

Supplemental Information for:

Quantitative assessment of reef foraminifera community from metabarcoding data

Elsa B. Girard^{a,b}, Emilie A. Didaskalou^c, Andi M. A. Pratama^d, Carolina Rattner^a, Raphaël Morard^e, Willem Renema^{a,b}

^a Naturalis Biodiversity Center, Darwinweg 2, 2333 CR Leiden, the Netherlands

^b IBED, University of Amsterdam, Sciencepark 904, 1098 XH Amsterdam, the Netherlands

^c CML, University of Leiden, Einsteinweg 2, 2333 CC Leiden, the Netherlands

^d Marine Science Department, Faculty of Marine Science and Fisheries, Hasanuddin University, Jl. Perintis Kemerdekaan Km. 10 Tamalenrea, Makassar 90245, Indonesia

^e MARUM, University of Bremen, Leobener Str. 8, 28359 Bremen, Germany

Running title: Quantitative metabarcoding for foraminifera

Corresponding author: Elsa B. Girard

Email: elsa.girard@naturalis.nl

MOLECULAR ECOLOGY RESOURCES

Supplementary figures

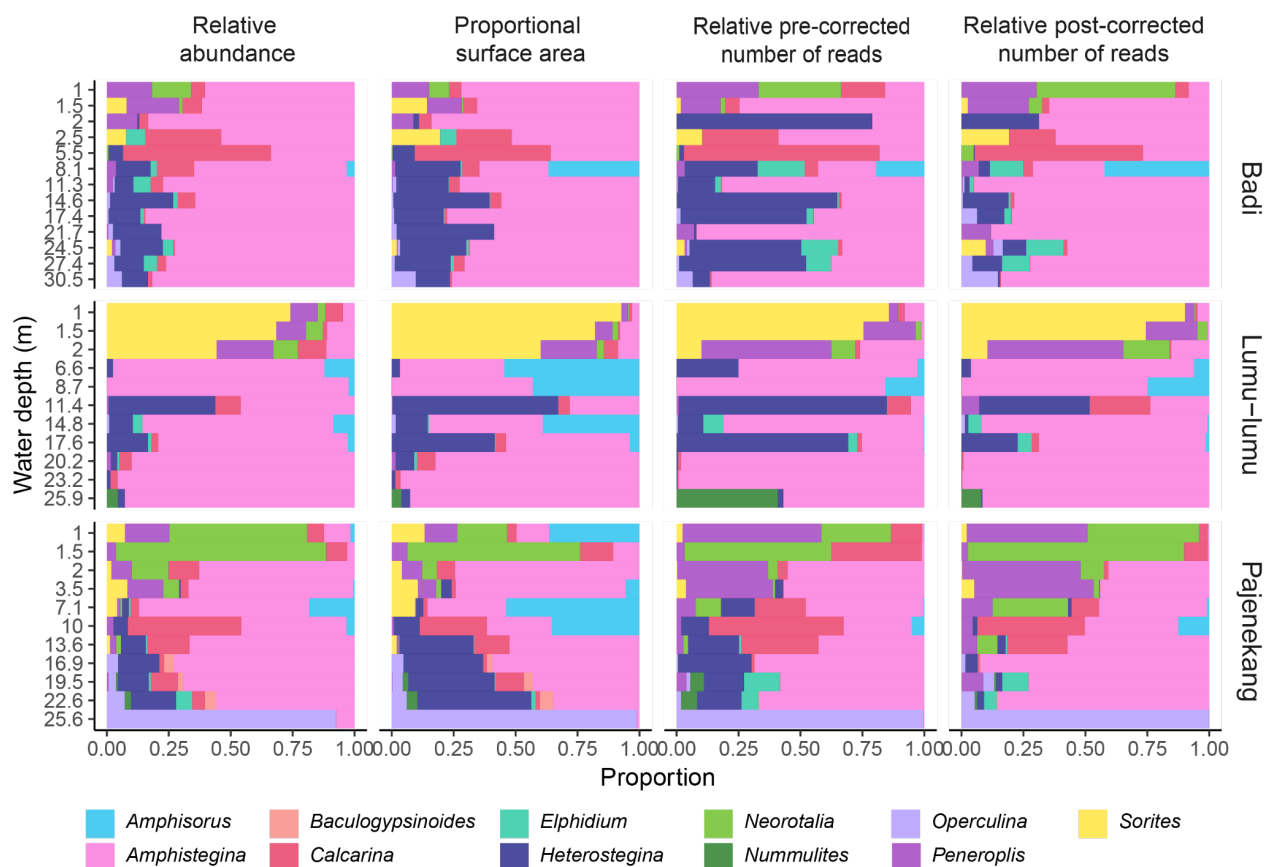


Figure S1. The similarity between all data types is shown by the species distribution along a depth gradient from the three islands studied, Badi, Lumu-lumu and Pajenekang. Proportion of genera for four data types (i.e. relative abundance, proportional surface area, relative pre- and post-corrected number of reads).

MOLECULAR ECOLOGY

RESOURCES

Supplementary tables

Table S1. Sampling details information, including category “Purpose” stating what the sample was used for (ddPCR, Micro-CT scan and metabarcoding).

Table S2. Results from the ddPCR triplicate testing. Results show that the number of gene copies quantified for all three replicates for each sample and for each species have the same order of magnitude. Error scales up in samples with lower number of gene copies. Based on those results, we decided to perform only one replicate per sample.

Table S3. Data processing from the raw sequencing output to the ESV table, using APSCALE. Settings and software versions are included.

Table S4. Custom-built reference sequences for large benthic foraminifera in the Spermonde Archipelago ([Girard et al. 2022](#)).

Table S5. Complete results on the fit of the two models (linear and logarithmic) tested in this study, between the number of gene copies and the measured surface area, including formulas.

Table S6. Welch Two Sample t-test and F-test analysis results showing whether the difference between estimations (raw reads and linearly corrected reads) and measurements (specimen counts and measured surface area) was statistically significant. High significance is shown at $p\text{-value} < 0.01$.

Table S7. Pairwise t-test analysis results showing no significant difference between the proportion of different metrics (specimen counts, measured area, raw reads and linearly corrected reads).

Table S8. Pairwise t-test analysis results showing no significant difference between the proportion of raw reads from unsorted and sorted samples and between the replicates.