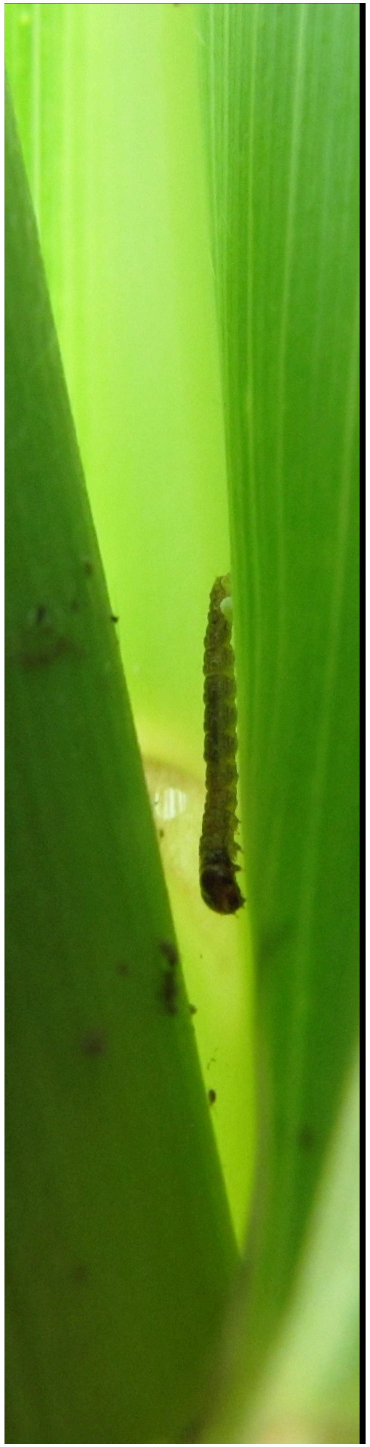


Plant adaptation and speciation in *Spodoptera frugiperda* (Fall Army Worm)

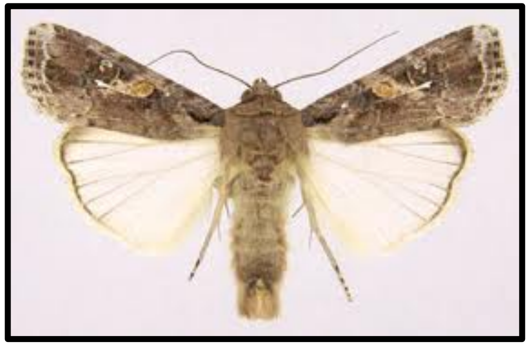
Kiwoong Nam & Nicolas Nègre

Laboratoire DGIMI

INRA, Université de Montpellier

Spodoptera frugiperda (Noctuidae)

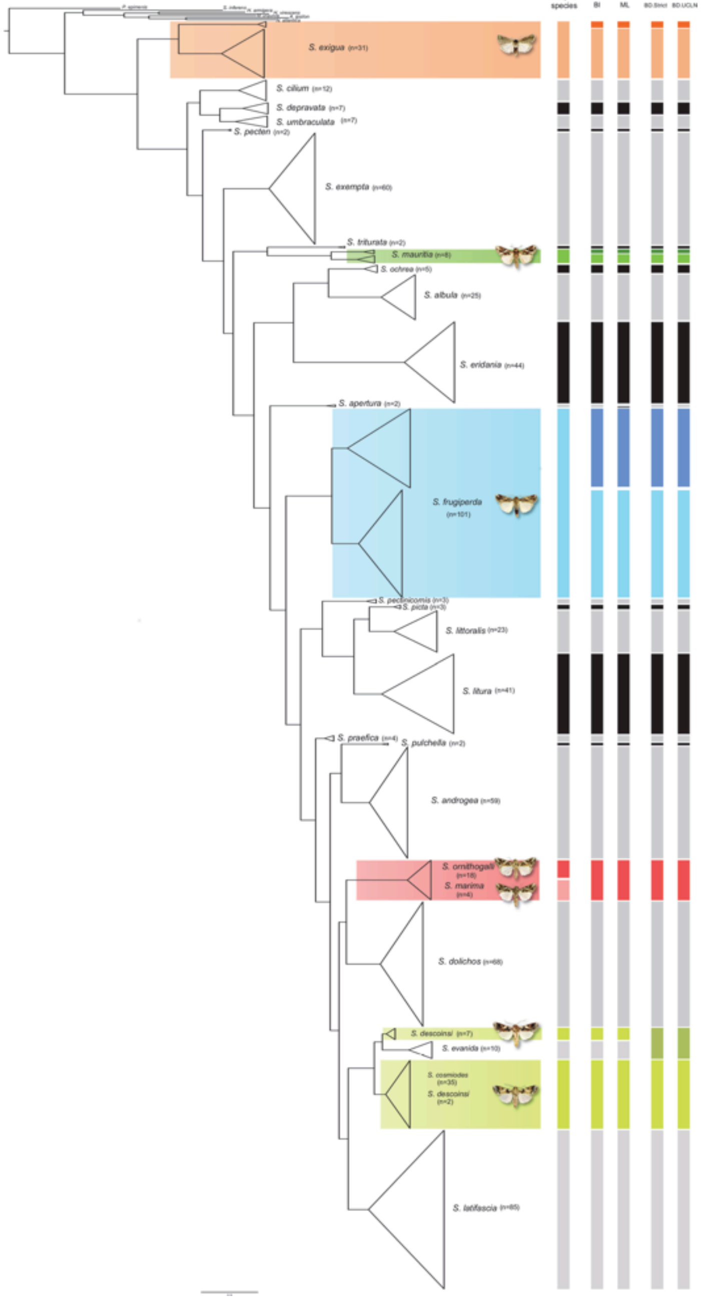
A pest moth of economical importance in North America, the Caribbean and South America.



2 strains morphologically undistinguishable

While polyphagous the two strains have a preferential host plant

Two species ?



Dumas et al., *Plos One*, 2015

Sf-Corn

Sf-Rice

"[…]host races are **genetically differentiated**, **sympatric** populations of parasites that use **different hosts** and between which there is appreciable **gene flow**."
in Berlocher & Feder, 2002

Populations

Host races

Species

A multi-level integrative analysis in order to understand the mechanisms of speciation in relation to the host-plant adaptation

A. **Genetic differences** between the two strains. What is the estimated divergence and is it linked to agriculture ?

B. **Phenotypic differences** between the two strains. Is the host plant a trigger mechanism of speciation or reinforcement mechanism ?

C. **Gene expression differences** in response to environment. Build a comprehensive transcriptome and test the adaptation of *S. frugiperda* to its host plant.

D. **Epigenetic differences** : is the genetic plasticity linked to different chromatinian landscapes of *S. fru* ? Can holocentrism help to find the balance between deleterious deregulation and adaptability ?

Question 1

What is a genetic basis of adaption to host-plant?

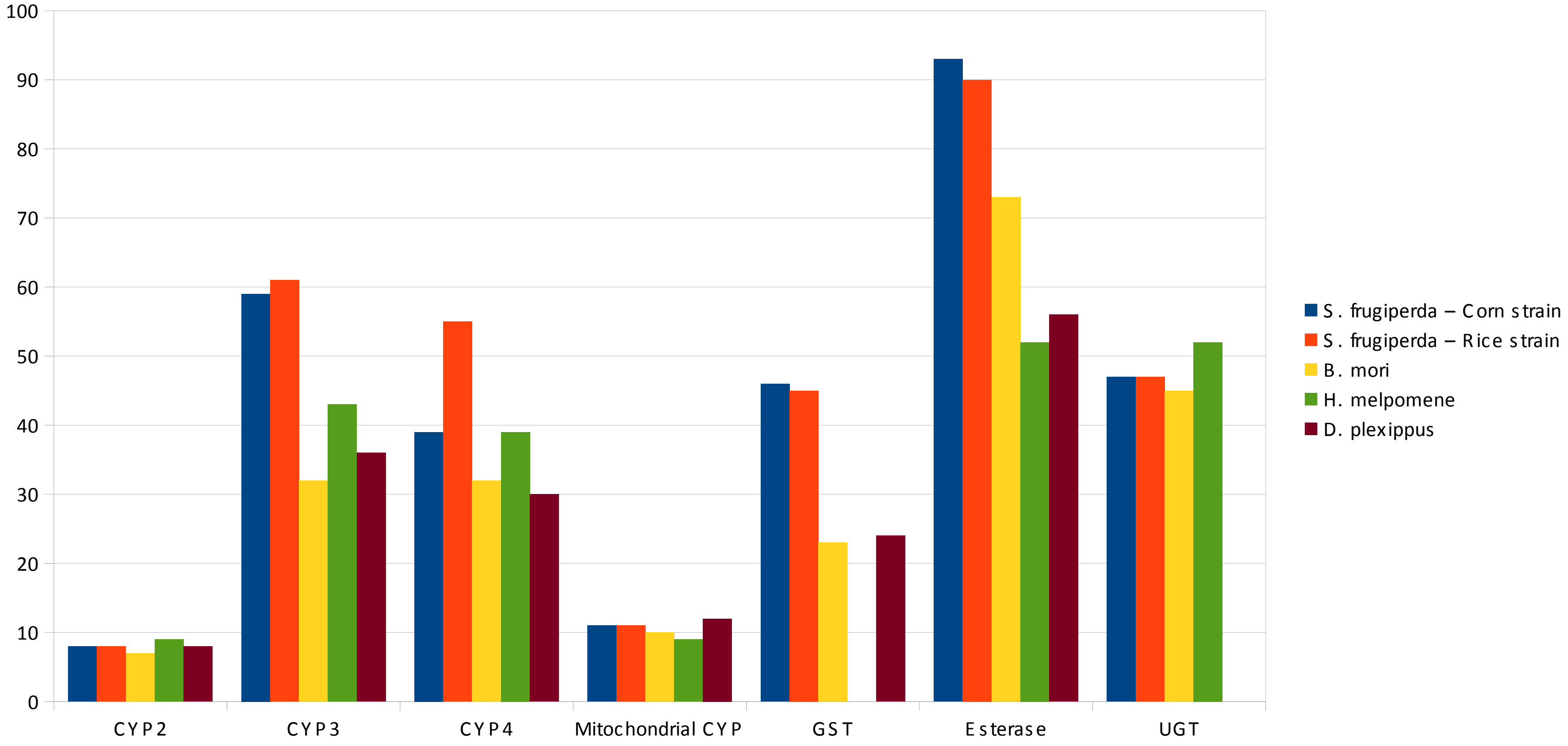
⇨ Comparaitive genomics analysis

1. Polyphagy versus Monophagy
2. Corn versus Rice strains

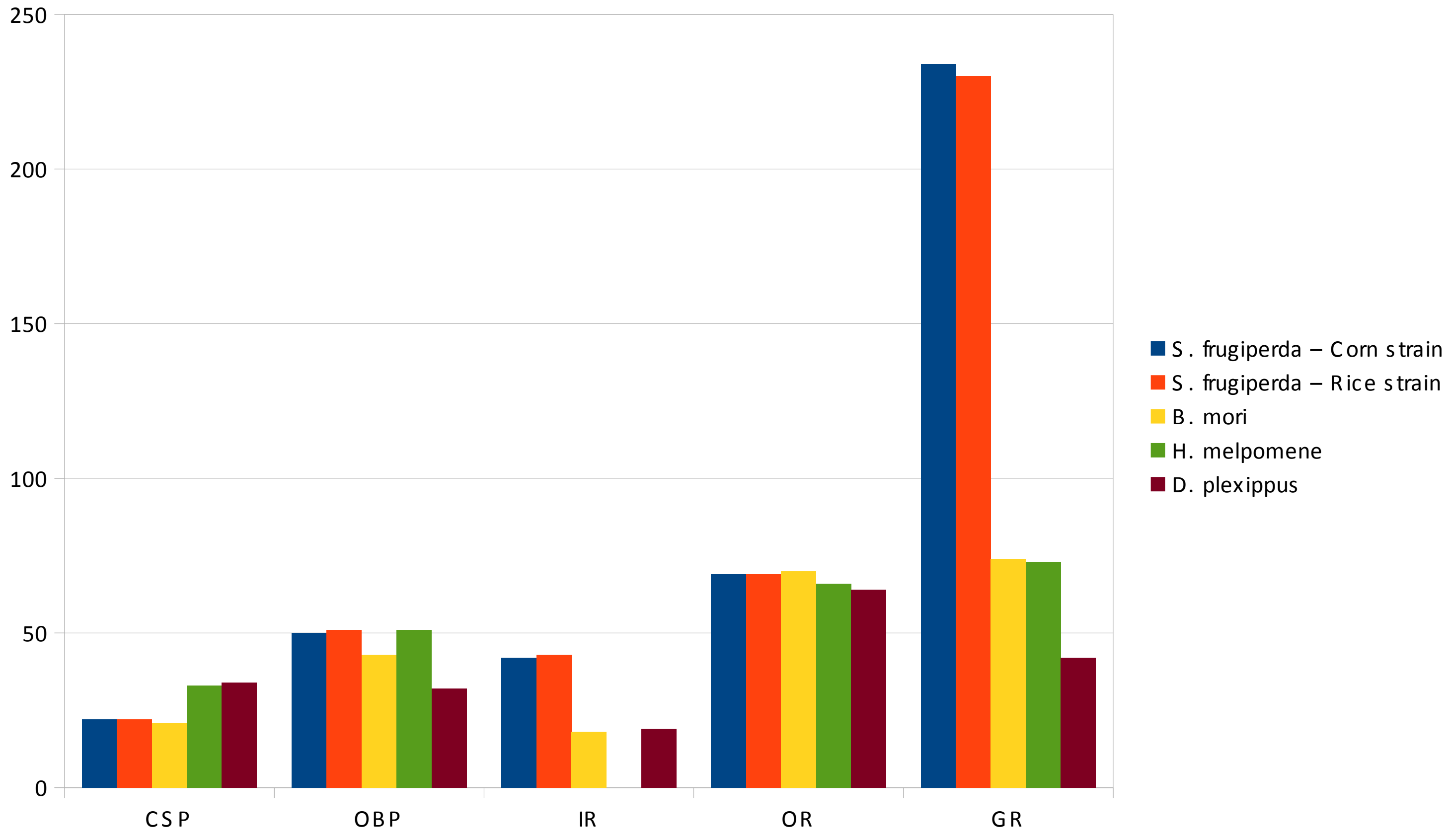
Whole genome sequencing

	Corn strain	Rice strain
Origin of the sample	Guadalupe	Florida
Assembly software	AllPaths-LG	Platanus
# of scaffolds	41,577	29,127
Total Size	437,873,293	371,020,023
N50	52.781 bp	28.526 bp
Number of N	11,379,916 (2.60 %)	130,481 (0.04%)

Detoxification genes

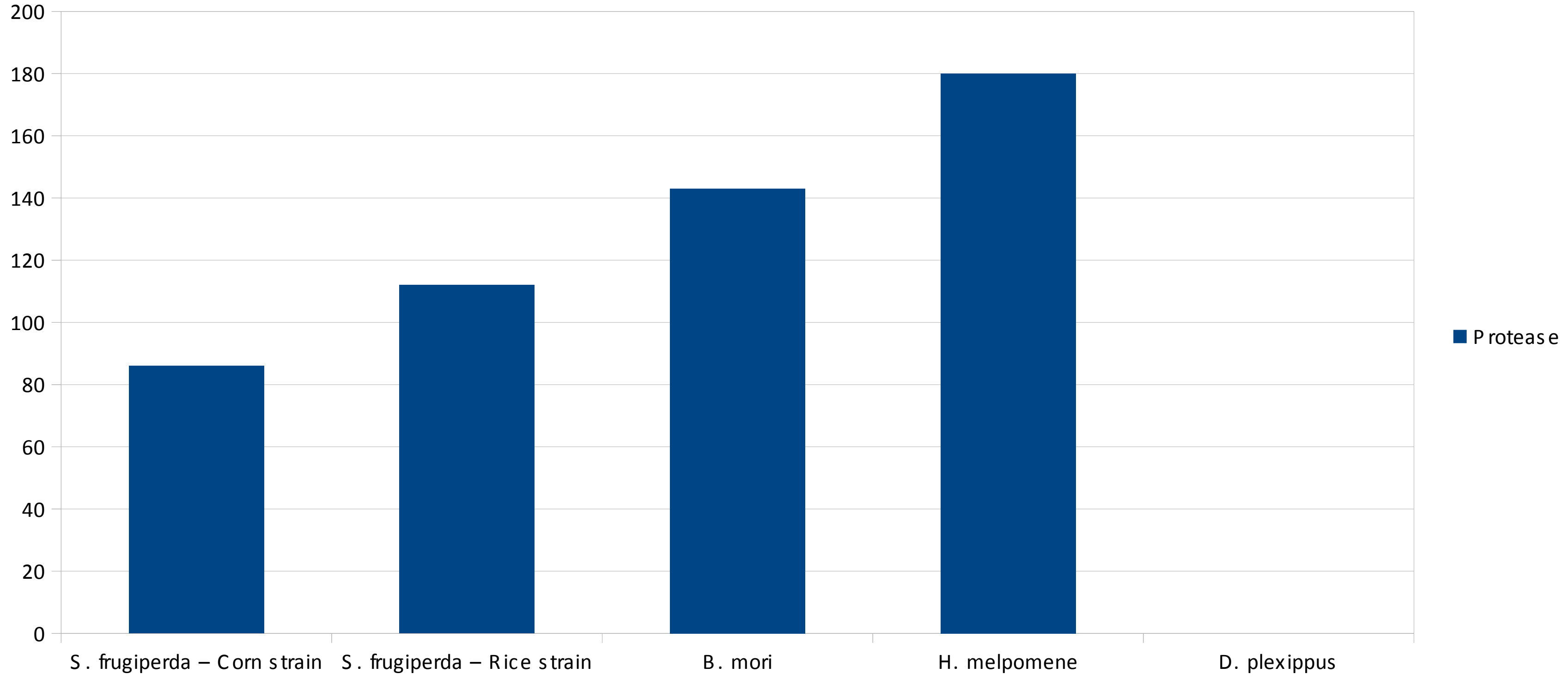


Chemosensory genes



Gene expansion of GR genes

Digestion genes

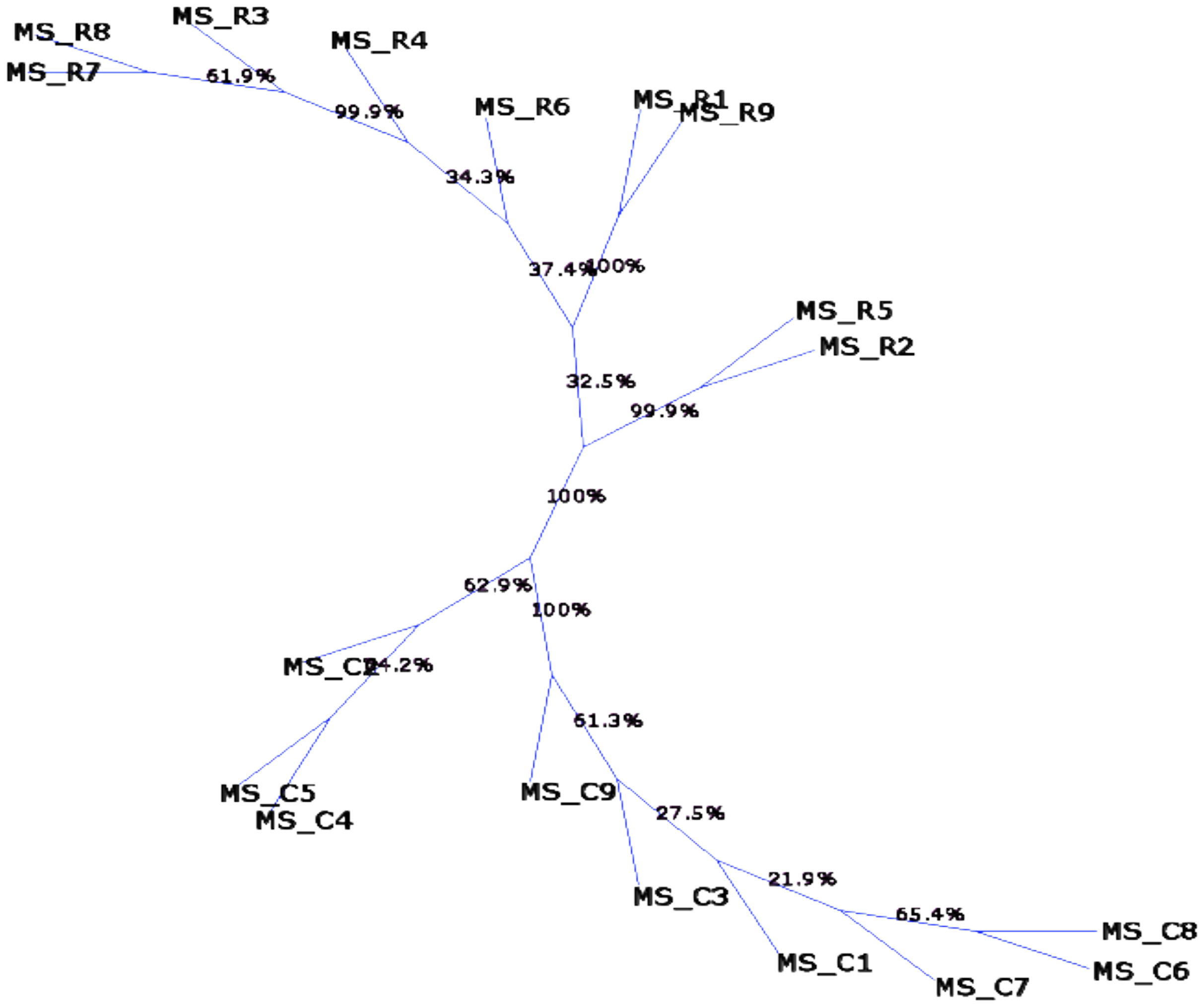


The comparison between Corn and Rice strains

1. Sampling Corn and Rice strains from Mississippi
2. Sequencing 9 corn and 9 rice (Illumina Hiseq)
3. Mapping (bowtie2) and variant calling (samtools)
4. Calculating genetic distance between pairs of individuals
5. Making a NJ tree

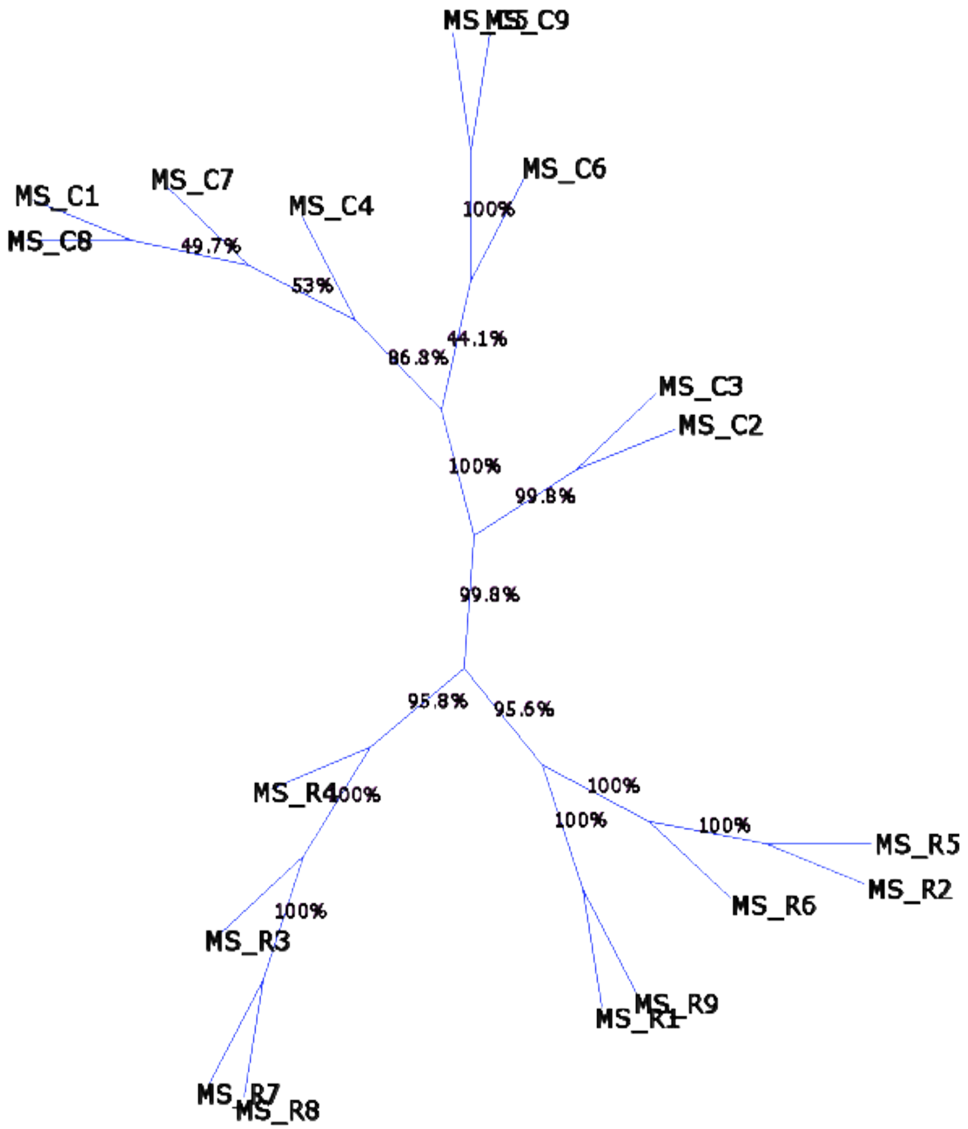
NJ tree

stru_mt



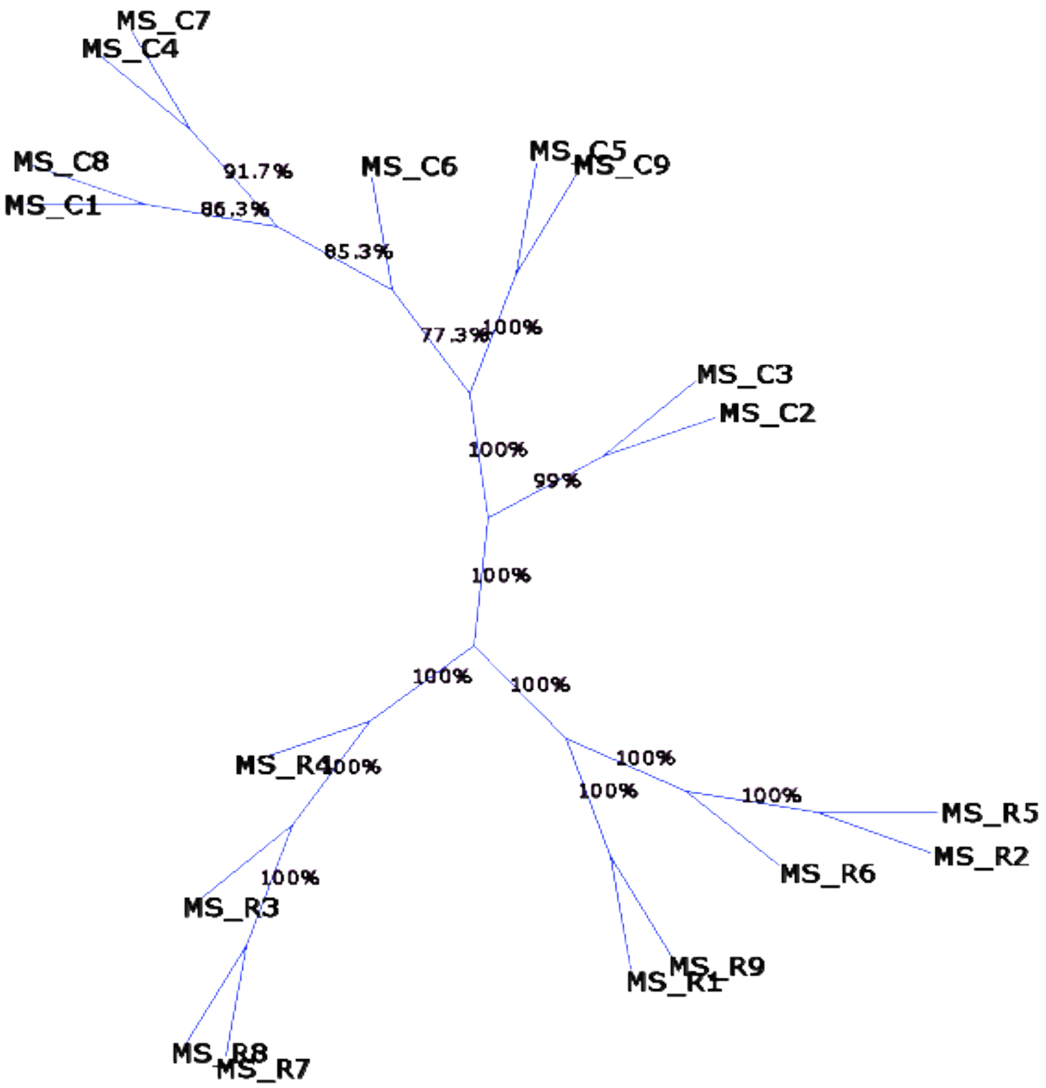
NJ tree

stru_C

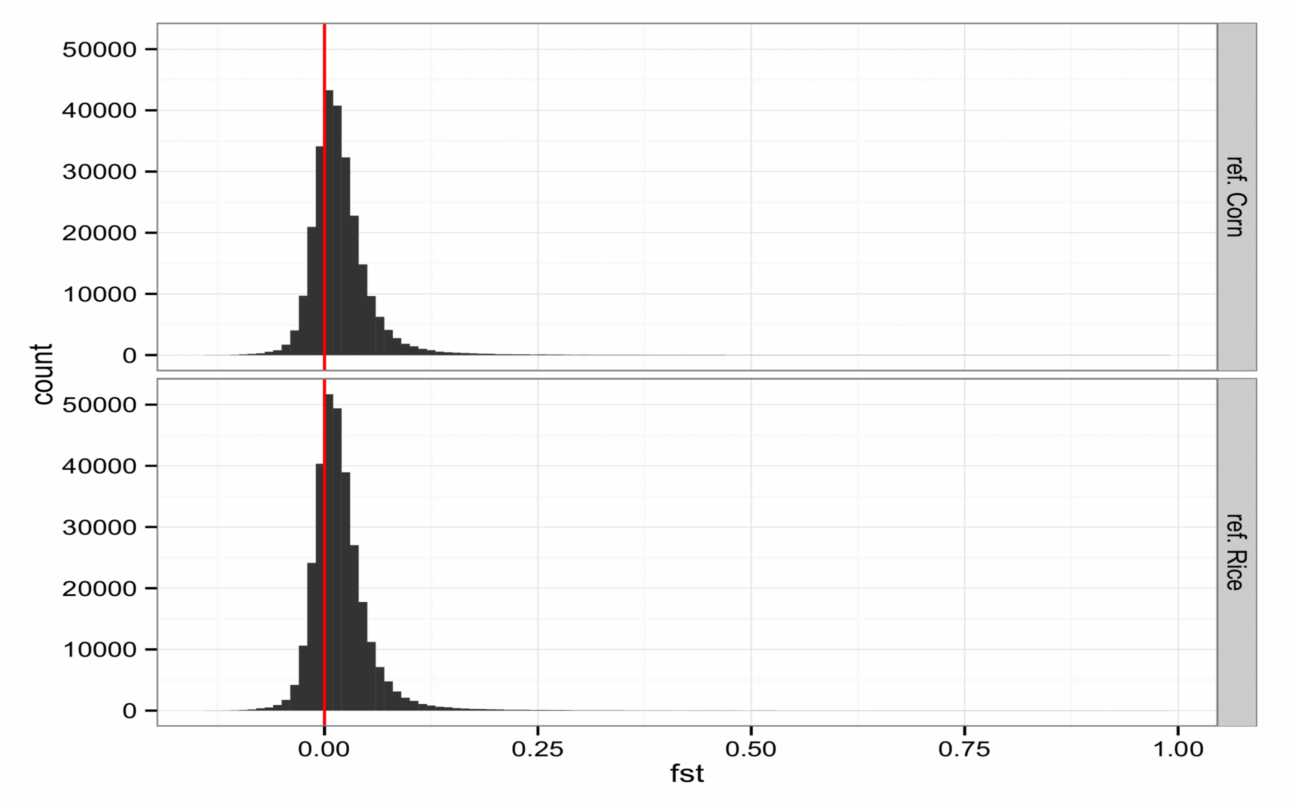


NJ tree

stru_R



Distribution of Fst



Average Fst = 0.019

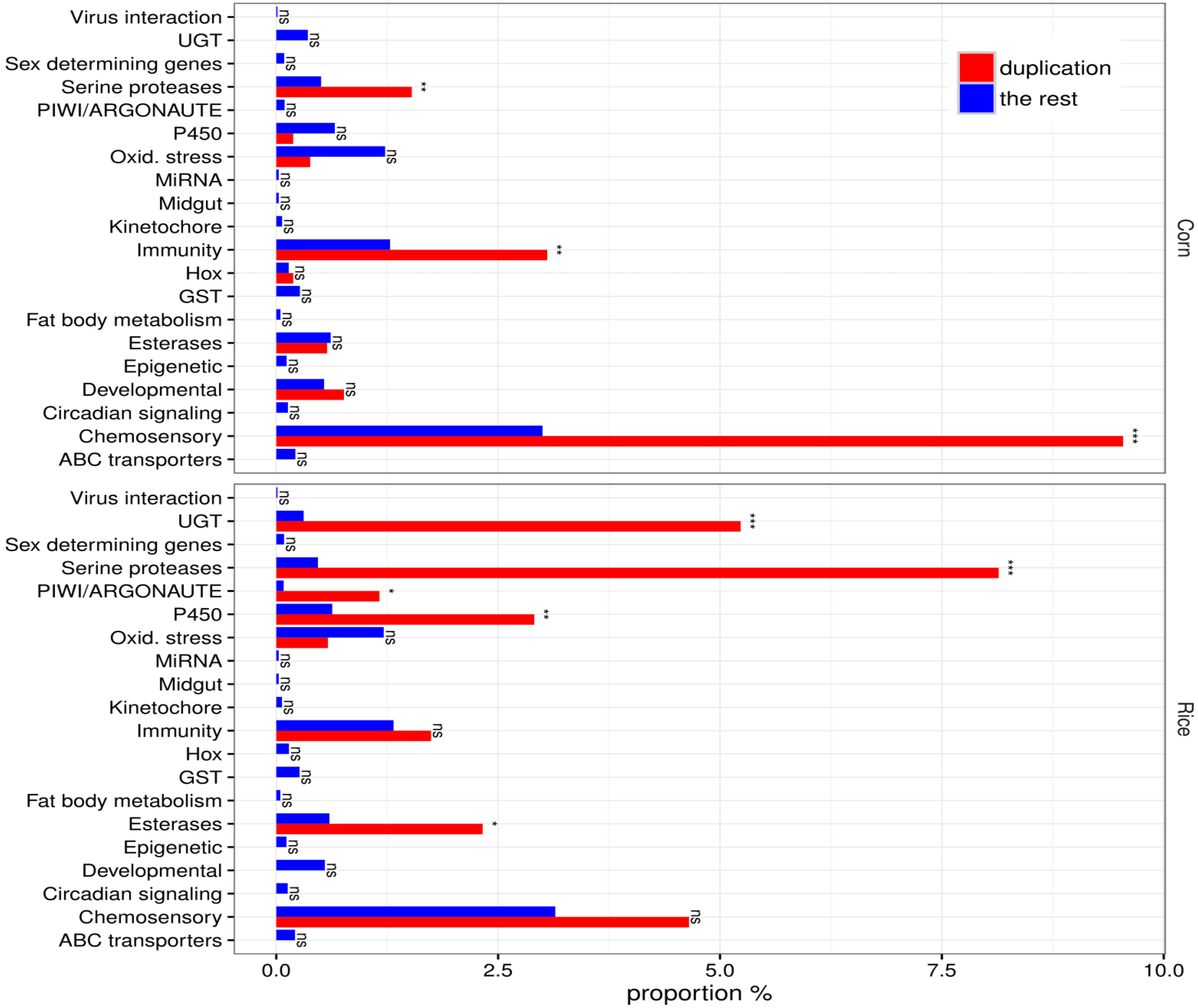
Significant genetic differentiation between corn and rice ($p < 0.005$)

Two genetic identities of corn and rice strains

Structural variation

	Insertion	Deletion	Copy number gain	Copy number loss	Inversion	Transposition
Number	1,108	1,009	475	417	49	271
Coverage	1.1 Mb	0.9 Mb	5.2 Mb	1.0 Mb	59 kb	345 kb

Structural variation



Question 1

What is a genetic basis of adaption to host-plant?

➡ Comparaitive genomics analysis

1. Polyphagy versus Monophagy
2. Corn versus Rice strains

➡ Copy number variation of detoxification, chemosensory, immune, and digestion genes

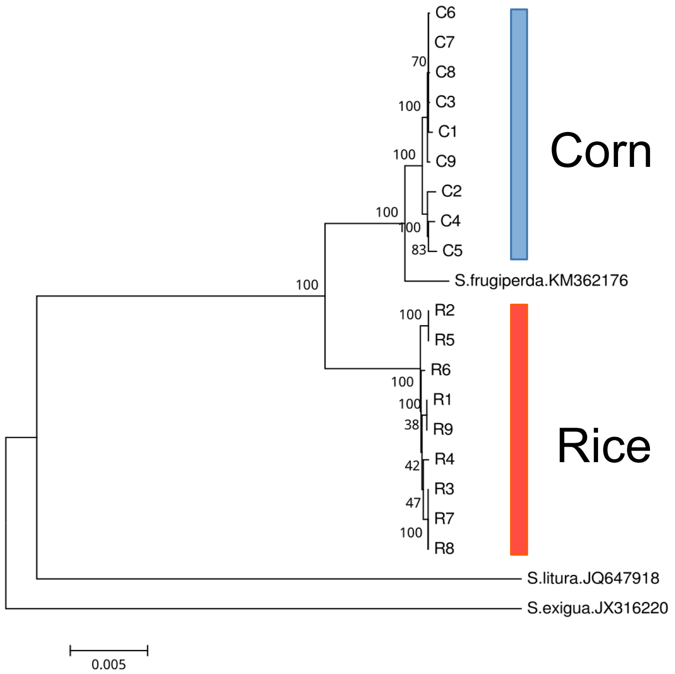
Remaining issue: What is the meaning of CNV, population level difference in CNV

Question 2

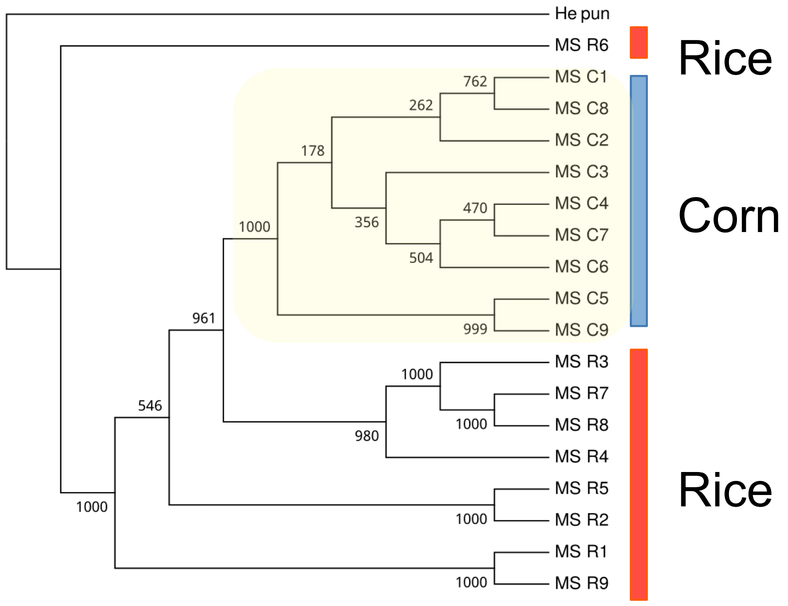
What is a genetic basis of speciation between corn and rice strains

➡ Population genetics analysis (preliminary)

Phylogenetic analysis



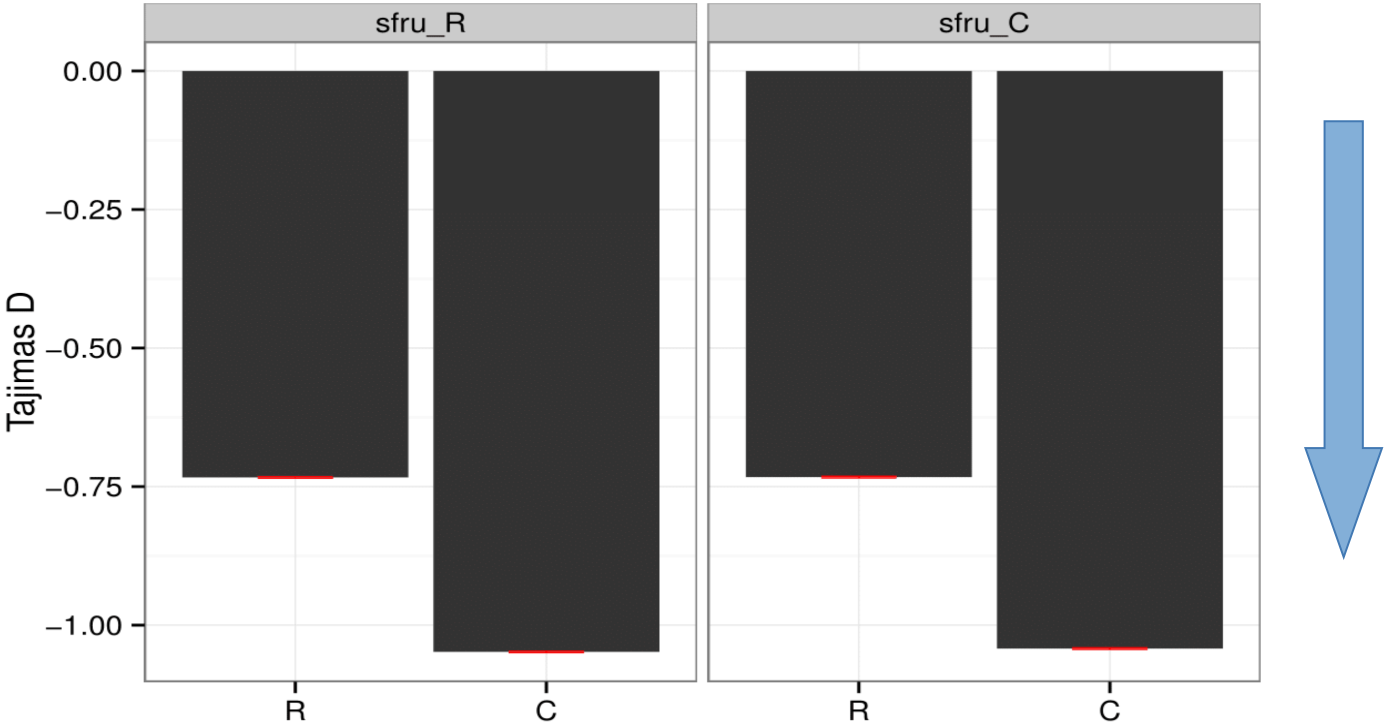
mtDNA (full sequences)



NucDNA (1,207,856 Positions)

Corn and Rice strains are the sister group of each other

Corn strain is a subgroup of the rice strain

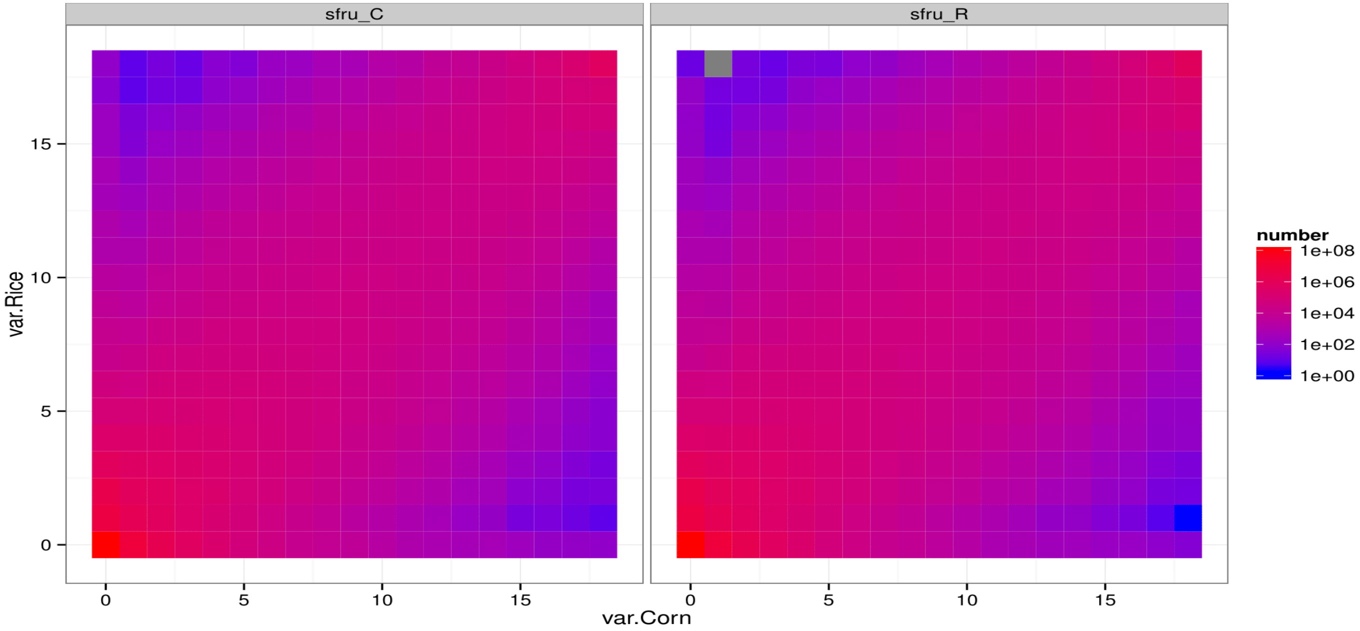


Faster
population
growth

Corn experiences faster rate of population expansion

Correlated allele frequency

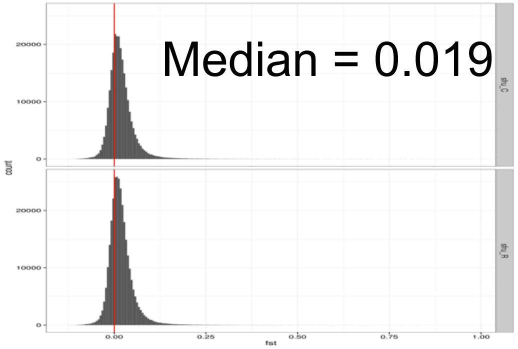
1. For each position, the numbers of alternative allele in the corn and the rice strains are counted
2. Calculating the frequency of these numbers



→ Shared SNPs → Recent diverging time

Mito-divergence predates nuc-divergence?

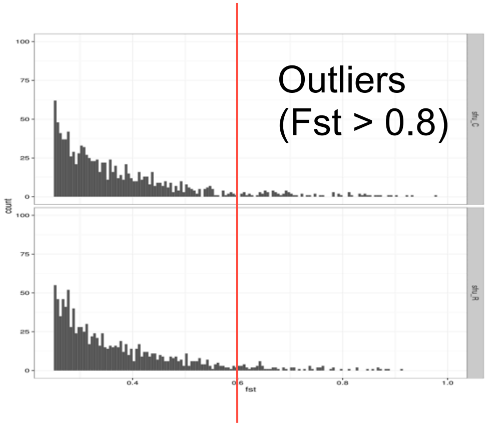
Distribution of Fst



Median = 0.019

Ref: C

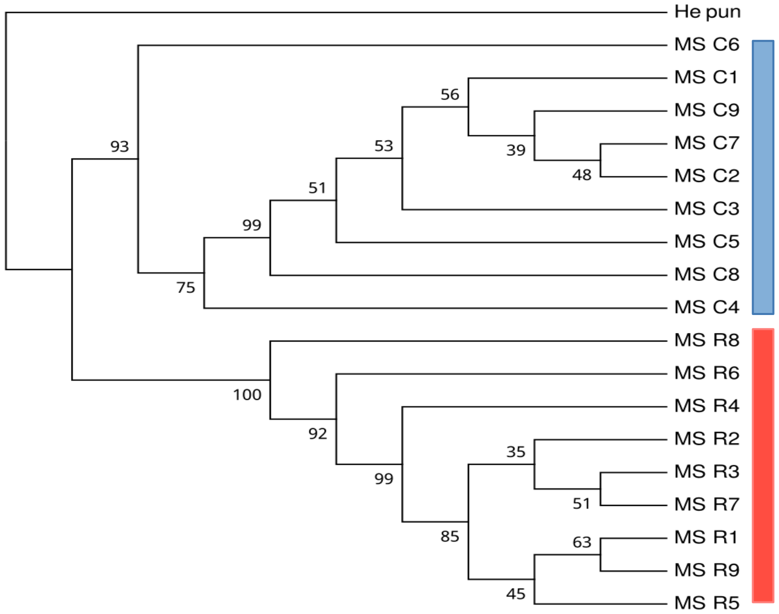
Ref: R



Outliers
(Fst > 0.8)

16 contigs

7 contigs

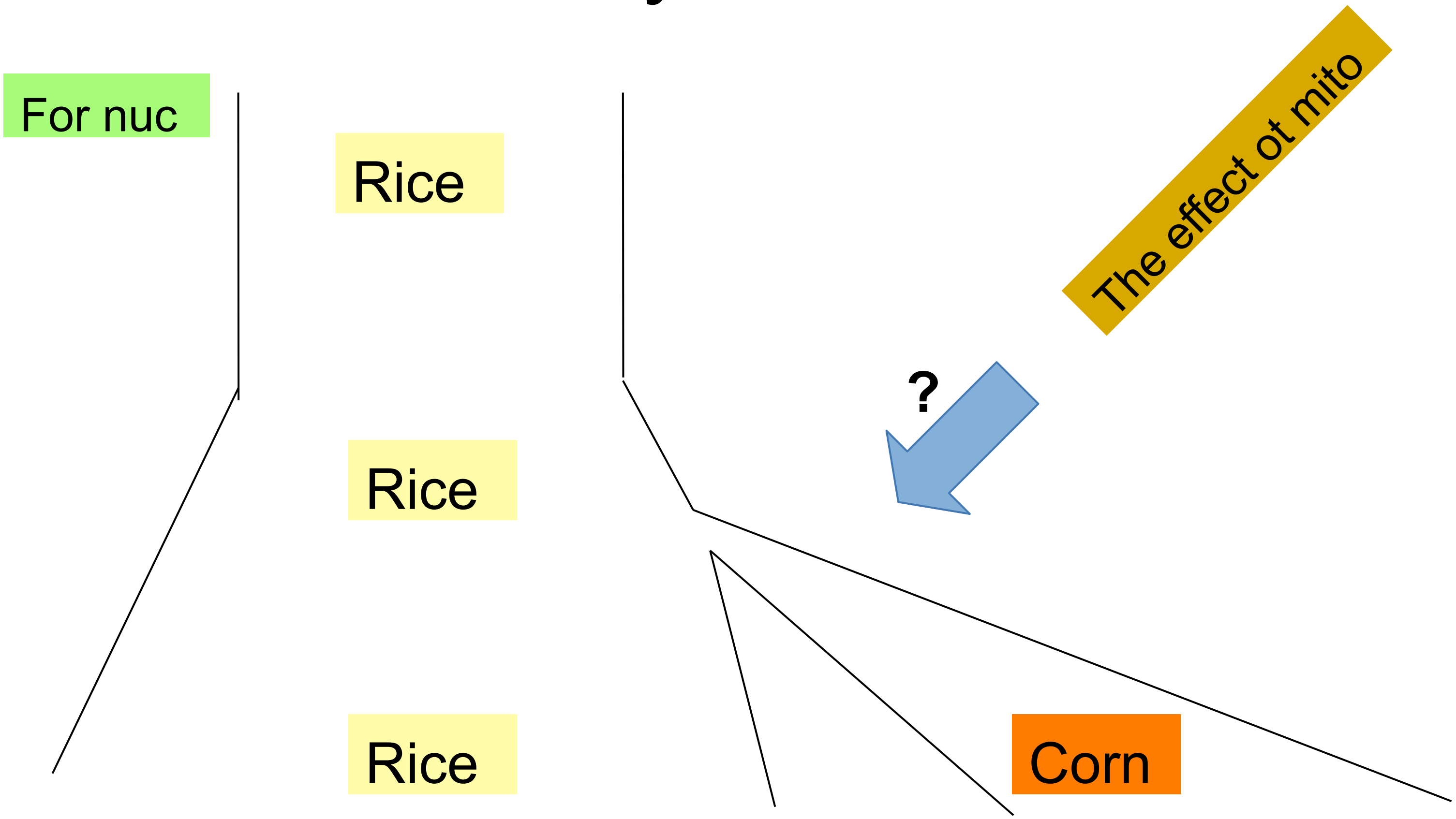


The same topology with mtDNA

98 SNPs

Mitochondria-divergence predates nuclear divergence

Most likely scenario



For nuc

Rice

Rice

Rice

Corn

The effect of mito

?

Question 2

What is a genetic basis of speciation between corn and rice strains

⇨ Population genetics analysis

⇨ Mt-nuclear incompatibility might be a driving force of speciation

Many things to be done...

A multi-level integrative analysis in order to understand the mechanisms of speciation in relation to the host-plant adaptation

A. **Genetic differences** between the two strains. What is the estimated divergence and is it linked to agriculture ?

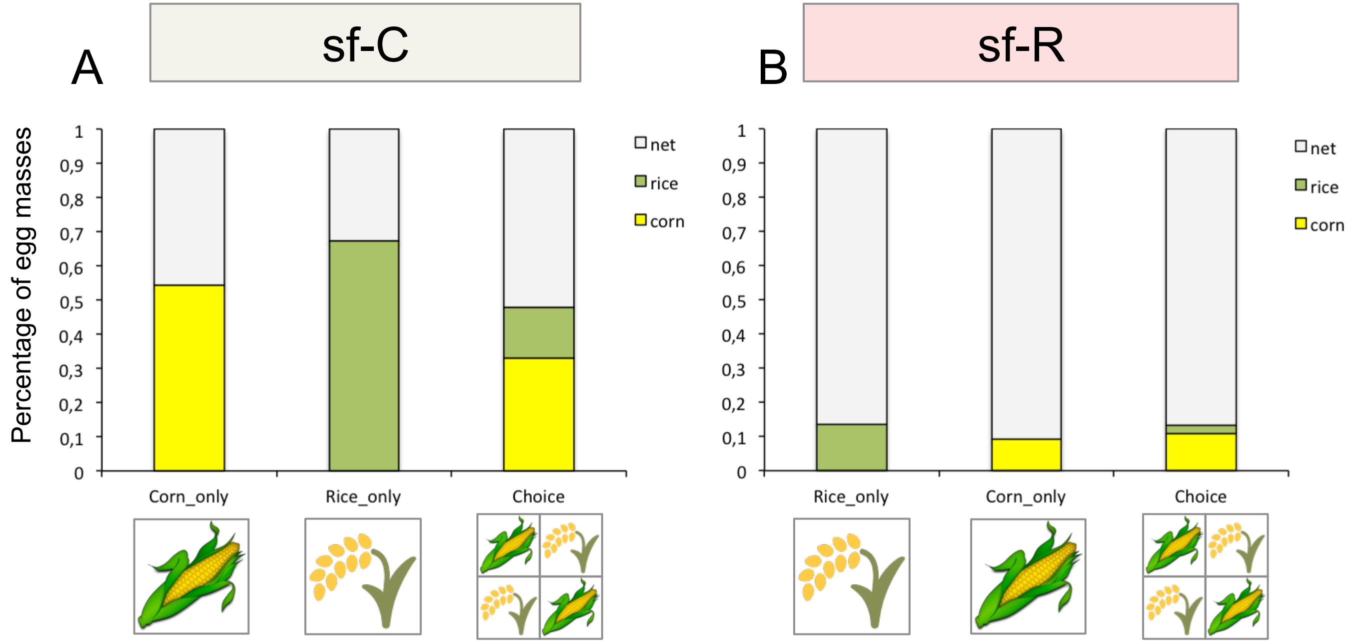
B. **Phenotypic differences** between the two strains. Is the host plant a trigger mechanism of speciation or reinforcement mechanism ?

C. **Gene expression differences** in response to environment. Build a comprehensive transcriptome and test the adaptation of *S. frugiperda* to its host plant.

D. **Epigenetic differences** : is the genetic plasticity linked to different chromatinian landscapes of *S. fru* ? Can holocentrism help to find the balance between deleterious deregulation and adaptability ?

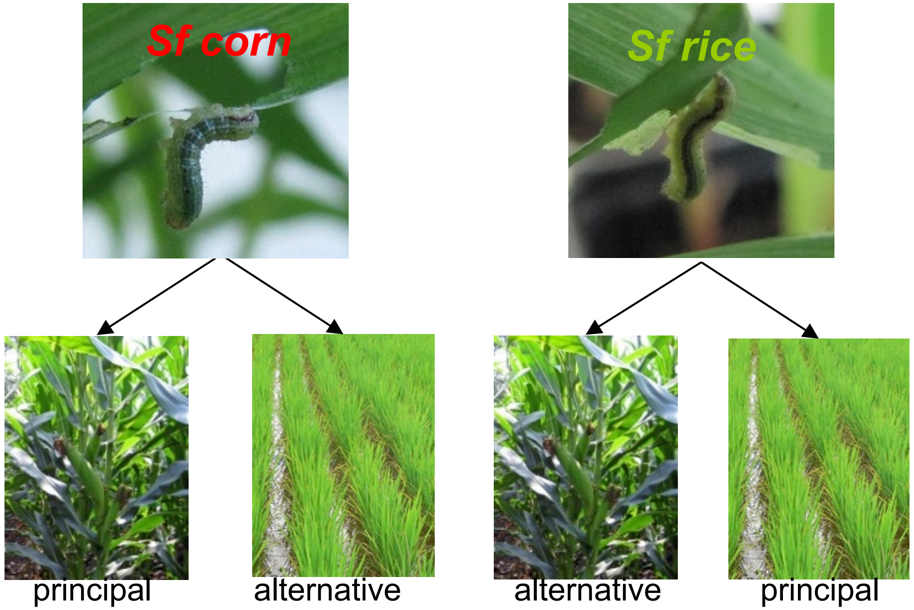
1: Oviposition preference in RT experiments

Oviposition preference can be an indicator of the fitness of species on different plants



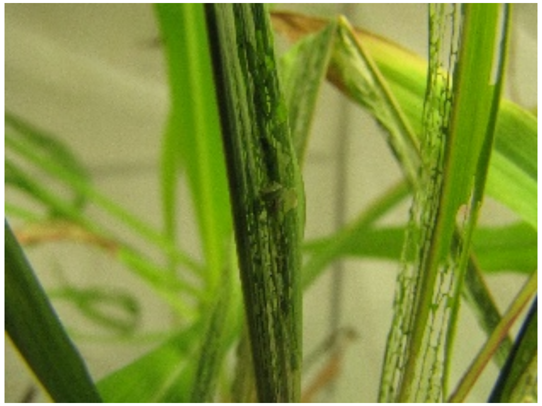
- Behavioral difference between strains
- No sign of specialization to plant
- Test fitness variations at larval stages

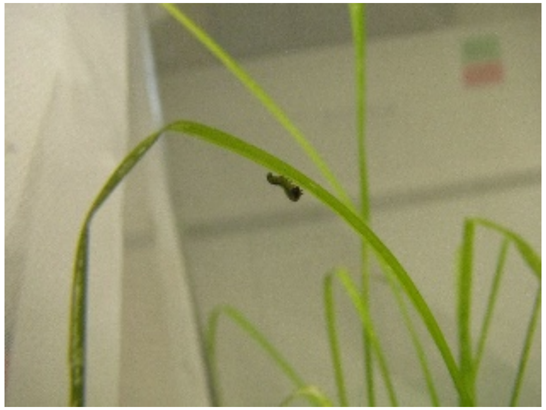
Experimental design



Reciprocal Transplant experiment of two laboratory strains on their principal or alternative host

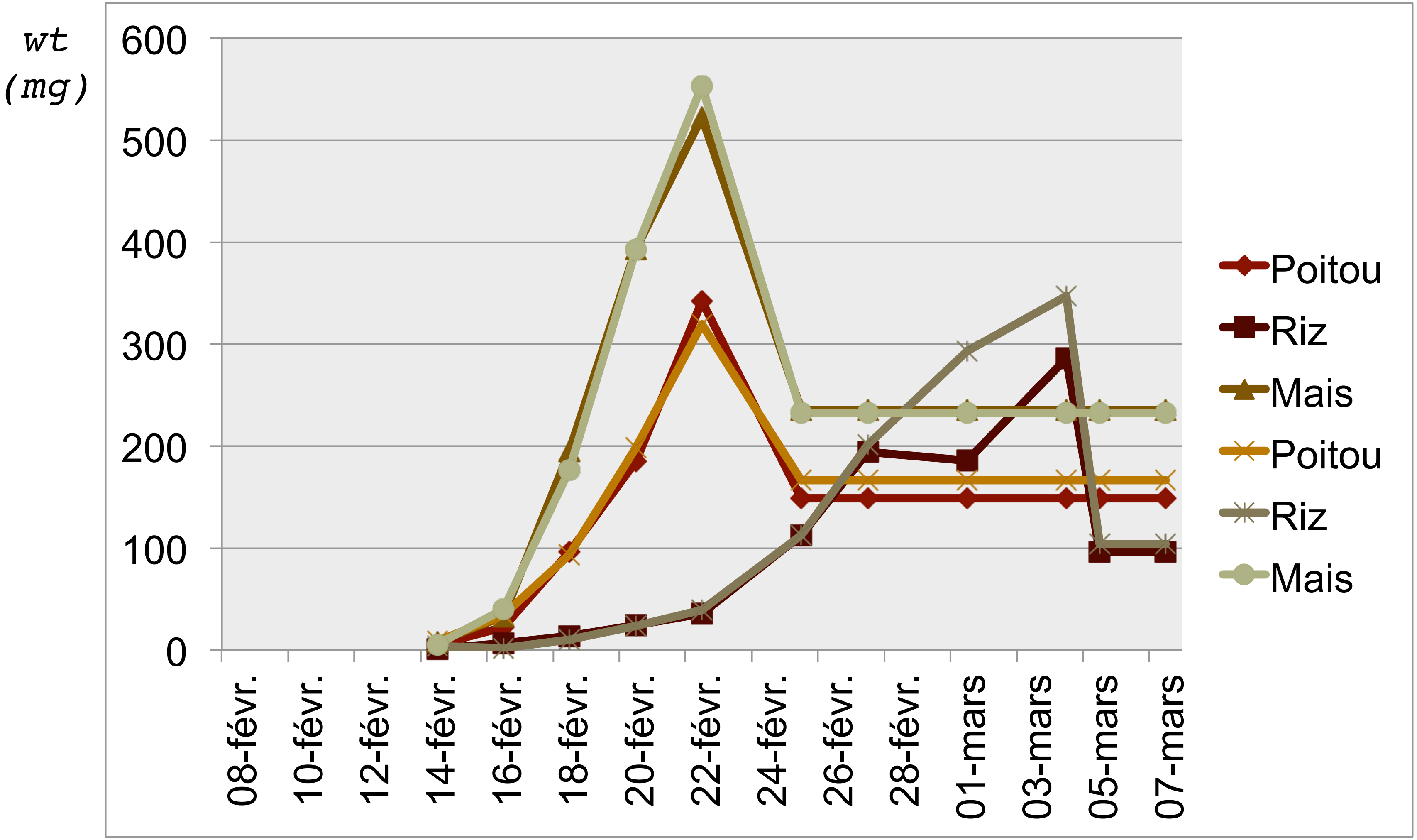
Experimental conditions





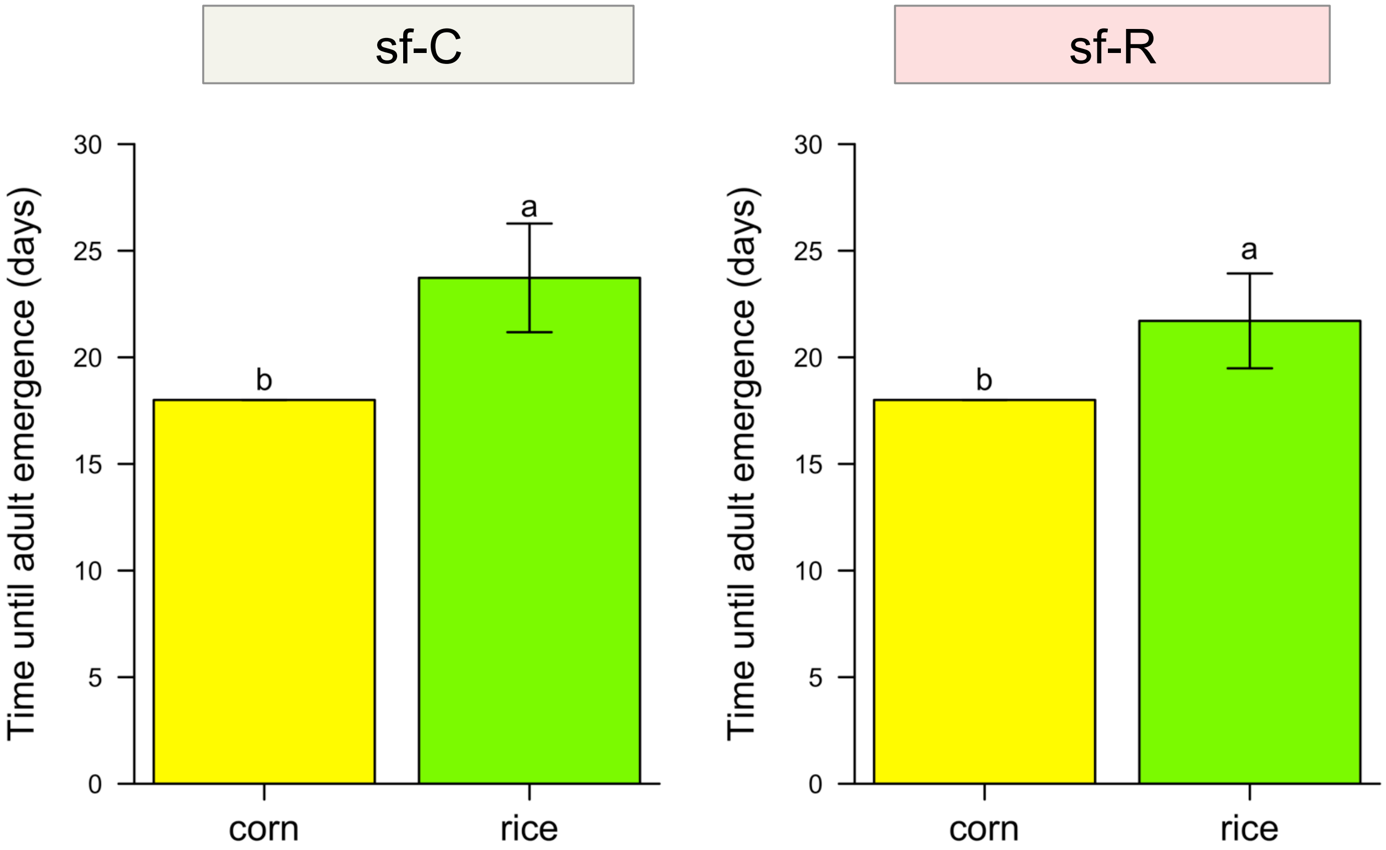
2 : developmental fitness in RT experiments

- 2 replicates of an RT experiment on corn, rice and Poitout
- parameters measured : weight (wt), developmental time (dt), survival (sv)



2 : developmental fitness in RT experiments

dt



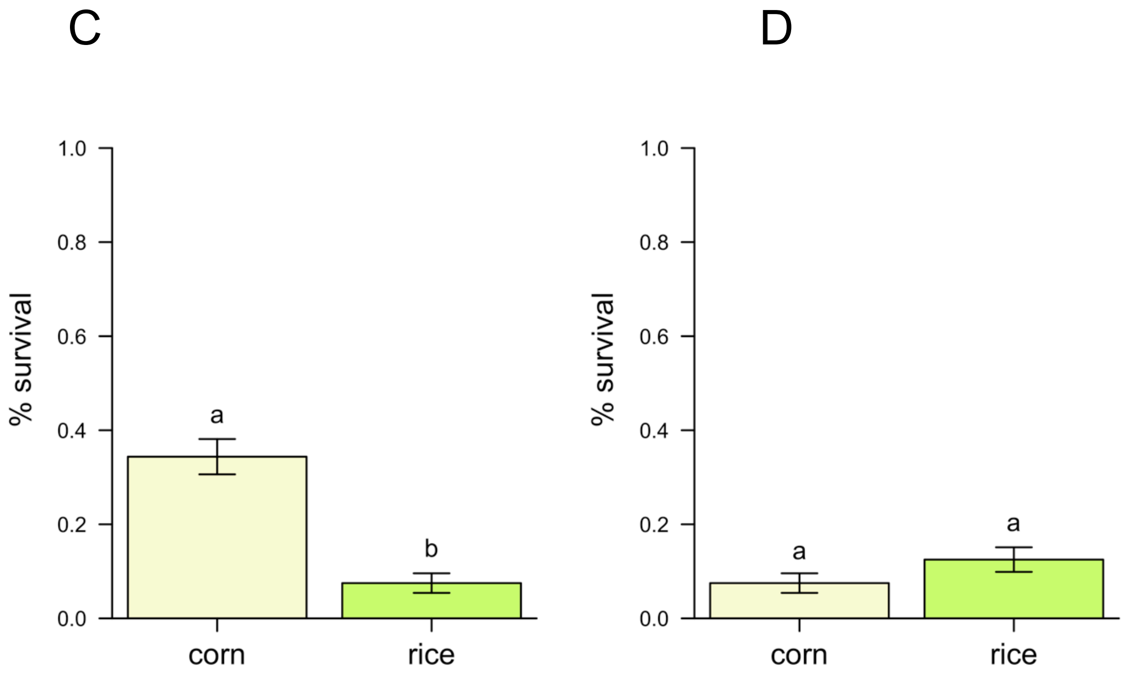
- Major diet effect on wt and dt
- Larvae of both strains gain more weight and develop faster on corn and Poitout than on rice. Is rice really a preferred host for *S. frugiperda* ?

2 : developmental fitness in RT experiments

sv

sf-C

sf-R



Better survival of sf-C on corn, maybe indicative of specialization

A multi-level integrative analysis in order to understand the mechanisms of speciation in relation to the host-plant adaptation

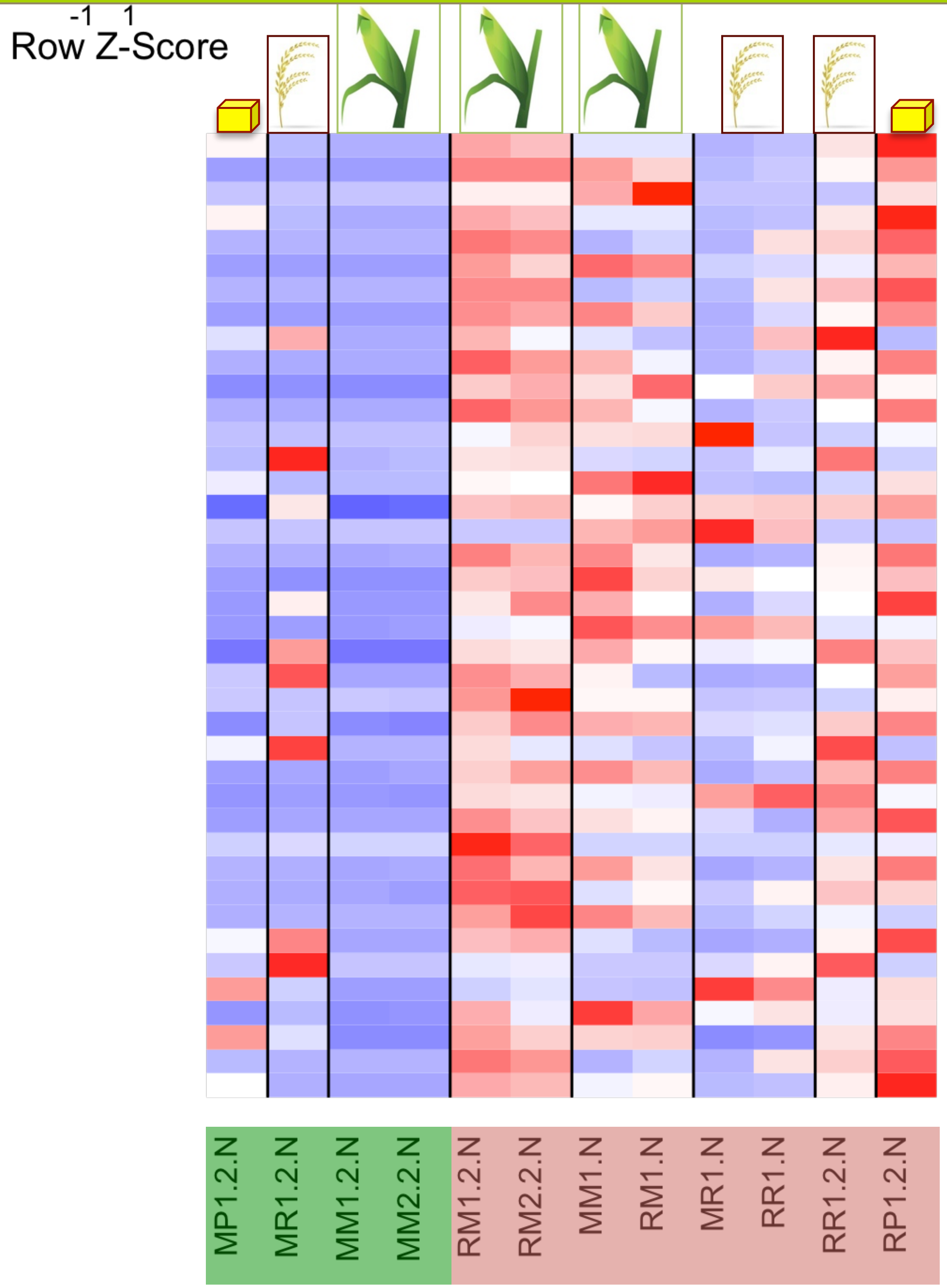
A. **Genetic differences** between the two strains. What is the estimated divergence and is it linked to agriculture ?

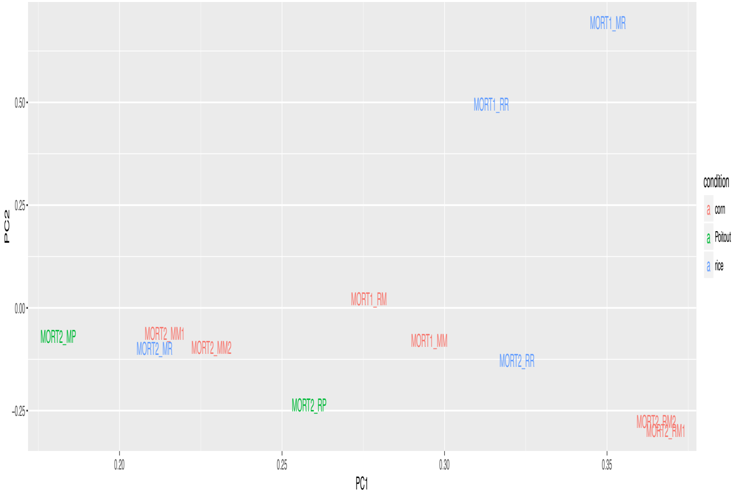
B. **Phenotypic differences** between the two strains. Is the host plant a trigger mechanism of speciation or reinforcement mechanism ?

C. **Gene expression differences** in response to environment. Build a comprehensive transcriptome and test the adaptation of *S. frugiperda* to its host plant.

D. **Epigenetic differences** : is the genetic plasticity linked to different chromatinian landscapes of *S. fru* ? Can holocentrism help to find the balance between deleterious deregulation and adaptability ?

Reciprocal transplant : RNAseq results

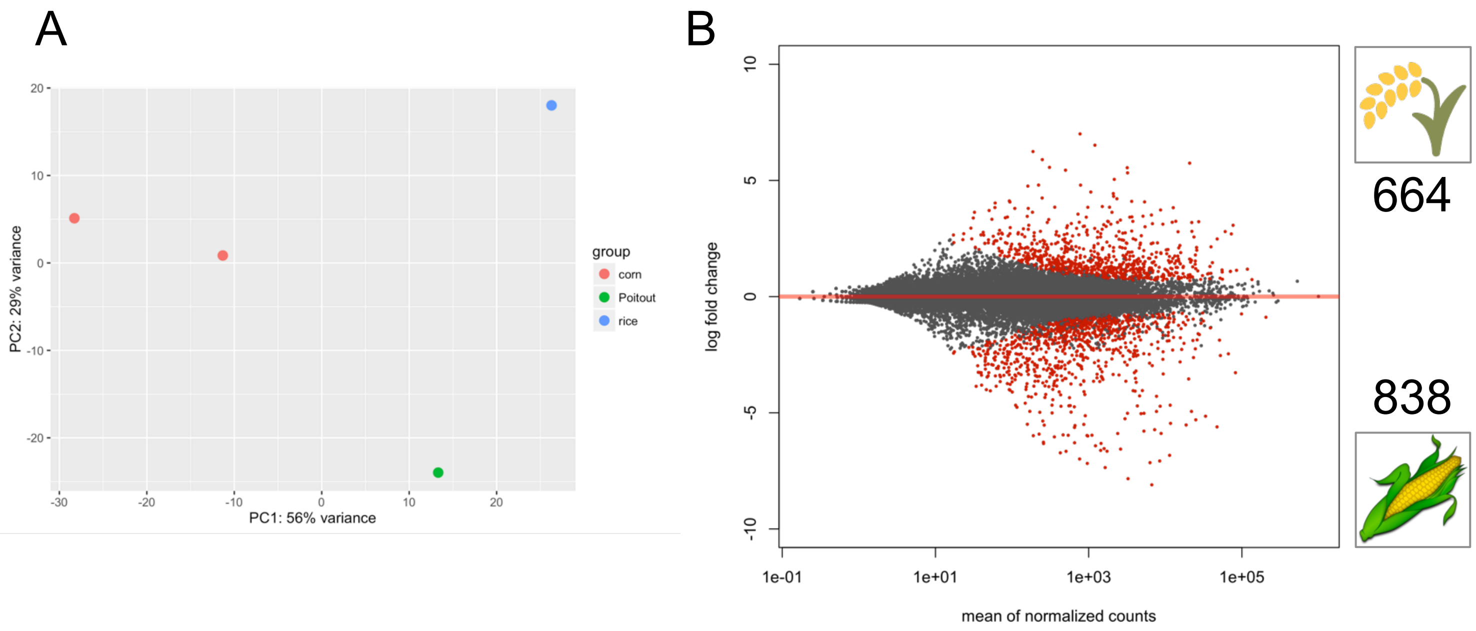




3 : sf-C adaptation to diet

RNAseq performed on pooled L4 larvae from the same RT experiment from LHT measurements.

DESeq2 analysis of rice vs corn at FDR 0.05



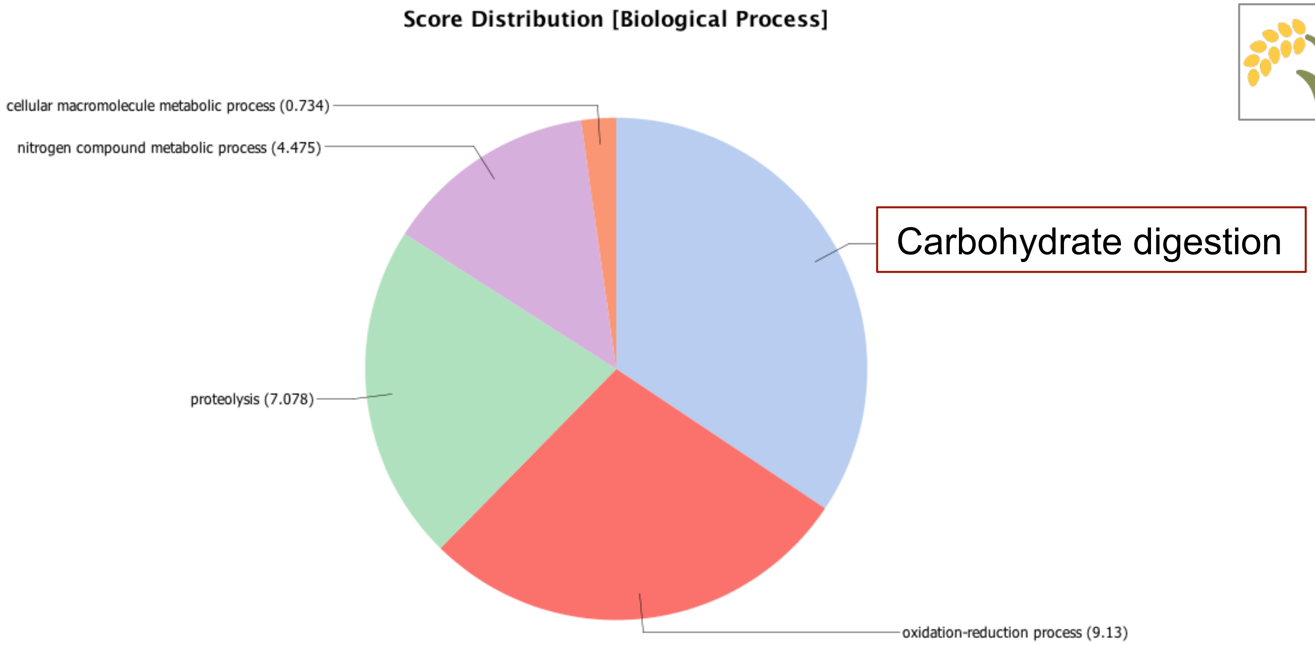
- Corn plant attracts most of the variation
- Poitout segregates with rice diet

- Manual annotation of 50UP and 50DOWN genes
- B2GO analysis on 100UP and 100DOWN genes

	baseMean	log2FoldChange	padj	TopHit (LepidoDB)	InterproScan
GSSPFG00007595001-RA	771,7308559	7,002580391	1,90E-43	peroxidase [Papilio polytes]	Haem peroxidase
GSSPFG00035209001.5-RA	1205,999892	6,517279966	8,37E-80	odorant degrading enzyme CXE3 [Sesamia inferens]	CCE
GSSPFG00034121001-RA	187,4611647	6,243434891	3,81E-26	363_100_1 protein [Mamestra configurata]	Mucin?
GSSPFG00024968001-RA	248,1256531	5,895962208	1,00E-29	insulin-like peptide 2 [Spodoptera exigua]	Insulin-like growth factor ?
GSSPFG00032729001.1-RA	20804,66838	5,744978051	2,90E-12	insect intestinal lipase 6 [Mamestra configurata]	Lipase
GSSPFG00018006001-RA	307,9337488	5,563220507	8,09E-33	glycogen synthase	glycogen synthase
GSSPFG00022903001-RA	3186,043324	5,547152817	1,25E-12	oxidase/peroxidase [Danaus plexippus]	Haem peroxidase
GSSPFG00035987001.2-RA	498,5874084	5,441953732	4,29E-25	Polycalin1	Polycalin 1
GSSPFG00016399001.2-RA	3217,700512	5,326524828	6,34E-83	zinc-containing alcohol dehydrogenase [Danaus plexippus]	D-arabinitol dehydrogenase 1
GSSPFG00019144001.1-RA	2159,161138	5,046075665	2,02E-65	phytanoyl-CoA dioxygenase, peroxisomal-like	Insects, such as the sumac flea beetle, are reported to use phytol and its metabolites (e.g. phytanic acid) as chemical deterrents against predation.[11] These compounds originate from host plants.
GSSPFG00026714001.5-RA	221,636688	4,800933492	1,64E-21	glutathione S-transferase epsilon 2 [Spodoptera litura]	GST - detox
GSSPFG00025431001-RA	783,3426892	4,798280771	6,86E-37	putative amine oxidase [Danaus plexippus]	
GSSPFG00032711001-RA	160,7988425	4,756047218	1,66E-10	Polycalin 1	Polycalin 1
GSSPFG00030891001-RA	1065,90604	4,501626838	3,57E-45	muscle M-line assembly protein unc-89-like	Ig like globin domains
GSSPFG00027822001.1-RA	2095,254553	4,436236977	6,06E-45	peritrophin type-A domain protein 2	receptor
GSSPFG00011317001-RA	2116,821517	4,315128134	2,47E-28	putative ecdysone oxidase	
GSSPFG00025591001-RA	247,1227549	4,243860424	3,96E-20	putative fatty acyl-CoA reductase CG5065 [Plutella xylostella]	
GSSPFG00018788001-RA	2335,127868	4,212089333	2,74E-45	trans-1,2-dihydrobenzene-1,2-diol dehydrogenase-like	
GSSPFG00023844001-RA	92,09295431	4,116683451	1,21E-08	/	Gcn5-related N-acetyltransferase (GNAT) dom
GSSPFG00004817001.2-RA	3641,996938	4,099821195	9,96E-30	polycalin [Mamestra configurata]	
GSSPFG00015277001.2-RA	5367,659831	4,075453282	1,40E-49	intestinal mucin SeM8 [Spodoptera exigua]	
GSSPFG00022138001-RA	93,6263094	4,072038784	2,09E-09	organic cation transporter protein-like [Bombyx mori]	transporter activity
GSSPFG00026250001-RA	60,15443759	4,027948368	1,34E-07	Bardet-Biedl syndrome 4 protein homolog [Bombyx mori]	cilium morphogenesis
GSSPFG00035966001.2-RB	1506,025918	3,876622747	1,01E-27	Polycalin1	Polycalin 1
GSSPFG00023437001-RA	360,7694025	3,84864175	7,85E-16	myrosinase 1-like [Bombyx mori]	Glycosyl hydrolase family 1
GSSPFG00035966001.2-RA	1871,022232	3,8478387	4,86E-20	Polycalin 1	Polycalin 1
GSSPFG00034818001-RA	974,2955837	3,844736745	3,55E-20	scarface, partial [Papilio xuthus]	
GSSPFG00002897001-RA	59,71326036	3,833745183	8,94E-07	putative inorganic phosphate cotransporter [Bombyx mori]	?
GSSPFG00033646001-RA	12709,69292	3,826954192	1,90E-43	beta-glucosidase precursor [Spodoptera frugiperda]	Glycosyl hydrolase family 1
GSSPFG00029358001-RA	162,6476021	3,796427064	1,88E-11	?	Prokaryotic membrane lipoprotein lipid attachment site profile.
GSSPFG00005312001-RA	108,5965601	3,795999276	8,15E-09	?	part of above ?
GSSPFG00004452001-RA	8673,792711	3,728632646	8,36E-34	maltase A1-like [Plutella xylostella]	Alpha amylase, catalytic domain
GSSPFG00016743001-RA	50,15548631	3,706311972	4,30E-06	None	None
GSSPFG00011130001.5-RA	184,6151307	3,630975894	2,79E-12	esterase E4-like [Plutella xylostella]	Carboxylesterase family
GSSPFG00005114001-RA	32,30722975	3,62847575	9,27E-05	glutamate--cysteine ligase regulatory subunit-like [Bombyx mori]	
GSSPFG00019553001-RA	2012,671731	3,628218759	4,81E-29	pancreatic triacylglycerol lipase-like [Bombyx mori]	Lipase
GSSPFG00013350001-RA	725,1343222	3,549141199	4,10E-25	putative aminopeptidase W07G4.4-like [Bombyx mori]	Cytosol aminopeptidase family
GSSPFG00033685001-RA	6594,281794	3,511858221	1,12E-21	None	None (putative TE)
GSSPFG00028708001.1-RA	5117,218721	3,460847581	3,24E-34	chitin binding domain 3 protein [Mamestra configurata]	eukaryotic translation initiation factor 3 ?
GSSPFG00001519001-RA	157,8094415	3,453198811	3,38E-10	antennae-specific aldehyde oxidase 2 [Amyelois transitella]	CO dehydrogenase flavoprotein
GSSPFG00004453001-RA	3664,654744	3,430405216	9,91E-27	sucrose hydrolase [Samia ricini]	MaltaseA1
GSSPFG00032884001-RA	87,01643164	3,425438649	3,89E-06		
GSSPFG00018427001-RA	3438,46307	3,420046724	1,46E-30		
GSSPFG00018639001-RA	930,2533342	3,41944504	1,96E-26		
GSSPFG00007332001-RA	81,49209275	3,393973201	0,000104096		
GSSPFG00026432001-RA	31,17865189	3,383948095	0,000391096		
GSSPFG00018007001-RA	54,54247622	3,327407346	3,97E-05		
GSSPFG00000142001-RA	300,9080501	3,315914051	1,94E-14		

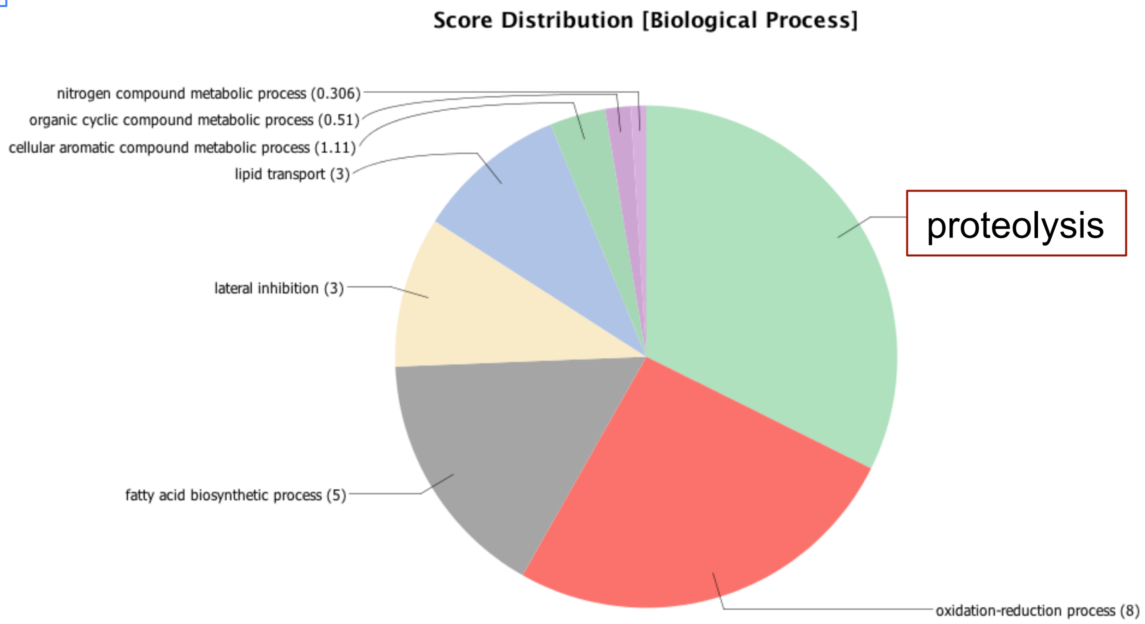
sf-C_RT2_100UP

= genes that are overexpressed on rice vs corn
OR
downregulated on corn vs rice



Carbohydrate digestion

sf-C_RT2_100DOWN

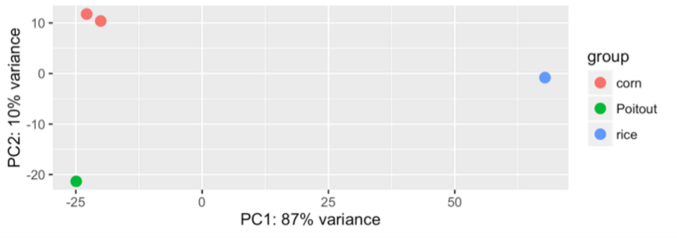


proteolysis

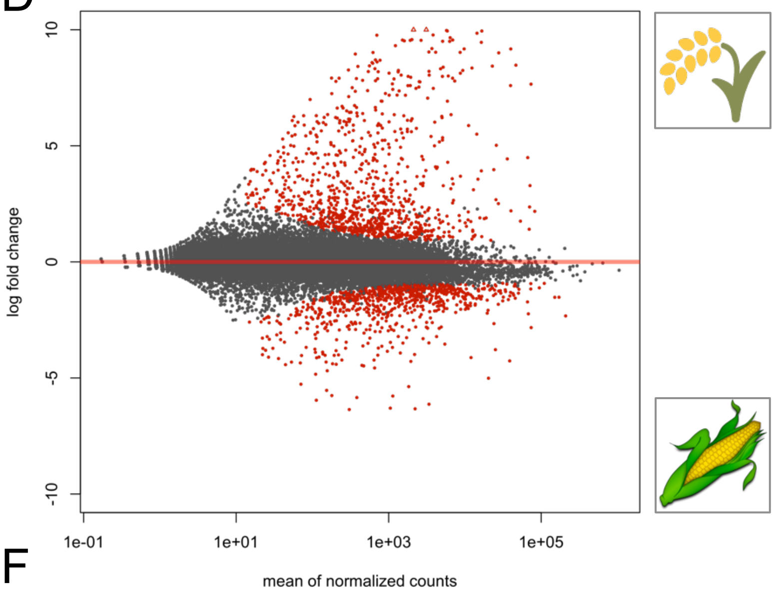
4 : sf-R adaptation to diet

RT2

C



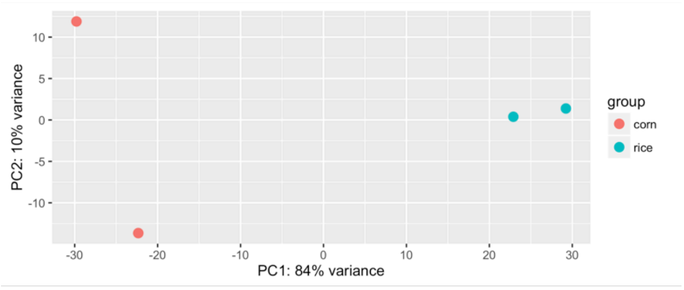
D



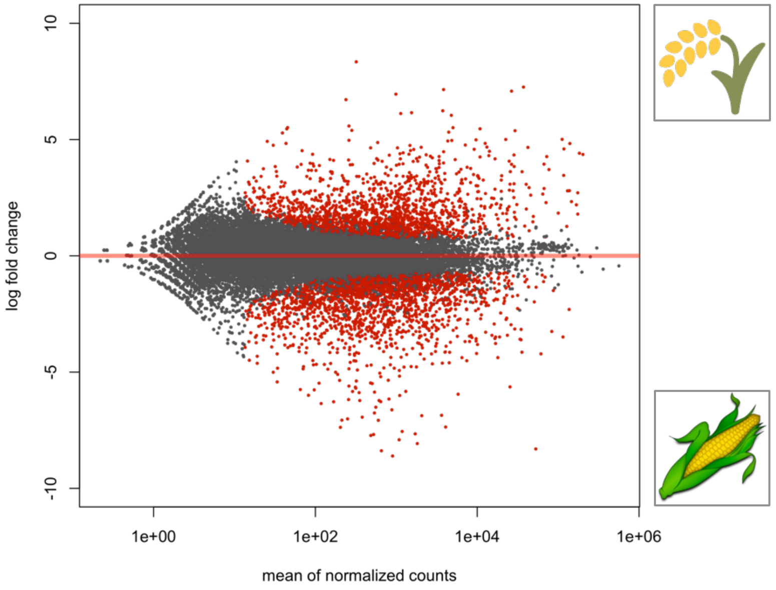
RT1

=2nd generation adaptation

E

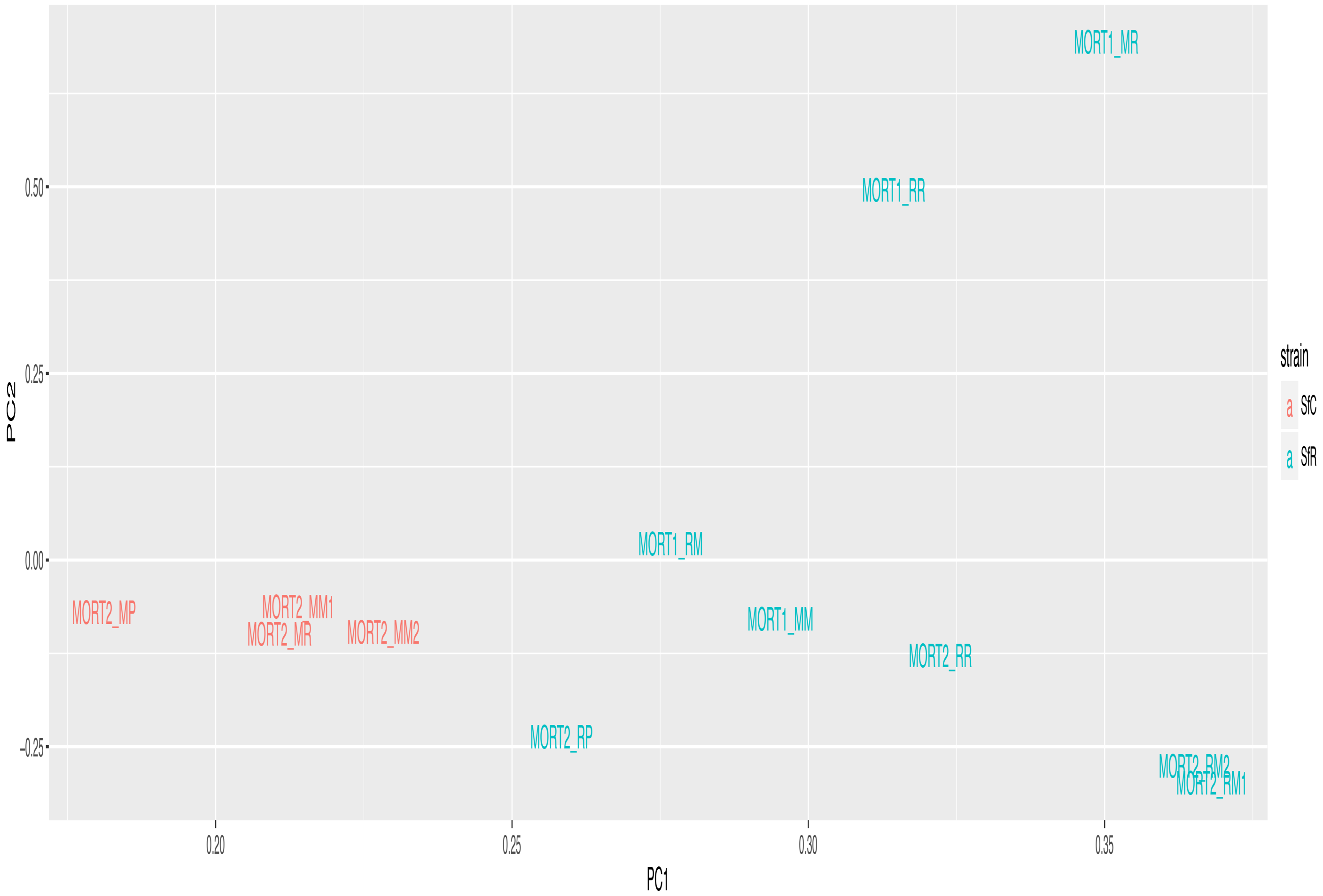


F



=>In progress

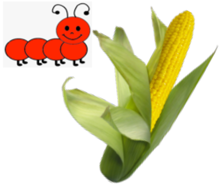
5: sf-C and sf-R transcriptional differences



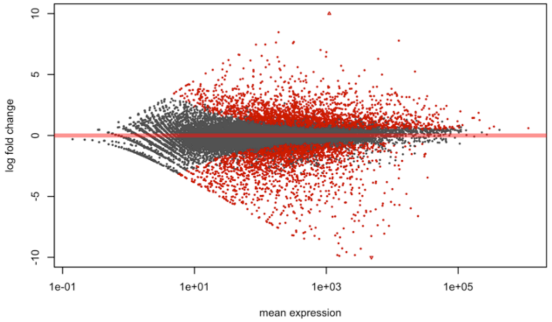
=> Manual annotation of sf-R vs sf-C

Reciprocal transplant : differential expression

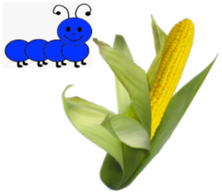
Spodoptera frugiperda

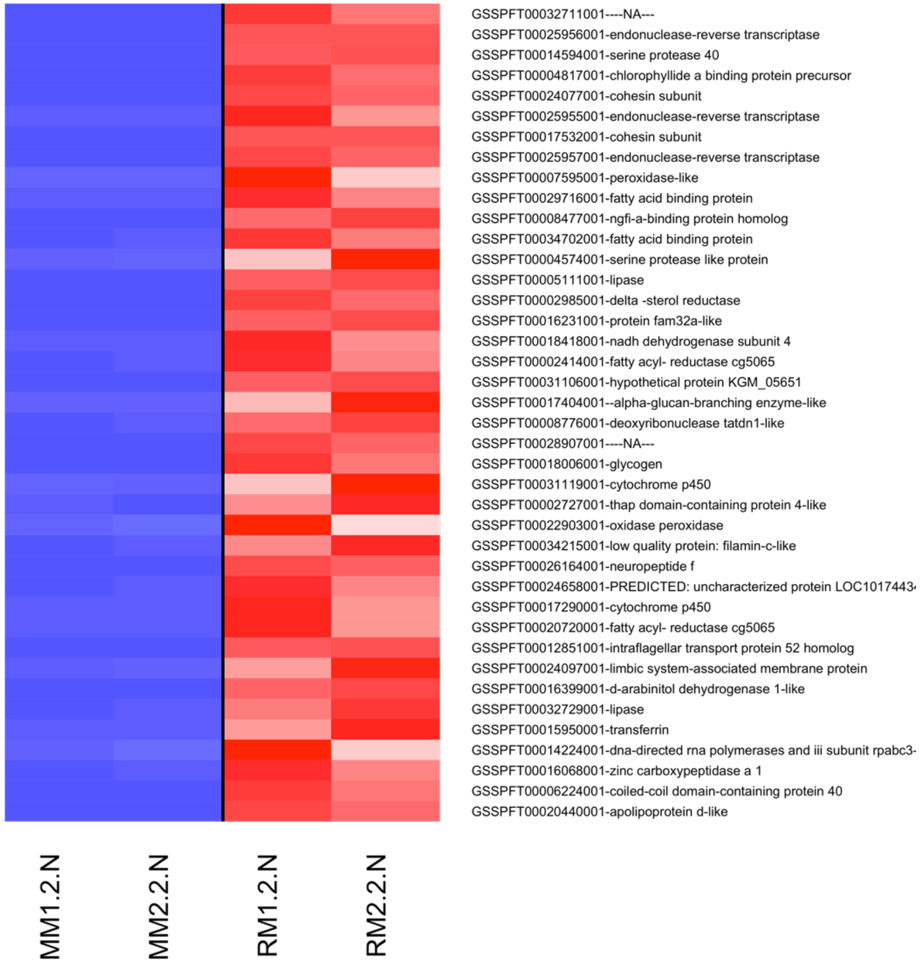


2476



2319





50UP in sfR

target_id	Annotation
GSSPFG00016764001-RA	cuticular protein hypothetical 15 precursor [Bombyx mori
GSSPFG00030121001.3-RA	chymotrypsin-like serine protease precursor, partial [Spodoptera frugiperda]
GSSPFG00026555001-RA	Glucose/ribitol dehydrogenase
GSSPFG00011516001-RA	vacuolar ATPase subunit C
GSSPFG00010065001-RA	unknown unsecreted protein
GSSPFG00029715001-RA	fatty acid-binding protein 2
GSSPFG00029716001.2-RA	fatty acid binding protein [Spodoptera litura]
GSSPFG00034702001-RA	fatty acid binding protein [Spodoptera litura]
GSSPFG00026559001-RA	Glucose/ribitol dehydrogenase family
GSSPFG00007853001-RA	muscle protein 20
GSSPFG00024881001-RA	proton-coupled folate transporter-like / adenylate cyclase
GSSPFG00018418001-RA	**mitochondrion part**
GSSPFG00034426001-RA	NADH:ubiquinone oxidoreductase subunit B14.5a (Complex I-B14.5a)
GSSPFG00019425001-RA	alpha-tocopherol transfer protein-like
GSSPFG00035519001.3-RA	Chymotrypsin serine protease family (S1)
GSSPFG00003434001-RA	uncharacterized protein LOC101736429 [Bombyx mori]
GSSPFG00016702001-RA	Asparaginase/glutaminase family
GSSPFG00010063001-RA	Xaa-Pro aminopeptidase
GSSPFG00028883001-RA	**ubiquitin conjugating enzyme 4**
GSSPFG00016763001-RA	cuticular protein hypothetical 16 precursor [Bombyx mori
GSSPFG00026310001-RA	methionine aminopeptidase 1
GSSPFG00021535001-RA	protein PTCD3 homolog, mitochondrial
GSSPFG00029999001-RA	differentially expressed in FDCP 8 homolog [Bombyx mori]
GSSPFG00001834001.5-RA	antennal esterase CXE3 [Spodoptera littoralis]
GSSPFG00023266001-RA	PREDICTED: lipopolysaccharide-induced tumor necrosis factor-alpha factor homolog [Bombyx mori]
GSSPFG00014594001.3-RA	Diverged serine protease type- 4
GSSPFG00004122001-RA	26S proteasome non-ATPase regulatory subunit 13 [Bombyx mori
GSSPFG00020932001-RA	Growth-Arrest-Specific Protein 2
GSSPFG00019591001-RA	PREDICTED: rRNA methyltransferase 3A, mitochondrial [Plutella xylostella]
GSSPFG00013608001-RA	lysosomal thiol reductase IP30
GSSPFG00004511001-RA	PREDICTED: uncharacterized protein LOC101743815 [Bombyx mori]
GSSPFG00029944001-RA	COP9/Signalosome and eIF3 complex-shared subunit 1 [EIF3M].
GSSPFG00020580001-RA	unknown secreted protein
GSSPFG00013284001-RA	putative protein kinase C and casein kinase substrate in neurons protein [Danaus plexippus]
GSSPFG00029721001-RA	membrane-bound protein
GSSPFG00004617001-RA	membrane-bound protein
GSSPFG00017532001-RA	putative cohesin subunit
GSSPFG00026557001-RA	Glucose/ribitol dehydrogenase family
GSSPFG00019552001-RA	PREDICTED: lipase member H-A-like [Plutella xylostella]
GSSPFG00024077001-RA	putative cohesin subunit
GSSPFG00016612001-RA	ATP-binding cassette, ABC transporter-type
GSSPFG00028809001-RA	Kelch-like protein diablo
GSSPFG00006189001-RA	Domain of unknown function (DUF4508)
GSSPFG00033014001-RA	CRAL/TRIO domain/ MSP (Major sperm protein) domain
GSSPFG00008997001-RA	PRMT5 arginine-N-methyltransferase
GSSPFG00034511001-RA	/
GSSPFG00021034001-RA	GrpE nucleotide exchange factor
GSSPFG00030981001-RA	Protein msta/ histone deacetylase binding
GSSPFG00025956001-RA	SlackLINE
GSSPFG00005936001-RA	stromal interaction molecule 1 precursor [Bombyx mori

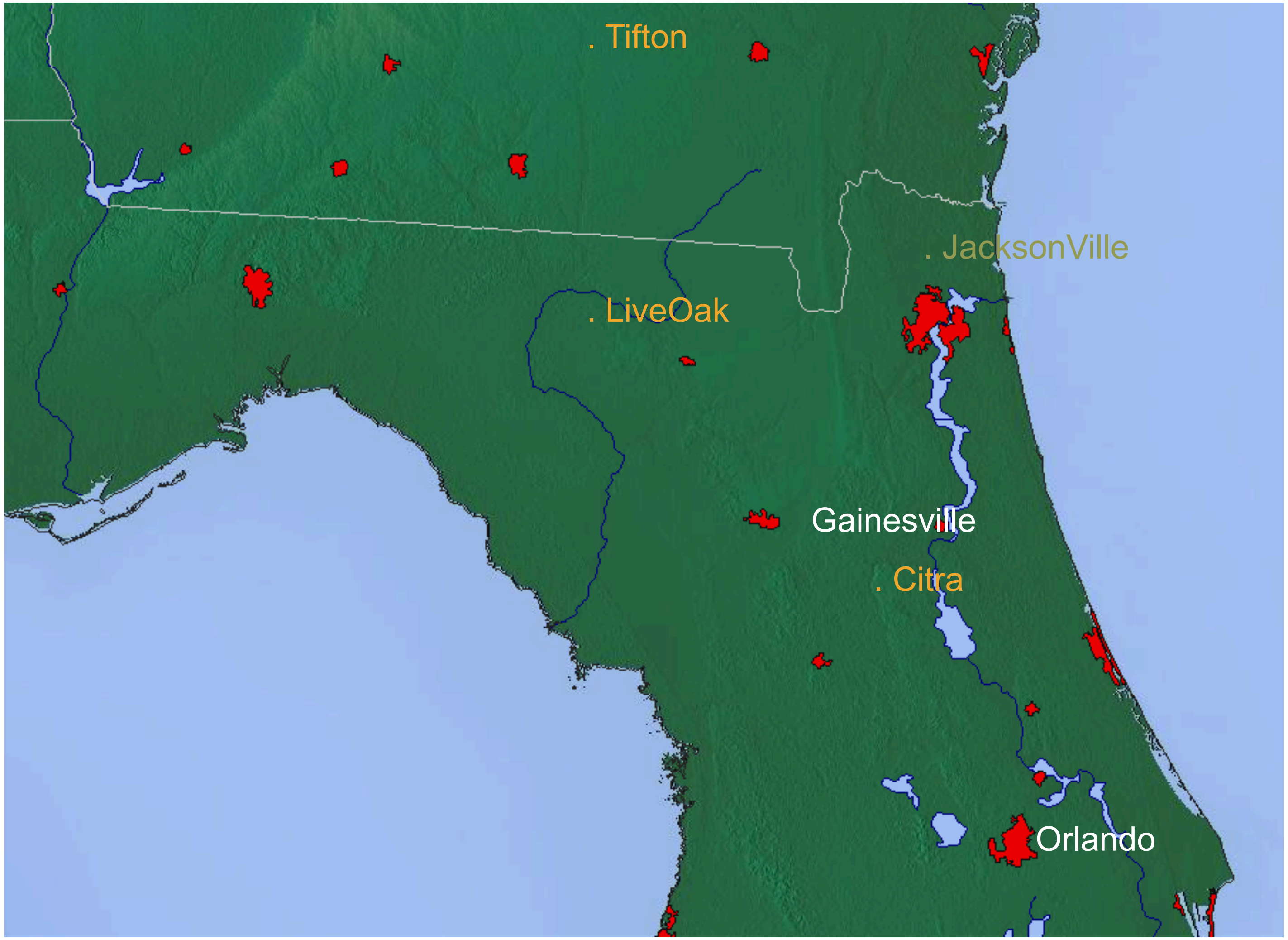
RNAseq validation in natural populations of *Spodoptera frugiperda*

. Citra
    sweet corn
    57 L4

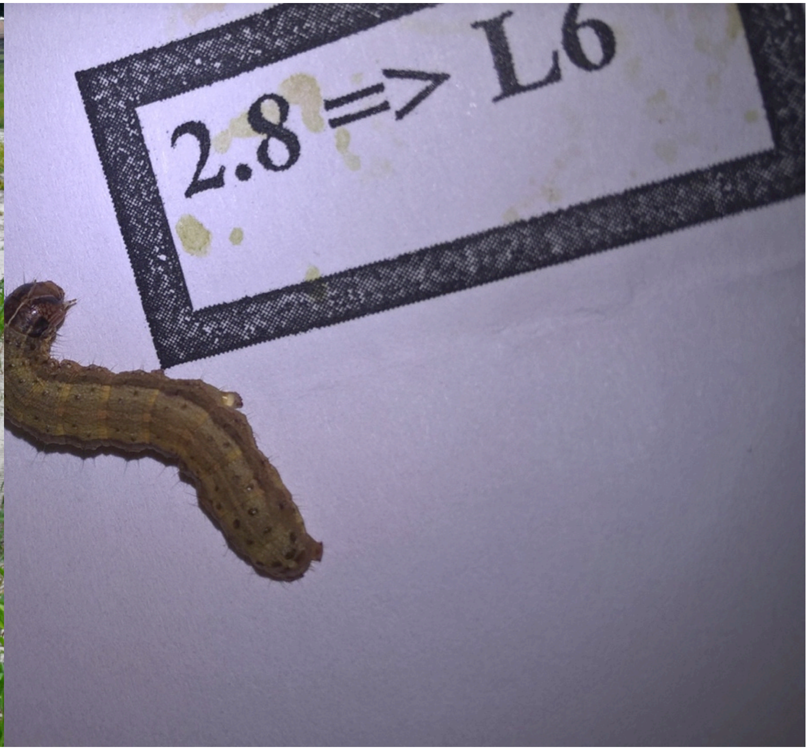
. LiveOak
    sorghum
    3 L4

. JacksonVille
    Pasture Grass
    21 L4

. Tifton
    Volunteer corn
    34 L4

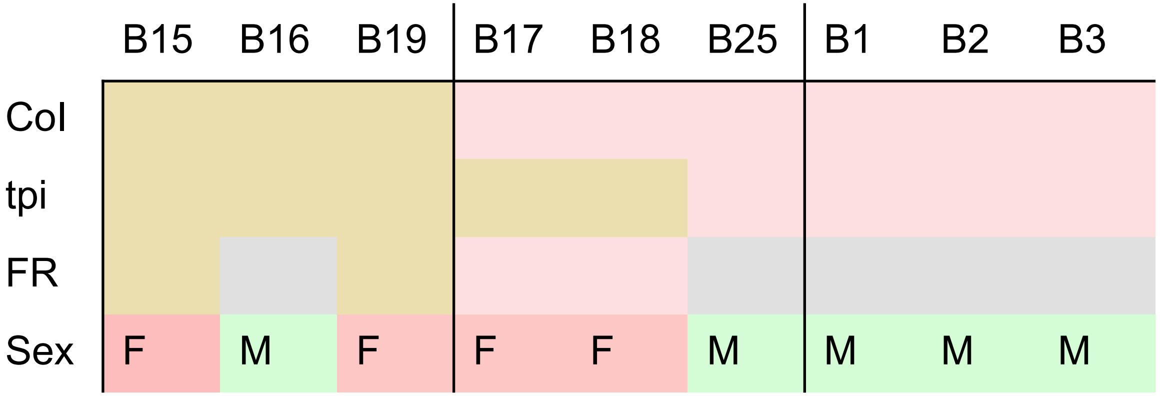


Field Collections : Sept 2015

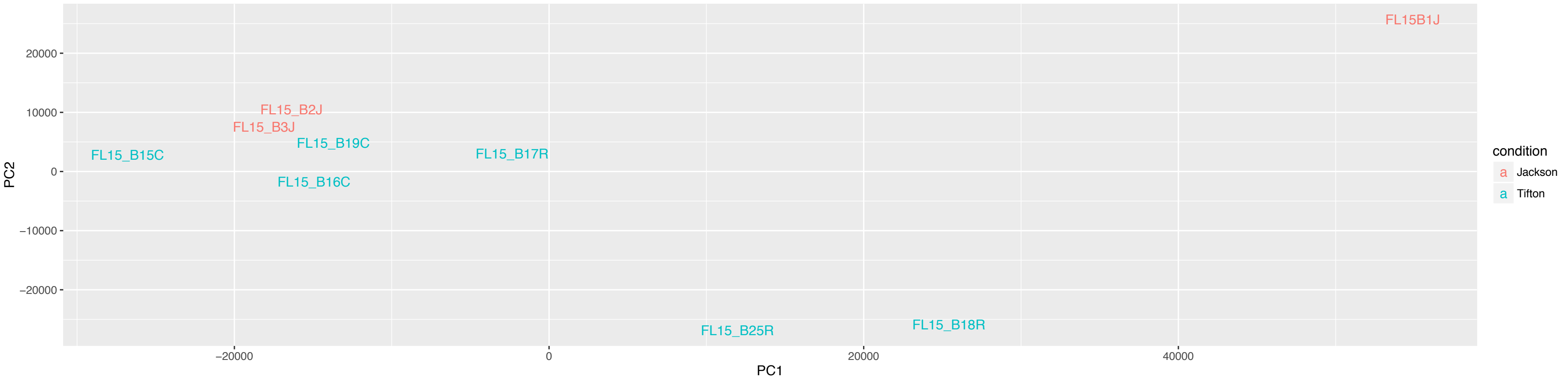


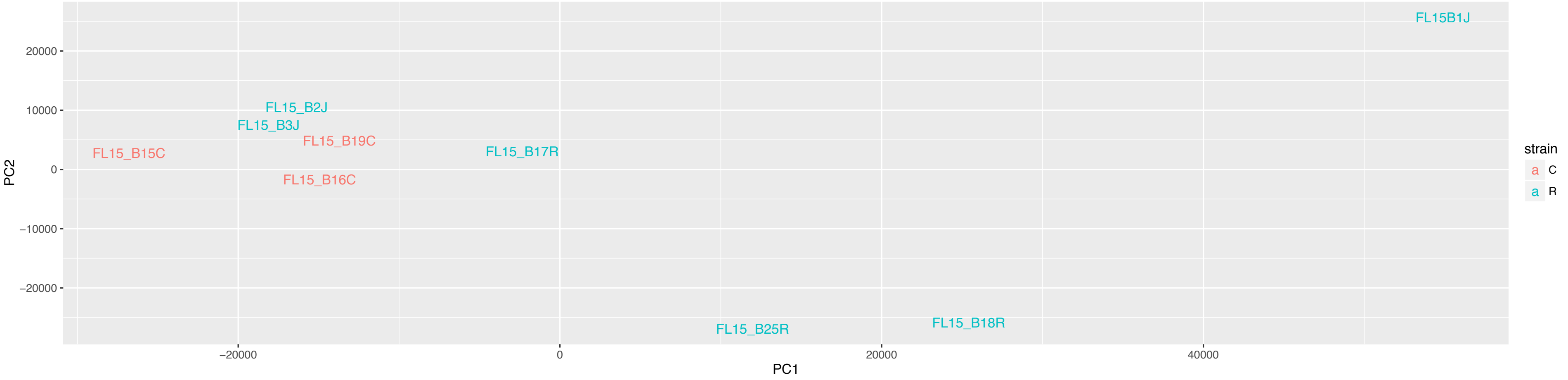
- Collect Sf larvae in dedicated experimental corn fields
- Check their stage based on the head capsule => L4
- Bring back frozen material in RNAlater
- Extract both DNA and RNA for ~50 individuals
- Genotype based on CoI and Tpi
- Investigate ~12 candidates by qPCR

RNAseq on 6 individuals from Tifton (corn field) and 3 from Jacksonville (pasture grass)



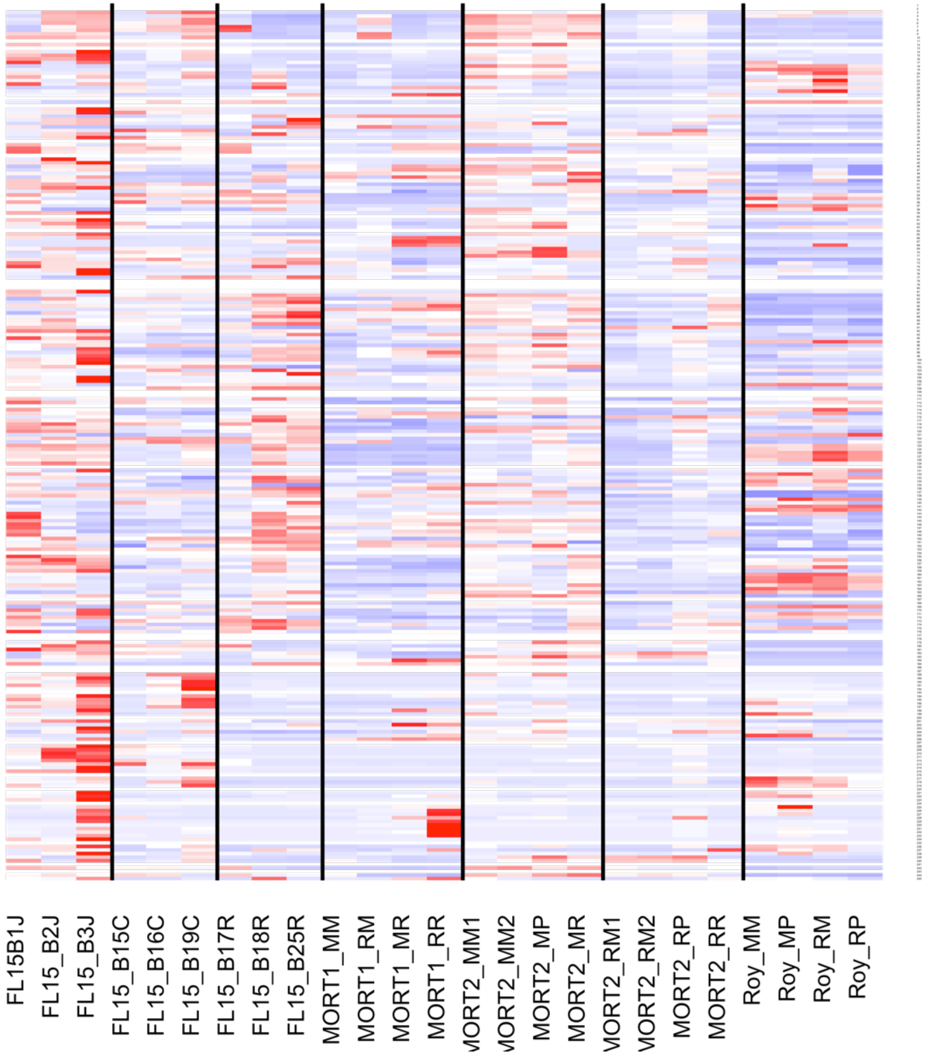
	B15	B16	B19	B17	B18	B25	B1	B2	B3
Col	🟨	🟨	🟨	🟥	🟥	🟥	🟥	🟥	🟥
tpi	🟨	🟨	🟨	🟨	🟨	🟥	🟥	🟥	🟥
FR	🟨	⬜	🟨	🟥	🟥	⬜	⬜	⬜	⬜
Sex	F	M	F	F	F	M	M	M	M



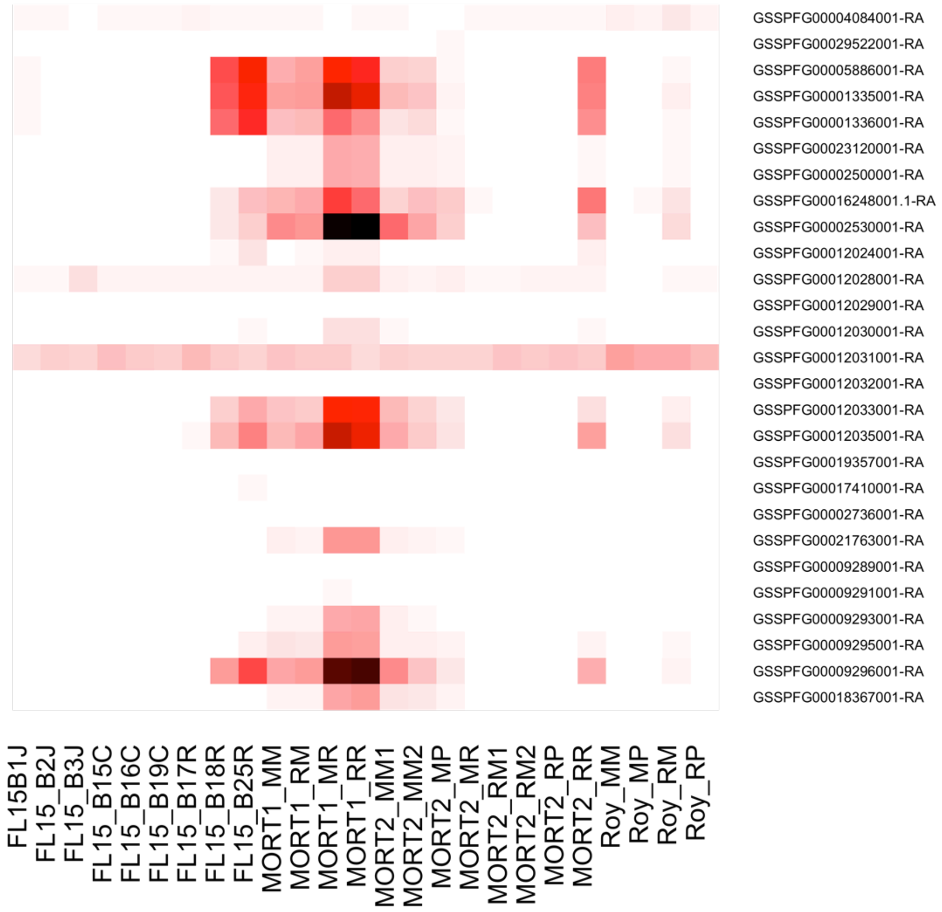


8: gene families

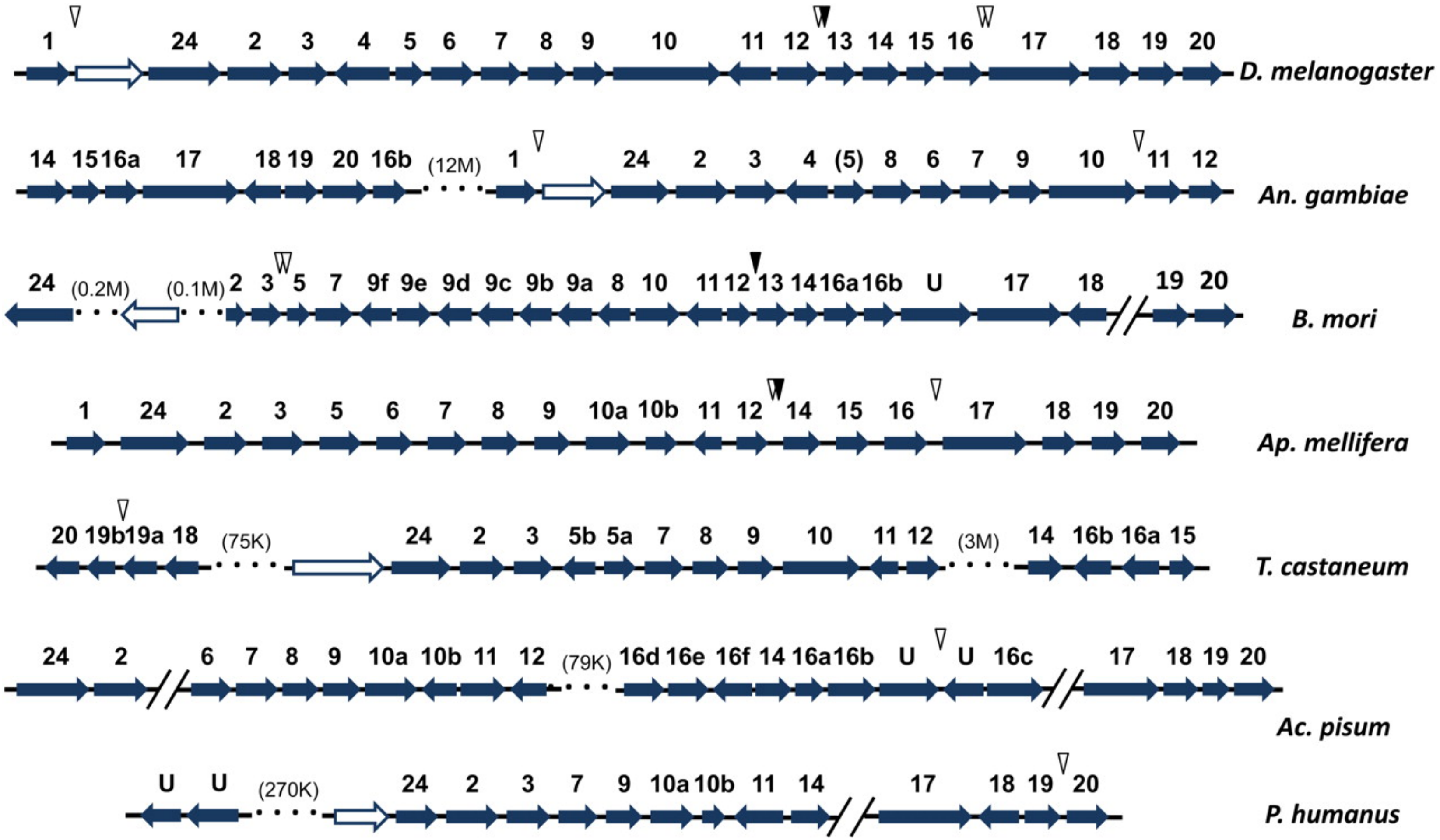
immunity genes



Osiris

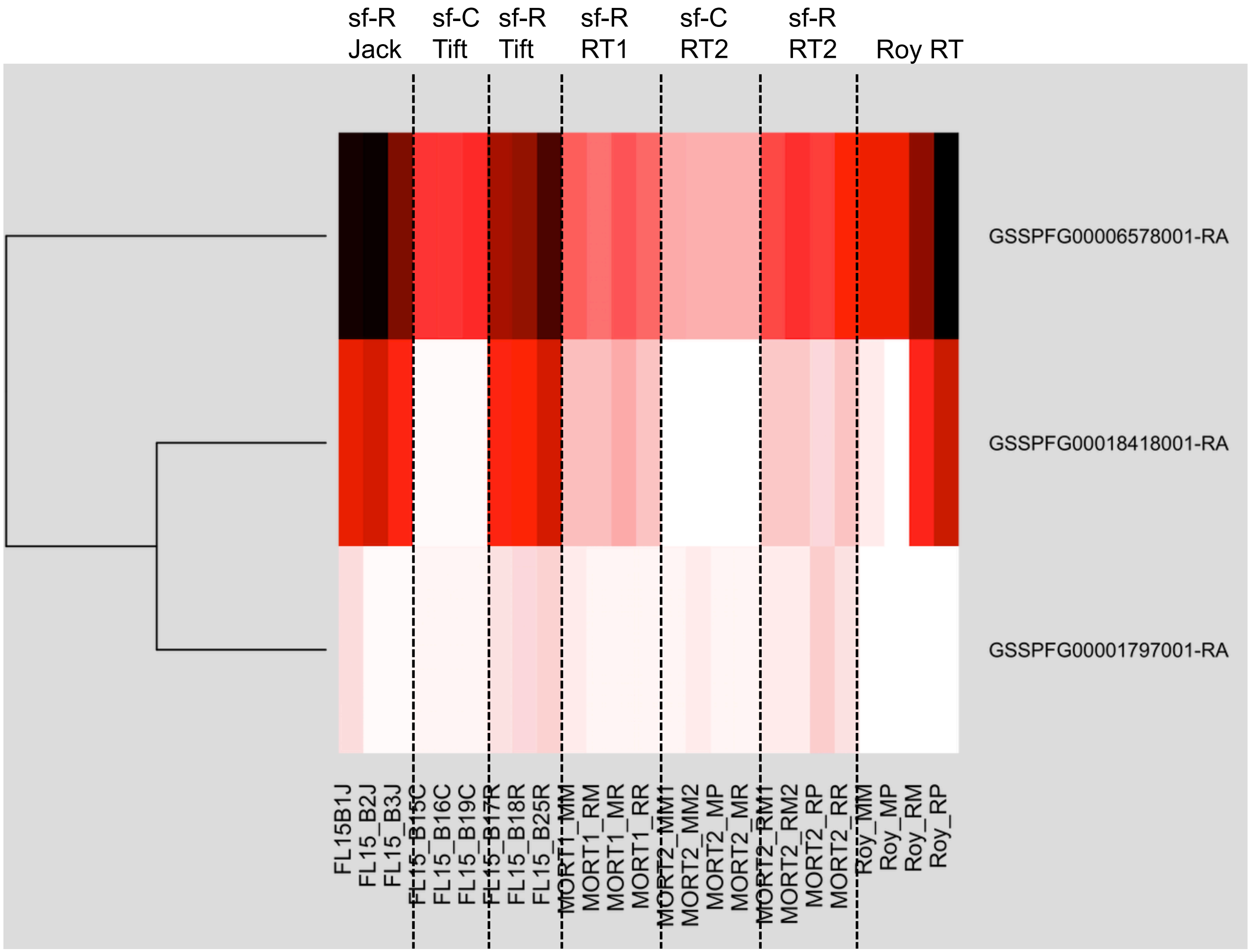

GSSPFG00004084001-RA
GSSPFG00029522001-RA
GSSPFG00005886001-RA
GSSPFG00001335001-RA
GSSPFG00001336001-RA
GSSPFG00023120001-RA
GSSPFG00002500001-RA
GSSPFG00016248001.1-RA
GSSPFG00002530001-RA
GSSPFG00012024001-RA
GSSPFG00012028001-RA
GSSPFG00012029001-RA
GSSPFG00012030001-RA
GSSPFG00012031001-RA
GSSPFG00012032001-RA
GSSPFG00012033001-RA
GSSPFG00012035001-RA
GSSPFG00019357001-RA
GSSPFG00017410001-RA
GSSPFG00002736001-RA
GSSPFG00021763001-RA
GSSPFG00009289001-RA
GSSPFG00009291001-RA
GSSPFG00009293001-RA
GSSPFG00009295001-RA
GSSPFG00009296001-RA
GSSPFG00018367001-RA
FL15B1J
FL15_B2J
FL15_B3J
FL15_B15C
FL15_B16C
FL15_B19C
FL15_B17R
FL15_B18R
FL15_B25R
MORT1_MM
MORT1_RM
MORT1_MR
MORT1_RR
MORT2_MM1
MORT2_MM2
MORT2_MP
MORT2_MR
MORT2_RM1
MORT2_RM2
MORT2_RP
MORT2_RR
Roy_MM
Roy_MP
Roy_RM
Roy_RP

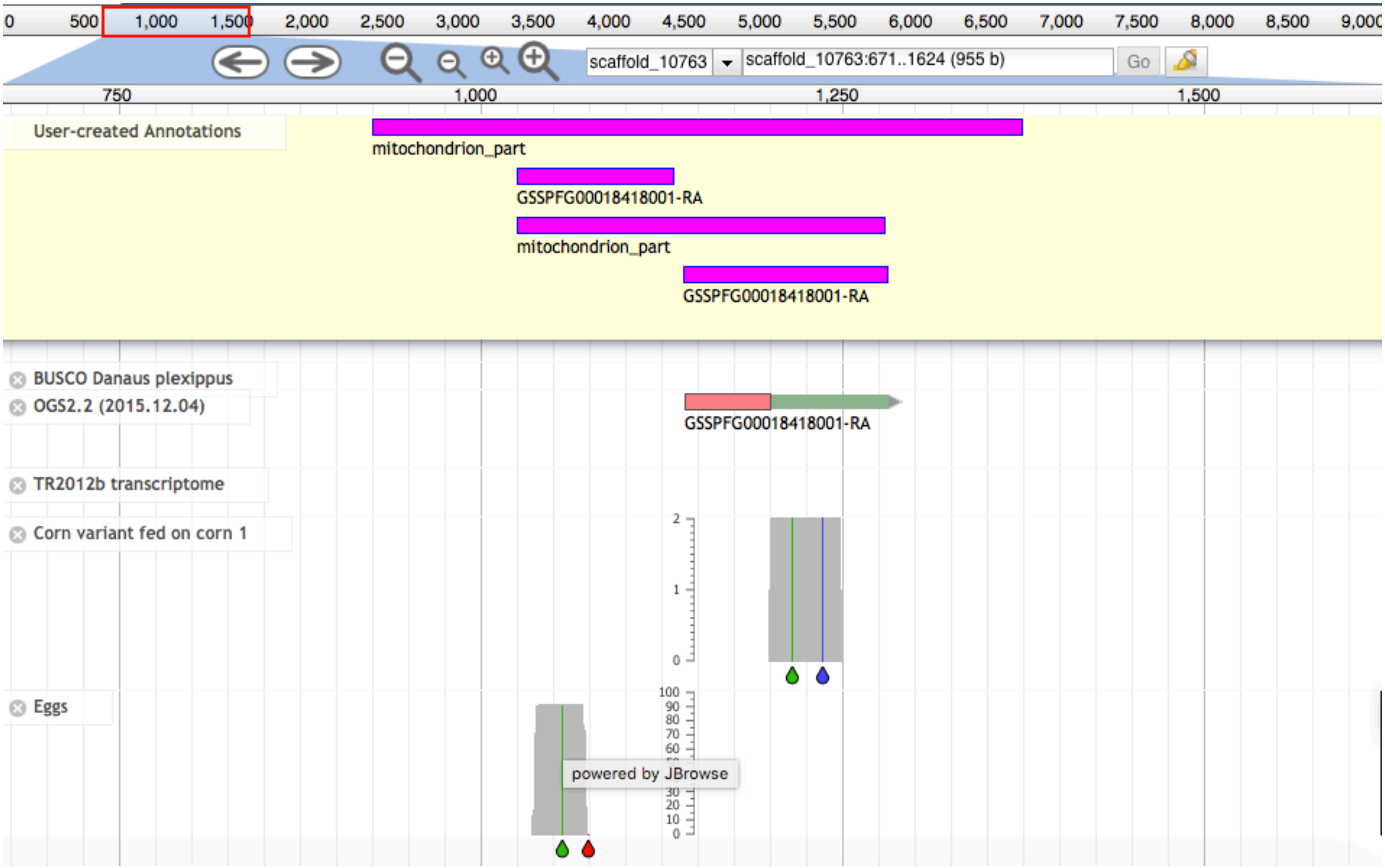
Synteny of the Osiris gene cluster.



Neethu Shah et al. G3 2012;2:313-319

3 genes overexpressed in sfR vs sfC





scaffold_10763:671..1624 (955 b)

Epigenetics, Holocentrism and Adaptation (EHA)

DGIMI

Diversité, Génomes & Interactions
Microorganismes Insectes

•**Emmanuelle d'Alençon**	DR2 INRA
•Réjane Streiff	CR1 INRA
•**Sylvie Gimenez**	TR INRA
•Imène Seninet	IE2 INRA
•***Marion Orsucci***	*Doct.*
•Sandra Nihm	Doct.
•Kiwoong Nam	CR2 INRA
•Rima Naït-Saidi	Stagiaire M1

Collaborators

Fabrice Legeai - BIPAA, Rennes
Anthony Bretaudeau - BIPAA, Rennes
Rob Meagher - ARS-USDA, Gainesville, Florida
Rod Nagoshi - ARS-USDA, Gainesville, Florida
Astrid Groot § Sabine Haenniger - Amsterdam
Gaël Kergoat - CBGP, Montpellier
Marie Frayssinet – DGIMI, Montpellier

http://www6.montpellier.inra.fr/dgimi/

https://dgimiflorida2015.wordpress.com/

A multi-level integrative analysis in order to understand the mechanisms of speciation in relation to the host-plant adaptation

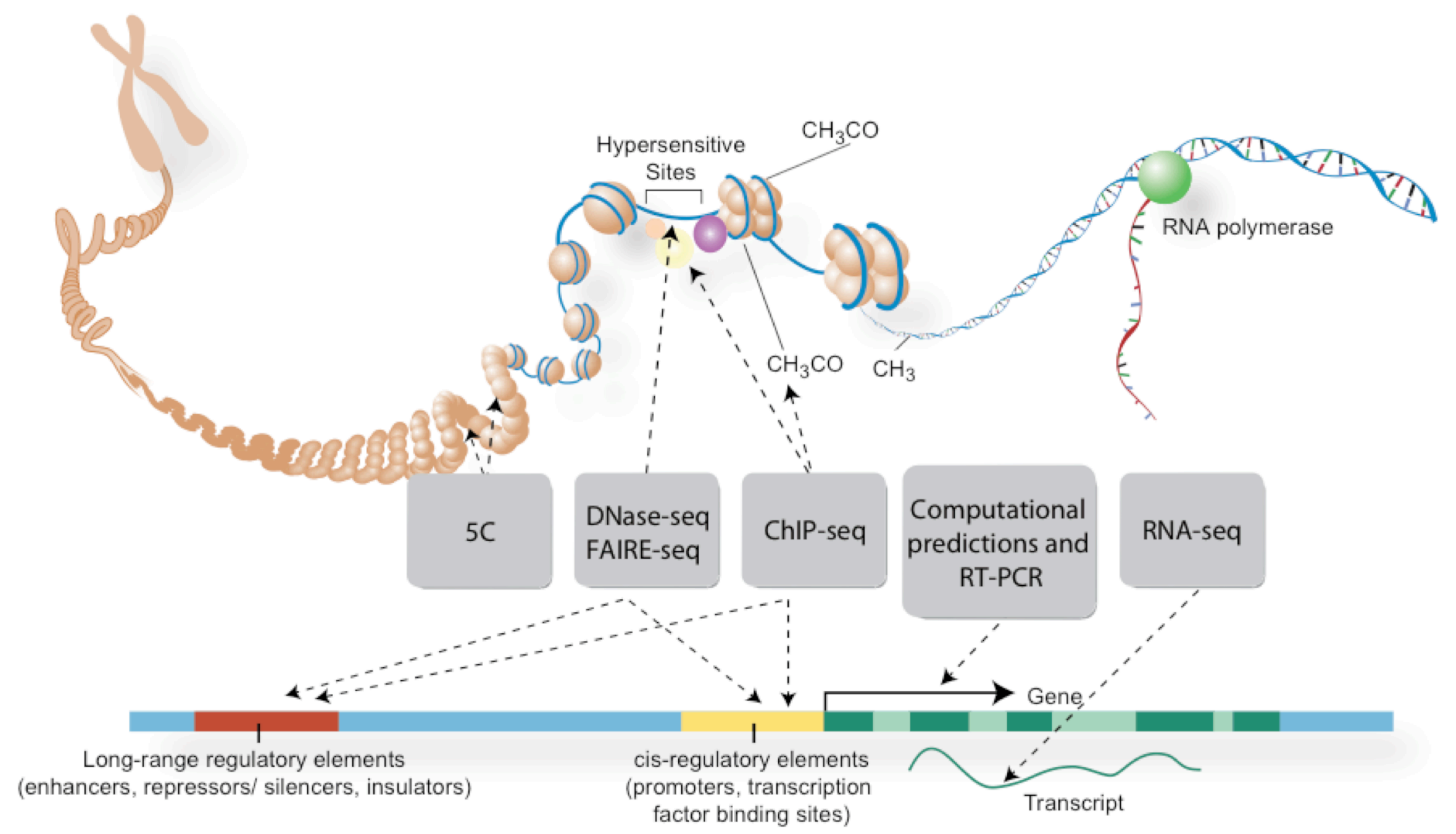
A. **Genetic differences** between the two strains. What is the estimated divergence and is it linked to agriculture ?

B. **Phenotypic differences** between the two strains. Is the host plant a trigger mechanism of speciation or reinforcement mechanism ?

C. **Gene expression differences** in response to environment. Build a comprehensive transcriptome and test the adaptation of *S. frugiperda* to its host plant.

D. **Epigenetic differences** : is the genetic plasticity linked to different chromatinian landscapes of *S. fru* ? Can holocentrism help to find the balance between deleterious deregulation and adaptability ?

FAIREseq identifies cis-regulatory element as well as long-range regulatory elements



http://genome.genetics.rutgers.edu/encode/

Use Histone marks to annotate regulatory elements of *Sf* genomes

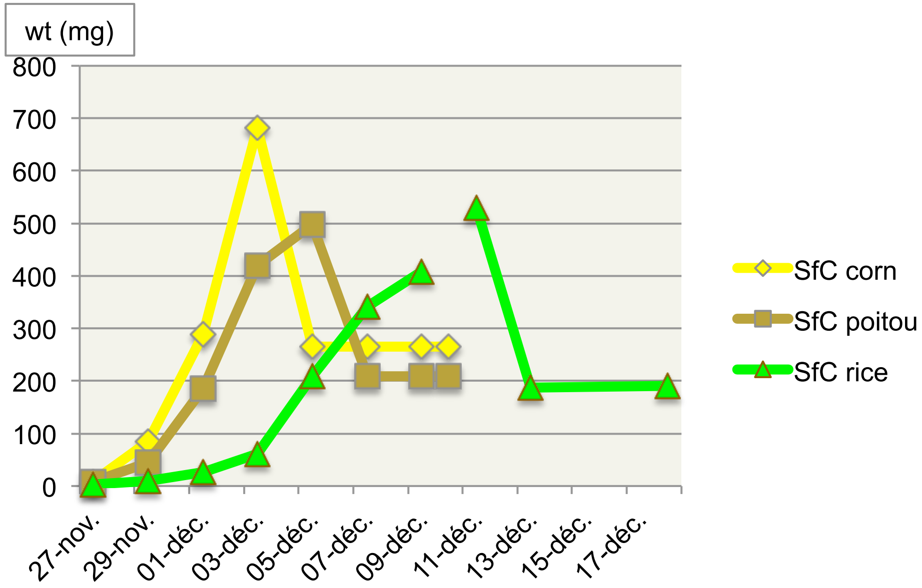
H3K9me2 on Sf9 cell line : associated to genes transcription unit.

2 : developmental fitness in RT experiments

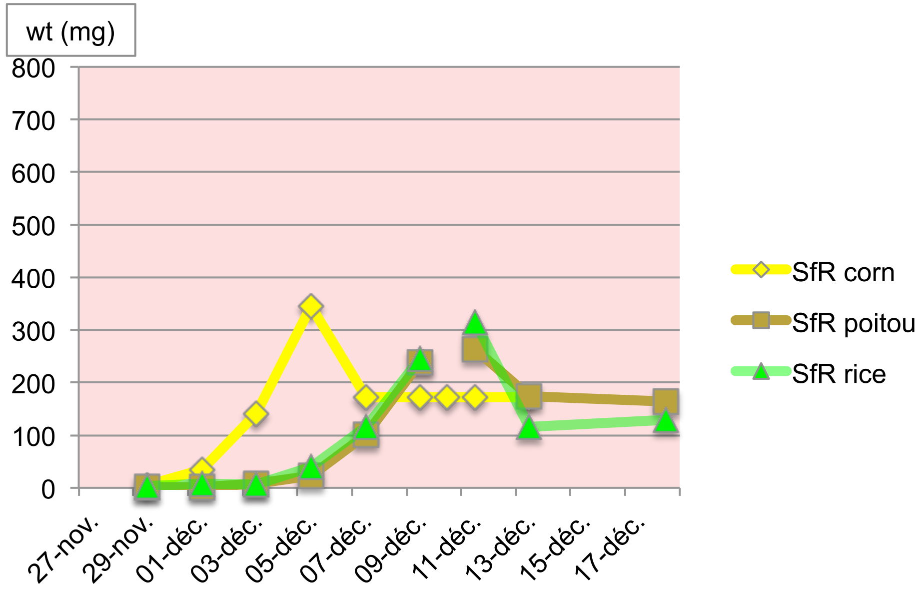
- 2 replicates of an RT experiment on corn, rice and Poitout
- parameters measured : weight (wt), developmental time (dt), survival (sv)

`wt`
Replicate 1

sf-C



sf-R



- Major diet effect on wt and dt
- Larvae of both strains gain more weight and develop faster on corn and Poitout than on rice. Is rice really a preferred host for *S. frugiperda* ?