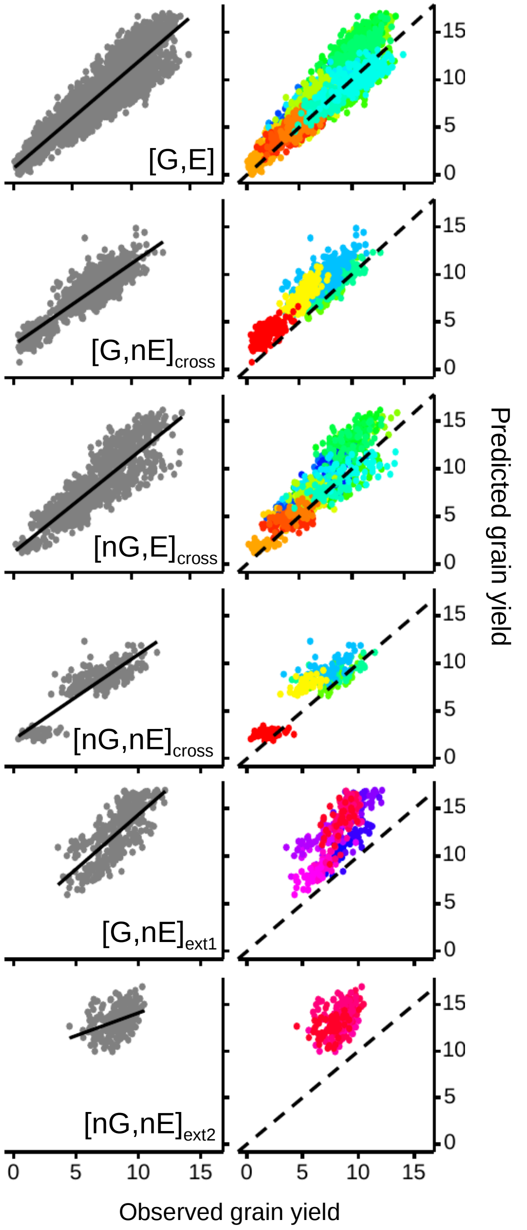
Millet et al. 2019 Nature Genetics

Prediction of grain yield in new experiments for new hybrids

Cross-validation and test of the model by sampling hybrids and experiments:
80% training – 20% test

	G (200)	nG (46)	nG_{ext} (56)
E (20)	**Training** Estimation of model parameters	**Test set** Prediction of genotypic parameters using marker profile	NA
nE (5)	**Test set** Prediction of main environmental effect using sensor data	**Test set** Prediction of genotypic and environmental parameters using marker profile and sensor data	NA
nE_{ext} (12)	**Test set** (subset of 32 Geno) Prediction of main environmental effect using sensor data	NA	**Test set** (subset of 4 Exp) Prediction of genotypic and environmental parameters using marker profile and sensor data

🟦 Cross-validation sets 🟧 External validation sets



Correlation coef per env:

0.60 – 0.91

0.43 –0 .85

0.21 – 0.71

0.20 – 0.74

0.38 – 0.80

0.32 – 0.39

⇒ A simple model allows a good estimation and prediction of grain yield

Millet et al. 2019 Nature Genetics

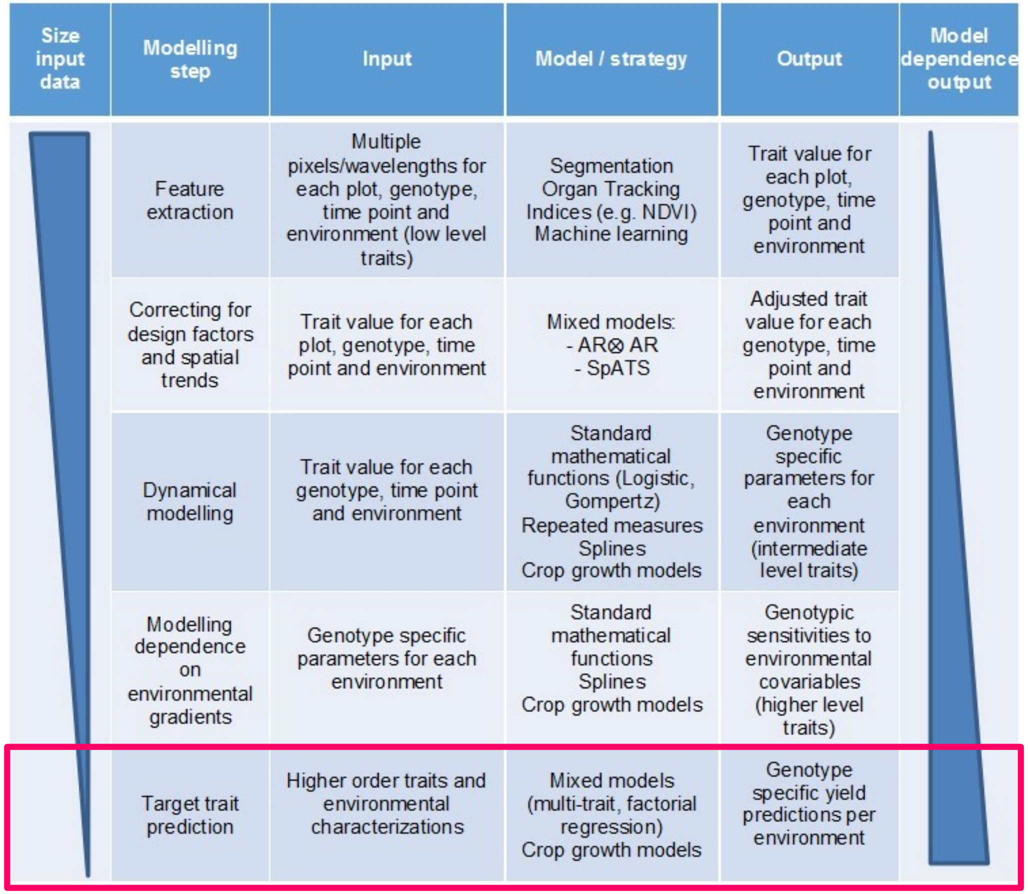
Dissecting G×E by combining field and platform: a summary

Using phenomic **AND** sensor network (enviromic?) we could dissect the G×E in the field by using precise phenology of each hybrid

Despite a period of sensitivity at flowering time, early period also matters

Simple G2P model allows good prediction accuracies

What if we use other variable available at the plant level in the platform to predict the field trait?



Size input data	Modelling step	Input	Model / strategy	Output	Model dependence output
	Feature extraction	Multiple pixels/wavelengths for each plot, genotype, time point and environment (low level traits)	Segmentation Organ Tracking Indices (e.g. NDVI) Machine learning	Trait value for each plot, genotype, time point and environment	
	Correcting for design factors and spatial trends	Trait value for each plot, genotype, time point and environment	Mixed models: - AR⊗AR - SpATS	Adjusted trait value for each genotype, time point and environment	
	Dynamical modelling	Trait value for each genotype, time point and environment	Standard mathematical functions (Logistic, Gompertz) Repeated measures Splines Crop growth models	Genotype specific parameters for each environment (intermediate level traits)	
	Modelling dependence on environmental gradients	Genotype specific parameters for each environment	Standard mathematical functions Splines Crop growth models	Genotypic sensitivities to environmental covariables (higher level traits)	
	Target trait prediction	Higher order traits and environmental characterizations	Mixed models (multi-trait, factorial regression) Crop growth models	Genotype specific yield predictions per environment	

van Eeuwijk et al. 2019 Plant Science

Prediction of grain yield in field using metabolites in platform

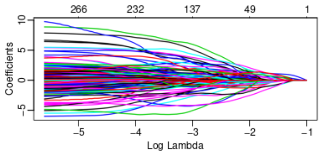
1 415 metabolic signatures + 11 identified metabolites
young plants in Phenoarch platform

Prigent S.

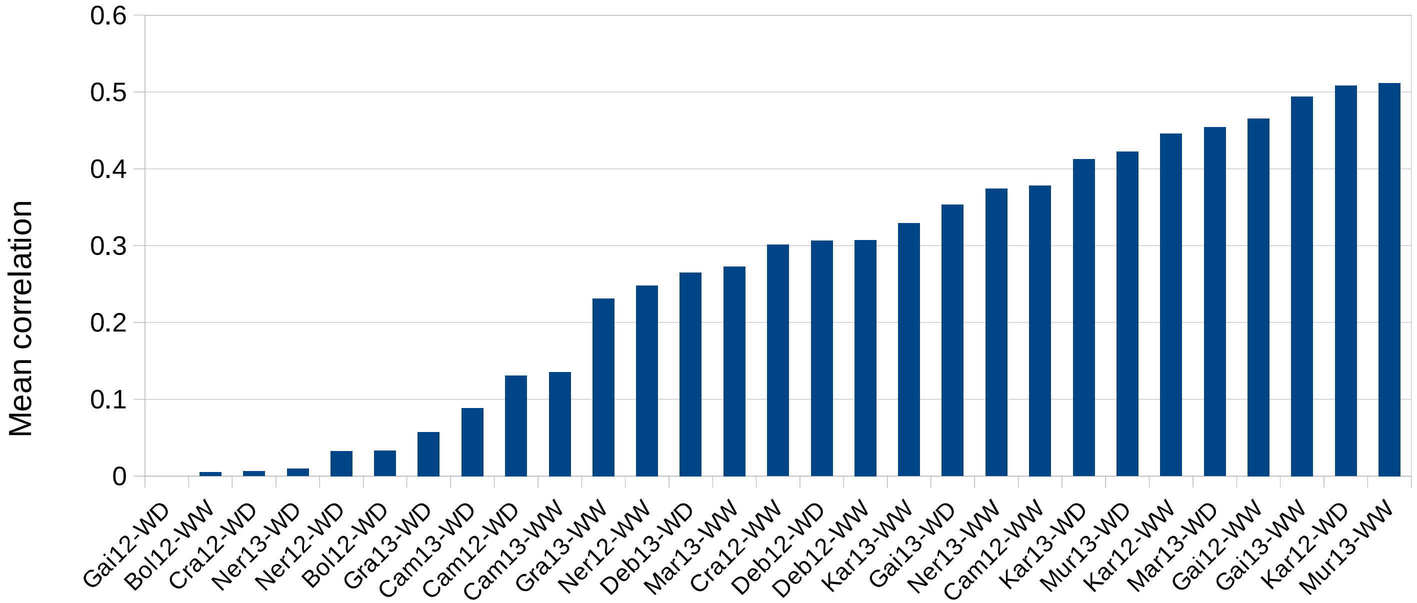
Gibon Y.

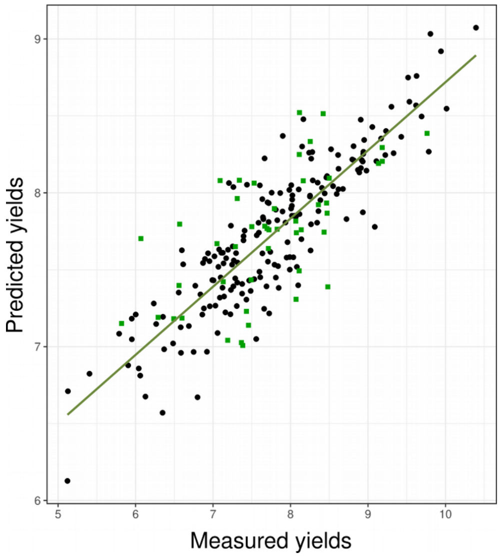
Creation of a linear model to link metabolic variables and yields: Generalized multi-linear models
GY = a_1 metabo$_1$ + … + a_n metabo$_n$

Elastic-net regression



With cross-validation: using 20% of the genotypes to test the model, repeat 50 times
Mean prediction accuracy per env:

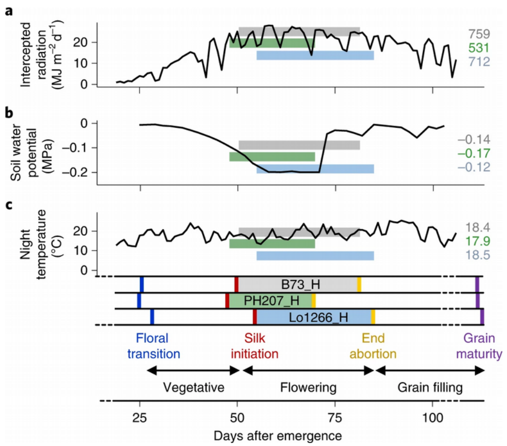




Combining metabolites in platform AND environmental covariates

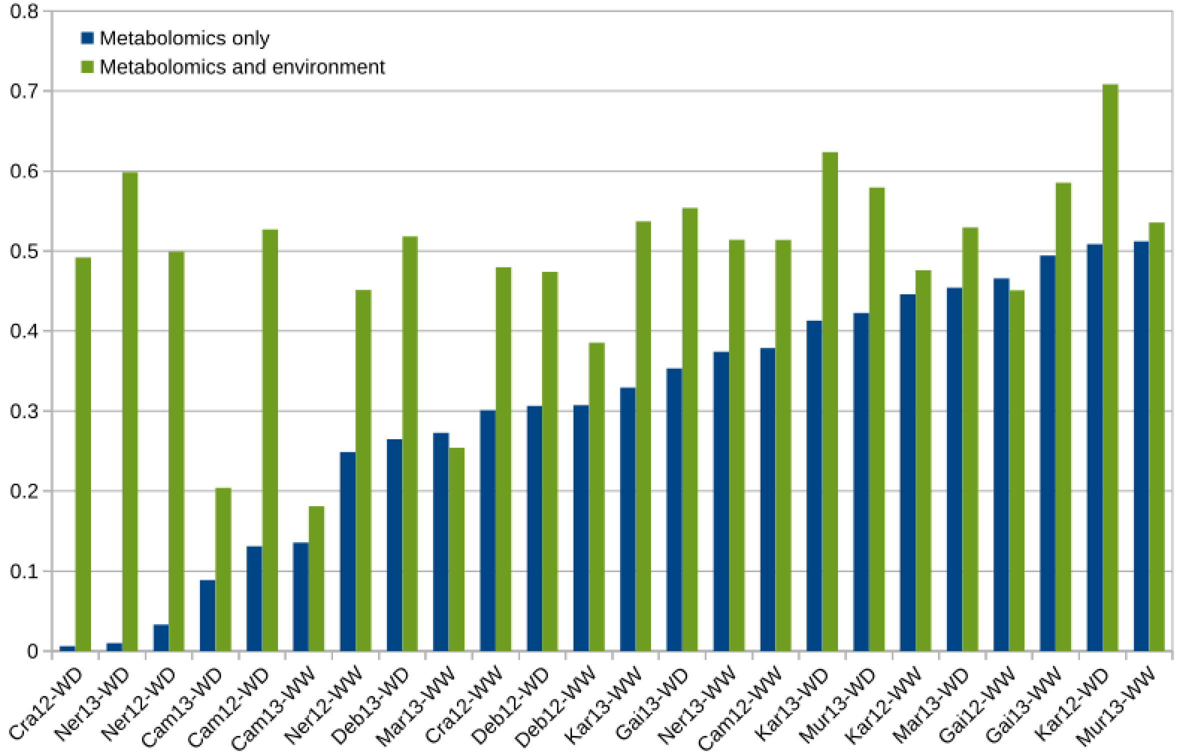
Using field environmental covariates: 3 periods x 4 variables = 12 covariables:

$GY = a_1 \text{ metabo}_1 + \ldots + a_n \text{ metabo}_n + b_1 \text{ cov}_1 + \ldots + b_l \text{ cov}_l$



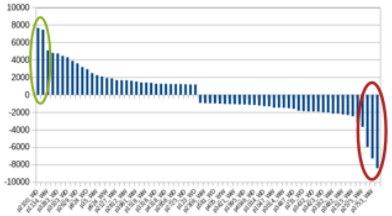
With cross-validation: using 20% of the genotypes to test the model, 50 times

Mean prediction accuracy per env:

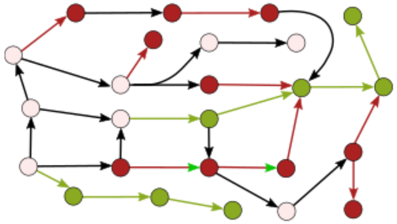


To do next:

Identification of biomarkers



Mapping biomarkers on metabolic maps



Summary

Original approach combining multiple sources of data and modelling methods

1. Using sensor network only we could classify QTLs and identify candidate genes (in field <u>and</u> in platform)

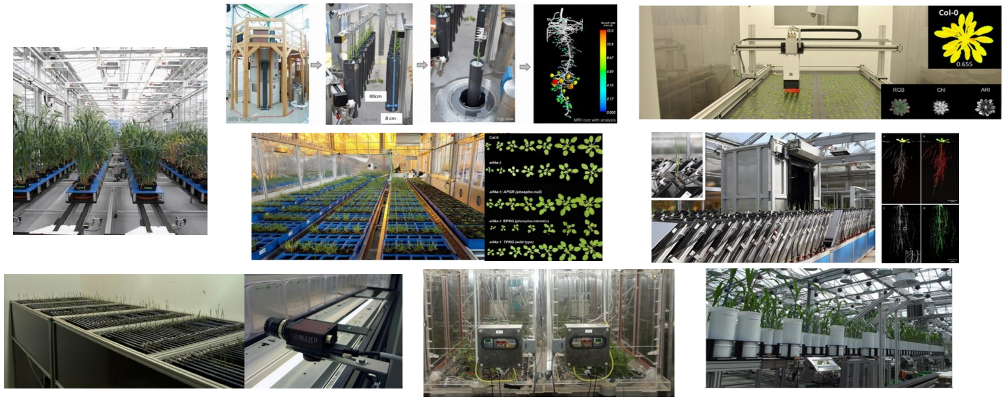
2. Combining traits measured in phenotyping platform and environmental characterization we could better dissect the G×E and predict the grain yield

Perspectives

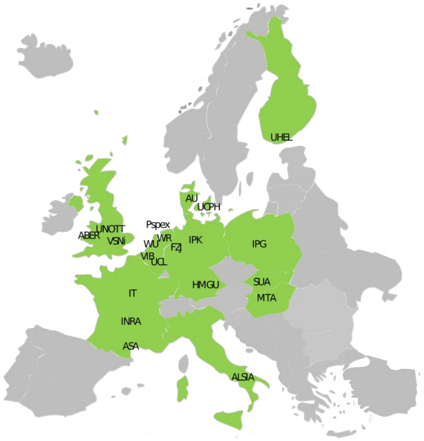
- breeding programs in regions with contrasted environmental conditions
- relevant for climate change (night temp and soil water)

Ongoing work

- Validation of the genomic regions: NILs under test (field experiments, C. Welcker)
- Root traits in platform (coll. X. Draye, UCLouvain)



European
Plant Phenotyping
Network
EPPN 2020



2017-2020
Access to 31 installations in Europe

https://eppn2020.plant-phenotyping.eu/

- Based on a simple application procedure
- Calls every 6 months (**7th call open in autumn 2019**)
- Full cost of projects covered by the project, including travel
- 20% access funds for non-European labs

WAGENINGEN
UNIVERSITY & RESEARCH

Acknowledgements

WAGENINGEN
UNIVERSITY & RESEARCH

F van Eeuwijk
W Kruijer

lepse
Montpellier

F Tardieu
C Welcker
A Coupel-Ledru

L Cabrera
Bosquet
S Alvarez Prado
S Lacube

LE MOULON
Génétique végétale UMR

A Charcosset
S Nicolas
S Negro
C Bauland

Thank you for your attention !

biogemma

S Praud
P Dubreuil
MH Tixier
C Richard
J Lopez
JL Olivier
G Lenne

Max-Planck-Institut
für Molekulare Pflanzenphysiologie

B Usadel
A Nagel

KWS

T Welz
T Presterl
E Lagzi
T Rettig

R Tuberosa
MA Canè

MTA•ATK

Z Bedo
T Spitkó
N Harnos

UCL
Université catholique de Louvain

X Draye

syngenta

N Ranc
V Ribière
V Schlegel
F Racz
Z Nagy
F Tanzi

European Plant Phenotyping Network
EPPN 2020

DROPS

amaizing