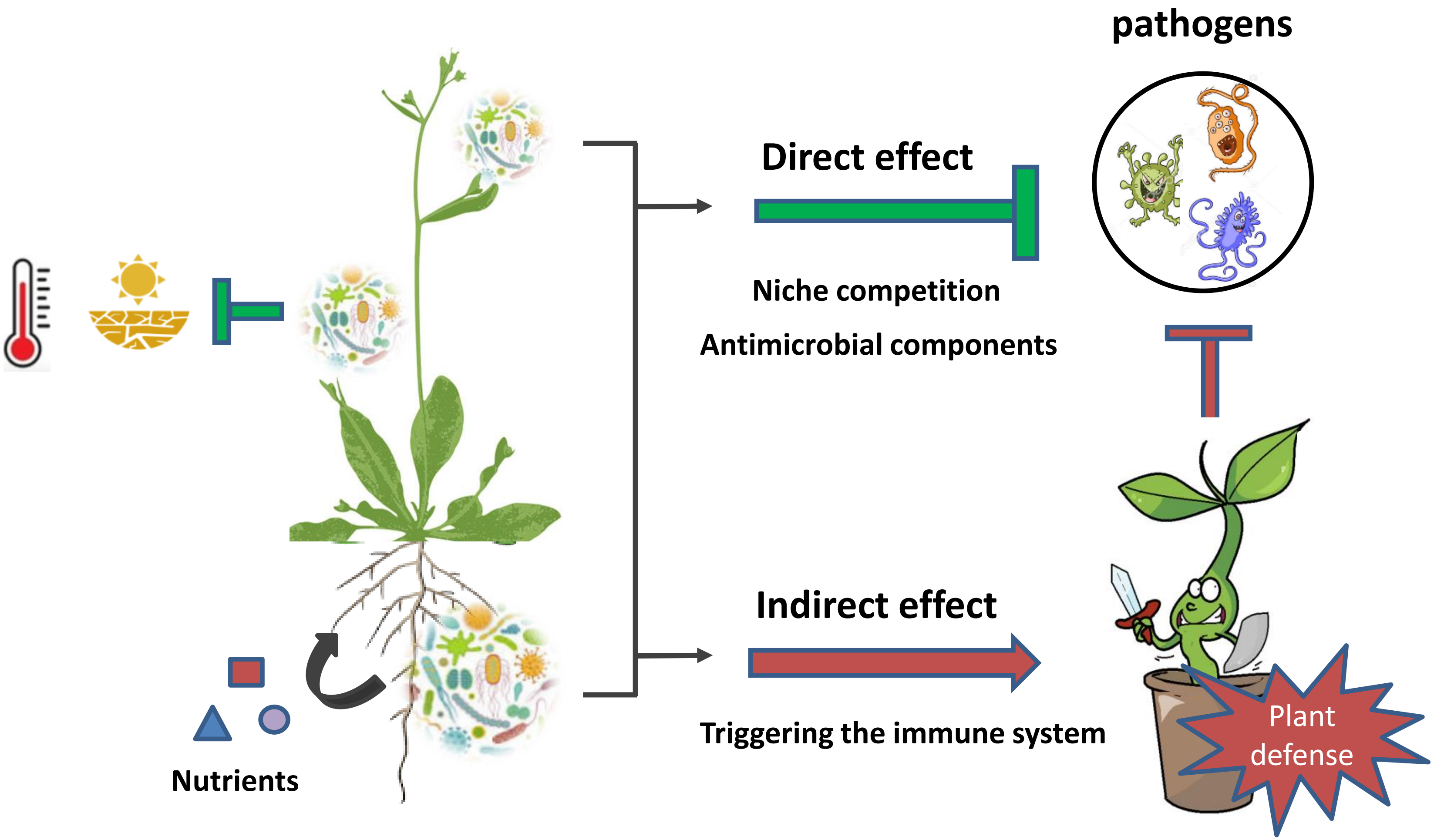
From ecological genomics to plant-microbiota co-selection

Fabrice Roux

Plant-Microbe-Environment Interactions lab
CNRS-INRAE, Toulouse, France

Plant-Microbiota interactions

Microbiota is crucial for **plant health** and thereby for **plant growth and survival**, due in large part to the presence of **beneficial microbes.**



Environmental effects on plant-microbiota interactions

Abiotic

Climate

(Fitzpatrick *et al.* Annual Rev Microbiol 2020)

Soil physical-chemical properties

(Bulgarelli *et al.* Annual Rev Plant Biol 2013)

Biotic

Insects

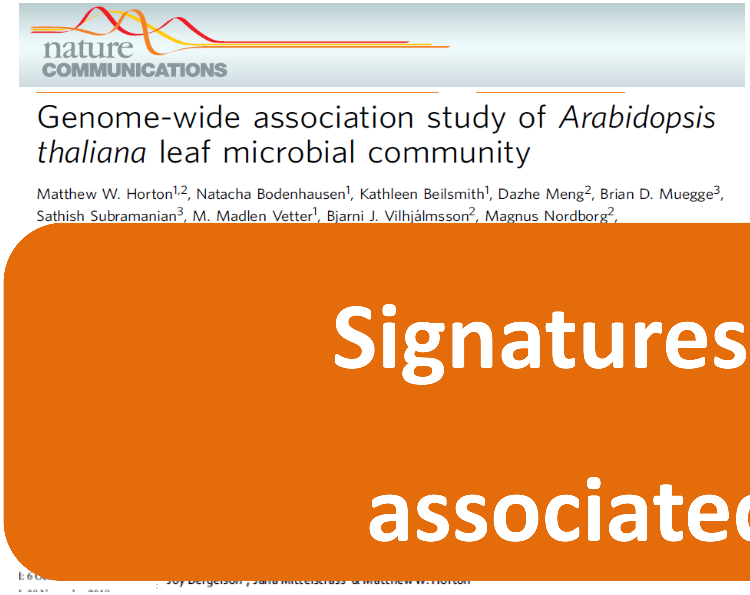
(Humphrey & Whiteman Nat Ecol Evol 2020)

Neighbouring plants

(Meyer et al. ISME 2022)

Host genetic effects on microbiota

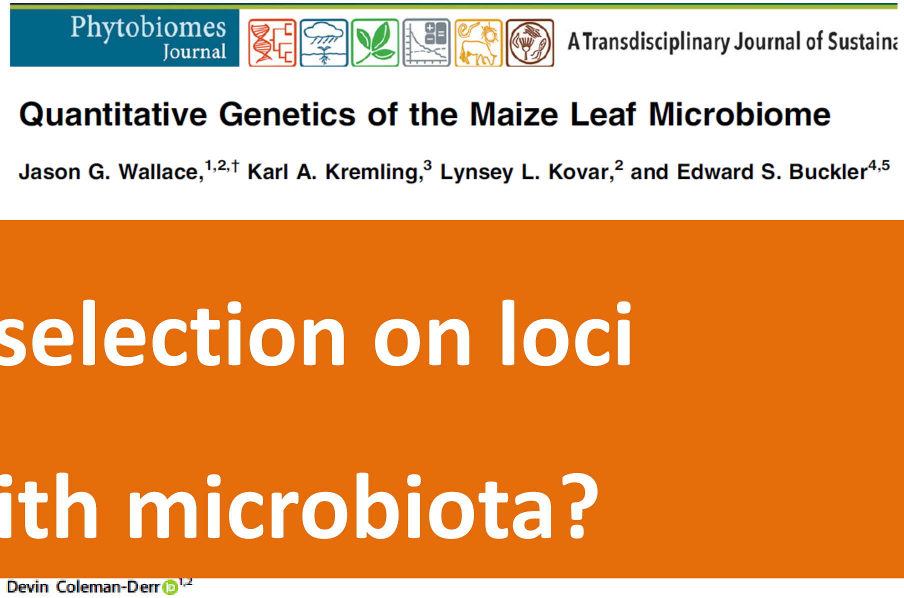
Very few GWA studies → polygenic architecture


nature COMMUNICATIONS

Genome-wide association study of *Arabidopsis thaliana* leaf microbial community

Matthew W. Horton[1,2], Natacha Bodenhausen[1], Kathleen Beilsmith[1], Dazhe Meng[2], Brian D. Muegge[3], Sathish Subramanian[3], M. Madlen Vetter[1], Bjarni J. Vilhjálmsson[2], Magnus Nordborg[2],

Joy Bergelson, Jana Mittelstrass & Matthew W. Horton

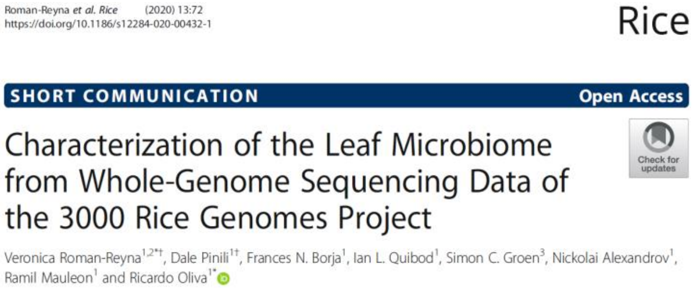

Phytobiomes Journal A Transdisciplinary Journal of Sustaina

Quantitative Genetics of the Maize Leaf Microbiome

Jason G. Wallace,[1,2,†] Karl A. Kremling,[3] Lynsey L. Kovar,[2] and Edward S. Buckler[4,5]

Devin Coleman-Derr[1,2]

Signatures of selection on loci associated with microbiota?

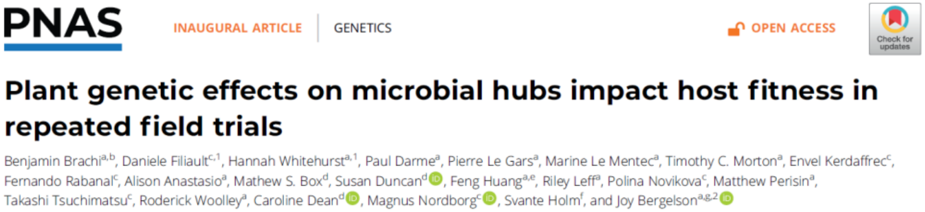

Roman-Reyna *et al. Rice* (2020) 13:72
https://doi.org/10.1186/s12284-020-00432-1

Rice

SHORT COMMUNICATION **Open Access**

Characterization of the Leaf Microbiome from Whole-Genome Sequencing Data of the 3000 Rice Genomes Project

Veronica Roman-Reyna[1,2*†], Dale Pinili[1†], Frances N. Borja[1], Ian L. Quibod[1], Simon C. Groen[3], Nickolai Alexandrov[1], Ramil Mauleon[1] and Ricardo Oliva[1*]


PNAS INAUGURAL ARTICLE | GENETICS OPEN ACCESS

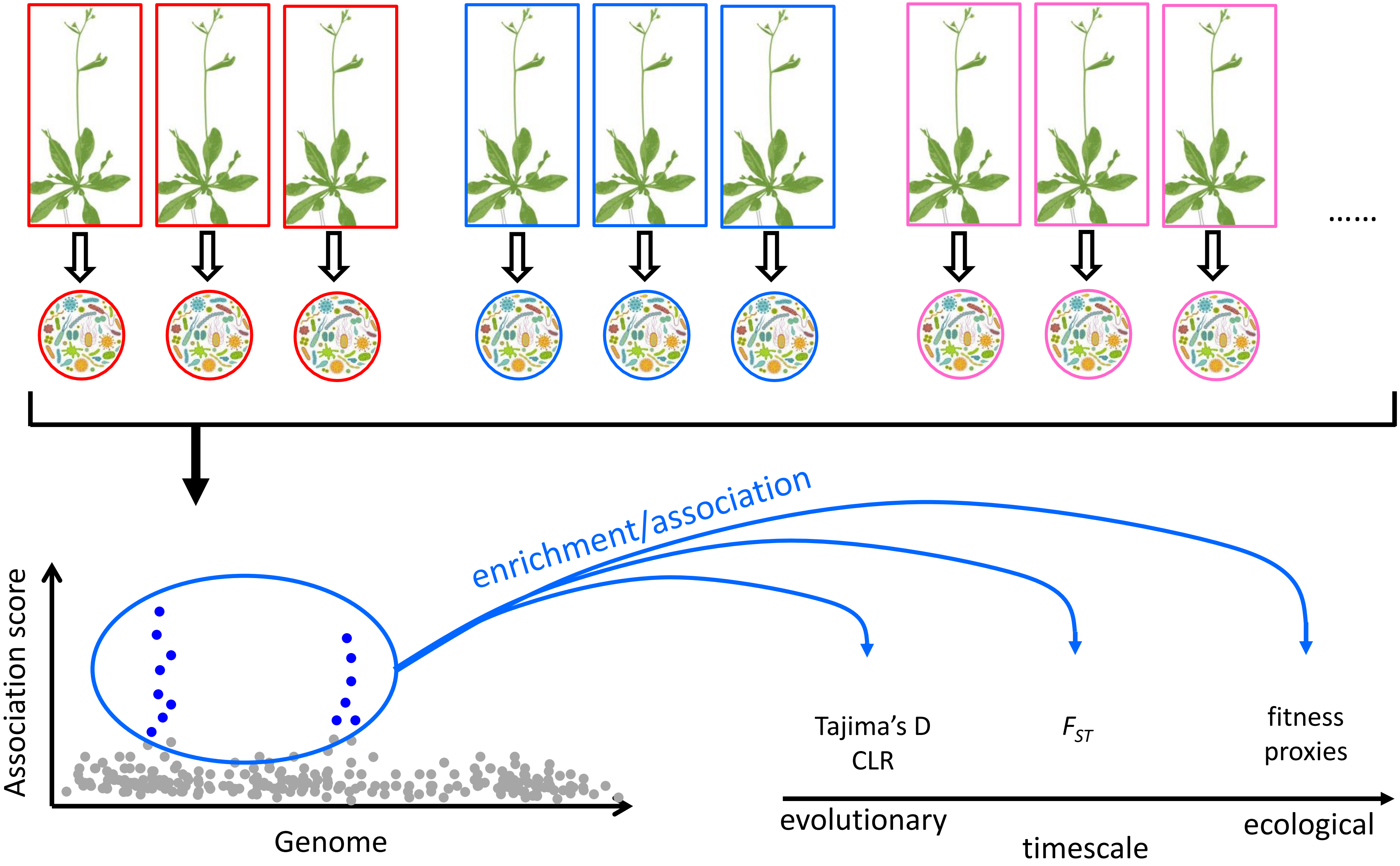
Plant genetic effects on microbial hubs impact host fitness in repeated field trials

Benjamin Brachi[a,b], Daniele Filiault[c,1], Hannah Whitehurst[a,1], Paul Darme[a], Pierre Le Gars[a], Marine Le Mentec[a], Timothy C. Morton[a], Envel Kerdaffrec[c], Fernando Rabanal[c], Alison Anastasio[a], Mathew S. Box[d], Susan Duncan[d], Feng Huang[a,e], Riley Leff[a], Polina Novikova[c], Matthew Perisin[a], Takashi Tsuchimatsu[c], Roderick Woolley[a], Caroline Dean[d], Magnus Nordborg[c], Svante Holm[f], and Joy Bergelson[a,g,2]

Bergelson, Brachi, Roux & Vailleau Curr Opin Biotech 2021

How to establish a genomic map of local adaptation?

I. Setting up of GWA study in a common garden



enrichment/association

Tajima's D
CLR

F_{ST}

fitness
proxies

evolutionary
timescale
ecological

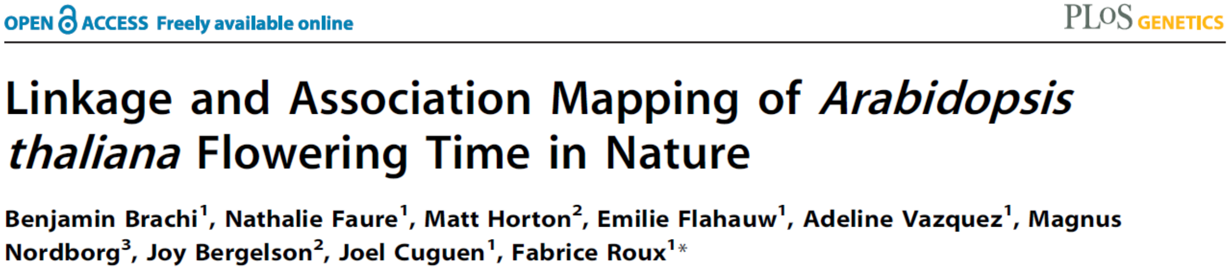
How to establish a genomic map of local adaptation?

I. Setting up of GWA study in a common garden

A. In which growing conditions?

VS

No common QTLs between greenhouse and more ecologically realistic conditions for the highly heritable phenotypic trait flowering time.



OPEN ACCESS Freely available online PLoS GENETICS

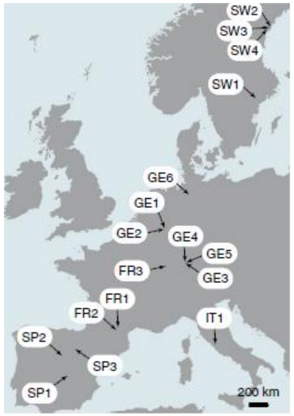
Linkage and Association Mapping of *Arabidopsis thaliana* Flowering Time in Nature

Benjamin Brachi[1], Nathalie Faure[1], Matt Horton[2], Emilie Flahauw[1], Adeline Vazquez[1], Magnus Nordborg[3], Joy Bergelson[2], Joel Cuguen[1], Fabrice Roux[1]*

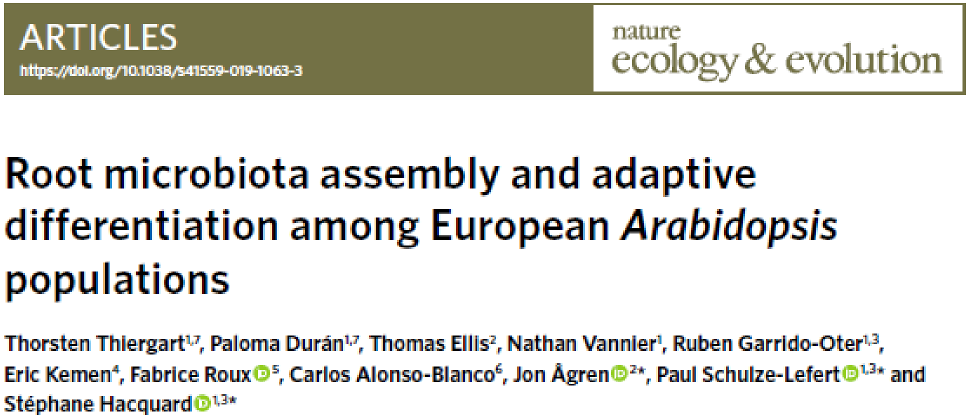
How to establish a genomic map of local adaptation?

I. Setting up of GWA study in a common garden

B. Using which soil/SynCom?



Strong geographic structuring of the soil biome at the European scale.



ARTICLES

https://doi.org/10.1038/s41559-019-1063-3

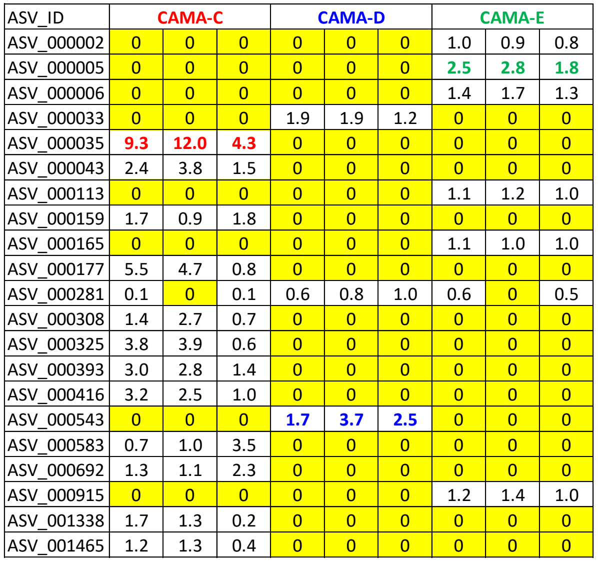
nature ecology & evolution

Root microbiota assembly and adaptive differentiation among European *Arabidopsis* populations

Thorsten Thiergart[1,7], Paloma Durán[1,7], Thomas Ellis[2], Nathan Vannier[1], Ruben Garrido-Oter[1,3], Eric Kemen[4], Fabrice Roux[5], Carlos Alonso-Blanco[6], Jon Ågren[2]*, Paul Schulze-Lefert[1,3]* and Stéphane Hacquard[1,3]*

Soil microbiota variation at the scale of few meters (168 populations – SW France)

Camarès



ASV_ID	CAMA-C			CAMA-D			CAMA-E		
ASV_000002	0	0	0	0	0	0	1.0	0.9	0.8
ASV_000005	0	0	0	0	0	0	**2.5**	**2.8**	**1.8**
ASV_000006	0	0	0	0	0	0	1.4	1.7	1.3
ASV_000033	0	0	0	1.9	1.9	1.2	0	0	0
ASV_000035	**9.3**	**12.0**	**4.3**	0	0	0	0	0	0
ASV_000043	2.4	3.8	1.5	0	0	0	0	0	0
ASV_000113	0	0	0	0	0	0	1.1	1.2	1.0
ASV_000159	1.7	0.9	1.8	0	0	0	0	0	0
ASV_000165	0	0	0	0	0	0	1.1	1.0	1.0
ASV_000177	5.5	4.7	0.8	0	0	0	0	0	0
ASV_000281	0.1	0	0.1	0.6	0.8	1.0	0.6	0	0.5
ASV_000308	1.4	2.7	0.7	0	0	0	0	0	0
ASV_000325	3.8	3.9	0.6	0	0	0	0	0	0
ASV_000393	3.0	2.8	1.4	0	0	0	0	0	0
ASV_000416	3.2	2.5	1.0	0	0	0	0	0	0
ASV_000543	0	0	0	**1.7**	**3.7**	**2.5**	0	0	0
ASV_000583	0.7	1.0	3.5	0	0	0	0	0	0
ASV_000692	1.3	1.1	2.3	0	0	0	0	0	0
ASV_000915	0	0	0	0	0	0	1.2	1.4	1.0
ASV_001338	1.7	1.3	0.2	0	0	0	0	0	0
ASV_001465	1.2	1.3	0.4	0	0	0	0	0	0

Mayjonade et al. (under revision)

How to establish a genomic map of local adaptation?

I. Setting up of GWA study in a common garden

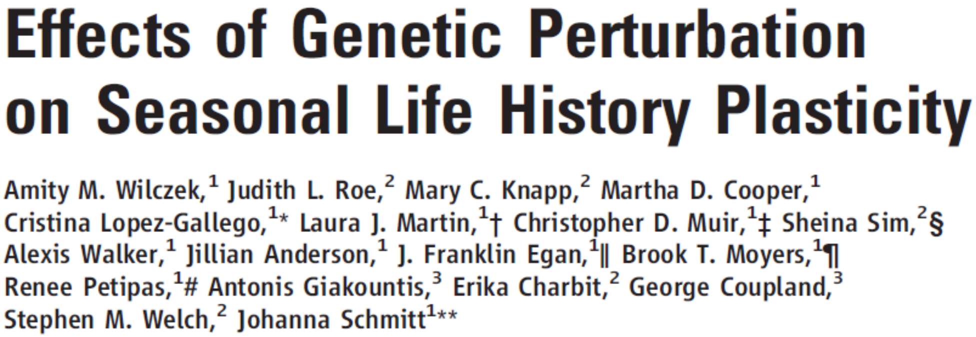
C. When planting seeds?

Main germination cohort


mid-Sept

late Sept

November


Effects of Genetic Perturbation on Seasonal Life History Plasticity

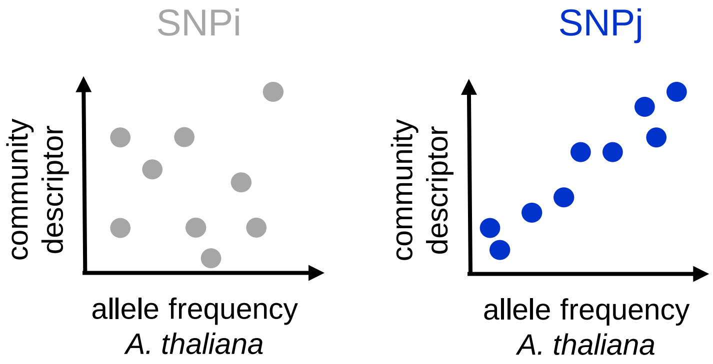
Amity M. Wilczek,[1] Judith L. Roe,[2] Mary C. Knapp,[2] Martha D. Cooper,[1]
Cristina Lopez-Gallego,[1]* Laura J. Martin,[1]† Christopher D. Muir,[1]‡ Sheina Sim,[2]§
Alexis Walker,[1] Jillian Anderson,[1] J. Franklin Egan,[1]‖ Brook T. Moyers,[1]¶
Renee Petipas,[1]# Antonis Giakountis,[3] Erika Charbit,[2] George Coupland,[3]
Stephen M. Welch,[2] Johanna Schmitt[1]**

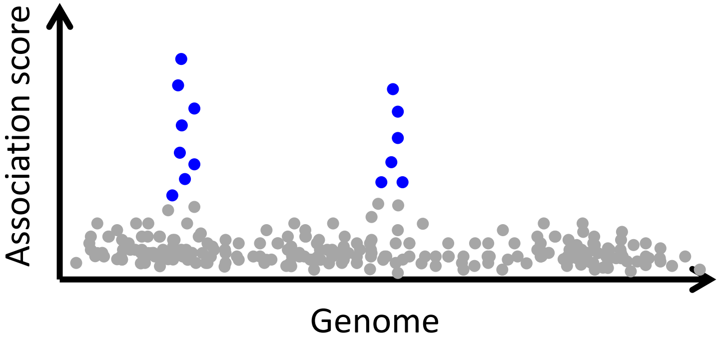
Science
AAAS

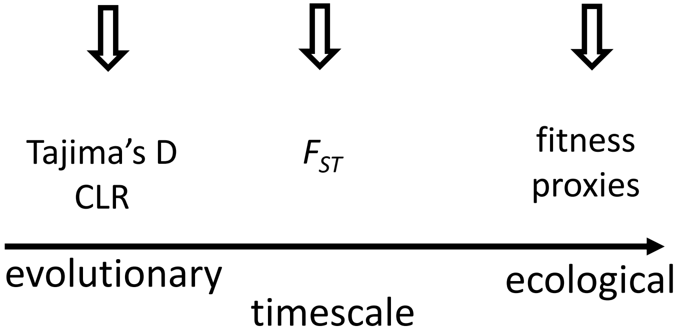

QTL effects can largely depend on sowing date.

How to establish a genomic map of local adaptation?

II. Setting up of Genome-Environment Association (GEA) study







Tajima's D
CLR

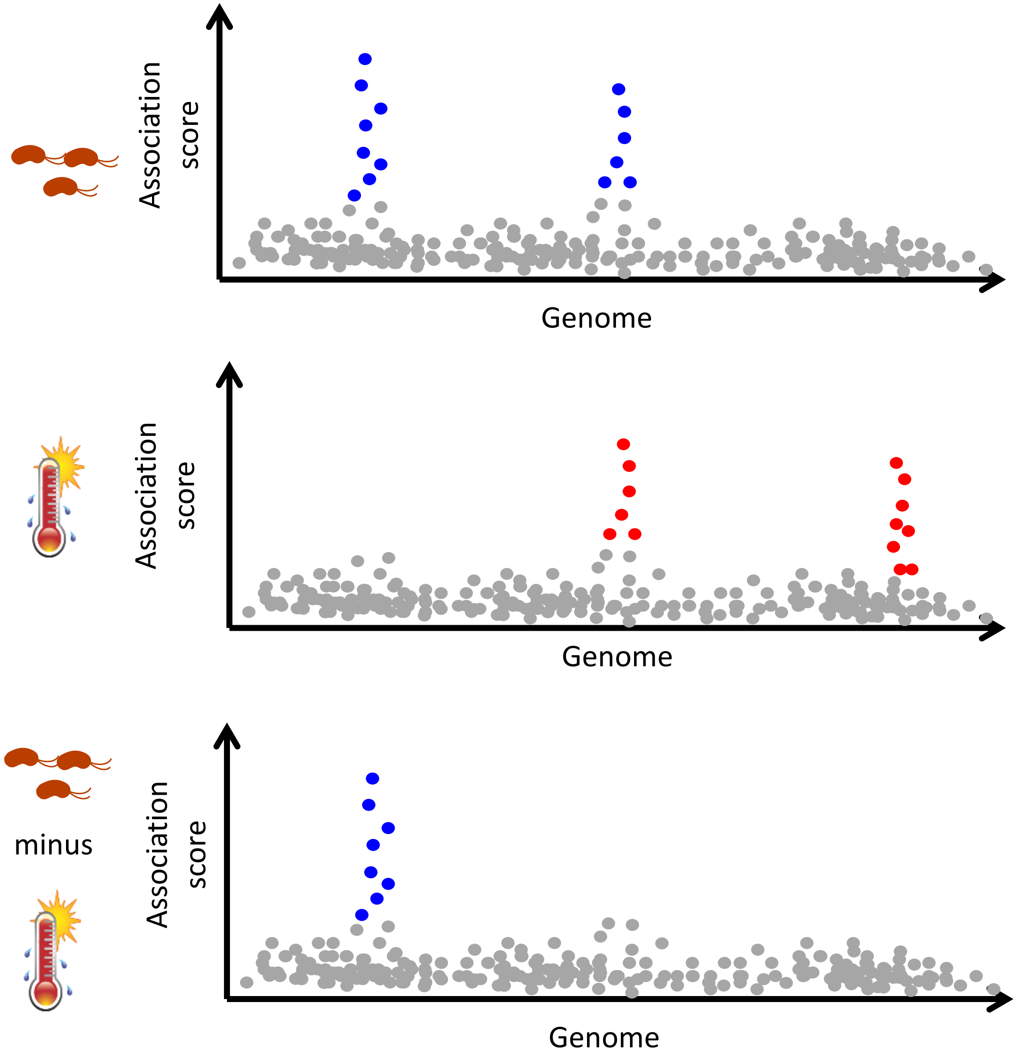
F_{ST}

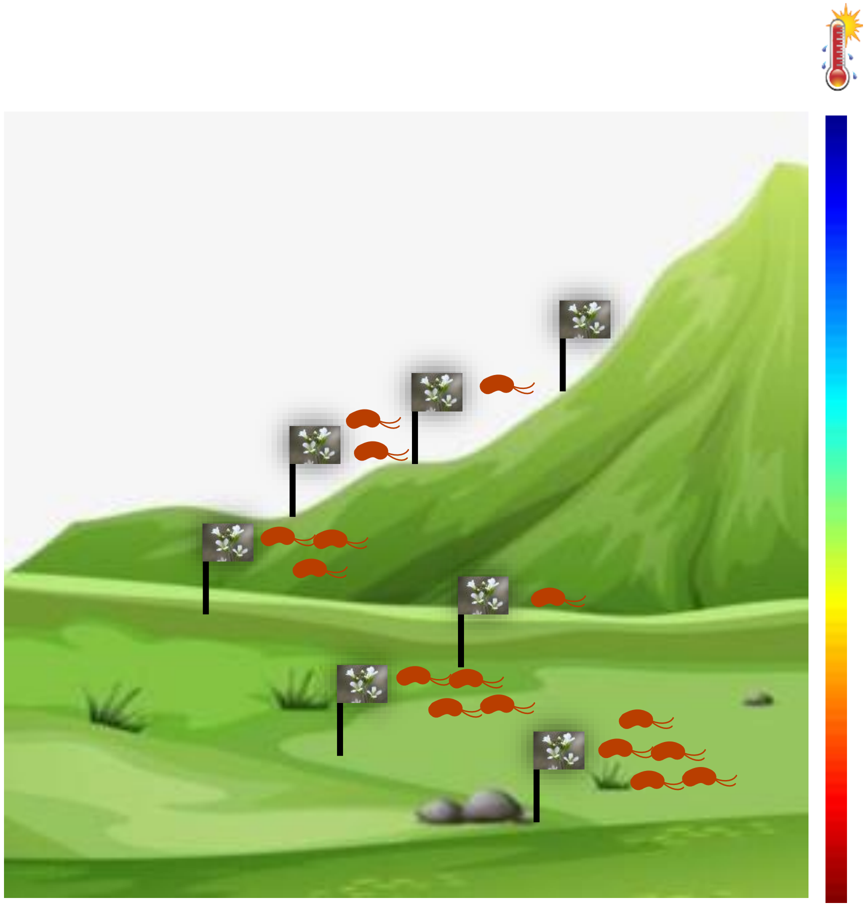
fitness proxies

evolutionary
timescale
ecological

How to establish a genomic map of local adaptation?

Main drawback of GEA: need for controlling for confounding ecological factors

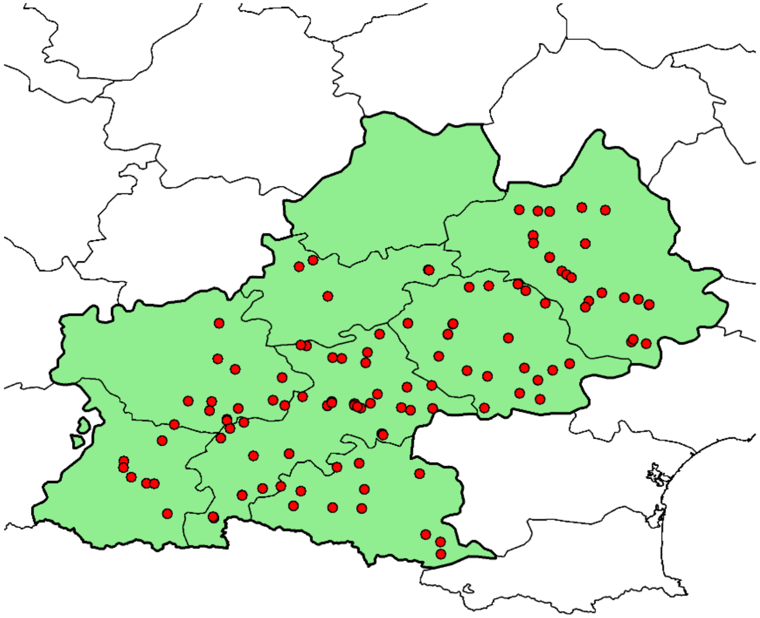




=> advantage: identification of ecological drivers

Setting up a Genome-Environment Association study

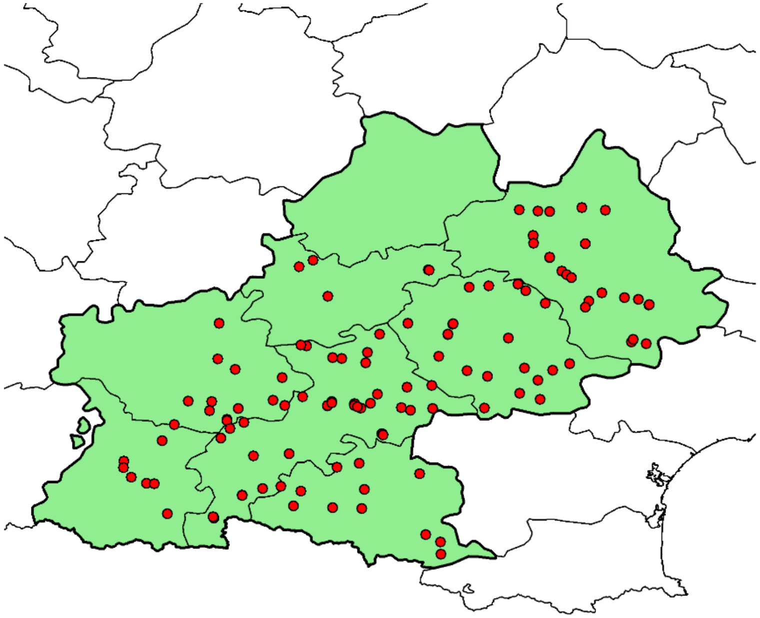
Working on 168 natural populations in the south-west of France



Large range of habitats

Setting up a Genome-Wide Association study

Working on 168 natural populations in the south-west of France



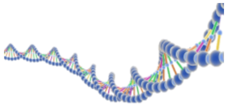
climate (n = 6)

soil (n = 14)

plant communities (n = 49)

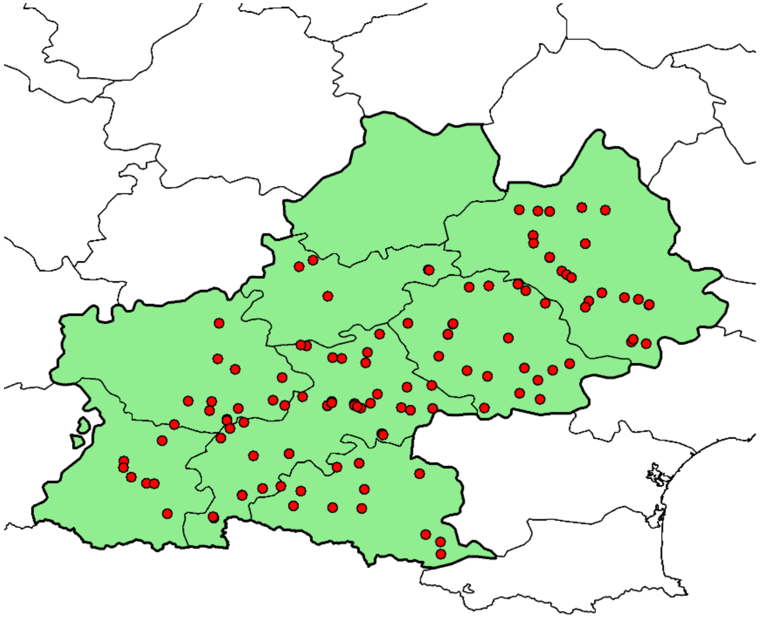
bacterial communities *(gyrB)*:
in situ, leaf & root (n = 194)

- PoolSeq approach: allele frequencies of ~ 4.8 million SNPs across the genome for each population => genomic information for 2,776 plants.
- Individual genome sequencing of 458 accessions: 5,386,423 bi-allelic SNPs



Setting up a Genome-Wide Association study

Working on 168 natural populations in the south-west of France



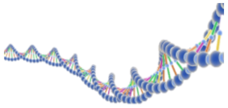
→ climate (n = 6)

→ soil (n = 14)

→ plant communities (n = 49)

→ bacterial communities *(gyrB)*:
in situ, leaf & root (n = 194)

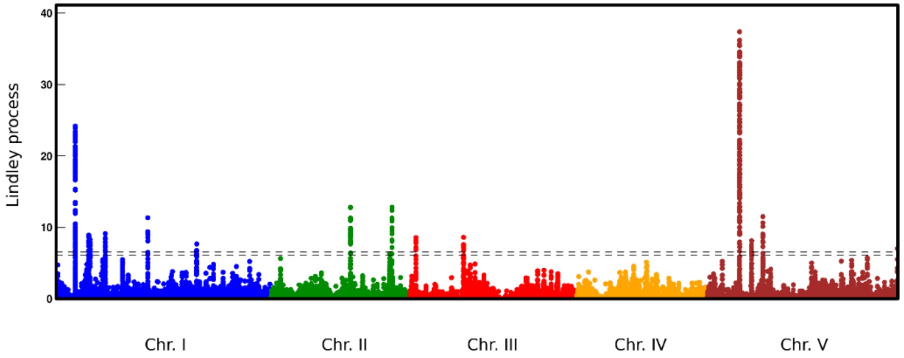
→
- PoolSeq approach: allele frequencies of ~ 4.8 million SNPs across the genome for each population => genomic information for 2,776 plants.
- Individual genome sequencing of 458 accessions: 5,386,423 bi-allelic SNPs



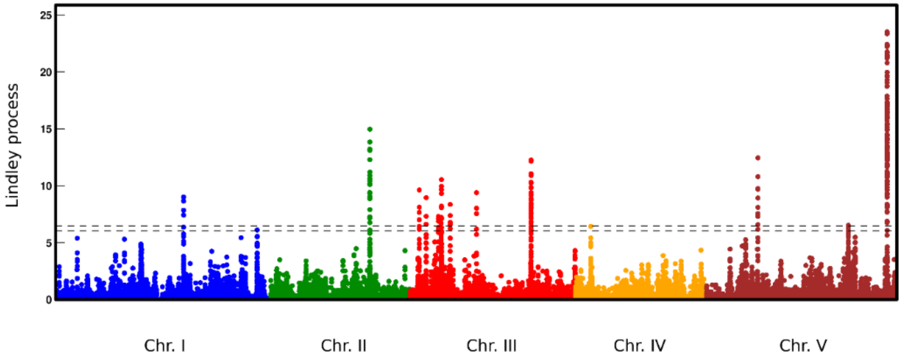
Detection of a polygenetic architecture associated with microbiota descriptors

Combining a Bayesian hierarchical model with a local score approach

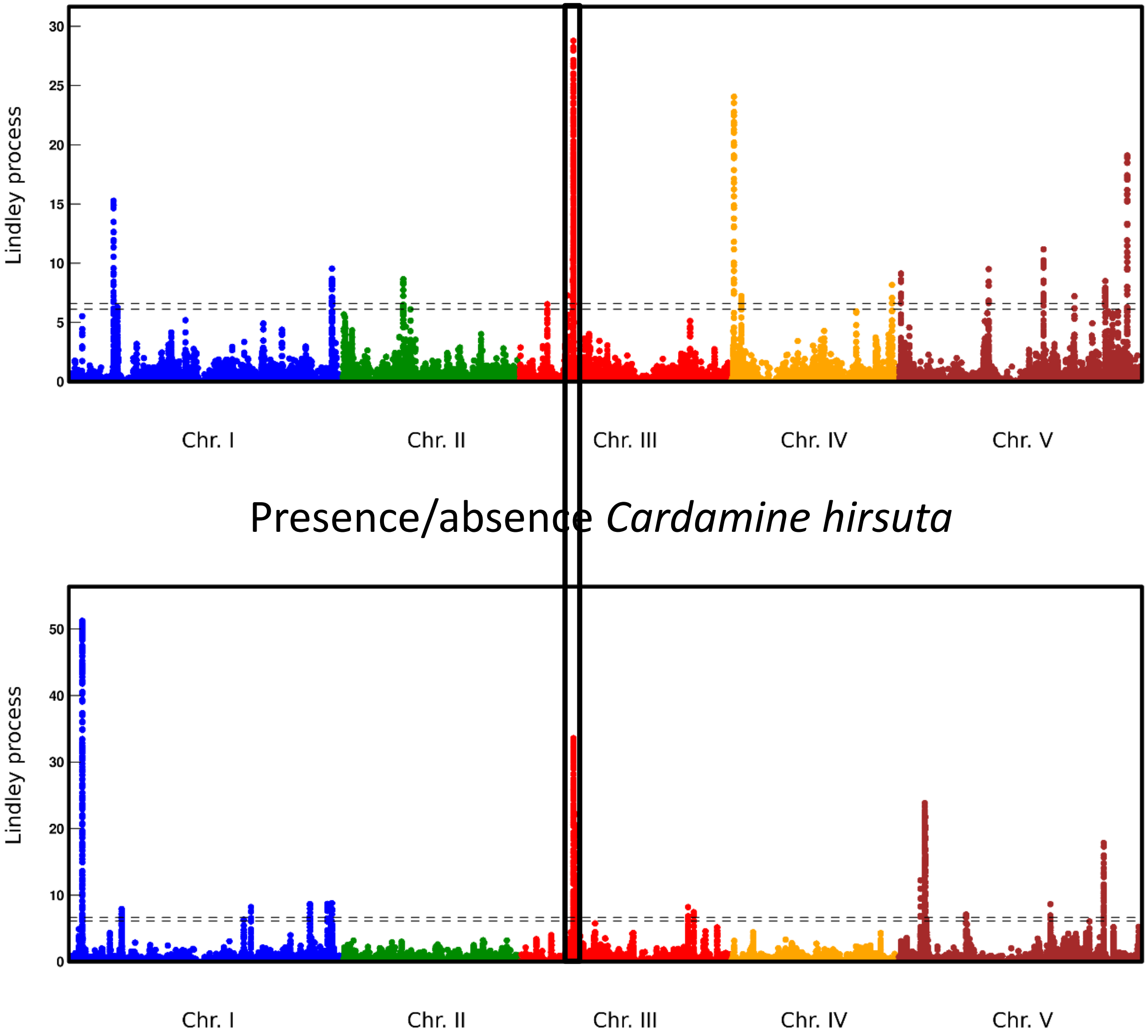
presence/absence OTU3



Microbiota composition (PCoA 2nd axis)



Common QTLs between microbiota descriptors and non-microbial ecological factors

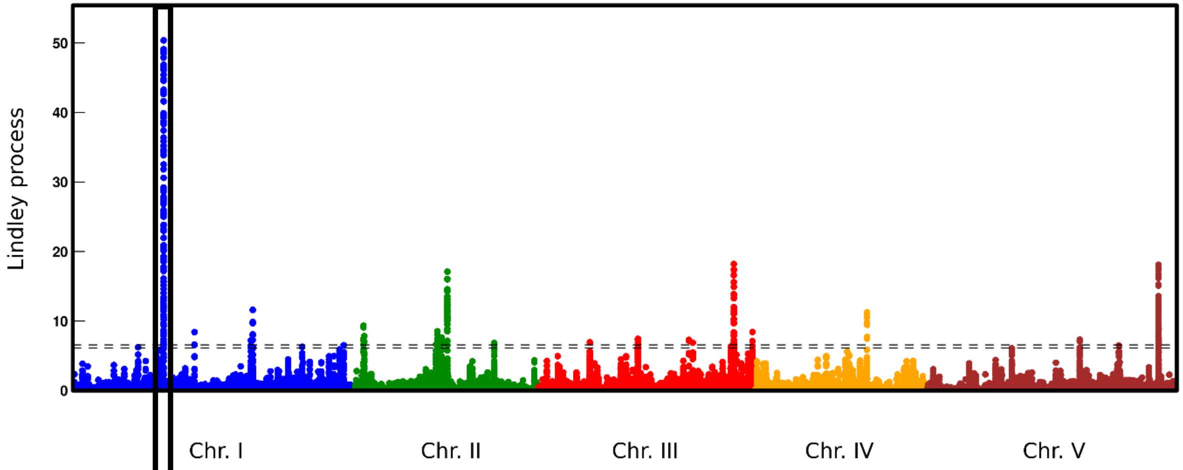


Presence/absence OTU6

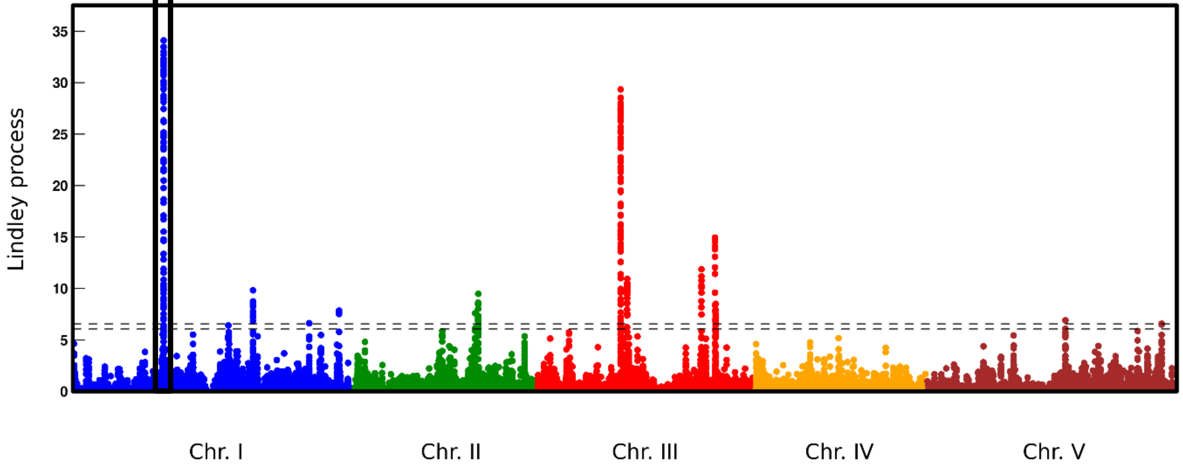
Presence/absence *Cardamine hirsuta*

Common QTLs between pathobiota descriptors and non-microbial ecological factors

Pathobiota composition (PCoA 2nd axis)



Plant community richness

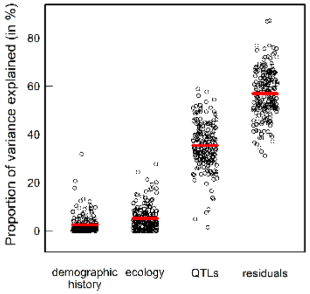


The Genetic Architecture of Adaptation to Leaf and Root Bacterial Microbiota in *Arabidopsis thaliana*

Fabrice Roux,*[1] Léa Frachon,[1,2] and Claudia Bartoli[1,3]

Molecular Biology & Evolution 2023

- A higher proportion of microbiota variance is explained by the cumulative effects of QTLs rather than the cumulative effects of ecological factors

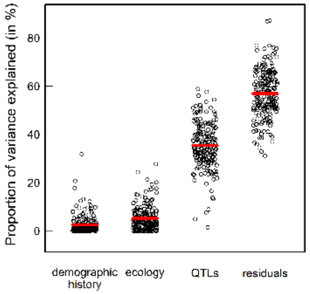


The Genetic Architecture of Adaptation to Leaf and Root Bacterial Microbiota in *Arabidopsis thaliana*

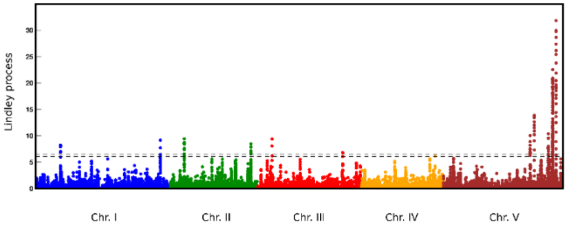
Fabrice Roux,*[1] Léa Frachon,[1,2] and Claudia Bartoli[1,3]

Molecular Biology & Evolution 2023

- A higher proportion of microbiota variance is explained by the cumulative effects of QTLs rather than the cumulative effects of ecological factors



- The candidate genes are highly dependent on the identity of microbiota - pathobiota descriptors

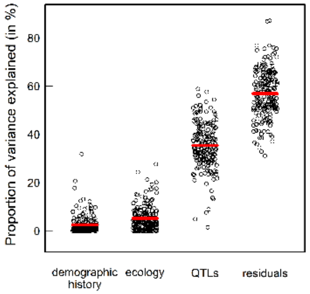


The Genetic Architecture of Adaptation to Leaf and Root Bacterial Microbiota in *Arabidopsis thaliana*

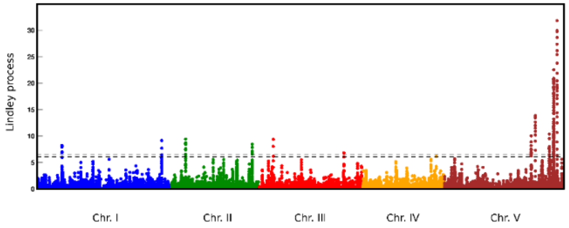
Fabrice Roux,*[1] Léa Frachon,[1,2] and Claudia Bartoli[1,3]

Molecular Biology & Evolution 2023

- A higher proportion of microbiota variance is explained by the cumulative effects of QTLs rather than the cumulative effects of ecological factors



- The candidate genes are highly dependent on the identity of microbiota - pathobiota descriptors



- Stronger signatures of local adaptation on candidate genes for root microbiota than on candidate genes for leaf microbiota

→ *ATMYB15*

nature plants

ARTICLES

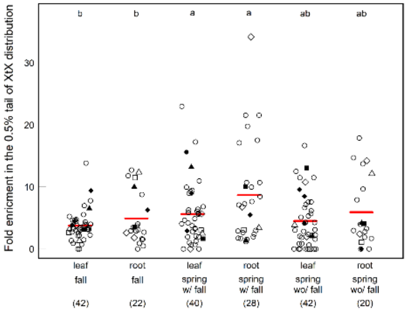
https://doi.org/10.1038/s41477-021-00920-2

Check for updates

OPEN

Coordination of microbe–host homeostasis by crosstalk with plant innate immunity

Ka-Wai Ma[1,6], Yulong Niu[1,6], Yong Jia[2,6], Jana Ordon[1], Charles Copeland[1], Aurélia Emonet[3], Niko Geldner[3], Rui Guan[1], Sara Christina Stolze[4], Hirofumi Nakagami[4], Ruben Garrido-Oter[1,5] ✉ and Paul Schulze-Lefert[1,5] ✉

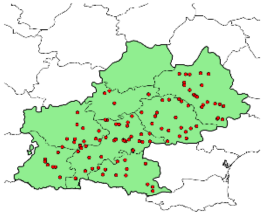


Significant overlapping with candidate genes from GWAS

GEA in south-west

of France

in situ characterization
microbiota - *gyrB*



Roux *et al.* MBE 2023

GWAS in Sweden

native habitats

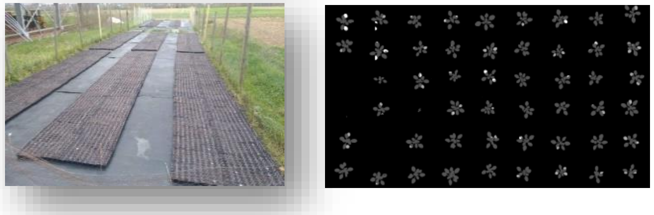
common garden
microbiota – 16S

Brachi *et al.* PNAS 2022

GWAS in the south-west of France

field conditions

common garden
response to 13 commensal strains



Ramirez-Sanchez, Duflos *et al.* Phytobiomes 2024
Duflos *et al.* in prep

Daniela Ramirez-Sanchez

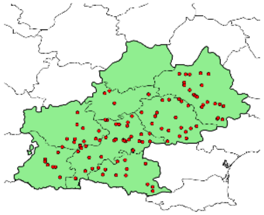
Rémi Duflos

Significant overlapping with candidate genes from GWAS

GEA in south-west

of France

in situ characterization
microbiota - *gyrB*



Roux *et al.* MBE 2023

GWAS in Sweden

native habitats

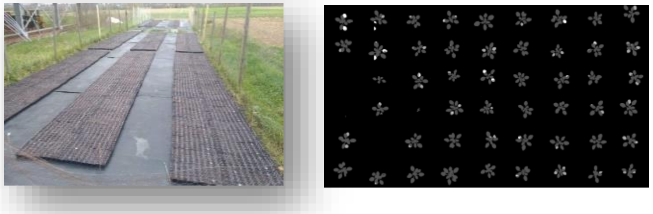
common garden
microbiota – 16S

Brachi *et al.* PNAS 2022

GWAS in the south-west of France

field conditions

common garden
response to 13 commensal strains



Ramirez-Sanchez, Duflos *et al.* bioRxiv 2022

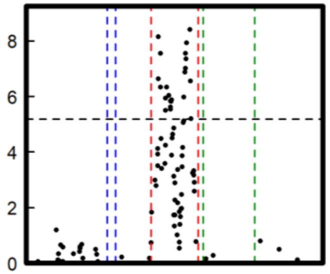
Short list of 50 genes common between GWAS and GEA

At5g46330 FLAGELLIN-SENSITIVE 2 (*FLS2*) with 39 top SNPs identified by GEA

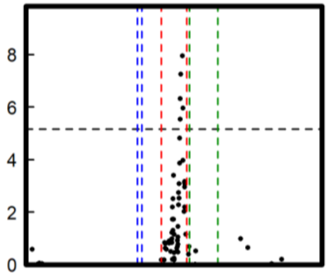
Significant overlapping with candidate genes from GWAS

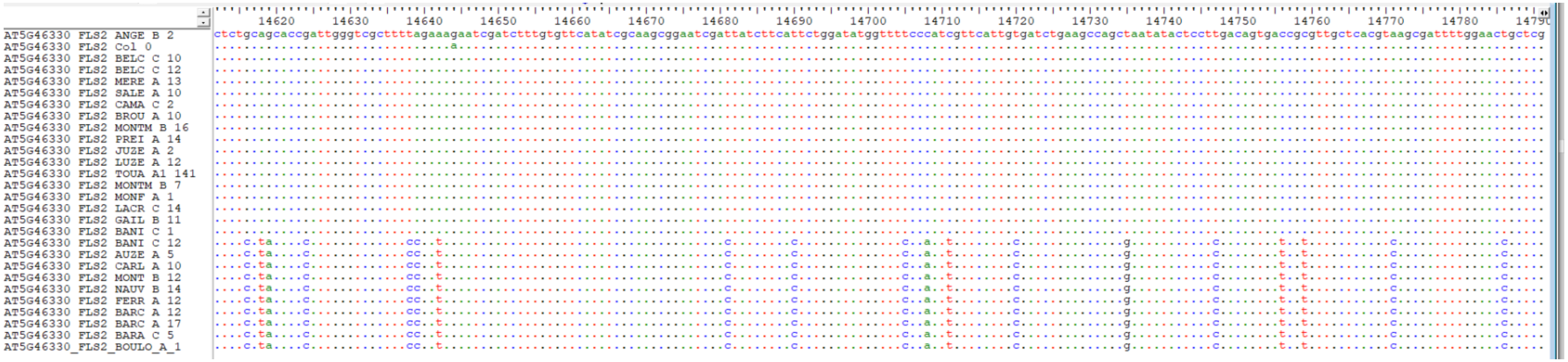
At5g46330 FLAGELLIN-SENSITIVE 2 (*FLS2*) with 39 top SNPs identified by GEA

Oxalobacteraceae bacterium

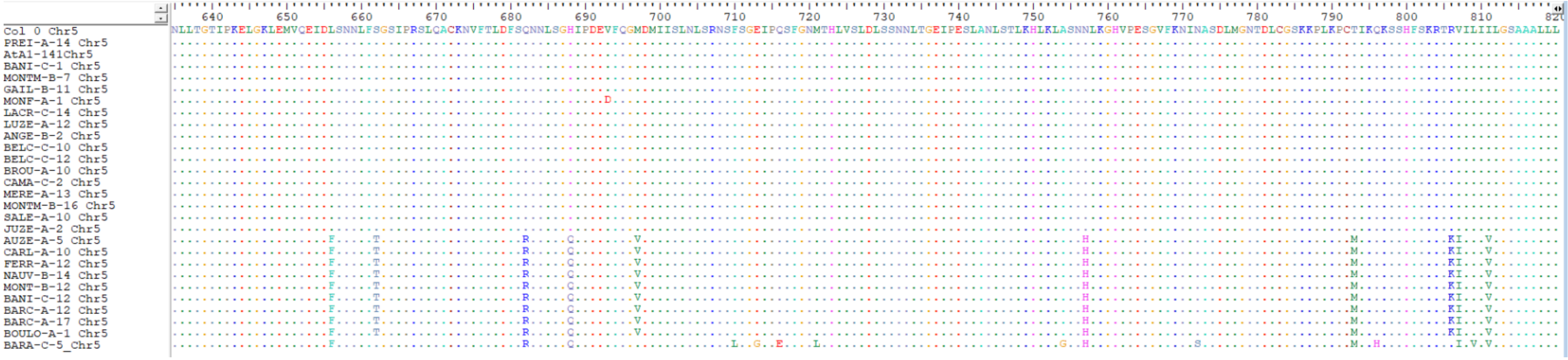


Methylobacterium sp.





```
14620 14630 14640 14650 14660 14670 14680 14690 14700 14710 14720 14730 14740 14750 14760 14770 14780 14790
AT5G46330 FLS2 ANGE B 2      ctctgcagcaccgattgggtcgcttttagaaagaatcgatctttgtgttcatatcgcaagcggaatcgattatcttcattctggatatggttttcccatcgttcattgtgatctgaagccagctaatatactccttgacagtgaccgcgttgctcacgtaagcgattttggaactgctcg
AT5G46330 FLS2 Col 0
AT5G46330 FLS2 BELC C 10
AT5G46330 FLS2 BELC C 12
AT5G46330 FLS2 MERE A 13
AT5G46330 FLS2 SALE A 10
AT5G46330 FLS2 CAMA C 2
AT5G46330 FLS2 BROU A 10
AT5G46330 FLS2 MONTM B 16
AT5G46330 FLS2 PREI A 14
AT5G46330 FLS2 JUZE A 2
AT5G46330 FLS2 LUZE A 12
AT5G46330 FLS2 TOUA A1 141
AT5G46330 FLS2 MONTM B 7
AT5G46330 FLS2 MONF A 1
AT5G46330 FLS2 LACR C 14
AT5G46330 FLS2 GAIL B 11
AT5G46330 FLS2 BANI C 1
AT5G46330 FLS2 BANI C 12
AT5G46330 FLS2 AUZE A 5
AT5G46330 FLS2 CARL A 10
AT5G46330 FLS2 MONT B 12
AT5G46330 FLS2 NAUV B 14
AT5G46330 FLS2 FERR A 12
AT5G46330 FLS2 BARC A 12
AT5G46330 FLS2 BARC A 17
AT5G46330 FLS2 BARA C 5
AT5G46330_FLS2_BOULO_A_1
```

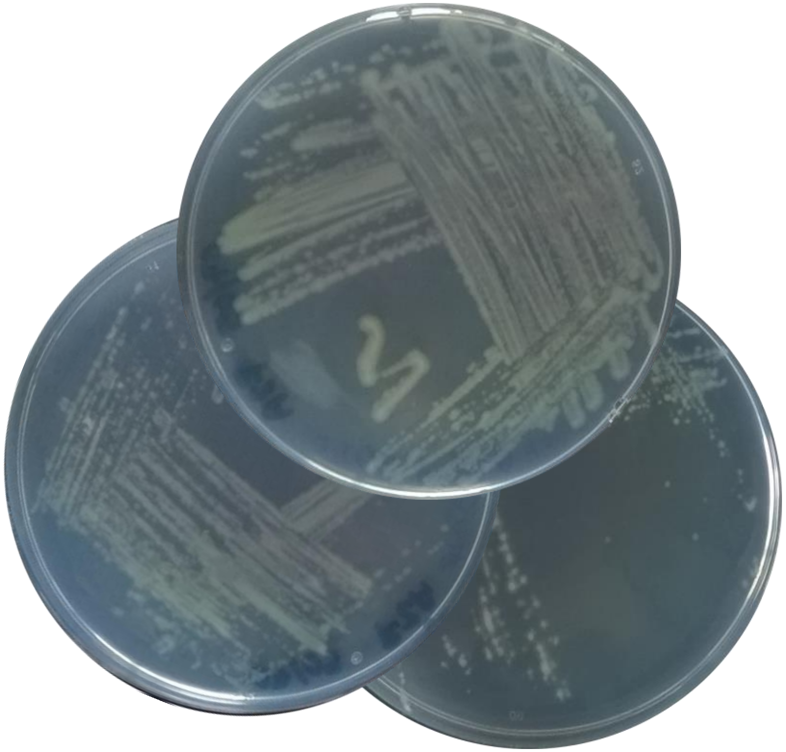


```
640 650 660 670 680 690 700 710 720 730 740 750 760 770 780 790 800 810 820
Col 0 Chr5        NLLTGTIPKELGKLEMVCEIDLSNNLFSGSIPRSLQACKNVFTLDFSQNNLSGHIPDEVFQGMDMIISLNLSRNSFSGEIPQSFGNMTHLVSLDLSSNNLTGEIPESLANLSTLKHLKLASNNLKGHVPESGVFKNINASDLMGNTDLCGSKKPLKPCTIKQKSSHFSKRTRVILIILGSAAALLL
PREI-A-14 Chr5
AtA1-141Chr5
BANI-C-1 Chr5
MONTM-B-7 Chr5
GAIL-B-11 Chr5
MONF-A-1 Chr5
LACR-C-14 Chr5
LUZE-A-12 Chr5
ANGE-B-2 Chr5
BELC-C-10 Chr5
BELC-C-12 Chr5
BROU-A-10 Chr5
CAMA-C-2 Chr5
MERE-A-13 Chr5
MONTM-B-16 Chr5
SALE-A-10 Chr5
JUZE-A-2 Chr5
AUZE-A-5 Chr5
CARL-A-10 Chr5
FERR-A-12 Chr5
NAUV-B-14 Chr5
MONT-B-12 Chr5
BANI-C-12 Chr5
BARC-A-12 Chr5
BARC-A-17 Chr5
BOULO-A-1 Chr5
BARA-C-5_Chr5
```

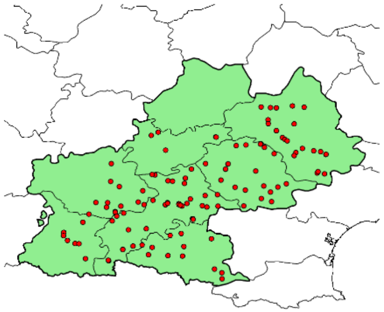
Wlodzimierz, Rabanal, Burns et al. Nature 2023

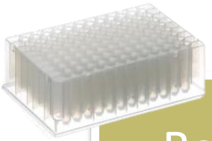
Genetic effects on plant-microbiota interactions

Exploring natural genetic variation of commensal bacteria?



Identifying genes at the interspecific and intraspecific levels





Bacterial collection based on a CBC approach (> 7,000 strains)

Interspecific level

- 942 bacterial species
- Genome & epigenome Nanopore sequencing (>90%)
- Development of a GWA mapping method adapted to interspecific genomic diversity

Intraspecific level

- 74 strains of *Pseudomonas siliginis*
- PacBio genome sequencing

Rémi Duflos

Why *Pseudomonas siliginis*?

- 6th most abundant and prevalent OTU across 168 natural populations of *A. thaliana* (Bartoli et al. ISME 2018)
- A biostimulant effect on *A. thaliana* (Ramirez-Sanchez et al. Front. Microbiol. 2022)
- Used in a biostimulant associated with another bacteria: NUELLO® iN

Pseudomonas siliginis
A strong atmospheric nitrogen fixing endophyte bacterial strain

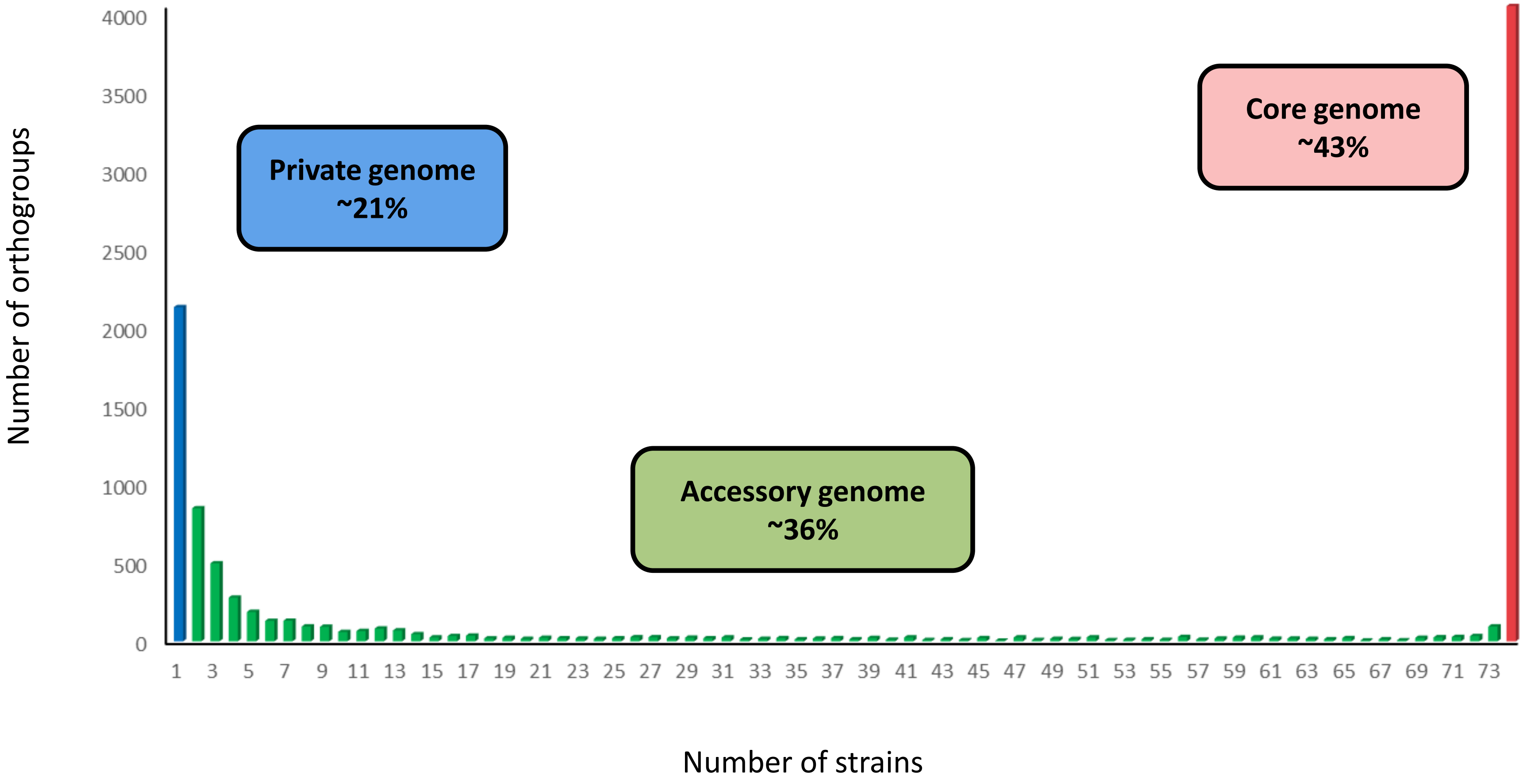
NUELLO iN

Curtobacterium salicaceae
A strong atmospheric nitrogen fixing and soil insoluble phosphate mobilising endophyte bacterial strain

- Potential biocontrol agent (Bartoli et al. ISME 2018)

Extensive genomic plasticity within *P. siliginis*

Identification of 10,189 genes

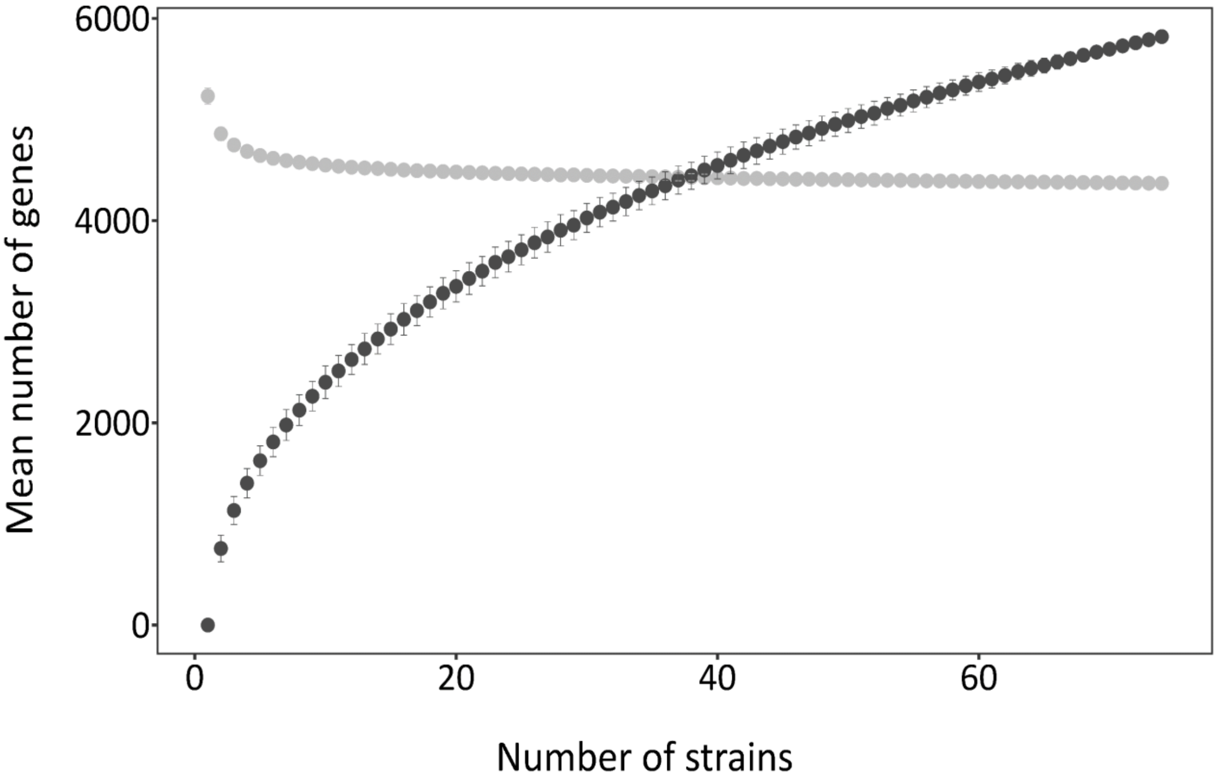


Private genome
~21%

Accessory genome
~36%

Core genome
~43%

Absence of saturation of the pan-genome size



Dispensable genome

(accessory and private)

Core genome

A very high density of genetic markers

Mapped on the longest genome of our collection by LIPME BioInformatics

Considering only bi-allelic positions on the core genome

→ Complete matrix : 656,256 genetic markers

→ Around 1 SNP every 9 bp

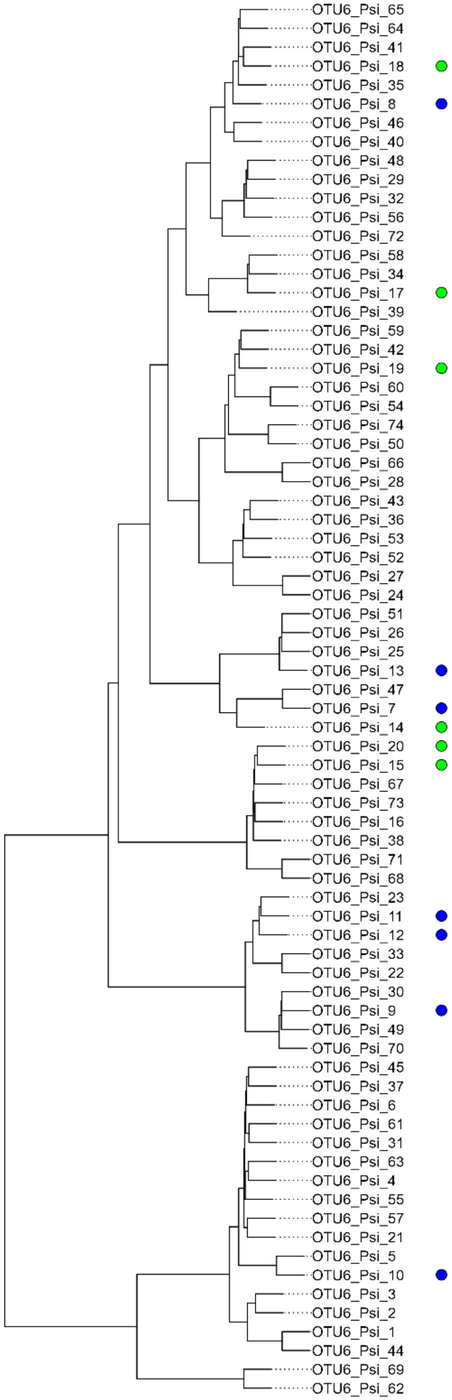
Small effect of the demographic history

Weak link between phylogenetic distance

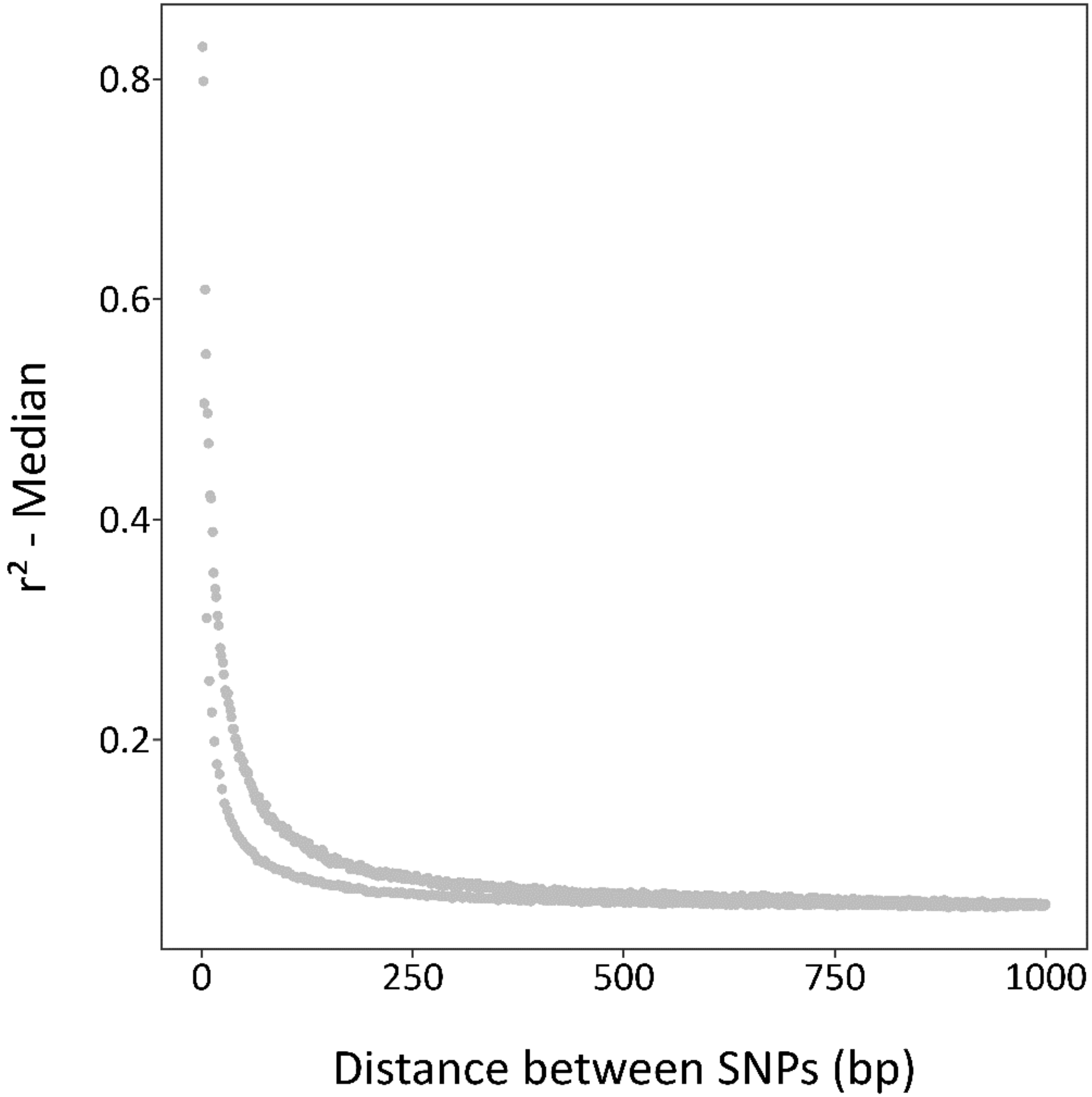
and geographic distance

● AMBR-A

● ANGE-B

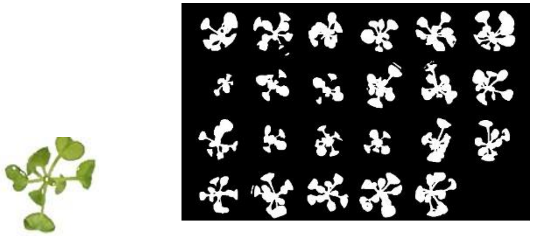


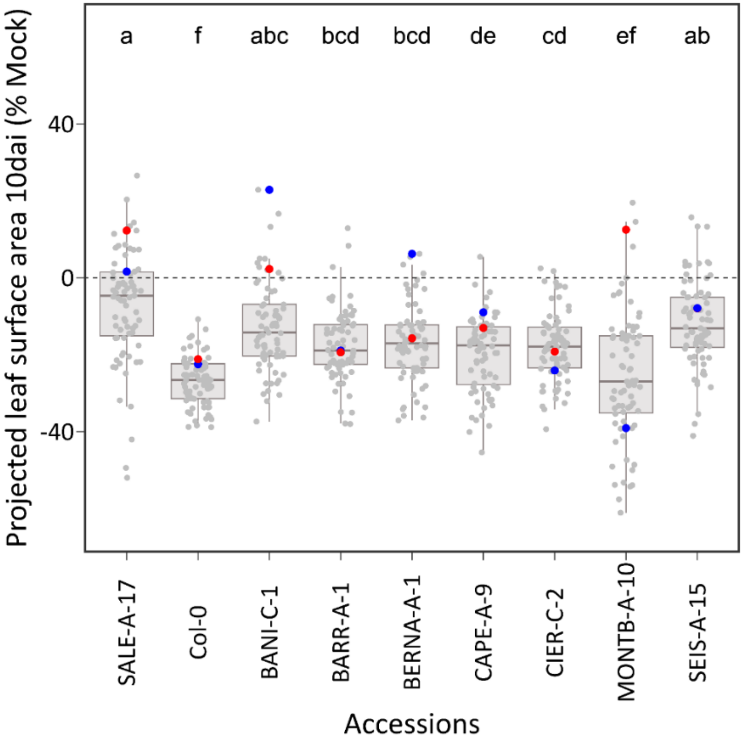
The extent of linkage disequilibrium is really short



The strain effect depends on the genotype of *A. thaliana*

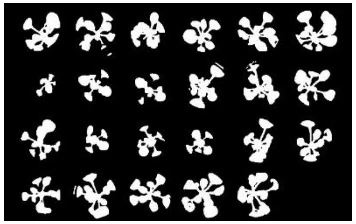
Anaïs BOTELLO
Master student

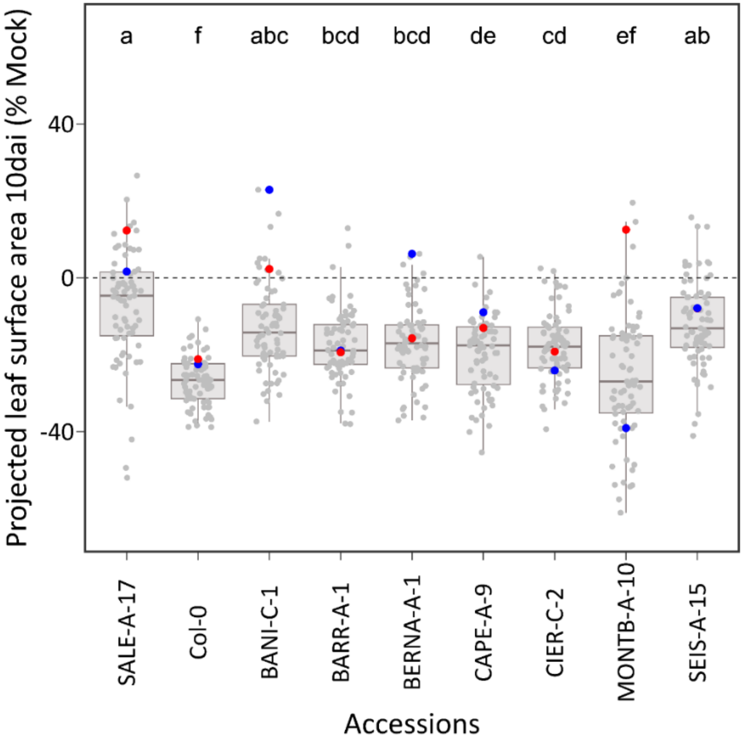




Most strains have a negative effect on plant growth

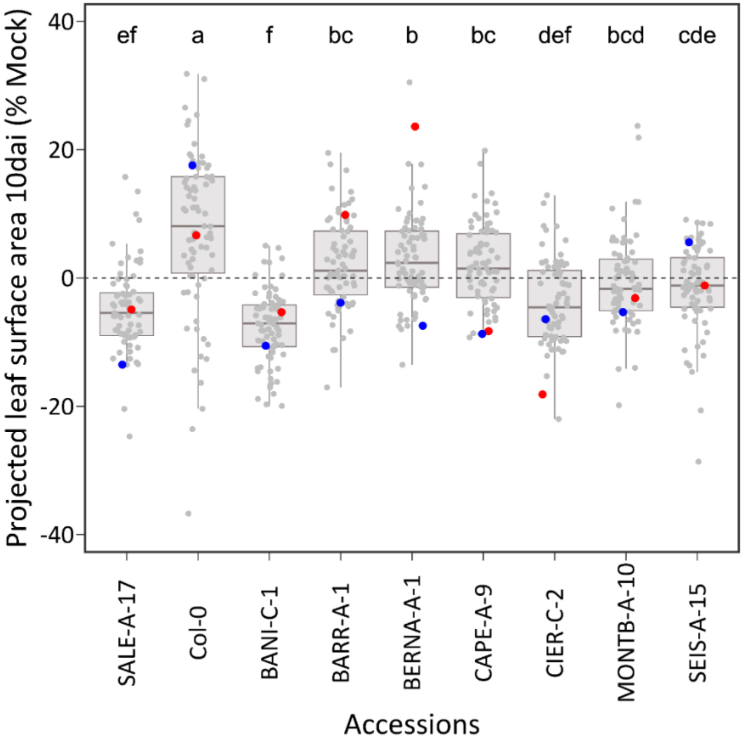
The strain effect depends on *A. thaliana* genotype & the growth conditions





Most strains have a negative effect on plant growth

Anaïs BOTELLO
Master student

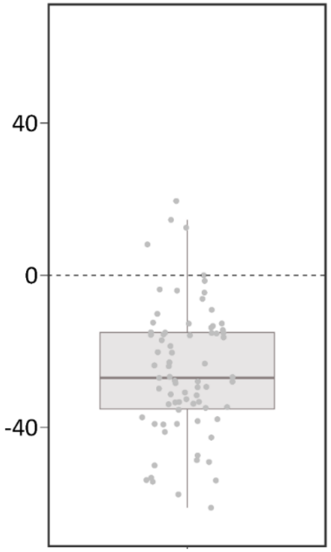


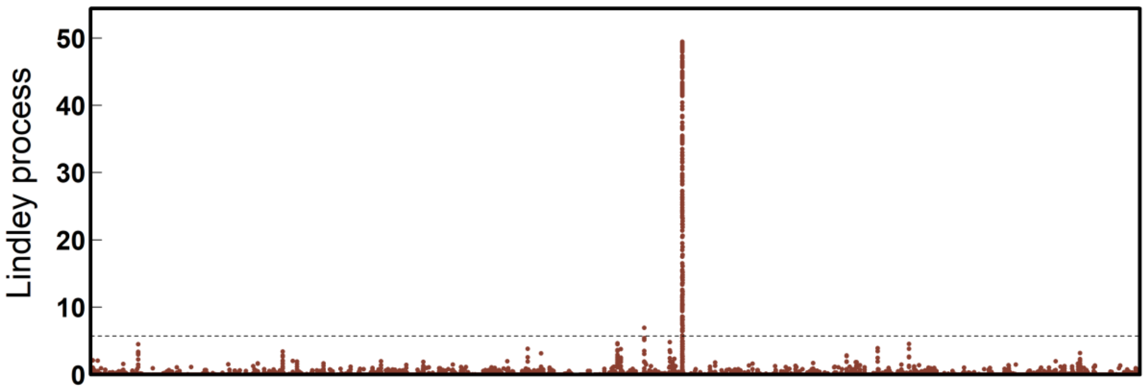
Contrasted effect among the nine accessions

Successful detection of QTLs of the effect of *P. siliginis*

GWA mapping: **mixed model** correcting for the effect of the demographic history +

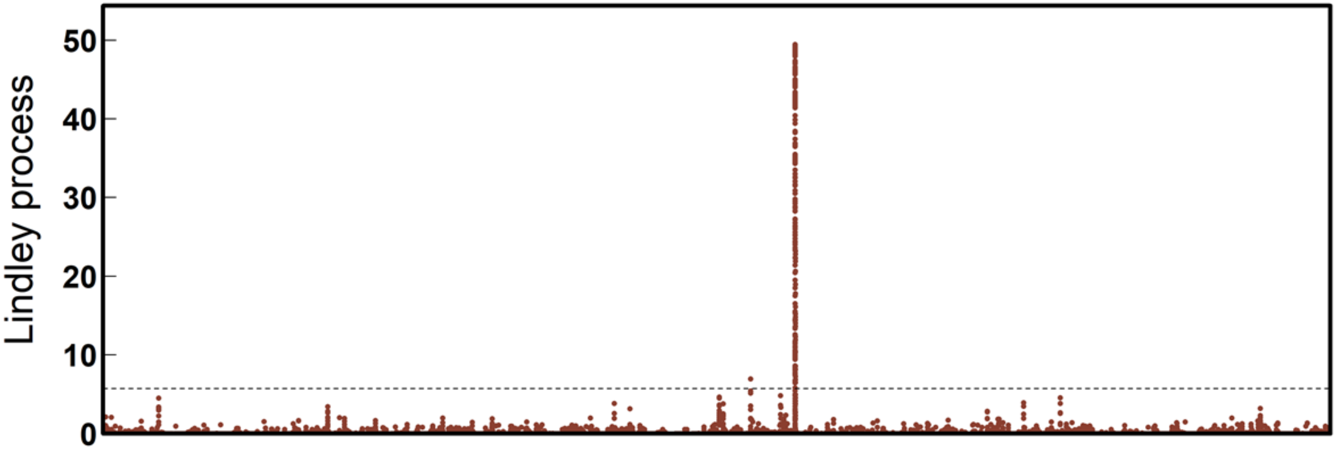
local score approach

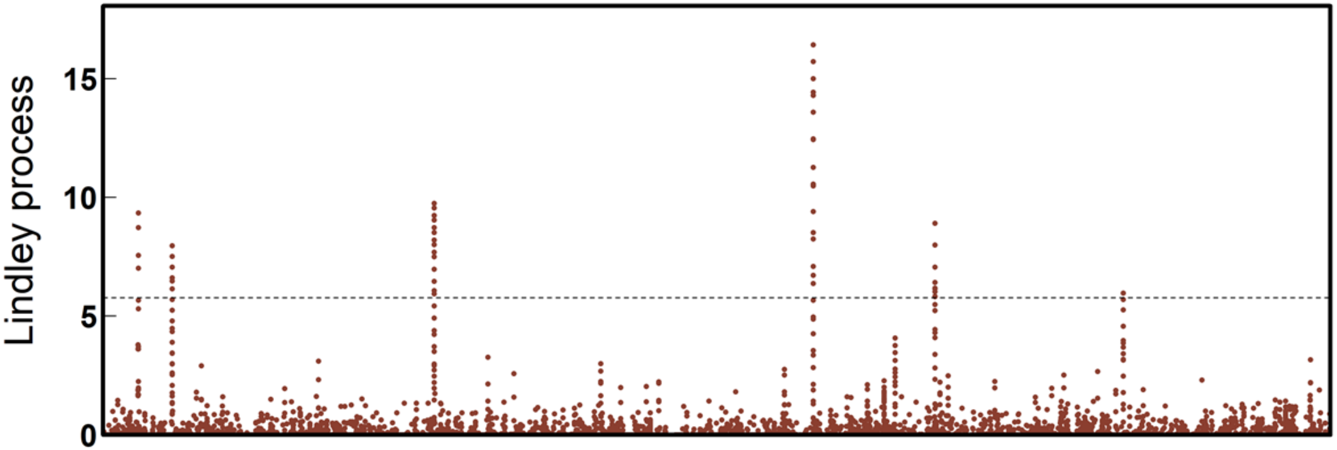




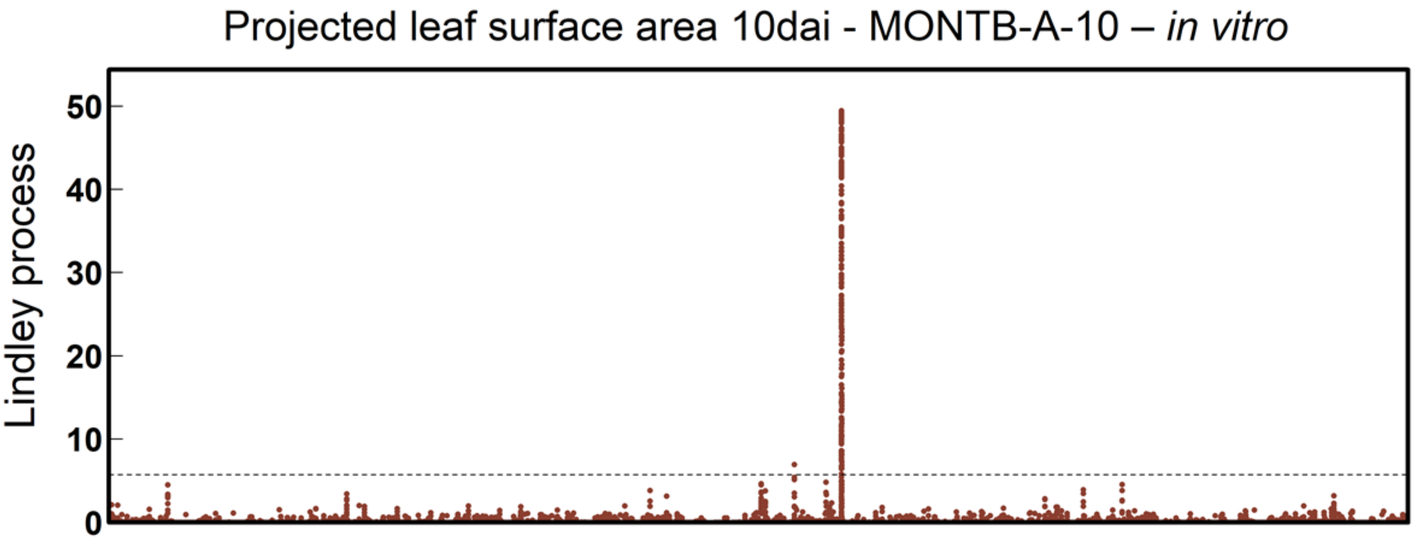
Projected leaf surface area 10dai - MONTB-A-10 – *in vitro*

The genetic architecture depends on the genotype of *A. thaliana*

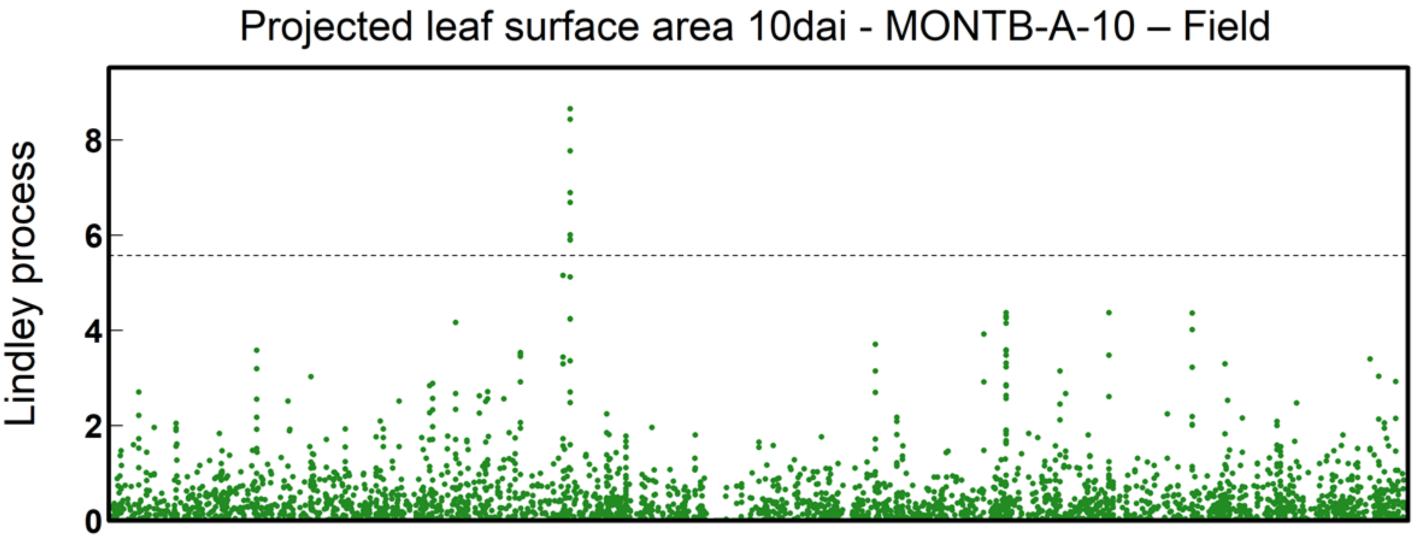

Projected leaf surface area 10dai - MONTB-A-10 – *in vitro*


Projected leaf surface area 10dai - Col-0 – *in vitro*

The genetic architecture depends on the growth conditions



Projected leaf surface area 10dai - MONTB-A-10 – *in vitro*



Projected leaf surface area 10dai - MONTB-A-10 – Field

Main conclusions

Host genotype

Identification of adaptive QTLs for many descriptors of bacterial communities

↓

diversity and composition

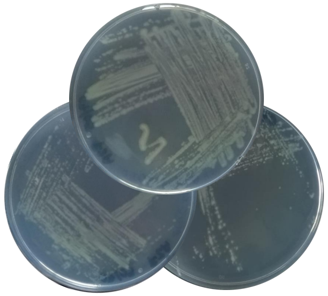
↓

Benefits of exploring biotic diffuse interactions and considering higher-order interactions

Similar conclusions reached on plant-plant interactions

(Baron *et al.* Funct Ecol 2015, Frachon *et al.* MBE 2019, Libourel *et al.* Front Plant Sci 2021, The international PLANTCOM network Trends Plant Sci 2023)

Main conclusions



Microbial genotype

Genomic variation within a commensal bacterial species + short LD + phenotypic variation

↓

Successful GWA mapping

↓

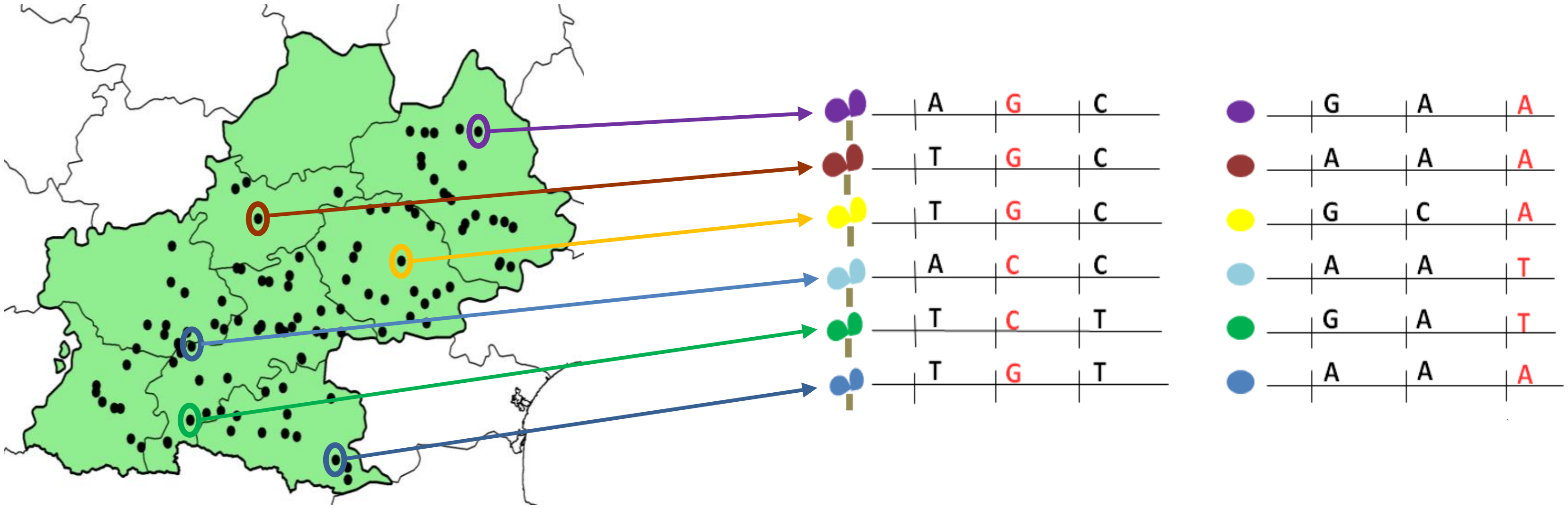
The genetic architecture of the effects of *P. siliginis* depends on the growth conditions and the genotype of *A. thaliana*

↓

Signatures of selection?

Performing free-phenotyping co-GWAS

Identifying genetic polymorphisms in strong LD across paired genomic data across space.



Innovative breeding schemes for plant-microbiota co-selection?

Bartoli & Roux *et al.* 2017 FiPS

Acknowledgments

erc
European Research Council
Established by the European Commission
Synergy Grants

INRAE

LabEx TULIP

CONACYT
Consejo Nacional de Ciencia y Tecnología

ECOGEN team

Botello Anaïs

Chalas Corentin

Duflos Rémi

Duran Paloma

Gabinaud Béatrice

Gibelin-Viala Chrystel

Falgous Léonie

Folletti Tifaine

Fuertes Coralie

Hanemian Mathieu

Huard-Chauveau Carine

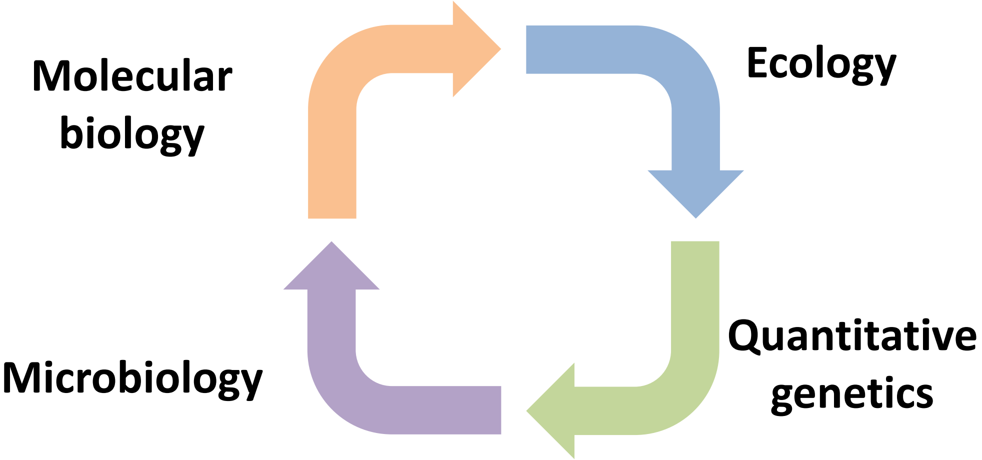
Lion Roxane

Mayjonade Baptiste

Fabrice Roux 500 QS

Silva Sonia

Vailleau Fabienne



Molecular biology

Ecology

Microbiology

Quantitative genetics

How to predict complex biotic interactions?

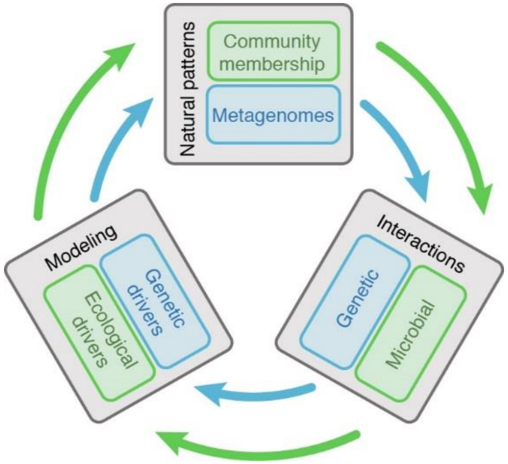
ERC Synergy ‘Understanding and predicting pathogen communities’

PATHOCOM

erc

Detlef
Weigel

Joy
Bergelson

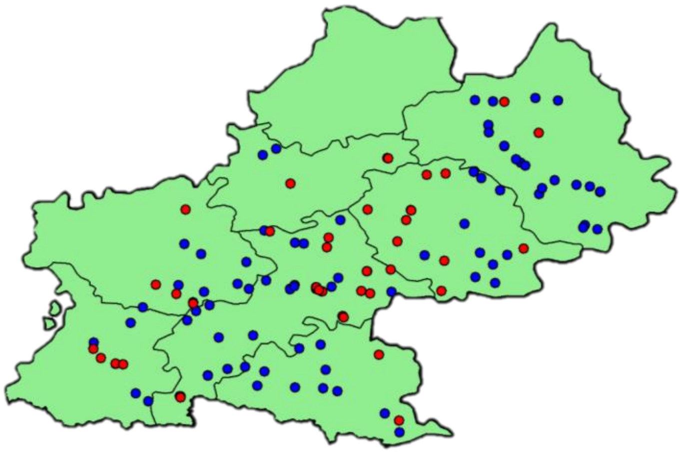


Three 3-year post-doc positions starting in April 2025:

- Testing the effects of microbiota genetics on pathogen-pathogen interactions
- Testing the effects of host genetics on pathogen-pathogen interactions (N = 2)

Performing GEA analyses on *P. siliginis*

74 strains from 45 populations

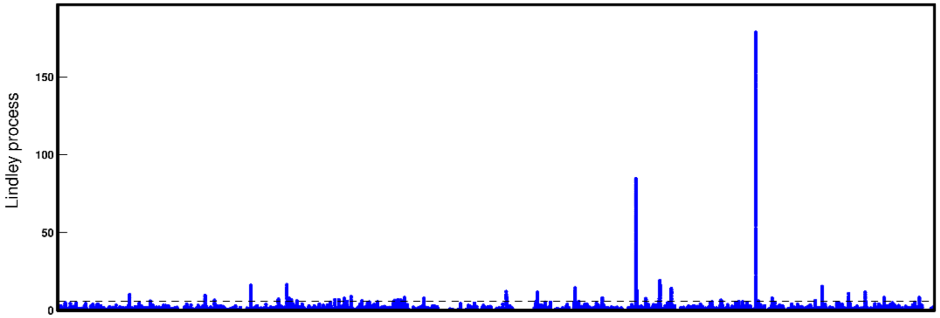


climate (n = 6)

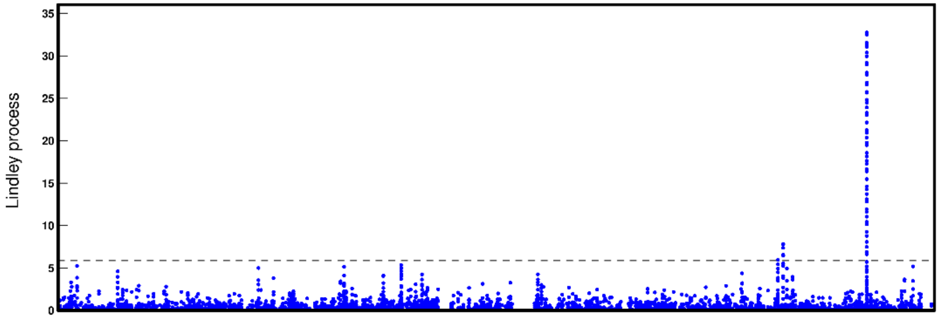
soil (n = 14)

plant communities (n = 49)

mean annual temperature



soil nitrogen content



Duflos *et al.* in prep

Setting up a Genome-Wide Association study

162 accessions x 13 strains*

GWAS in field conditions / mono-infection

*isolated from 7 out of the 12 most abundant and prevalent non-pathogenic bacterial species in the south-west of France

Whole-genome sequence for 22 strains

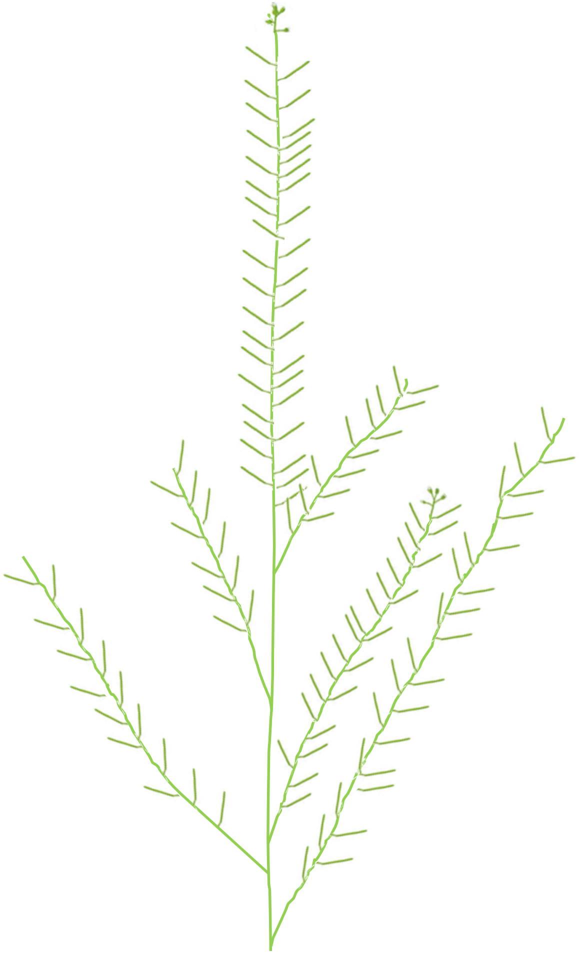
OTUs	Genus/ Species
OTU 2	*Paraburkholderia fungorum*
OTU 3	*Oxalobacteraceae* bacterium
OTU 4	*Comamonadaceae* bacterium
OTU 5	*Pseudomonas moraviensis*
OTU 6	*Pseudomonas siliginis*
OTU 13	*Methylobacterium* sp.
OTU 29	*Sphingomonadaceae* bacterium

Daniela Ramirez-Sanchez

Rémi Duflos

Ramirez-Sanchez *et al.* Front Microb 2022

Phenotyping of reproductive traits

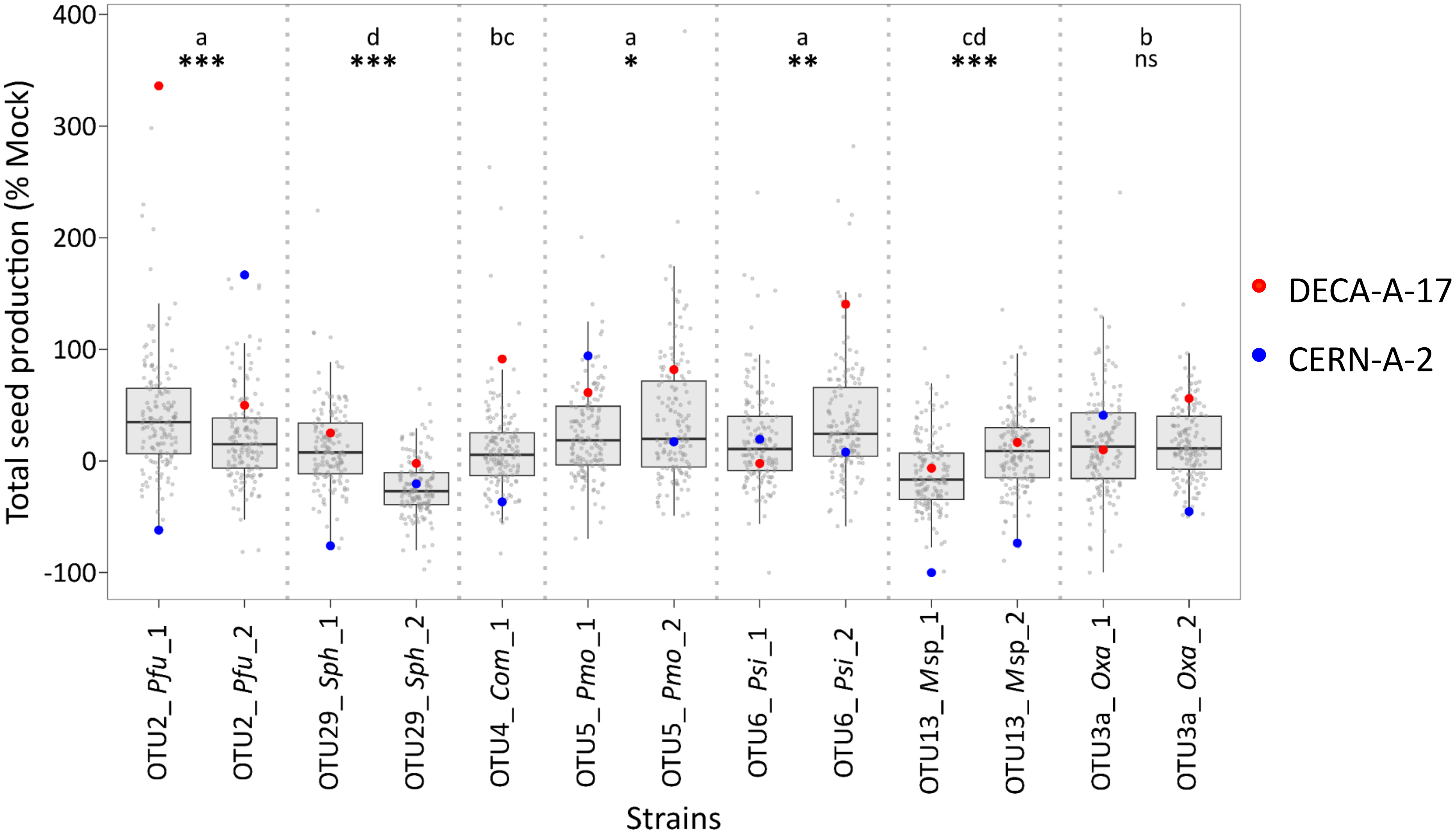


Total seed production:

- Number of fruits
- Mean fruit length (strongly correlated with the number of seeds per fruit)

N = 13,450 plants

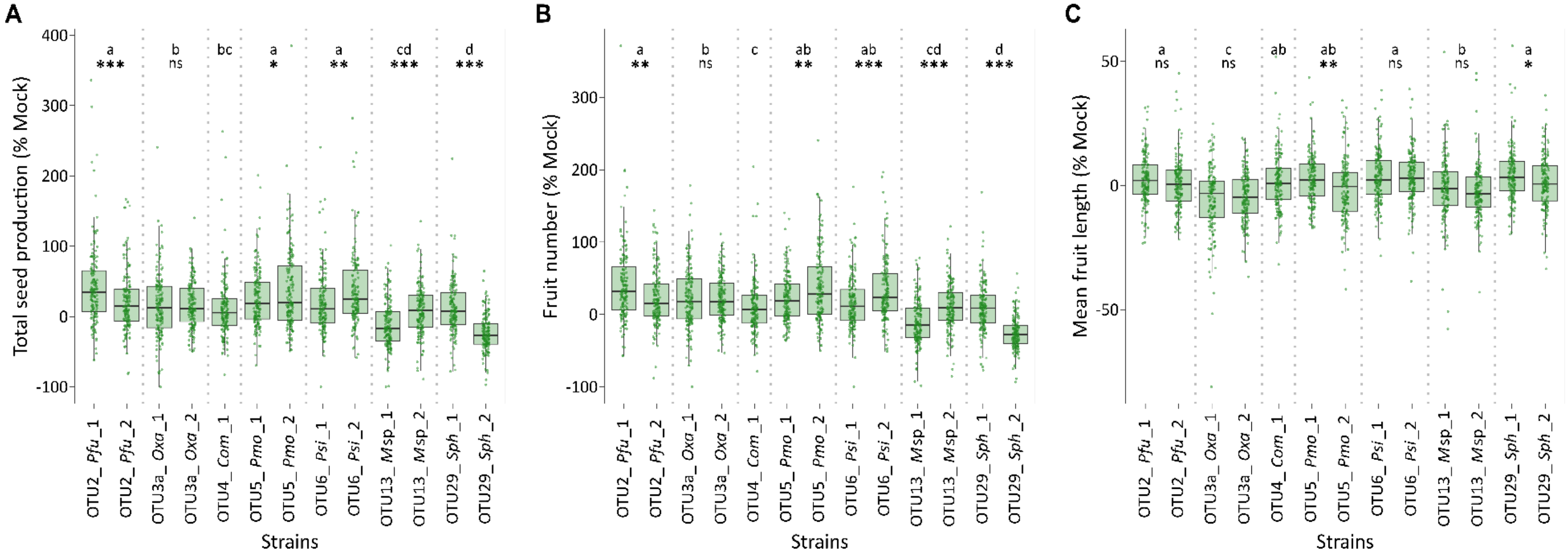
Extensive genetic variation observed in response to commensals



Duflos *et al.* (in prep)

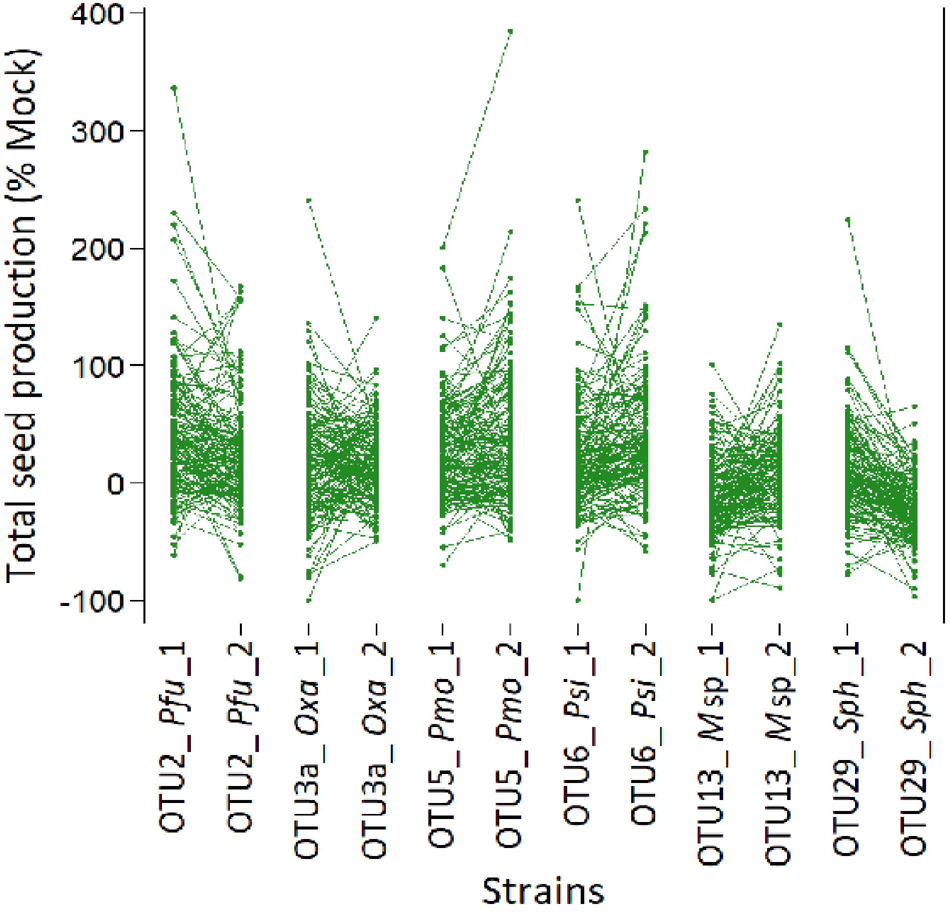
Extensive genetic variation observed in response to commensals

Total seed production   Fruit number   Mean fruit length



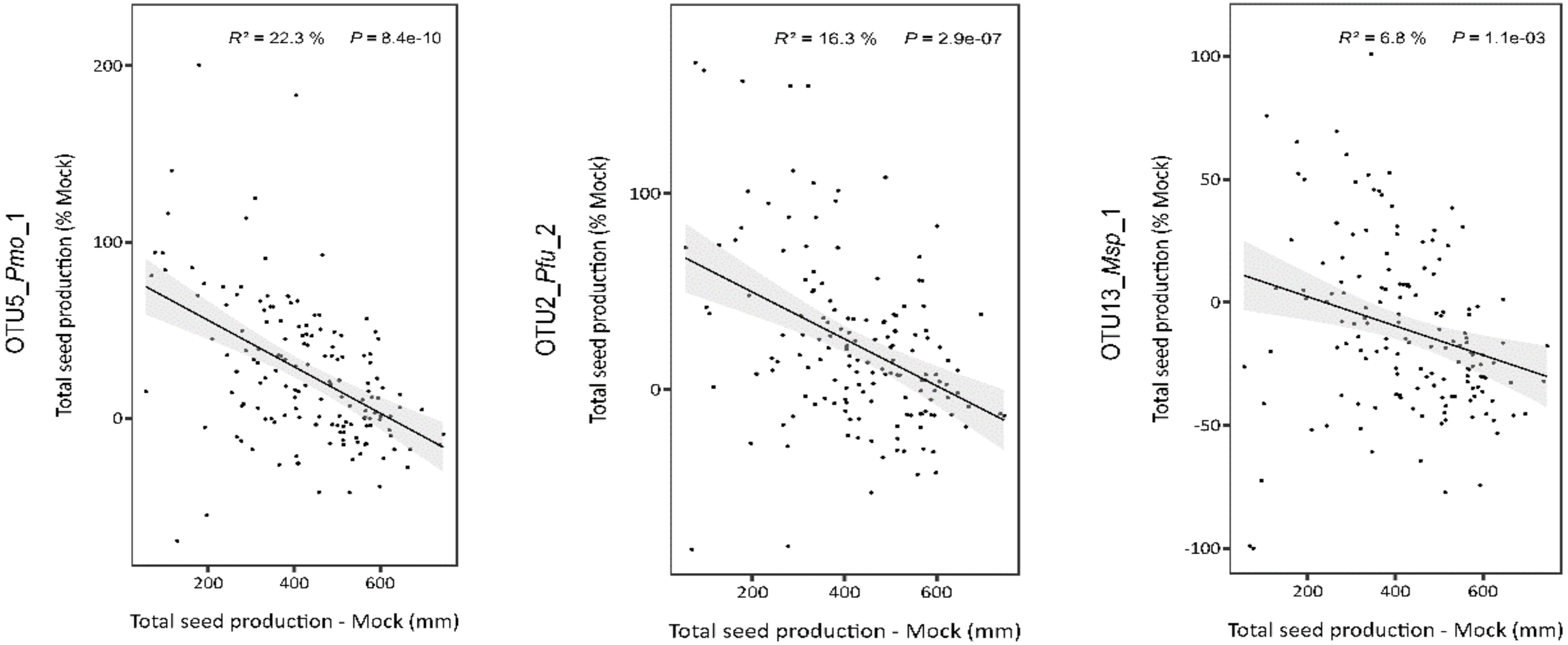
Duflos *et al.* (in prep)

Strong Genotype$_{host}$ x Genotype$_{commensal}$ interactions



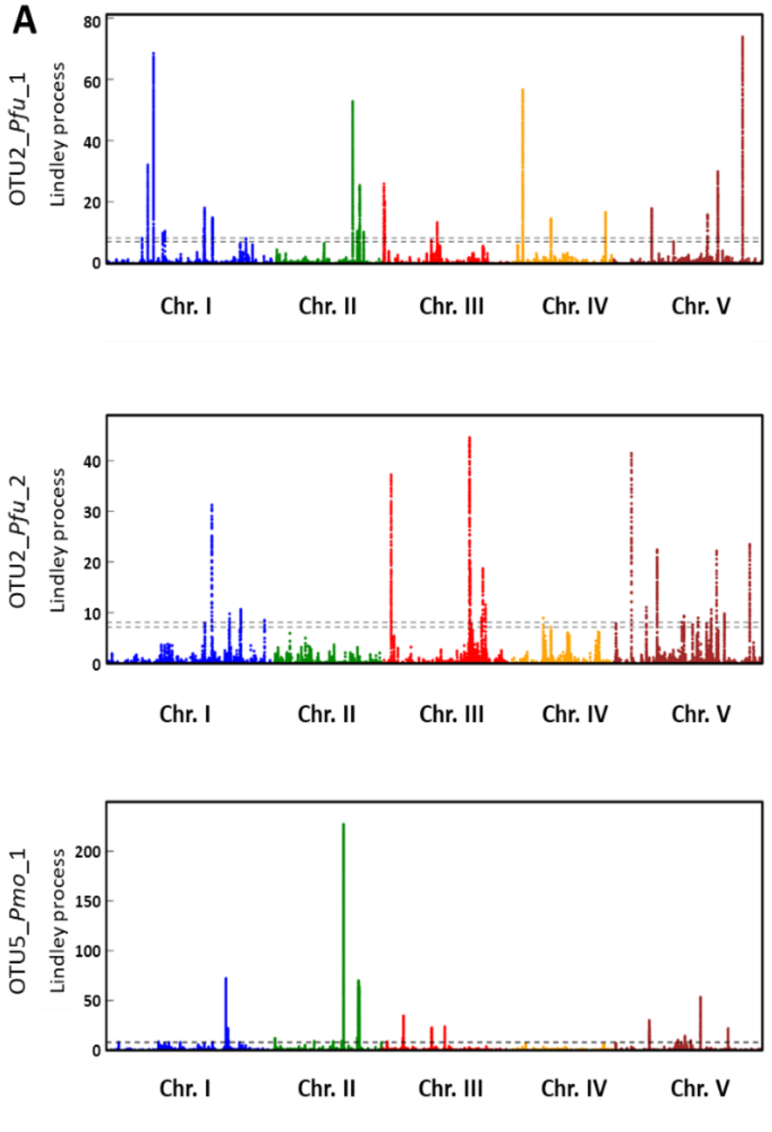
Duflos *et al.* (in prep)

Negative trade-off between the response to a commensal strain and total seed production in absence of inoculation



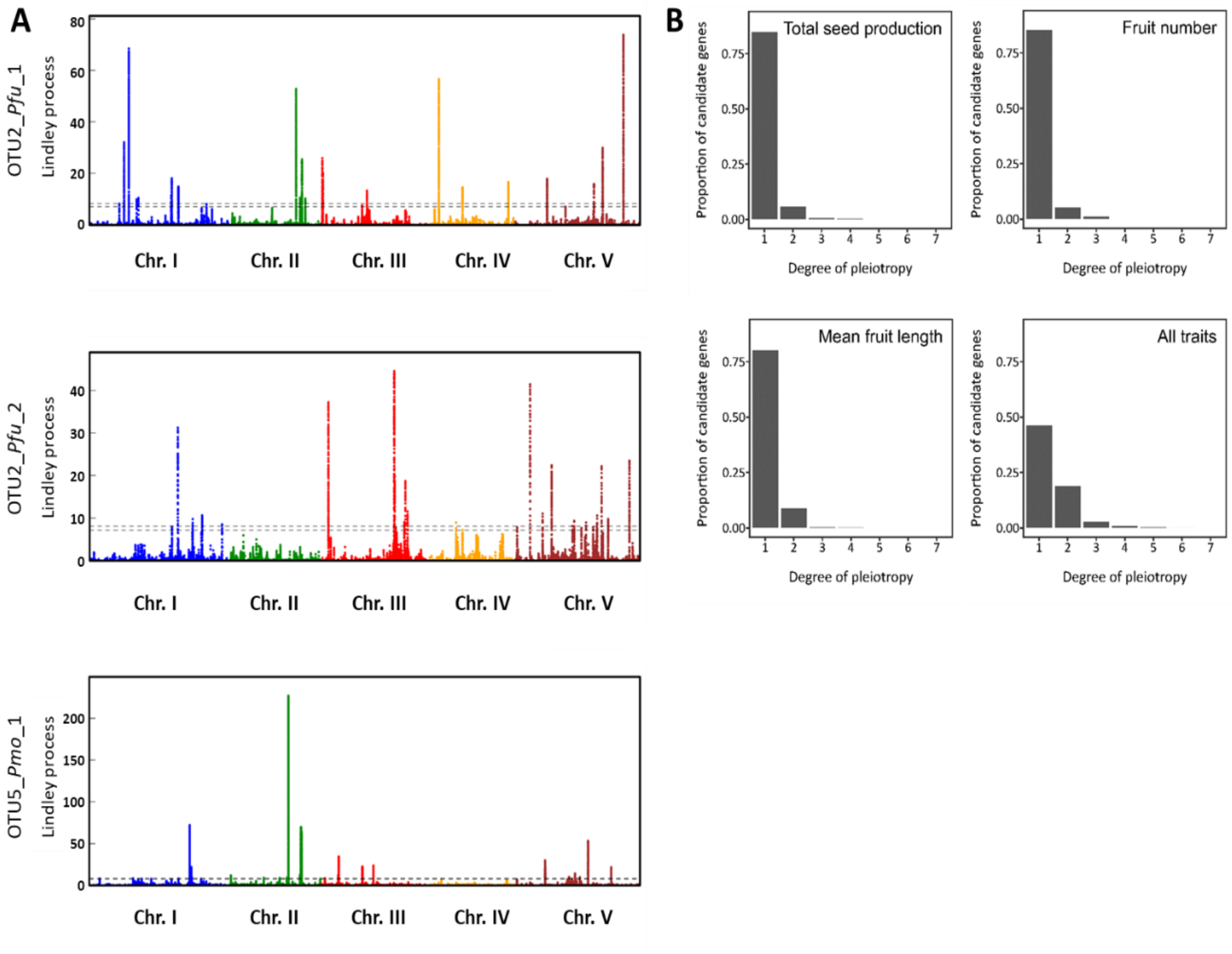
Duflos *et al.* (in prep)

A polygenic and flexible genetic architecture



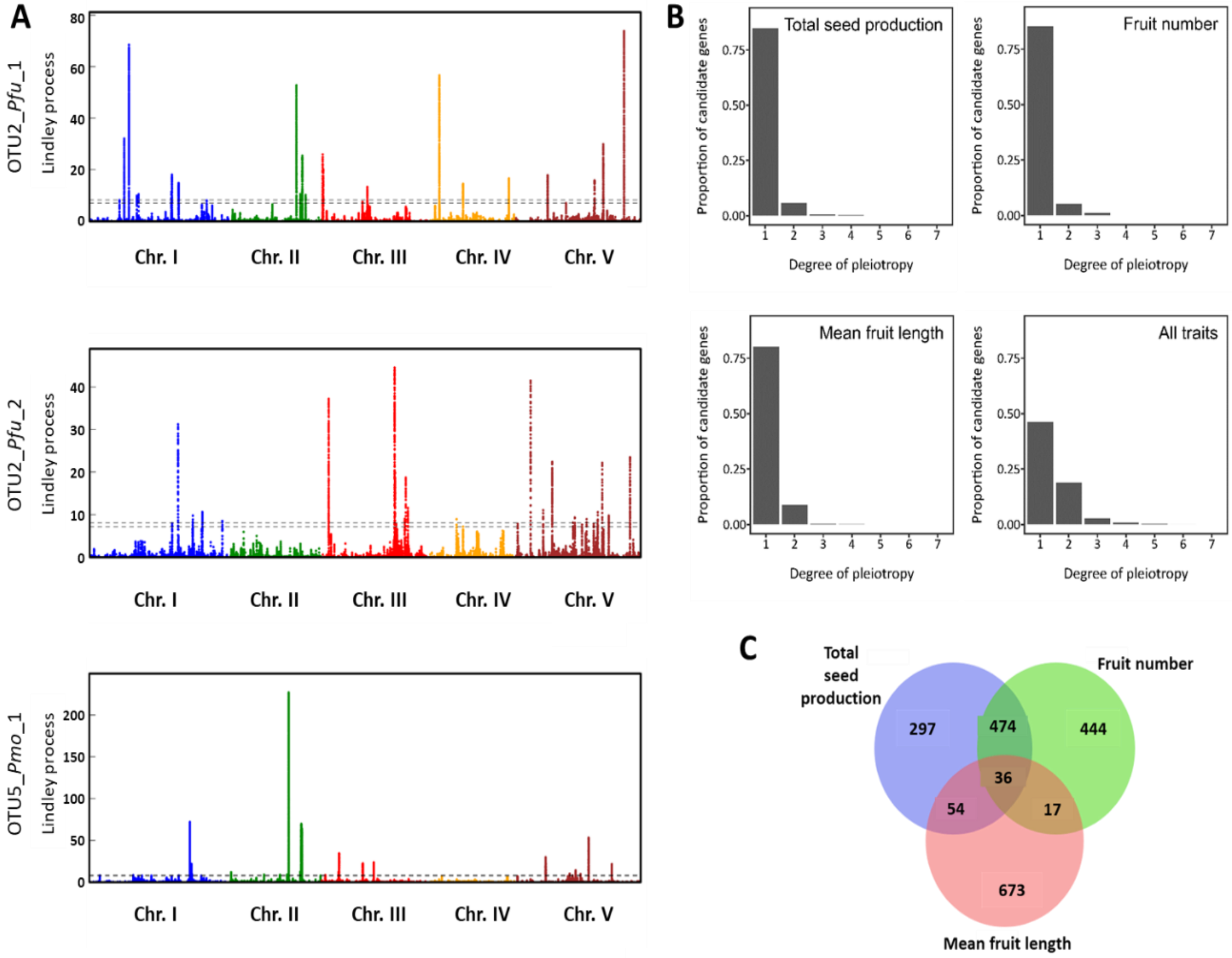
Duflos *et al.* (in prep)

A polygenic and flexible genetic architecture



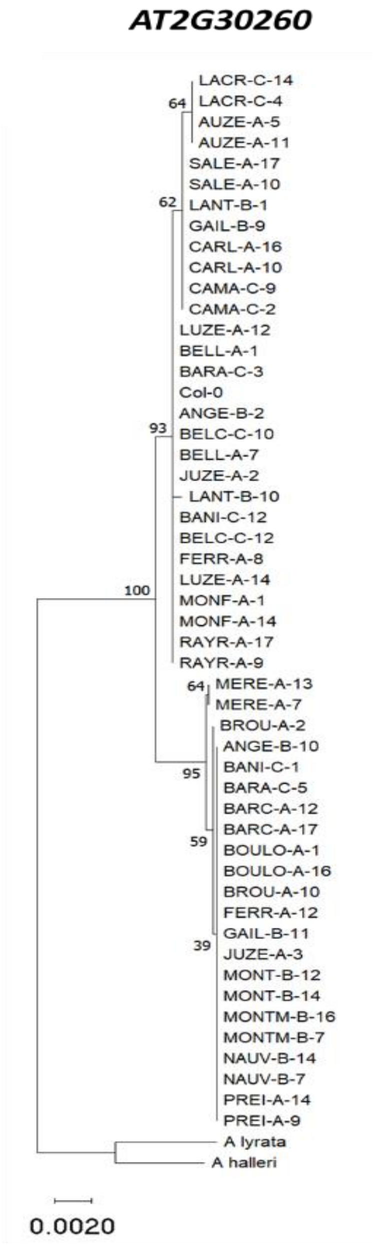
Duflos *et al.* (in prep)

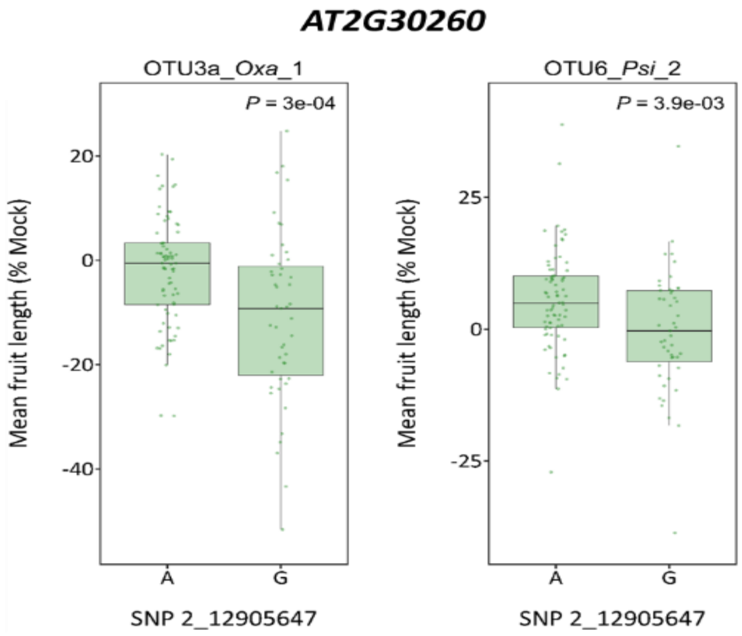
A polygenic and flexible genetic architecture



Duflos *et al.* (in prep)

Identification of candidate genes with two contrasted haplotypes



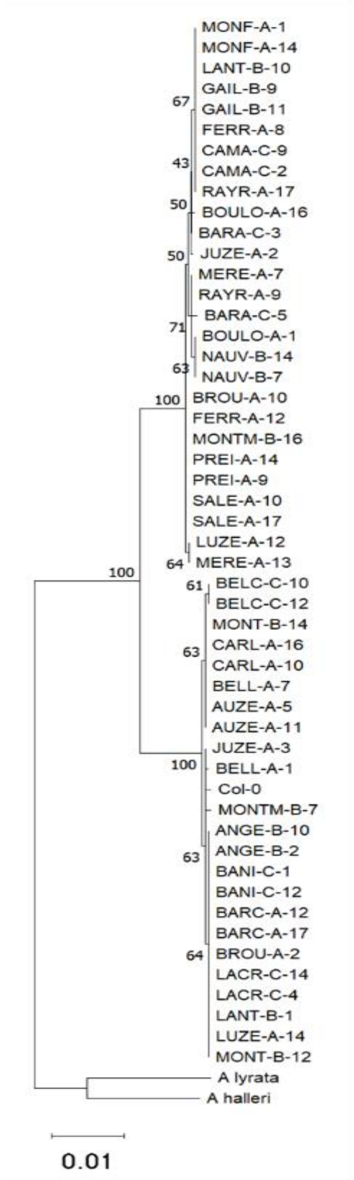


U2 small nuclear ribonucleoprotein B, which is a direct target of MYC2, a protein involved in a microbiota–root–shoot circuit to boost plant growth

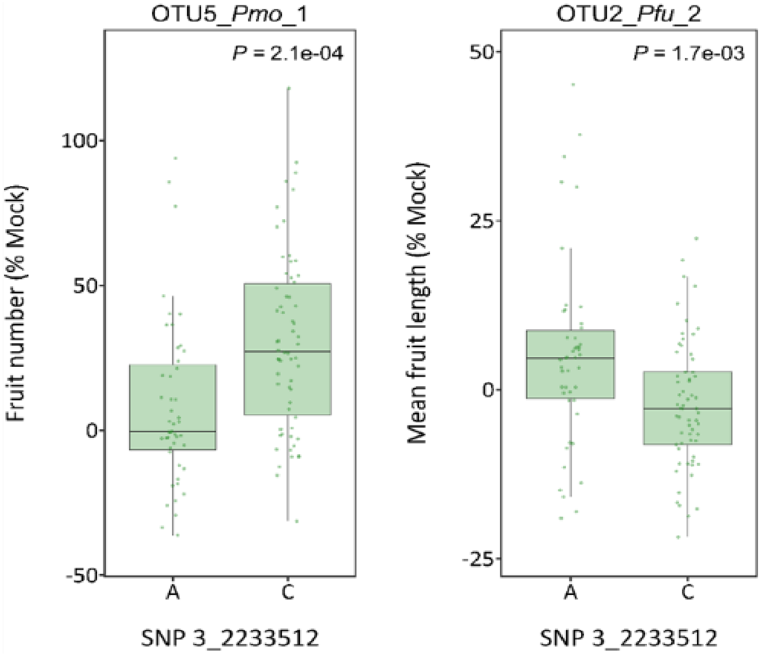
Duflos *et al.* (in prep)

Identification of candidate genes with two contrasted haplotypes

AT3G07060



AT3G07060



NHL domain-containing protein, Embryo defective 1974, required for embryo development.

Duflos *et al.* (in prep)