

# Characterization of adaptation mechanisms in sorghum using a multi-reference back-cross nested association mapping design and envirotyping

Vincent Garin, Chiaka Diallo, Mohamed Lamine Tekete, Korotimi Thera, Baptiste Guitton, Karim Dagno, Abdoulaye G. Diallo, Mamoutou Kouressy, Willmar Leiser, Fred Rattunde, Ibrahima Sissoko, Aboubacar Toure, Baloua Nebie, Moussa Samake, Jana Kholova, Julien Frouin, David Pot, Michel Vaksman, Eva Weltzien, Niaba Teme, and Jean-Francois Rami



**Sorghum in the 21st Century**  
**Global Sorghum Conference**  
Resiliency and Sustainability in the Face of Climate Change

June 5-9 The Corum Event Center,  
**2023** Montpellier, France

# Consortium

---

## **IER**

Niaba Teme, Mohamed Lamine Tekete, Korotimi Thera, Karim Dagno, Abdoulaye G. Diallo, Mamoutou Kouressy

## **CIRAD**

Baptiste Guitton, David Pot, Julien Frouin, Michel Vaksman, Jean-Francois Rami

## **ICRISAT (Mali)**

Eva Weltzien (Wisconsin University), Chiaka Diallo (Institut polytechnique rural), Willmar Leiser (KWS), Fred Rattunde (Wisconsin University), Ibrahima Sissoko, Aboubacar Toure, Baloua Nebie (CIMMYT)

## **ICRISAT (Inde)**

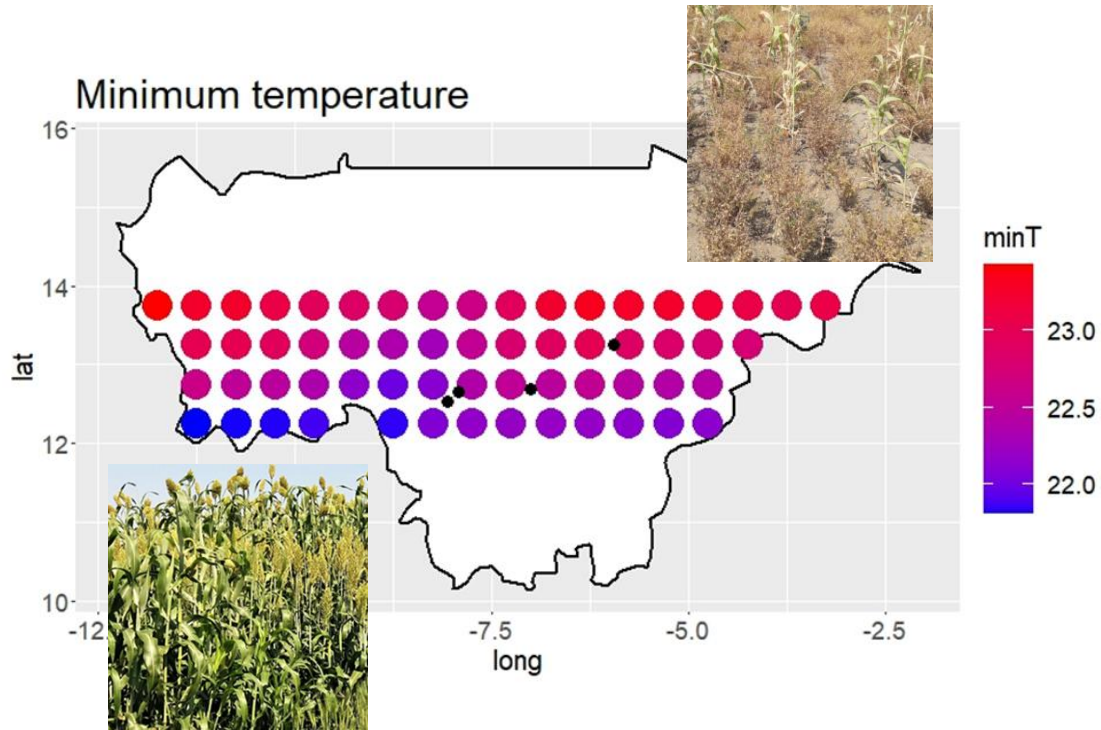
Jana Kholova

## **USTTB**

Moussa Samake



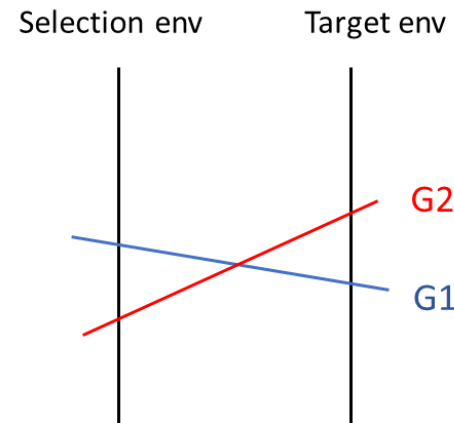
# The challenge of the genotype by environment (GxE) interaction



Plants, especially sorghum are experiencing strong GxE interaction.

Sorghum shows a wide range of adaptation (plasticity).

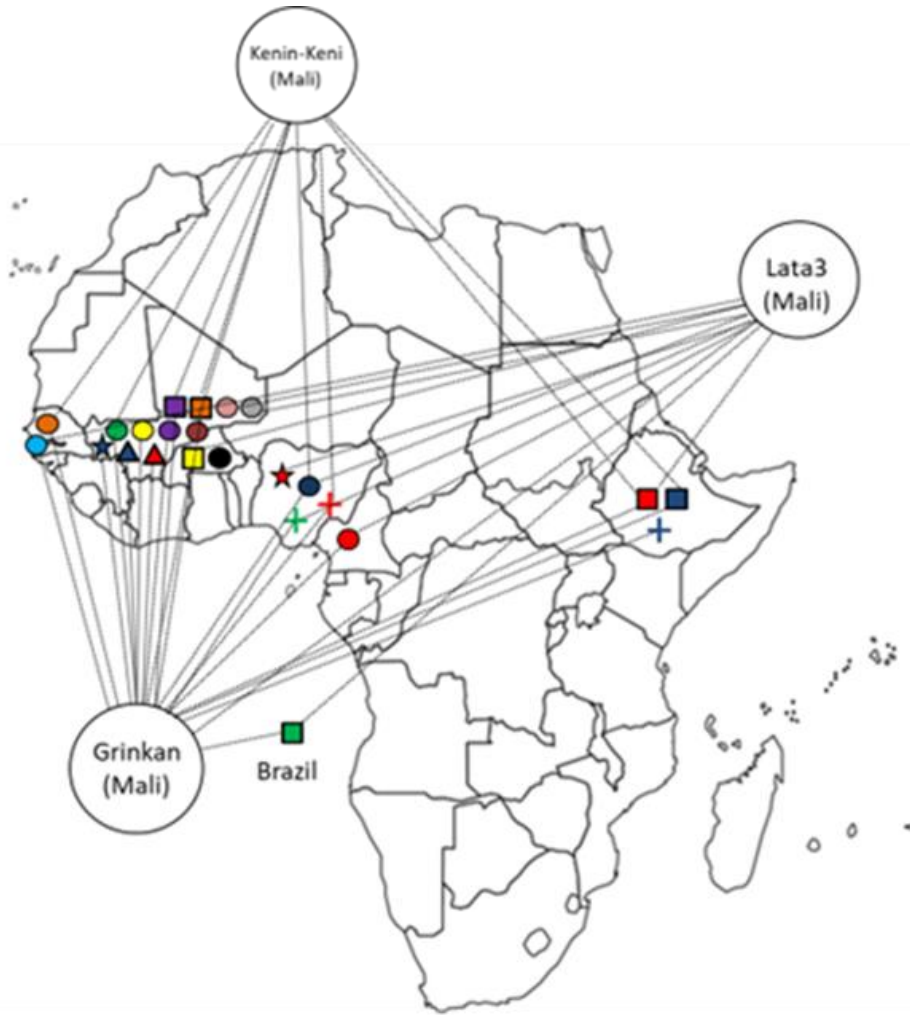
GxE is a challenge for breeders, especially in the context of marker assisted selection



**How to meet the demand of farmers that experience extremely diverse cultivation conditions?**

**-> Advanced genetic resources + Envirotyping + statistical methodology + decision support tools**

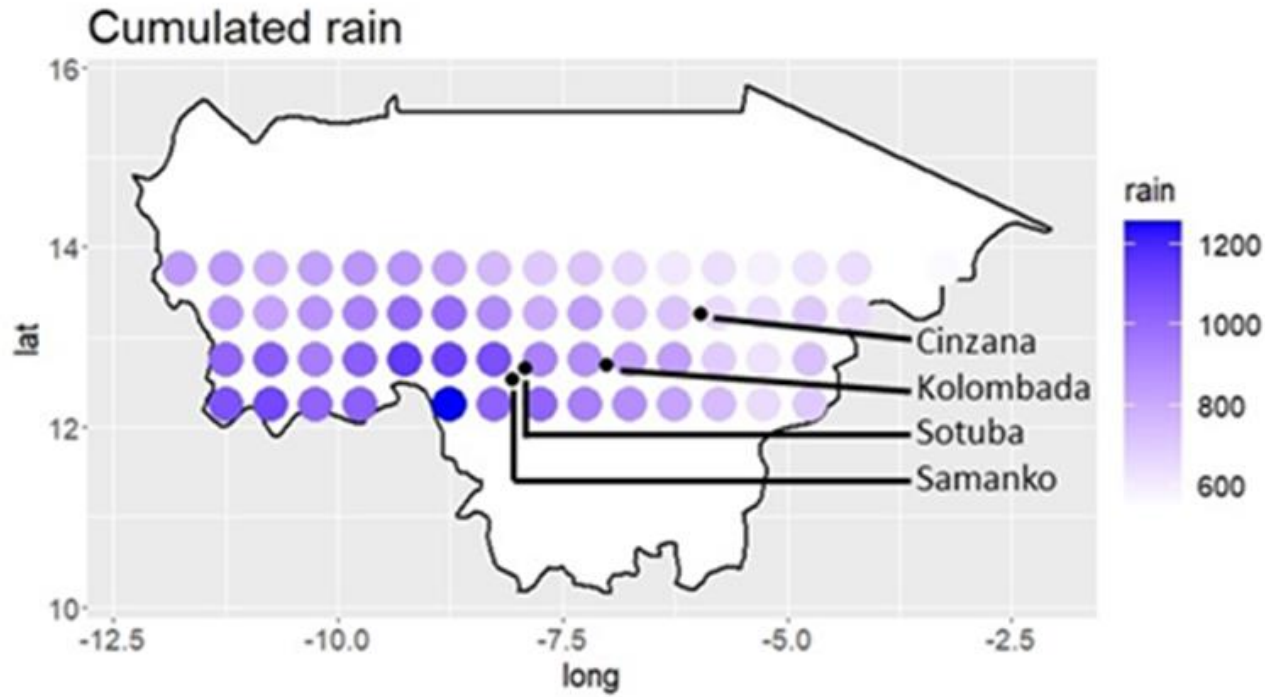
# Advanced genetic resources: West and Central Africa (WCA) BCNAM population



	Parent	PH	Mat	Kp3	Specific advantage	Reference	
Guinea	Fara-Fara (Nigeria)	Grinkan	av	av	av	Elite line	Guitton et al. (2018)
	IS15401 (Cameroun)	Kenin-Keni	av	av	av	Elite Line	Leroy et al. (2014)
	CSM417 (Mali)	Lata3	+	av	av	Elite Line	Diallo et al. (2019)
	CSM63 (Mali)	Fara-Fara	+	+	+	Diversity	Andrews (1973)
	CSM388 (Mali)	E36-1	-	av	-	Drought tolerance	Mahalakshmi et al. (2002)
	BimbG (Guinea)	IS15401	+	+	+	Striga resistance	FAO (2008)
	Hafijeka (Gambia)	IS23540	-	av	-	Sweet stem	FAO (2023)
	Sangatigui (Mali)	B35	-	-	-	Drought tolerance	Rama Reddy et al. (2014)
	DouaG (Mali)	Konotene	+	+	-	Grain weight	Clément et al. (1980)
	N'Golofing (Mali)	SC566-14	-	-	-	AI tolerance	Magalhaes et al. (2004)
	Gnossiconi (Burkina Faso)	Framida	-	+	-	Striga resistance	Hausmann et al. (2001)
	Caudatum	CSM417	+	+	+	Grain quality	Clément et al. (1980)
E-361 (Ethiopia)		CSM63	av	-	-	Precocity	Chantereau et al. (1998)
IS23540 (Ethiopia)		CSM388	+	+	+	Grain quality	Folliard et al. (2004)
SC566-14 (Brazil)		Gadiaba Dié	+	+	+	Grain weight	Clément et al. (1980)
Framida (Burkina Faso)		W. Kaura	-	+	+	Diversity	Goma et al. (2012)
Kalaban (Mali)		V33/08	av	+	av	Grain quality	Soumaré et al. (2008)
Malisor (Mali)		Kalaban	-	av	-	Productivity	FAO (2008)
Konotene (Mali)		Malisor 84-7	-	-	+	Head bug resistance	Ratnadass et al. (2002)
Gadiaba (Mali)		BimbG	+	+	+	Grain quality	Sagnard et al. (2013)
B35 (Ethiopia)		Hafijeka	+	+	+	Grain quality	Folkertsma et al. (2005)
Du/Caud Durra	S. Kaura	+	+	+	Diversity	Kassam et al. (1975)	
	White Kaura (Nigeria)	Sangatigui	av	av	av	Diversity	CEDEAO-UEMOA-CILSS (2016)
	Short Kaura (Nigeria)	DouaG	+	+	+	Low-P adaptation	Kante et al. (2017)
	V33/08 (Mali)	Gnossiconi	av	av	av	Yield stability	vom Brocke et al. (2014)
		Ngolofing	+	av	+	Grain quality	Clément et al. (1980)
	Sambalma (Nigeria)	Sambalma	+	+	+	AI tolerance	Kante et al. (2019)

WCA-BCNAM populations: 41 crosses between central parent Grinkan, Kenin-Keni, and Lata3 (N=3901) and 24 donors parents with a wide diversity of traits.

# Envirotyping: Multi-environment phenotyping



Multi-environments characteristics:

Four testing sites: Sotuba, Cinzana, Samanko, Kolombada

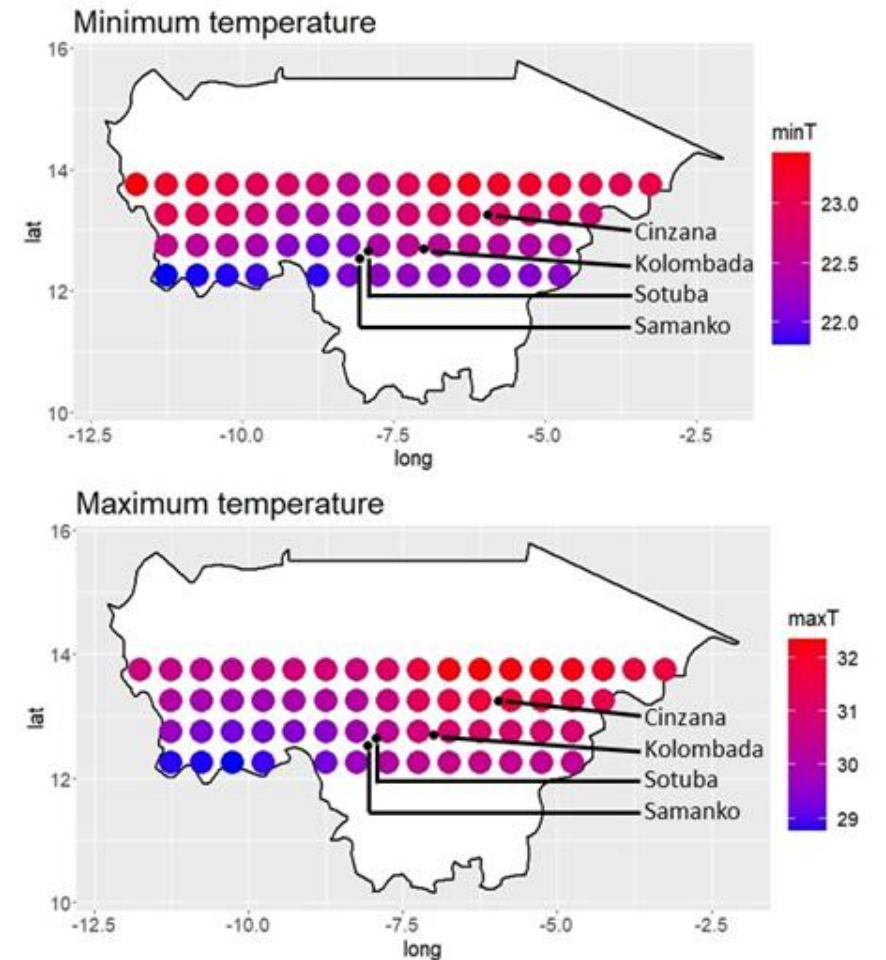
Two sowing dates (Sot, Cin): End June, End July

Location	Cumulated rain season [mm]	Temperature [°C]	Photoperiod (June) [h]	Photoperiod (July) [h]
Sotuba	570 – 851	27	12.3	12
Cinzana	338 – 708	28.5	12.15	11.95
Samanko	967	27.5	12.3	
Kolombada	438	26.7	12.3	

Trait: Flag leaf, plant height, number of internodes, average internode length, peduncle length, panicle length, yield

# Envirotyping: 15 environmental covariates (EC)

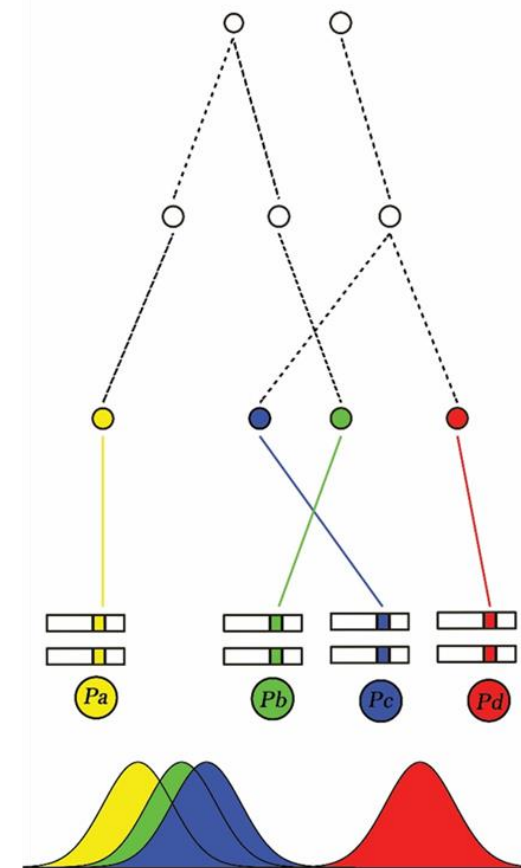
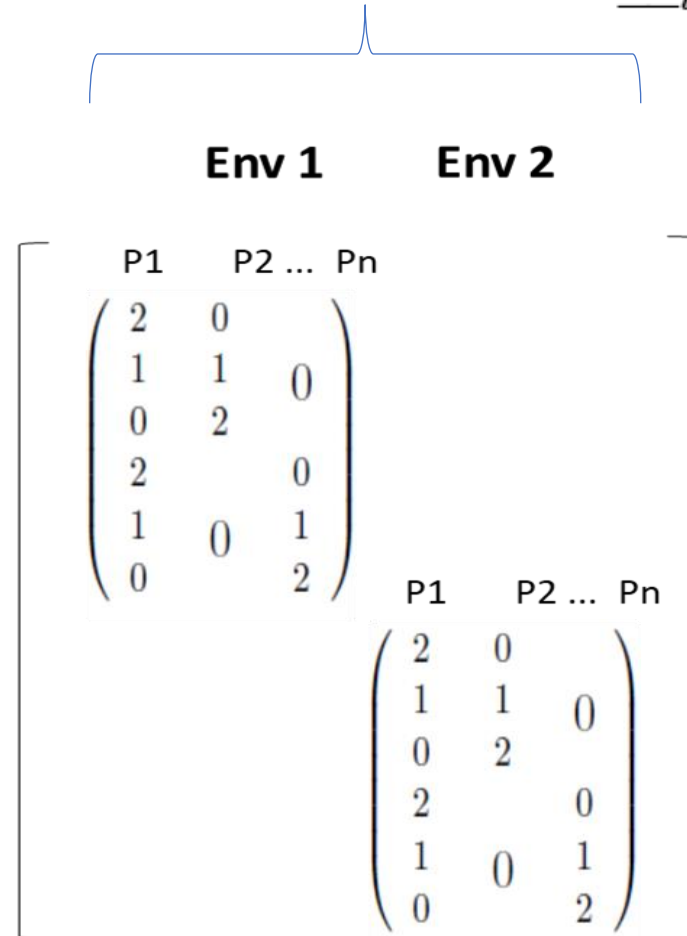
Category	EC	Unit	Observed/ inferred	Sum/mean
Atmospheric	cumulated rain	mm	obs	sum
	humidity	%	obs	mean
	vapour pressure deficit	kPa	inf	mean
	slope of saturation VP curve	kPa/d	inf	mean
	potential evapotranspiration	mm/day	inf	mean
	deficit by precipitation	mm/day	inf	mean
Temperature	minimum temperature	d	obs	mean
	maximum temperature	d	obs	mean
	temperature range	d	obs	mean
	cumulated degree day	dd	obs	sum
	T effect on radiation use efficiency	0-1	inf	mean
Radiation	cumulated hour of sun	h	obs	sum
	photoperiod	h	inf	mean
	solar radiation	MJ/m <sup>2</sup> /day	inf	sum
Photothermal	Photothermal (photoperiod * DD)	h*dd	obs	sum



EnvRtype: Costa-Neto et al. (2020)  
Nasapower: Sparks (2018)

# Statistical methodology: MPP-ME QTL detection model

$$\underline{y}_{icj} = \mu + e_j + c_{cj} + \underline{x}_{ia} * \beta_{aj} + \underline{ge}_{icj} + \underline{\epsilon}_{icj} \quad [1]$$



We assumed that each parent carries a different allele at the QTL position that can have a different effect in each environment (N\_par x N\_Env effect) estimated compared to reference (central parent).

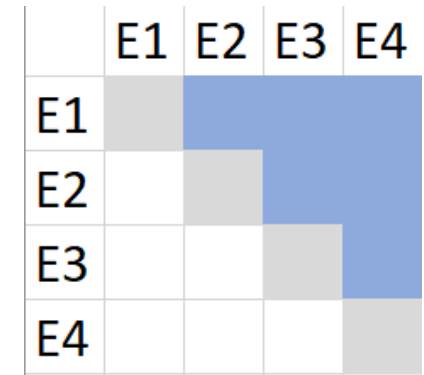
# Statistical methodology: Variance covariance structure

$$\underline{y}_{icj} = \mu + e_j + c_{cj} + x_{ia} * \beta_{aj} + \underline{ge}_{icj} + \underline{\epsilon}_{icj} \quad [1]$$

$$V(\underline{ge}_{icj} + \underline{\epsilon}_{icj})$$

Compound symmetry (CS)

$$V \begin{bmatrix} y_{i11} \\ y_{i'21} \\ y_{i12} \\ y_{i'22} \\ y_{i13} \\ y_{i'23} \end{bmatrix} = \begin{bmatrix} \sigma_g^2 + \sigma_\epsilon^2 & 0 & \sigma_g^2 & 0 & \sigma_g^2 & 0 \\ 0 & \sigma_g^2 + \sigma_\epsilon^2 & 0 & \sigma_g^2 & 0 & \sigma_g^2 \\ \sigma_g^2 & 0 & \sigma_g^2 + \sigma_\epsilon^2 & 0 & \sigma_g^2 & 0 \\ 0 & \sigma_g^2 & 0 & \sigma_g^2 + \sigma_\epsilon^2 & 0 & \sigma_g^2 \\ \sigma_g^2 & 0 & \sigma_g^2 & 0 & \sigma_g^2 + \sigma_\epsilon^2 & 0 \\ 0 & \sigma_g^2 & 0 & \sigma_g^2 & 0 & \sigma_g^2 + \sigma_\epsilon^2 \end{bmatrix}$$

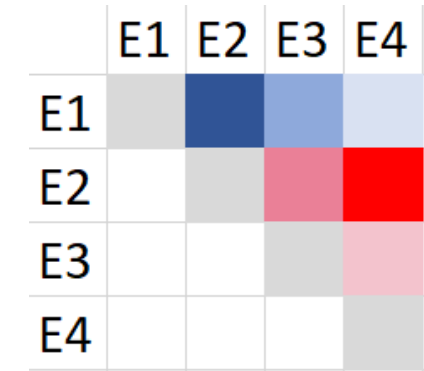


$\sigma_g^2$

(uniform gene action)

Unstructured (UN)

$$V \begin{bmatrix} y_{i11} \\ y_{i'21} \\ y_{i12} \\ y_{i'22} \\ y_{i13} \\ y_{i'23} \end{bmatrix} = \begin{bmatrix} \sigma_{g_1}^2 + \sigma_\epsilon^2 & 0 & \sigma_{g_{12}} & 0 & \sigma_{g_{13}} & 0 \\ 0 & \sigma_{g_1}^2 + \sigma_\epsilon^2 & 0 & \sigma_{g_{12}} & 0 & \sigma_{g_{13}} \\ \sigma_{g_{12}} & 0 & \sigma_{g_2}^2 + \sigma_\epsilon^2 & 0 & \sigma_{g_{23}} & 0 \\ 0 & \sigma_{g_{12}} & 0 & \sigma_{g_2}^2 + \sigma_\epsilon^2 & 0 & \sigma_{g_{23}} \\ \sigma_{g_{13}} & 0 & \sigma_{g_{23}} & 0 & \sigma_{g_3}^2 + \sigma_\epsilon^2 & 0 \\ 0 & \sigma_{g_{13}} & 0 & \sigma_{g_{23}} & 0 & \sigma_{g_3}^2 + \sigma_\epsilon^2 \end{bmatrix}$$



$\sigma_{g_{jj'}}^2$

(Set of gene act differently in each E)



# Statistical methodology: QTLxEC model

$$\underline{y_{icj}} = env_j + cross_{cj} + \underbrace{x_{ip} * \beta_{pj}}_{\text{QTL term}} + \underline{GE_{icj}} + \underline{e_{icj}} \quad (2)$$

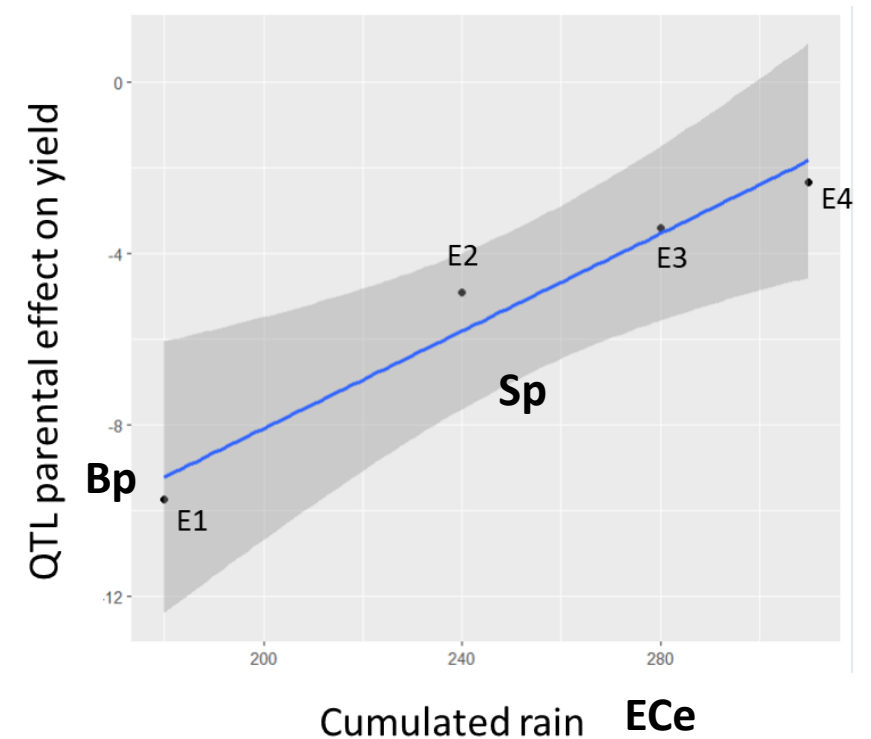
QTL term

$$x_{ip} * (\beta_p + EC_e * S_p + l_{pe})$$

**B<sub>p</sub>**: parental allele main effect across environment

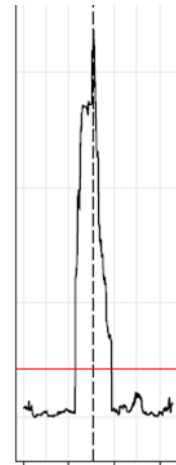
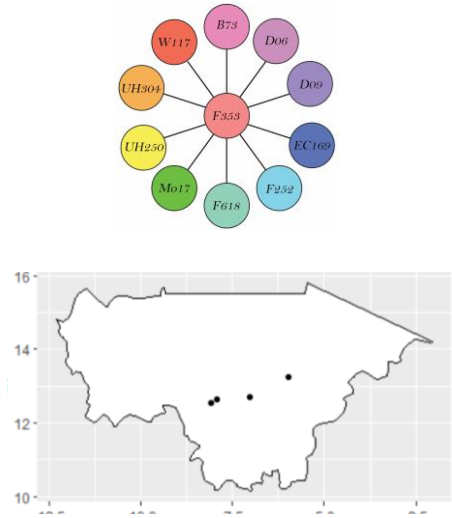
**EC<sub>e</sub>**: value of the environmental covariate in environment E

**S<sub>p</sub>**: Sensitivity of the parental QTL allelic effect to the EC



# Statistical methodology– QTL allele effect projection in the Malian environment

## a) Estimation of the QTL allele sensitivity to the EC

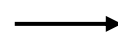


$$Q_{eff} = \beta_p + EC_e * S_p$$

$$\hat{Q}_{eff} = 250 + EC_e * 0.05$$

## b) Projection of the QTL effect given new EC values

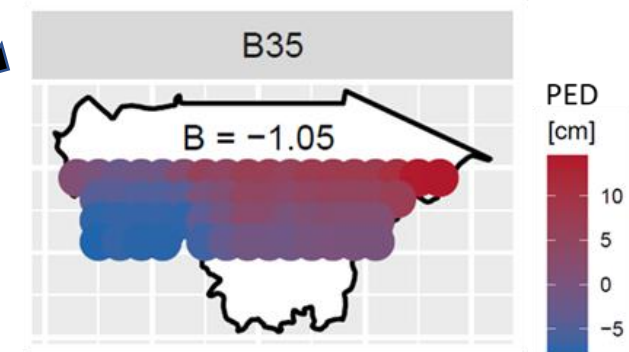
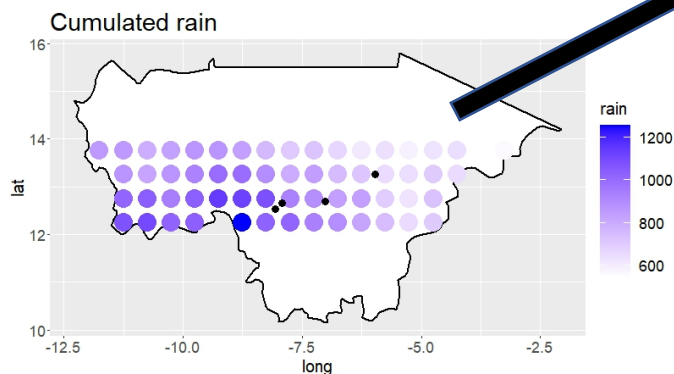
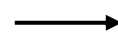
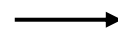
$$\hat{Q}_{eff}(lat,lon) = 250 + EC_e * 0.05$$



$$\hat{Q}_{eff}(12.5,-5)$$

[...]

$$\hat{Q}_{eff}(13.5 -10)$$



# Statistical methodology – mppR R package

---

## mppR: Multi-Parent Population QTL Analysis

Analysis of experimental multi-parent populations to detect regions of the genome (called quantitative trait loci, QTLs) influencing phenotypic traits measured in unique and multiple environments. The population must be composed of crosses between a set of at least three parents (e.g. factorial design, 'diallel', or nested association mapping). The functions cover data processing, QTL detection, and results visualization. The implemented methodology is described in Garin, Wimmer, Mezouk, Malosetti and van Eeuwijk (2017) <[doi:10.1007/s00122-017-2923-3](https://doi.org/10.1007/s00122-017-2923-3)> as well as in Garin, Malosetti and van Eeuwijk (2020) <[doi:10.1007/s00122-020-03621-0](https://doi.org/10.1007/s00122-020-03621-0)>.

Version: 1.4.0  
Depends: R (≥ 3.1.0)  
Imports: [dplyr](#), [ggplot2](#), graphics, grDevices, [igraph](#), [Matrix](#), methods, [nlme](#), parallel, [qtl](#), stats, utils  
Suggests: [knitr](#), [rmarkdown](#), [testthat](#)  
Published: 2023-01-05  
Author: Vincent Garin  [aut, cre], Valentin Wimmer [aut], Dietrich Borchardt [ctb, dtc], Fred van Eeuwijk [ctb, ths], Marcos Malosetti [ctb, ths]  
Maintainer: Vincent Garin <vincent.garin6 at gmail.com>  
BugReports: <https://github.com/vincentgarin/mppR/issues>  
License: [GPL-3](#)  
URL: <https://github.com/vincentgarin/mppR>  
NeedsCompilation: no  
Citation: [mppR citation info](#)  
Materials: [README](#) [NEWS](#)  
CRAN checks: [mppR results](#)

### Documentation:

Reference manual: [mppR.pdf](#)  
Vignettes: [QTL detection in multiparental populations characterized in multiple environments](#)  
[mppR: An R Package for QTL Analysis in Multi-parent Populations](#)

## Results – General QTL detection results

---

	FLAG	PH	NODE_N	NODE_L	PED	PAN	GWGH	YIELD
Grinkan 2012	6 (48.9)	3 (48)	4 (17.6)	2 (47.1)	6 (31.2)	5 (8.5)	3 (13.5)	3 (5.4)
Grinkan 2013	4 (32.1)	4 (44.5)	1 (9.6)	2 (42.6)	2 (21.1)	1 (4.6)	0	2 (7.3)
Kenin-Keni 2012	6 (53.4)	2 (9.6)	2 (16.8)	3 (18.8)	3 (14.3)	2 (5.2)	2 (7.9)	1 (3.9)
Kenin-Keni 2013	2 (35.5)	2 (12.7)	2 (22.3)	0	2 (16.9)	1 (5.8)	0	1 (5.9)
Lata3	4 (50.3)	2 (20.1)	2 (13.7)	1 (11)	2 (10.2)	0	6 (30.4)	4 (14.3)

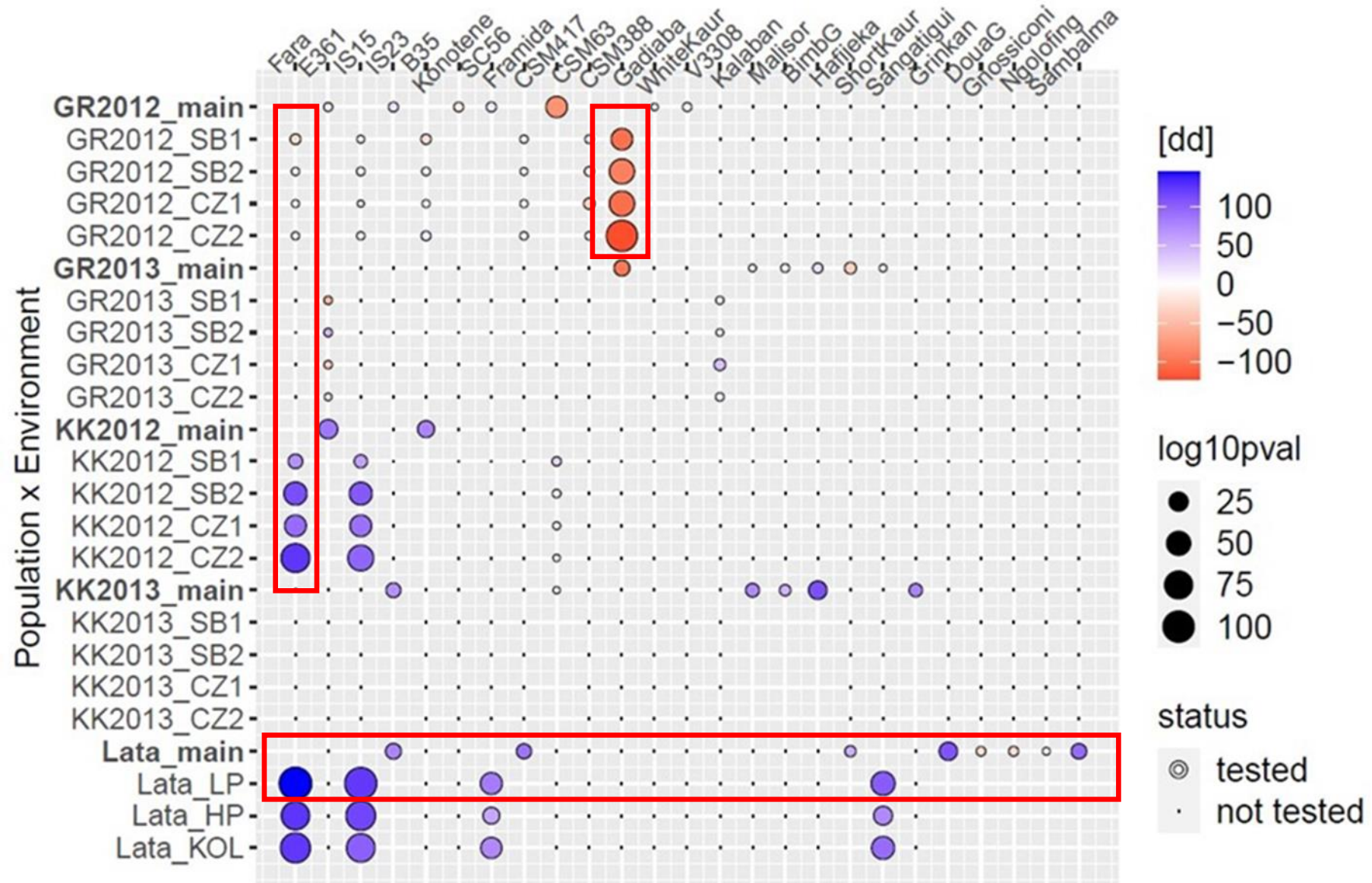
Good explanation (R<sup>2</sup>) of FLAG, PH, NODE\_N and PED. Less QTL and lower R<sup>2</sup> for PAN and YIELD.

## Results – Large QTL effect with candidate genes

QTL ID	trait	chr	range [cM]	range [Mbp]	R2	QxE range	Npar	Candidate genes
Q_FL_3_78	FLAG	3	77.34-78.36	5.11-5.15	17.1	[-123; 144] [dd]	24	Elf3, SbCN12
Q_FL_6_3	FLAG	6	1.49-2.94	0.04-0.08	19.4	[-178; 130] [dd]	24	Ma6
Q_FL_6_38	FLAG	6	36.32-39.4	4.04-4.12	6.3	[-192; -27] [dd]	19	Ma1
Q_FL_9_105	FLAG	9	103.7-106.7	5.46-5.54	2.7	[-40;91] [dd]	23	SbFL9.1
Q_PH_7_76	PH	7	74.28-76.69	5.47-5.52	21.9	[-37; 69] [cm]	20	qHT7.1, (Dw3)
Q_PH_7_106	PH	7	102-108.3	5.94-6.07	7.6	[-14;58] [cm]	15	Dw3
Q_NN_3_78	NODE_N	3	78.13-78.75	5.14-5.17	9.3	[-1.8;2.1] [n]	16	Elf3, SbCN12
Q_NN_6_2	NODE_N	6	1.49-2.73	0.04-0.08	7.9	[-2.6;1.8] [n]	19	Ma6
Q_NL_7_78	NODE_L	7	76.29-79.59	5.51-5.58	29.9	[-0.1;4.2] [cm]	16	qHT7.1, (Dw3)
Q_NL_7_98	NODE_L	7	96.1-100.7	5.83-5.91	11.8	[-5.8-4.2] [cm]	8	Dw3
Q_PED_7_78	PED	7	74.8-82.1	54.84-56.26	12.6	[-8; 10] [cm]	24	qHT7.1, (Dw3)

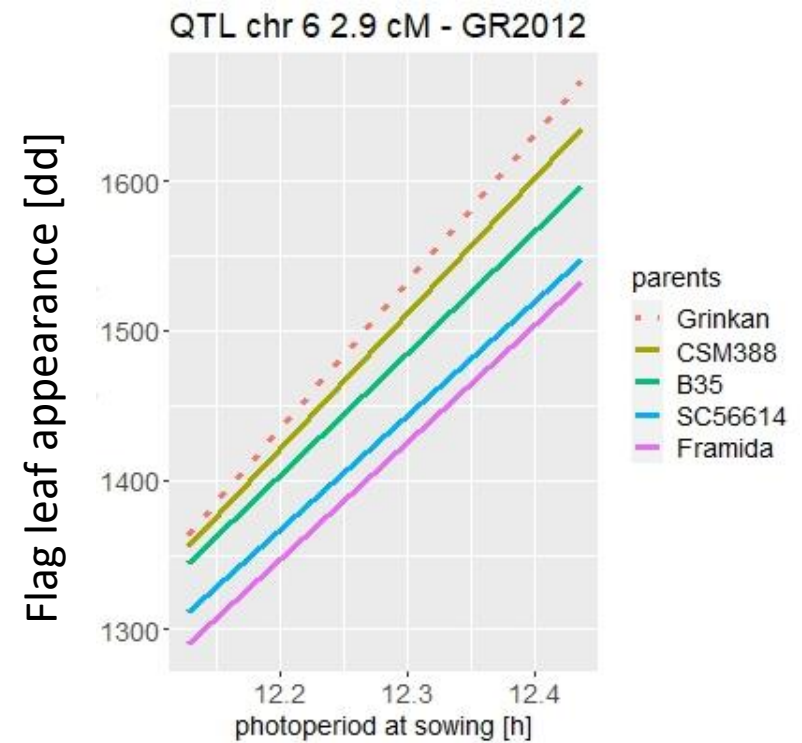
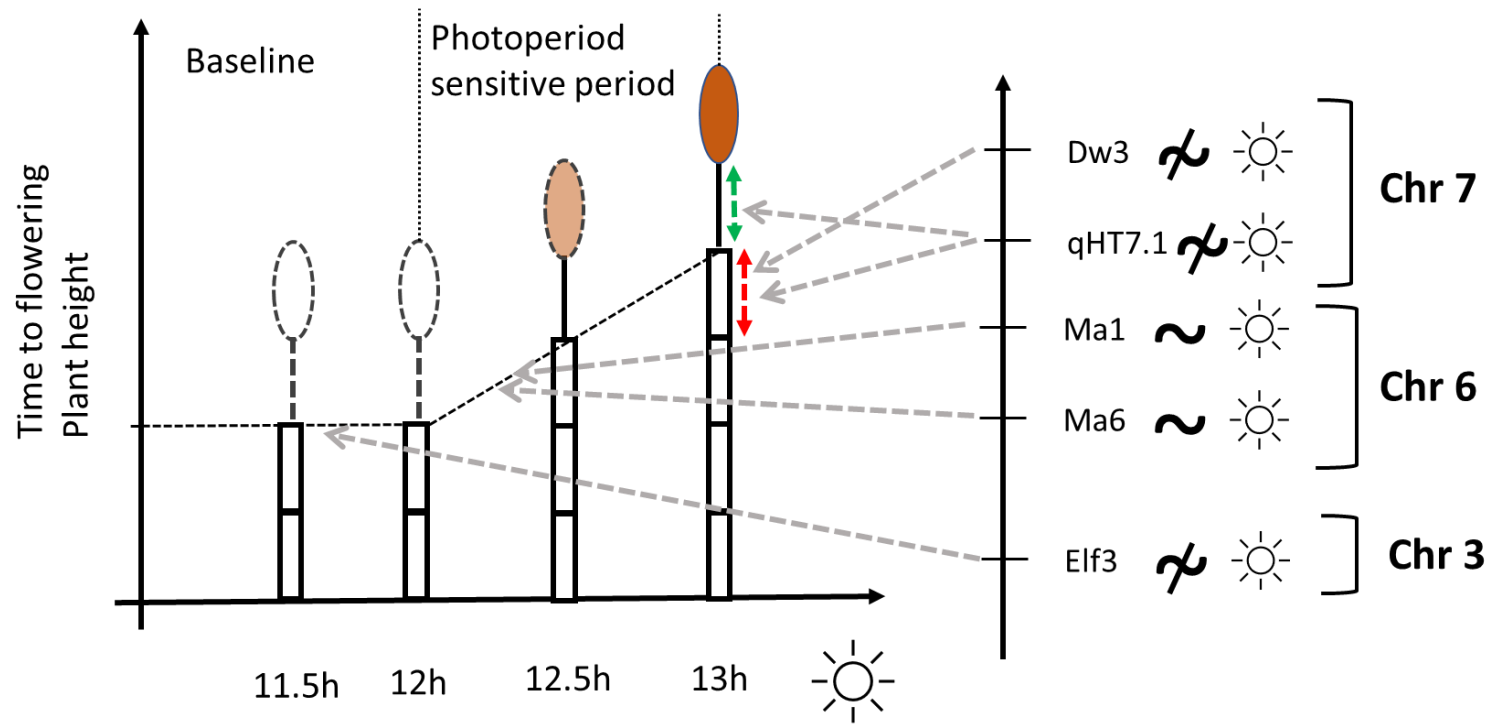
Large QTL effect with underlying candidate genes for FLAG and PH

# Results – Complex genetic pattern at large effect QTL positions

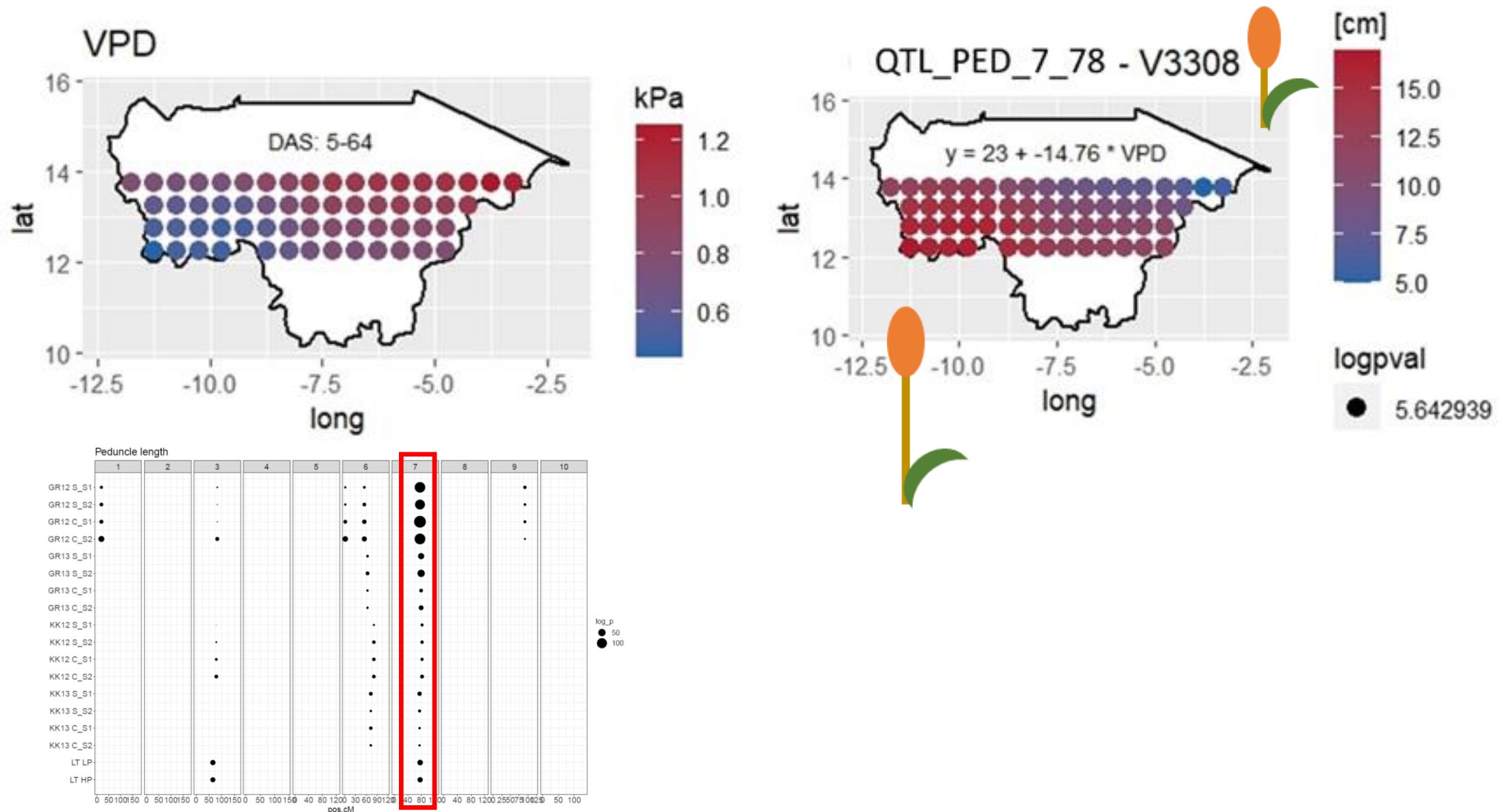


Parental QTL allelic effects are modulated by genetic background and environment

# Results – Environmental sensitivity of FLAG and PH large QTLs



# Results – Projection of expected QTL allelic effect in the Malian environment





# Decision support tools – R shiny application for interactive QTL effect database

**Description:** Search for QTL positions in the database using trait, chromosome, and position information. QTL effect can also be subsetted according to R2 value

Select a trait

Flag leaf appearance

Select chromosome(s)

6

Select physical boundaries ...

Lower boundary [bp]

0

Upper boundary [bp]

100000000

... or Genetic boundaries

Lower boundary [cM]

0

Upper boundary [cM]

1000

Filter by

Physical distance

Genetic distance

Optional filter by R2 value [0-100]

minimum R2 value

0

Search

QTL effect significance [-log10(p-value)]

QTL	chr	bp	cM	logp_av	logp_min	logp_max
QTL_FLAG_6_3	6.00	758633.94	2.60	47.20	15.90	96.50
QTL_FLAG_6_38	6.00	40649927.66	37.30	28.90	7.50	65.70

QTL R2

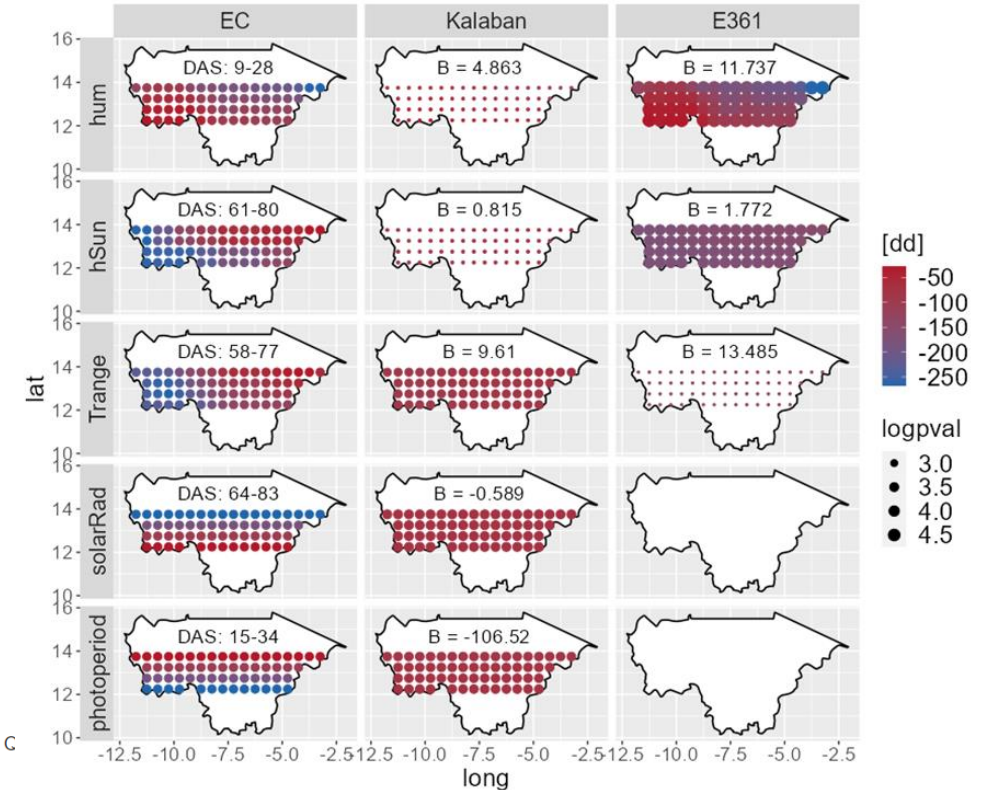
QTL	chr	bp	cM	R2_av	R2_min	R2_max
QTL_FLAG_6_3	6.00	758633.94	2.60	19.40	14.40	24.00
QTL_FLAG_6_38	6.00	40649927.66	37.30	6.30	2.20	11.10

QTL range of additive effects

QTL	chr	bp	cM	Eff_min	Eff_max
QTL_FLAG_6_3	6.00	758633.94	2.60	-178.51	130.16
QTL_FLAG_6_38	6.00	40649927.66	37.30	-191.67	26.98

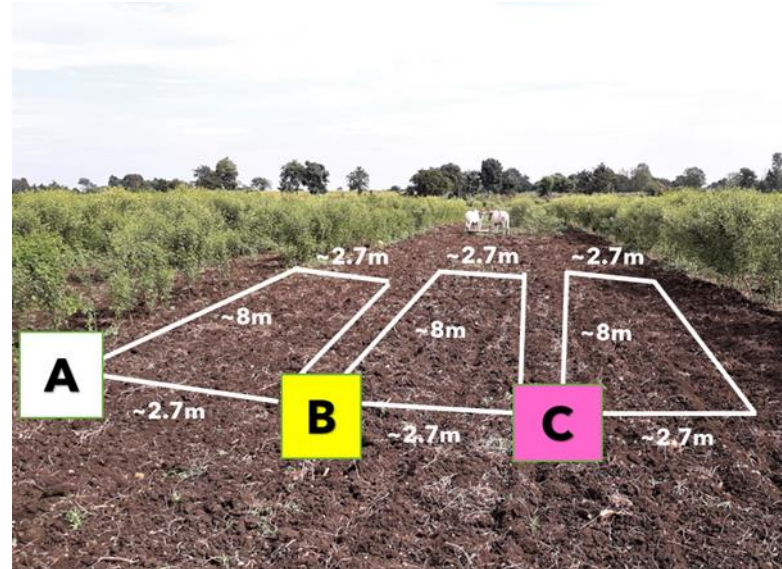
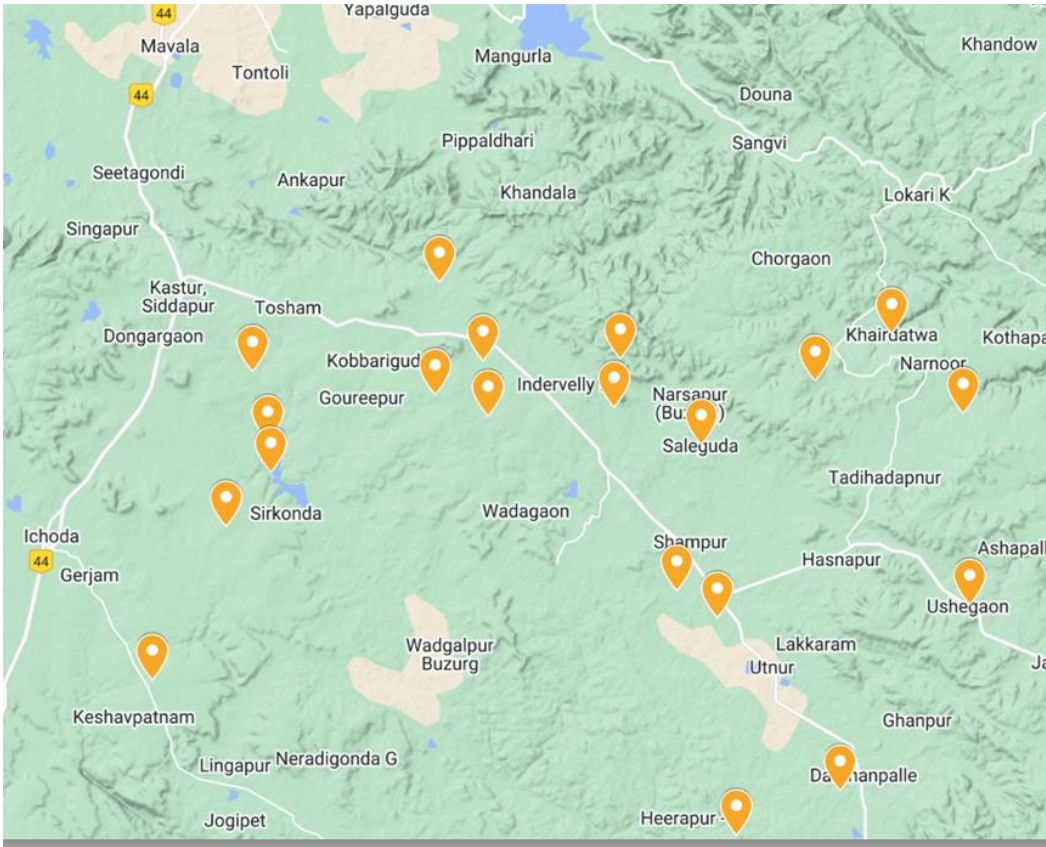
List of unique QTL position that can be copied to get extra information about the QTL effect and the C  
QTL\_FLAG\_6\_3 QTL\_FLAG\_6\_38

GR2013 Flag leaf - chr:6 cM:2.91 (S6\_838602)



<https://github.com/vincentgarin/SQE>

# The last mile – Integrating better farmers needs and context



Tricot experiment (van Etten et al. 2019) for sorghum participatory variety selection with 200 farmers from 20 villages of the Adilabad district (India)

# The last mile – Integrating better farmers needs and context

## ExM constraint selection

Trait selection  
General appreciation

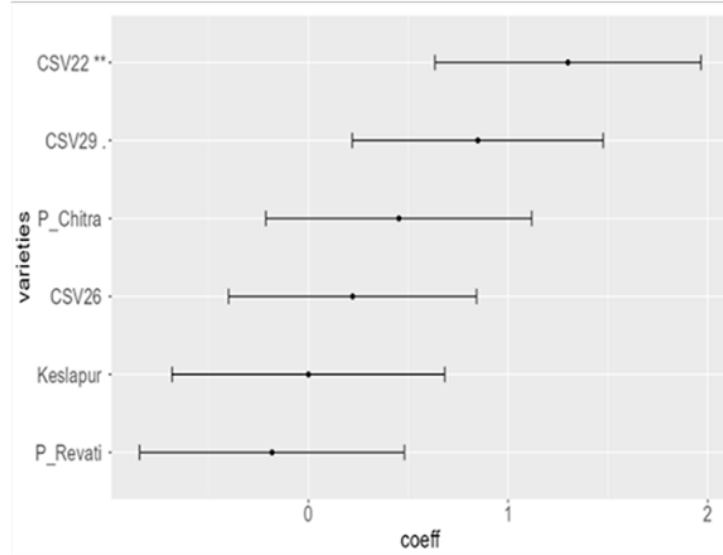
Soil type  
Black

**Soil quality**  
Good - Very good

Irrigation  
Overall

Fertilisation  
Fertilisation

## Farmer variety preference



## ExM constraint selection

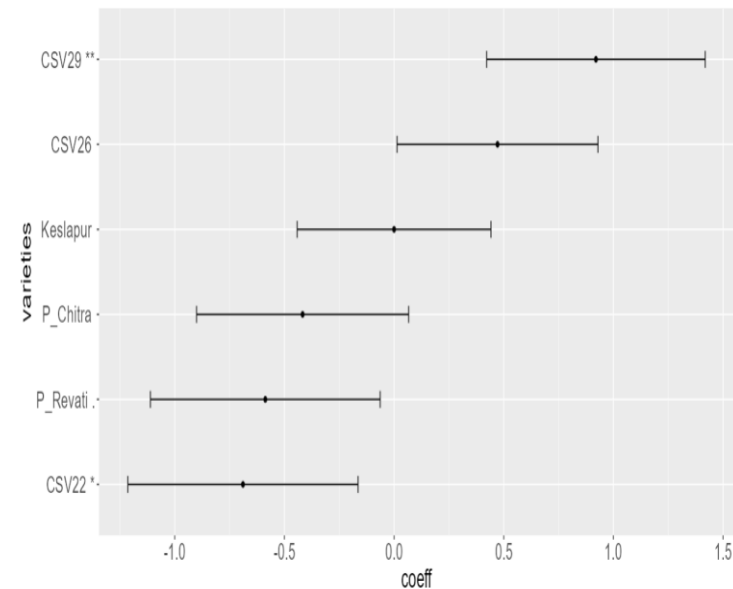
Trait selection  
General appreciation

Soil type  
Overall

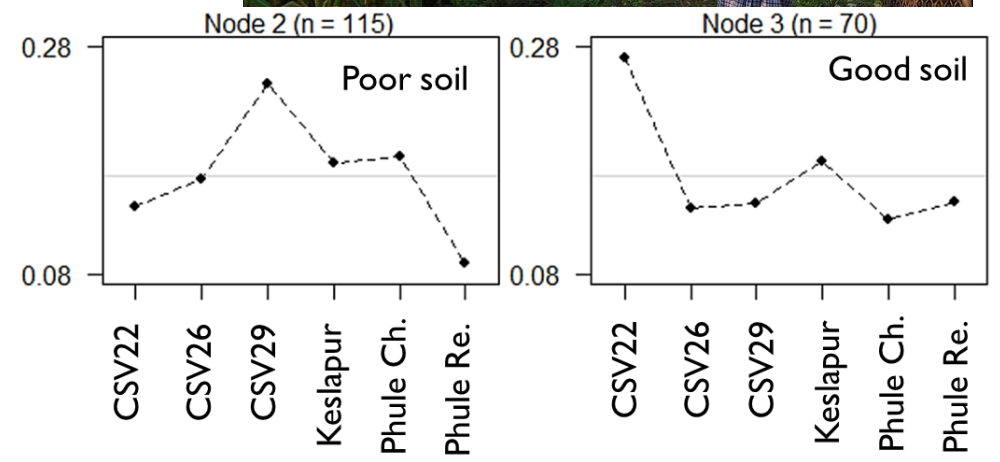
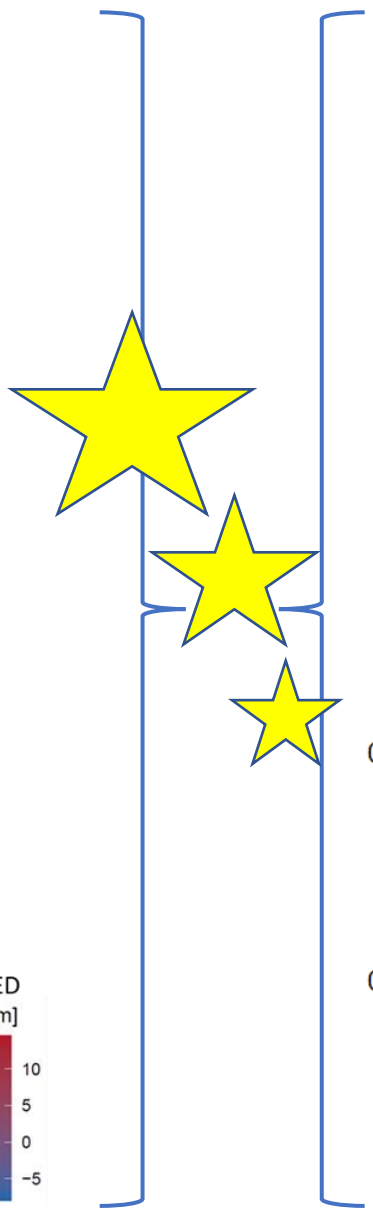
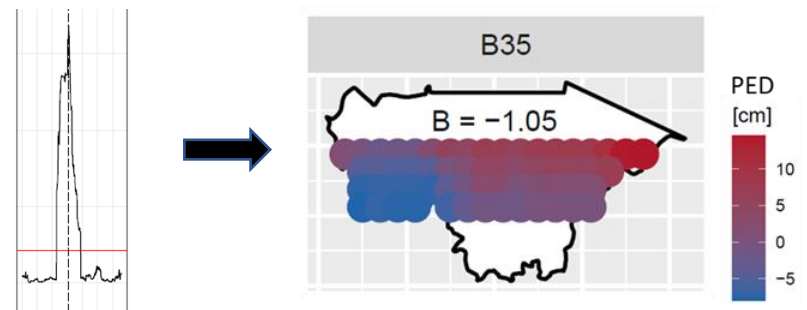
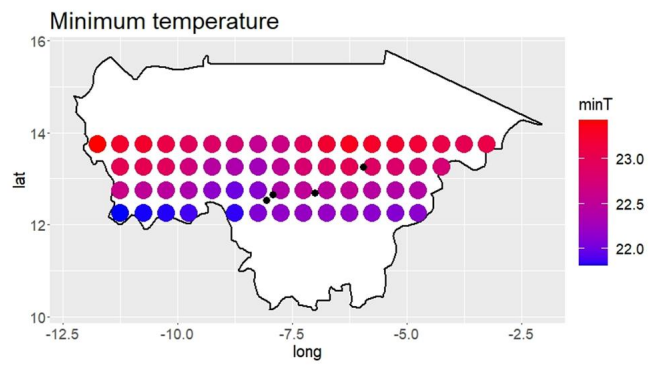
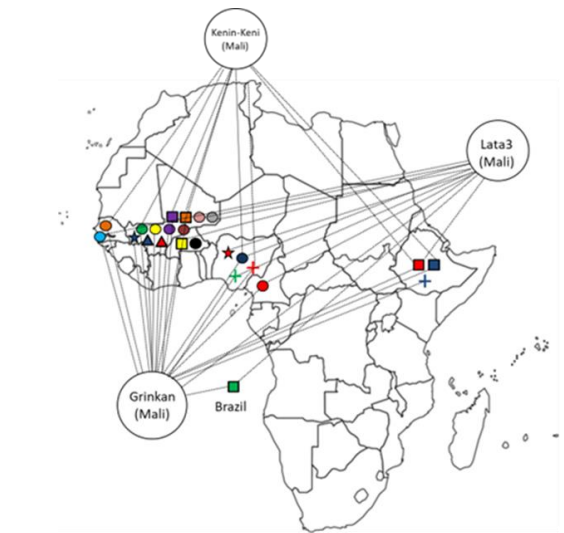
**Soil quality**  
Poor - Medium

Irrigation  
Overall

Fertilisation  
Overall



# Conclusion – Better integration of developed technologies and farmers need



# Questions

---



# References

---

Costa-Neto, G., Galli, G., Carvalho, H. F., Crossa, J., & Fritsche-Neto, R. (2021). EnvRtype: a software to interplay enviromics and quantitative genomics in agriculture. *G3*, *11*(4), jkab040

Sparks, A. H. (2018). nasapower: a NASA POWER global meteorology, surface solar energy and climatology data client for R. *Journal of Open Source Software*, *3*(30), 1035.

Van Etten, J., Beza, E., Calderer, L., Van Duijvendijk, K., Fadda, C., Fantahun, B., ... & Zimmerer, K. S. (2019). First experiences with a novel farmer citizen science approach: crowdsourcing participatory variety selection through on-farm triadic comparisons of technologies (tricot). *Experimental Agriculture*, *55*(S1), 275-296.