Invertebrate community composition in the Falklands using environmental DNA-metabarcoding

Preliminary Report April 2022

Sofia Consuegra

Department of Biosciences, College of Science, Swansea University,

Swansea SA2 8PP, United Kingdom; email: s.consuegra@swansea.ac.uk

Background

Environmental DNA (eDNA) extracted from water samples is a rapid, cost effective tool for monitoring species distributions (Lodge et al., 2012) reducing the need for destructive sampling (Bylemans, Furlan, Pearce, Daly, & Gleeson, 2016). Metabarcoding analysis of eDNA can be useful for whole community/broad range assessment (Bohmann et al., 2014; Lacoursière-Roussel, Côté, Leclerc, & Bernatchez, 2016), and also performs well for target species monitoring (Harper et al., 2018), offering a potential alternative to traditional monitoring methods. The analysis of environmental DNA (eDNA) is becoming widely used for the detection of invasive species (Biggs et al., 2015; Thomsen & Willerslev, 2015) and protocols are being refined to increase its accuracy and reliability (Goldberg et al., 2016; Wilson, Wozney, & Smith, 2016). The aim of this pilot study was to assess the ability of eDNA-metabarcoding to identify the invertebrate community composition of 5 different water bodies in the Falkland Islands.

Methods

Sample origin, DNA extraction and library preparation

Samples consisted of filters collected in the field after filtering different volumes of water at five different locations in the Falklands (Table 1). The filters were refrigerated and transported to Swansea University in august 2021 where they were kept at 4°C for further processing. DNA was extracted from the refrigerated filters using the QIAGEN PowerSoil kit, using a homogenization step (a Precellys 24 tissue homogenizer, Bertin Instruments). Extracted DNA was stored at -80°C until library preparation.

Table 1. Location of sampling sites at the Falklands, volume of water filtered, number of filters used, properties of the water (pH, conductivity, dissolved oxygen-DO, temperature and salinity) and number of invertebrate samples taken for comparison purposes (no. vials).

Date	Location	Farm	Long	Lat	Blank Water	Volume sampled (ml)				No. filters					Conductivity (µS/cm)	DO (mg/L)	Temp (°C)	Salinity (PSU)	No. vials invert
						Blank	R1	R2	R3	Blank	R1	R2	R3	pH					
14/07/2021	Yorke Bay Pond	Stanley Common	-57.78375	-51.68061	invitrogen	1000	650	600	600	2	4	4	4	7.15	499	11.0	1.2	0.2	4
15/07/2021	Round Pond	Stanley Common	-57.872137	-51.7279	Millipore T	200	55	50	55	1	10	10	10	6.49	1264	13.9	1.2	0.6	2
20/07/2021	Sand Pond	MPA / Fitzroy	-58.468329	-51.8261	Millipore T	1000	900	900	940	2	6	5	5	7.15	1103	11.8	2.7	0.6	5
21/07/2021	Malo River	Riverview	-58.327963	-51.62211	Autoclaved	945	1000	1000	1000	2	2	2	2	4.73	105	12.8	2.7	0.0	3
01/08/2021	Tern Hill Stream	Weddell Island	-60.924816	-51.87451	Mix of Mill	1000	1000	950	1000	2	4	3	3	4.52	215	10.7	5.5	0.0	5

Extraction blanks (where no filter was added), were carried through all steps of the library preparation and bio-informatic analysis. Initially, all samples were amplified in triplicate, in 25ul PCR reactions, with 2µl of template and 0.5µl of 16S RNA primers (Ins16S_1__Ri7 + Inse01_F; targeting mainly insects), for 25 cycles, with an annealing temperature of 55°C. Subsequently, 5µl of this first reaction was used as a template for a second reaction, using the 16S primers with the addition of overhang adaptors for subsequent Nextera indexing, using identical PCR conditions, for 10 more cycles. PCR triplicates were then pooled, using 10µl of each triplicate from the nested PCR reaction and purified using Agencourt AMPure XP beads (Beckman Coulter). Subsequently, amplicons were indexed using the Nextera XT Index Kit v2 (Illumina, Inc., San Diego, California, USA), and DNA concentration of each reaction was quantified via Qubit dsDNA HS Assay (Invitrogen) and pooled in equal molar concentrations. Filtration blanks for each site, extraction blanks for each round of extractions and PCR blanks were carried through all steps of library preparation, including triplication in the first PCR, and bioinformatics processing.

Water samples and all blanks were subjected to the same bioinformatics pipeline and processed simultaneously. The amplicon sequence variant (ASV) approach was used because it enables detection of single nucleotide differences (Callahan, McMurdie, & Holmes, 2017) and therefore provides a higher resolution than a traditional OTU approach. Qiime2 (version 2019.1)(Bolyen et al., 2019) was used to process de-multiplexed paired end sequences. DADA2 (Callahan et al., 2016) within Qiime2 was used for de-noising steps. Based on read quality scores, the first 10bp of each sequence were trimmed and all sequences truncated to 100bp in length. Subsequently, sequencing errors were corrected where possible, and chimeric sequences removed, paired end reads were joined and sequences de-replicated using the default DADA2 settings in Qiime2. Initial taxonomic assignment was conducted using a custom database. Initially, reads were classified against a full database that included all taxa available on NCBI amplified *in silico* with the 16S primers, using the KNN method in mothur (Schloss et al., 2009).

Results

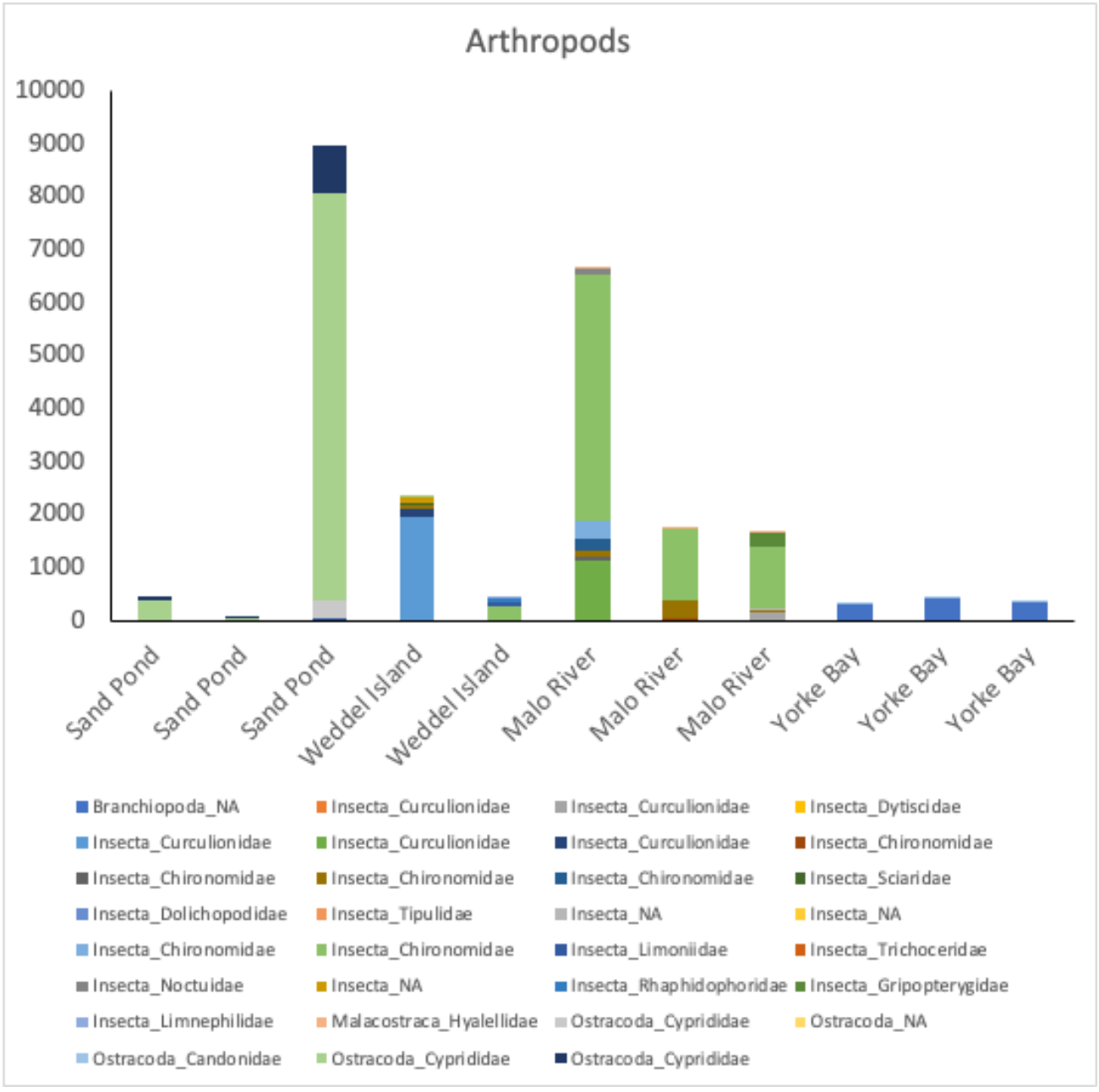
A total of 57,046 raw paired end reads were generated from the five sampling sites, of which 37,808 remained for downstream analyses after filtering. Sufficient number of reads were obtained from four of the five sites sampled (Table 2; the number of reads for Round Pond was too small to be considered). Of these, 23,709 reads corresponded to invertebrates (arthropods and annelids). ASVs matched to 15 known species and 20 genera, belonging to classes Insecta, Ostracoda, Citellata, Brachiopoda and Malacostraca (Figures 1 and 2). A number of ASVs could not be identified; this can be due to two reasons: on the one hand many of the species in the Falklands may not having been described and on the other hand the coverage of the 16S database maybe limited for some taxa. Using additional primers in the COI and 12S region could help matching more of the reads to particular species or at least genera.

Table 2. Number of total reads from eDNA-16S metabarcoding of Falklands water samples.

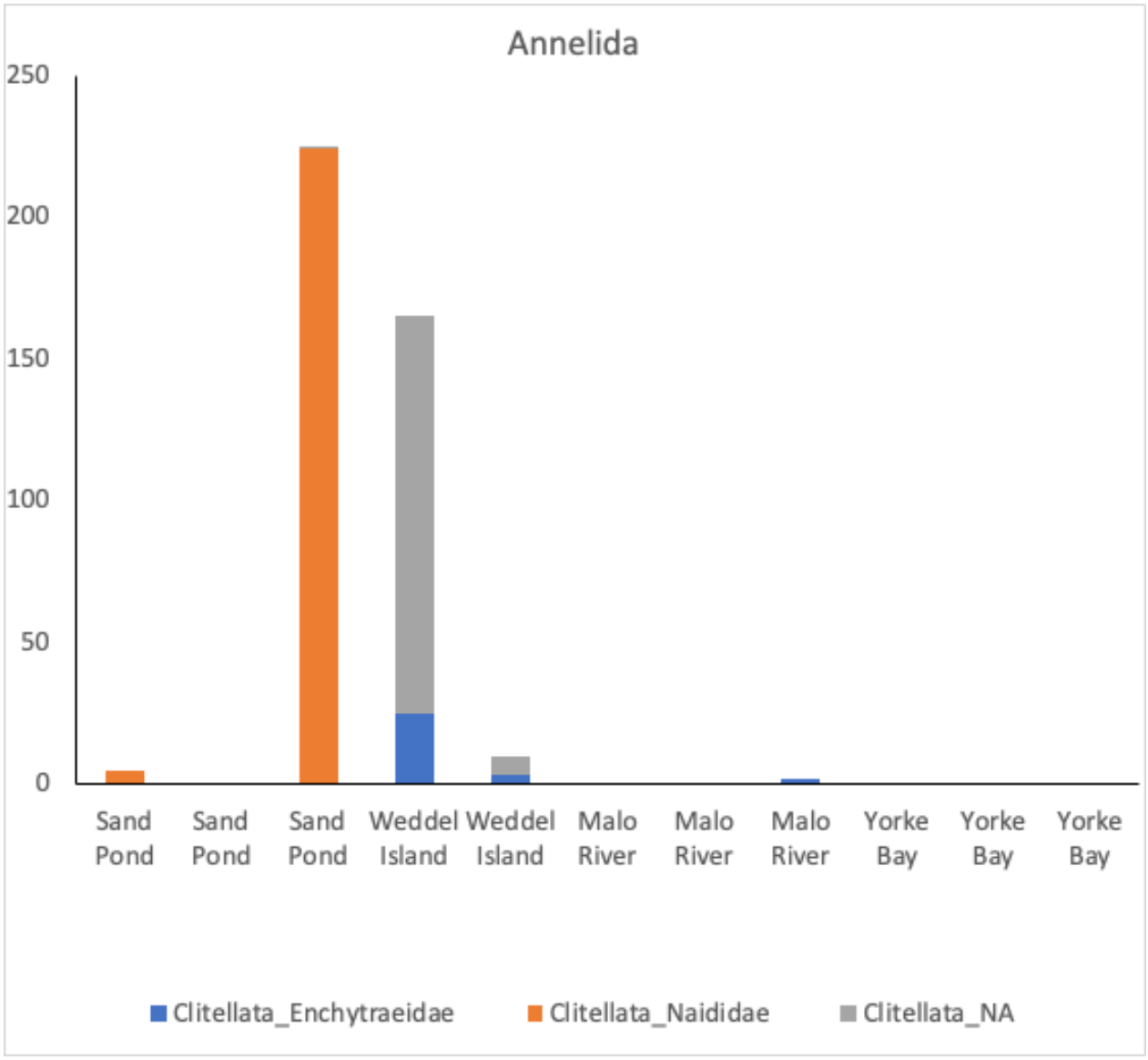
	Working Replicates	Total number of reads (after quality control)	Number of reads of invertebrates
San Pond	3	9884	9676
Weddel Island	2	9996	2905
Malo River	3	13325	10034
Yorke Bay	3	4595	1094
Round Pond	0	0	0

Figure 1. Invertebrate species (a. arthropods and b. annelids) distribution by sampling site based on eDNA-metabarocoding from water samples.

a)



b)



References

Biggs, J., Ewald, N., Valentini, A., Gaboriaud, C., Dejean, T., Griffiths, R. A., . . . Dunn, F. (2015). Using eDNA to develop a national citizen science-based monitoring programme for the great crested newt (Triturus cristatus). *Biological Conservation, 183*, 19-28. doi:http://dx.doi.org/10.1016/j.biocon.2014.11.029

Bohmann, K., Evans, A., Gilbert, M. T. P., Carvalho, G. R., Creer, S., Knapp, M., . . . de Bruyn, M. (2014). Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology & Evolution, 29*(6), 358-367.

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., . . . Asnicar, F. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology, 37*(8), 852-857.

Bylemans, J., Furlan, E. M., Pearce, L., Daly, T., & Gleeson, D. M. (2016). Improving the containment of a freshwater invader using environmental DNA (eDNA) based monitoring. *Biological Invasions, 18*, 3081-3089.

Callahan, B. J., McMurdie, P. J., & Holmes, S. P. (2017). Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *The ISME journal, 11*(12), 2639-2643.

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods, 13*(7), 581-583.

Goldberg, C. S., Turner, C. R., Deiner, K., Klymus, K. E., Thomsen, P. F., Murphy, M. A., . . . Cornman, R. S. (2016). Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution*.

Lacoursière-Roussel, A., Côté, G., Leclerc, V., & Bernatchez, L. (2016). Quantifying relative fish abundance with eDNA: a promising tool for fisheries management. *Journal of Applied Ecology, 53*, 1148-1157.

Lodge, D. M., Deines, A., Gherardi, F., Yeo, D. C. J., Arcella, T., Baldridge, A. K., . . . Zeng, Y. (2012). Global introductions of crayfishes: evaluating the impact of species invasions on ecosystem services. *Annual Review of Ecology, Evolution and Systematics, 43*, 449-472.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., . . . Robinson, C. J. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology, 75*(23), 7537-7541.

Thomsen, P. F., & Willerslev, E. (2015). Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation, 183*, 4-18. doi:http://dx.doi.org/10.1016/j.biocon.2014.11.019

Wilson, C. C., Wozney, K. M., & Smith, C. M. (2016). Recognizing false positives: synthetic oligonucleotide controls for environmental DNA surveillance. *Methods in Ecology and Evolution, 7*(1), 23-29.