**Supporting Information**

**DNA metabarcoding marker choice skews perception of marine eukaryotic biodiversity**

Jordan M. Casey, Emma Ransome, Allen G. Collins, Angka Mahardini,Eka M. Kurniasih, Andrianus Sembiring,Nina M. D. Schiettekatte, Ni Kadek Dita Cahyani, Aji Wahyu Anggoro, Mikaela Moore, Abby Uehling, Mahdi Belcaid, Paul H. Barber, Jonathan B. Geller, Christopher P. Meyer

**Supporting Tables**

|  |  |
| --- | --- |
| **Primer Label** | **Primer Sequence (5'-3')** |
| m1COIintF\_Tag1 | **AGACGC**GGWACWGGWTGAACWGTWTAYCCYCC |
| m1COIintF\_Tag2 | **AGTGTA**GGWACWGGWTGAACWGTWTAYCCYCC |
| m1COIintF\_Tag3 | **ACTAGC**GGWACWGGWTGAACWGTWTAYCCYCC |
| m1COIintF\_Tag4 | **ACAGTC**GGWACWGGWTGAACWGTWTAYCCYCC |
| m1COIintF\_Tag5 | **ATCGAC**GGWACWGGWTGAACWGTWTAYCCYCC |
| m1COIintF\_Tag6 | **ATGTCG**GGWACWGGWTGAACWGTWTAYCCYCC |
| m1COIintF\_Tag7 | **ATAGCA**GGWACWGGWTGAACWGTWTAYCCYCC |
| jgHCO\_Tag1 | **AGACGC**TAIACYTCIGGRTGICCRAARAAYCA |
| jgHCO\_Tag2 | **AGTGTA**TAIACYTCIGGRTGICCRAARAAYCA |
| jgHCO\_Tag3 | **ACTAGC**TAIACYTCIGGRTGICCRAARAAYCA |
| jgHCO\_Tag4 | **ACAGTC**TAIACYTCIGGRTGICCRAARAAYCA |
| jgHCO\_Tag5 | **ATCGAC**TAIACYTCIGGRTGICCRAARAAYCA |
| jgHCO\_Tag6 | **ATGTCG**TAIACYTCIGGRTGICCRAARAAYCA |
| jgHCO\_Tag7 | **ATAGCA**TAIACYTCIGGRTGICCRAARAAYCA |

**Table S1.** Seven-tailed primer pairs of MICOIintF and jgHCO2190. Six base pair tags are bolded within the primer sequence.

|  |  |  |
| --- | --- | --- |
| **Rank** | **COI** | **18S** |
| 11 | Sipuncula | Ochrophyta |
| 12 | Nemertea | Entoprocta |
| 13 | Ascomycota | Nematoda |
| 14 | Amoebozoa | Miozoa |
| 15 | Nematoda | Xenacoelomorpha |
| 16 | Chlorophyta | Nemertea |
| 17 | Chaetognatha | Retaria |
| 18 | Platyhelminthes | Ciliophora |
| 19 | Pseudofungi | Sipuncula |
| 20 | Choanozoa | Chlorophyta |
| 21 | Entoprocta | Bigyra |
| 22 | Kinorhyncha | Cercozoa |
| 23 | Sulcozoa | Kinorhyncha |
| 24 | Haptophyta | Gastrotricha |
| 25 | Rotifera | Brachiopoda |
| 26 | Cercozoa | Pseudofungi |
| 27 | Hemichordata | Charophyta |
| 28 | Basidiomycota | Ctenophora |
| 29 | Xenacoelomorpha | Ascomycota |
| 30 | Bigyra | Choanozoa |
| 31 | Picozoa | Amoebozoa |
| 32 | Miozoa | Hemichordata |
| 33 | Gastrotricha | Haptophyta |
| 34 | Rhizaria | Sulcuzoa |
| 35 | Placozoa | Basidiomycota |
| 36 | Retaria | Heliozoa |
| 37 | Proteobacteria | Loukozoa |
| 38 | Chytridiomycota | Tardigrada |
| 39 |  | Metamonada |
| 40 |  | Zygomycota |
| 41 |  | Acanthocephala |
| 42 |  | Rotifera |
| 43 |  | Sulcozoa |
| 44 |  | Placozoa |
| 45 |  | Cryptista |
| 46 |  | Chytridiomycota |
| 47 |  | Chaetognatha |
| 48 |  | Orthonectida |
| 49 |  | Glomeromycota |
| 50 |  | Loricifera |
| 51 |  | Hermichordata |

**Table S2.** List of phyla beyond the ten phyla with the highest average relative abundances across the ARMS for the COI and 18S datasets, ranked by average relative abundance. These phyla are summarized as the “Other” category in Figure 4.

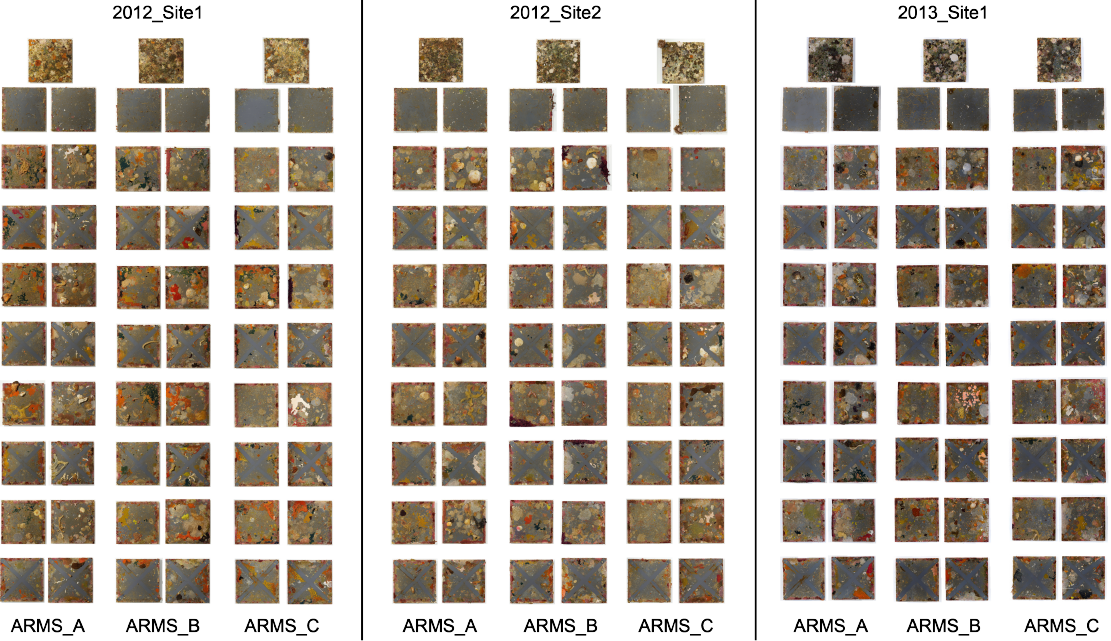
|  |  |  |
| --- | --- | --- |
| **Phylum** | **COI** | **18S** |
| Amoebozoa | 0.04% [0.02, 0.07] | <0.01% [<0.01, 0.01] |
| Annelida | 13.73% [7.68, 24.56] | 14.7% [8.47, 24.81] |
| Arthropoda | 14.41% [8.12, 25.37] | 17.6% [10.04, 30.71] |
| Ascomycota | 0.09% [0.05, 0.16] | 0.01% [0.01, 0.02] |
| Bryozoa | 12.67% [7.25, 22.59] | 2.73% [1.51, 4.97] |
| Chlorophyta | 0.03% [0.02, 0.06] | 0.2% [0.11, 0.34] |
| Choanozoa | 0.01% [<0.01, 0.01] | 0.01% [<0.01, 0.01] |
| Chordata | 2.85% [1.55, 5.25] | 15.44% [8.68, 27.4] |
| Cnidaria | 6.95% [3.91, 12.18] | 9.22% [5.2, 16.59] |
| Echinodermata | 1.83% [0.99, 3.33] | 1.73% [0.91, 3.2] |
| Entoprocta | 0.01% [<0.01, 0.01] | 0.39% [0.2, 0.75] |
| Haptophyta | <0.01% [<0.01, <0.01] | <0.01% [<0.01, <0.01] |
| Mollusca | 2.94% [1.65, 5.28] | 3.18% [1.82, 5.63] |
| Nematoda | 0.03% [0.02, 0.06] | 0.7% [0.39, 1.23] |
| Nemertea | 0.15% [0.08, 0.31] | 0.31% [0.16, 0.6] |
| Ochrophyta | 1.61% [0.9, 2.83] | 1.43% [0.82, 2.53] |
| Porifera | 15.14% [8.72, 26.37] | 7.54% [4.26, 13.2] |
| Pseudofungi | 0.01% [0.01, 0.03] | 0.02% [0.01, 0.03] |
| Rhodophyta | 14.86% [8.31, 26.12] | 9.69% [5.49, 16.79] |
| Sipuncula | 0.5% [0.27, 0.96] | 0.19% [0.1, 0.34] |

**Table S3**. Means ± 95% credible intervals from a Bayesian mixed model that assesses the effect of COI *versus* 18S marker choice on the relative abundances of 20 co-occurring phyla (see Figure 5).

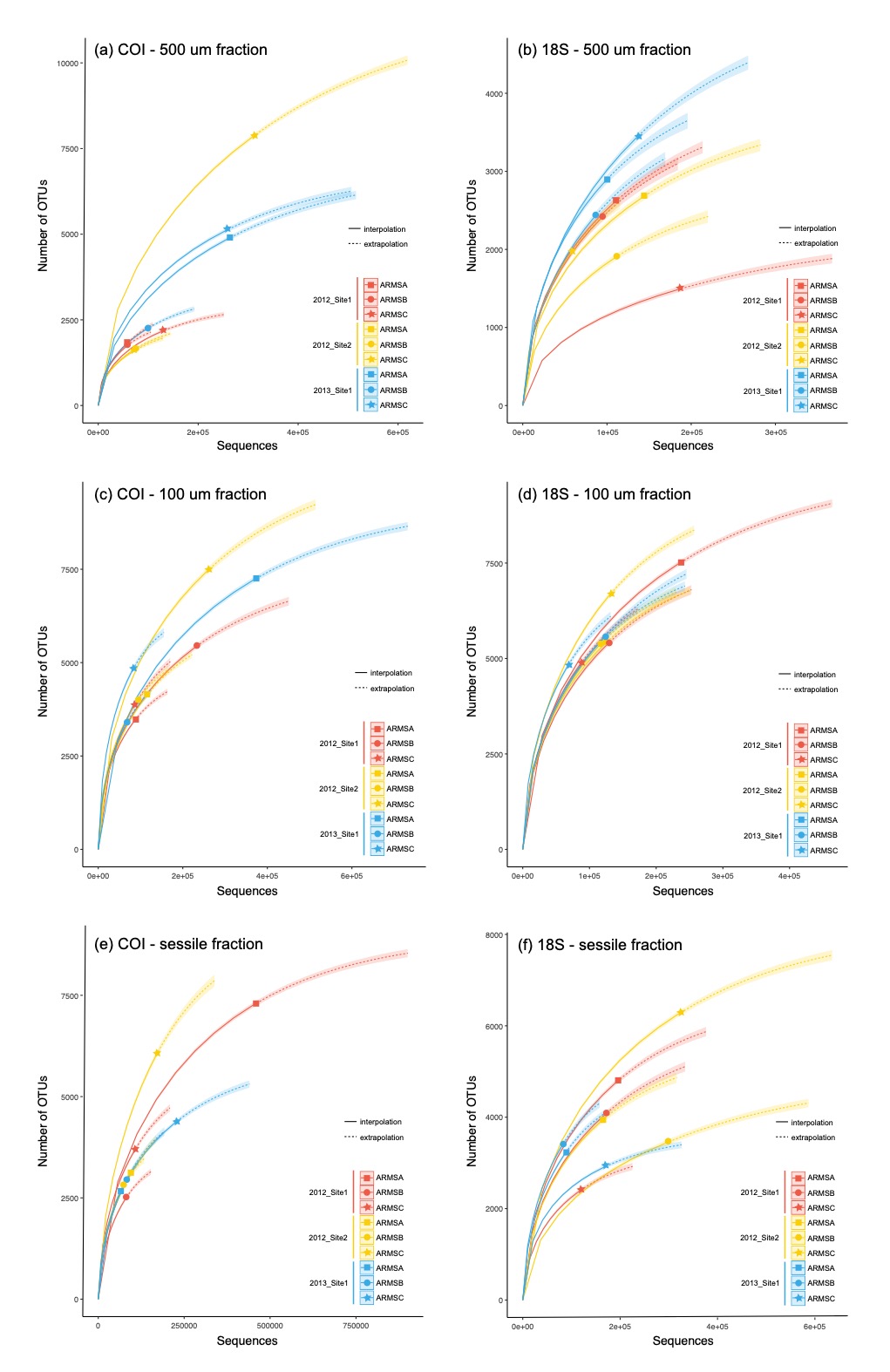
|  |  |  |  |
| --- | --- | --- | --- |
| **Phylum** | **COI** | **18S** | **CoralNet** |
| Annelida | 1.36% [0.62, 3.06] | 2.77% [1.32, 5.88] | 2.52% [1.24, 5.14] |
| Bryozoa | 22.33% [10.98, 46.59] | 5.15% [2.4, 11.05] | 27.54% [13.48, 57.21] |
| Chlorophyta | 0.04% [0.02, 0.08] | 0.38% [0.18, 0.77] | 1.55% [0.72, 3.21] |
| Chordata | 4.44% [2.06, 9.33] | 35.95% [16.66, 72.78] | 3.19% [1.54, 6.57] |
| Mollusca | 2.35% [1.08, 5.2] | 2.95% [1.37, 6.47] | 2.4% [1.09, 5.28] |
| Porifera | 32.87% [15.76, 66.71] | 12.86% [5.83, 26.6] | 38.56% [18.47, 77.73] |
| Rhodophyta | 19.07% [9.34, 38.93] | 18.32% [8.8, 36.47] | 21.11% [10.28, 42.99] |

**Table S4.** Means ± 95% credible intervals from a Bayesian mixed model that assesses the effect of COI marker choice, 18S marker choice, and visual image analysis of ARMS plates (CoralNet) on the relative abundances of seven co-occurring phyla (see Figure 6).

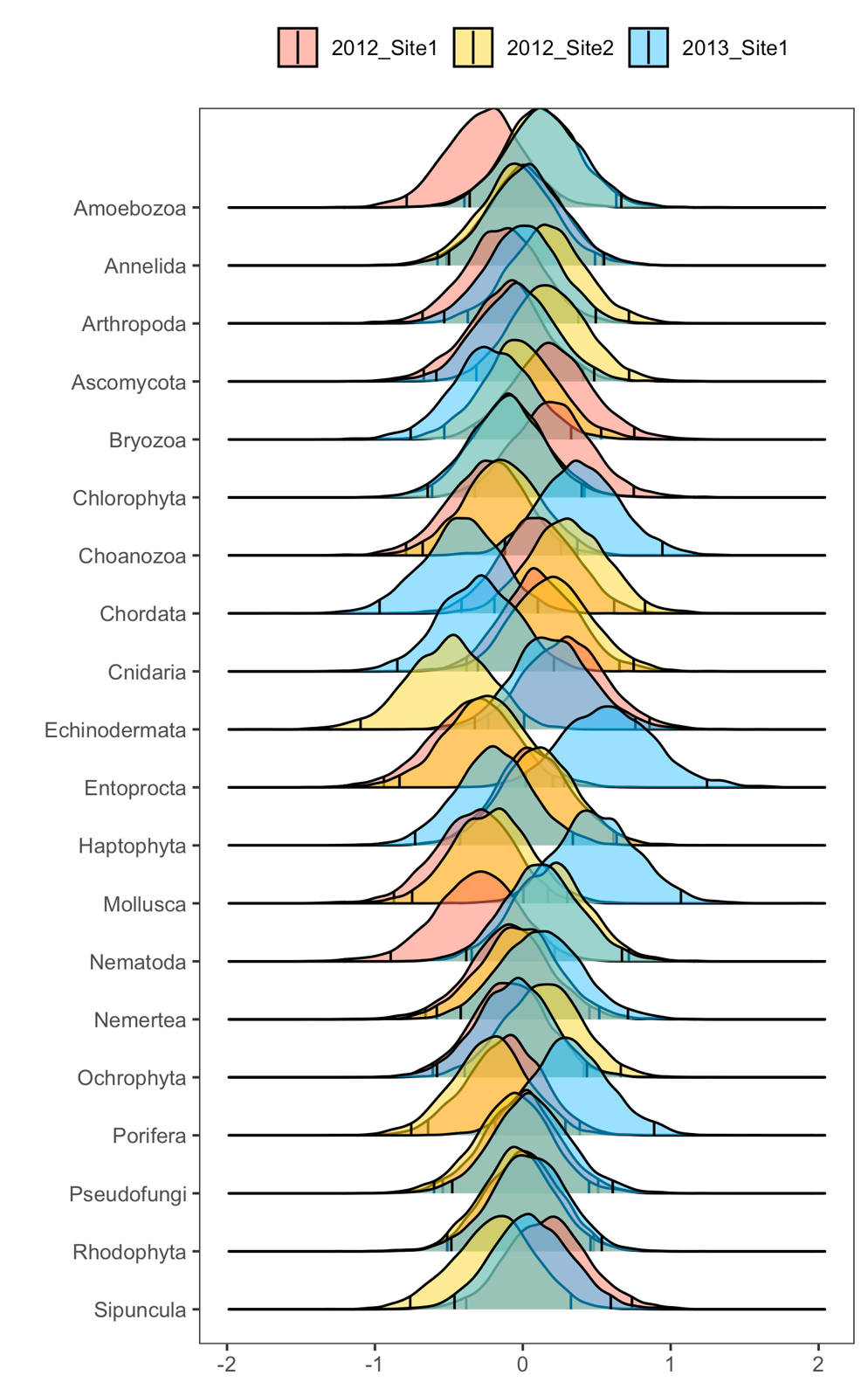
**Supporting Figures**

****

**Figure S1.** Compilation of high quality plate images from the nine autonomous reef monitoring structures (ARMS) across the sampling years and sites (Site 1 in 2012, Site 2 in 2012, and Site 1 in 2013). For each ARMS, the images on the left-hand side are the tops of the plates, and the photos on the right-hand side are the bottoms of the plates. High quality images of each individual plate are also available on the California Digital Library (<https://n2t.net/ark:/21547/DXy2>).



**Figure S2**. Rarefaction curves for the (a,c,e) COI and (b,d,f) 18S markers showing operational taxonomic unit (OTU) richness by total number of sequences for each fraction of each autonomous reef monitoring structure (ARMS) in each year and site. Fractions include the 500 μm – 2 mm fraction (a,b; 500 μm fraction), the 500 μm – 106 μm fraction (c,d; 100 μm fraction), and the sessile fraction (e,f). Solid lines indicate interpolated values, and dotted lines indicate predicted extrapolated values.



**Figure S3**. The relative effect of the spatio-temporal variables (year and site) on the predicted relative abundances of the 20 co-occurring phyla across the COI and 18S datasets. The mean posterior distributions (±95% highest density intervals (HDIs) are designated by vertical black bars within each curve) from a Bayesian mixed model. When the HDIs between the two curves do not overlap, this indicates strong differences among quantification techniques.