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

Discovery & Filtering
- Arlequin
- PCAdapt
- Bayescan

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

Genetic Structure
Inferring genetic structure within runs

**Putative neutral markers**

Used to infer genetic differentiation ($F_{st}$)

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

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

Both methods indicate genetic structure between sampled populations of the West Gulf of Mexico and Florida.
<table>
<thead>
<tr>
<th>Sample</th>
<th>16 OUTLIERS GOM WEST/EAST DIFFERENCES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Port Lavaca, Sisal (Yucatán), Cedar Key (6)</td>
</tr>
</tbody>
</table>
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
Inferences on genetic structure by region (run 1)

- 13 localities; 67 samples
- 2014 and 2015

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)

<table>
<thead>
<tr>
<th></th>
<th>EAST</th>
<th>WEST</th>
<th>SOUTH</th>
</tr>
</thead>
<tbody>
<tr>
<td>EAST</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WEST</td>
<td>0.01073</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOUTH</td>
<td>0.01022</td>
<td>0.00639</td>
<td></td>
</tr>
</tbody>
</table>

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)

52 outliers

<table>
<thead>
<tr>
<th>Location</th>
<th>Code</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tampa (FL)</td>
<td>TAM</td>
<td>13</td>
</tr>
<tr>
<td>Weeks Bay (AL)</td>
<td>WEK</td>
<td>9</td>
</tr>
<tr>
<td>Slidell (LA)</td>
<td>SLI</td>
<td>4</td>
</tr>
<tr>
<td>Galveston (TX)</td>
<td>GAL</td>
<td>14</td>
</tr>
<tr>
<td>Port Lavaca (TX)</td>
<td>POL</td>
<td>17</td>
</tr>
<tr>
<td>Rockport (TX)</td>
<td>ROC</td>
<td>10</td>
</tr>
</tbody>
</table>

STRUCTURE Non-admixture correlated LOCPRIOR
Inferences on genetic structure with putative *neutral* markers (run 3)

<table>
<thead>
<tr>
<th></th>
<th>JVI</th>
<th>MEL</th>
<th>CEK</th>
</tr>
</thead>
<tbody>
<tr>
<td>JVI</td>
<td></td>
<td></td>
<td>-0.00063</td>
</tr>
<tr>
<td>MEL</td>
<td>-0.00002</td>
<td>0.0089</td>
<td></td>
</tr>
<tr>
<td>CEK</td>
<td>0.0004</td>
<td>0.00905</td>
<td>-0.0039</td>
</tr>
<tr>
<td>PC</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2,276 loci (Stacks); 20 PCA; 3 DA
Inferences on genetic structure with putative *neutral* markers (run 3)

<table>
<thead>
<tr>
<th></th>
<th>JVI</th>
<th>MEL</th>
<th>CEK</th>
</tr>
</thead>
<tbody>
<tr>
<td>JVI</td>
<td>-0.00066</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEL</td>
<td>0.00079</td>
<td>0.00928</td>
<td></td>
</tr>
<tr>
<td>CEK</td>
<td></td>
<td></td>
<td>-0.00425</td>
</tr>
<tr>
<td>PC</td>
<td>0.00036</td>
<td>0.00881</td>
<td></td>
</tr>
</tbody>
</table>

2,352 loci (VCFtools); 50 PCA; 3 DA
Inferences on genetic structure with putative selective markers (run 3)

30 loci (Stacks); 20 PCA; 3 DA

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.