



Range wide population genetics update

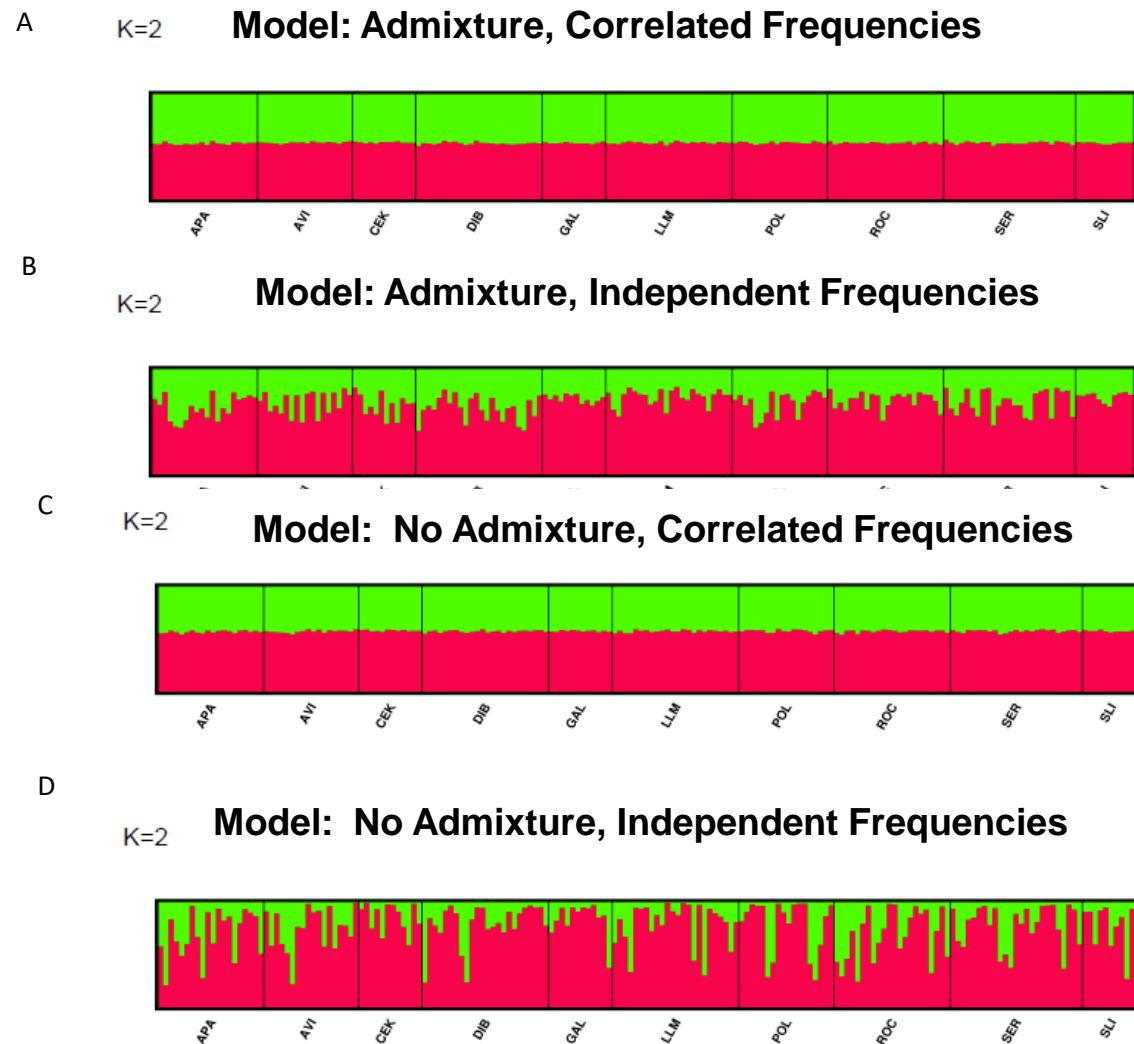
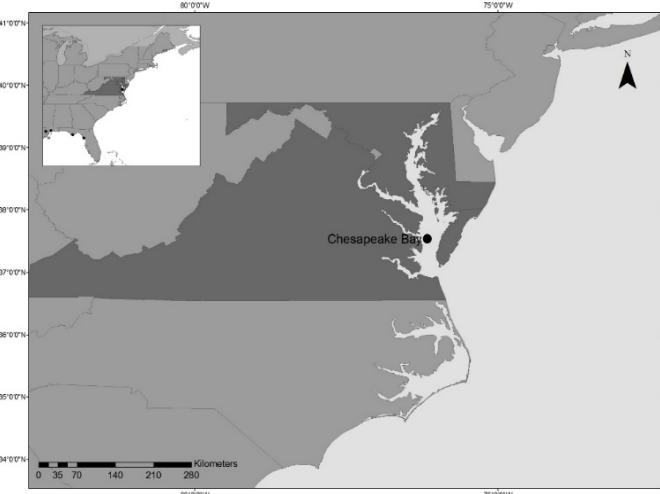
Luis Hurtado, Isabel Caballero, Mariana Mateos



Population genetics and historical demographic inferences of the blue crab *Callinectes sapidus* in the US based on microsatellites

Danielle Macedo¹, Isabel Caballero¹, Mariana Mateos¹,
Raphael Leblois², Shelby McCay¹ and Luis A. Hurtado¹

No genetic structure is detected in the US with microsatellites



Recommendation

- Detection of population structure in marine organisms characterized by extremely large populations and high dispersal potential, and/or with recently diverged populations may be difficult using neutral markers, such as microsatellites, because genetic drift is weak in large populations, and even low levels of gene flow can homogenize populations.
- It is important, then, to examine non-neutral markers, which are more sensitive for detecting population structure in this case.

RAD-seq methodology

Library & Sequencing

- Illumina HiSeq 2500
- 125 bp reads (run 1 & 2), 160 bp reads (run 3)

Assembly

- *De novo*
- Stacks v.1.48 (denovo map, rxstacks)

SNPs Discovery & Filtering

- Stacks v.1.48 (populations)
- VCFtools for filtering and thinning

Outlier analyses

- Arlequin
- PCAdapt
- Bayescan

Genetic Structure

- F_{ST}
- Bayesian Clustering (Structure)
- DAPC (Discriminant Analyses of Principal Components)

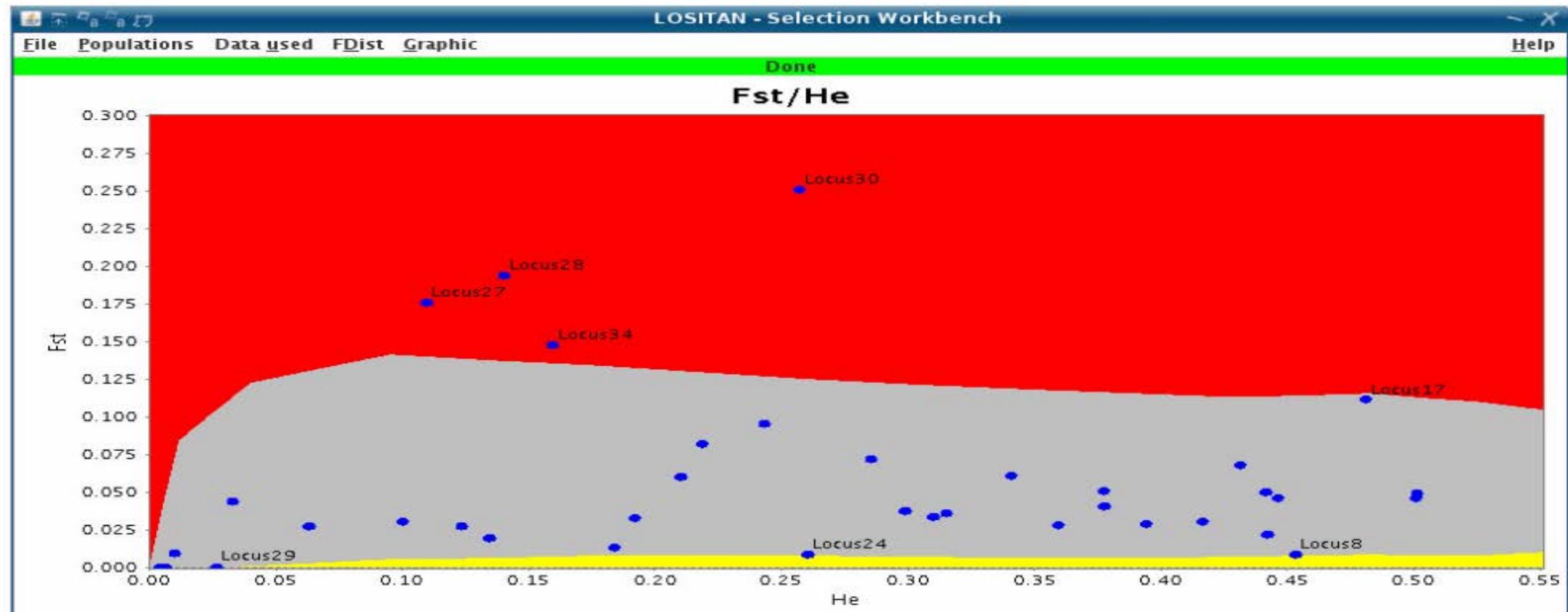
Inferring genetic structure within runs

Putative
neutral
markers

Used to infer genetic
differentiation (Fst)

Putative
selective
markers

Can be more informative
in cases where
populations are (1)
extremely large, (2) there
is high dispersal and (3)
recent divergence

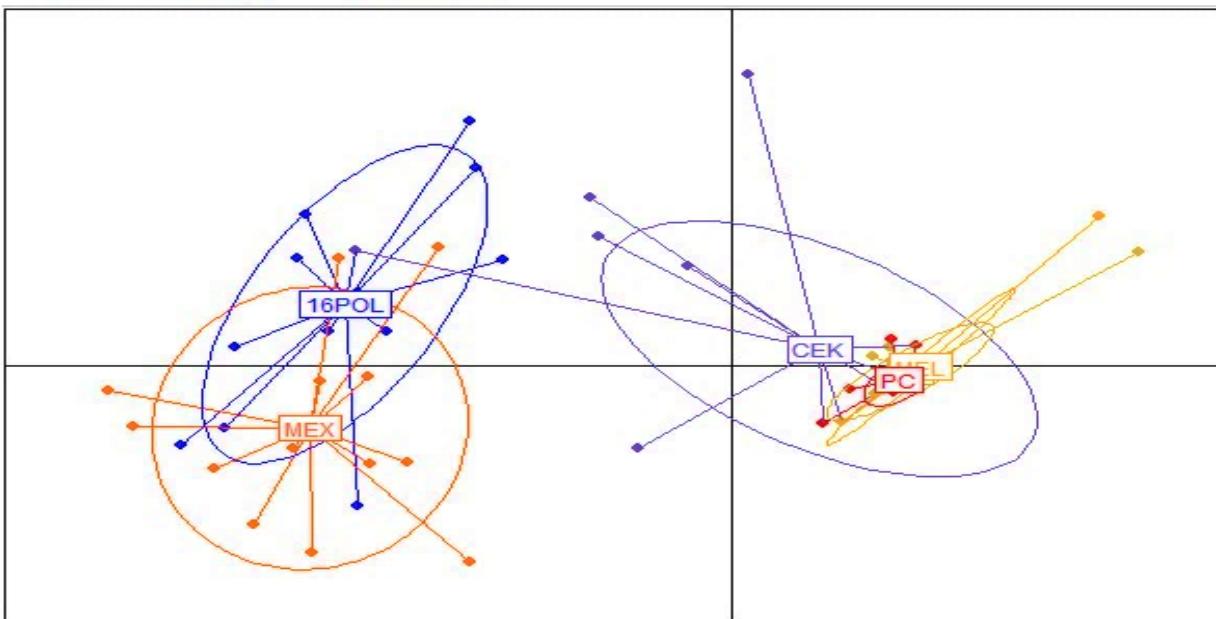
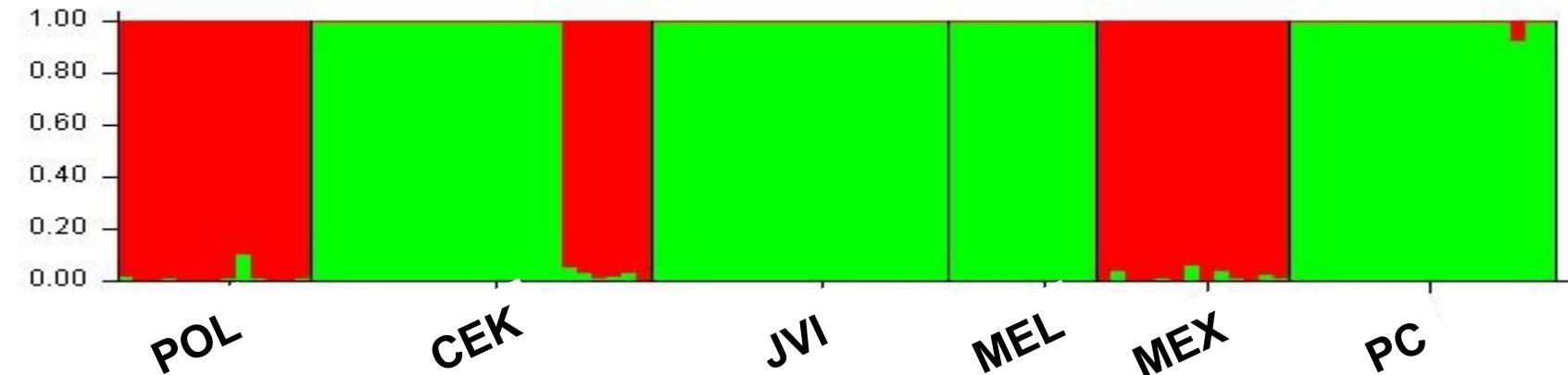


Outlier detection



Two genetic clusters were detected with STRUCTURE and DAPC using 16, 70 and 86 outliers

Inferences on genetic structure with putative selective markers (run 3)



Both methods indicate genetic structure between sampled populations of the West Gulf of Mexico and Florida

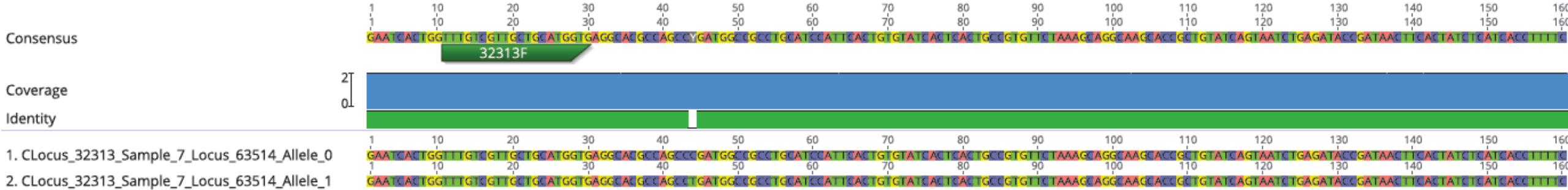
16 OUTLIERS GOM WEST/EAST DIFFERENCES

sample	6083_5	32326_15	14805_143	32211_43	77205_8	36244_120	33381_55	98024_47	48302_96	31238_20	10319_62	22380_75	38416_75	45525_143	63081_23	71006_15
16POL_16POL01	T	G	A	T	G	T	A	G	T	A	C	T	C	C	G	T
16POL_16POL01	C	A	A	T	C	C	T	A	T	A	C	T	T	A	T	T
16POL_16POL02	T	G	A	T	G	T	A	G	T	A	C	T	C	C	?	T
16POL_16POL02	C	A	G	C	C	C	T	A	T	A	C	T	T	?	C	?
16POL_16POL04	T	G	?	?	G	T	A	G	T	A	C	T	T	?	C	?
16POL_16POL04	C	A	?	?	C	C	T	A	G	T	C	T	C	C	G	?
16POL_16POL05	T	G	A	T	G	T	A	G	T	A	C	T	C	C	G	T
16POL_16POL05	C	A	G	C	C	C	T	G	T	A	C	T	C	C	G	T
16POL_16POL06	T	G	A	T	G	T	A	G	T	A	C	T	C	C	G	T
16POL_16POL06	C	A	G	C	C	C	T	A	T	A	C	T	G	T	A	T
16POL_16POL07	T	G	A	?	G	T	A	G	T	A	C	T	C	C	G	T
16POL_16POL07	C	A	?	?	C	C	T	A	G	A	C	T	T	A	C	?
16POL_16POL08	T	G	A	T	G	T	A	G	T	A	C	T	C	C	G	T
16POL_16POL08	C	A	?	?	C	C	T	T	G	A	C	T	G	T	A	C
16POL_16POL09	?	?	A	T	G	T	A	G	T	A	C	T	C	C	G	T
16POL_16POL09	?	?	G	C	C	C	T	A	T	A	C	C	G	T	A	C
16POL_16POL10	T	G	A	T	G	T	A	G	T	A	C	T	C	C	G	?
16POL_16POL10	T	G	A	T	G	T	A	G	T	A	C	T	G	C	A	?
16POL_16POL11	T	G	A	T	?	?	A	G	T	A	C	T	C	C	?	?
16POL_16POL11	C	A	G	C	?	?	T	G	G	A	C	T	T	T	?	?
16POL_16POL12	T	G	A	T	G	T	A	G	T	A	C	T	C	C	G	T
16POL_16POL12	C	A	G	C	C	C	T	A	T	A	C	T	T	A	C	?
16POL_16POL14	T	G	A	T	G	T	A	G	T	A	C	T	C	C	G	T
16POL_16POL14	C	A	G	C	C	C	T	A	T	A	C	T	G	T	A	C
16POL_16POL17	T	G	A	T	G	T	A	G	T	A	C	T	C	C	G	T
16POL_16POL17	C	A	G	C	C	C	T	G	G	A	C	T	T	A	C	?
MEX_MEX01	T	G	A	T	G	T	A	G	T	A	C	T	C	C	G	T
MEX_MEX01	C	A	G	C	C	C	T	G	G	A	C	T	T	A	C	?
MEX_MEX02	T	G	A	T	G	T	A	G	T	A	C	T	C	C	G	T
MEX_MEX02	C	A	?	?	T	G	T	A	A	T	C	T	C	C	G	T
MEX_MEX03	T	G	A	T	?	?	A	G	T	A	C	T	C	C	G	T
MEX_MEX03	C	A	G	C	?	?	T	A	G	A	C	T	G	T	A	C
MEX_MEX05	T	G	?	?	G	T	A	G	T	A	C	T	C	C	G	T
MEX_MEX05	C	A	?	?	C	C	T	A	G	A	C	T	T	A	C	?
MEX_MEX06	?	?	?	?	G	T	A	G	T	A	C	T	C	C	G	T
MEX_MEX06	?	?	?	?	C	C	T	A	G	A	C	T	T	A	T	?
MEX_MEX07	T	G	A	T	G	T	A	G	T	A	C	T	C	C	G	T
MEX_MEX07	C	A	?	?	T	C	C	T	A	G	A	T	T	A	C	?
MEX_MEX08	T	G	A	T	G	T	A	G	?	A	C	T	C	C	G	T
MEX_MEX08	T	G	G	T	G	T	T	A	?	A	C	T	T	A	C	?
MEX_MEX09	T	G	A	T	?	?	A	?	T	A	C	T	C	C	G	T
MEX_MEX09	C	A	G	C	?	?	T	?	G	A	C	T	G	T	A	C
MEX_MEX10	T	G	A	T	G	T	A	G	T	A	C	T	C	C	G	T
MEX_MEX10	T	G	A	T	C	C	A	A	T	A	C	T	T	A	T	?
MEX_MEX11	T	G	A	T	T	G	T	A	G	?	A	T	C	C	G	?
MEX_MEX11	C	A	G	C	G	T	T	A	?	A	C	T	T	A	C	?
MEX_MEX13	T	G	A	T	G	T	?	G	T	A	C	T	C	C	G	T
MEX_MEX13	C	A	G	C	G	C	?	G	G	A	C	T	T	A	C	?
MEX_MEX14	T	G	A	T	G	T	A	G	T	A	C	T	C	C	G	T
MEX_MEX14	C	A	A	T	C	C	T	G	G	A	C	T	T	A	C	?
MEX_MEX15	T	G	A	T	G	T	A	G	T	A	C	?	C	C	G	T
MEX_MEX15	C	A	A	T	C	C	A	A	G	A	C	?	T	T	A	C
CEK_CEK23	?	G	A	T	T	G	T	A	G	T	C	T	C	C	G	T
CEK_CEK23	?	G	A	T	T	C	C	T	A	T	G	C	T	T	A	C
CEK_CEK24	T	G	A	T	G	T	T	A	?	T	A	C	T	C	G	T
CEK_CEK24	C	A	A	T	T	C	C	T	?	T	G	?	T	A	T	?
CEK_CEK25	T	G	A	T	G	T	A	G	T	A	C	?	T	C	G	T
CEK_CEK25	C	A	G	C	G	G	T	T	G	T	C	?	G	?	A	T
CEK_CEK26	?	?	A	T	G	T	A	?	T	A	C	?	T	T	A	C
CEK_CEK26	?	G	C	C	C	T	?	T	A	C	G	T	C	G	T	?
CEK_CEK27	T	G	A	T	?	?	A	G	T	T	A	C	?	C	C	G
CEK_CEK27	C	A	A	T	?	?	A	G	T	A	C	?	T	T	A	C
CEK_CEK28	T	G	A	T	G	T	A	G	T	A	C	?	T	C	G	T
CEK_CEK28	C	A	G	C	C	C	T	A	T	A	C	T	G	T	A	C

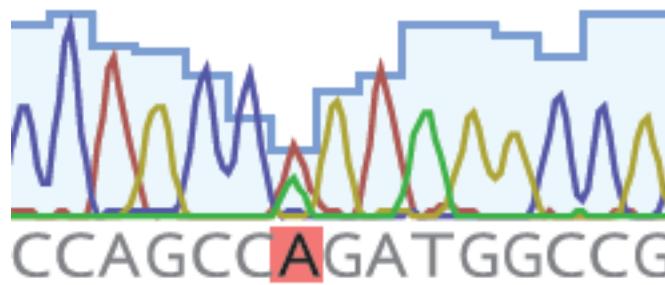
Port Lavaca, Sisal (Yucatán), Cedar Key (6)

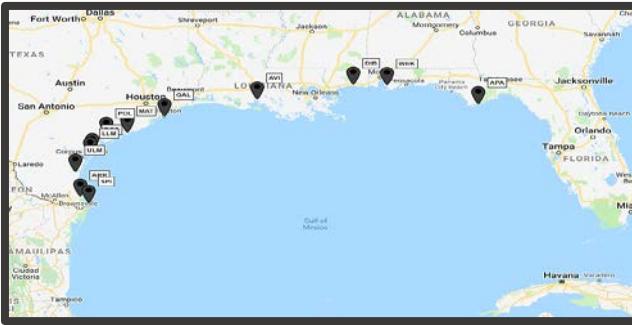
Florida, including Cedar Key (17)

PRIMERS FOR SEQUENCING WERE DEVELOPED TO IDENTIFY POLYMORPHIC SITES



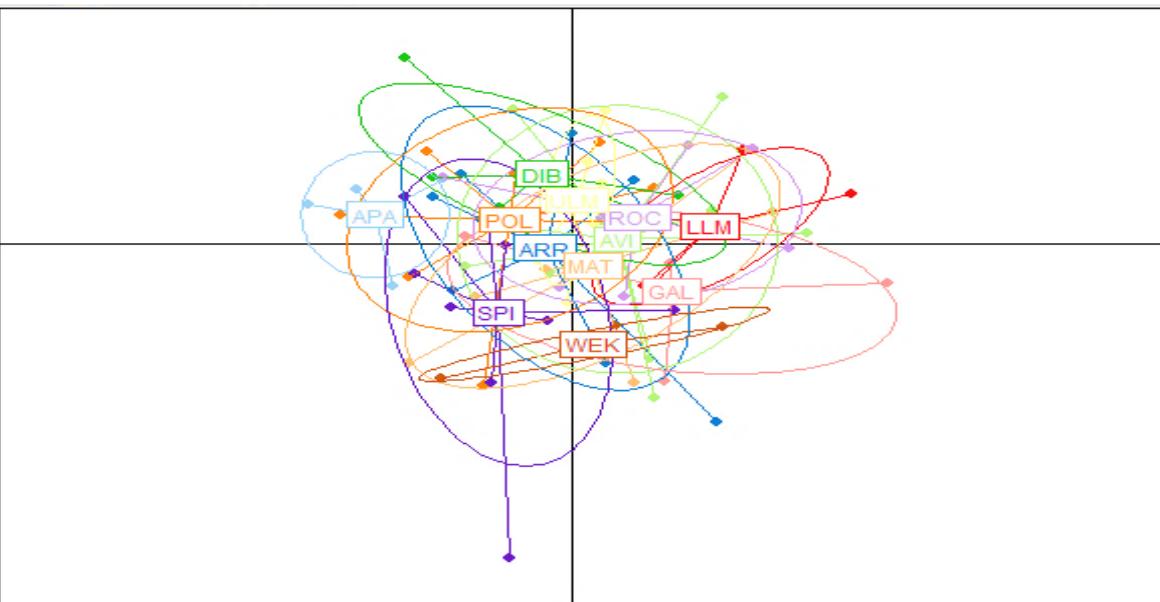
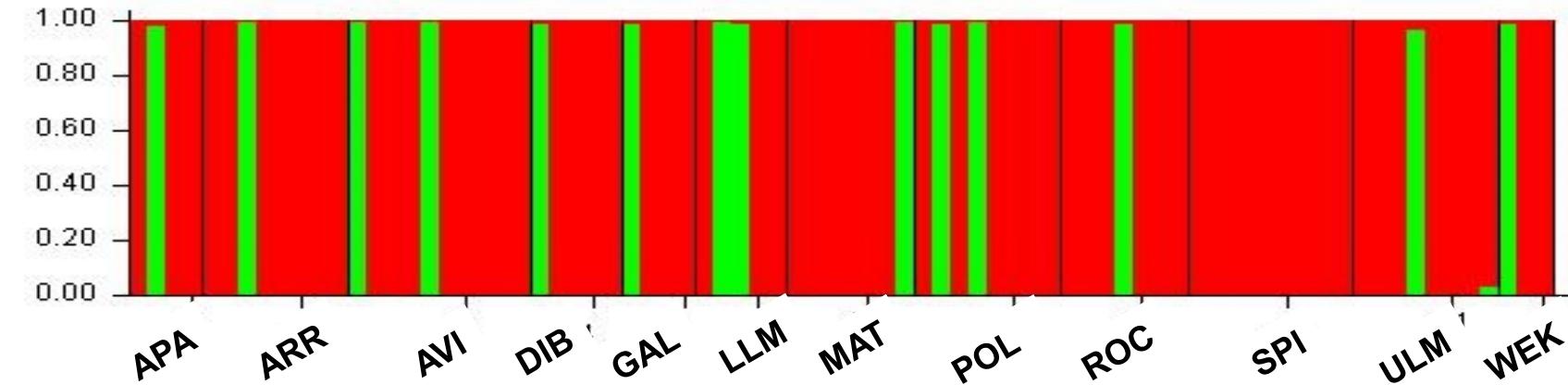
Sanger sequencing allows heterozygote detection





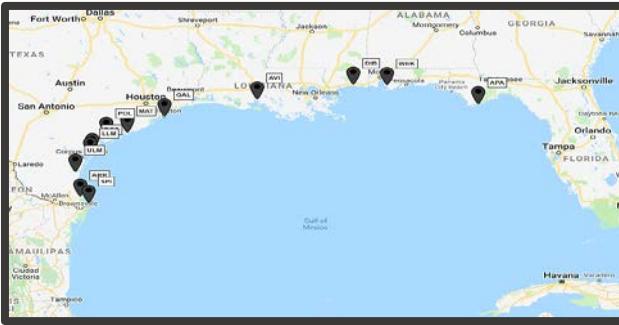
Inferences on genetic structure with putative neutral markers (run 1)

- 12 localities; 78 samples
- 2014 and 2015
- Very low and no significant pairwise Fst
- No genetic structure detected with STRUCTURE and DAPC



Same patterns were observed using 29 putative selective markers

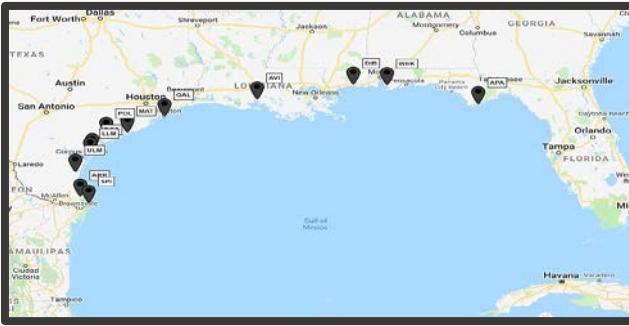
No evidence of genetic structure in sampled populations of the Gulf of Mexico



- **13 localities; 67 samples**
- **2014 and 2015**

Inferences on genetic structure by region (run 1)

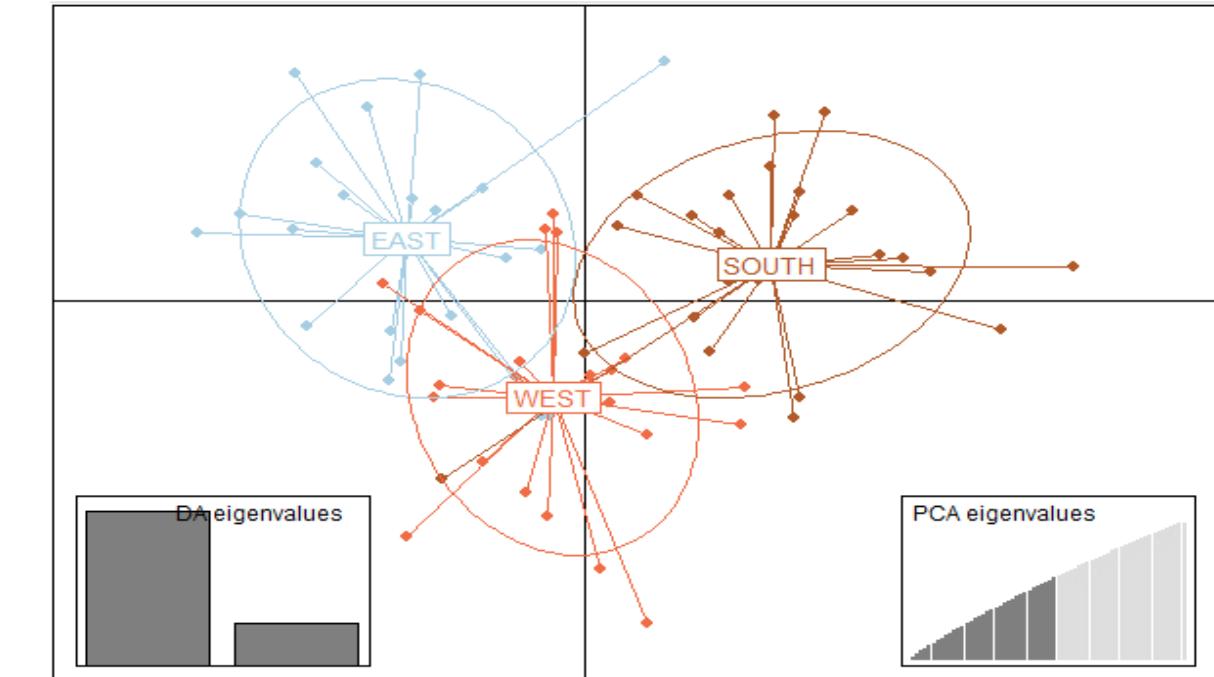
Apalachicola (FL)	APA	3	
Weeks Bay (AL)	WEK	3	
D'Iberville (MS)	DIB	4	
Slidell (LA)	SLI	2	
Avery Island (LA)	AVI	9	EAST = 21
Galveston (TX)	GAL	3	
Matagorda (TX)	MAT	6	
Port Lavaca (TX)	POL	8	
Rockport (TX)	ROC	5	WEST = 22
Upper Laguna Madre (TX)	ULM	7	
Arroyo City (TX)	ARR	5	
Lower Laguna Madre (TX)	LLM	4	
South Padre Island (TX)	SPI	8	SOUTH = 24
TOTAL		67	



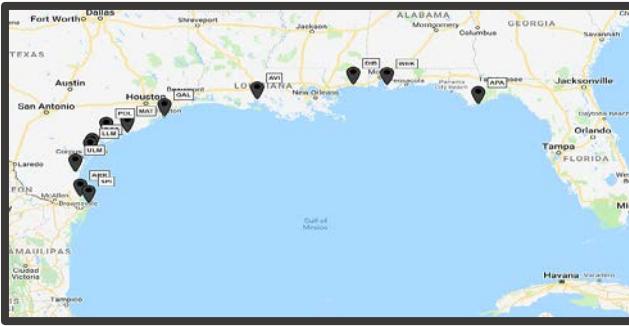
Inferences on genetic structure with putative neutral markers by region (run 1)

	EAST	WEST
EAST		
WEST	-0.00277	
SOUTH	-0.00139	-0.00277

Small, non-significant Fst values



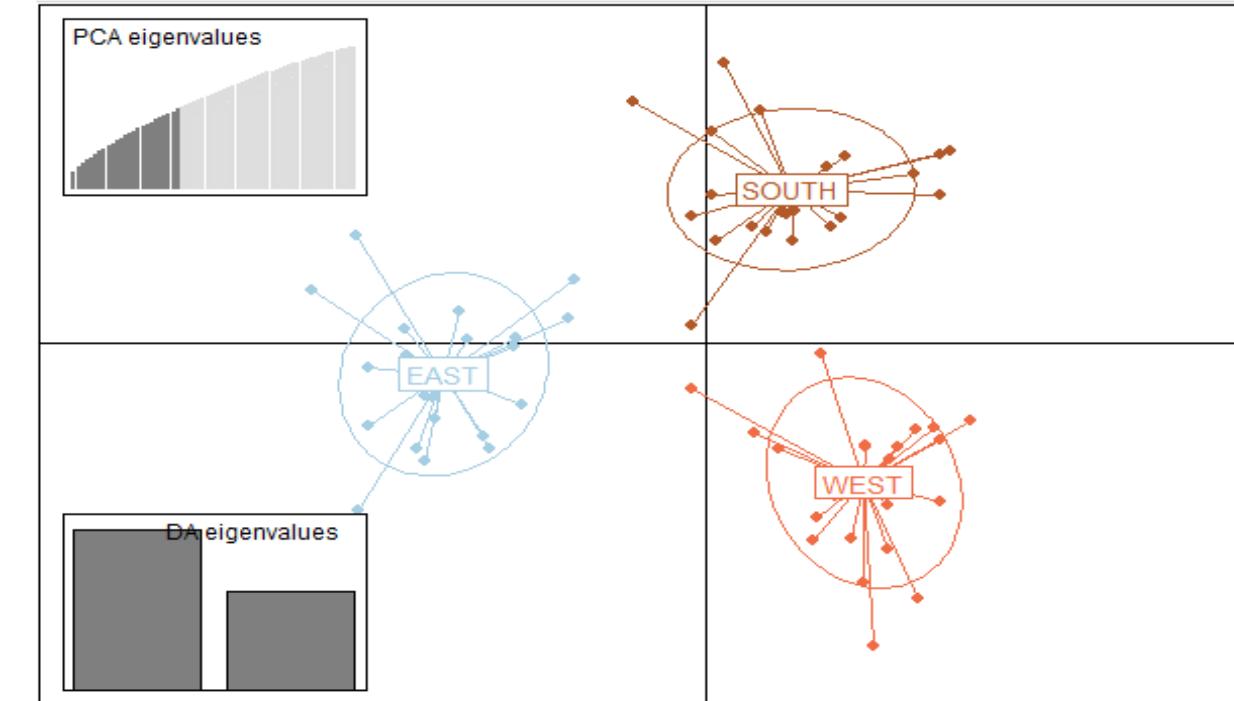
9,409 loci (Stacks); 35 PCA; 2 DA



Inferences on genetic structure with outliers markers by region (run 1)

	EAST	WEST	SOUTH
EAST			
WEST	0.01073		
SOUTH	0.01022	0.00639	

Larger, significant Fst values

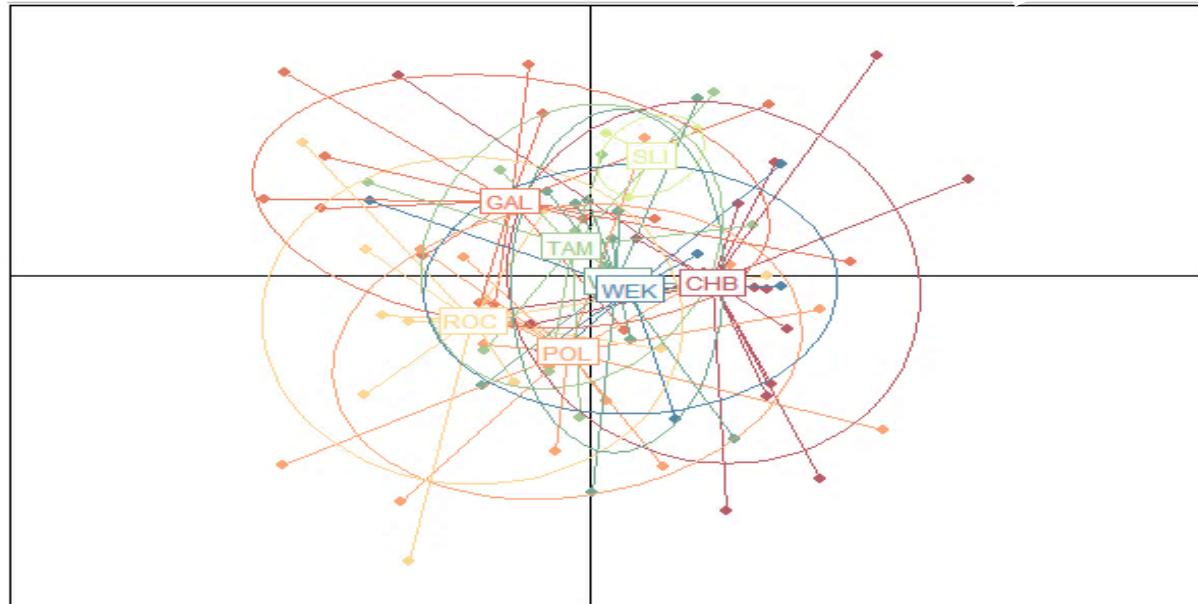
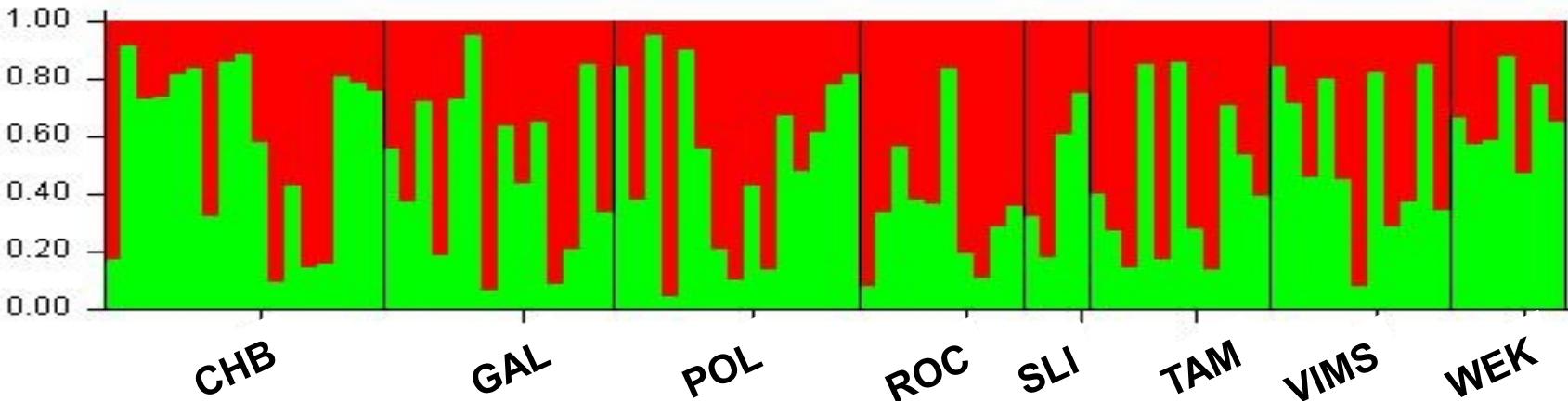


1,396 loci (Stacks); 25 PCA; 2 DA

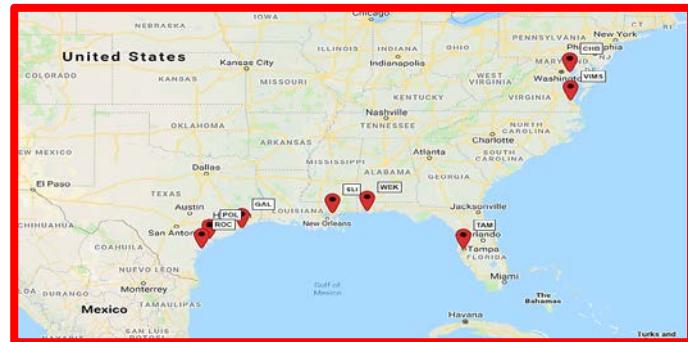


Inferences on genetic structure with putative neutral markers (run 2)

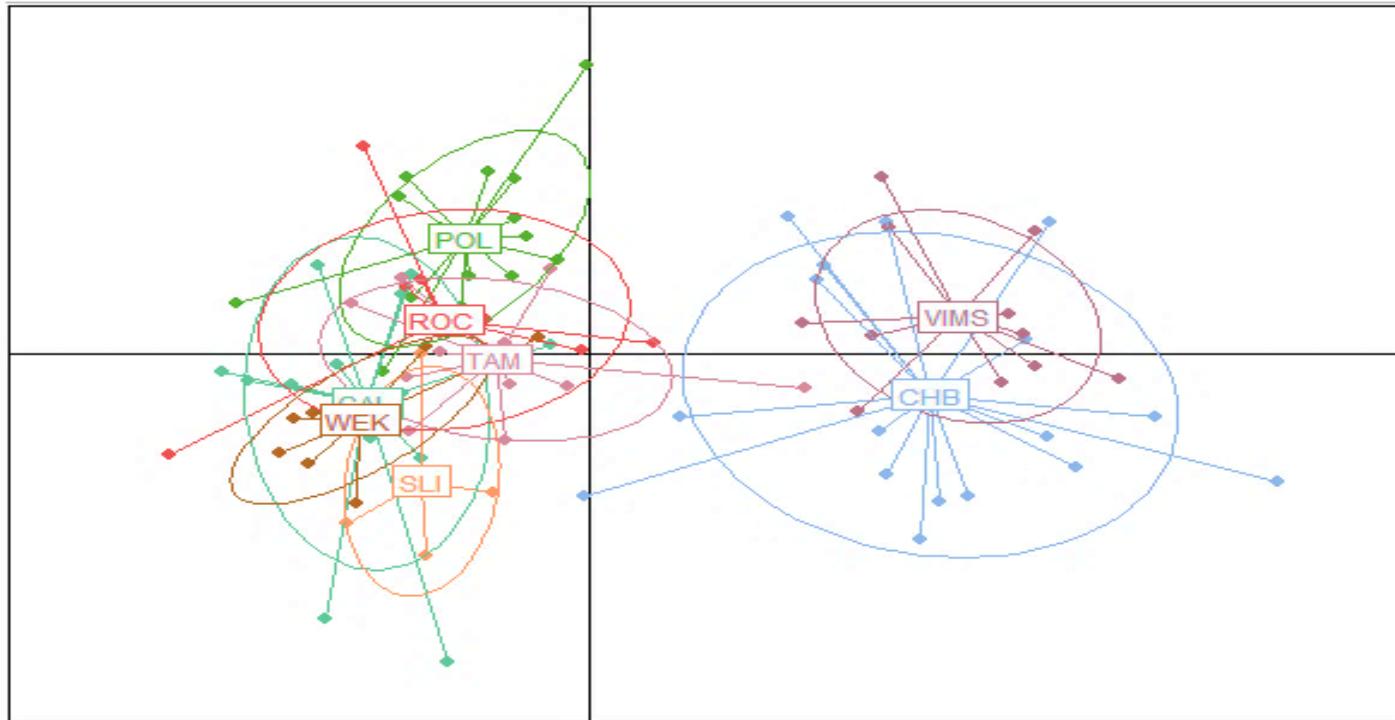
- 8 localities; 89 samples
- 2014 and 2015
- No significant pairwise Fst
- No genetic structure detected with STRUCTURE and DAPC



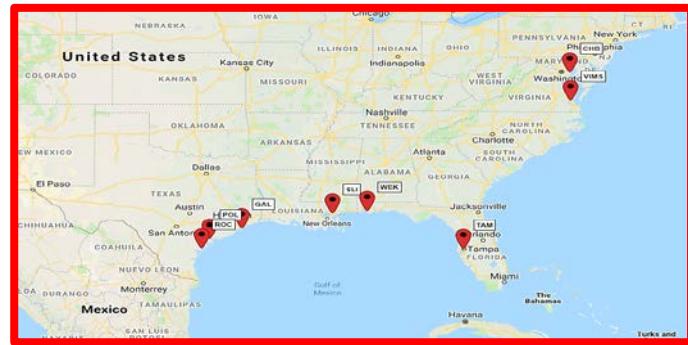
No genetic structure in sampled populations of the Gulf of Mexico and the Atlantic



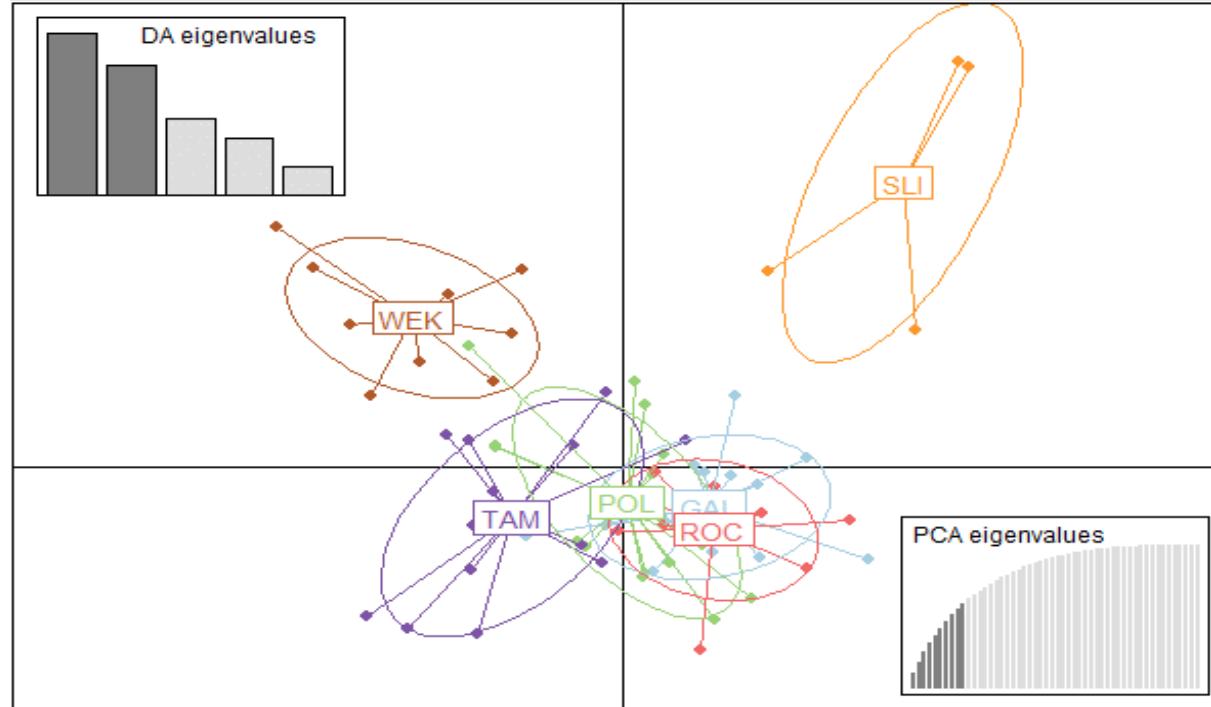
Inferences on genetic structure with putative selective markers (run 2)



However, DAPC with 54 putative selective markers shows evidence of genetic structure in sampled populations of the Gulf of Mexico and the Atlantic

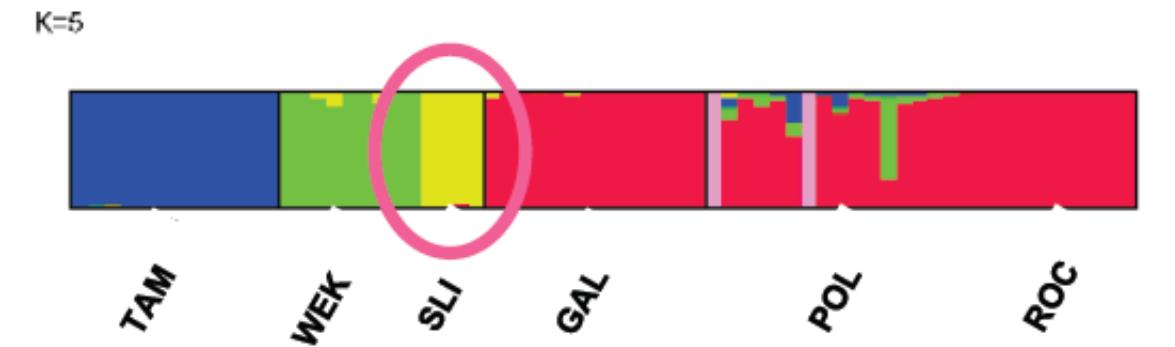


Inferences on genetic structure with putative selective markers GOM (run 2)

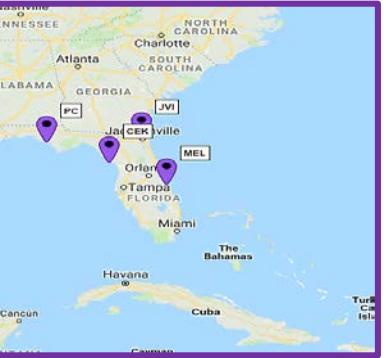


Tampa (FL)	TAM	13
Weeks Bay (AL)	WEK	9
Slidell (LA)	SLI	4
Galveston (TX)	GAL	14
Port Lavaca (TX)	POL	17
Rockport (TX)	ROC	10

52 outliers

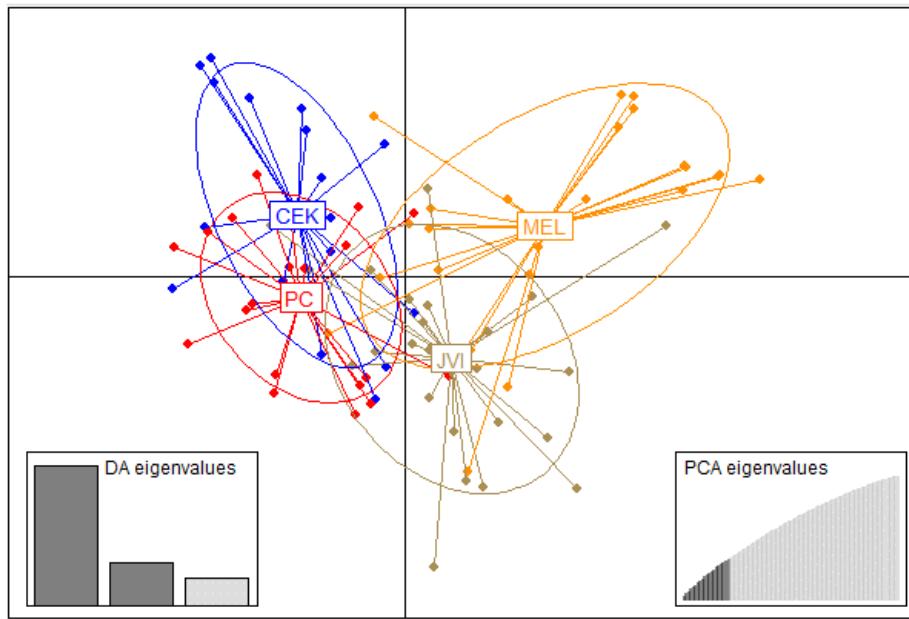


STRUCTURE Non-admixture correlated LOCPRIOR



Inferences on genetic structure with putative neutral markers (run 3)

	JVI	MEL	CEK
JVI			
MEL	-0.00063		
CEK	-0.00002	0.0089	
PC	0.0004	0.00905	-0.0039

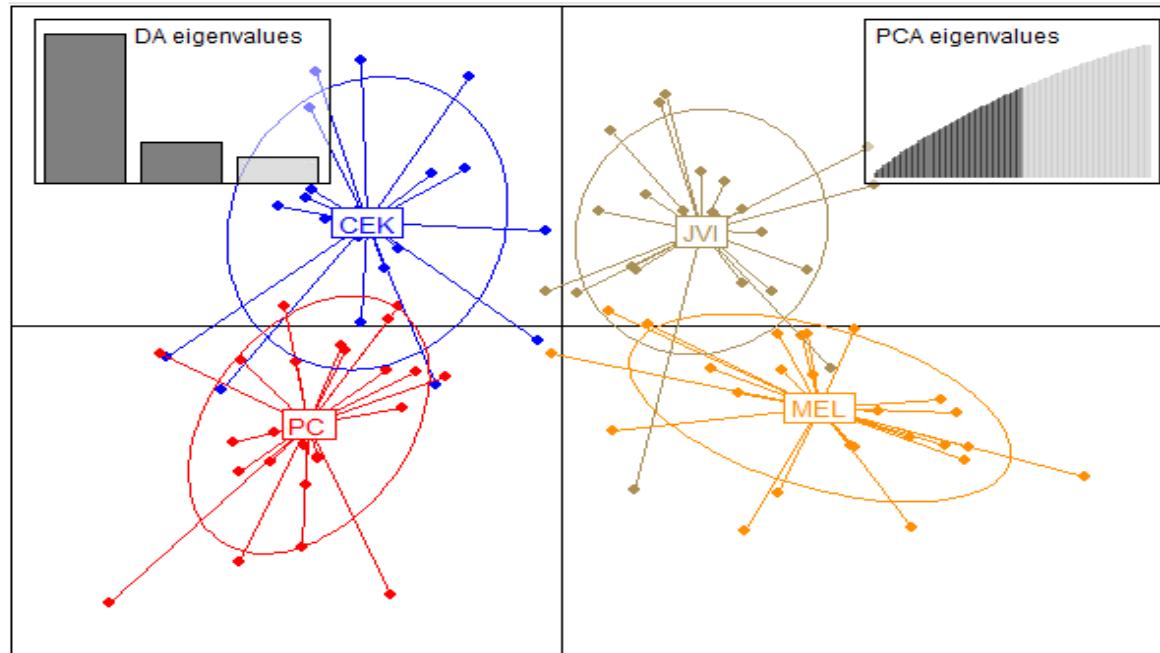


2,276 loci (Stacks); 20 PCA; 3 DA



Inferences on genetic structure with putative neutral markers (run 3)

	JVI	MEL	CEK
JVI			
MEL	-0.00066		
CEK	-0.00079	0.00928	
PC	-0.00036	0.00881	-0.00425

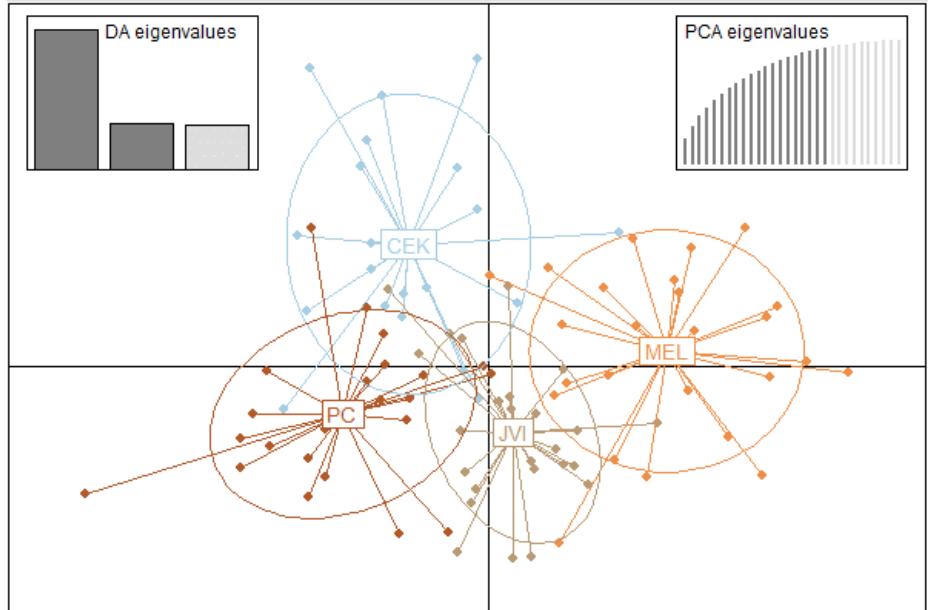


2,352 loci (VCFtools); 50 PCA; 3 DA



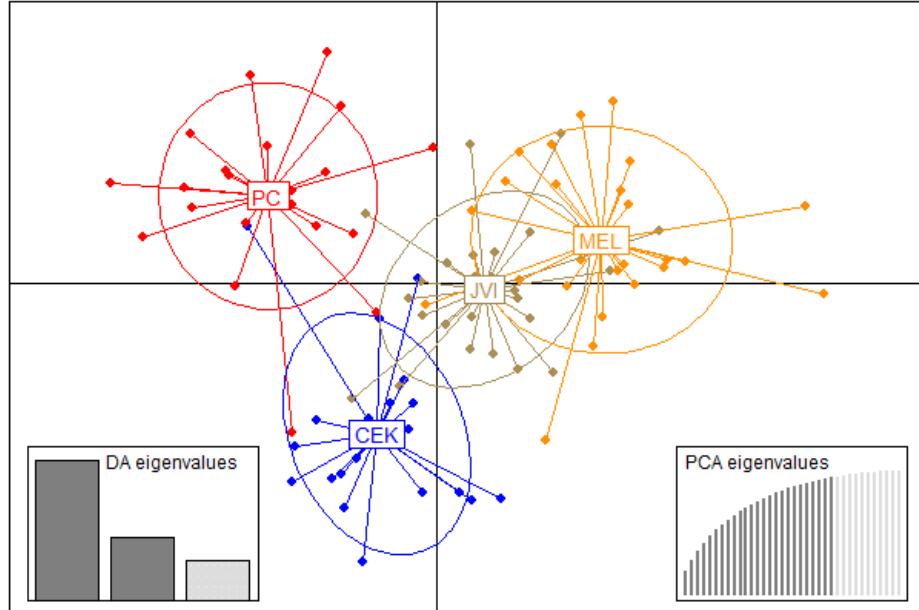
Inferences on genetic structure with putative selective markers (run 3)

	JVI	MEL	CEK
JVI			
MEL	0.06276		
CEK	0.07219	0.1356	
PC	0.04259	0.11549	0.01497



30 loci (Stacks); 20 PCA; 3 DA

	JVI	MEL	CEK
JVI			
MEL	0.05122		
CEK	0.08506	0.14099	
PC	0.05499	0.12489	0.0233



36 loci (VCFtools); 25 PCA; 3 DA

Conclusions and recommendations

- There is genetic structure in the US.
- This structure is detected mainly with putative non-neutral markers.

Acknowledgements



Smithsonian Environmental
Research Center



**INSTITUTE FOR GENOME
SCIENCES AND SOCIETY**
TEXAS A&M UNIVERSITY

TEXAS A&M HIGH PERFORMANCE RESEARCH COMPUTING



Genomics and Bioinformatics Service

Providing Genomics and Bioinformatics Services to the Texas A&M System, Texas, and the World



QUESTIONS?