Application of high-throughput sequencing to population differentiation in the cupped oysters Crassostrea angulata/C. gigas

P-A. Gagnaire, F. Cornette, S. Heurtebise, M-T. Augé, N. Bierne, S. Lapègue

1 Sampling:
- SNP genotyping: 9 populations, 29-50 individuals per population, 384 SNPs
- RAD sequencing: 4 populations, 25 individuals per population, 18,000 SNPs

2 Principle of RAD sequencing:

3 Genetic differentiation indices estimated from SNP and RAD datasets:

<table>
<thead>
<tr>
<th>Comparison Type</th>
<th>No. of RAD SNPs</th>
<th>Average F_S, RADs</th>
<th>Minimum F_S, RADs</th>
</tr>
</thead>
<tbody>
<tr>
<td>KEE/JAP within species</td>
<td>8259</td>
<td>0.009</td>
<td>0.028</td>
</tr>
<tr>
<td>JAP/LAF within species</td>
<td>8499</td>
<td>0.004</td>
<td>0.019</td>
</tr>
<tr>
<td>KEE/JAP between species native area</td>
<td>6479</td>
<td>0.177</td>
<td>0.058</td>
</tr>
<tr>
<td>SAD/LAF between species introduced area</td>
<td>6148</td>
<td>0.177</td>
<td>0.071</td>
</tr>
</tbody>
</table>

- Within species differentiation between native and introduced areas is stronger in C. angulata for both SNP and RAD datasets. Average within species genetic differentiation is higher from RAD data.
- Between species differentiation is stronger in Europe than in Asia when assessed with RAD data. Average between species genetic differentiation estimated with the SNP panel dataset is higher than with the RAD dataset.
- Differences among datasets for the average differentiation values may be explained by the difficulty to assemble highly divergent RAD tags between species.

4 Distribution of RAD pairwise FST values:

5 Spatial frequency patterns of between-species outlier SNPs:
Highly differentiated SNPs may be used to discriminate species and understand spatial introgression patterns.

6 Comparison of PCA from SNP and RAD data:
- RAD data increase the power to discriminate native and introduced populations.
- Stronger differentiation in C. angulata may reflect the older introduction time.

Conclusions:
- Genome-wide polymorphism datasets significantly improve our ability to discriminate closely related species that still exchange genes in nature.
- The amount of information allows the identification of native and introduced stocks, even when genetic differentiation is weak and introduction is recent.
- Mapping SNPs on the high-density linkage map and the reference genome will reveal the location of the genomic regions involved in reproductive isolation.
- Chromosomal patterns of genetic differentiation will be compared with the map of recombination rate to test the effect of reduced recombination in maintaining the species barrier.