

# Validation of a universal set of primers to study animal-associated microeukaryotic communities

Javier del Campo  <sup>1,2,3\*</sup> Maria J. Pons,<sup>4</sup>

Maria Herranz,<sup>1,5</sup> Kevin C. Wakeman,<sup>6,7</sup>

Juana del Valle,<sup>8,9</sup> Mark J. A. Vermeij,<sup>10,11</sup>

Brian S. Leander<sup>1,5</sup> and Patrick J. Keeling<sup>1</sup>

<sup>1</sup>Department of Botany and Biodiversity Research Centre, University of British Columbia, Vancouver, British Columbia, Canada.

<sup>2</sup>Department of Marine Biology and Oceanography, Institut de Ciències del Mar-CSIC, Barcelona, Catalonia.

<sup>3</sup>Department of Marine Biology and Ecology, Rosenstiel School of Marine and Atmospheric Science - University of Miami, 4600 Rickenbacker Causeway, Miami 33149, FL.

<sup>4</sup>Laboratorio de Microbiología Molecular y Genómica Bacteriana, Universidad Científica del Sur, Lima, Peru.

<sup>5</sup>Department of Zoology and Biodiversity Research Centre, University of British Columbia, Vancouver, British Columbia, Canada.

<sup>6</sup>Institute for the Advancement of Higher Education, Hokkaido University, Sapporo 060-0808 Japan.

<sup>7</sup>Graduate School of Science, Hokkaido University, Sapporo 060-0810 Japan.

<sup>8</sup>School of Medicine, Research Center and Innovation of the Health Sciences Faculty, Universidad Peruana de Ciencias Aplicadas, Lima, Peru.

<sup>9</sup>Instituto de Investigación Nutricional, Lima, Peru.

<sup>10</sup>Aquatic Microbiology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Science Park 700, 1098 XH Amsterdam, The Netherlands.

<sup>11</sup>CARMABI Foundation, PO Box 2090, Piscaderabaai z/n, Willemstad, Curaçao.

## Summary

The application of metabarcoding to study animal-associated microeukaryotes has been restricted because the universal barcode used to study microeukaryotic ecology and distribution in the environment, the Small Subunit of the Ribosomal RNA gene

(18S rRNA), is also present in the host. As a result, when host-associated microbial eukaryotes are analysed by metabarcoding, the reads tend to be dominated by host sequences. We have done an *in silico* validation against the SILVA 18S rRNA database of a non-metazoan primer set (primers that are biased against the metazoan 18S rRNA) that recovers only 2.6% of all the metazoan sequences, while recovering most of the other eukaryotes (80.4%). Among metazoans, the non-metazoan primers are predicted to amplify 74% of Porifera sequences, 4% of Ctenophora, and 15% of Cnidaria, while amplifying almost no sequences within Bilateria. *In vivo*, these non-metazoan primers reduce significantly the animal signal from coral and human samples, and when compared against universal primers provide at worst a 2-fold decrease in the number of metazoan reads and at best a 2800-fold decrease. This easy, inexpensive, and near-universal method for the study of animal-associated microeukaryotes diversity will contribute to a better understanding of the microbiome.

## Introduction

Since the publication of the human microbiome consortium paper (Huttenhower *et al.*, 2012), there has been an acceleration in the study of host-associated microbes using metabarcoding methods in many different animal systems. While the term “microbe” encapsulates a wide range of microbial organisms, most studies focus exclusively on bacteria and archaea (McFall-Ngai *et al.*, 2013), some focus on fungi (Nash *et al.*, 2017), and only a few focus on other microeukaryotes (microbial unicellular eukaryotes) like protists (Andersen *et al.*, 2013). However, protists are a part of a healthy microbiota and play a relevant role in the mammal gut ecosystem (including humans) (Scanlan *et al.*, 2014), altering the diversity and composition of the gut communities as a whole (Morton *et al.*, 2015), interacting directly with the host immune system, and even contributing to mucosal immunity conferring disease protection (Chudnovskiy *et al.*, 2016). In other animals, beyond mammals, microeukaryotes contribute to key host processes, like cellulose digestion in termites (Brune, 2014), and in some cases, like the

Received 3 March, 2019; revised 9 May, 2019; accepted 3 July, 2019. \*For correspondence. E-mail jdelcampo@rsmas.miami.edu; Tel. 1 305-421-5092.

zooxanthellae in corals, are crucial for the survival of the host (Glynn, 1993). Apart from these many beneficial roles, microeukaryotes are also well-known parasites, and their action can have dramatic effects on the fitness of the host (Stensvold and van der Giezen, 2018).

There are several reasons why the microeukaryotic members of the microbiome have been largely overlooked so far. For one, bacteria are, in general, the primary or only focus of most microbiological work, not only in host-associated environments but also in other systems (Keeling and Campo, 2017). But in the case of host-associated microeukaryotes, there is an additional technical issue that complicates the study of microeukaryotes using metabarcoding. The most common markers used for metabarcoding are 16S rRNA for bacteria, ITS for fungi and 18S rRNA for the microeukaryotes in general, including fungi. All of them have been applied in free-living environments but their success differs when used in animal-associated ones. The main issue is that the sequences recovered from PCR using universal 18S primers are mostly from the host and not its microeukaryotic microbiome (Parfrey et al., 2014; Wilmes et al., 2017; Wilcox and Hollocher, 2018), sometimes by many orders of magnitude. One solution to this problem is to use blocking primers in the PCR reaction, which are chemically modified primers (with a C3 spacer) that target the host 18S rRNA and will prevent the extension during the PCR when using universal primers. But because metazoans are so diverse, blocking primers have to be specifically designed case by case for each host (Boessenkool et al., 2012; Wilcox et al., 2014; Hino et al., 2016; Belda et al., 2017), so this might be suitable for a subject of intensive study like humans, but not for any broader study of animal microbiome diversity. Another option is to use specific primers for eukaryotes that avoid a particular host (Waidele et al., 2017), this approach appears to be easier than the use of blocking primers, but once again requires a specific set of primers for every metazoan group. Similarly, one particular lineage of eukaryotes can be investigated using primers specific for that group and excluding animals (like those used for fungi), but in this case, no other aspect of the eukaryotic diversity can be assessed.

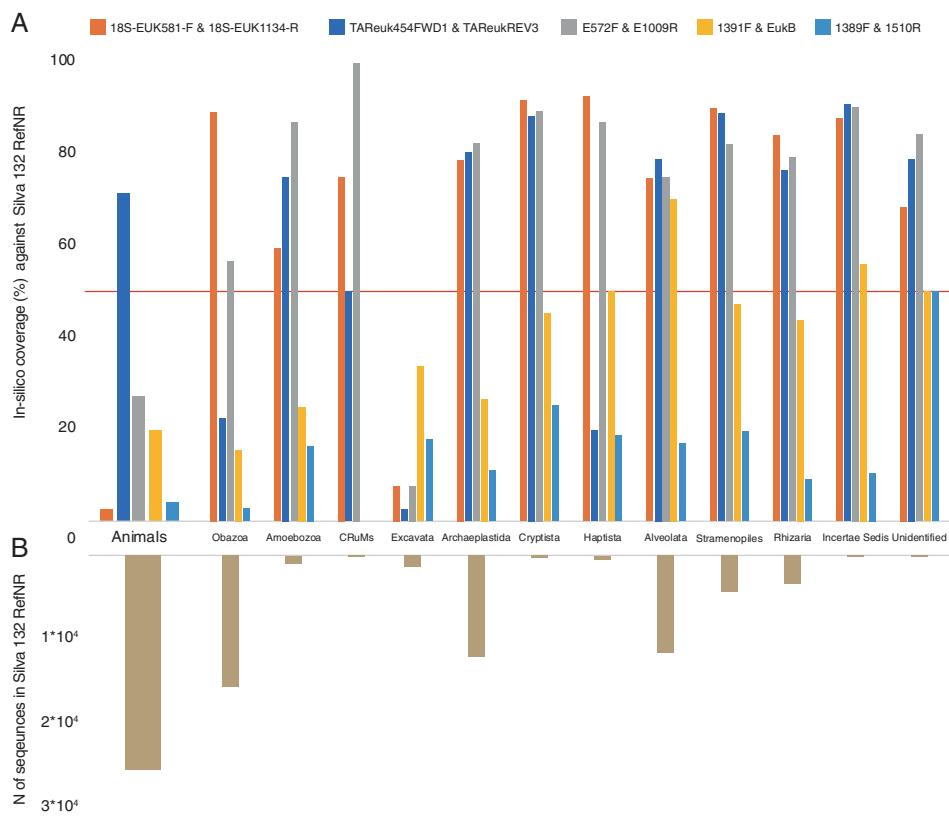
We sought to test a universally applicable approach to recover the eukaryotic component of the microbiome, the eukaryome (Lukeš et al., 2015), from as many animals as possible, and in the easiest and least expensive manner. To this end, we screened the parasitology literature for primers that have been used to detect parasites in animals, and we focused on studies that screened for parasites in a wide range of hosts. In one study, by Bower and colleagues (2004), they assessed primers in "mock communities" consisting of metazoan tissue mixed with parasites from culture and identified a pair of primers

(18S-EUK581-F and 18S-EUK1134-R) to use in what they defined as a universal non-metazoan polymerase chain reaction (UNonMet-PCR). These non-metazoan primers are biased against the amplification of the metazoan signal but they still amplify the rest of eukaryotes. Bower et al. screened for several kinds of parasites in a broad range of metazoans, from mollusks, to nematodes, to fish. We considered these primers a potential candidate for the study of the eukaryome as a whole using a metabarcoding approach and in this study we test them *in silico* and *in vivo*.

## Results

Based on the results of Bower et al. and our own amplifications, the product of the UNonMet-PCR primers was typically 600 bp, and when aligned against the *Saccharomyces cerevisiae* 18S rRNA gene, the amplicon covers the V4 and flanking regions. We first tested the UNonMet-PCR primers *in silico* against the SILVA 18S rRNA Reference Database (Yilmaz et al., 2014). The SILVA database contains all the 18S rRNA sequences longer than 1200 bp present in GenBank, and it is considered a reliable source for probe and primer design. Testing the UNonMet-PCR primers against SILVA recovered only 2.6% of the metazoan 18S rRNA reads that were present, whereas they still performed well with the rest of the Eukaryotic groups (80.4%), except among excavates, which will be discussed below (Fig. 1). Examining the distribution of the metazoan recovered diversity in more detail reveals that most of the reads came from sponges—the primers recover up to 74% of the sponge reads present in SILVA—as well as a small number of ctenophores (4%) and cnidarians (15%). Fewer than 1% of reads from bilaterians are recovered (Fig. 2A).

Next, we compared these results with *in silico* tests of the most commonly used primers in microeukaryotic metabarcoding (Supporting Information Supplementary Table 1). For the comparison, we selected 4 sets of universal 18S primers, based on their relevance in current publications and potential to limit metazoans. TAREuk454FWD1 and TAREukREV3 amplify the V4 region and were used in the European coastal study BioMarkS (Stoeck et al., 2010; Massana et al., 2015) and the global ocean survey Malaspina (Pernice et al., 2016), and have subsequently become the most widely used V4 primers; E572F and E10009R (Amaral-Zettler et al., 2009) are the V9 universal primers suggested by the Earth Microbiome Project (Gilbert et al., 2014); 1391F and EukB also amplify the V9 region and were used by the Tara Oceans Consortium (de Vargas et al., 2015). Both sets of V9 primers are widely used. Additionally, we analysed a set of V4 universal primers E572F and E1009R developed by Comeau et al. (2011) as the published V4 universal primers that recover fewer



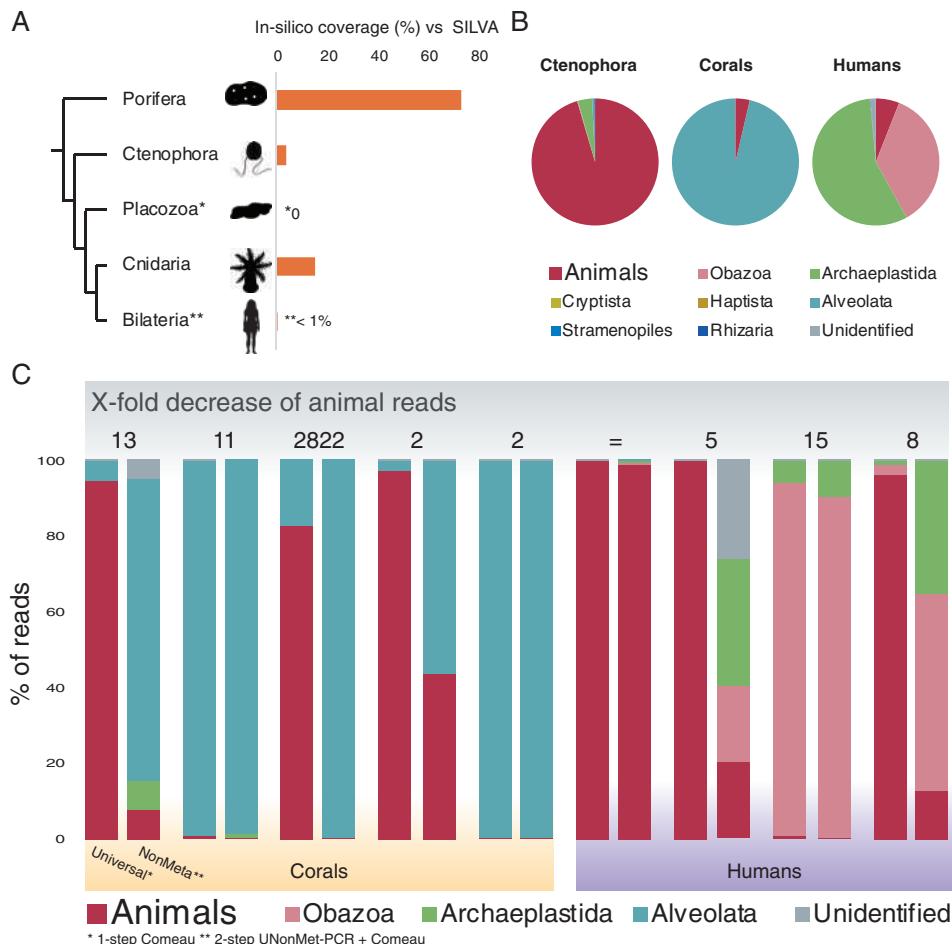
**Fig. 1.** Comparison the *in silico* 18S rRNA diversity recovered from SILVA 132 RefNR using UNonMet-PCR (18S-EUK581-F & 18S-EUK1134-R) primers versus the most used V4 and V9 18S rRNA primers for microeukaryotic metabarcoding.

A. Percentage of recovered diversity from the metazoans and each of the major eukaryotic groups.  
 B. Number of reads for each group present in SILVA 132 RefNR.

metazoan signal. Focusing on the percentage of metazoans reads recovered, the UNonMet-PCR primers recover the fewest and TAReuk recovers the most. The Tara Oceans primers also recover relatively few metazoans; however, their performance with the rest of eukaryotic groups is poor, failing to recover more than 30% of the diversity within most of groups of protists (Fig. 1). Within the rest of the eukaryotes, the UNonMet-PCR primers perform as well or in some cases even better than the analysed universal primers, except in the case of the excavates, where their performance is poor. This is generally the case with V4 universal primers and has been previously reported as a known limitation (Pawlowski *et al.*, 2011).

To see whether these *in silico* analyses reflected real performance *in vivo*, we tested the UNonMet-PCR primers on a range of animal-associated communities. The expected amplicon size (600 bp) is at the high end of the size limit that the most recent iterations MiSeq Illumina technology can handle in the best-case scenario ( $2 \times 300$  bp). Usually the quality of the sequences at the end of the amplicon pairs tends to be low and those fragments are trimmed, so the realistic amplicon size of MiSeq reads is typically around 500 bp ( $2 \times 250$  bp). To enable the merging of paired-ends, we therefore used a two-step approach to proceed with the *in vivo* test. We amplified the samples with the UNonMet-PCR primers in order to reduce the metazoan signal present and

reamplified these products using the primers described by Comeau *et al.* (2011) that only retrieve 20% of metazoans. The samples analysed came from Ctenophora, Cnidaria (corals) and Bilateria (humans). We did not examine Porifera because the *in silico* results clearly suggest they are unlikely to work. The resulting read distribution shows that the primers did not perform well with the two species of ctenophores analysed (5 samples), where most of the recovered signal correspondent to the host (Fig. 2B). In contrast, read mapping from corals (11 samples) and humans (19 samples) resulted in less than 10% of the signal coming from animals (Fig. 2B), allowing us to detect or increase the signal of various groups of microeukaryotes associated with these metazoan hosts, or other eukaryotic signal likely associated with the diet of the analysed individuals (most obviously in the high proportion of plant sequences derived from the human samples). A subset of the coral and human samples tested using the UNonMet-PCR primers were reanalyzed using only the Comeau primers to determine the impact of the non-metazoan primers (Fig. 2C). In 8 out of 9 compared samples, there was a significant decrease in the load of metazoan reads, even in those samples where the initial animal signal was very low. In the worst-case scenario, there was a 2-fold reduction of the metazoan reads, whereas the most dramatic reduction was 2,800 times lower.



**Fig. 2.** A. Percentage of *in silico* metazoan 18S rRNA diversity recovered from SILVA 132 RefNR using UNonMet-PCR primers (18S-EUK581-F & 18S-EUK1134-R). A more detailed analysis at Phylum level is available at Supplementary Table 2.

B. Relative abundance of 18S rRNA reads in Ctenophora (5 samples), Cnidaria (corals) (11 samples) and Bilateria (Humans) (19 samples) using UNonMet-PCR primers.

C. Comparison of the relative abundance of 18S rRNA reads in corals and humans using “1-step Comeau Universal” vs “2-step UNonMet-PCR + Comeau Universal” V4 primers. Details on the origin of the samples can be found at Supplementary Table 3.

## Discussion

Both the *in silico* analysis and *in vivo* results show that the UNonMet-PCR primers work well to significantly reduce the metazoan signal from HTS data sets and constitute a suitable approach to access and study the microeukaryotic communities associated with animal hosts. This pair of primers not only excludes the metazoan signal, but they perform as well or sometimes even better than the most commonly used universal 18S rRNA primers in terms of their ability to amplify V4 from a known range of microeukaryotes. Despite the reduction of metazoan signal, there always remains some host signal that we attribute to non-specific amplification because of the amount of host biomass. The degree of this signal is not sufficient to limit the analysis of microeukaryotes, and indeed we see this as an advantage because it gives an independent identification of the animal host.

The use of non-metazoan primers has important advantages over other methods. First, it eliminates the need to discard most of the sequence data as is the case universal primers are used alone. It is true that sequencing to greater depth can still allow access to eukaryome

diversity, but this is inefficient and expensive (Wilcox and Hollocher, 2018). Second, blocking primers have been successfully used to access the eukaryome, but there is no universal blocking primer that will eliminate all or even most animals, so blocking primers tend to be specific to a particular study (Belda *et al.*, 2017). The UNonMet-PCR primers can be used with most animals, except sponges and perhaps ctenophores, and we show *in vivo* are effective with corals (cnidarian) and humans (bilaterians), two groups that are far apart in the animal tree of life (Dunn *et al.*, 2014). Third, the 18S rRNA fragment recovered by the UNonMet-PCR covers the V4 region, which provides reasonably good phylogenetic resolution, so the barcoding can be more precise than V9-based data (Pawlowski *et al.*, 2011). A better phylogenetic resolution makes the results useful not only for microbial ecology, but eventually also for diagnosis. Fourth, this method is both relatively simple and inexpensive, making it broadly available, and the data accordingly better comparable between different studies.

The main concern about this approach is the need for a two-step amplification. The use of a two-step amplification also known as nested PCR is not new, it has been

used since the 1980s, mostly for pathogen detection, including parasitic eukaryotes (Snounou *et al.*, 1993). Thus, we are familiar with the methodology pitfalls and limitations. A nested PCR may lead to additional sequencing errors (although we should note the second round involves limited rounds of amplification so it is not clear if this is a significant problem), and two stages of potential amplification bias. The bias is an issue with all methods, and generally samples analysed in the same way give a relative picture, while samples treated in different ways are hard to compare, and this method is much the same in that regard (Hadziavdic *et al.*, 2014). However, this is only a limitation because the size of the amplicon generated by the UNonMet-PCR is slightly larger than can be handled by current Illumina technologies (Bradley *et al.*, 2016). It is not unreasonable to assume that these limitations will soon be eliminated by changes in the sequencing platforms and chemistry that will allow the UNonMet-PCR to be used directly for Illumina libraries (or other chemistries) (van Dijk *et al.*, 2014), which would also make the process even simpler and less expensive.

## Conclusions

We have identified and tested a relatively simple solution to an issue that presents a very obvious limitation on microbiome research since the rise of high-throughput metabarcoding was applied to the study of the animal-associated microbial communities. Using the UNonMet-PCR primers will allow research on metazoan host-associated environments to more easily include analyses of eukaryome, providing a much fuller picture of the community and finally begin to fill the knowledge gap in microbiome studies of not only humans, but most animals.

## Methods

### *In silico validation*

The *in silico* test was done using the TestPrime 1.0 (Klindworth *et al.*, 2013) against SILVA 132 RefNR (Quast *et al.*, 2012) allowing no mismatches.

### *Sampling*

Corals were collected from several locations in Curaçao between April 19 and 21, 2015, and Okinawa (Japan) between April 24 and 26, 2015. Whole samples, including skeleton and tissue, were homogenized using mortar and pestle. DNA was extracted with the RNA Powel Soil Total RNA Isolation Kit coupled with the DNA Elution Accessory Kit (MoBio) in the case of the Curaçao samples and DNeasy Blood & Tissue Kit (Qiagen) in the case of the

Japan samples. Ctenophores were collected from Calvert Island (British Columbia, Canada) between June 5 and 7, 2015. The ctenophore gut and body were extracted and homogenized, and DNA was extracted with DNA Power Soil (MoBio). Human stool samples were obtained from children less than 5 years of age who were hospitalized with acute gastroenteritis from January 2011 to April 2014 in the "Hospital Docente Regional de Cajamarca" located in Cajamarca Department. Stool samples from healthy and Cryptosporidium-infected children (6 to 8 years) have been isolated in a screening parasites campaign in school of "Baños del Inca" located also in Cajamarca Department in rural Northern Peru. Genetic material from faecal suspensions was extracted using the High Pure RNA Isolation Kit (Roche Applied Science) in accordance with the manufacturer's instructions. DNA concentration was quantified on a Qubit 2.0 Fluorometer (Thermo Fisher Scientific Inc.). A complete list of samples is available at Supporting Information Supplementary Table 3.

### *18S rRNA amplification and Illumina sequencing*

The expected amplicon size is at the high end of the size limit that the most recent iterations MiSeq Illumina technology can handle in the best-case scenario ( $2 \times 300$  bp). Usually the quality of the sequences at the end of the amplicon pairs tends to be low and those fragments are trimmed, so the realistic amplicon size of MiSeq reads is typically around 500 bp ( $2 \times 250$  bp). To enable the merging of paired-ends, we therefore used a two-step approach to proceed with the *in vivo* test. We amplified the samples with the UNonMet-PCR primers in order to reduce the metazoan signal present and reamplified these products using the primers by Comeau *et al.* (2011). Eukaryotic microbiome amplicons were prepared using PCR with high-fidelity Phusion polymerase (Thermo Fisher Scientific Inc.), using primers that target the V4 region of 18S rRNA gene, but which exclude metazoan sequences (18S-EUK581-F 5'-GTGCCAG CAGCCGCG-3', 18S-EUK1134-R 5'-TTAAGTTTCAG CCTTGC-3') (Bower *et al.*, 2004). PCR was performed using the following protocol: 30 s at 98 °C, followed by 35 cycles each consisting of 10 s at 98 °C, 30 s at 51.1 °C, and 1 min at 72 °C, ending with 5 min at 72 °C. PCR products were visually inspected for successful amplification using gel electrophoresis with 1% agarose gels. PCR products were then cleaned using the QIAquick PCR Purification Kit (Qiagen) and quantified on a Qubit 2.0 Fluorometer. Amplicon sequencing was performed by the Integrated Microbiome Resource facility at the Centre for Comparative Genomics and Evolutionary Bioinformatics at Dalhousie University. PCR amplification from template DNA was performed in duplicate using high-fidelity Phusion polymerase. A single round of PCR

was done using “fusion primers” (Illumina adaptors + indices + specific regions) targeting the V4 region of the Eukaryotic 18S rRNA gene (primer set E572F + E1009R (Comeau *et al.*, 2011);  $\sim$ 440 bp fragment) with multiplexing. PCR products were verified visually by running a high-throughput Invitrogen 96-well E-gel. The duplicate amplicons from the same samples were pooled in one plate, then cleaned-up and normalized using the high-throughput Invitrogen SequalPrep 96-well Plate Kit. The samples were then pooled to make one library, which was then quantified fluorometrically before sequencing on an Illumina MiSeq using a 300 bp paired-end read design. The 18S rRNA amplicon reads are deposited in the European Nucleotide Archive (PRJEB24453, PRJEB29965) and the NCBI Sequence Read Archive (PRJNA482746).

#### Amplicon analysis

Amplicon reads were processed (dereplication, chimera detection and singleton removal) using VSEARCH (Rognes *et al.*, 2016). Operational taxonomic units (OTU) were clustered at 97% identity using VSEARCH and analysed using QIIME v1.9.1 (Caporaso *et al.*, 2010). The taxonomic identity of each OTU was assigned based on the SILVA 132 RefNR database (Yilmaz *et al.*, 2014), using the assign\_taxonomy function in QIIME. OTU that were “unassigned” were inspected using BLAST against the GenBank nr database and manually reassigned to the closest hit if possible. OTU represented by fewer than four reads were removed, as were OTU that were identified as metazoan 18S rDNA or mitochondria. Samples with fewer than 1500 reads were excluded from downstream analysis.

#### Acknowledgements

We thank Forest L. Rohwer, Ben Knowles, Emma E. George and Jan Janouškovec for assistance and advice in coral sample retrieval and processing. This work was funded by a grant from the Canadian Institutes for Health Research (MOP-42517). JdC was supported by a grant from the Tula Foundation to the Centre for Microbial Biodiversity and Evolution and the Marie Curie International Outgoing Fellowship FP7-PEOPLE-2012-IOF-331450 CAARL.

#### References

Amaral-Zettler, L.A., McCliment, E.A., Ducklow, H.W., and Huse, S.M. (2009) A method for studying Protistan diversity using massively parallel sequencing of V9 Hypervariable regions of small-subunit ribosomal RNA genes. *PLoS One* **4**: e6372.

Andersen, L.O., Vedel Nielsen, H., and Stensvold, C.R. (2013) Waiting for the human intestinal Eukaryotome. *ISME J* **7**: 1253–1255.

Belda, E., Coulibaly, B., Fofana, A., Beavogui, A.H., Traore, S.F., Gohl, D.M., *et al.* (2017) Preferential suppression of *Anopheles gambiae* host sequences allows detection of the mosquito eukaryotic microbiome. *Sci Rep* **7**: 3241.

Boessenkool, S., Epp, L.S., Haile, J., Bellemain, E.P., Edwards, M.E., Coissac, E., *et al.* (2012) Blocking human contaminant DNA during PCR allows amplification of rare mammal species from sedimentary ancient DNA. *Mol Ecol* **21**: 1806–1815.

Bower, S.M., Carnegie, R.B., Goh, B., Jones, S.R., Lowe, G. J., and Mak, M.W. (2004) Preferential PCR amplification of parasitic protistan small subunit rDNA from metazoan tissues. *J Eukaryot Microbiol* **51**: 325–332.

Bradley, I.M., Pinto, A.J., and Guest, J.S. (2016) Design and evaluation of illumina MiSeq-compatible, 18S rRNA gene-specific primers for improved characterization of mixed phototrophic communities. *Appl Environ Microbiol* **82**: 5878–5891.

Brune, A. (2014) Symbiotic digestion of lignocellulose in termite guts. *Nat Rev Microbiol* **12**: 168–180.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335–336.

Chudnovskiy, A., Mortha, A., Kana, V., Kennard, A., Ramirez, J.D., Rahman, A., *et al.* (2016) Host-protozoan interactions protect from mucosal infections through activation of the Inflammasome. *Cell* **167**: 444–456.e14.

Comeau, A.M., Li, W.K.W., Tremblay, J.-É., Carmack, E.C., and Lovejoy, C. (2011) Arctic ocean microbial community structure before and after the 2007 record sea ice minimum. *PLoS One* **6**: e27492.

van Dijk, E.L., Auger, H., Jaszczyzyn, Y., and Thermes, C. (2014) Ten years of next-generation sequencing technology. *Trends Genet* **30**: 418–426.

Dunn, C.W., Giribet, G., Edgecombe, G.D., and Hejnol, A. (2014) Animal phylogeny and its evolutionary implications. *Annu Rev Ecol Evol Syst* **45**: 371–395.

Gilbert, J.A., Jansson, J.K., and Knight, R. (2014) The earth microbiome project: successes and aspirations. *BMC Biol* **12**: 69.

Glynn, P.W. (1993) Coral reef bleaching: ecological perspectives. *Coral Reefs* **12**: 1–17.

Hadziavdic, K., Lekang, K., Lanzén, A., Jonassen, I., Thompson, E.M., and Troedsson, C. (2014) Characterization of the 18S rRNA gene for designing universal eukaryote specific primers. *PLoS One* **9**: e87624.

Hino, A., Maruyama, H., and Kikuchi, T. (2016) A novel method to assess the biodiversity of parasites using 18S rDNA Illumina sequencing; parasitome analysis method. *Parasitol Int* **65**: 572–575.

Huttenhower, C., Gevers, D., Knight, R., Abubucker, S., Badger, J.H., Chinwalla, A.T., *et al.* (2012) Structure, function and diversity of the healthy human microbiome. *Nature* **486**: 207–214.

Keeling, P.J., and del Campo, J. (2017) Marine Protists are not just big bacteria. *Curr Biol* **27**: R541–R549.

Klindworth, A., Pruesse, E., Schneemann, T., Peplies, J., Quast, C., Horn, M., and Glöckner, F.O. (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for

classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* **41**: e1.

Lukeš, J., Stensvold, C.R., Jirků-Pomajbíková, K., and Parfrey, L.W. (2015) Are human intestinal eukaryotes beneficial or commensals? *PLOS Pathog* **11**: e1005039.

Massana, R., Gobet, A., Audic, S., Bass, D., Bittner, L., Boute, C., et al. (2015) Marine protist diversity in European coastal waters and sediments as revealed by high-throughput sequencing. *Environ Microbiol* **17**: 4035–4049.

McFall-Ngai, M., Hadfield, M.G., Bosch, T.C.G., Carey, H.V., Domazet-Lošo, T., Douglas, A.E., et al. (2013) Animals in a bacterial world, a new imperative for the life sciences. *Proc Natl Acad Sci* **110**: 3229–3236.

Morton, E.R., Lynch, J., Froment, A., Lafosse, S., Heyer, E., Przeworski, M., et al. (2015) Variation in rural African gut microbiota is strongly correlated with colonization by Entamoeba and subsistence. *PLoS Genet* **11**: e1005658.

Nash, A.K., Auchtung, T.A., Wong, M.C., Smith, D.P., Gesell, J.R., Ross, M.C., et al. (2017) The gut mycobiome of the human microbiome project healthy cohort. *Microbiome* **5**: 153.

Parfrey, L.W., Walters, W.A., Lauber, C.L., Clemente, J.C., Berg-Lyons, D., Teiling, C., et al. (2014) Communities of microbial eukaryotes in the mammalian gut within the context of environmental eukaryotic diversity. *Front Microbiol* **5**: 1–13.

Pawlowski, J., Christen, R., Lecroq, B., Bachar, D., Shahbazkia, H.R., Amaral-Zettler, L.A., and Guillou, L. (2011) Eukaryotic richness in the abyss: insights from Pyrotag sequencing. *PLoS One* **6**: e18169.

Pernice, M.C., Giner, C.R., Logares, R., Perera-Bel, J., Acinas, S.G., Duarte, C.M., et al. (2016) Large variability of bathypelagic microbial eukaryotic communities across the world's oceans. *ISME J* **10**: 945–958.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2012) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**: D590–D596.

Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016) VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**: e2584.

Scanlan, P.D., Stensvold, C.R., Rajilić-Stojanović, M., Heilig, H.G.H.J., De Vos, W.M., O'Toole, P.W., and Cotter, P.D. (2014) The microbial eukaryote *Blastocystis* is a prevalent and diverse member of the healthy human gut microbiota. *FEMS Microbiol Ecol* **90**: 326–330.

Snounou, G., Viriyakosol, S., Zhu, X.P., Jarra, W., Pinheiro, L., do Rosario, V.E., et al. (1993) High sensitivity of detection of human malaria parasites by the use of nested polymerase chain reaction. *Mol Biochem Parasitol* **61**: 315–320.

Stensvold, C.R., and van der Giezen, M. (2018) Associations between gut microbiota and common luminal intestinal parasites. *Trends Parasitol* **34**: 369–377.

Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M.D.M., Breiner, H.-W., and Richards, T.A. (2010) Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol Ecol* **19**: 21–31.

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahe, F., Logares, R., et al. (2015) Eukaryotic plankton diversity in the sunlit ocean. *Science* **348**: 1261605.

Waidele, L., Korb, J., Voolstra, C.R., Künzel, S., Dedeine, F., and Staubach, F. (2017) Differential ecological specificity of Protist and bacterial microbiomes across a set of termite species. *Front Microbiol* **8**: 1–13.

Wilcox, J.J.S., and Hollocher, H. (2018) Unprecedented Symbiont eukaryote diversity is governed by internal trophic webs in a wild non-human primate. *Protist* **169**: 307–320.

Wilcox, T.M., Schwartz, M.K., McKelvey, K.S., Young, M.K., and Lowe, W.H. (2014) A blocking primer increases specificity in environmental DNA detection of bull trout (*Salvelinus confluentus*). *Conserv Genet Resour* **6**: 283–284.

Wilmes, P., Wampach, L., Heintz-Buschart, A., Hogan, A., Muller, E.E.L., Narayanasamy, S., et al. (2017) Colonization and succession within the human gut microbiome by Archaea, bacteria, and microeukaryotes during the first year of life. *Front Microbiol* **8**: 1–21.

Yilmaz, P., Parfrey, L.W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., et al. (2014) The SILVA and “all-species living tree project (LTP)” taxonomic frameworks. *Nucleic Acids Res* **42**: D643–D648.

## Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

**Supplementary Table 1.** List of primers tested, region, amplicon size, publication and SILVA 132 RefNr coverage.

**Supplementary Table 2.** Percentage of *in silico* metazoan 18S rRNA diversity recovered from SILVA 132 RefNr using the UNonMet-PCR primers at Phylum level.

**Supplementary Table 3.** List of samples including basic taxonomic information and the sampling site.