Using environmental metabarcoding to assess change associated with marine structures

Structures in the Marine Environment (SIME2019) - 17th May 2019

Tom Wilding, Mark Coulson (SAMS/RLI-UHI)
Structure

• What is metabarcoding?
• Fives steps of metabarcoding
• Initial results
• Regulatory context
What is metabarcoding?

• Barcoding is the identification of taxa via their DNA sequence
• Meta – is the sequencing of numerous (millions) of DNA fragments, simultaneously.
• Metabarcoding is a rapidly evolving technology that enables ~high resolution taxonomic identification across broad taxonomic groups in ~100 samples simultaneously.
Five steps in metabarcoding

1. Sample collection
2. ‘Wet lab’
3. Sequencing
4. Bioinformatics
5. Data interpretation (statistical modelling)
Benthic sample collection

• Sediment or filter papers from water samples
• Typically 5 – 20 g per sediment sample
• Filter ~litres of water
Wet-lab

1. DNA extraction, purification, marker PCR, tagging and quantification

'Library' preparation
Wet-lab

- Extracted DNA is a mixture of bacterial/protistan/meiobenthic cellular DNA and eDNA from all groups (including macro- and mega-benthos)
- From the extracted DNA can target regions within:
  - 16S (bacteria), 18S (general eukaryotes), COI (metazoa) others for plants, fungi and specialised groups
- Regions within markers are ‘targeted’ by primers, which can be bespoke designed, none are truly ‘universal’ (e.g. none cross all taxa)
- Advantages/disadvantages to all regions/markers and associated primers
- Sequencing - ‘Illumina’ or ‘Oxford nanopore’ dominate
Bioinformatics

• Shift from identifying ‘operational taxonomic units’ (OTUs) by clustering sequences to ASVs (amplicon sequence variants)
• Annotate sequences by comparison with databases
• Databases of various ‘quality’ – most charismatic species (e.g. whales, fish) sequenced for most markers
• Most meiobenthos not sequenced
• Databases are not necessarily well curated
• Bacterial databases are good and well curated
In a regulatory context, the question is:

- Can we move from this:
Super-abundance of the polychaete *Capitella sp*
• And this ...
More diverse, larger, less abundant benthic infauna
• Via:
Typical metabarcoding data from 100-sample (taxa barplot)
To effective regulatory compliance assessments?
Preliminary results

• Variability between replicates within grab
Assess within grab variability as a function of marker

Ideally

- Low variation (reps within grab)
- Clear gradient
- No overlapping stations
- Consistent trend between Sites
Within grab variability depends on the marker - COI

- Lots of noise
- Overlapping
- But still some intuitive trends

COI = metazoa

Figure 64 Marker: COI, Taxon: Species, Remove unassigned: YES, Presence/absence: NO. Stress=0.2078
Within grab variability depends on the marker – 18S

- Reasonable resolution
- Clear patterns of ‘distance-effect’
- Some overlap between stations

18S = general eukaryotes

Figure S1 Marker 18S, Taxon: Species, Make Unique: FALSE, Remove unassigned: YES, Presence/absence: NO, Stress=0.1286
Within grab variability depends on the marker - 16S

- Very tight clustering of replicates from within the same grab
- Stations very clearly separated
- Low stress and intuitive patterns
How good are bacteria at predicting traditional morphology-based metrics and what taxa are driving these patterns?
Supervised machine learning - RandomForest

- Ideally a 1:1 relationship between traditional and predicted, with no error
- Already reasonable predictive power (prediction interval shown)
- Primary drivers include member of sulphur metabolising families (Granulosicoccaceae), Xanthomonadales JTB255 group, Flammeovirgaceae and ‘unknowns’
• Collaboration with Thorsten Stoeck
• Clear seasonal cycle at the reference stations.
• Natural cycles do exist
What does metabarcoding offer?

• Molecular approaches are cheaper, faster, more comprehensive and objective cf macrobenthos
• Can target any species or group (e.g. whales to bacteria)
• Increasingly useful concurrence between macrobenthic and molecular indices (pattern-matching)

Will it be used in regulation/compliance monitoring?
• Yes, molecular will largely replace traditional approaches over the next 3 years.
My thanks to the following for supporting this research

Questions?