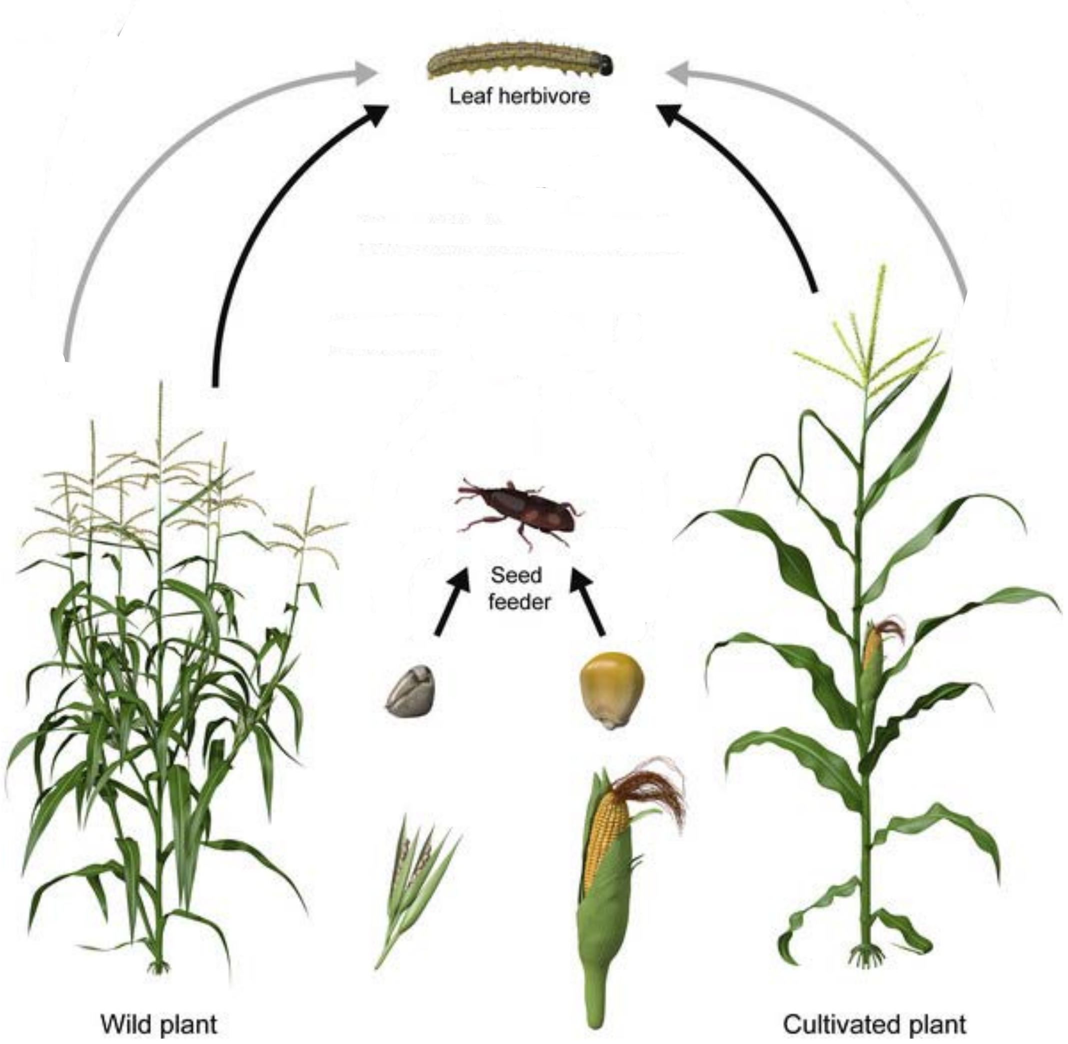
Influence of historical land use and modern agricultural expansion on the spatial and ecological divergence of sugarcane borer, *Diatraea saccharalis* (Lepidoptera: Crambidae) in Brazil

PhD Genetics and Molecular Biology - UNICAMP
Maria Imaculada Zucchi - UNICAMP
Celso Omoto - ESALQ/ USP
Andrew P. Michel - The Ohio State University

Vitor Pavinato
Postdoc

Plant domestication, **introduction** of non-native crop plants and **insect invasion** to new habitats has greatly shaped the evolutionary history of many phytophagous insect species by providing new niches to explore.[1]



---

[1] Kim and McPheron 1993

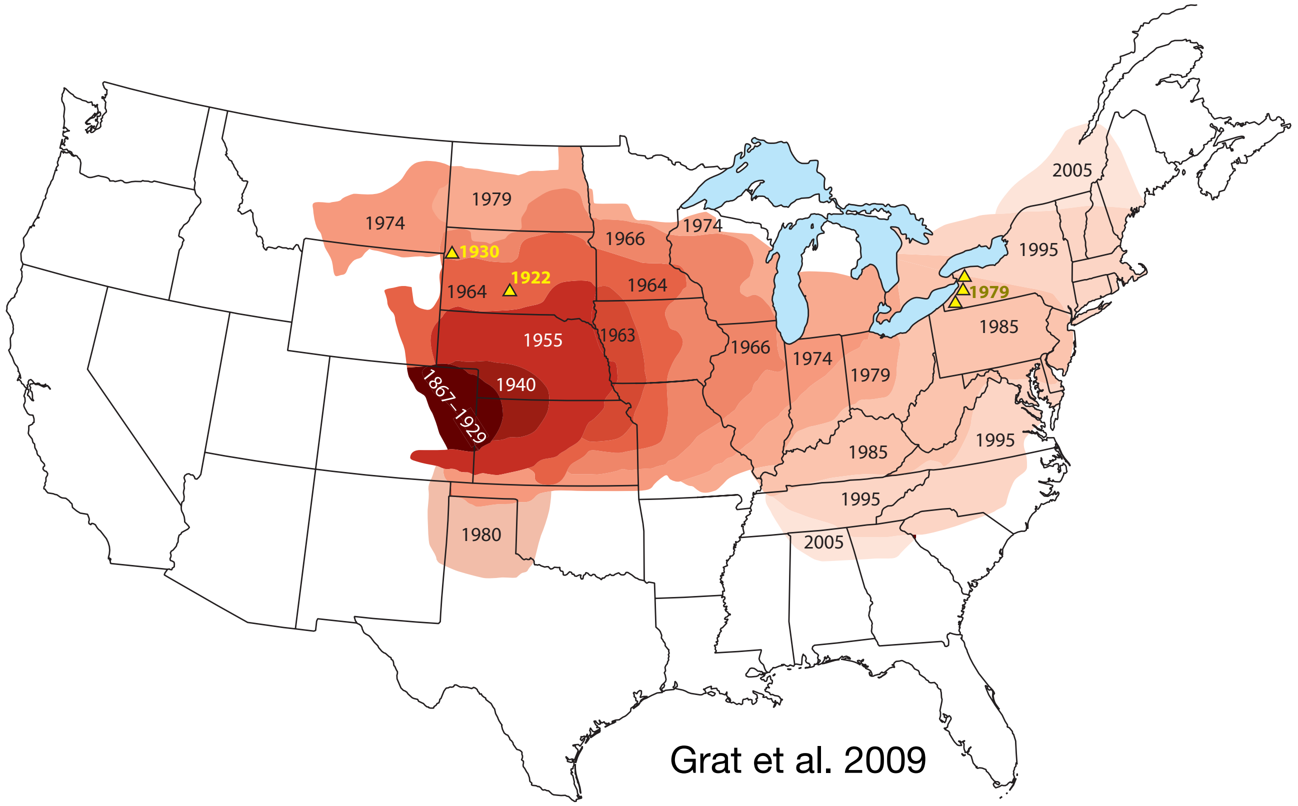
Plant domestication provides new niches as this process involves tradeoffs among human-desired characteristics and plant defenses[1,2] - aka the domestication syndrome[3,4]

Breeding may facilitate **insect adaptation**

The western corn rootworm - *Diabrotica virgifera virgifera*

Original range expanded with maize domestication in **Mexico**

Later with the **intense** maize cropping in the **US**



Grat et al. 2009

---

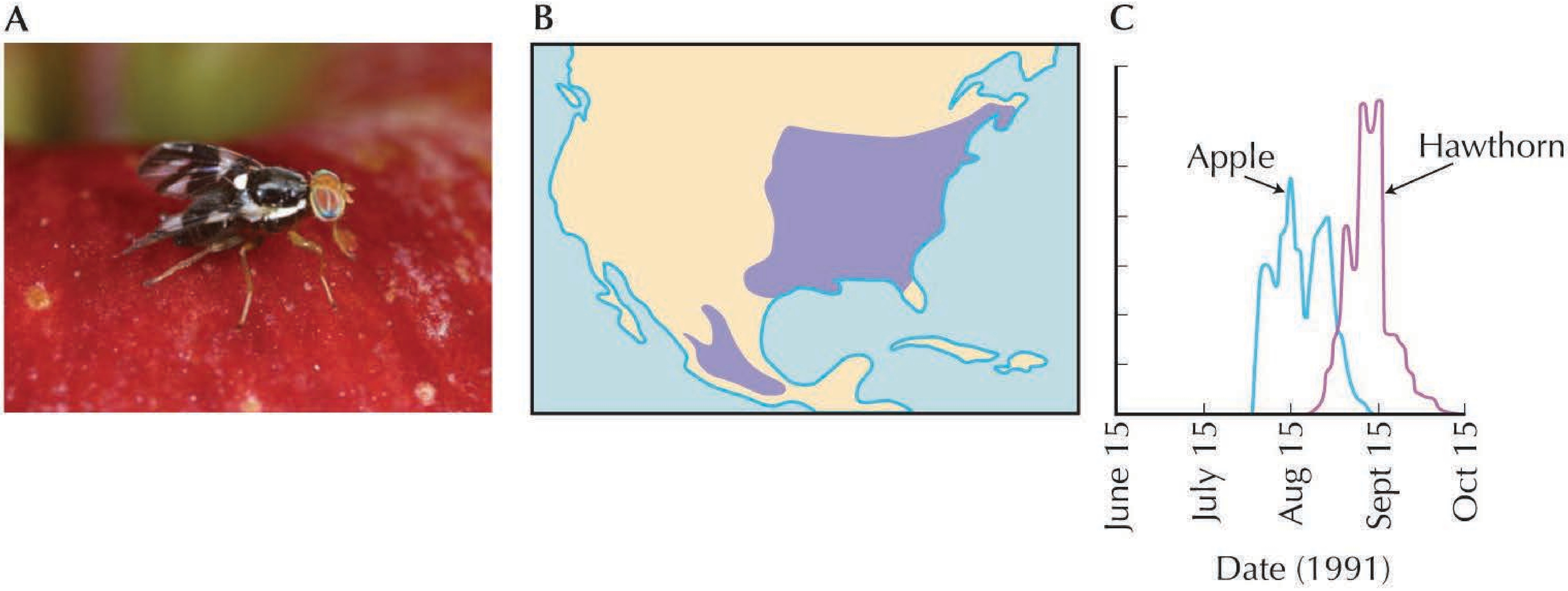
[1]Chen et al. 2015; [2]Mitchell et al. 2016; [3]Hammer 1984; [4]Gross and Olsen 2010

The introduction of domesticated plants to novel areas provides new niches, leading to **adaptation**, **host shifting** and subsequent **host-associated differentiation (HAD)** by native insects.

The apple maggot - *Rhagoletis pomonella*

Shifted from its original host-plant: hawthorn (*Crataegus* spp. *mollis*)

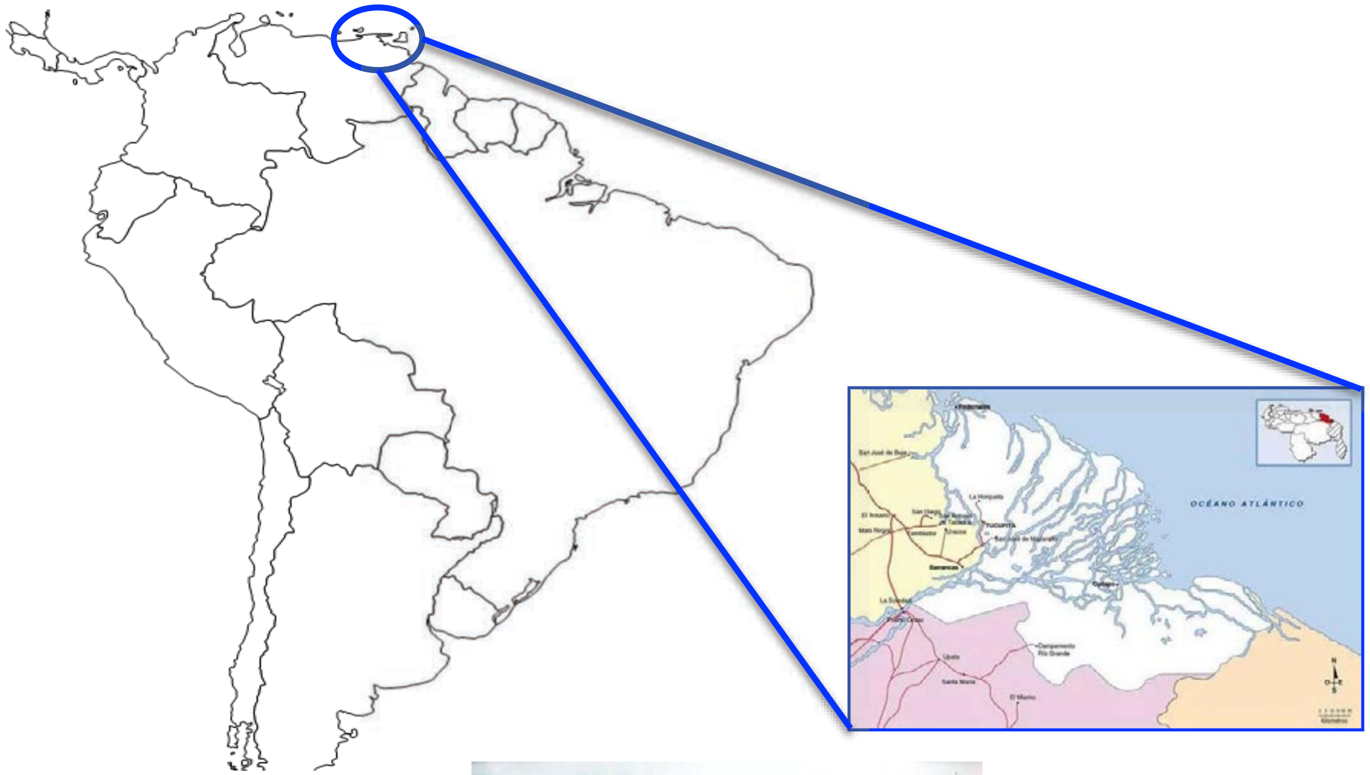
160 years ago - introduction of apple (*Malus domestica*)



Evolution © 2007 Cold Spring Harbor Laboratory Press

Sugarcane borer *Diatraea saccharalis*

Economically damages on:



North - South America

Center of Origin: South/Central America

Delta of Orinoco river
Venezuela flat lands,
Lakes of great Antilles

Delta del Orinoco

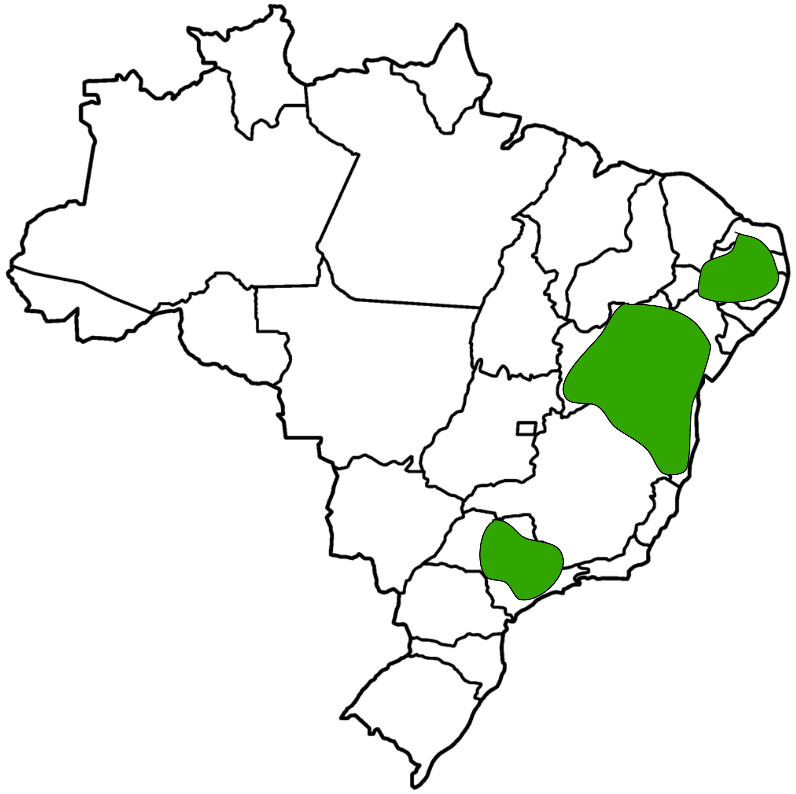
Original Host-plant:
Aquatic and semi-aquatic

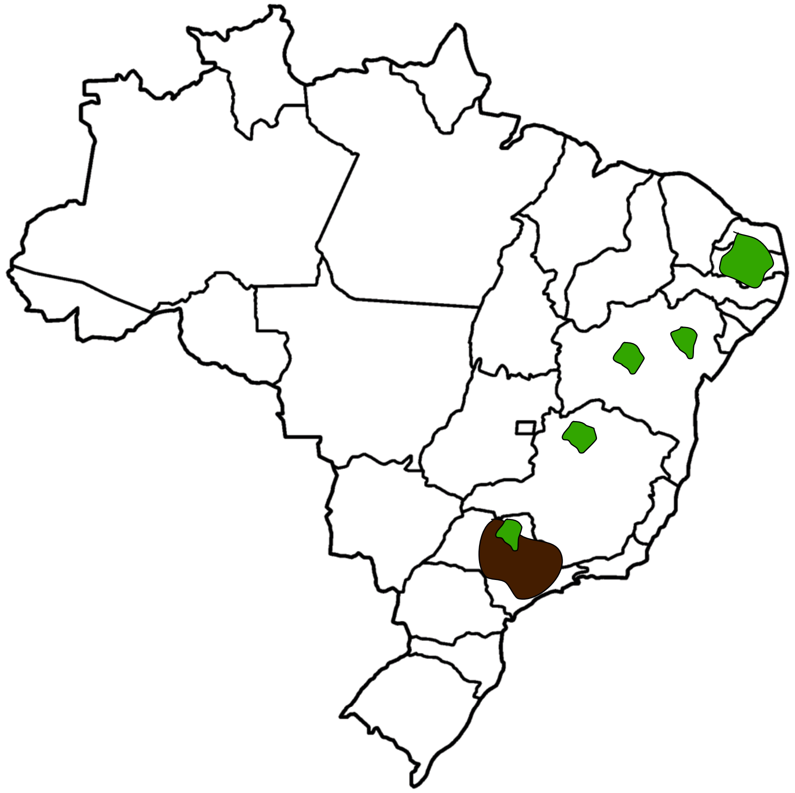
Widespread South America
following European colonization

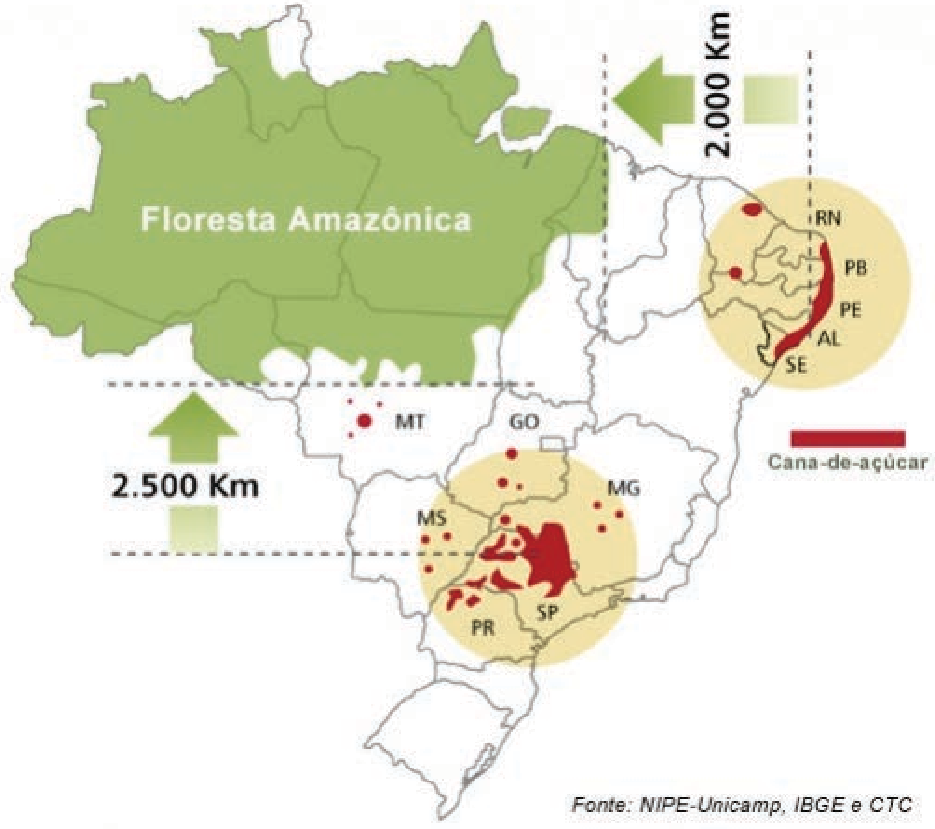
Domestication / spread of maize

Introduction of sugarcane

Sugarcane in Brazil







Fonte: NIPE-Unicamp, IBGE e CTC

Until 1700

1800

1930

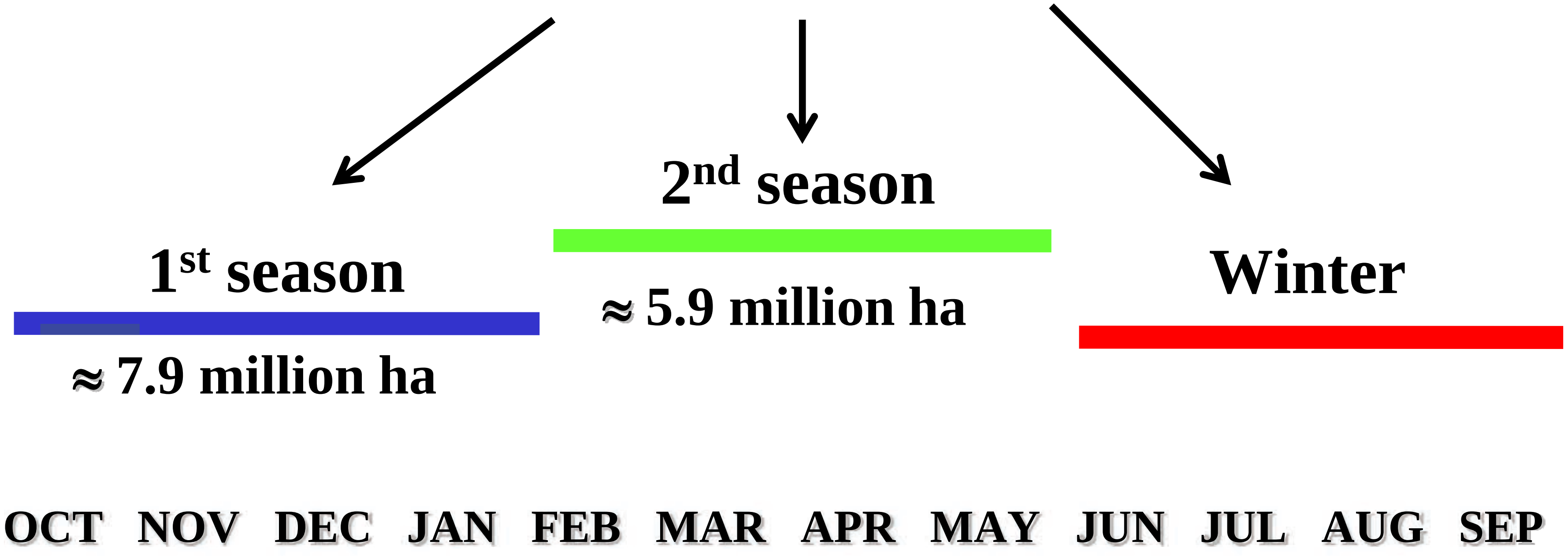
Todays distribution

Maize in Brazil - before 1990


OCT NOV DEC JAN FEV MAR APR MAY JUN JUL AUG SEP

Maize in Brazil - after 1990

Corn-Growing Seasons in Brazil


2nd season

≈ 5.9 million ha

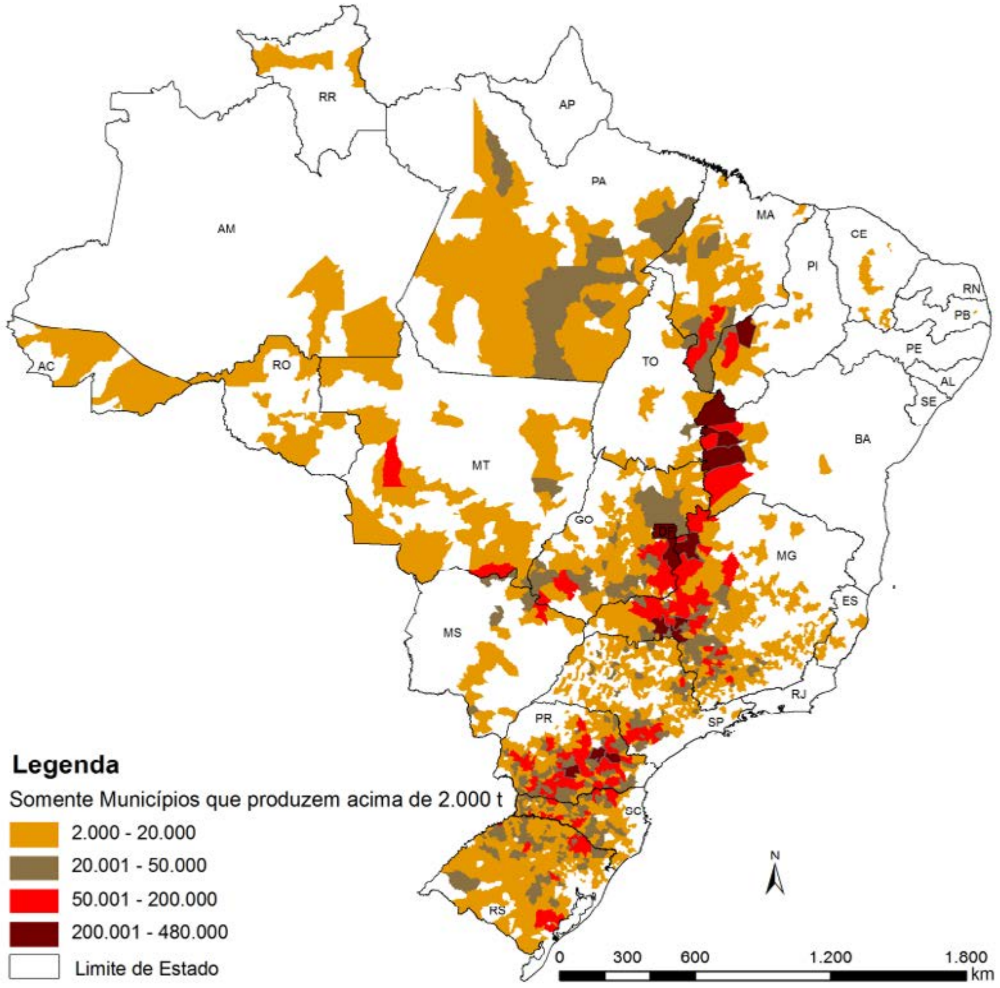
1st season

≈ 7.9 million ha

Winter

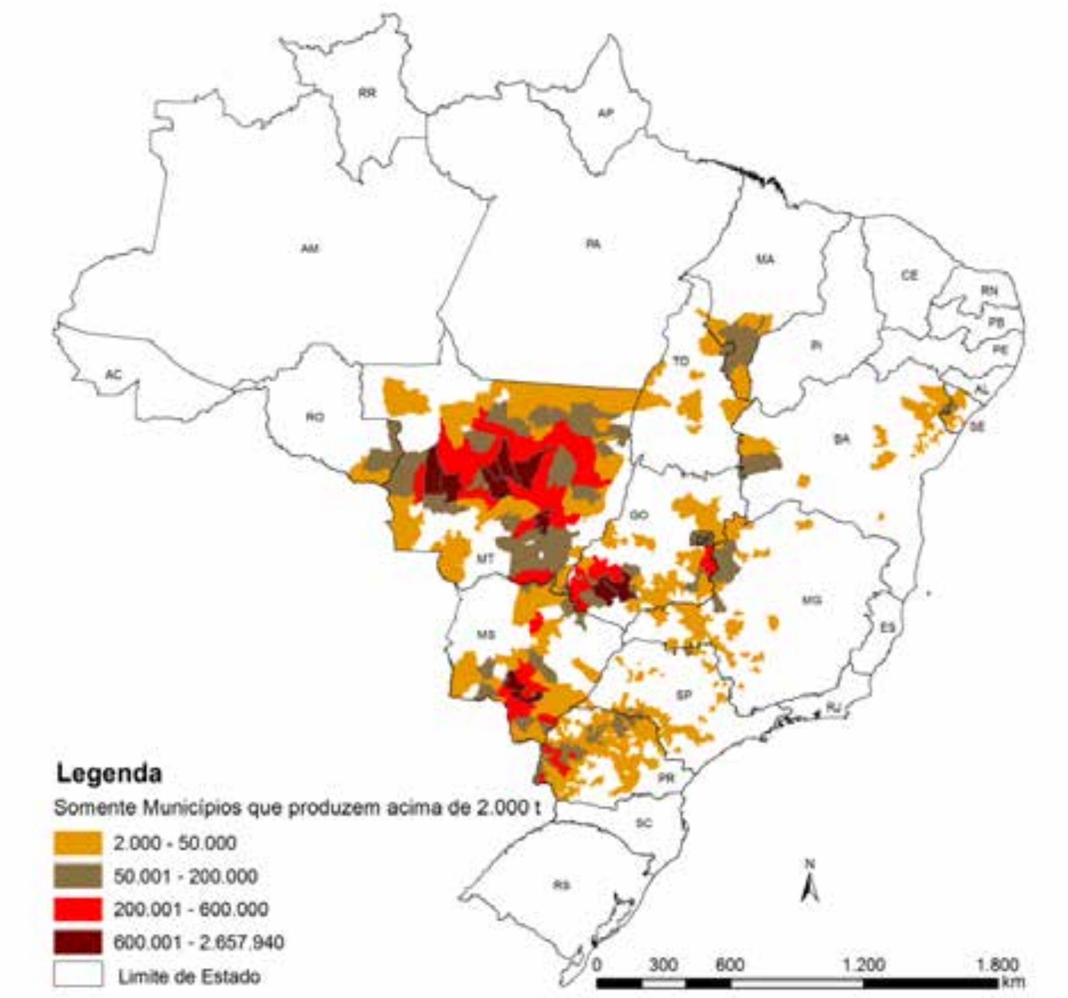
OCT NOV DEC JAN FEB MAR APR MAY JUN JUL AUG SEP

Maize in Brazil - after 1990



Fonte: Conab/IBGE.

1st season

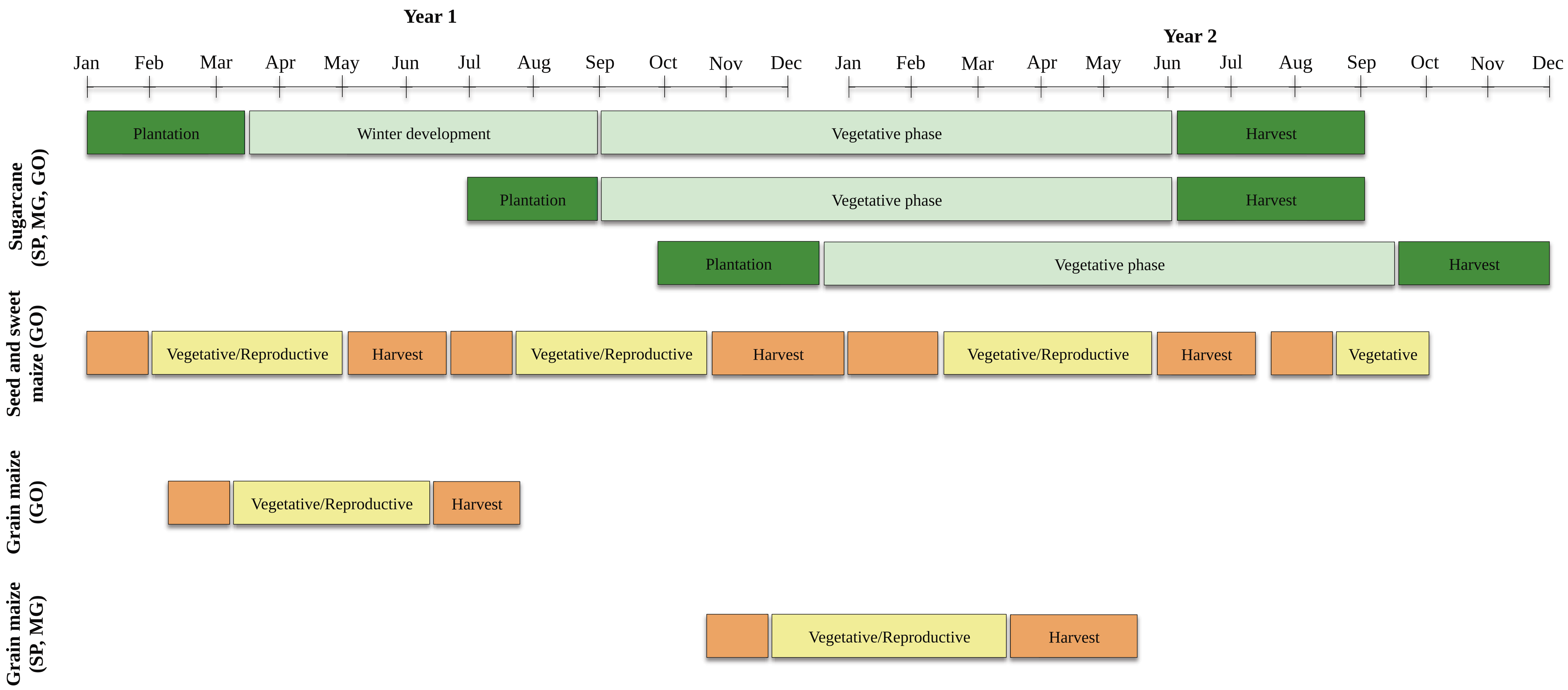


Fonte: Conab/IBGE.

2nd season

Agricultural season

Sugarcane and maize season



Year 1

Year 2

Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec

Sugarcane (SP, MG, GO)

Plantation | Winter development | Vegetative phase | Harvest

Plantation | Vegetative phase | Harvest

Plantation | Vegetative phase | Harvest

Seed and sweet maize (GO)

Vegetative/Reproductive | Harvest | Vegetative/Reproductive | Harvest | Vegetative/Reproductive | Harvest | Vegetative

Grain maize (GO)

Vegetative/Reproductive | Harvest

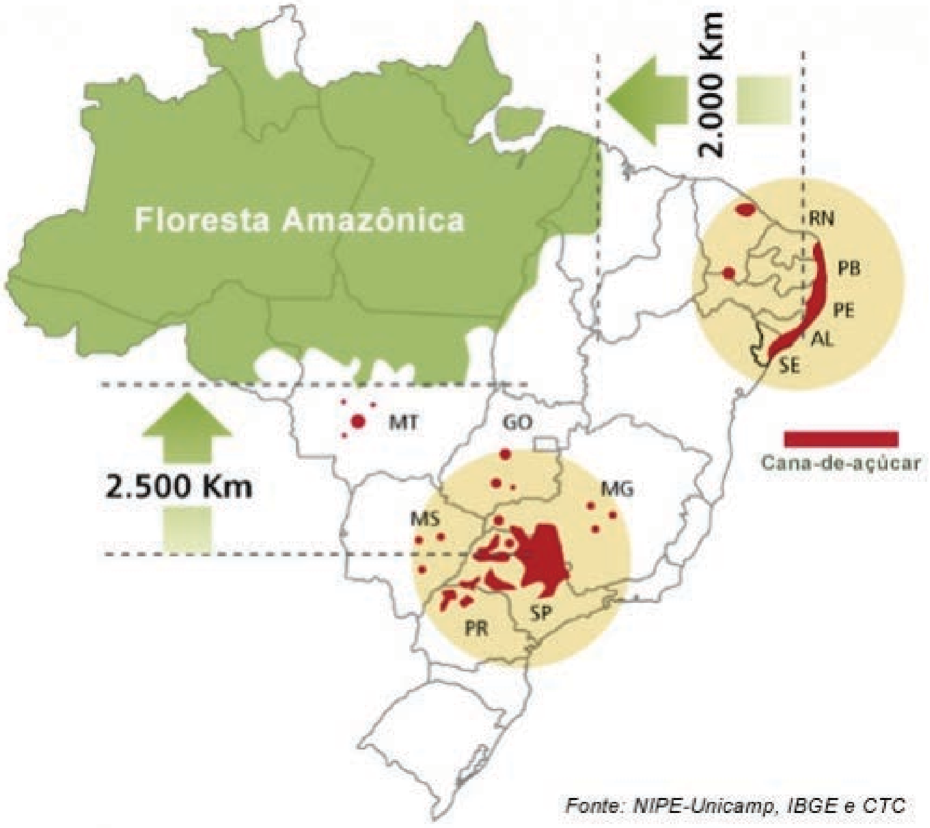
Grain maize (SP, MG)

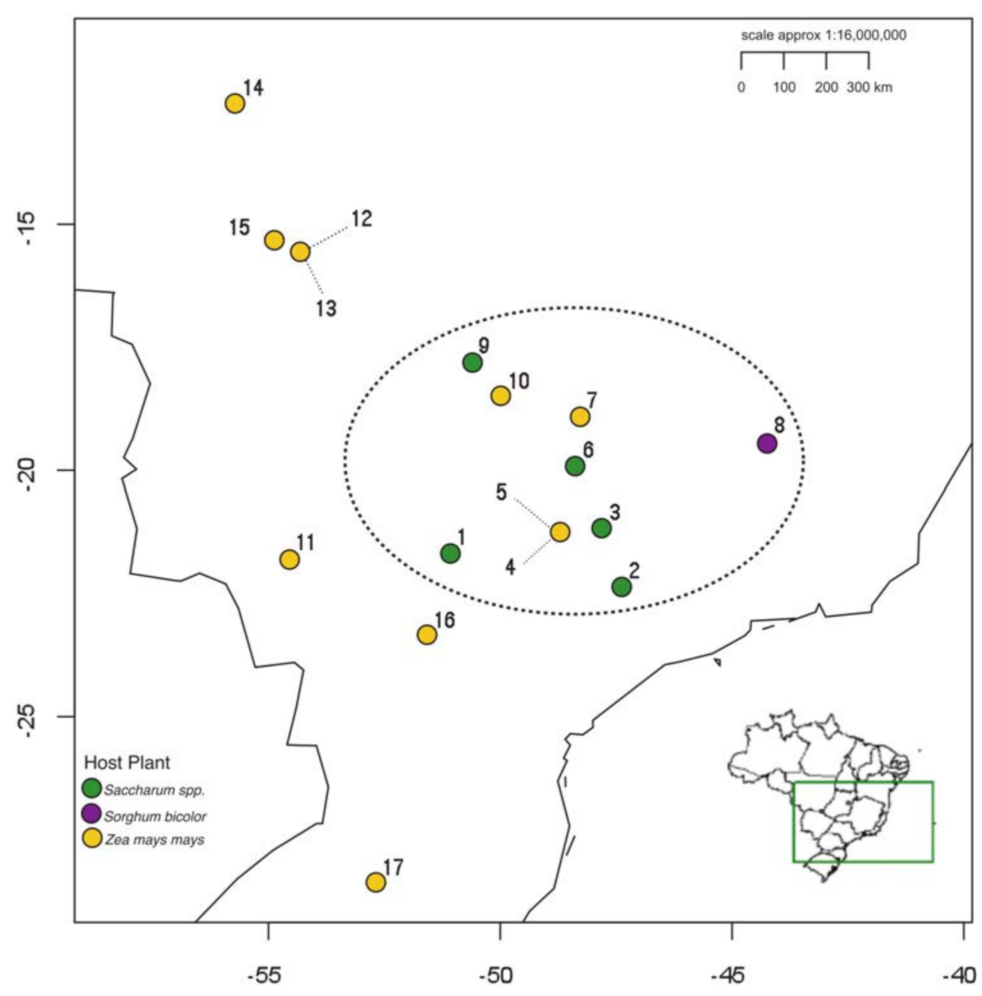
Vegetative/Reproductive | Harvest

Changes in Brazilian agriculture likely impacted insect pest evolutionary genetics; however, this is not completely understood.

We posit that the evolution of sugarcane borer is associated with the changes in the Brazilian landscape from **historical and modern agricultural activity**

We hypothesized that the modification of large areas of sugarcane during Portuguese colonization facilitated a **host shift to sugarcane**, whereas more recent agricultural **expansion of maize and sugarcane production** have provided sufficient evolutionary pressure leading to **population structure and HAD**

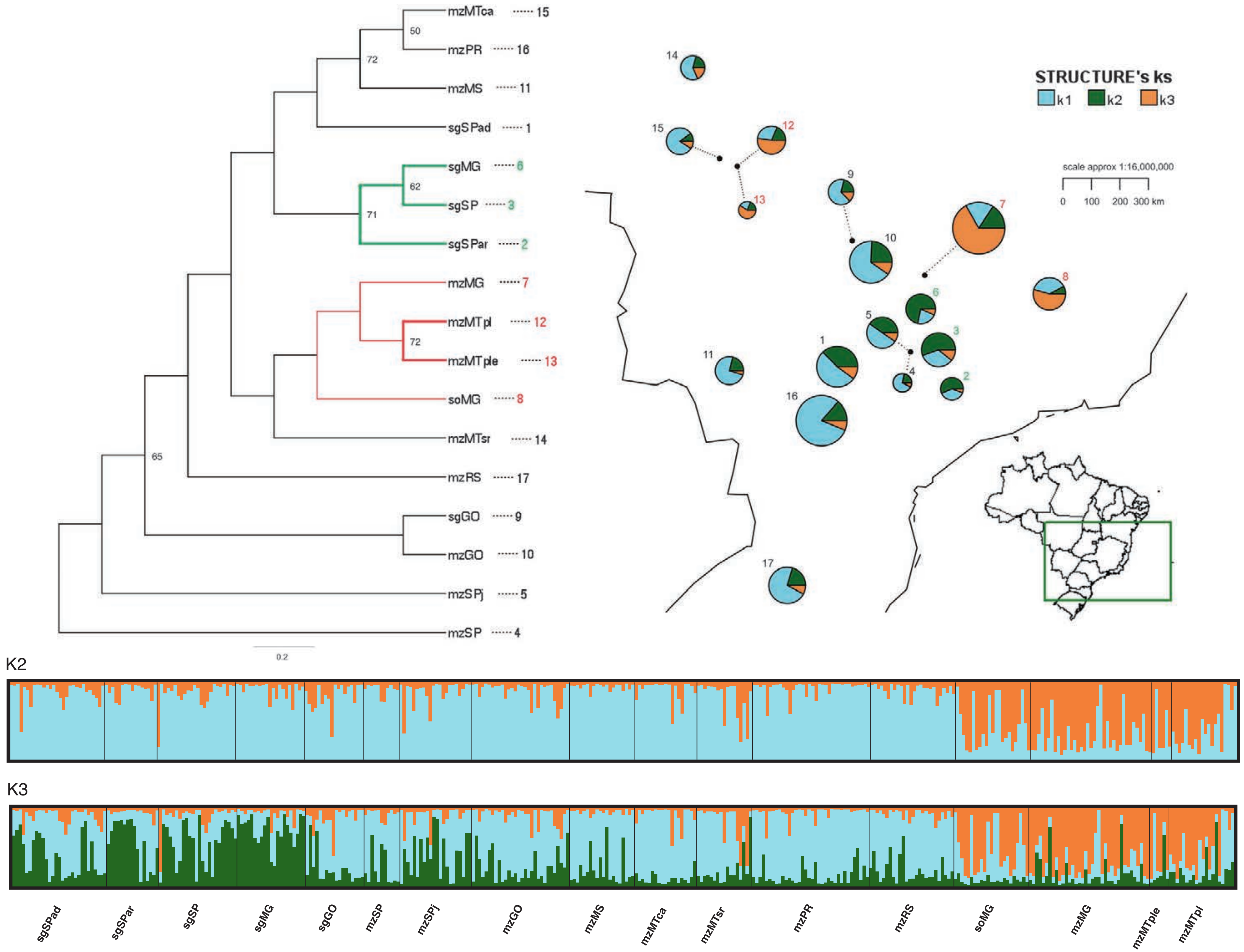




10 Microsatellite markers[1]

[1] Pavinato et al. 2013

Genetic structure

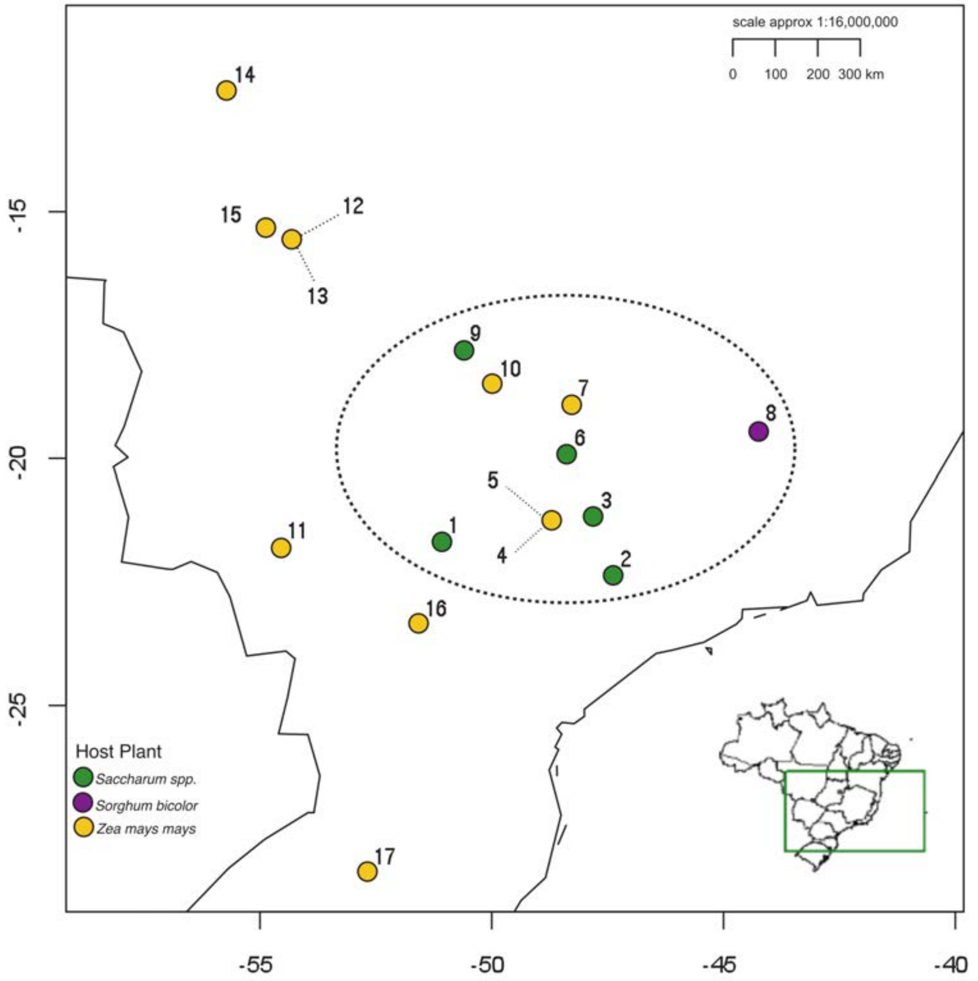


K=2: Spatial differentiation:

1) demography and colonization events;
2) agricultural system - shift towards de center

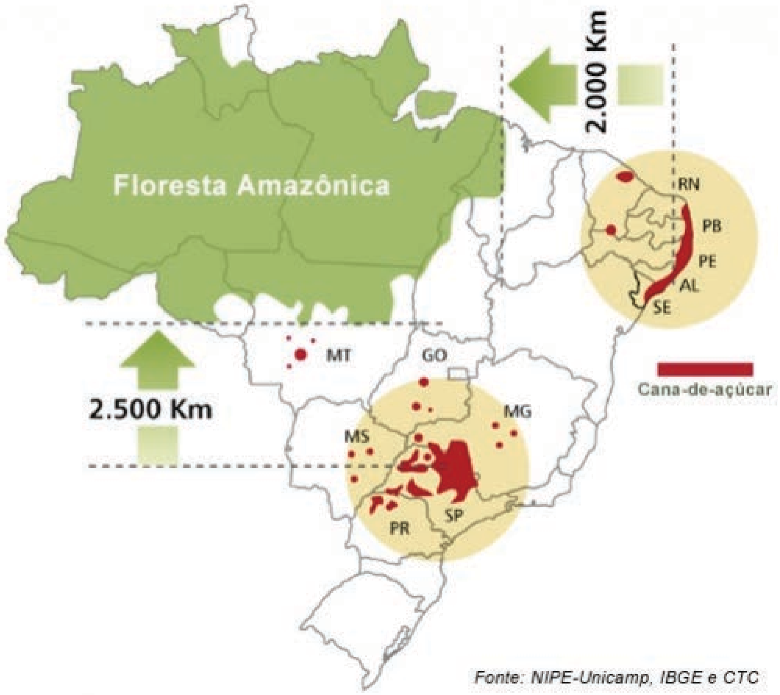
K=3: Incipient host differentiation:

Pavinato et al. 2018



Partial Mantel test
Causal modeling[1]

Model	r^2	p-value
pFST ~ host | geo dist	0.279	0.026
pFST ~ geo dist | host	-0.004	0.475



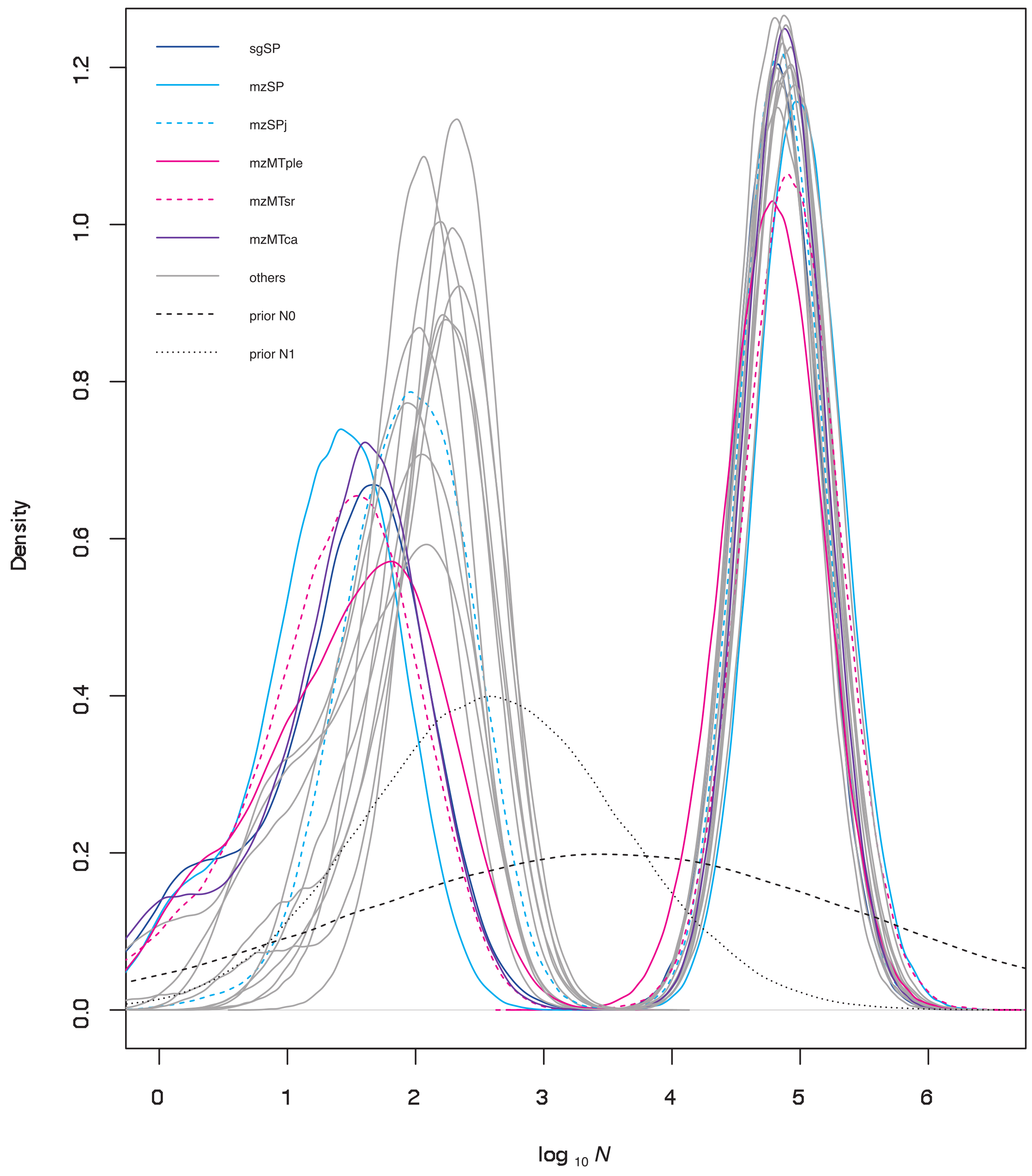
Fonte: NIPE-Unicamp, IBGE e CTC

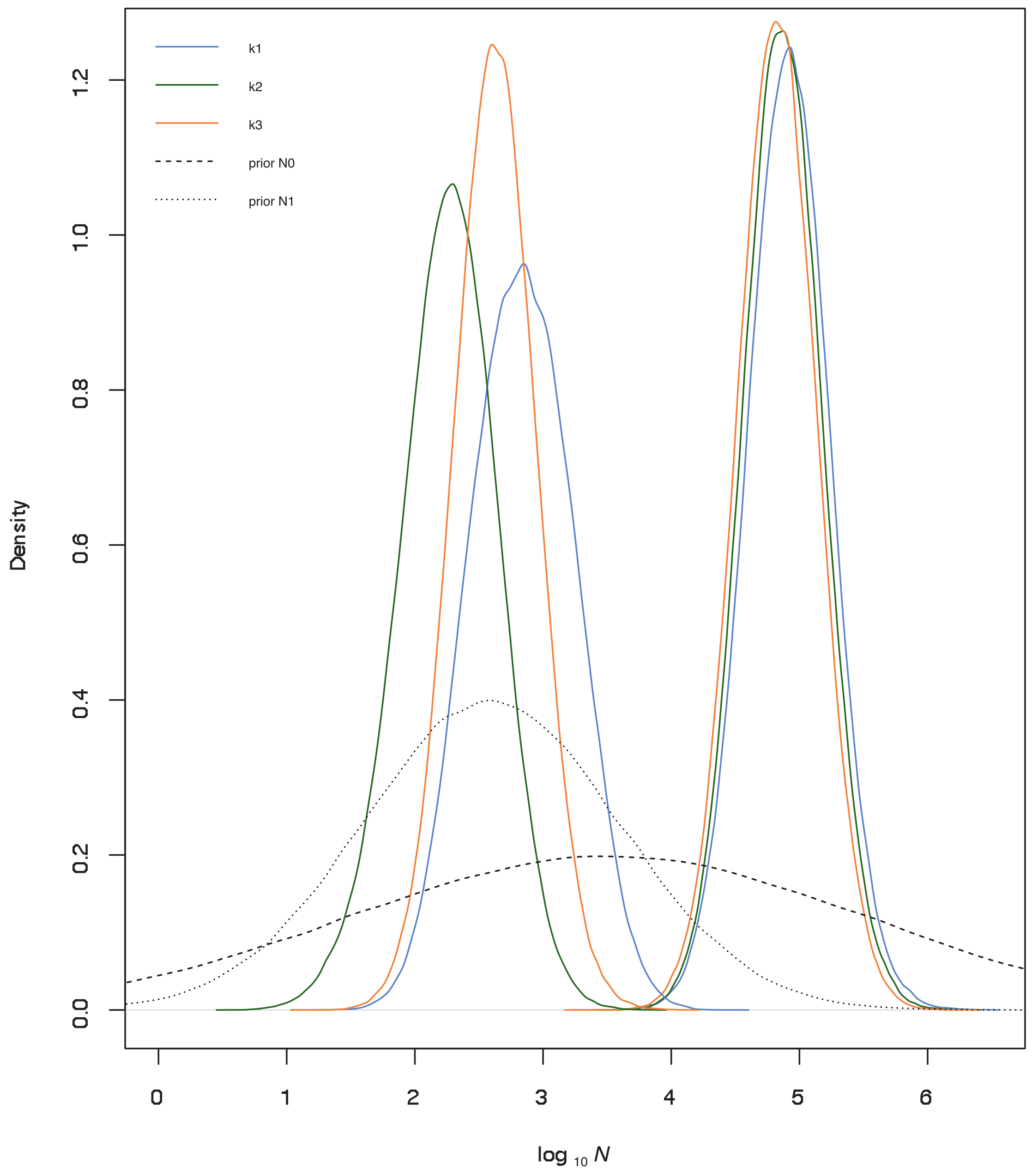
Incipient HRA - Ecological divergence

1) Ancestral signal of association;
2) Recent HRA

[1] Cushman et al. 2006; Castillo et al. 2014

Population size changes





N0 ~ 100
Na ~ 100.000

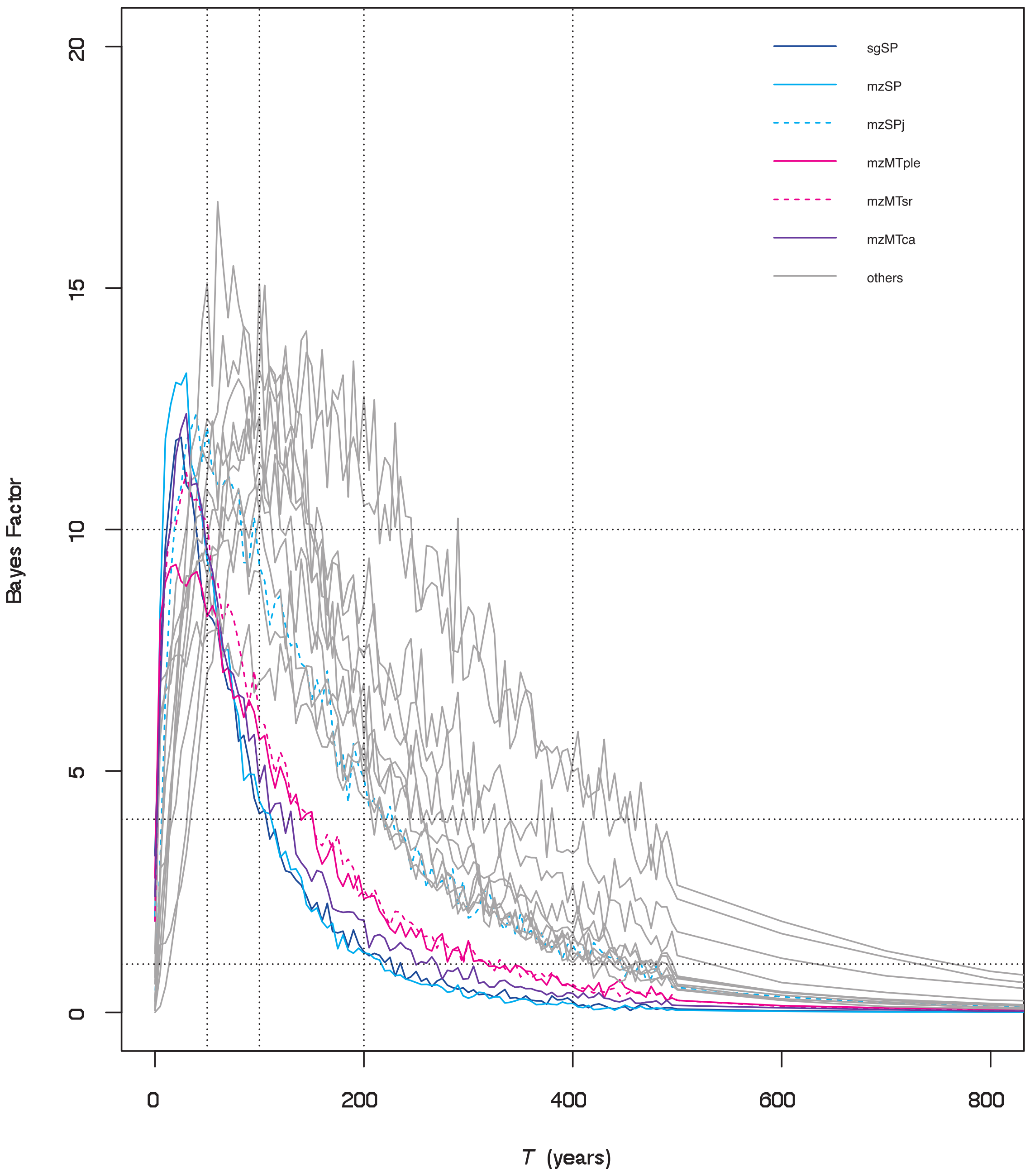
Population size changes

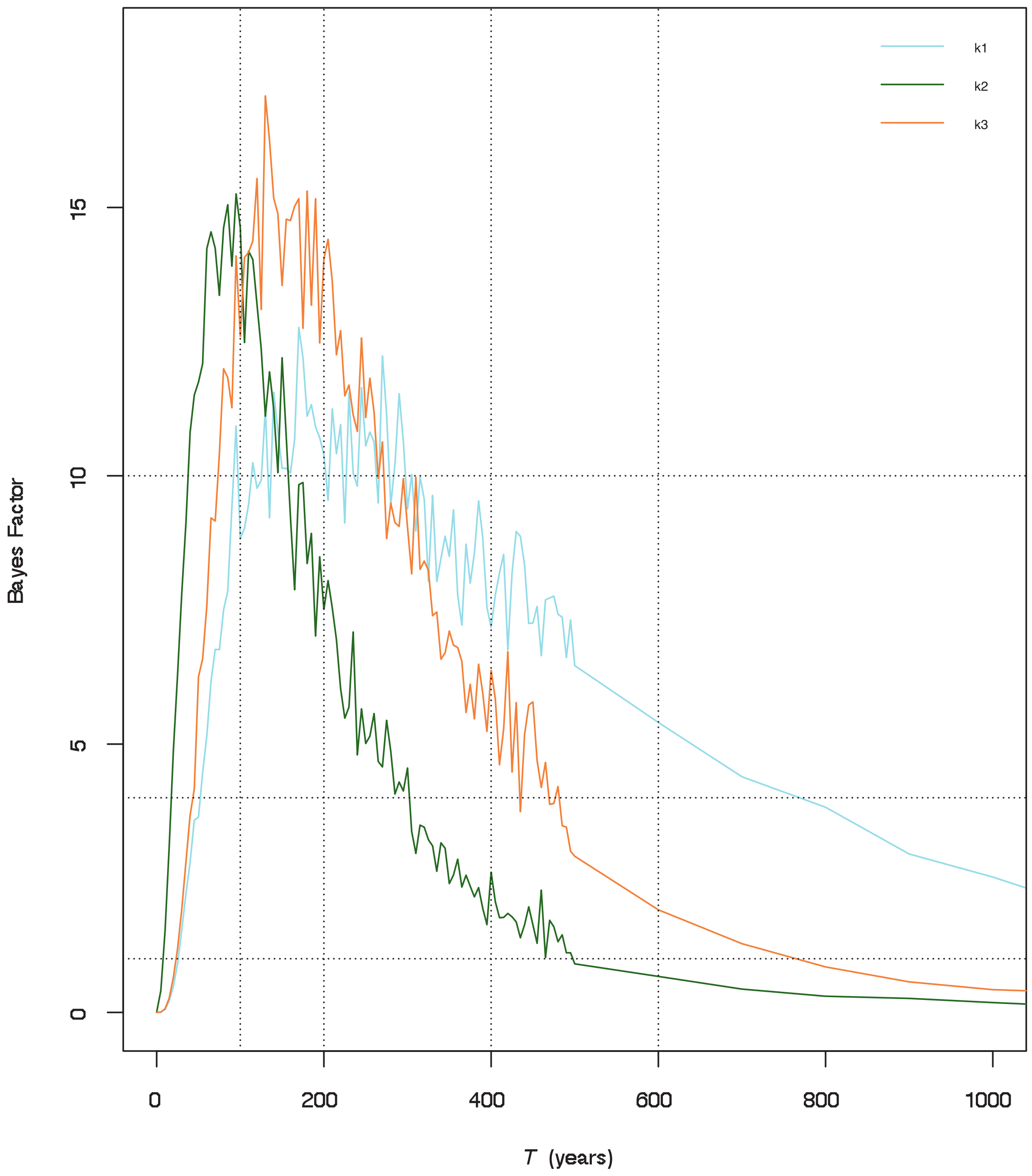
Test the time for N_e changes:

(1) changes imposed by modern agricultural practices in the last 60 years;

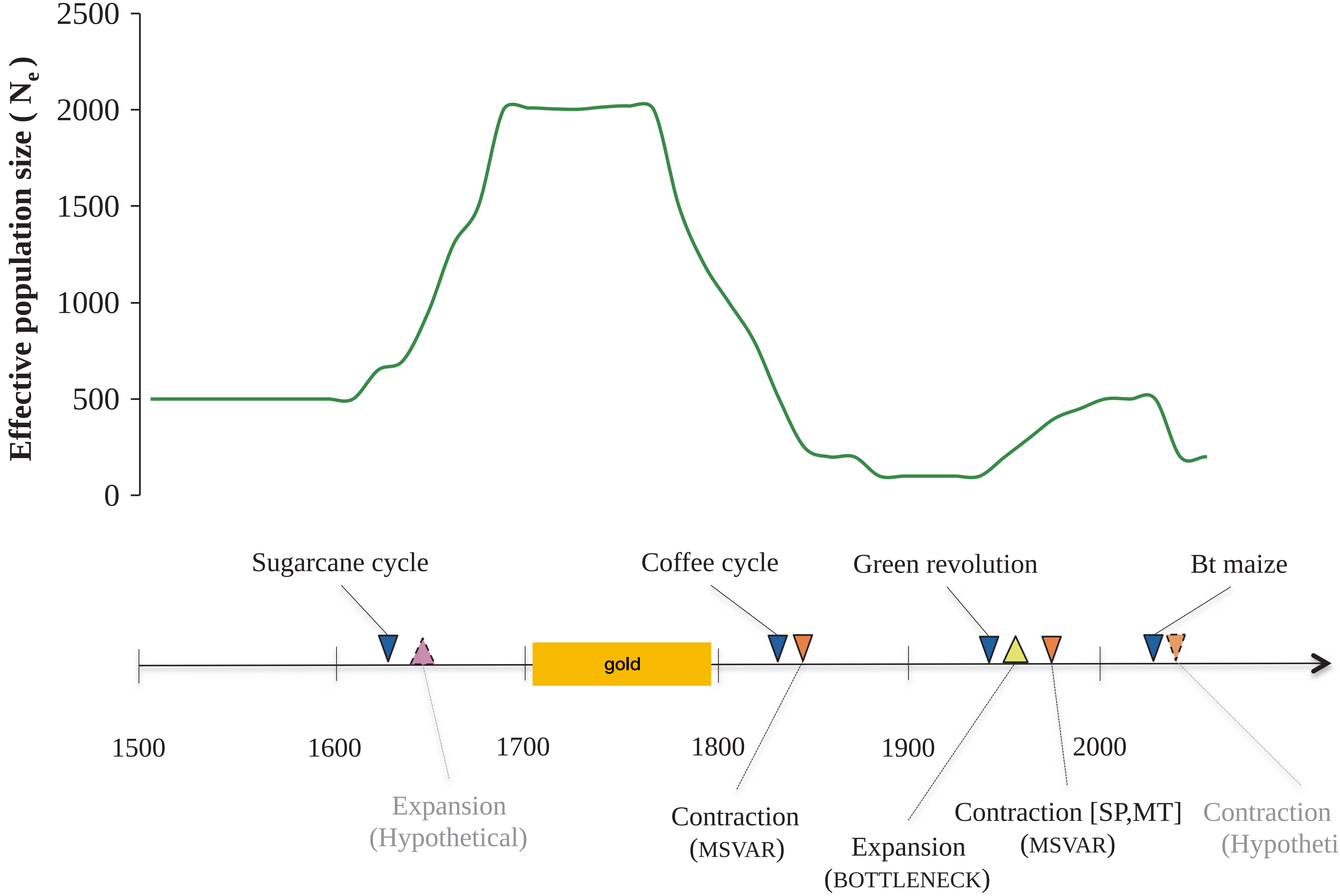
(2) changes ≈60–500 years ago;

(3) ancestral population





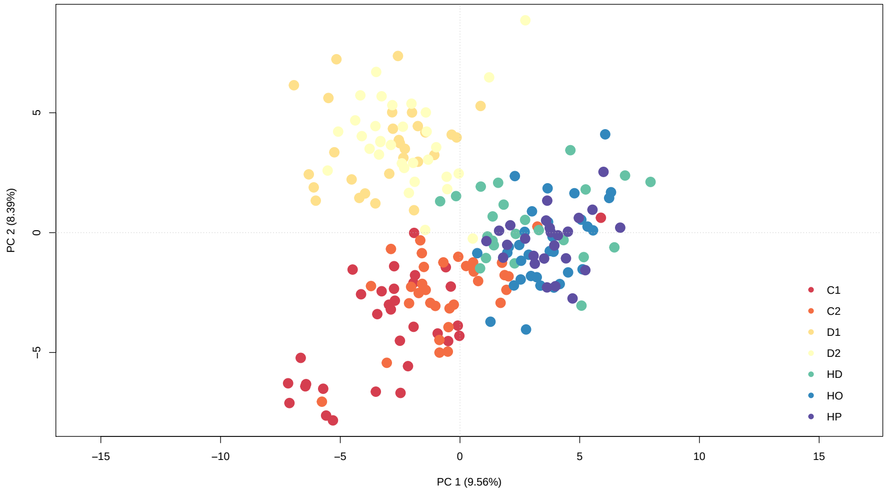
Genetic of SCB and agricultural system



Historical **introduction of sugarcane** production along with the more **recent expansion of maize and sugarcane fields**, impacted the genetic diversity and evolutionary dynamics of the sugarcane borer in Brazil:

demographic events (bottleneck followed by an expansion) coincided with known agricultural events in Brazil

We also observed signals of **incipient divergence** among hosts (sugarcane and maize).



Winter survival mechanisms and adaptive genetic variation in an Antarctic insect

Drew Spatch - The Ohio State University, OBEE
David Denlinger - The Ohio State University, OBEE
Andrew P. Michel - The Ohio State University, Entomology

Antarctic midge, *Belgica antarctica*

Antarctic midge, *B. antarctica* is an endemic insect species of the Antarctica. This is a **wingless and terrestrial** species of Chironomidae family (Diptera). This species developed several physiological adaptations to **cold and dry** conditions found at Antarctica environment.

Model to study adaptation to extreme conditions and the response to climate changes.

Main Project Goal:

This project focuses on **deciphering the physiological and molecular mechanisms** that enable the Antarctic midge *Belgica antarctica* to survive environmental stress and the loss of most of its body water in the desiccating polar environment.

1) Evaluating the role of aquaporins (water channel proteins) in the rapid removal of water from the body by studying expression of their genes during dehydration;

2) Investigating the mechanism of metabolic depression and the role of autophagy (controlled breakdown of cellular components) as a mediator of stress tolerance by studying expression of the genes responsible for autophagy during the dehydration process; and

3) **Evaluating the population structure, gene flow, and adaptive variation in physiological traits associated with stress tolerance using a genetic approach that takes advantage of the genomic sequence available for this species coupled with physiological and environmental data from the sampled populations and their habitats.**

Targeted Sequencing

National Science Foundation
WHERE DISCOVERIES BEGIN

Targeted Sequencing

nature COMMUNICATIONS

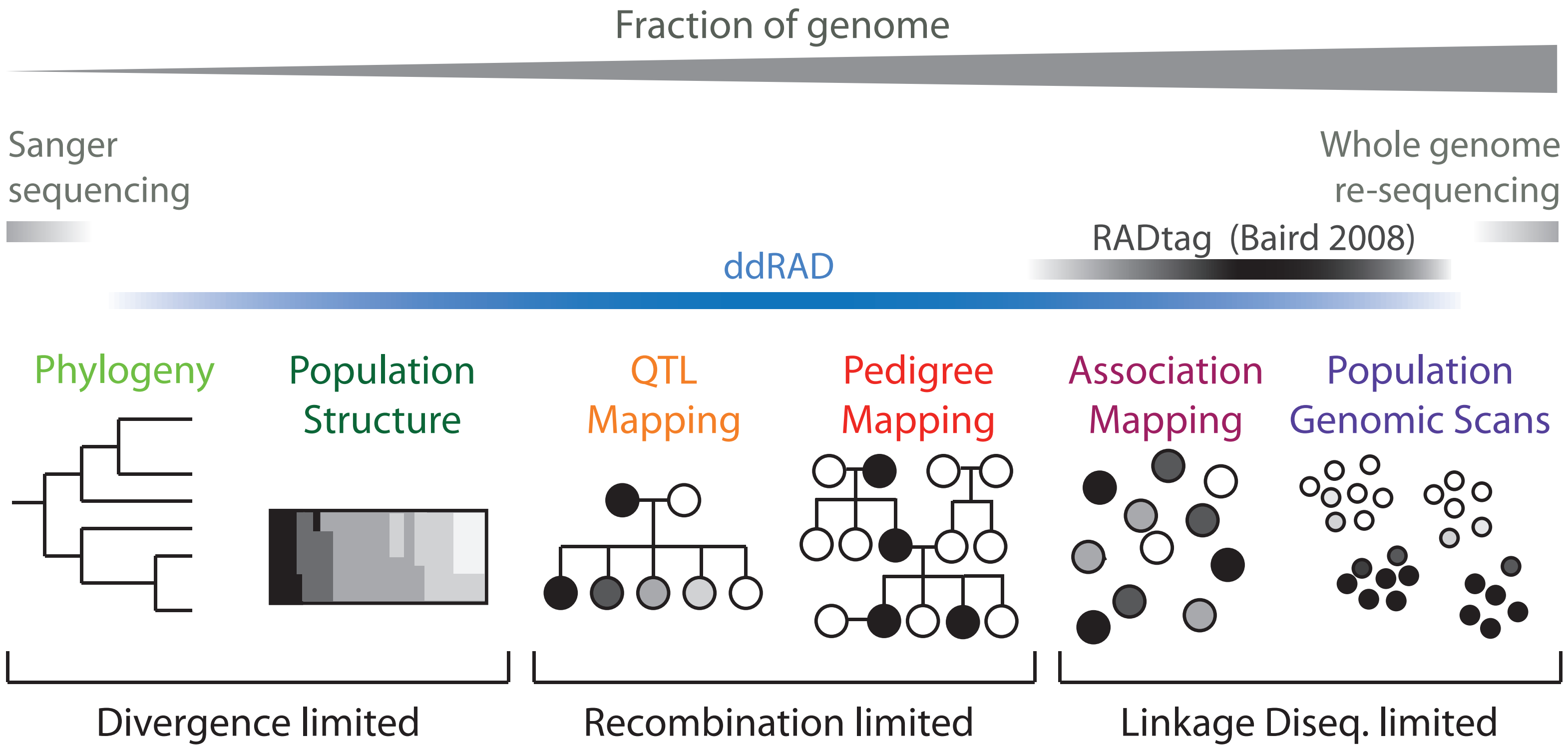
ARTICLE

Received 10 Dec 2013 | Accepted 8 Jul 2014 | Published 12 Aug 2014 DOI: 10.1038/ncomms5611 OPEN

Compact genome of the Antarctic midge is likely an adaptation to an extreme environment

Joanna L. Kelley[1,2], Justin T. Peyton[3,4,*], Anna-Sophie Fiston-Lavier[5,6,*], Nicholas M. Teets[3,7], Muh-Ching Yee[1,8], J. Spencer Johnston[9], Carlos D. Bustamante[1], Richard E. Lee[10] & David L. Denlinger[3,4]

Cost effective genotype several individuals



+

Demographic inference

Targeted enrichment method - PCR based

- Discovery and isolate informative SNPs for the genotype of Antarctic midges populations;

- Develop a pipeline to automate the processing of the sequencing data;

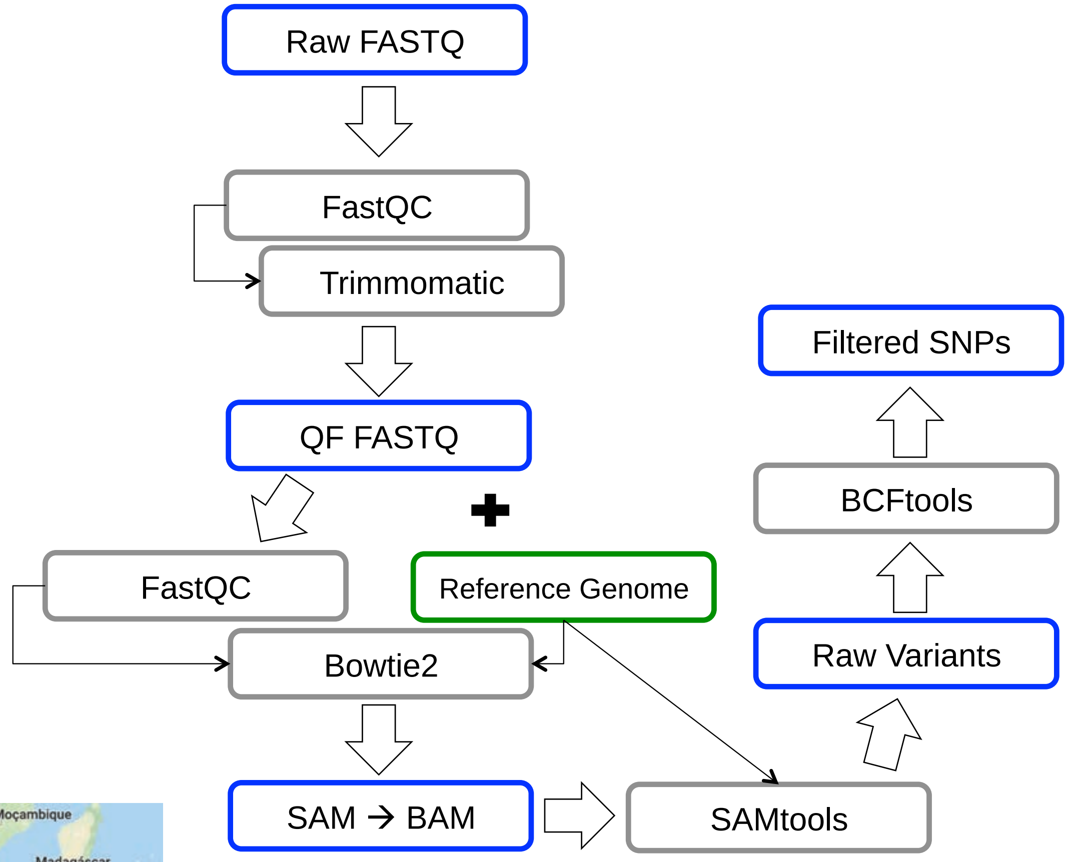
Whole Genome Sequencing Data

Discovered SNPs

Targeted SNPs

Fluidigm panel (double PCR protocol)

WGS 24 individuals



Raw FASTQ

FastQC

Trimmomatic

QF FASTQ

✚

FastQC

Reference Genome

Bowtie2

SAM → BAM

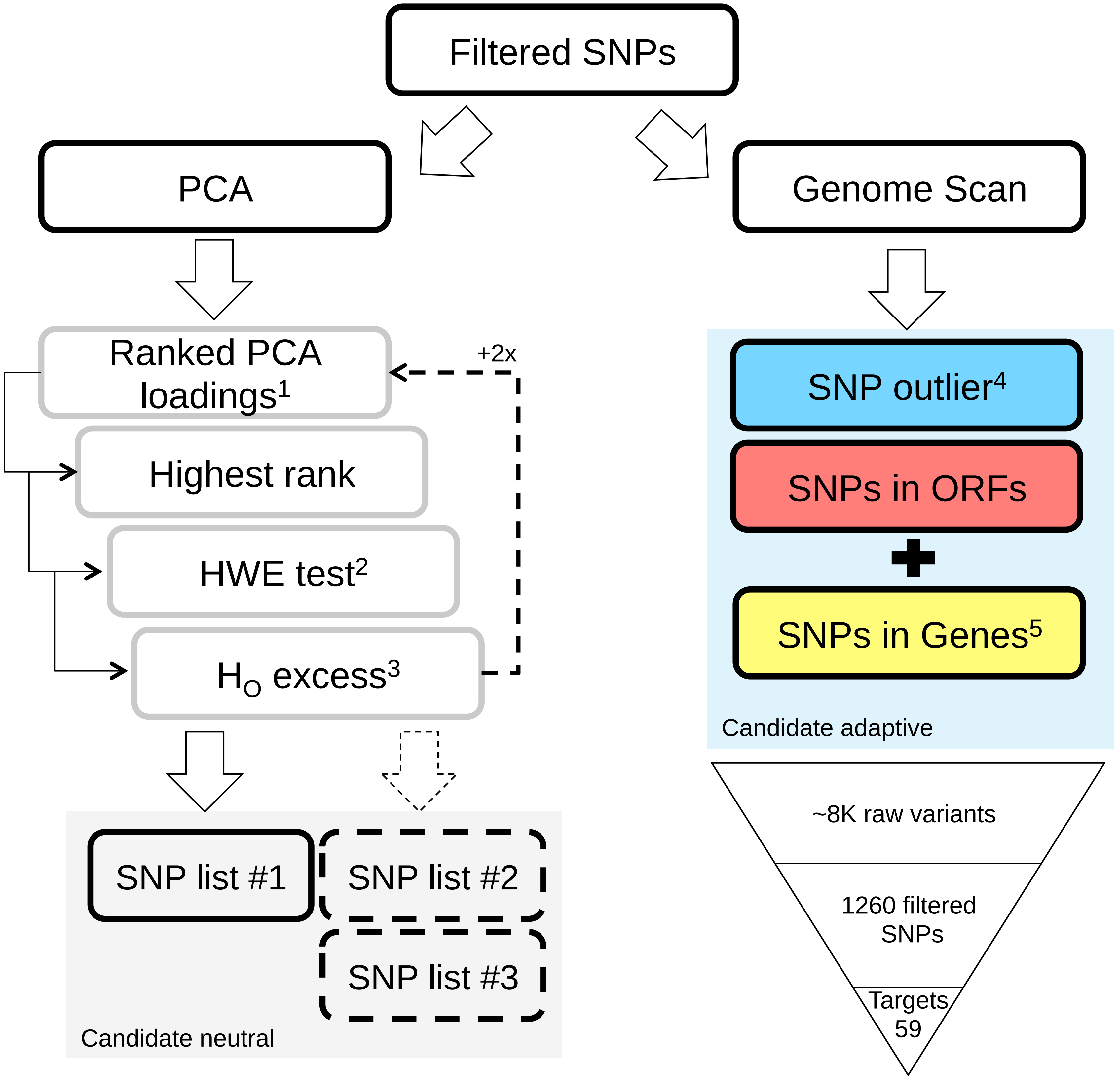
SAMtools

Raw Variants

BCFtools

Filtered SNPs

Framework to isolate informative SNPs



Filtered SNPs

PCA

Genome Scan

Ranked PCA loadings[1]

Highest rank

HWE test[2]

H_O excess[3]

+2x

SNP list #1

SNP list #2

SNP list #3

Candidate neutral

SNP outlier[4]

SNPs in ORFs

✚

SNPs in Genes[5]

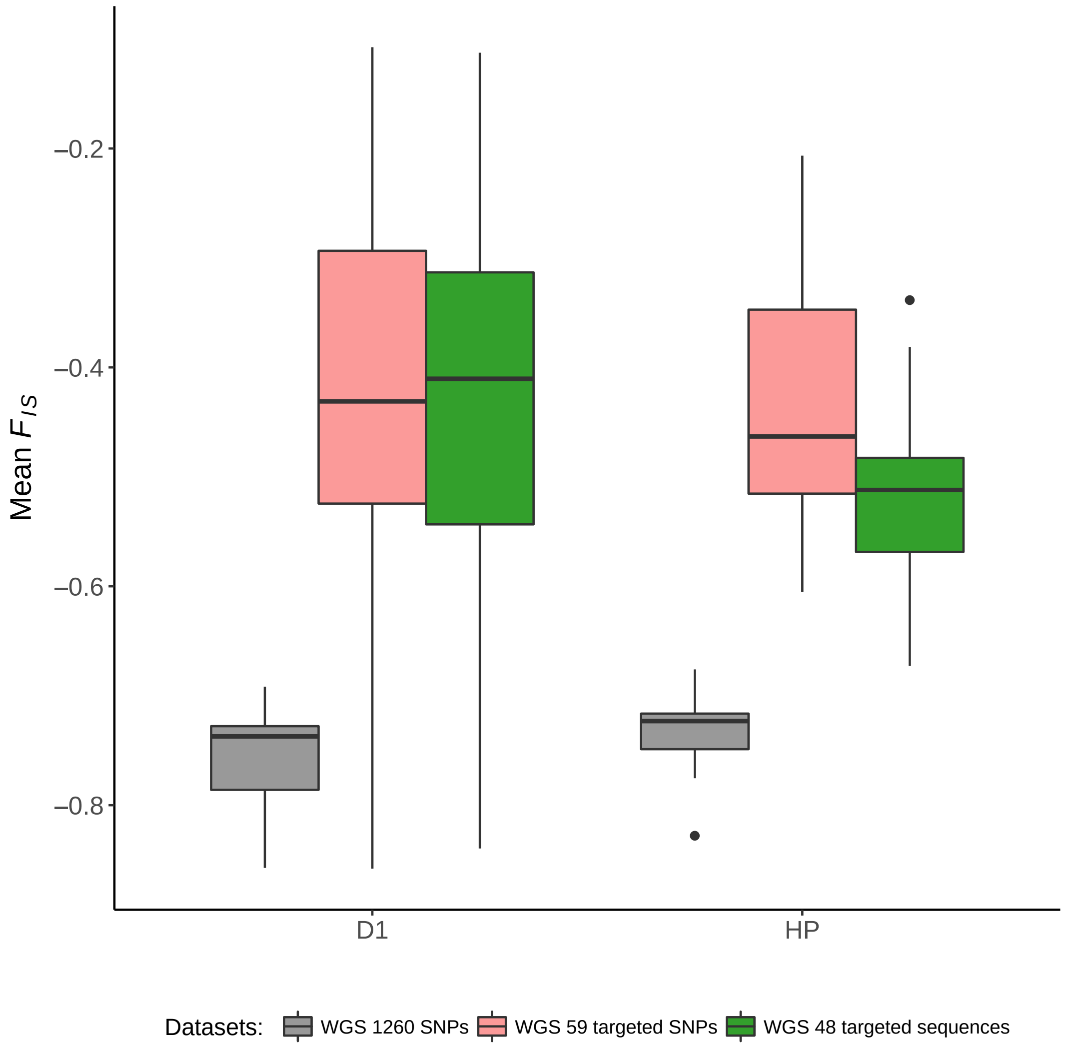
Candidate adaptive

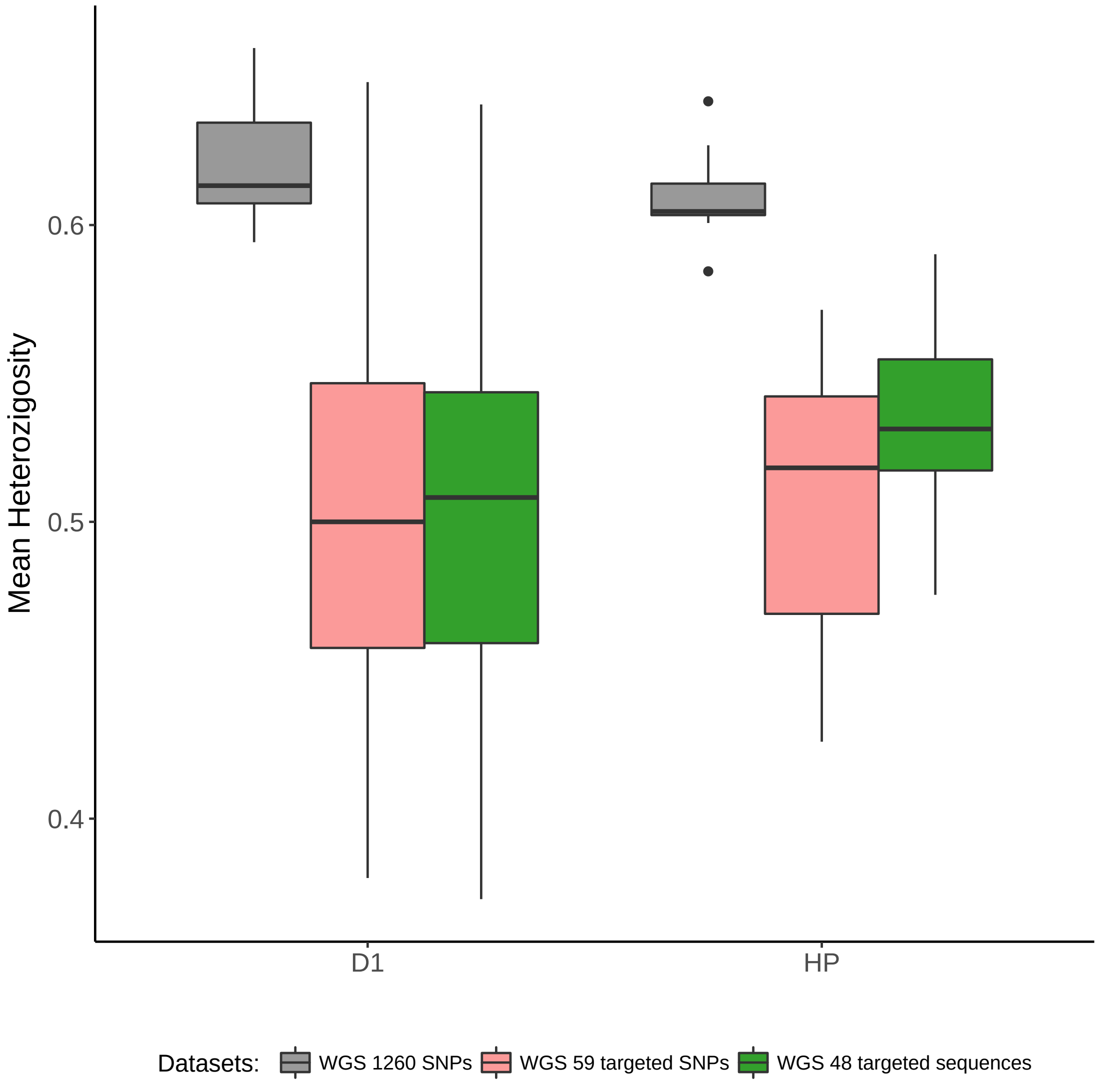
~8K raw variants

1260 filtered SNPs

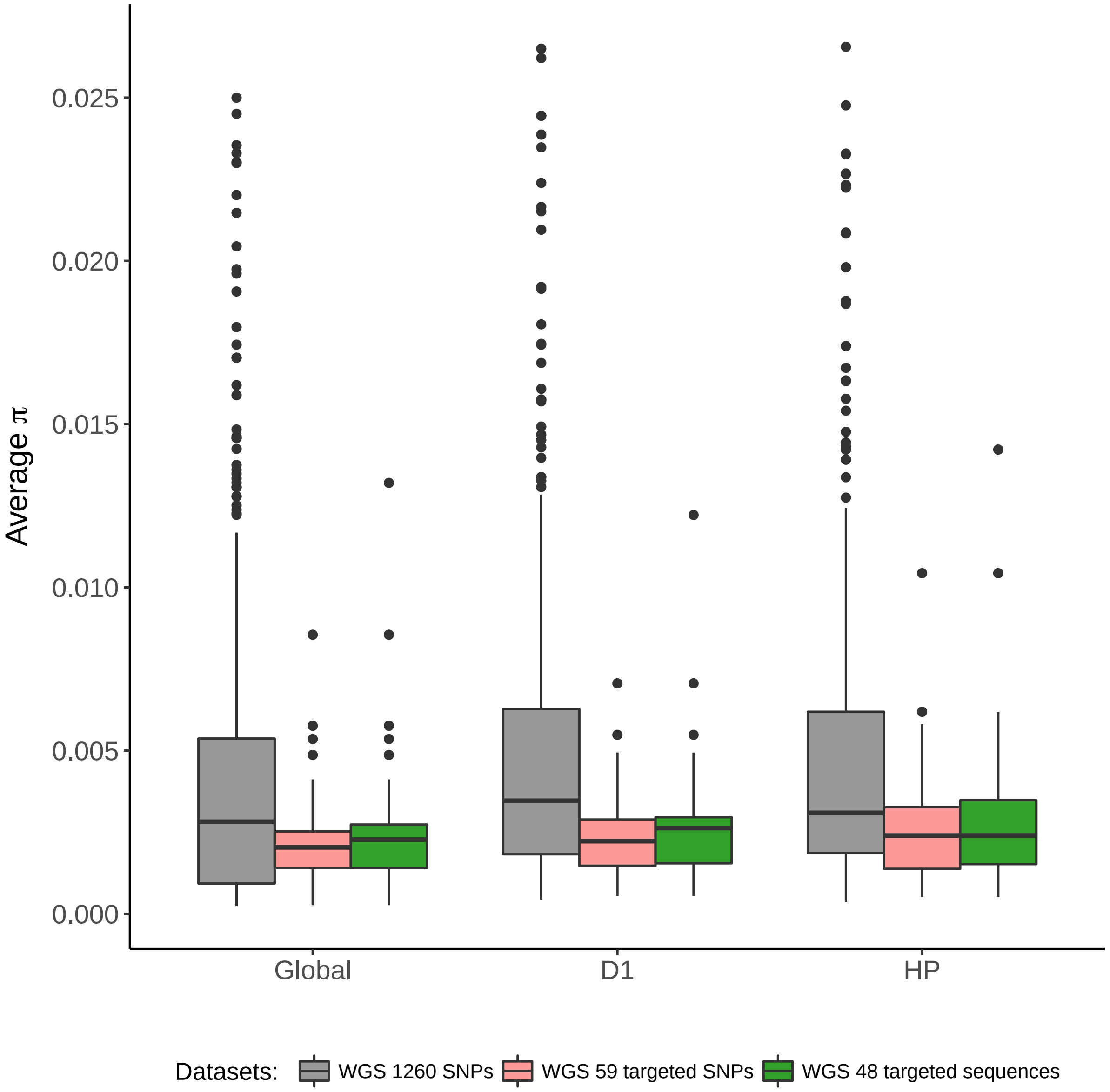
Targets 59

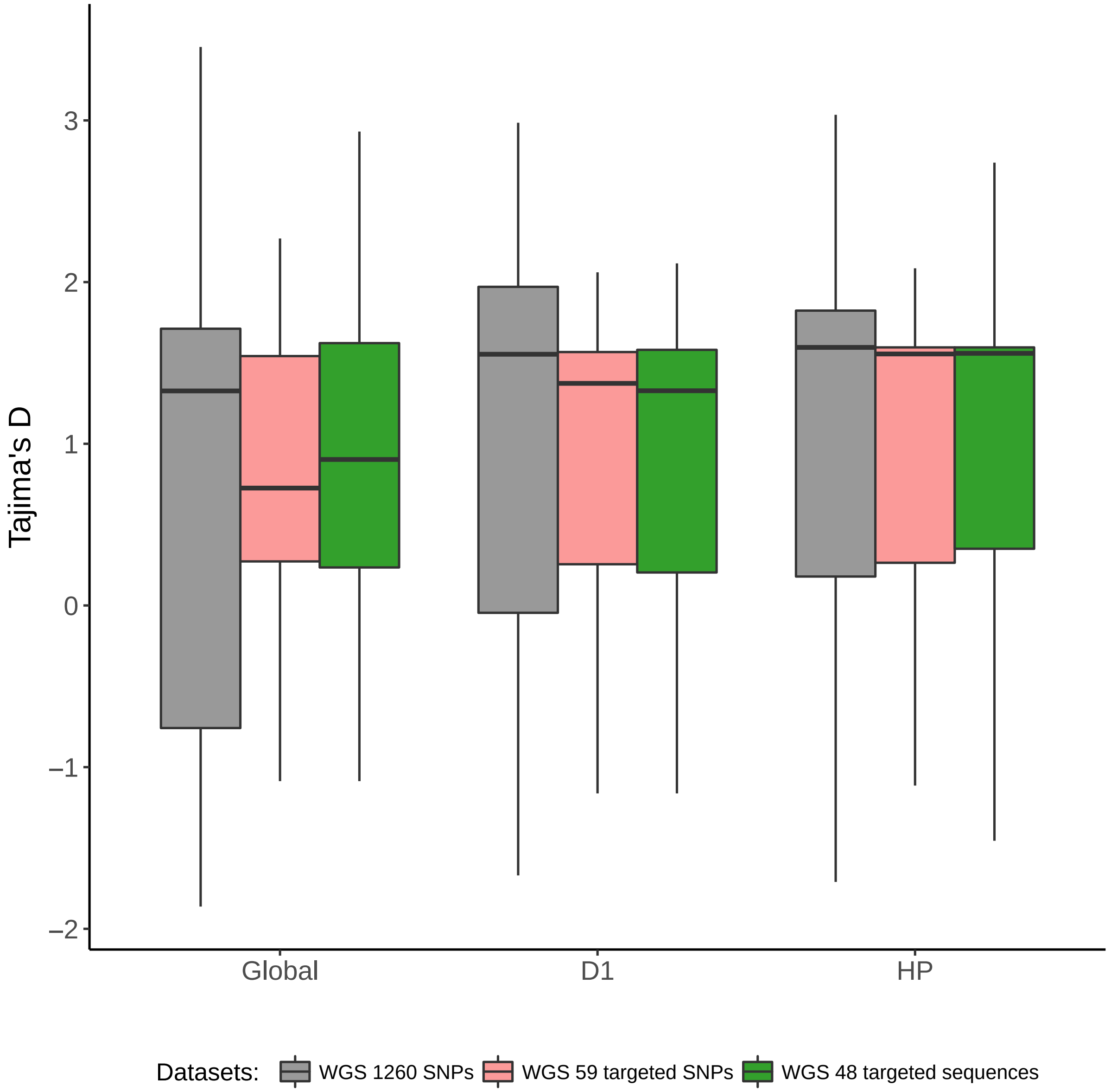
In silico comparisons



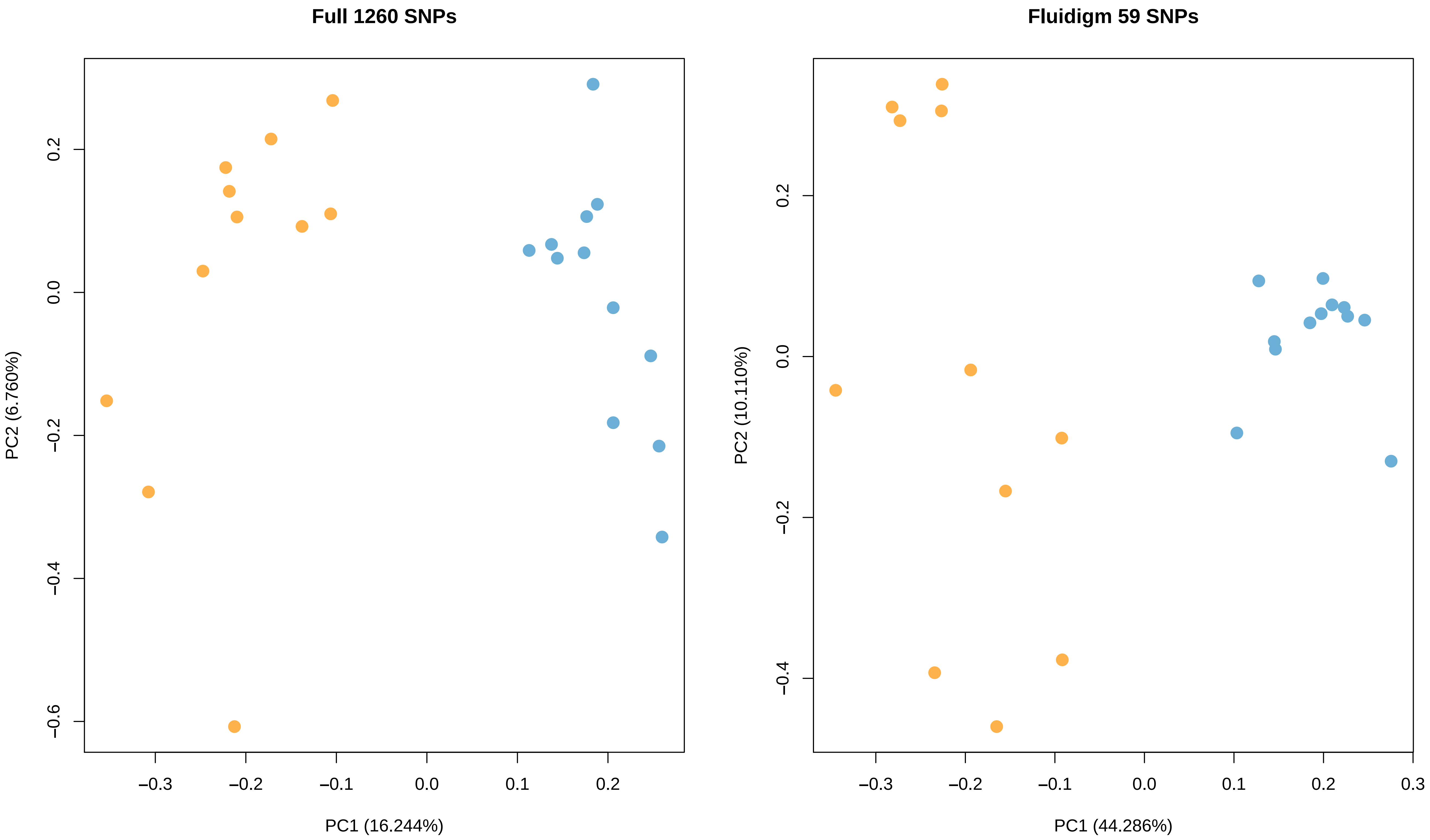


In silico comparisons

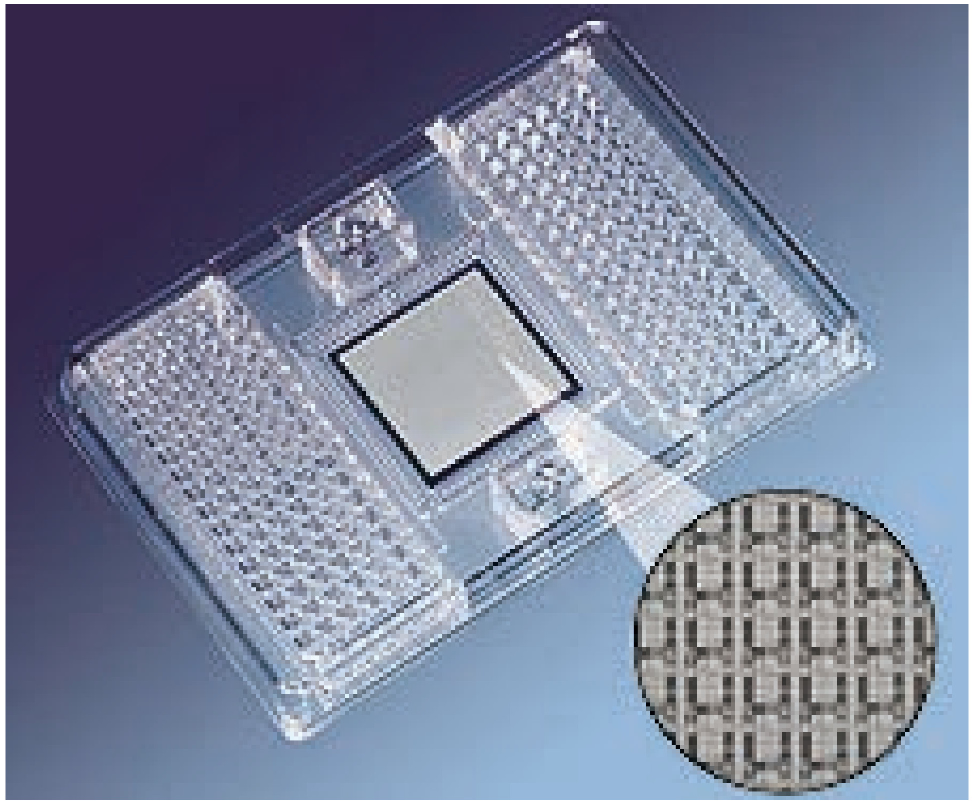


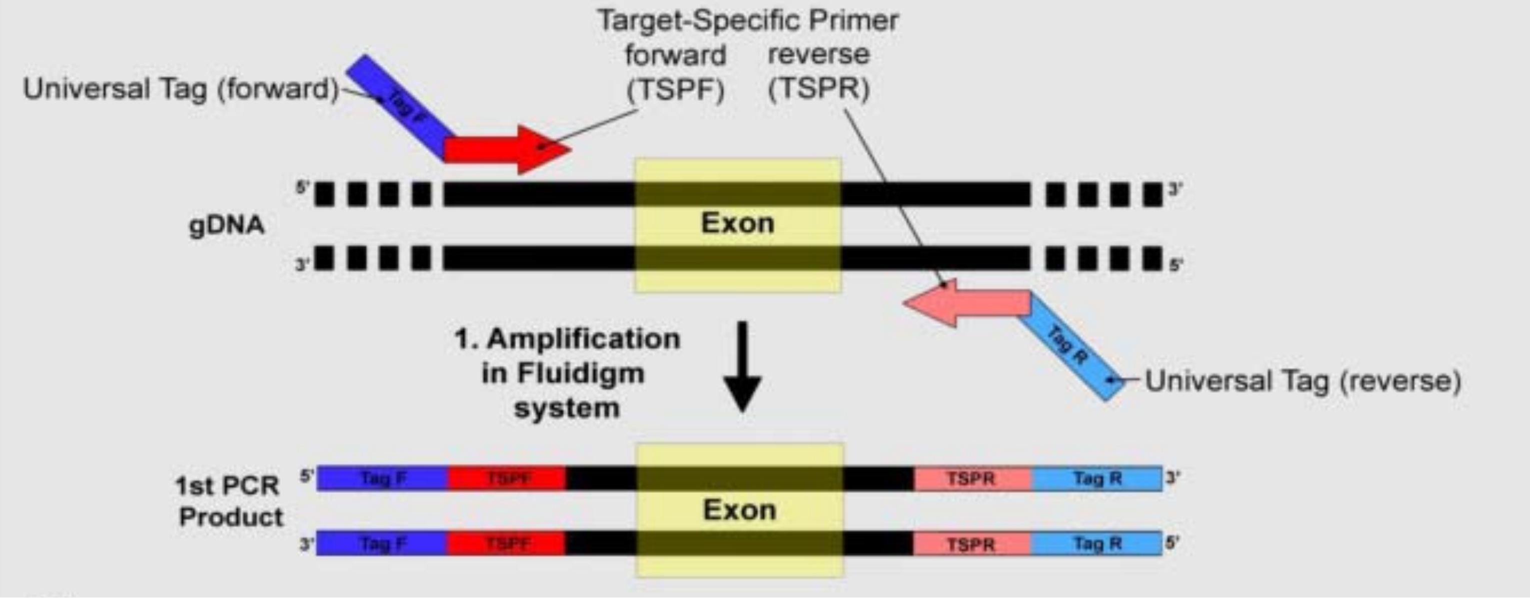


In silico comparisons

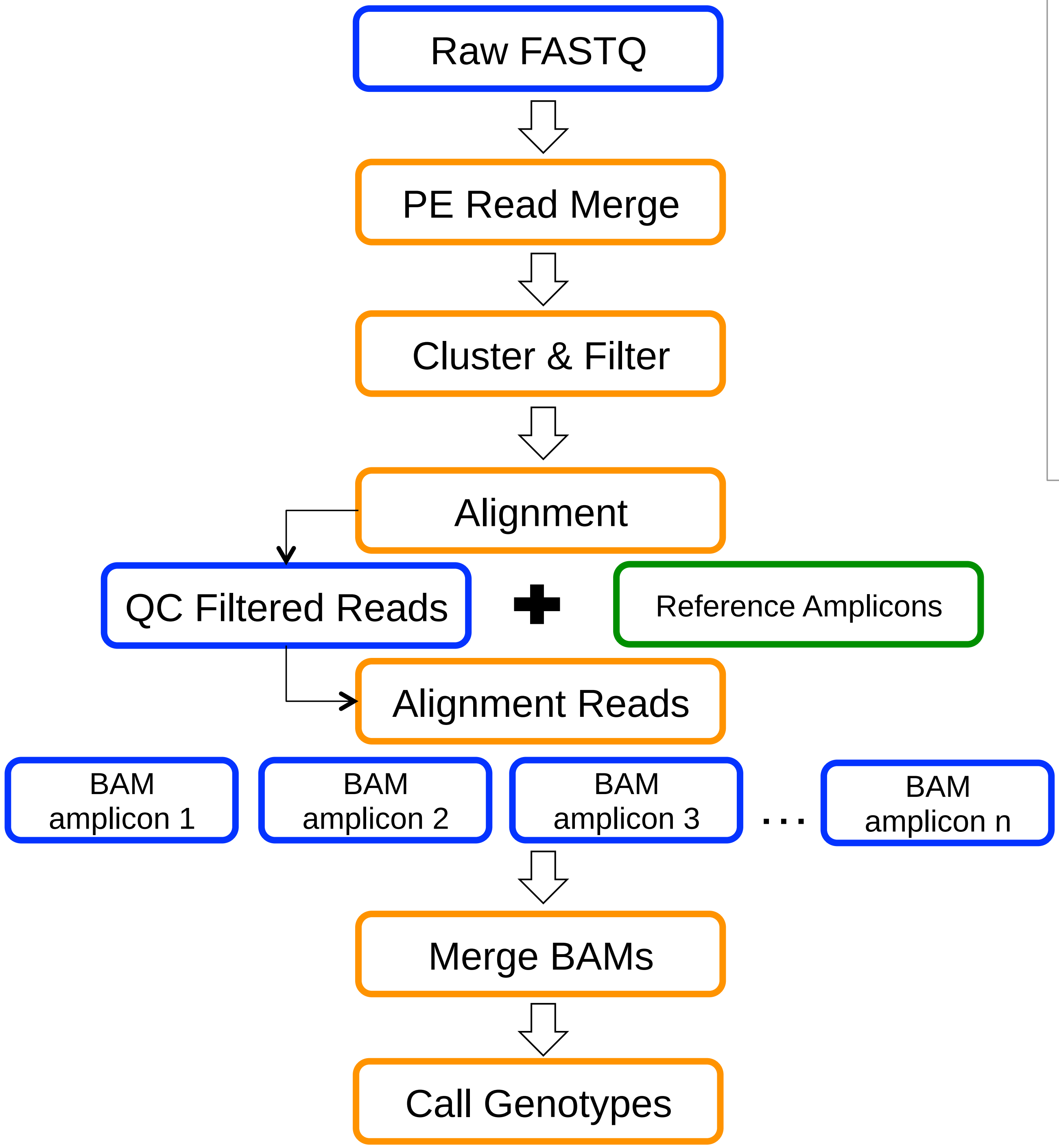


PCR-based Target Enrichment - Fluidigm





Pipeline for processing Amplicon data - PypeAmplicon



Raw FASTQ

PE Read Merge

Cluster & Filter

Alignment

QC Filtered Reads ✚ Reference Amplicons

Alignment Reads

BAM amplicon 1

BAM amplicon 2

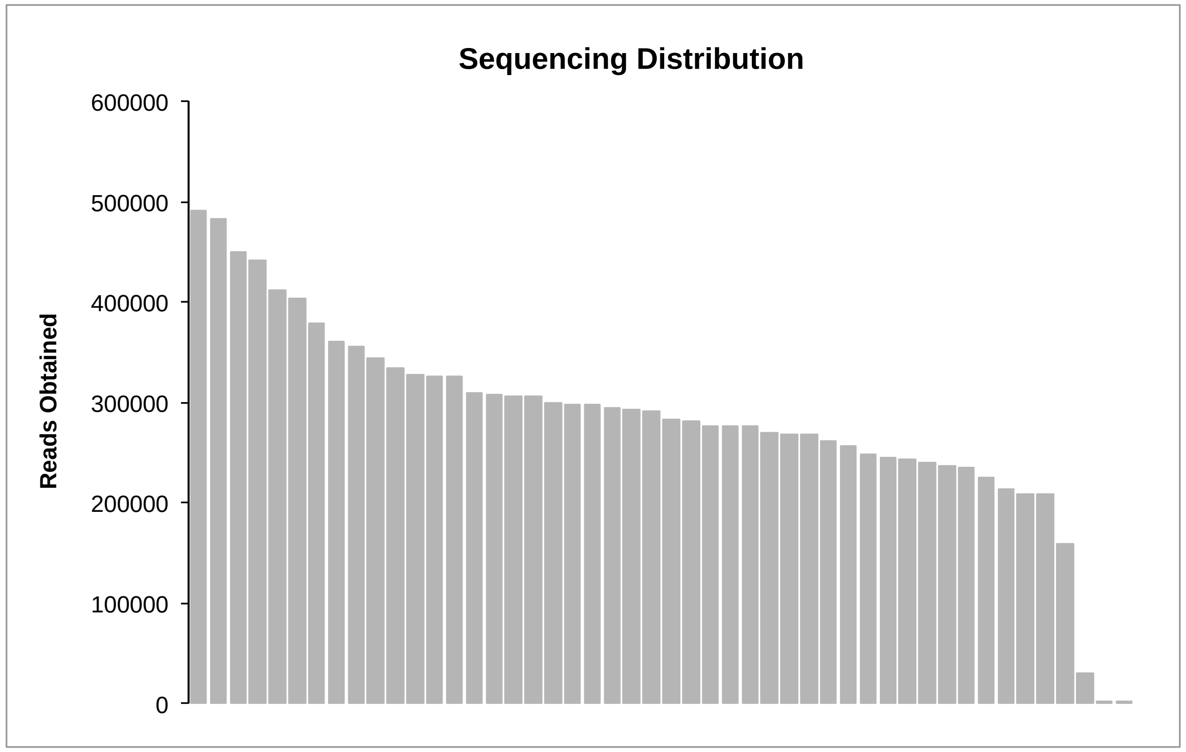
BAM amplicon 3

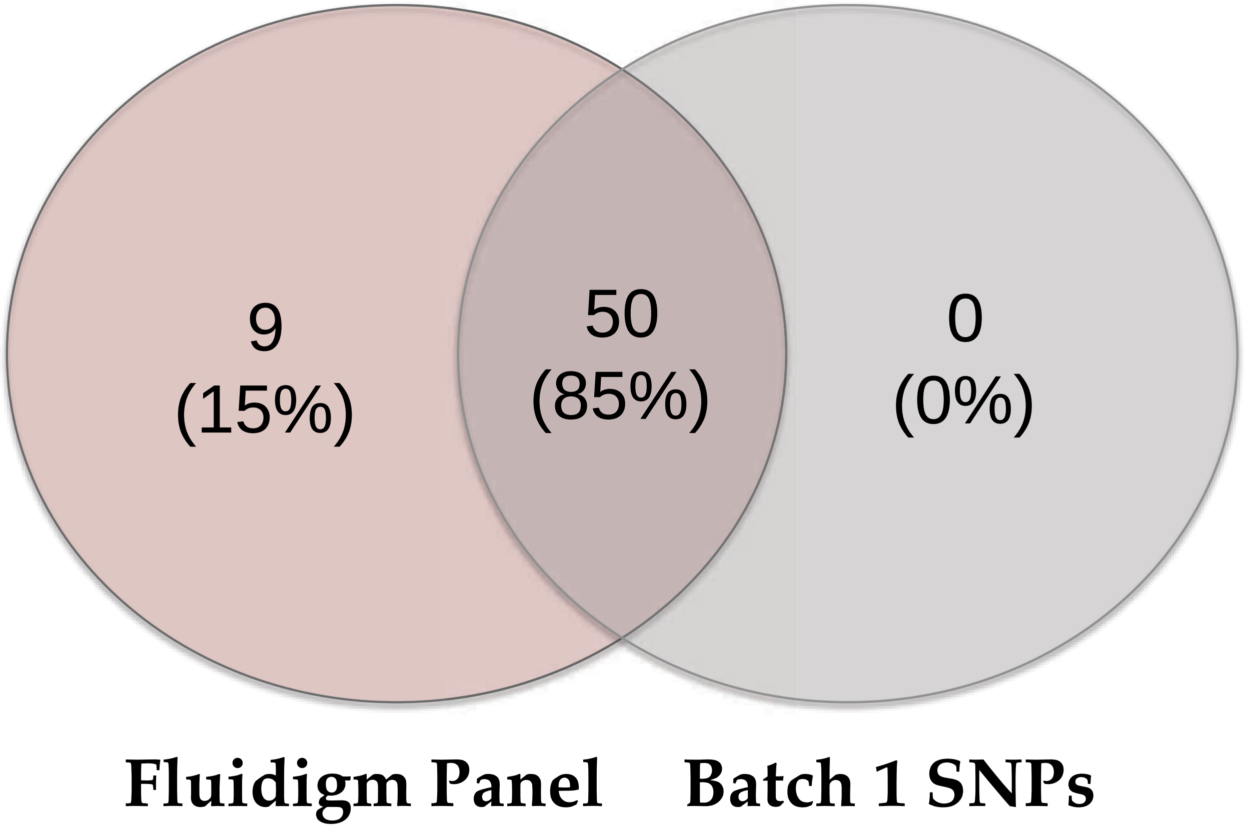
. . .

BAM amplicon n

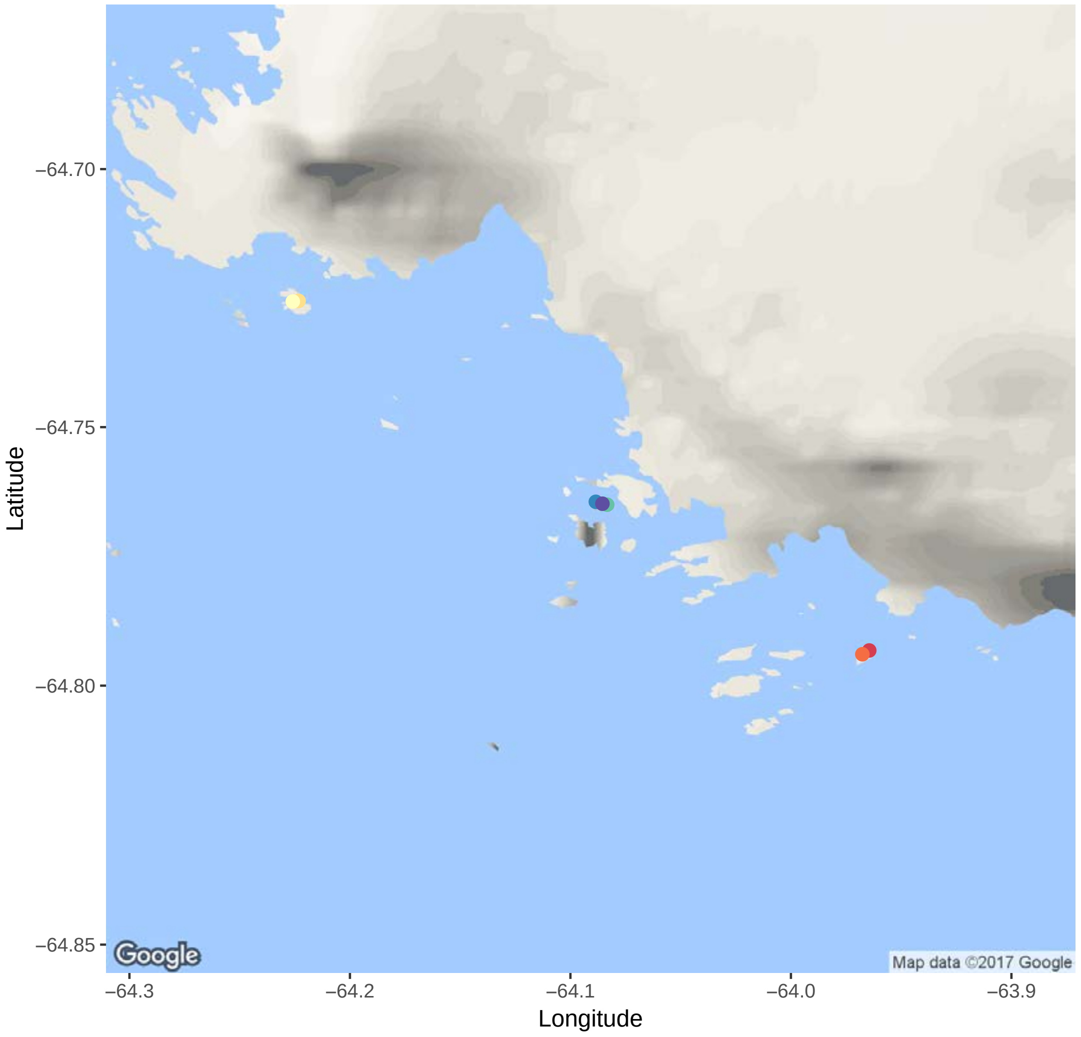
Merge BAMs

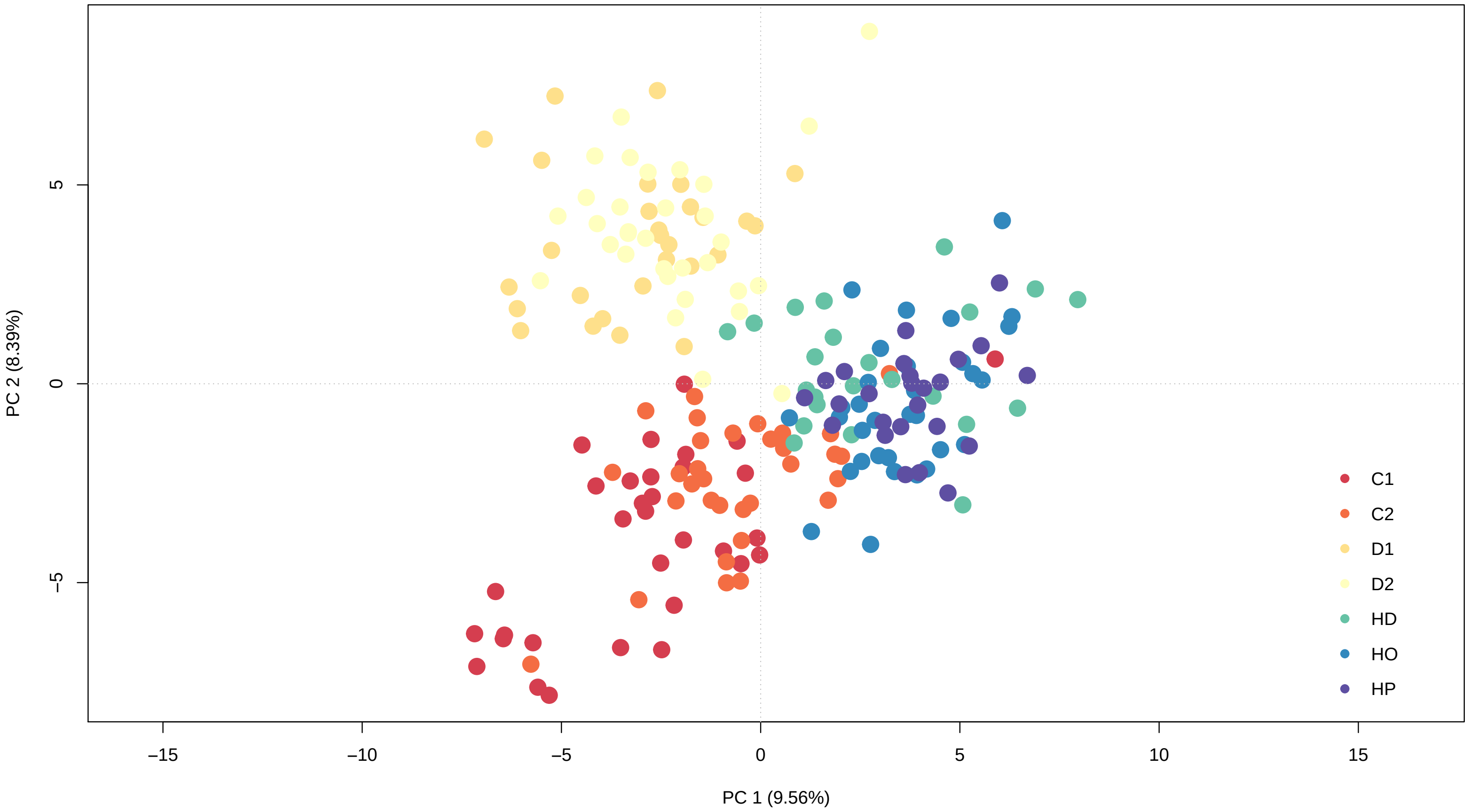
Call Genotypes





Preliminary results

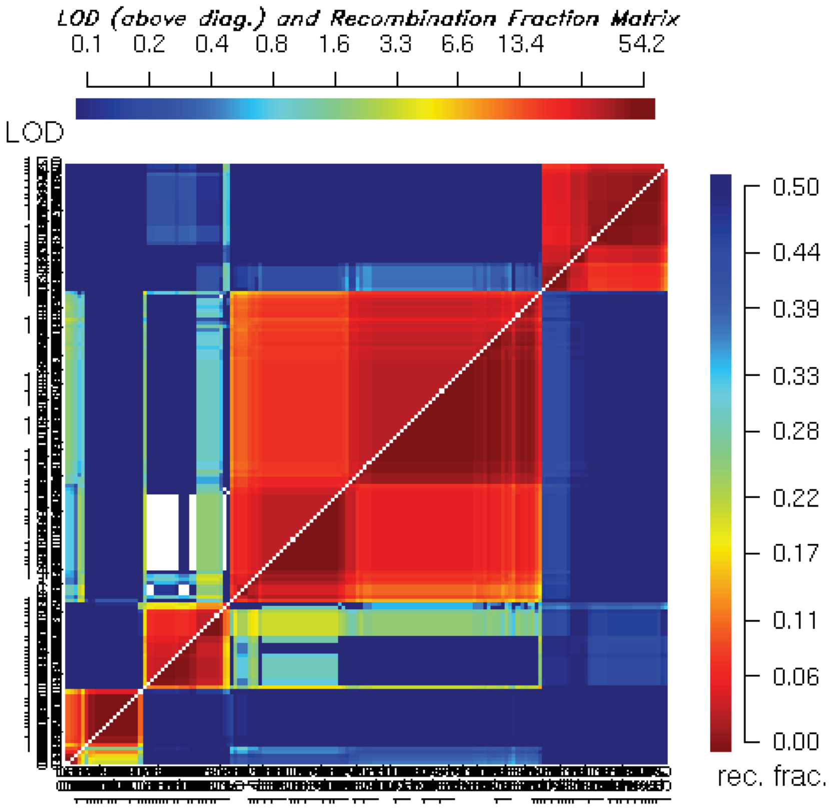


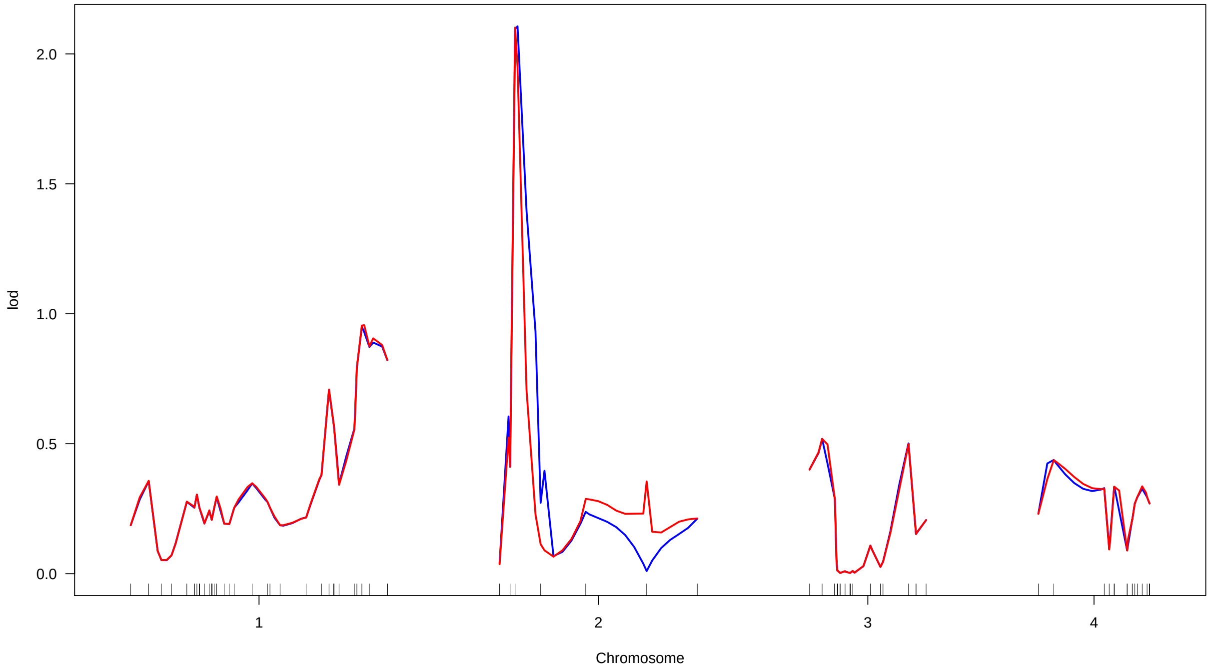


Challenges and future work

- Demographic Inference: WGS and amplicon data;

- Population Genomics





Determination of the inheritance and SNP markers associated with the virulence of soybean aphids on soybean

Doris Lagos-Kutz and Glen Hartman - University of Illinois
Anitha Chirumamilla - South Dakota State University
Andrew P. Michel - The Ohio State University

Most important pest of soybean (*Glycine max*) in USA[1,2]

Disrupt the photosynthetic process – even with low population densities[2]

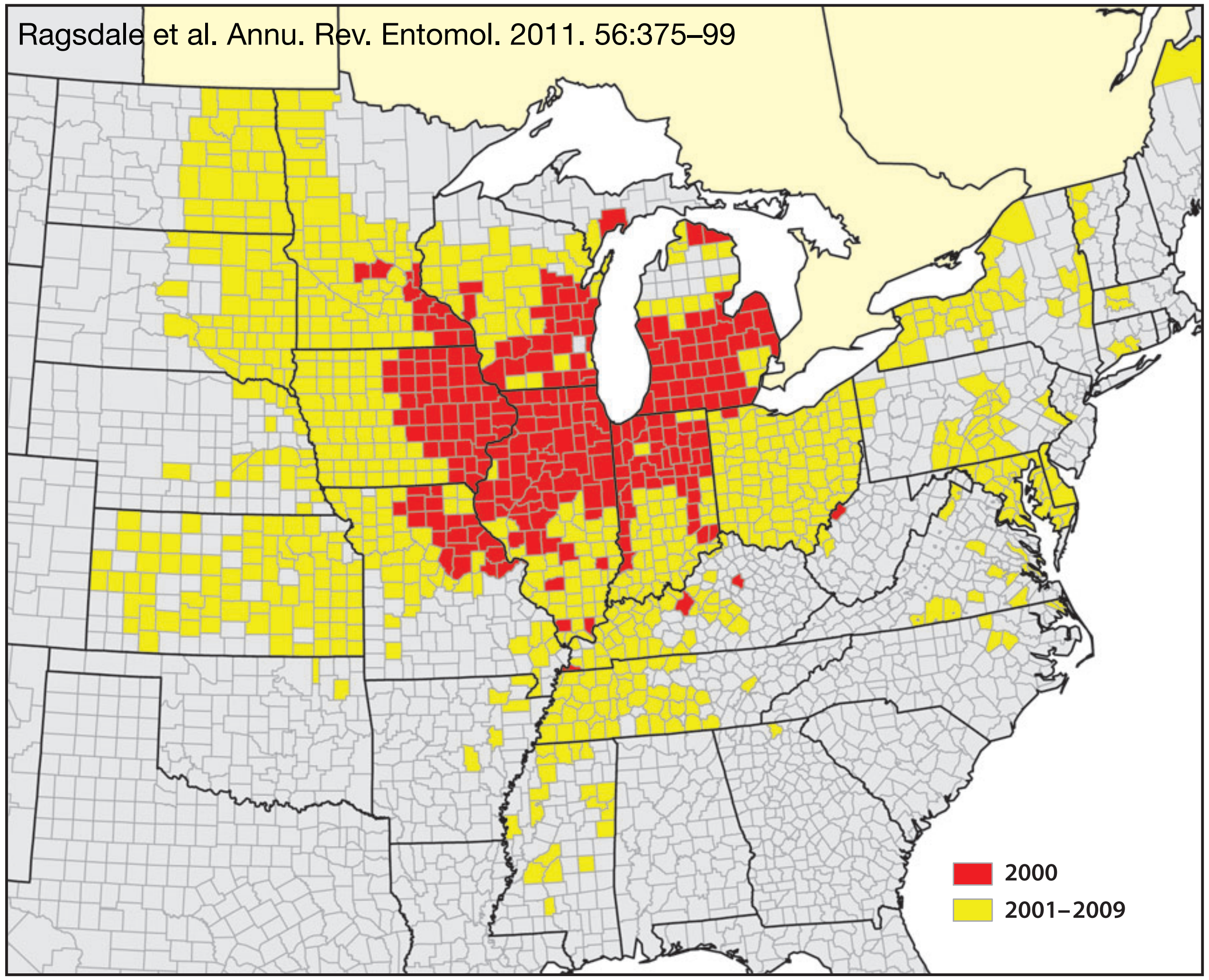
Vector of plant viruses[3]

Nebraska extension

[1]Heimpel et al. 2010; [2]Ragsdale et al. 2007; [3]Macedo et al. 2003; [4]Davis et al. 2005

Soybean Aphid - *Aphis glycines*

- Invasive species from Asia[1]
- It was first detected in North America in July 2000[2,3]
- It rapidly spread across 10 northcentral US states[4] (end of the summer same year);
- Primary host: common buckthorn - introduced from northern Europe;
- Secondary host: soybean

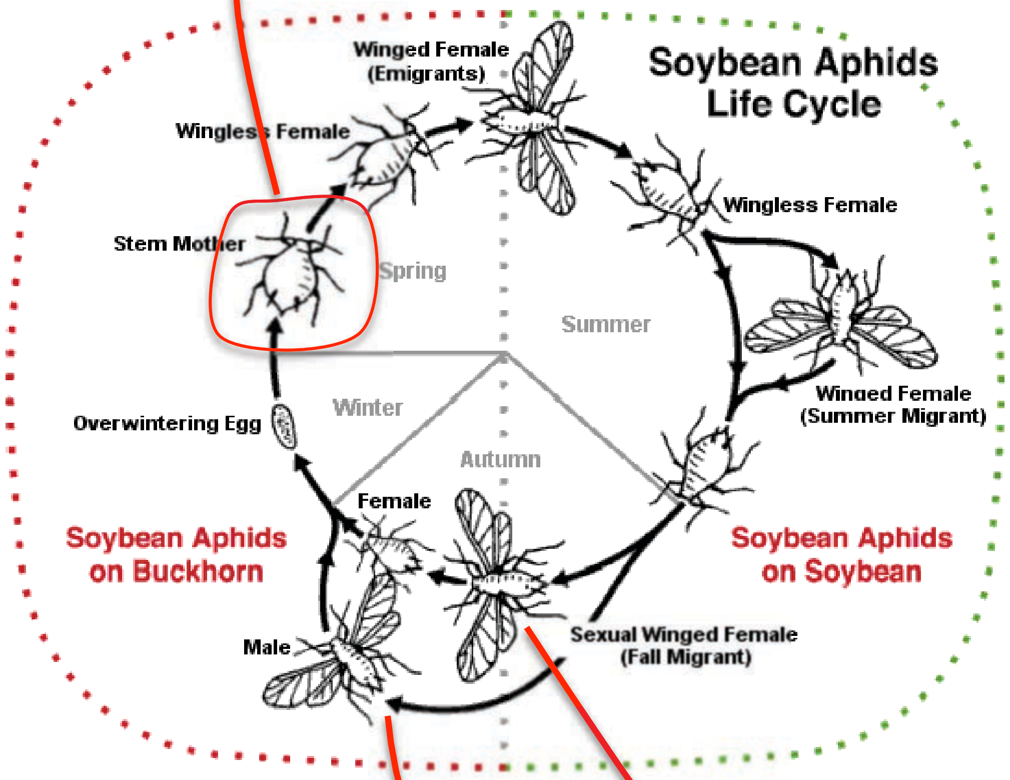


Ragsdale et al. Annu. Rev. Entomol. 2011. 56:375–99

[1]Venette & Ragsdale 2004; [2]Ragsdale et al. 2004; [3]Alleman; [4]Ragsdale et al. 2011

SBA Reproduction and Ecology

[relevant to obtain mapping population]



Fundatrices

Sexual Female migrate to Buckhorn

Sexual Female and Male

Heteroecious and **holocyclic** species

Sexual reproduction on primary host **[buckthorn trees]**

Asexual reproduction on secondary

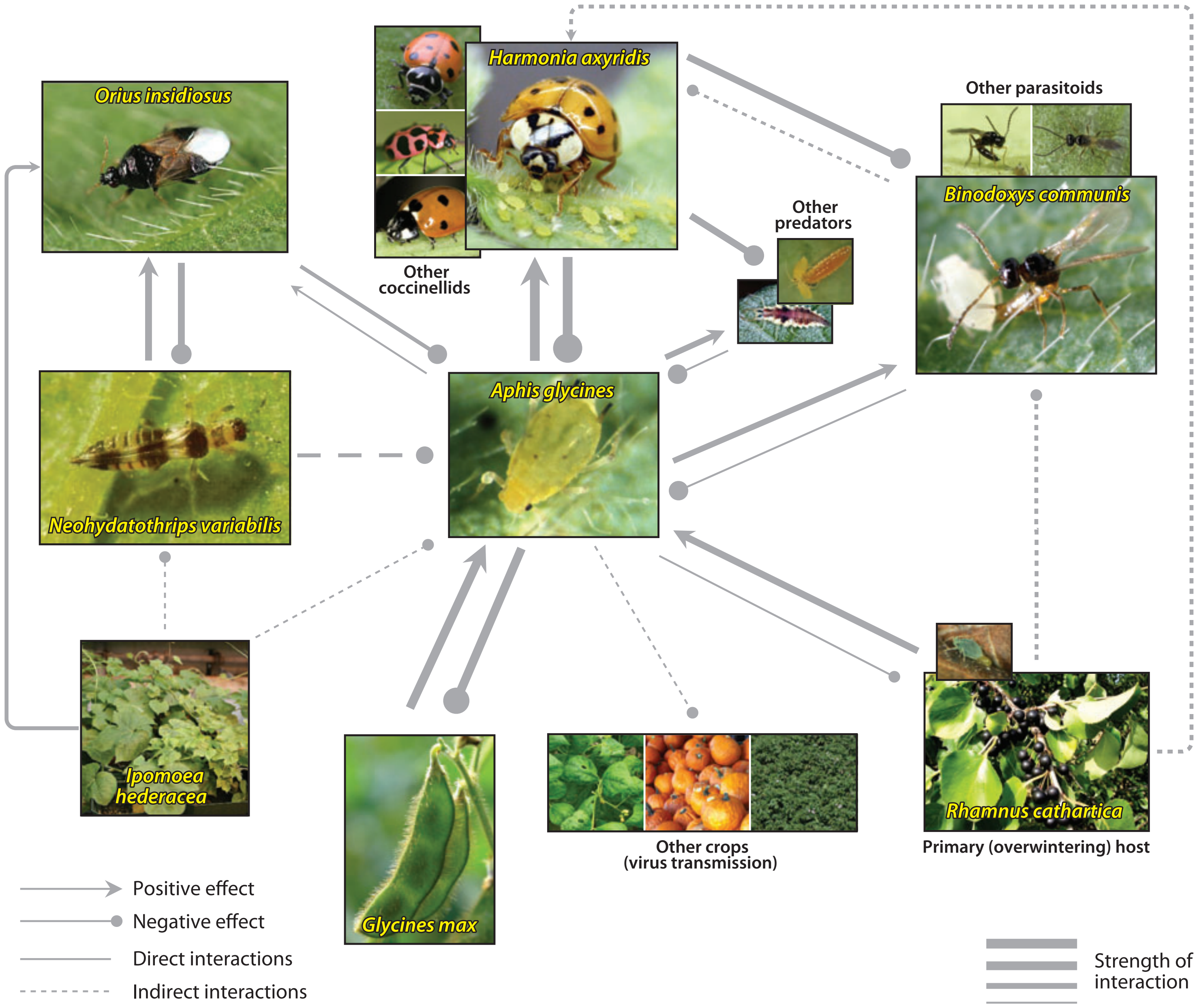
Common buckthorn (*Rhamnus cathartica*)

Management:

- Insecticides;

- **Natural Enemies**

- **Host-plant Resistance** (HPR) to SBA[1]



Ragsdale et al. Annu. Rev. Entomol. 2011. 56:375–99

---

[1] Hill et al. 2004

Several varieties and plant introductions (PIs) had been identified as been resistant to SBA:

- antibiosis;
- antixenosis;
- tolerance as mechanisms of resistance

"Rag" genes:

8 genes have been identified – HPR phenotypes

Jackson and Dowling varieties - resistant via antibiosis
PI 71506 - resistant via **antixenosis**[1]

Resistance in Jackson and Dowling:
1 single dominant gene *Rag1* - mapped linkage group M[2,3]

Resistance PI 243540:
1 single dominant gene *Rag2* - mapped linkage group F[4]

[1] Hill & Hartman 2004; [2] Hill & Hartman 2006a; [3] Hill & Hartman 2006b; [4] Mian et al. 2008.

Aphids adapted to the HPR[1]:

As early as 2005 aphids were observed that overcame resistance (i.e. virulent)

	15 days after infestation	
Res. Lines	IL	OH
Dowling (*Rag*1)	15	639
Jackson (*Rag*1?)	11	414
LD05-16611	29	627
Suscept. Lines		
Dwight	468	398
Williams 82	726	574

Biotypes	Rag1	Rag2	Rag1/Rag2
Biotype1	Avirulent	Avirulent	Avirulent
Biotype 2	Virulent	Avirulent	Avirulent
Biotype 3	Avirulent	Virulent	Avirulent
Biotype 4	Virulent	Virulent	Virulent

---

[1] Kim et al. 2008

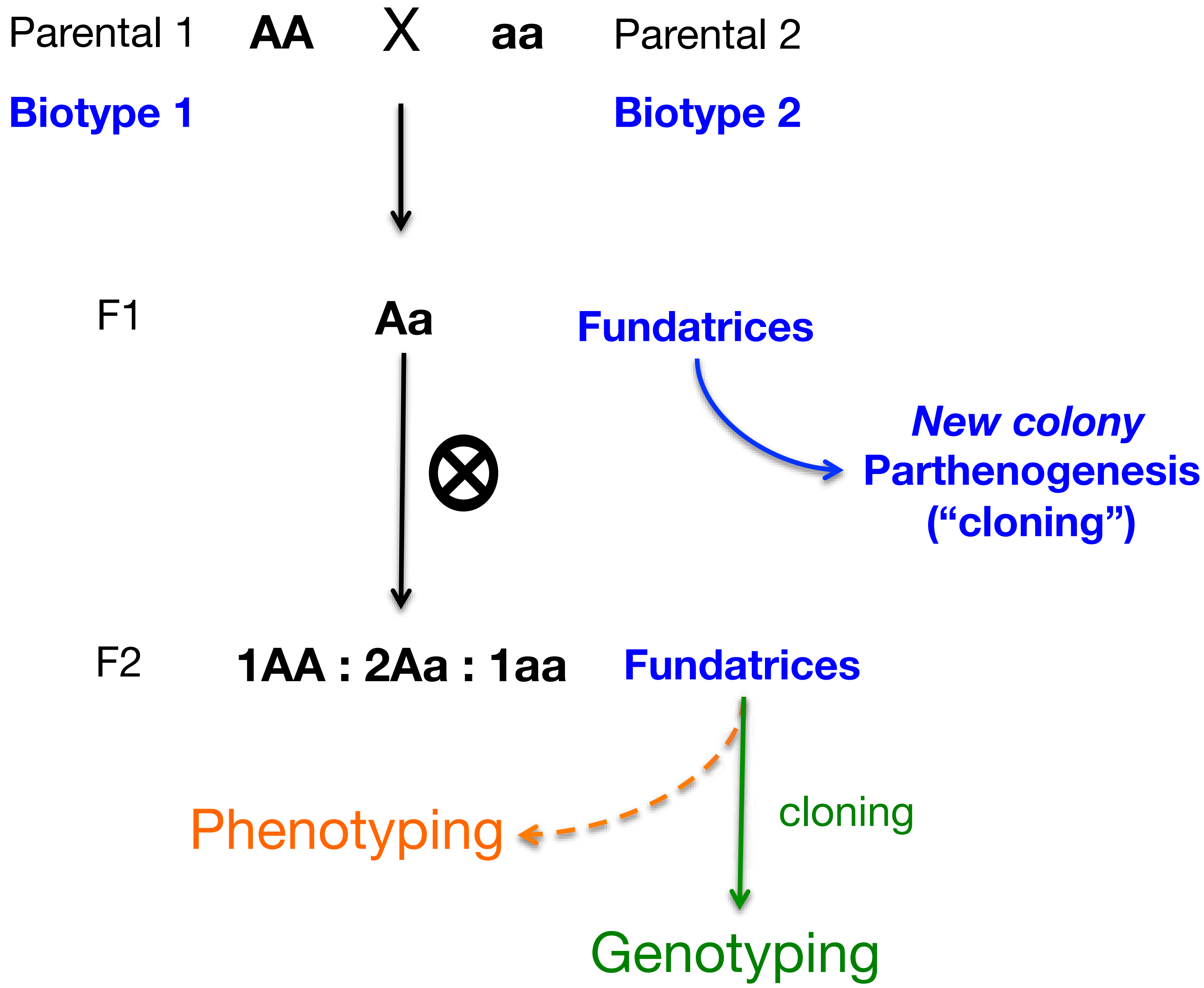
Finding the underlying genetic basis of SBA virulence
(if it is genetic)

- **Comparative Genome Scan** - PoolSeq (Wenger et al. *in preparation*):
  - 84 SNPs;

- **Differential gene expression analysis** (Yates et al *in preparation*):
  - Constitutive differences between Biotype 1 and Biotype 4.

- **Genetic-Phenotype Association - QTL mapping**

Mapping population: F2 intercross



Parental 1 **AA** X **aa** Parental 2

Biotype 1 **Biotype 2**

F1 **Aa** **Fundatrices**

⊗

New colony
Parthenogenesis
(“cloning”)

F2 **1AA : 2Aa : 1aa** **Fundatrices**

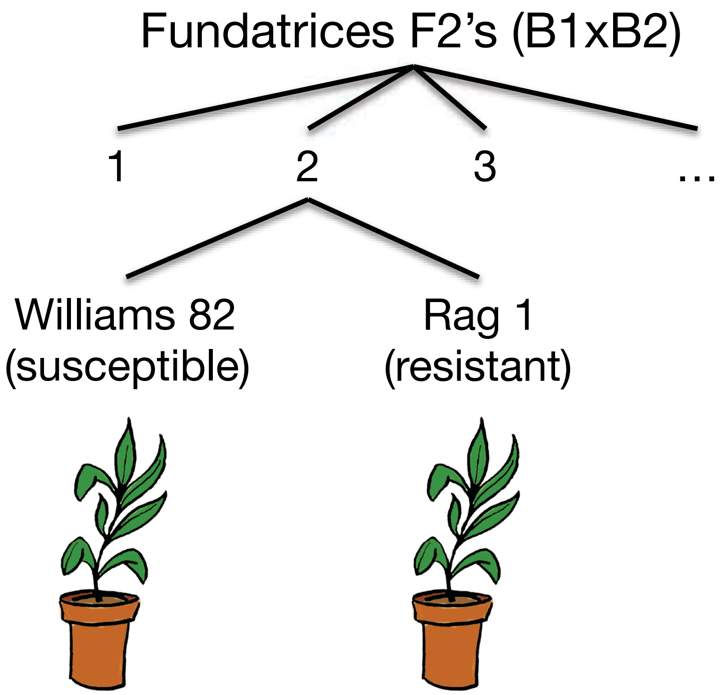
Phenotyping

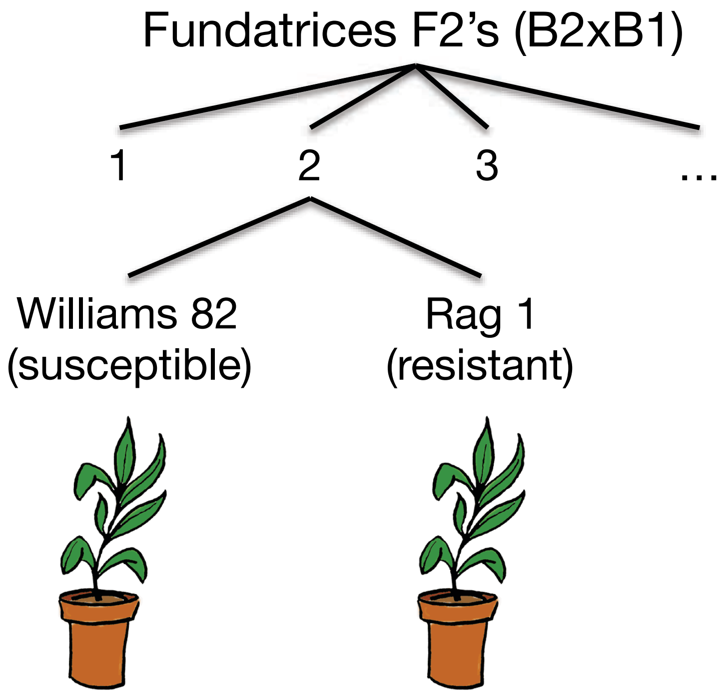
cloning

Genotyping

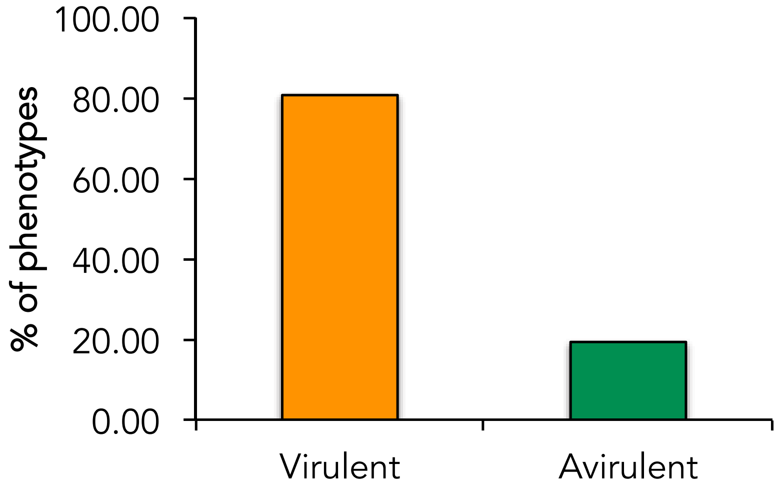
Phenotypic data

	Williams 82	Rag 1	Phenotype
Avirulent	1	0	0
Virulent	1	1	1

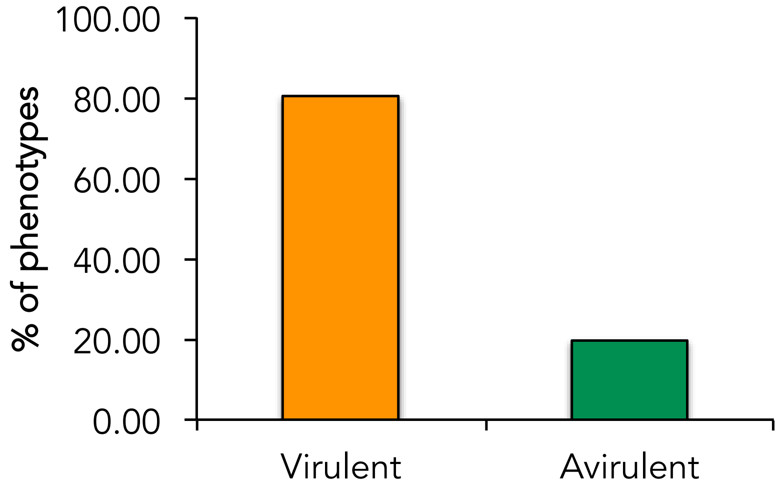




♀B1 x ♂B2 crosses



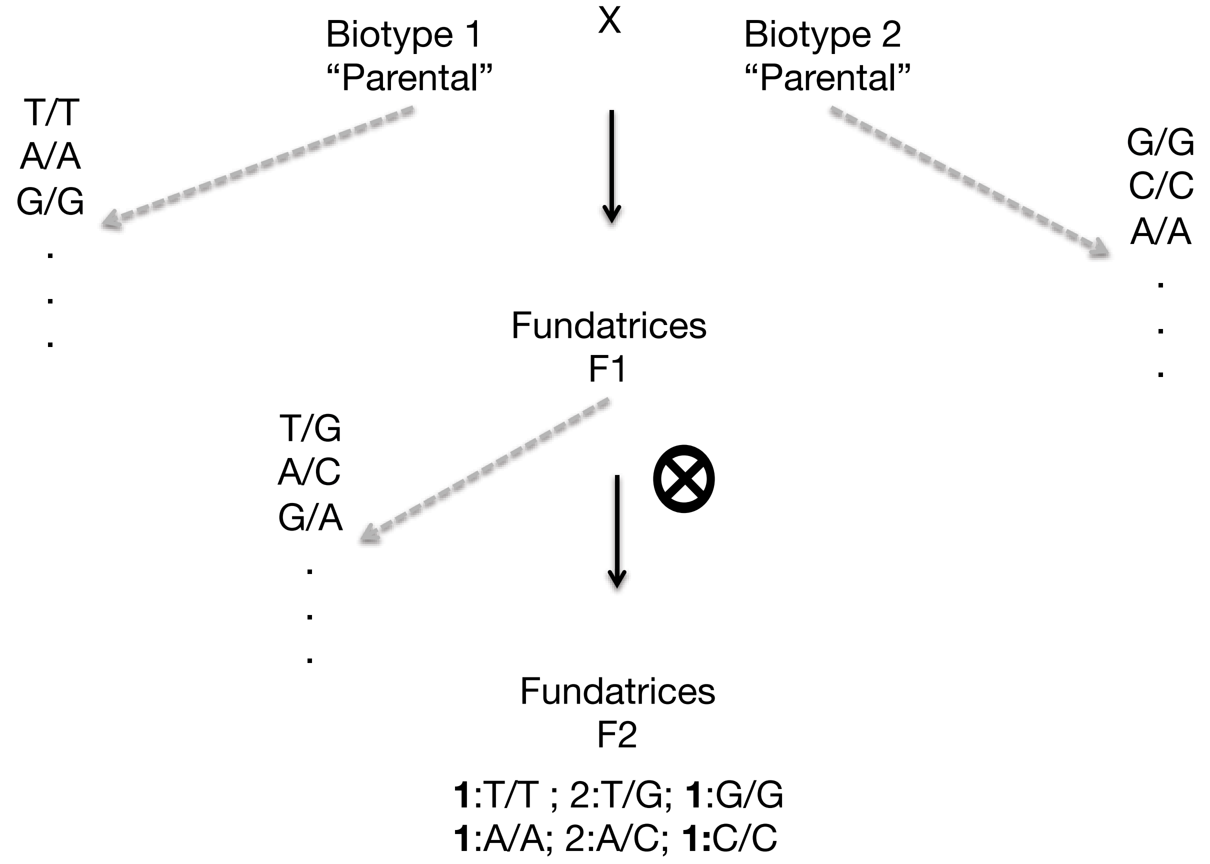
♀B2 x ♂B1 crosses



Marker discovery - panel design - Genotypic data

WGS of biotype pools
(B1, B2 and B3)

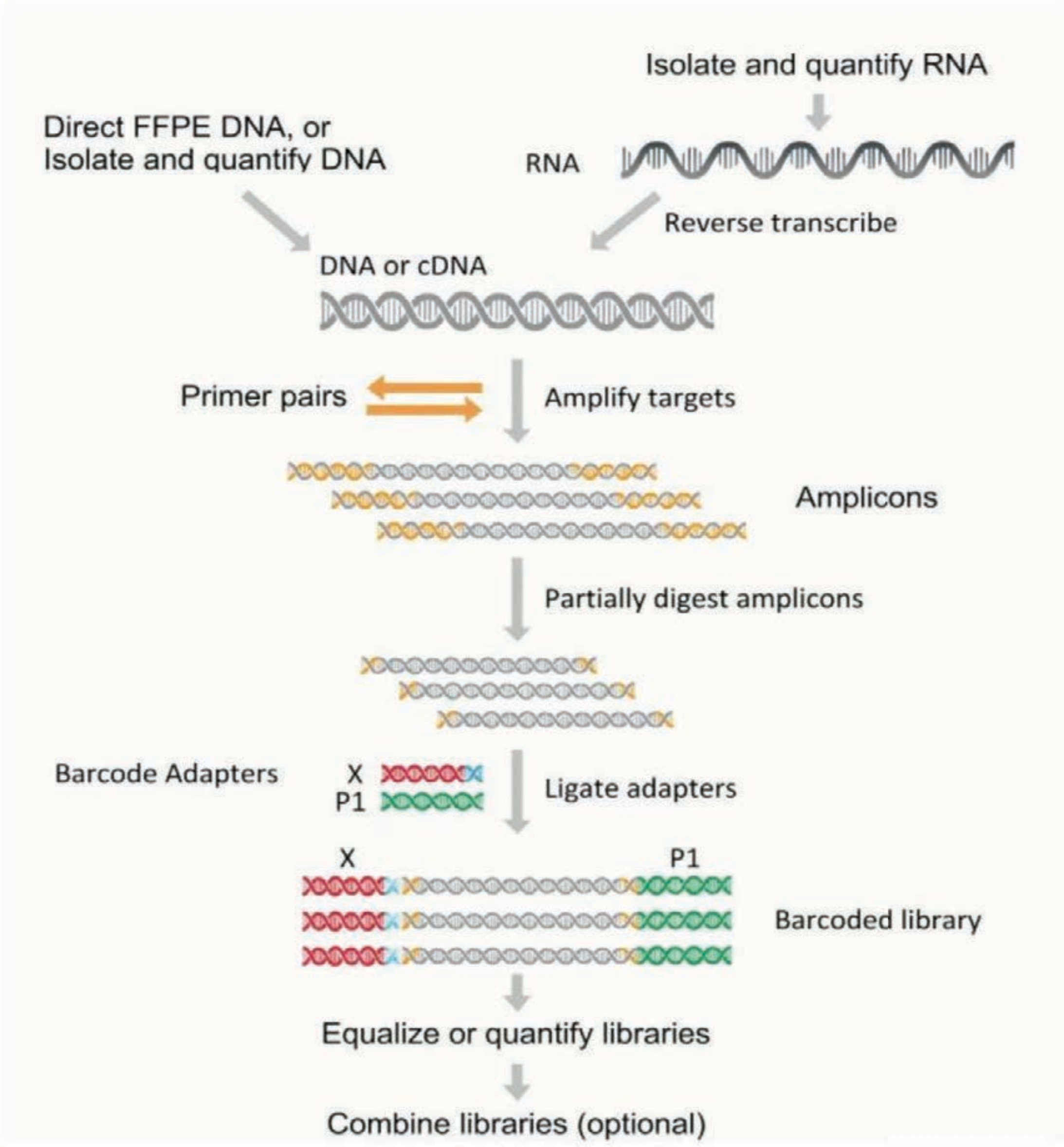
Reference-based SNP calling
First draft of the genome (Wenger et al 2017)



ION AMPLISEQ™

FLORAGENEX
Your partner from DNA to data

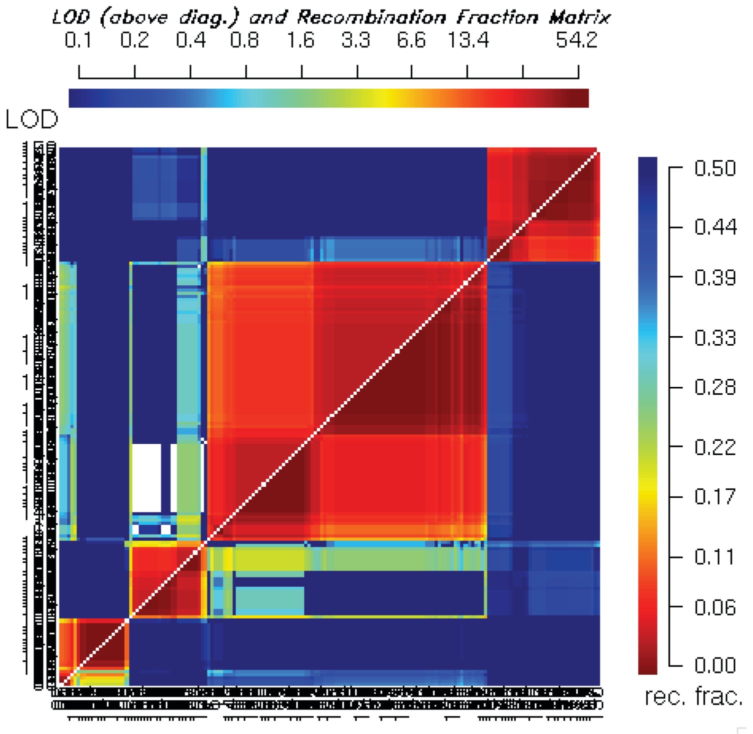
391 targeted SNPs
378 amplicons

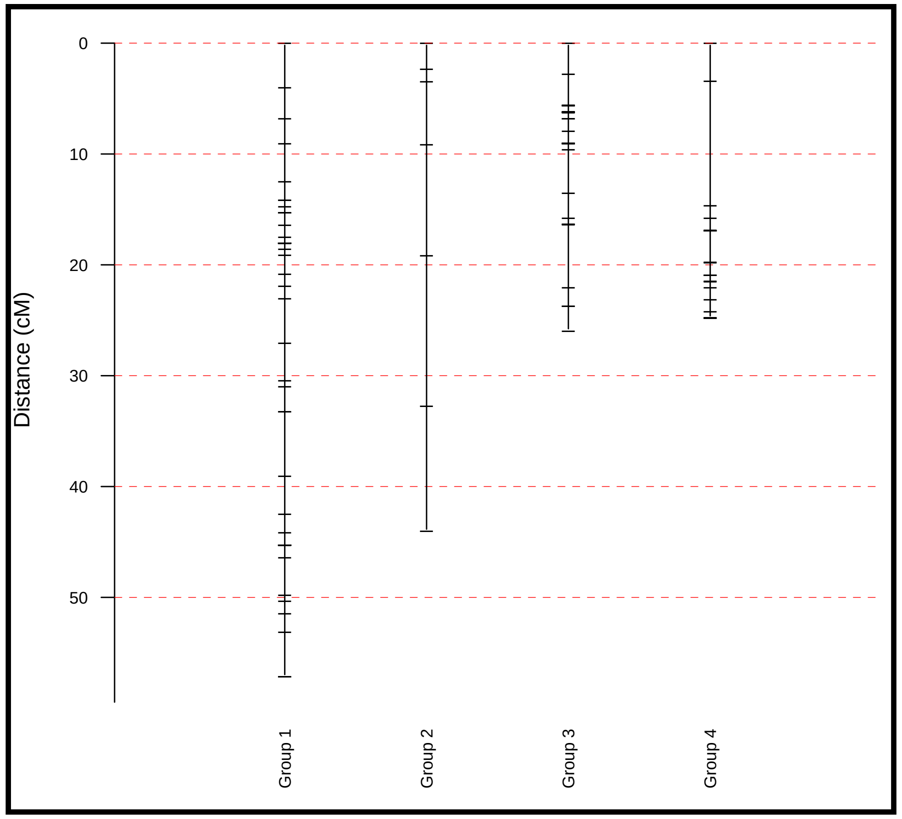


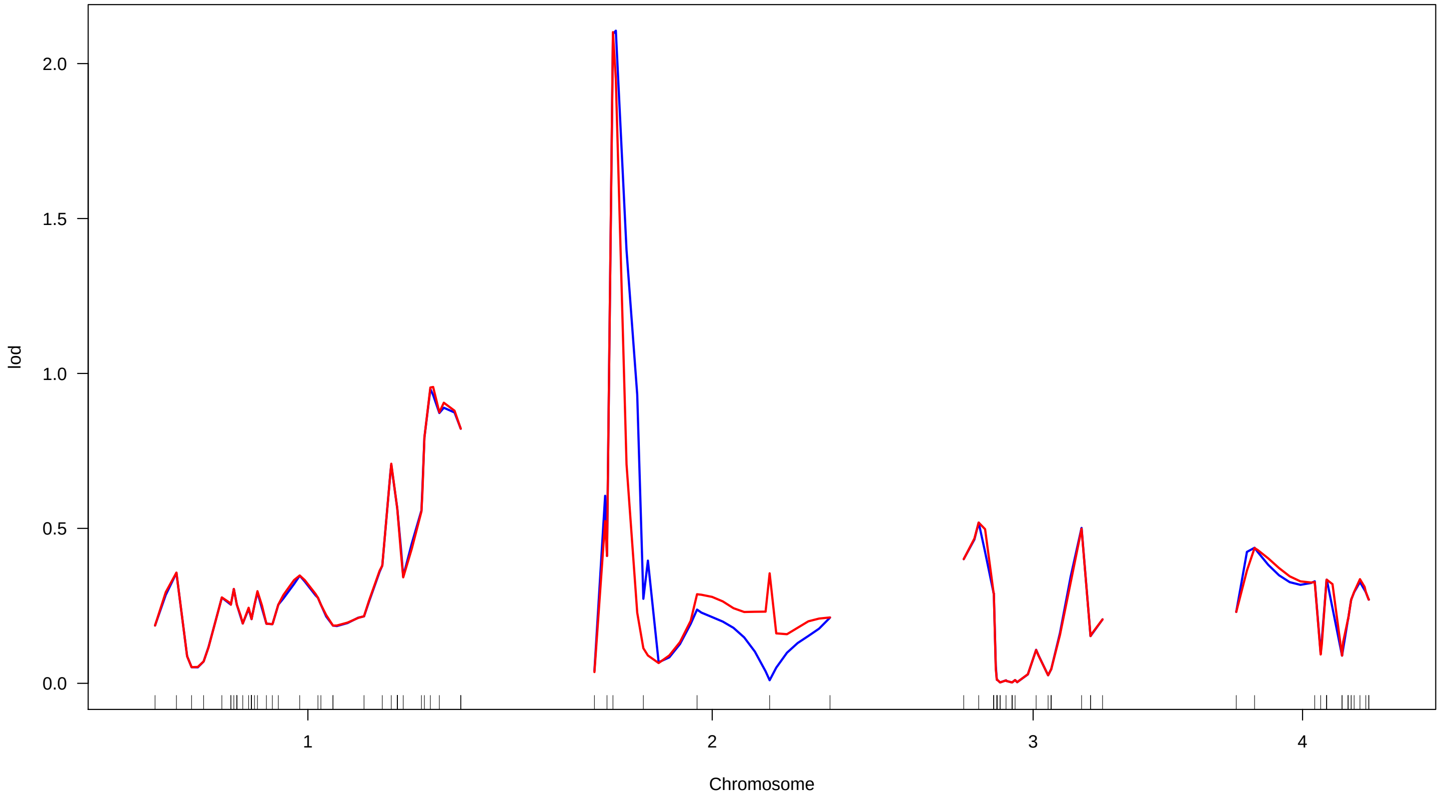
240 targeted SNPs

980 N-targeted SNPs

Mapping


LOD (above diag.) and Recombination Fraction Matrix
0.1
0.2
0.4
0.8
1.6
3.3
6.6
13.4
54.2
LOD
0.50
0.44
0.39
0.33
0.28
0.22
0.17
0.11
0.06
0.00
rec. frac.


0
10
20
30
40
50
Distance (cM)
Group 1
Group 2
Group 3
Group 4


2.0
1.5
1.0
0.5
0.0
lod
1
2
3
4
Chromosome

Challenges and future work

High percentage of targeted and non-targeted markers with segregation distortion

Inconsistency in the genotypic data of F1

- Reduced heterozygosity in parental lines;

- Not account for variation in the analysis of pool data;

- Reproductive isolation between biotypes (or partial RI as indicate the SG data);
  - Symbionts (*Wolbachia*);

  - Incompatibility - Hybridization / Introgression with other aphid species

Tracking selection in time-series population genomic data using ABC random forests

Miguel Navascués - UMR CBGP, INRA
Jean-Michel Marin - UMR IMAG, Université de Montpellier

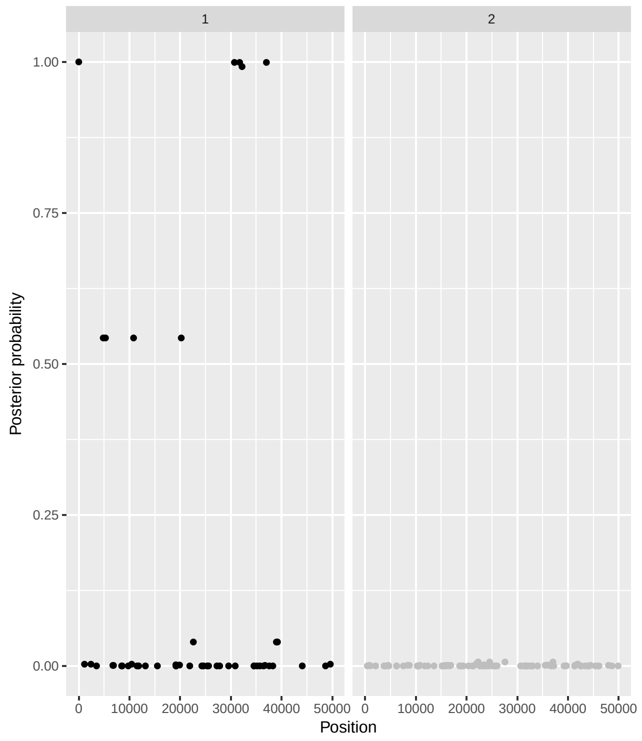
Postdoc InterLabex - ABCSelection

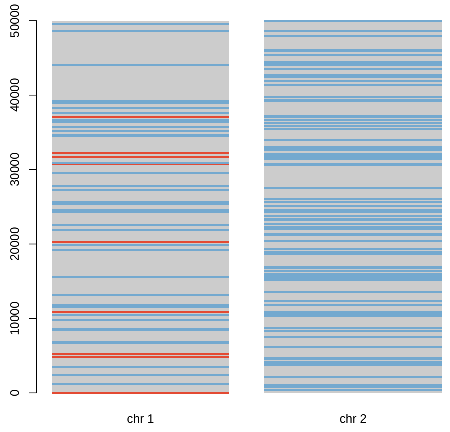
Labex AGRO

Labex CeMEB

Labex NUMEV

Framework: ABC-RF in Times-series PopGenomics Data Demography and Selection





Genetic relationship of biotypes

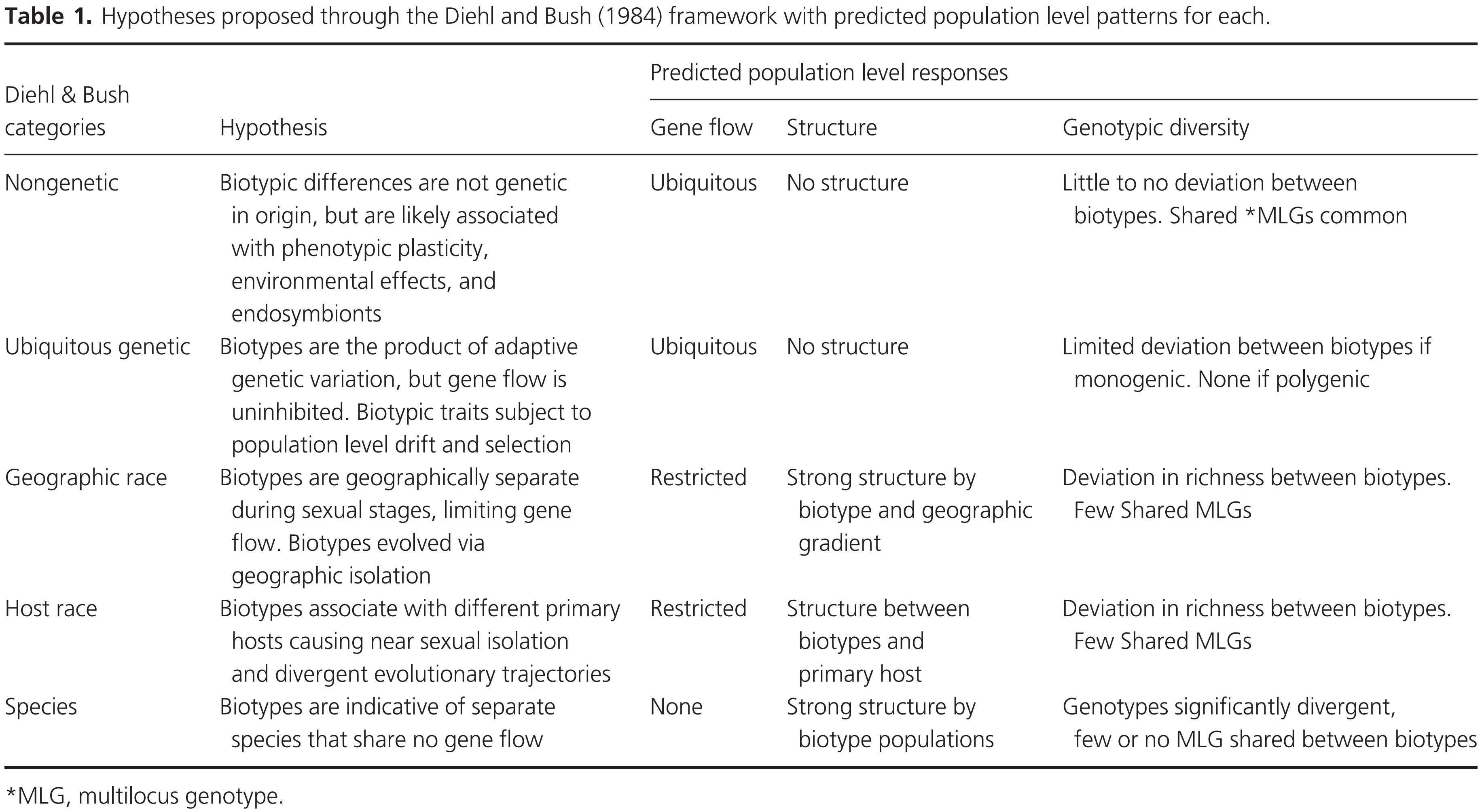


Table 1. Hypotheses proposed through the Diehl and Bush (1984) framework with predicted population level patterns for each.

Diehl & Bush categories	Hypothesis	Predicted population level responses		
		Gene flow	Structure	Genotypic diversity
Nongenetic	Biotypic differences are not genetic in origin, but are likely associated with phenotypic plasticity, environmental effects, and endosymbionts	Ubiquitous	No structure	Little to no deviation between biotypes. Shared *MLGs common
Ubiquitous genetic	Biotypes are the product of adaptive genetic variation, but gene flow is uninhibited. Biotypic traits subject to population level drift and selection	Ubiquitous	No structure	Limited deviation between biotypes if monogenic. None if polygenic
Geographic race	Biotypes are geographically separate during sexual stages, limiting gene flow. Biotypes evolved via geographic isolation	Restricted	Strong structure by biotype and geographic gradient	Deviation in richness between biotypes. Few Shared MLGs
Host race	Biotypes associate with different primary hosts causing near sexual isolation and divergent evolutionary trajectories	Restricted	Structure between biotypes and primary host	Deviation in richness between biotypes. Few Shared MLGs
Species	Biotypes are indicative of separate species that share no gene flow	None	Strong structure by biotype populations	Genotypes significantly divergent, few or no MLG shared between biotypes

*MLG, multilocus genotype.

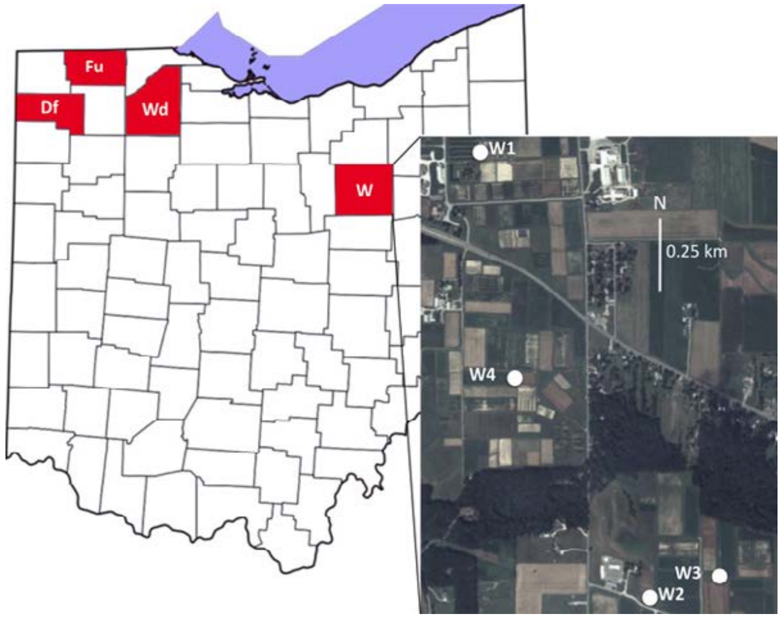
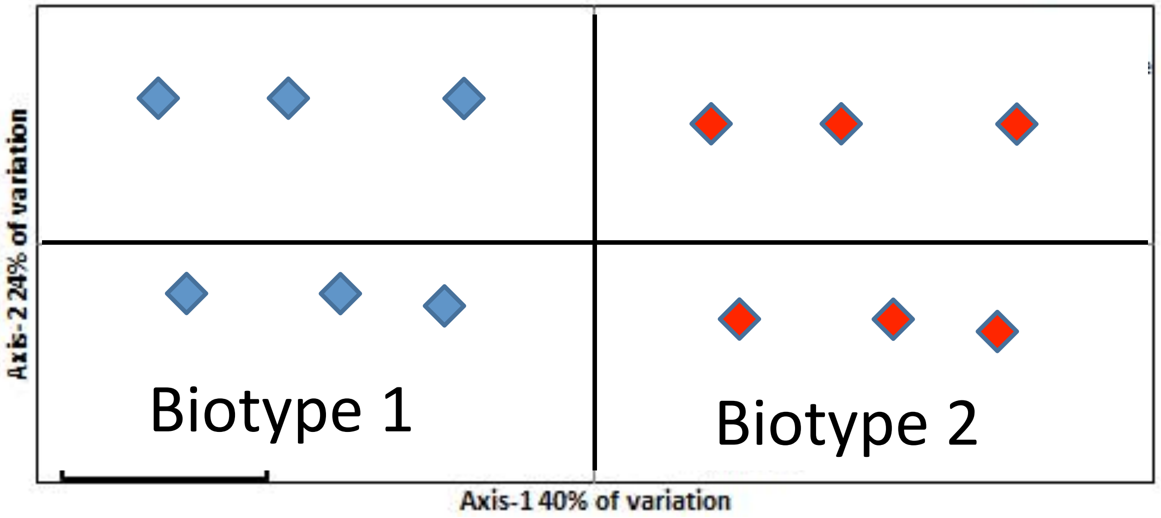


Figure 1 Geographic location of collections. Abbreviations and GPS coordinates available in Table 2.



Diehl & Bush 1984; Wenger & Michel 2013

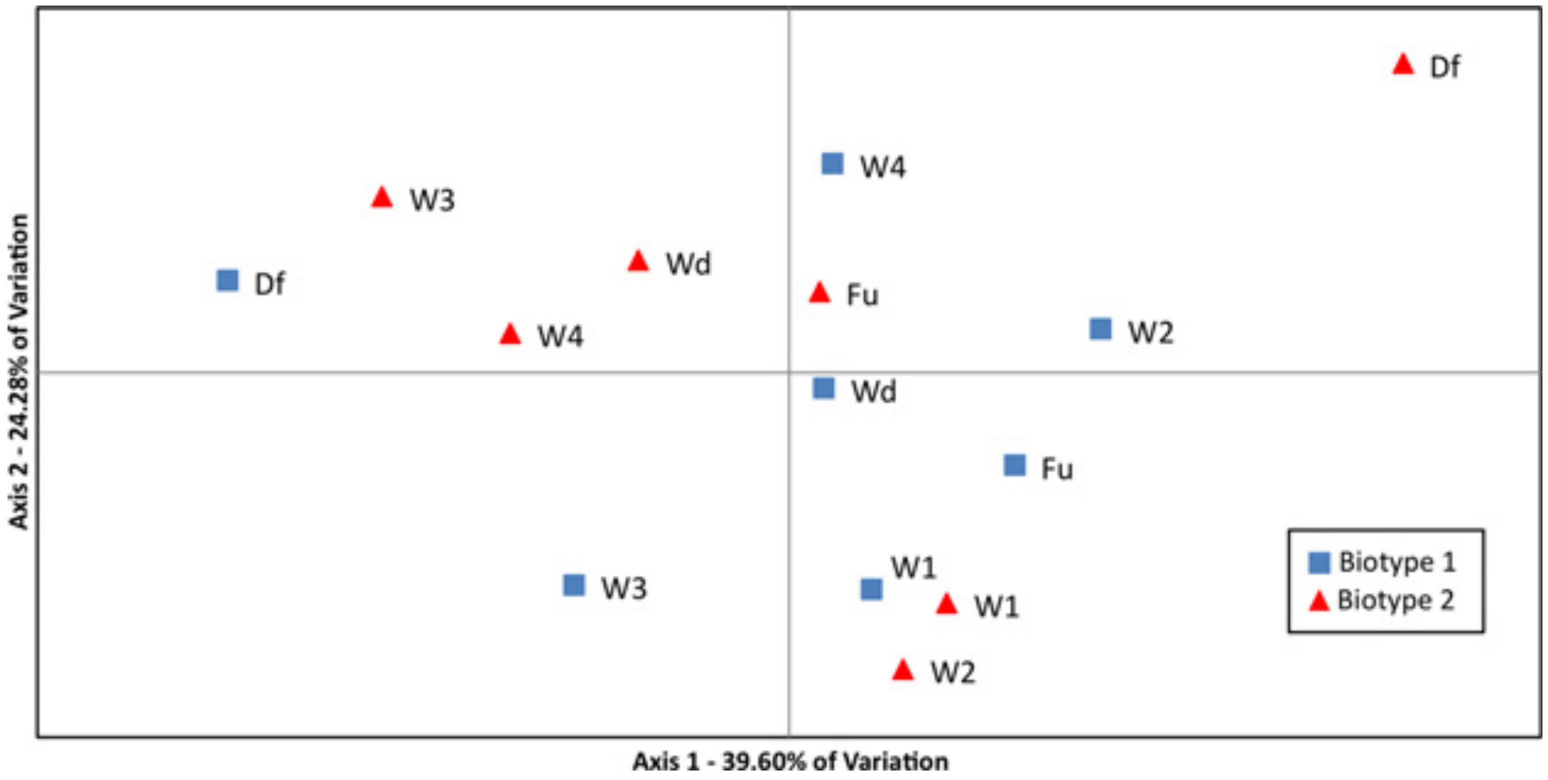


Table 5. Predicted patterns of gene flow, population structure, and genotypic diversity per Diehl & Bush (1984) category and sub-category. Categories in bold are consistent with data in the present study.

Diehl & Bush category	Sub-category	Genetic differentiation	Population structure	Genotypic diversity
Nongenetic	Endosymbiont	Between biotype	Minor structure biotype or geography	Biotype 2 less diverse
Nongenetic	**Phenotypic Plasticity**	**None**	**No clustering or by geography**	**No significant difference between biotypes**
Ubiquitous genetic	**Gene for Gene**	**None**	Minor structure biotype or geography	Biotype 2 less diverse
Ubiquitous genetic	**Epistasis**	**None**	**No clustering or by geography**	**No significant difference between biotypes**
Geographic race	N/A	Strong between biotype	Cluster by biotype	Sig Dif in diversity
Host race	N/A	Strong between biotype	Cluster by biotype	Significant difference in diversity
Species	N/A	Strong between biotype	Cluster by biotype	Significant difference in diversity

Threshold Model