Minlon: towards on-board assessments of microorganisms and marine megafauna

Judith van Bleijswijk & Julia Engelmann, Royal NIOZ, Texel
need for on board analyses!

- information directed sampling
- efficient (work during transit)
- shorter time between sampling and analyses

Porelab-MinIon

- small portable platform
  - on board analyses
- in situ read lengths
  - mixed amplicon analyses
- long read sequencing
  - high resolution identification
  - metagenomes
Porelab project: Environmental DNA of marine megafauna

eDNA:
2L water eDNA on 0.2 µm filter
PCR COI and 12S gene

Visual observation:
Visual observation:

- Bandwing flying fish
- Sailfin flying fish
- Tropical two-wing flying fish
- Atlantic spotted dolphin

Porelab project: Environmental DNA of marine megafauna

eDNA:
- on board?
- mixed amplicons?
- output?
Porelab project: marine fungi

- Role of fungi in algae breakdown
- CO$_2$?
- CH$_4$?
- Global warming!

In situ fungi community analyses
- PCR of rRNA operon 2500 bp
- Nanopore sequencing

Cultures
- Rate of algae breakdown
The first nanoPore North Sea metagenome
Microbial communities in the marine food web

Microbial community sequencing

Extract genomes or marker gene

Identify species

### Species abundance table

<table>
<thead>
<tr>
<th>Sample 1</th>
<th>Sample 2</th>
<th>Sample 3</th>
<th>Sample 4</th>
<th>...</th>
</tr>
</thead>
</table>
| Archi 1  | 50       | 12       | 78       | 22  | ...
| Bacti 1  | 30       | 43       | 32       | 45  | ...
| Bacti 2  | 23       | 65       | 11       | 33  | ...
| ...      | ...      | ...      | ...      | ... | ... |
- 3.6 mio MinION reads, 2.7 reads >=1000 bp
- 18 386 contigs,
  - mean length 19 427 bp
  - max length 1 626 368 bp,
  - N50: 30720 bp
  - L50 2149 bp
Assigned proteins

A. Domain Level

B. Phylum Level

< 0.3 % of total proteome
Assigned contigs

A Domain Level

B Phylum Level

Unassigned: no hit with blastp
equal percentage total
not classified at phylum

Assigned contigs (%)

- Bacteria: 62.7%
- Viruses: 29.3%
- Archaea: 4.0%
- Eukaryota: 3.8%

< 0.3% of total contigs
### Metagenome annotation on contig level

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contigs without blastp hit</td>
<td>5,378</td>
</tr>
<tr>
<td>Avg length of contigs without known protein</td>
<td>7,954 bp</td>
</tr>
<tr>
<td>Longest unassigned contig</td>
<td>107,818 bp</td>
</tr>
<tr>
<td>Unassigned</td>
<td>29.25%</td>
</tr>
<tr>
<td>Total protein number</td>
<td>467,050</td>
</tr>
<tr>
<td>Proteins without blast hit</td>
<td>237,253</td>
</tr>
<tr>
<td>Unassigned proteins</td>
<td>50.80%</td>
</tr>
<tr>
<td>Assigned TIGRs</td>
<td>32,104</td>
</tr>
<tr>
<td>Unassigned TIGRs</td>
<td>437,289</td>
</tr>
<tr>
<td></td>
<td>93.13%</td>
</tr>
</tbody>
</table>
Future plans

Time series of microbial communities in the North Sea to study causal relationships between microbial species

Combine long read metagenomes with marker gene data for better coverage and resolution

Better understand the role of microbes in the marine food web
Acknowledgements

• Porelab: three projects granted
  warm welcome, coffee, tea

• Floyd Wittink: suggestions fungi PCR
  tips bioinformatics and hardware

• Nikola Petrushevski: dedicated lab work
  sharing tips and tricks
### Most frequent classes

<table>
<thead>
<tr>
<th>Class</th>
<th>Hits</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gammaproteobacteria</td>
<td>3290</td>
<td>17.90</td>
</tr>
<tr>
<td>Alphaproteobacteria</td>
<td>3198</td>
<td>17.39</td>
</tr>
<tr>
<td>Flavobacteriia</td>
<td>1748</td>
<td>9.51</td>
</tr>
<tr>
<td>Bacteroidales</td>
<td>255</td>
<td>1.39</td>
</tr>
<tr>
<td>Betaproteobacteria</td>
<td>248</td>
<td>1.35</td>
</tr>
<tr>
<td>Planctomycetia</td>
<td>236</td>
<td>1.28</td>
</tr>
<tr>
<td>Verrucomicrobiae</td>
<td>198</td>
<td>1.08</td>
</tr>
<tr>
<td>Acidimicrobiia</td>
<td>186</td>
<td>1.01</td>
</tr>
</tbody>
</table>
Advantages of long reads

- Better metagenome assemblies:
  - Span repetitive regions
  - Larger contigs
  - Close more gaps
  - Easier to assign contigs to genomes

- Can sequence the complete 16S (1.5 kb) → higher resolution

Kellog 2015, Nature Plants