



INRAE

Adaptation of GWAS models for plant resistance to viruses: from rediscovering major genes to highlighting of new complex traits in breeding germplasm.



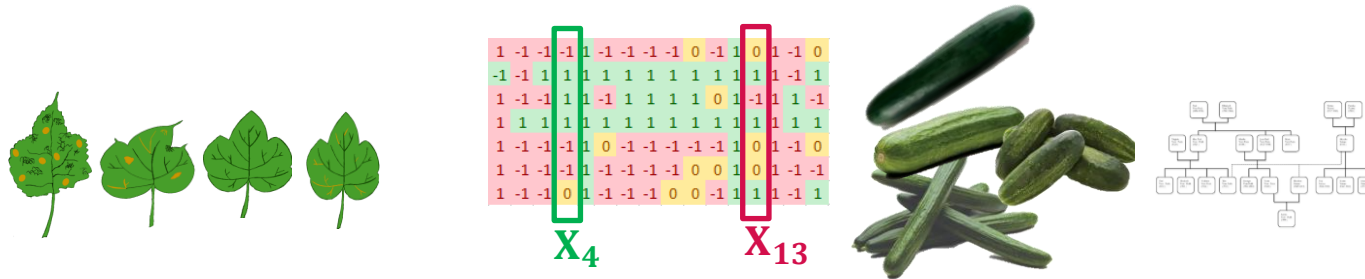
Eucarpia cucurbit 2021

May 2021– Séverine MONNOT*, Laurence Moreau, Tristan Mary-Huard, Mélissa Cantet, Nathalie Boissot



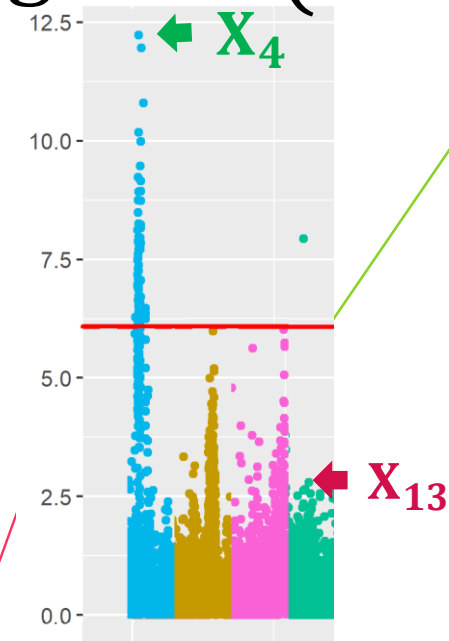
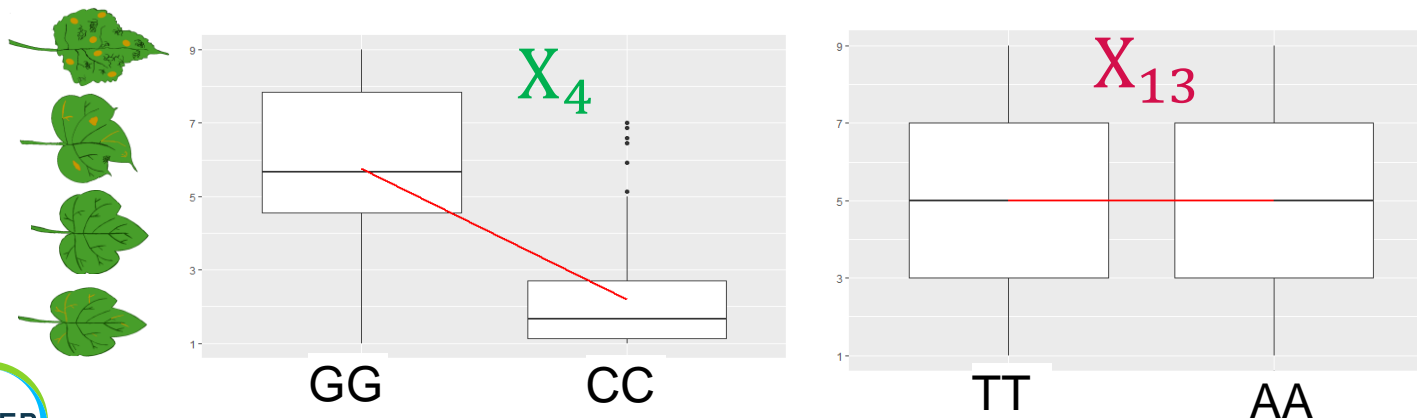
Genome Wide Association Studies (GWAS): some explanations

Definition : Statistical method that associates a phenotypic variation with a genetic variation



$$Y = \mu + X_M\beta + X_Wv + Z_Gg + e, \quad g \rightarrow N(0, \sigma_K^2 I)$$

The regression is performed for **each** marker :



Develop a cucumber diversity panel and phenotype its response variability

Discover new native resistances by using original GWAS, developed post-GWA analysis and to set up GWS.



Cucumber diversity in the PhD population



Mechanical inoculation trial – greenhouse conditions



August 2020 – openfield trial

Cucumber breeding panel :

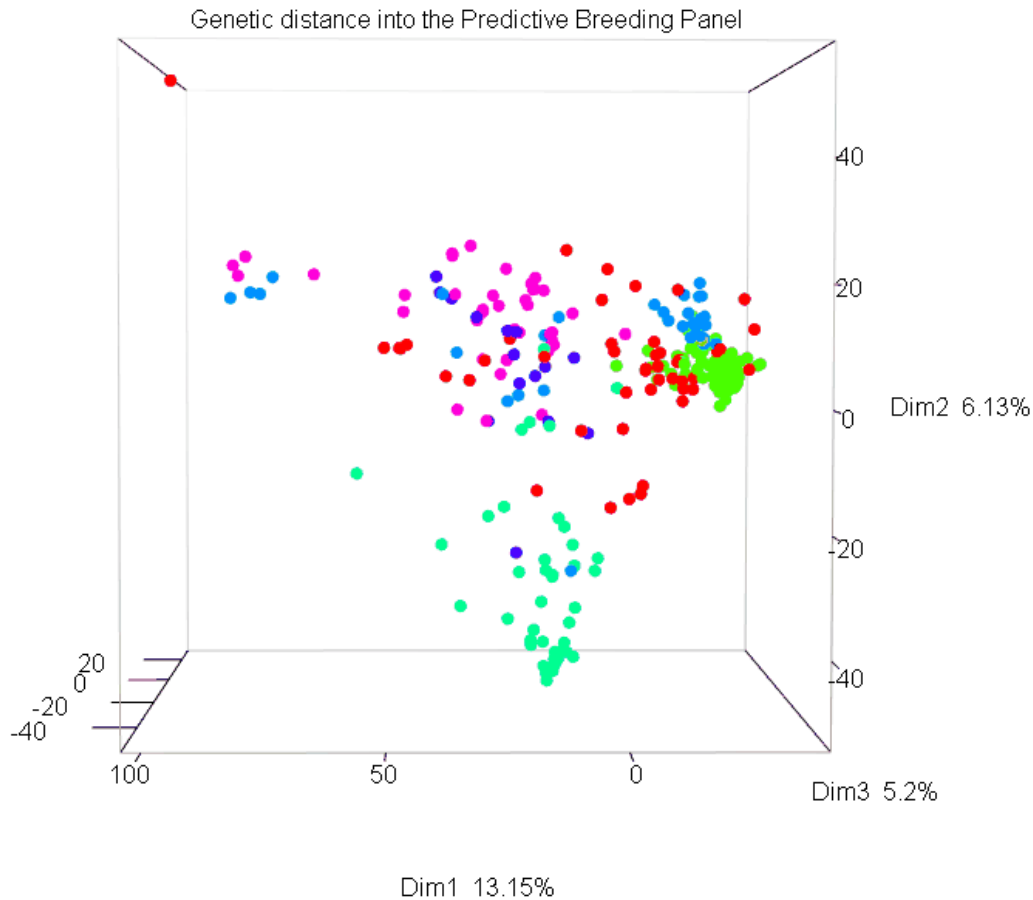
- // 220 old elite lines from all programs
(EUR, BAP, SLI, Gu...)
- // 35 landraces
- // 25 F1

Tested for a set of 7 viruses by

- // Mechanical inoculation
 - // ZYMV
 - // Polygenic virus
- // Natural infection in openfield condition

SNPs as genetic variation highlight population structure

1 million Single Nucleotide Polymorphism aligned on the 3rd version from the reference genome CCL (Li et al, 2019)



Europeans have low genetic diversity

Pickles and Slicers are tightly linked.

Gurkins are isolated.

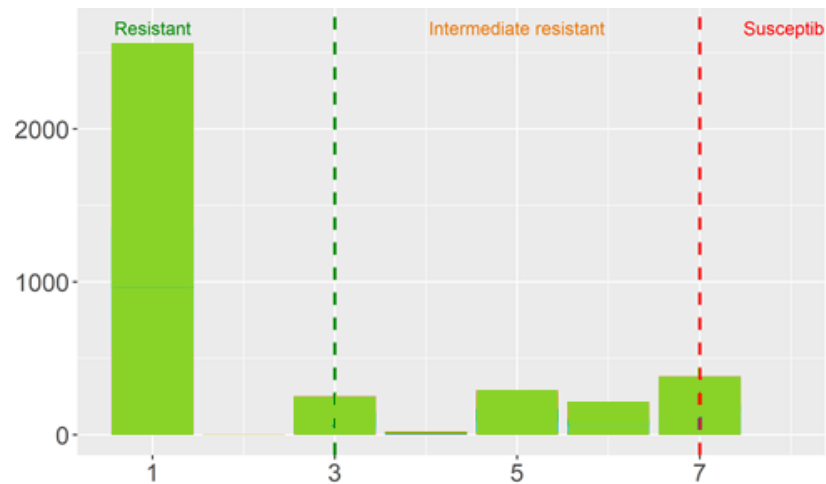
Beitalphas = Bridge between programs

2 groups of landraces

GWAS hypothesis and genetic architecture of resistance to viruses

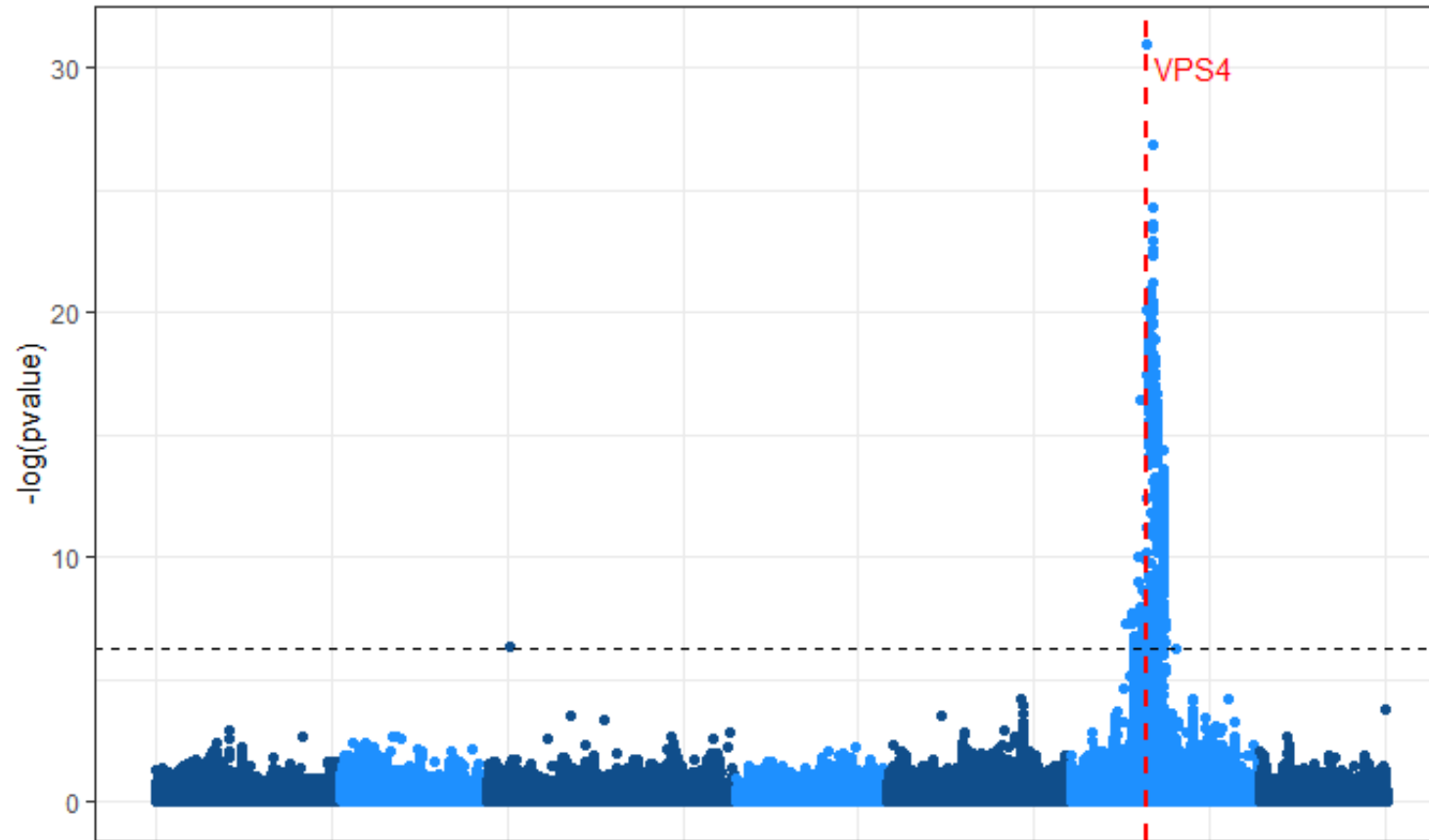
Resistance architecture

- // Quantitative (**Multigenic** with minor effect genes)
- // Qualitative (**Monogenic** with major effect gene)
- // Gaussian distribution not observed → One major gene widely used (ZYMV)



ZYMV phenotypes
Binomial distribution

Modele under the rare variant hypothesis detected VPS4 for ZYMV



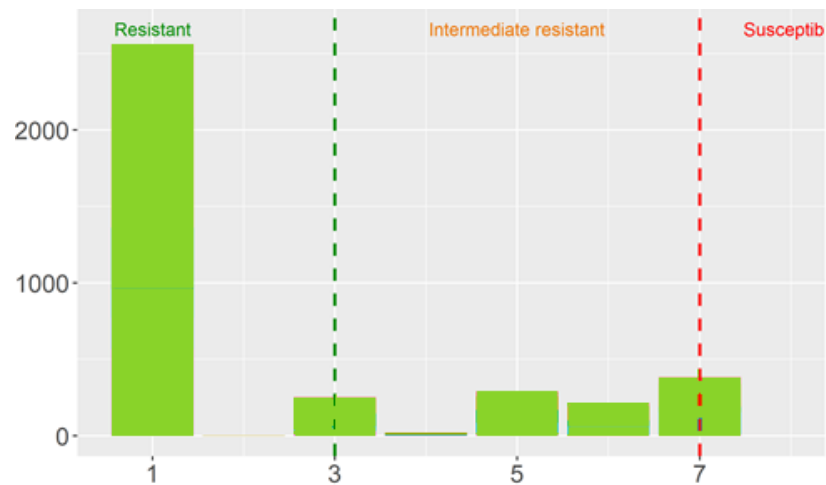
- The resistance for ZYMV in our breeding panel mapped on the chromosome 6
- Close to the VPS4 gene, known to be involved in ZYMV resistance in cucumber

(Ramirez-Madera & Havey, 2017)

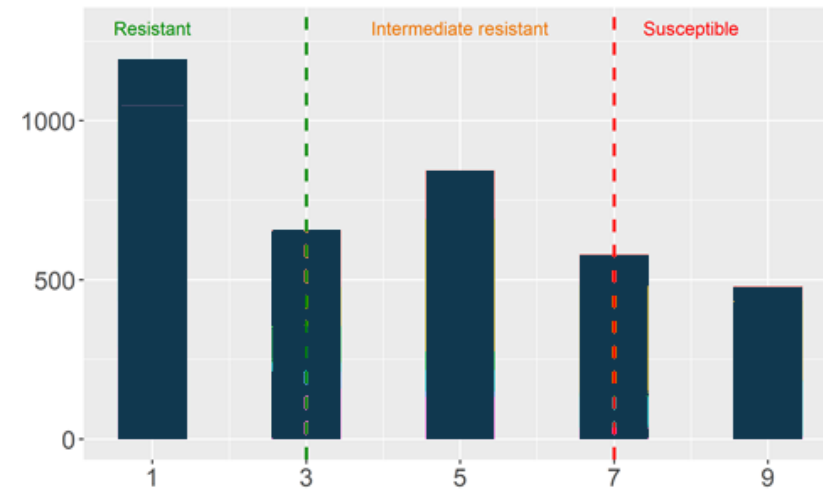
GWAS hypothesis and genetic architecture of resistance to viruses

Resistance architecture

- // Quantitative (**Multigenic** with minor effect genes)
- // Qualitative (**Monogenic** with major effect gene)
- // Gaussian distribution not observed → One major gene widely used (ZYMV)
- // Hide minor genes effect → Several resistances available (Virus A)



ZYMV phenotypes
Binomial distribution

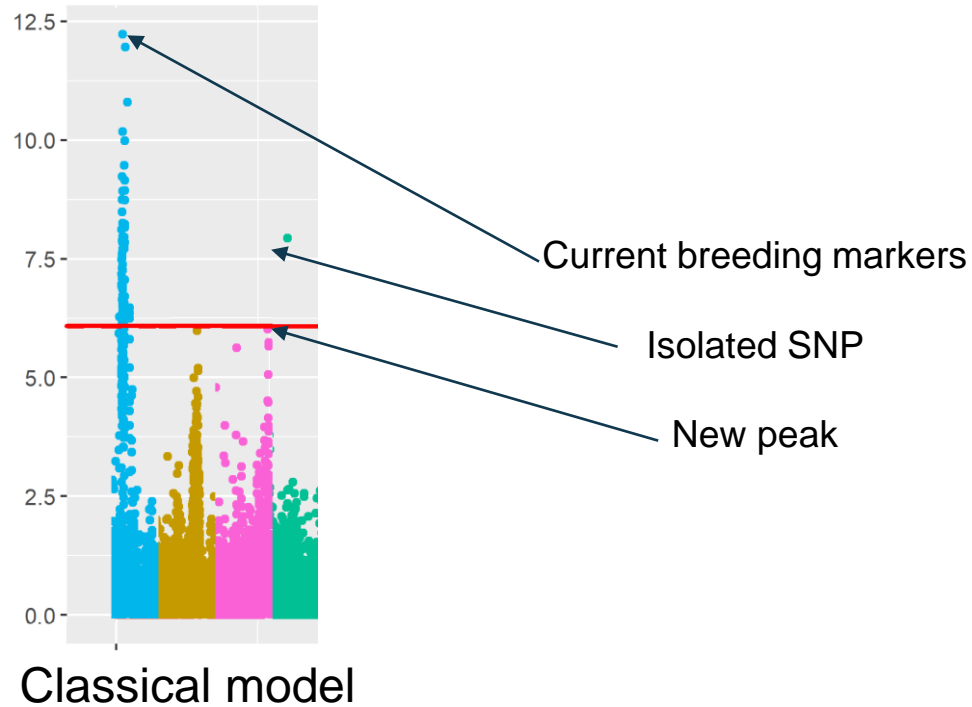


Polygenic virus phenotypes
Pseudo-Gaussian distribution

Find minor genes hidden by a major gene

A breeding marker located in a major peak

GWAS results using Y – classical model



// Peaks located on breeding markers

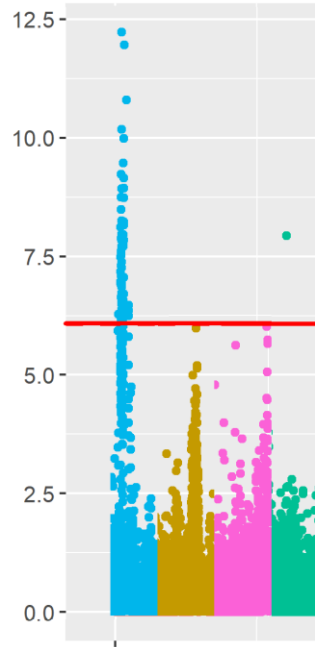
// New peaks

// Isolated SNPs



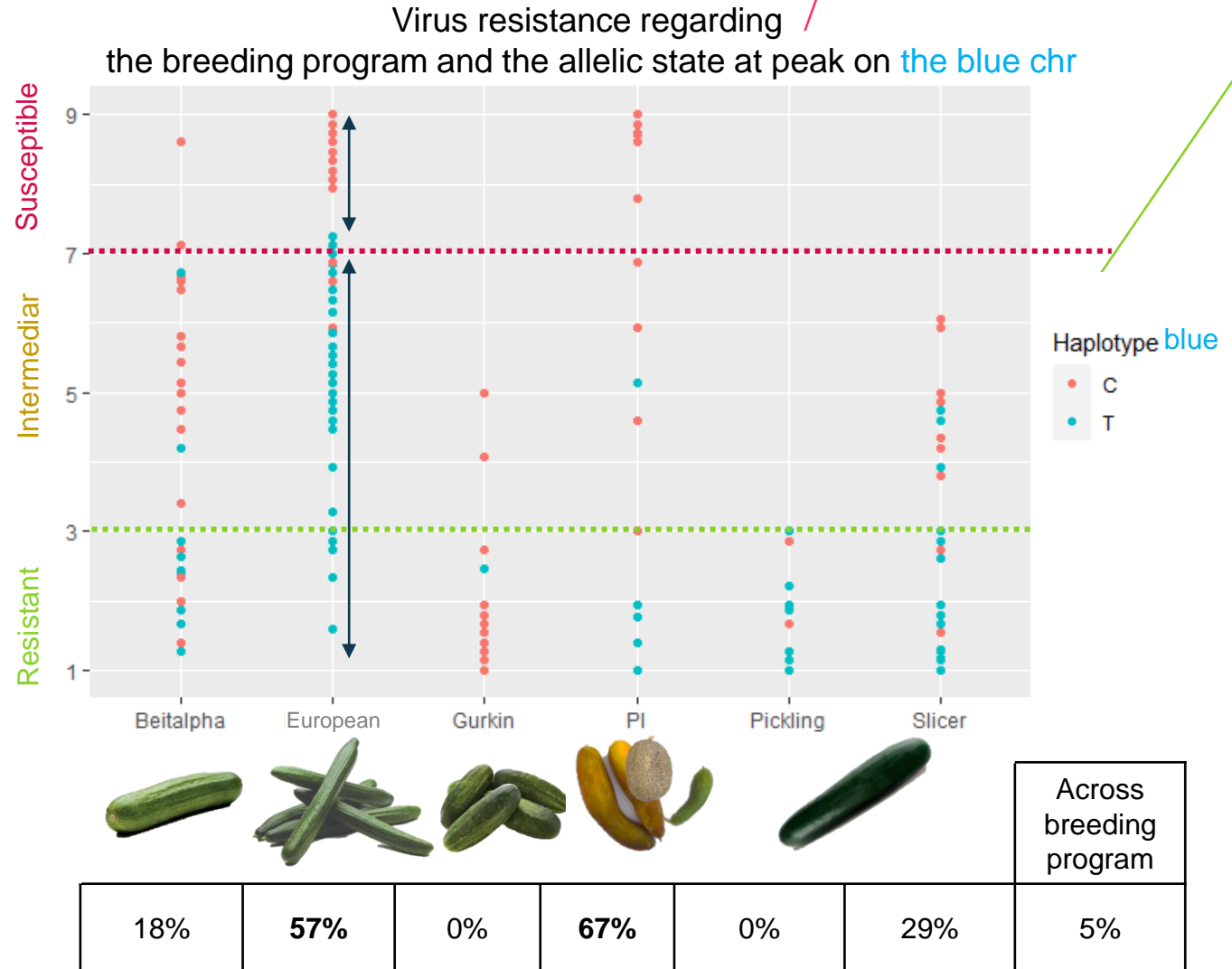
Find minor genes hidden by a major gene

A breeding marker located in a major peak **but not predictive for our breeding germplasm**



A breeding marker is available

- // OK for European and landraces
- // Not working for other breeding programs
- // How can we find the other causative genes?

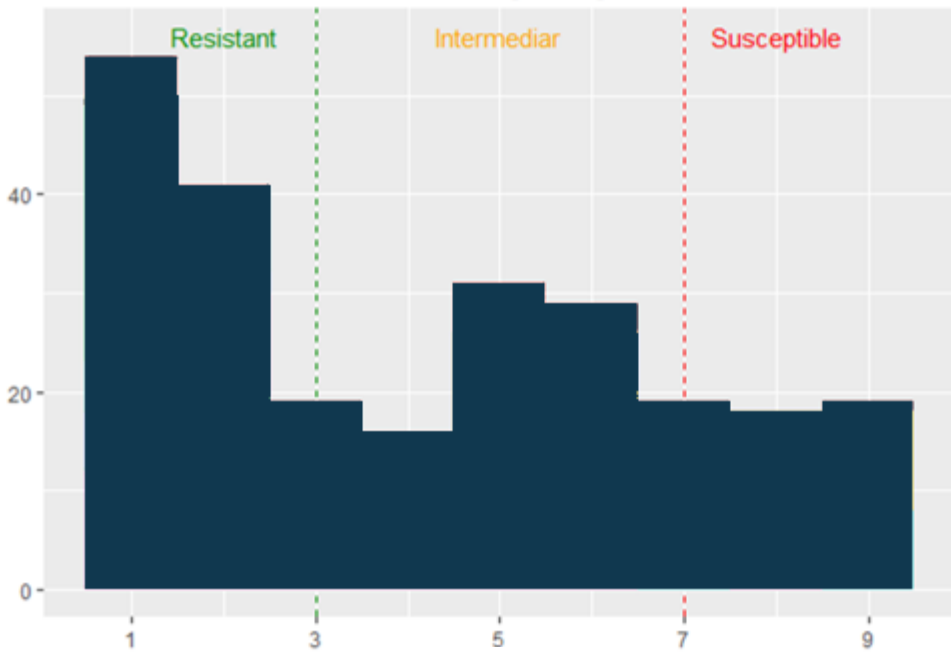


Find minor genes hidden by a major gene

Build new phenotypes corrected with the breeding marker effect

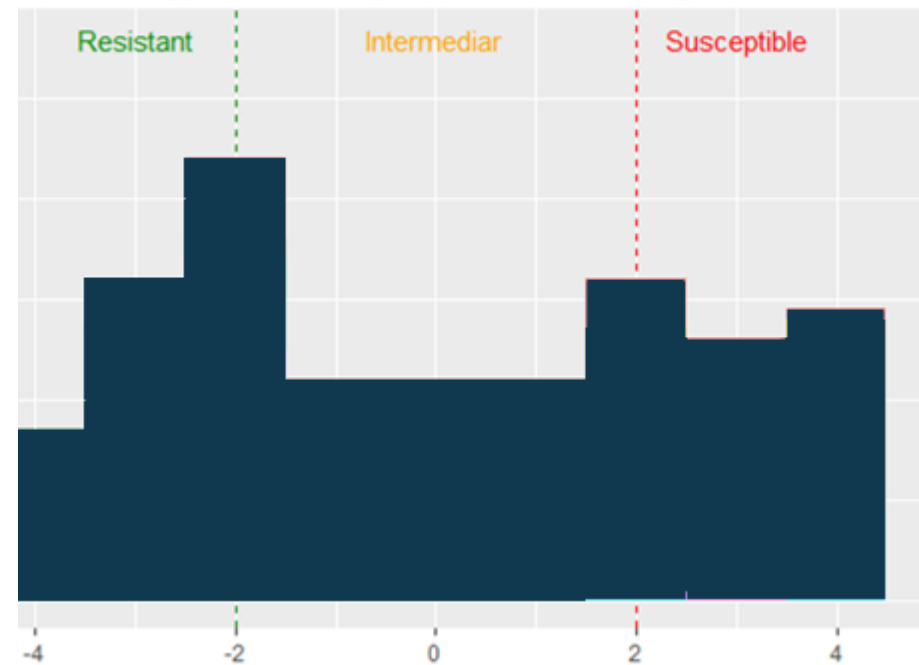
$$Y = \mu + X_{\text{Breed mark}}\beta_{\text{Breed mark}} + e = Y_2$$

Distribution of the 280 accessions regarding their VirA resistance score



Distribution of the 280 accessions regarding their VirA resistance score

Phenotypes corrected by the effect of the breeding marker



Y

Y₂ = residuals

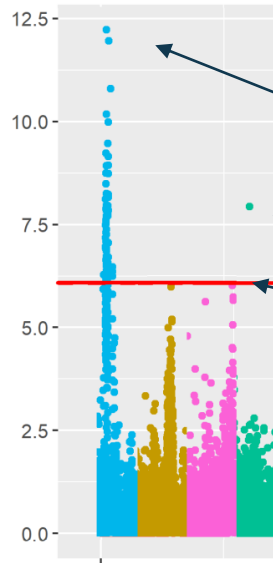
$$Y \text{ or } Y_2 = \mu + X_M\beta + X_Wv + Z_Gg + e$$



Find minor genes hidden by a major gene

Breeding marker correction highlighted new loci

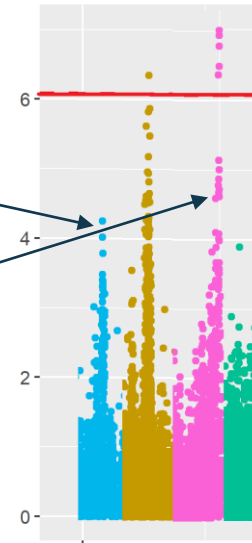
GWAS results using Y – classical model



Classical model

- // Peaks located on breeding markers
- // New peaks
- // Isolated SNPs

GWAS results using Y2 – corrected by other gene effect

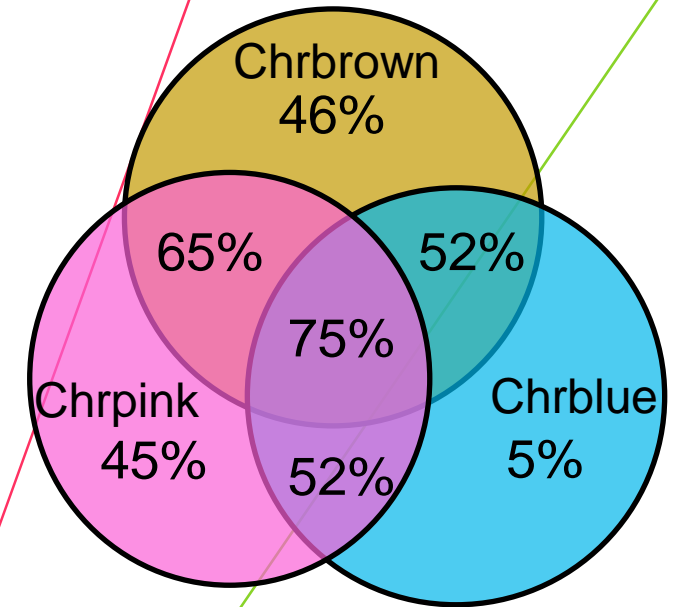
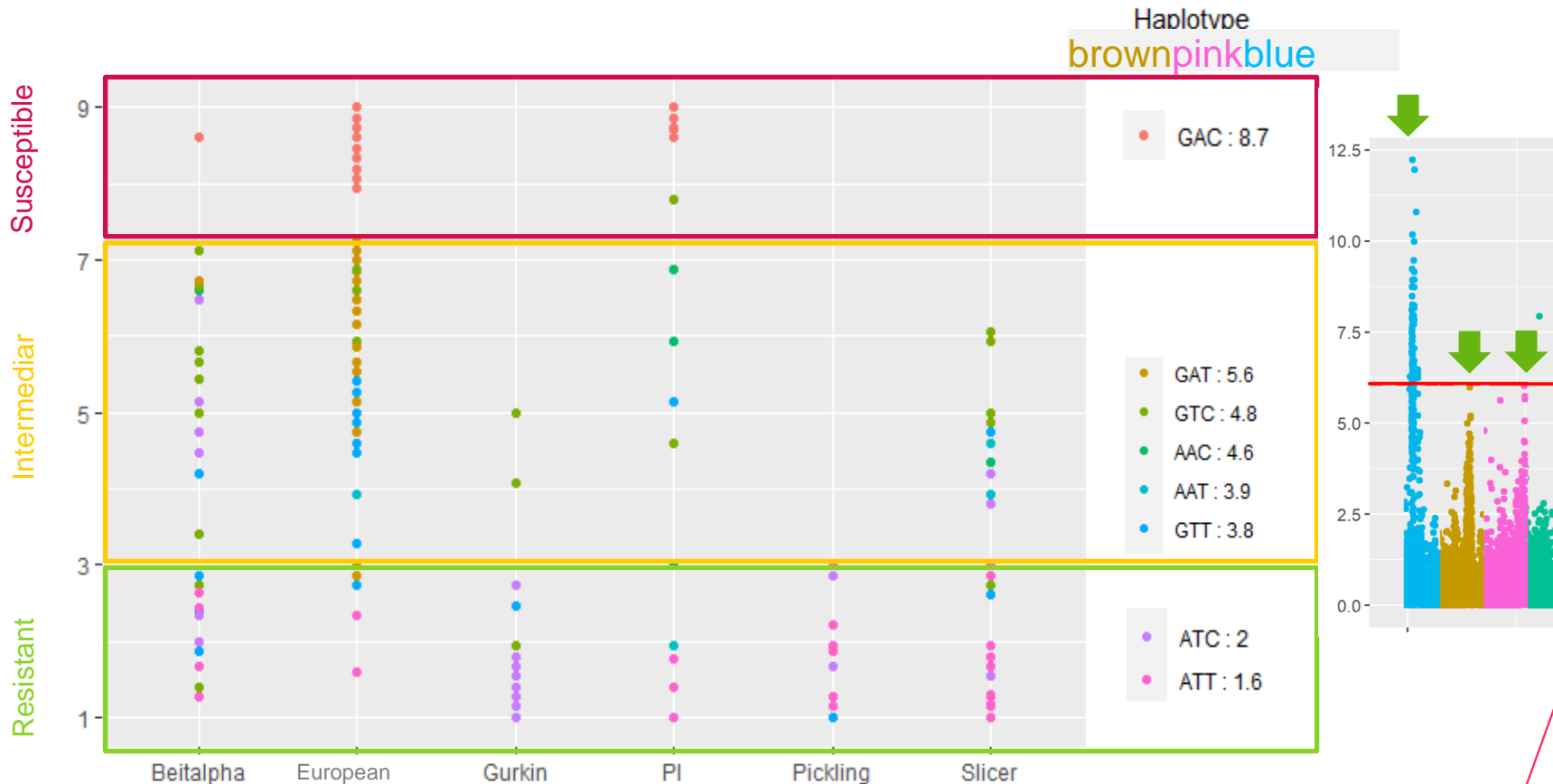


Model corrected by the breeding marker on chrblue

- // The peak located on the chrblue breeding marker disappeared
- // The other peaks remained
- // The isolated SNPs disappeared → False positive

Find minor genes hidden by a major gene

Build haplotypes combining the 3 main peak allelic states


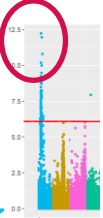


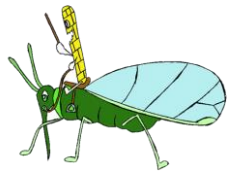



Percent of the phenotypic variation explained by 1 marker or by combination of markers

→ Encouraging results for the GWS set up

What's next?

1. Understand vector role with openfield data and aphid resistance

$$Y = \mu \left[\begin{array}{l} + X_{\text{blue}} \beta_{\text{blue}} \\ + X_{\text{resVect}} \delta_{\text{resVect}} \end{array} \right] + X_M \beta + X_W v + Z_G g + e$$

$g \rightarrow N(0, \sigma_K^2 I)$
 $e \rightarrow N(0, \sigma_e^2 I)$

2. Compare GWAS and GWS prediction

3. Synteny with *Arabidopsis thaliana*

Take home messages

GWAS is a powerful tool to map quantitative and qualitative native traits

New models can pinpoint new virus resistances with the same dataset

postGWAS analysis can highlight gene combination and help to adapt GWS models

Packages used :

// Gaussian and corrected: MM4LMM & mlmm

// Generalized linear model – Poisson, Binomial, etc. : glmnet





Do you have any question?

*Special thanks to
INRAE workteam: Pascale, Anaïs, Karine, Marie,
Salomé, petit Pierre, Nathalie, Christophe*

*Bayer for the funding of my PHD and the acquisition
of phenotypic and genotypic data. Special thanks to
Maurine, Agnès, Rebecca and their teams*



*My PhD supervisors Nathalie, Mélissa, Valérie,
Laurence et Tristan*

For further exciting conversation feel free to contact
me at severine.monnot@inrae.com

INRAE

PHD context

**Selection for oligogenic resistance to multiple viruses in cucumber:
from implementing genomic selection methods to in-depth investigation of gene repertoires involved in
resistance**

Observation :

Monogenic resistances with strong effects and easy to study have already been mapped and used in breeding programs.

Can we find new ways to improve level and durability of plant virus resistances?

Goal 1: Apply new GWAS models more adapted to the detection of locus linked with plant virus resistance in order to understand the genetic architecture of the different traits involved.

Goal 2: Determine if Genome Wide Selection can improve plant virus resistance breeding, regarding the trait genetic architecture and the population structure.

Goal 3: From prediction to explanation, use the plant model *Arabidopsis thaliana* with synteny studies.

