Population genomics of European lobster: SNP discovery and developing a SNP panel

Tom L. Jenkins1, Charlie D. Ellis2, Jamie R. Stevens3.

1. Biosciences, College of Life and Environmental Sciences, University of Exeter, EX4 4QD, UK
2. National Lobster Hatchery, South Quay, Padstow, Cornwall, PL28 8BL, UK

Introduction

Why study European lobster?

European lobster (Homarus gammarus) is a crustacean typically found sheltering amongst hard substrate in coastal waters of the northeast Atlantic and Mediterranean. The market value of the species makes it a highly prized seafood, therefore healthy stocks are of great importance to those fishing communities they support. However, populations across its range are threatened by overexploitation. Natural connectivity between populations plays a role in recruitment and maintaining population persistence, however, precise estimates of these patterns are still relatively unknown at fine and regional spatial scales across the species range. In some areas, releases of hatchery-reared juveniles have attempted to reduce the deficit caused from overfishing. There is concern, however, that admixture of hatchery stock may erode the genetic architecture of wild lobsters and compromise the genetic diversity accumulated by discrete populations. It is, therefore, crucial that contemporary genetic structure is known to ensure hatchery stock are genetically compatible with natural stock.

What can population genomics tell us?

Genetic structure and the degree of gene flow can be used as a proxy for estimating population connectivity. Previous molecular methods have suggested little structuring in H. gammarus. However, these patterns may be a product of the often limited resolution associated with traditional molecular markers in studies of marine organisms. By using newer high-resolution genome-wide genetic markers such as single nucleotide polymorphisms (SNPs), this increases the power to detect fine-scale patterns of genetic structure in non-model organisms. With these techniques, researchers can expect to discover hundreds to thousands of genetic markers as opposed to tens of markers or less using microsatellites.

Key project aims:

- Develop a panel of 192 SNP markers for use in population-level analyses.
- Use SNP panel to explore fine and regional scale genetic structure and connectivity in H. gammarus.

Questions to be addressed by this project:

- Are there genetically distinct populations of lobsters? Are natural patterns of population connectivity able to facilitate replenishment of threatened stocks?
- Which wild lobster stocks can be supplemented with hatchery-reared juveniles? Can the origin of unknown caught lobsters be determined using their genetic profile?

Sampling coverage

Why is sampling across the geographical range important?

To maximise the discovery of informative SNPs for the panel, it is vital to incorporate as much of the contemporary genetic variation present in the study species as possible. Therefore, in this study, we attempted to collect samples from multiple areas across the range of H. gammarus (Fig. 1).

RAD sequencing and SNP genotyping

RAD sequencing

Restriction site associated DNA sequencing (RAD-seq) was used to discover SNPs for the panel. This method uses restriction enzymes to cut genomic DNA at specific sites and SNPs are identified in regions adjacent to the cut sites. Libraries were prepared and sequenced by the University of Exeter Sequencing Service.

Preliminary library

A preliminary experiment was carried out using seven samples (Table 1) to validate the methods and quality of the input DNA.

First Library

The first library contains 48 lobsters from 25 sites.

Bioinformatics

The sequencing data is analysed using the software Stacks, which has been specifically designed to deal with RAD-seq data.

SNP panel

The SNP panel comprises 192 of the most informative SNPs. Primers are designed to amplify each SNP and a Fluidigm EP1 will genotype each SNP across all individuals in the sampling dataset (Fig. 2).

This high-throughput approach also enables new samples to be screened relatively easily and avoid costly library prep and Illumina sequencing lanes.

Preliminary results

Table 1: Raw reads obtained from the Illumina MiSeq 2500 and the percentage of reads obtained after quality control using the procseq, redcap function in the Stacks software. The coverage, no. of loci, no. of SNPs, no. of polymorphic loci, percentage of loci that were polymorphic for each sample after processing with the Stacks, map program in Stacks are shown. Samples originate from the following sites: FL, Padstow; KA, Kame; LD, Lostwithiel; CR, Charlestown; NR, Northumberland; OR, Cheyne.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Raw reads</th>
<th>Retained reads (%)</th>
<th>Coverage (x)</th>
<th>No. of loci</th>
<th>No. of SNPs</th>
<th>No. of polymorphic loci</th>
<th>Loci polymorphic (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FL5</td>
<td>5.7e6</td>
<td>97.8</td>
<td>36</td>
<td>16,806</td>
<td>11,251</td>
<td>12.7</td>
<td>15.2</td>
</tr>
<tr>
<td>KA4</td>
<td>6.0e6</td>
<td>96.9</td>
<td>38</td>
<td>20,528</td>
<td>12,859</td>
<td>15.2</td>
<td>16.8</td>
</tr>
<tr>
<td>KA8</td>
<td>4.9e6</td>
<td>97.8</td>
<td>37</td>
<td>17,712</td>
<td>11,246</td>
<td>15.7</td>
<td>17.0</td>
</tr>
<tr>
<td>1LO31</td>
<td>4.8e6</td>
<td>95.5</td>
<td>23</td>
<td>20,226</td>
<td>13,827</td>
<td>13.1</td>
<td>15.7</td>
</tr>
<tr>
<td>LR4</td>
<td>3.2e6</td>
<td>81.7</td>
<td>17</td>
<td>15,550</td>
<td>9,972</td>
<td>13.3</td>
<td>14.5</td>
</tr>
<tr>
<td>1NH8</td>
<td>3.5e6</td>
<td>92.7</td>
<td>19</td>
<td>16,506</td>
<td>11,418</td>
<td>13.9</td>
<td>15.7</td>
</tr>
<tr>
<td>1OR6</td>
<td>4.6e6</td>
<td>97.8</td>
<td>26</td>
<td>19,964</td>
<td>13,701</td>
<td>14.7</td>
<td>16.8</td>
</tr>
</tbody>
</table>

Applications of research

A SNP panel will provide us with the enhanced resolution previously unavailable for H. gammarus and will enable the most powerful assessment of genetic structure and connectivity to date. This research will likely contribute to:

- Improving management of lobster stocks.
- Monitoring the genetic diversity and structure of stocks over time and space.
- Evaluating the fate and contribution of hatchery-reared lobsters.
- Estimating the origin of any given lobster using their genetic profile.