Distribution and adaptation of *Xylella fastidiosa* in Californian grapevines

a genomic perspective

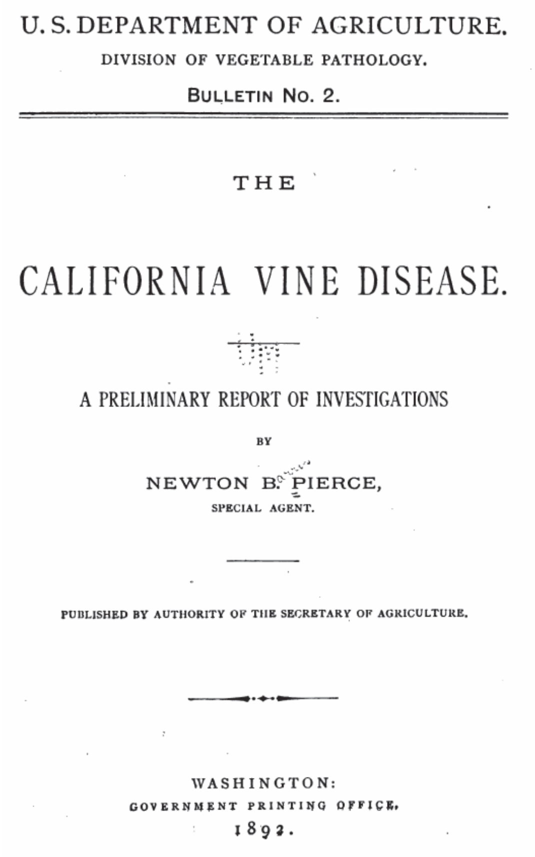
Mathieu Vanhove
Dept. Environmental Science, Policy, and Management
University of California, Berkeley

ACKNOWLEDGMENTS

- Lab Members
  - Anne Sicard
  - Michael A Voeltz
  - Jeffery Ezennia
  - Alexander Purcell
- Rodrigo Almeida
- …and the rest of the lab and collaborators

XF ACTORS
Xylella Fastidiosa Active Containment Through a multidisciplinary-Oriented Research Strategy
CALIFORNIA PD/GWSS BOARD
Partnership for Winegrape Pest Solutions
MARIE CURIE ACTIONS

“CALIFORNIA VINE DISEASE”: FIRST DETECTED IN ANAHEIM IN 1884



U. S. DEPARTMENT OF AGRICULTURE.
DIVISION OF VEGETABLE PATHOLOGY.
BULLETIN No. 2.

THE

CALIFORNIA VINE DISEASE.

A PRELIMINARY REPORT OF INVESTIGATIONS

BY

NEWTON B. PIERCE,
SPECIAL AGENT.

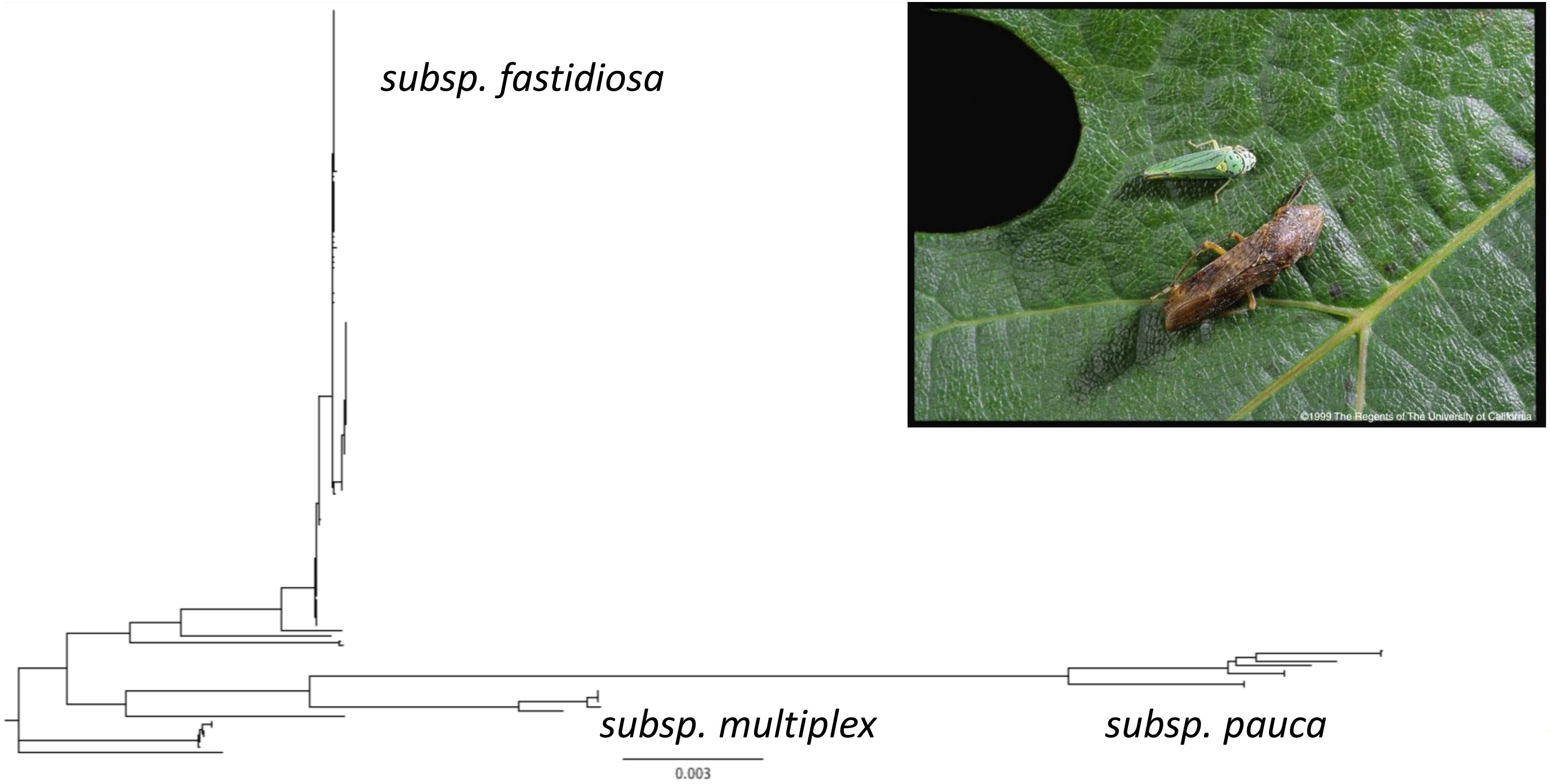
PUBLISHED BY AUTHORITY OF THE SECRETARY OF AGRICULTURE.

WASHINGTON:
GOVERNMENT PRINTING OFFICE.
1892.

Newton B. Pierce

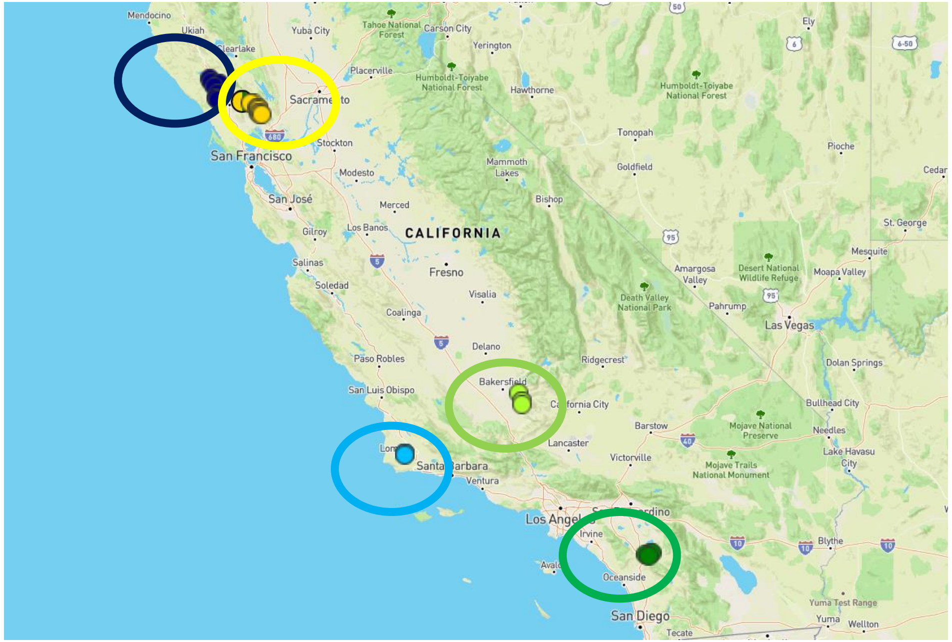
- Pierce’s disease costs California US$104 million per year
(Tumber et al. 2014. Calif. Agric. 68, 20-29)

Xf subsp. fastidiosa constitute a monophyletic group



Uncovering subsp. *fastidiosa* genetic diversity

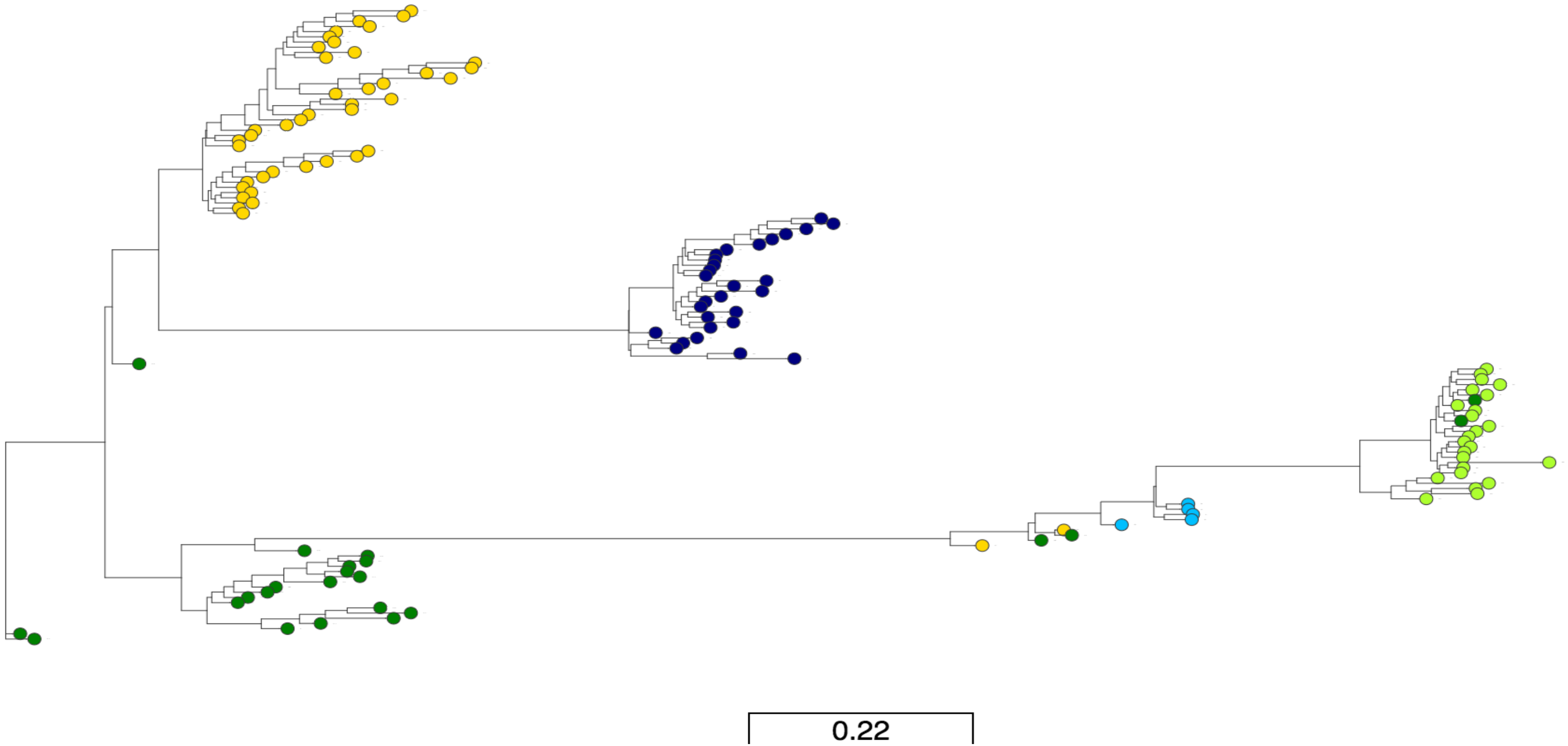
- Sequencing of 122 isolates
  - Santa Barbara: 5
  - Bakersfield: 25
  - Temecula: 23
  - Napa: 41
  - Sonoma: 28



RESEARCH QUESTIONS

- How much diversity is present in Californian grapevine-associated *X. fastidiosa*?
- Can we find a particular pathogen genotype-grapevine variety association?
- Is there any spatial pattern in pathogen association?
- Is there evidence of adaptation?
- What is the evolutionary history of *X. fastidiosa* subsp. *fastidiosa* in California?

POPULATIONS ARE GENETICALLY STRUCTURED



UNEXPECTED AMOUNT OF GENETIC DIVERSITY

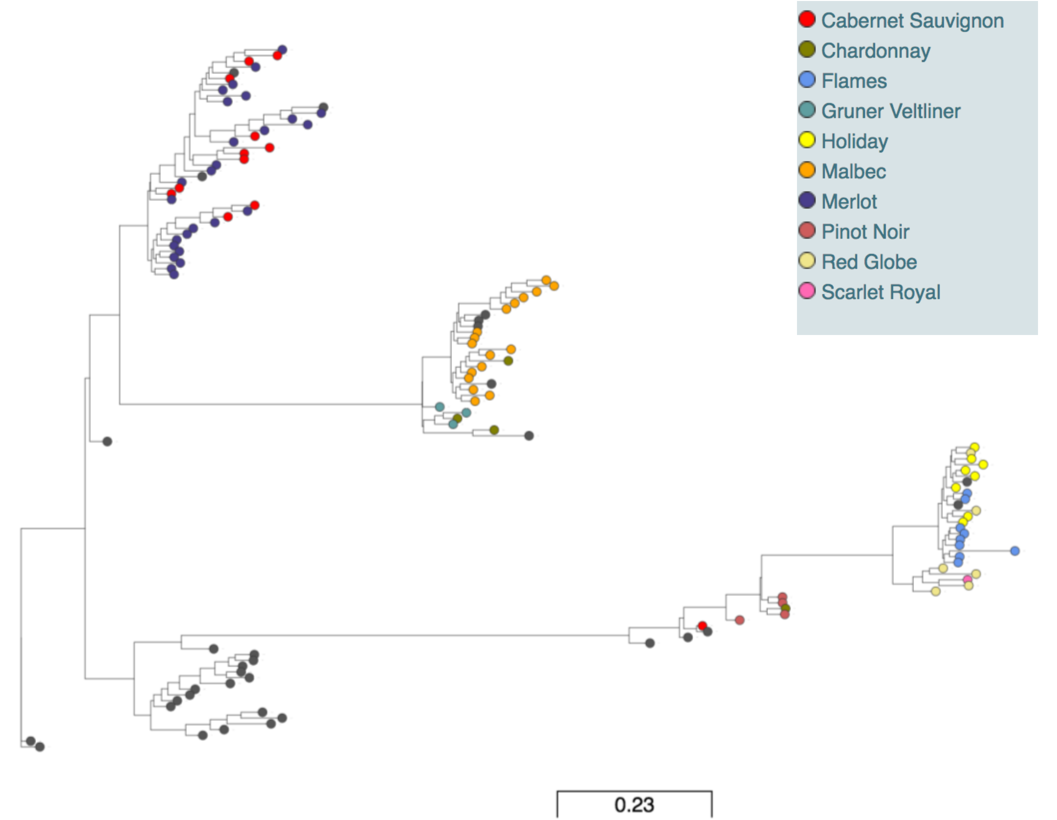
Populations	SNPs
California (n=123)	**5,218**
Cluster 1: Bakersfield (n=30)	1,561
Cluster 2: Temecula (n=18)	2,120
Cluster 3: Napa (n=40)	2,357
Cluster 4: Sonoma (n=30)	2,920

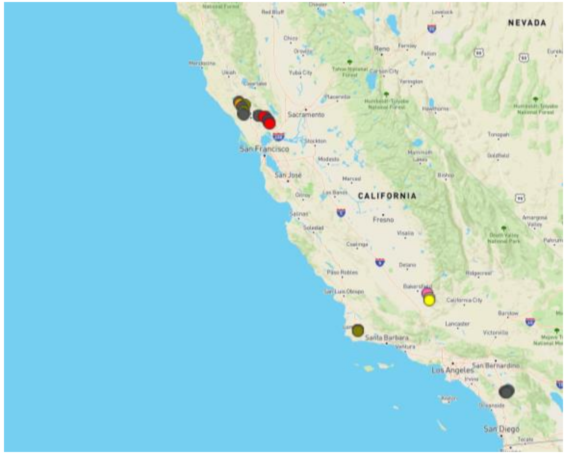
In Italy: average 250 SNPs

Anne Sicard

Postdoctoral researcher

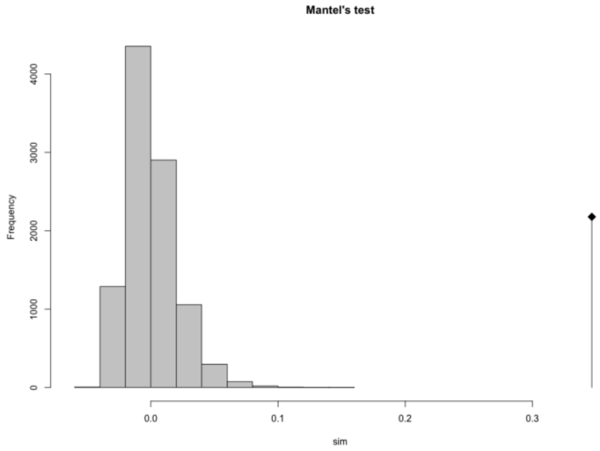
NO GENETIC CLUSTERING AMONG GRAPEVINE VARIETIES





EVIDENCE OF ISOLATION BY DISTANCE

Mantel test = 0.347 ($p < 0.001$)

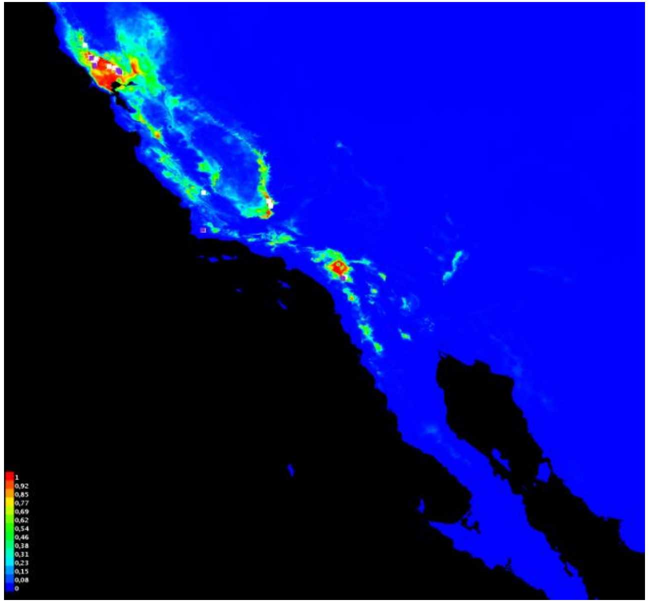


F_{ST}	Cluster 1: Bakersfield (n=30)	Cluster 2: Temecula (n=18)	Cluster 3: Napa (n=40)	Cluster 4: Sonoma (n=30)
Cluster 1: Bakersfield (n=30)	-	0,162	0,197	0,190
Cluster 2: Temecula (n=18)	0,162	-	0,131	0,151
Cluster 3: Napa (n=40)	0,197	0,131	-	0,139
Cluster 4: Sonoma (n=30)	0,190	0,151	0,139	-

PREDICTION OF ECOLOGICAL NICHE IN CALIFORNIA

- Ecological niche modelling
- Using 19 WorldClim layers

Variable	Contribution to the model	Permutation importance
Bio18: Precipitation of Warmest Quarter	27.3%	0.0%
Bio 19: Precipitation of Coldest Quarter	22.0%	12.2%
Elevation	15.6%	0.6%
Bio 6: Min Temperature of Coldest Month	10.1%	55.3%
Bio 3: Isothermality	6.4%	0.4%
Bio 16: Precipitation of Wettest Quarter	5.0%	2.9%



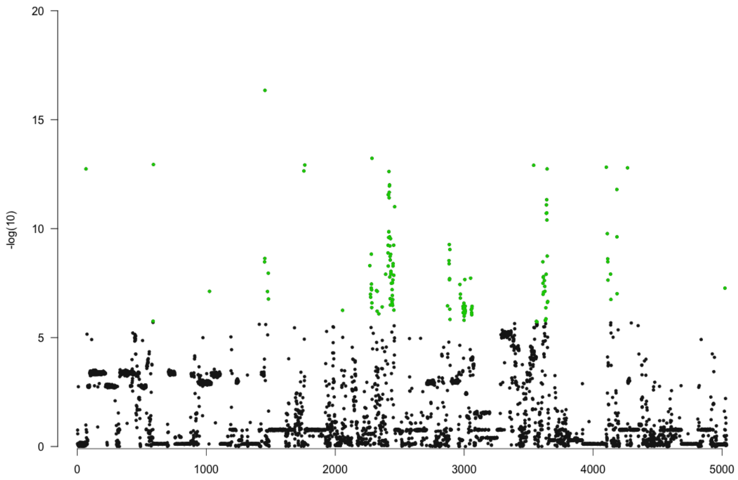
MaxEnt– Phillips *et al.* (2008)

IDENTIFYING ENVIRONMENTAL AND LANDSCAPE DRIVERS OF PATHOGEN SPREAD

- Example Bio01 – Annual Mean Temperature
  - 145 SNPs
    - 17 genes
    - 58 intergenic
    - 33 NYS mutations
    - 54 SYN mutation

Genes with known function:

PD_0058	fimbrial adhesin precursor
PD_0388	cardiolipin synthase
PD_0744	surface protein
PD_0952	conjugal transfer protein
PD_1023	truncated conjugal transfer protein
PD_1088	phage-related tail protein
PD_1094	phage-related contractile tail sheath protein
PD_1348	conjugal transfer protein
PD_1350	conjugal transfer protein
PD_2118	hemagglutinin-like secreted protein

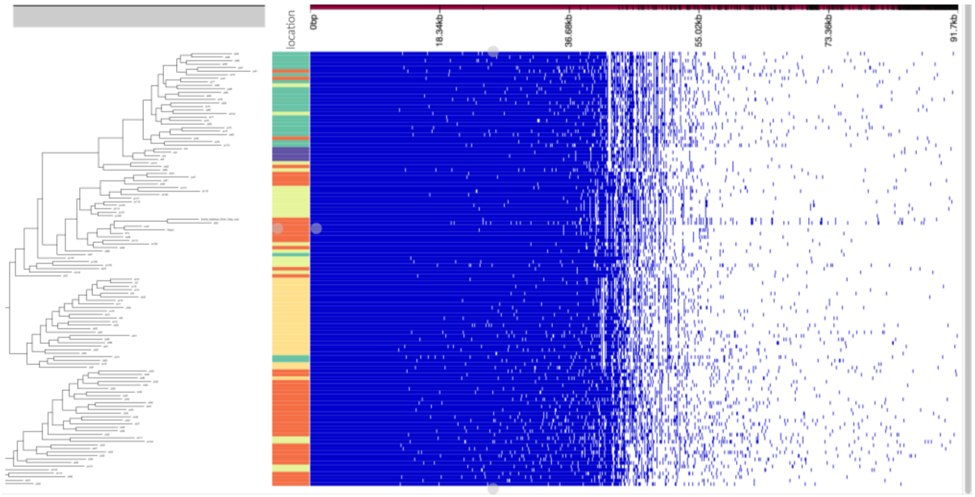


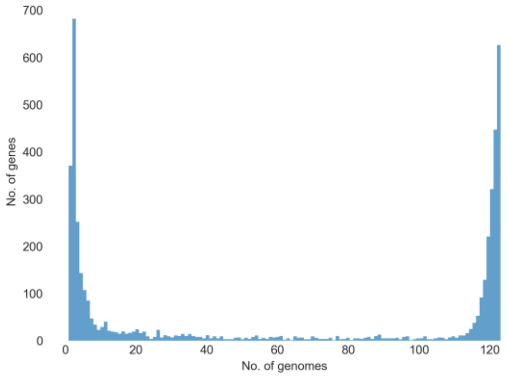
LFMM– Frichot *et al.* (2013)

ADAPTATION TO CLIMATE
A LANDSCAPE GENOMICS APPROACH

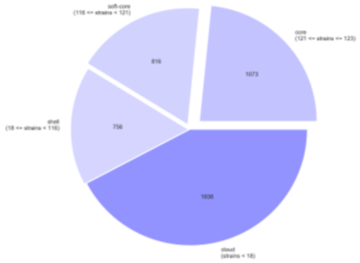
Variables	Name	Bonferroni correction		FDR	
		p = 0.01	p = 0.05	FDR = 0.001	FDR = 0.01
Elevation	Elevation	150	189	226	282
bio1	Annual Mean Temperature	145	248	432	1409
bio2	Mean Diurnal Range (Mean of monthly (max temp - min temp))	560	584	704	900
bio3	Isothermality (BIO2/BIO7) (* 100)	8	24	8	92
bio4	Temperature Seasonality (standard deviation *100)	59	76	415	663
bio5	Max Temperature of Warmest Month	121	153	654	1227
bio6	Min Temperature of Coldest Month	20	31	21	67
bio7	Temperature Annual Range (BIO5-BIO6)	93	120	540	956
bio8	Mean Temperature of Wettest Quarter	126	157	293	511
bio9	Mean Temperature of Driest Quarter	95	119	495	947
bio10	Mean Temperature of Warmest Quarter	98	129	556	966
bio11	Mean Temperature of Coldest Quarter	188	230	331	516
bio12	Annual Precipitation	x	x	x	x
bio13	Precipitation of Wettest Month	90	110	125	205
bio14	Precipitation of Driest Month	89	111	126	204
bio15	Precipitation Seasonality (Coefficient of Variation)	4	7	1	1
bio16	Precipitation of Wettest Quarter	91	337	646	800
bio17	Precipitation of Driest Quarter	86	109	125	216
bio18	Precipitation of Warmest Quarter	52	91	100	232
bio19	Precipitation of Coldest Quarter	82	105	115	212

KEY DIFFERENCES IN CORE AND ACCESSORY GENOME



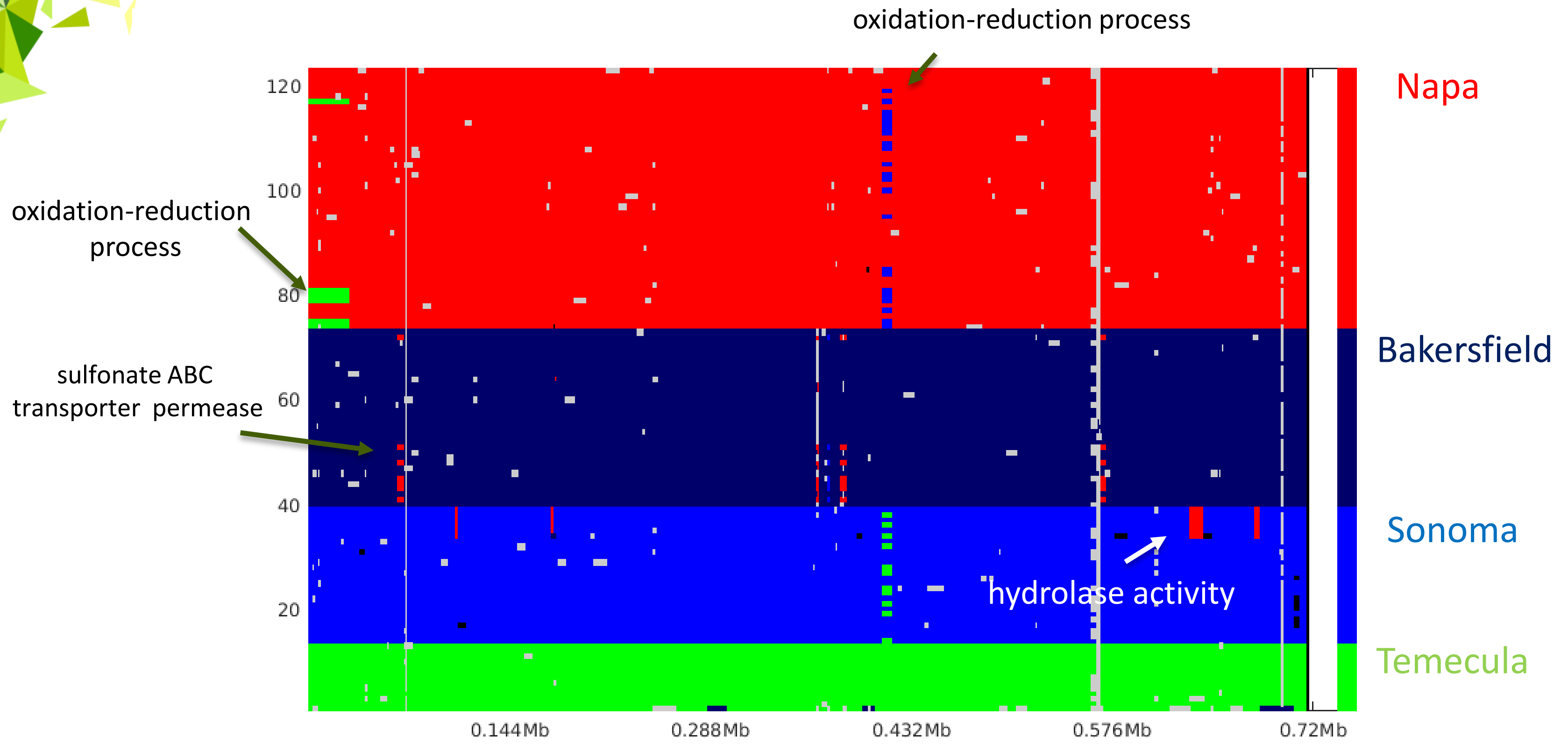


Isolates	Core	soft-core	shell	cloud
California (n=123)	1,073	816	756	1,938
Bakersfield (n=30)	1,200	559	755	730
Temecula (n=18)	1,644	0	902	171
Napa (n=45)	1,013	774	811	896
Sonoma (n=25)	1,473	472	549	249



Roary – Page *et al.* (2016)

EVIDENCE OF RECOMBINATION AMONG GENETIC CLUSTERS



fastGEAR – Mostowy *et al.* (2017)

Investigation signs of selection among clusters (MCDONALD–KREITMAN (MK) TEST)

Cluster	Description	Length	GO Names
Bakersfield	hypothetical protein XFEB_00363	519	C:integral component of membrane
Bakersfield	phosphomannomutase phosphoglucomutase	1404	P:carbohydrate metabolic process; F:intramolecular transferase activity, phosphotransferases
Bakersfield	conserved hypothetical protein	888	
Bakersfield	phage major capsid	1875	
Temecula, Napa	4-phosphopantetheinyl transferase	594	F:magnesium ion binding; F:holo-[acyl-carrier-protein] synthase activity
Napa, Sonoma	DNA polymerase III subunit delta	972	F:DNA-directed DNA polymerase activity; P:DNA biosynthetic process
Napa, Sonoma	cell surface	6159	P:pathogenesis; C:outer membrane
Napa, Sonoma	exodeoxyribonuclease III	789	P:DNA repair; F:nuclease activity
Temecula	hypothetical protein	1131	
Napa, Sonoma	DNA primase	2460	
Sonoma	carbamoyl-phosphate synthase small subunit	1125	P:'de novo' pyrimidine nucleobase biosynthetic process; P:glutamine metabolic process; F:carbamoyl-phosphate synthase (glutamine-hydrolyzing) activity
Sonoma	multifunctional CCA tRNA nucleotidyl transferase 2 3 -cyclic phosphodiesterase 2 nucleotidase phosphatase	1251	F:tRNA adenylyltransferase activity; F:ATP binding; F:RNA binding; P:tRNA 3'-terminal CCA addition
Sonoma	lytic murein transglycosylase	2133	C:periplasmic space; F:hydrolase activity, hydrolyzing O-glycosyl compounds

GENETIC EPIDEMIOLOGY AND DISEASE MANAGEMENT

- Reconstruct the evolutionary history of *Xylella fastidiosa (Xf)* subs. *fastidiosa*
- Identify particular genotype or route of infection
- Predict a potential spillover into susceptible host populations as well as uninfected regions
- A deeper understanding into the biology of the disease
- Identify signs of adaptation and response to biotic and abiotic factors

THANK YOU FOR LISTENING