

METHODOLOGY ARTICLE

Open Access



Unveiling protist diversity associated with the Pacific oyster *Crassostrea gigas* using blocking and excluding primers

Camille Clerissi^{1,2*} , Laure Guillou³, Jean-Michel Escoubas⁴ and Eve Toulza¹

Abstract

Background: Microbiome of macroorganisms might directly or indirectly influence host development and homeostasis. Many studies focused on the diversity and distribution of prokaryotes within these assemblages, but the eukaryotic microbial compartment remains underexplored so far.

Results: To tackle this issue, we compared blocking and excluding primers to analyze microeukaryotic communities associated with *Crassostrea gigas* oysters. High-throughput sequencing of 18S rRNA genes variable loops revealed that excluding primers performed better by not amplifying oyster DNA, whereas the blocking primer did not totally prevent host contaminations. However, blocking and excluding primers showed similar pattern of alpha and beta diversities when protist communities were sequenced using metabarcoding. Alveolata, Stramenopiles and Archaeplastida were the main protist phyla associated with oysters. In particular, *Codonellopsis*, *Cyclotella*, *Gymnodinium*, *Polarella*, *Trichodina*, and *Woloszynskia* were the dominant genera. The potential pathogen *Alexandrium* was also found in high abundances within some samples.

Conclusions: Our study revealed the main protist taxa within oysters as well as the occurrence of potential oyster pathogens. These new primer sets are promising tools to better understand oyster homeostasis and disease development, such as the Pacific Oyster Mortality Syndrome (POMS) targeting juveniles.

Keywords: Holobiont, Metabarcoding, Microbiome, Ostreoida, *Crassostrea*

Background

The farmed oyster *Crassostrea gigas* is affected by the Pacific Oyster Mortality Syndrome (POMS) targeting juveniles [1, 2]. This disease is multifactorial due to abiotic factors (water temperature, seawater quality) [3], and biotic factors (development stage, interactions with microorganisms) [4, 5]. Among biotic factors, some are microbial pathogens (*Ostreid herpesvirus* OsHV-1 μ Var, *Vibrio* spp.), but other microbes

might play a positive role [6]. Indeed, oysters are associated with various microorganisms (called microbiota), and recently resistance to disease was linked to characteristics of the prokaryotic compartment [4].

Most analyses of oyster microbiota corresponded to prokaryotes, but very little is known concerning microbial eukaryotes (hereafter named protists). This lack of knowledge is mainly explained by methodological issues. Indeed, while the 16S rRNA gene is successfully used to characterize prokaryotic assemblages in metabarcoding surveys, the related rRNA marker gene for eukaryotes (18S) mostly amplified the abundant host DNA rather than associated protists. However, protists might play a role in oyster homeostasis. For example, *C. gigas* oysters

* Correspondence: camille.clerissi@ephe.psl.eu

¹IHPE, Univ. Montpellier, CNRS, Ifremer, Univ. Perpignan Via Domitia, Perpignan, France

²PSL Université Paris: EPHE-UPVD-CNRS, USR 3278 CRIOBE, Université de Perpignan, 52 Avenue Paul Alduy, 66860 Perpignan Cedex, France

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

are infected by several protistan parasites [7–11]. Among them, *Marteilia* and *Pseudoperkinsus* genera have negative impacts on the mollusc aquaculture production [12, 13]. Moreover, other protists such as *Alexandrium minutum* are detrimental for oysters [14], but also for human health [15].

A universal non-metazoan (UNonMet) primer set was developed in 2004 [16], but the expected product size (~ 600 bp) exceeded the size limit of MiSeq Illumina technology (2×300 bp, requiring overlap between read pairs). Recently, a study highlighted high performances of this primer set using in silico analyses, and proposed to use nested PCR (i.e., two-step PCR that consists in amplifying a shorter amplicon after a first PCR using the UNonMet primers) to tackle the amplicon size issue [17]. Although nested PCR might bias the abundance of environmental sequences, the comparison of UNonMet with universal primer sets (non-nested PCR) revealed however similar protist assemblages in this study.

Instead of developing and/or using primer sets that directly exclude the amplification of metazoans, another strategy would use a combination of a universal primer set and a blocking primer that specifically targets a region of the universal reverse primer. Blocking primers are modified at the 3'-end with a Spacer C3 CPG (3 hydrocarbons), thus the elongation is prevented during PCR and the targeted sequences are not amplified. Such an approach has the advantage of being very specific (excluding only sequences similar to the blocking primer), and has proven to be effective in the study of fish and krill gut contents [18, 19], coral-associated protists [20], and in the removal of metazoa sequences from seawater community samples [21].

As a consequence, in order to reveal protist diversity associated with *C. gigas* oysters, we developed a blocking primer associated with a commonly used primer set targeting the V4 loop of the 18S rRNA gene [22–28]. First, we computed in silico analyses to compare the blocking primer to the UNonMet primer set, but also to a primer set predicted to exclude most *C. gigas* sequences [20]. Secondly, we performed metabarcoding sequencing using a heterogeneous dataset of oyster samples (multiple origins and transplantations). This in vivo comparison was not done using the UNonMet primer set, because of issues related to large amplicon sizes when this study was done. This study aims at comparing different primer sets to describe protist diversity within *C. gigas*

microbiota, and to discuss advantages and disadvantages of these different types of primers.

Results

In silico specificity of blocking and excluding primers

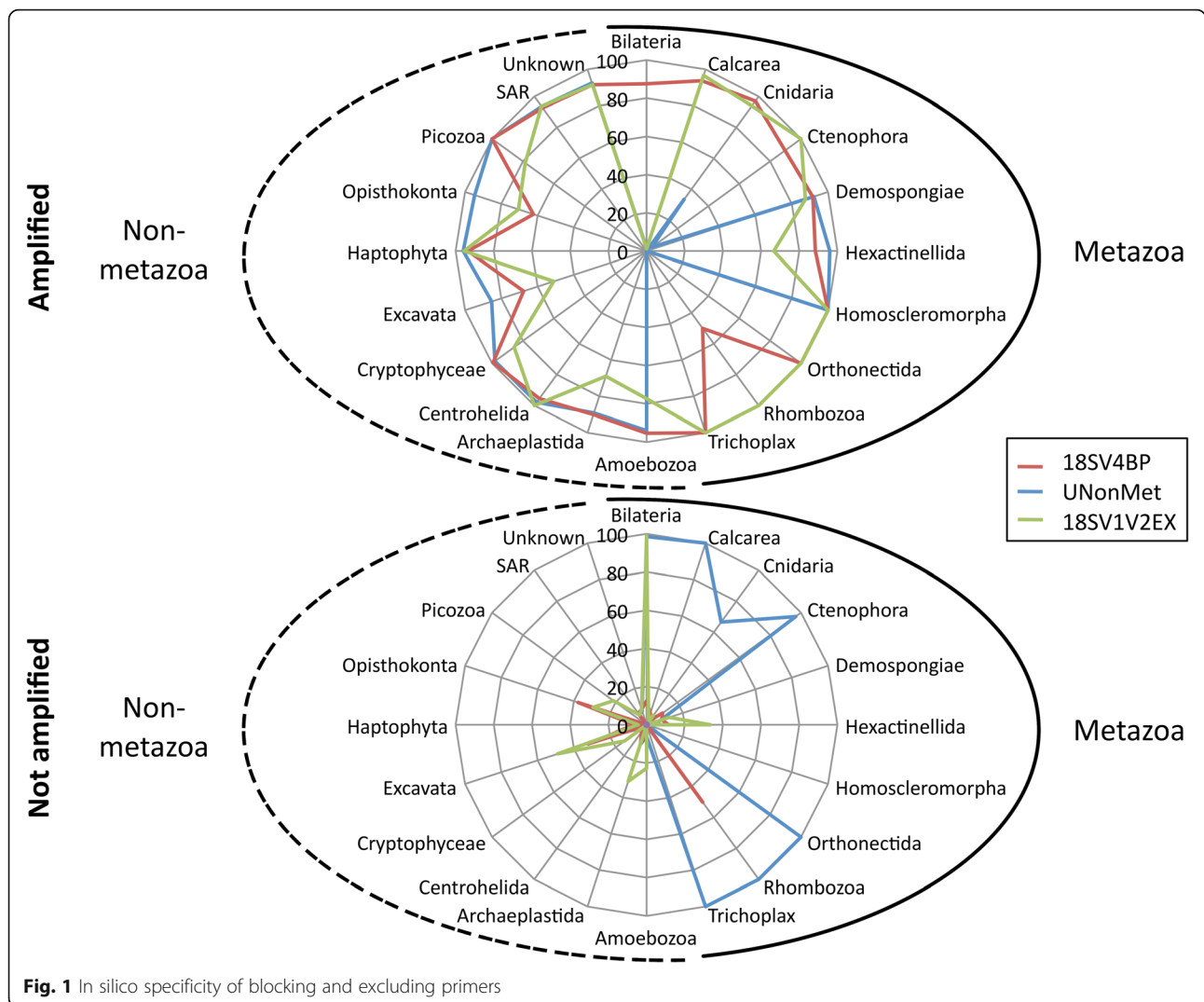
In order to describe protist diversity associated with oysters, a preliminary sequencing test was performed for a sample of *C. gigas* using a commonly used primer set targeting the V4 region of the 18S rRNA gene (Table 1). Because primers for 18SV4 were designed previously to amplify all eukaryotes [22], these sequencing tests showed as expected an excess of amplicons from *C. gigas*, representing ~ 99.7% of sequences (for a total of 2696 cleaned sequences).

Thus we designed blocking primers that specifically target the *Crassostrea* genus (Table 1) using the Silva SSU database (see Methods for more details). Briefly, to estimate the specificity of this blocking primer, we identified sequences of the Silva database that matched with both the primer set and the blocking primer. Only Ostreoida (oyster order) sequences from four genera (*Crassostrea*, *Hyotissa*, *Ostrea*, and *Saccostrea*) were removed by the blocking primer. Although not all (82%) Ostreoida sequences matched with the blocking primer, we found a very high in silico specificity for the *Crassostrea* genus (100%). In addition, this blocking primer did not remove protists found in the Silva database. Furthermore, we estimated the specificity of this blocking primer in association with the 18SV4 primer set (Table 1) against metazoan and non-metazoans sequences from the SSU Silva database. The term 18SV4BP was used hereafter for sequences obtained using the blocking primer targeting 18SV4. While the blocking primer excluded four Ostreoida genera, biases were identified for the 18SV4 primer set, particularly for Excavata and Opisthokonta (Fig. 1).

Then we compared 18SV4BP to two already published excluding primer sets: the UNonMet primers (also targeting the V4 loop of the 18S rRNA gene) [16], and a primer set (18SV1V2EX) that targeted the variable loops V1 and V2 of the 18S rRNA gene, predicted to prevent amplification of *C. gigas* (according to sequence comparison between primers and *C. gigas*) [20]. For metazoan sequences, these analyses highlighted that (i) UNonMet performed well to exclude most metazoan sequences (except for Cnidaria, Demospongiae, Hexactinellida, and Homoscleromorpha), and (ii) 18SV1V2EX mostly excluded Bilateria. For non-metazoan sequences,

Table 1 Blocking and excluding primers used for metabarcoding analyses

Marker region	Target	Forward (5' -> 3')	Reverse (5' -> 3')	Blocking primer (5' -> 3')
18SV4	Eukaryota	CCAGCASCYGGGTAATTC	ACTTTCGTTCTTGATYRA	TCTTGACTAATGAAAACATGCTTG
18SV1V2	Non-metazoa	ACCTGGTTGATCCTGCCA	GTARKCCWMTAYMYTACC	No blocking primer



18SV4BP performed better than 18SV1V2EX, but UNonMet tended to amplify a higher diversity than the two other primers.

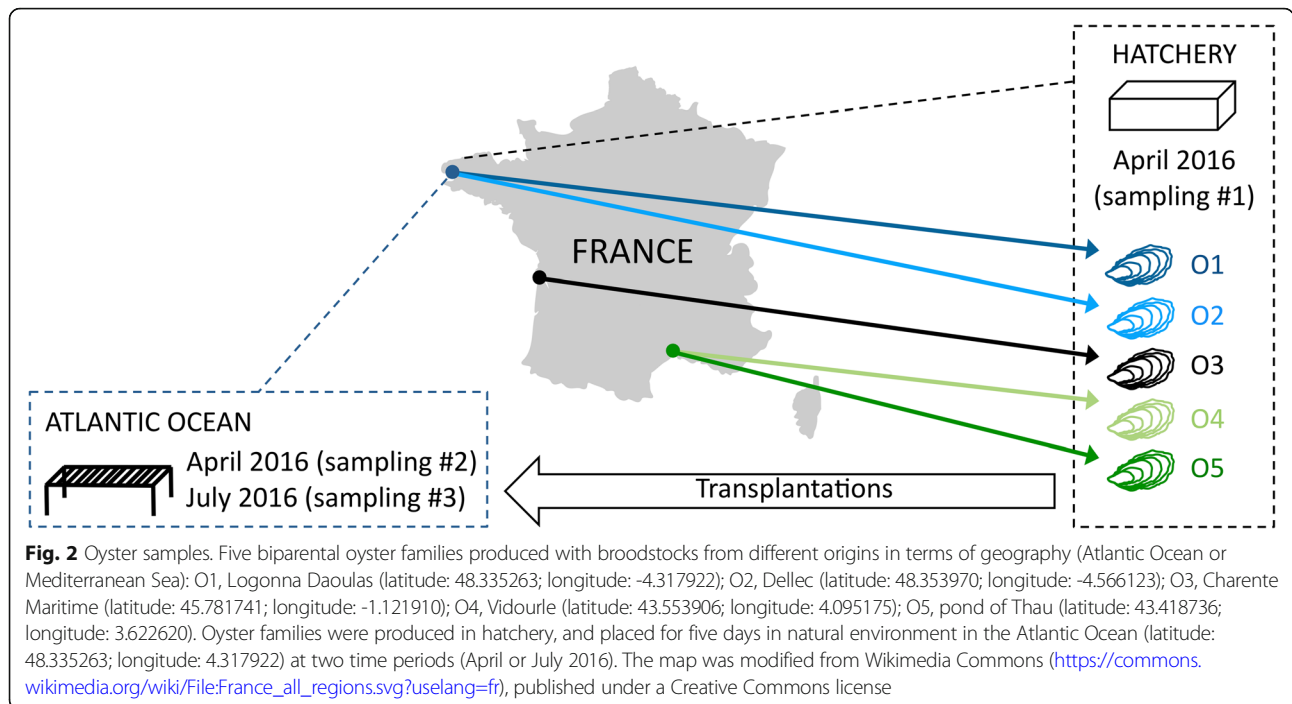
To conclude, in silico analyses of blocking and excluding primers revealed that UNonMet was powerful to describe oyster-associated protists. Moreover, the blocking primer 18SV4BP was highly specific and removed only oyster sequences. While 18SV1V2EX excluded *Crassostrea* sequences, they might also exclude some protist groups (Amoebozoa, Archaeplastida, Cryptophyceae, Excavata, Opisthokonta, and Picozoa).

Metabarcoding analyses of protist assemblages using biparental oysters families

Although the UNonMet excluding primers performed well in the in silico analyses, the amplicon size (~ 600 bp) was still a limitation for amplicon sequencing using MiSeq when this study was done. As a consequence, we

only performed in vivo comparisons between 18SV4BP and 18SV1V2EX. Five biparental oyster families (O1-O5) of *C. gigas* were used in this study (Fig. 2 and Additional file 1: Table S1). They were produced within a hatchery (Argenton, France) using genitors from different origins (Atlantic Ocean or Mediterranean Sea). Protist assemblages were sampled from the first oyster generation kept in the hatchery (sampling #1), or placed in the environment at two different time periods (sampling #2 and #3). This dataset was thus heterogeneous and represented a high diversity of oyster-associated protists in order to compare blocking and excluding primer sets.

First, we compared both marker regions for the abundances of total sequences at high taxonomic ranks: oyster, Embryophyceae, protists, and others (i.e., other metazoa and multi-affiliation). Similar numbers of sequences were obtained using 18SV1V2EX and 18SV4BP



(Table 2). However, while host DNA was missing from 18SV1V2EX dataset, the use of 18SV4BP still displayed host contaminations (Fig. 3, Additional file 2: Fig. S1 and Additional file 3: Table S2). Accordingly, protist fractions were significantly higher for 18SV1V2EX than 18SV4BP (Table 2).

Nucleotide and alpha diversities of protist sequences

In order to perform rigorous comparisons of protist sequences between 18SV1V2EX and 18SV4BP, we decided to keep only samples having more than 5160 protist sequences. Moreover, the dataset was rarefied to this minimal value. While the whole 18SV1V2EX samples had more than 5160 protist sequences, only 34 out of 54 18SV4BP samples were kept for subsequent analyses (Additional files 4, 5, 6: Fig. S2, Tables S3-S4).

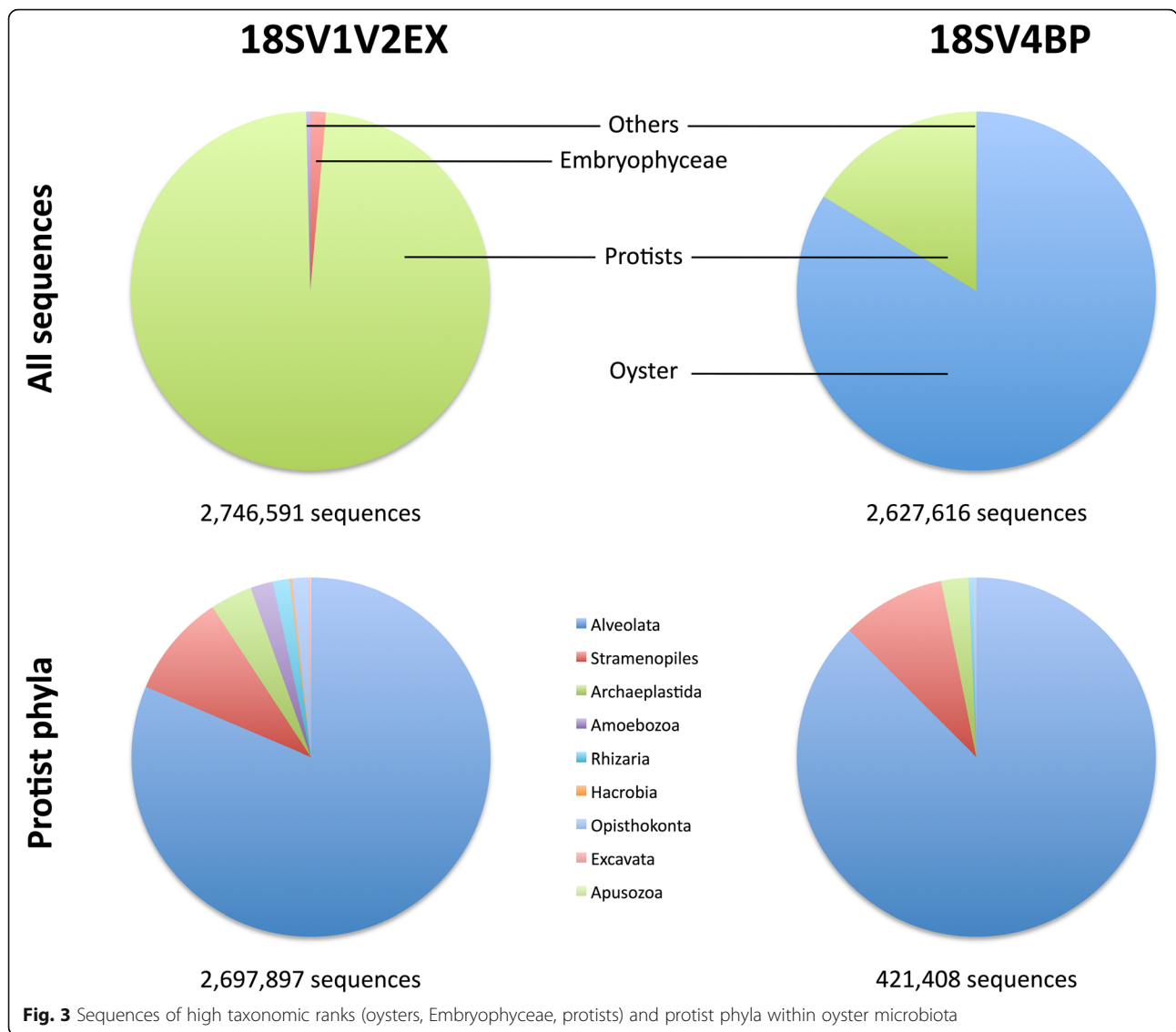
Table 2 Comparison of the abundances of total sequences and high taxonomic groups between 18SV1V2EX and 18SV4BP. Numbers into brackets are *p*-values. NS: not significant. 18SV1V2EX and 18SV4BP indicate which marker had significantly higher values

Taxa	18SV1V2EX vs. 18SV4BP
Total sequences	NS (<i>P</i> = 0.211)
Oyster	18SV4BP (<i>P</i> < 0.001)
Embryophyceae	18SV1V2EX (<i>P</i> < 0.001)
Protists	18SV1V2EX (<i>P</i> < 0.001)
Others	18SV4BP (<i>P</i> < 0.001)

The analysis of protist sequences showed that 18SV4BP amplicons (377 bp) were longer than 18SV1V2EX (304 bp) in average (Additional file 7: Fig. S3), and that nucleotide diversity of 18SV1V2EX (0.152) was higher than 18SV4BP (0.090). It suggested that although amplicons of 18SV1V2EX were shorter than 18SV4BP, the obtained diversity was higher for 18SV1V2EX. Because the same samples were analyzed using 18SV1V2EX or 18SV4BP, these observations might be the result of either a better resolution of the V1 and V2 loops of 18S rRNA gene or amplification of a more diverse set of protists. Furthermore, we compared both 18SV1V2EX and 18SV4BP for different alpha diversity indices (Additional file 8: Table S5), but no significant differences were found between both markers (Fig. 4).

Composition of protist communities within oyster microbiota

First, we used the rarefied dataset and clustering analyses to study protist assemblages at the OTU scale. Both 18SV1V2EX and 18SV4BP showed similar patterns for protist assemblages (Fig. 5), and Bray-Curtis dissimilarities were significantly correlated ($r = 0.86$, $p = 0.001$, Mantel test). In particular, we found that protist assemblages were significantly linked to environmental conditions (S1-S3) ($p = 0.001$) rather than oyster families (O1-O5) ($p > 0.4$) for both 18SV4BP and 18SV1V2EX using PERMANOVA. Mantel tests were then computed at each taxonomic rank (from genus to phylum), and suggested that 18SV1V2EX and 18SV4BP mostly gave similar results (Fig. 6). For example, this study revealed that the main protist phyla were



Alveolata, Stramenopiles and Archaeplastida (Fig. 3). Nevertheless, the lower correlation between 18SV1V2EX and 18SV4BP was obtained at the genus level ($r = 0.44$, $p = 0.001$, Mantel test). Surprisingly, both marker regions identified different dominant genus both within Alveolata (Additional files 5-6: Tables S3-S4). Indeed, *Codonellopsis* was the most abundant genus for 18SV1V2EX (10.30% and undetected for 18SV1V2EX and 18SV4BP, respectively), whereas *Woloszynskia* dominated protist assemblages using 18SV4BP (0.22 and 46.92% for 18SV1V2EX and 18SV4BP, respectively).

Dominant genera within protist assemblages

We computed heatmaps for both genetic markers to describe the distribution of dominant genera within oyster samples (Fig. 7). In addition to *Codonellopsis* and

Woloszynskia, the main genera were *Cyclotella*, *Gymnodinium*, *Polarella*, and *Trichodina*. Notably, the potential pathogen *Alexandrium* was also found in high abundances within some samples (Fig. 7), and we also identified the potential pathogen *Pseudoperkinsus* at much lower abundances (Additional file 6: Table S4). Then, we compared the nucleotide sequences of these dominant genera to the nucleotide collection of NCBI using BLASTn, and we computed phylogenetic reconstructions using the first hits (Fig. 8). These analyses revealed the phylogenetic diversity of protists within oysters, and particularly that *Pseudoperkinsus* and *Trichodina* genera were similar to isolates already identified within different bivalve species (*Adipiocola pacifica*, *Crassostrea gigas*, *Mizuhopecten yessoensis*, *Mytilus* sp., and *Venerupis philippinarum*) (Table 3).

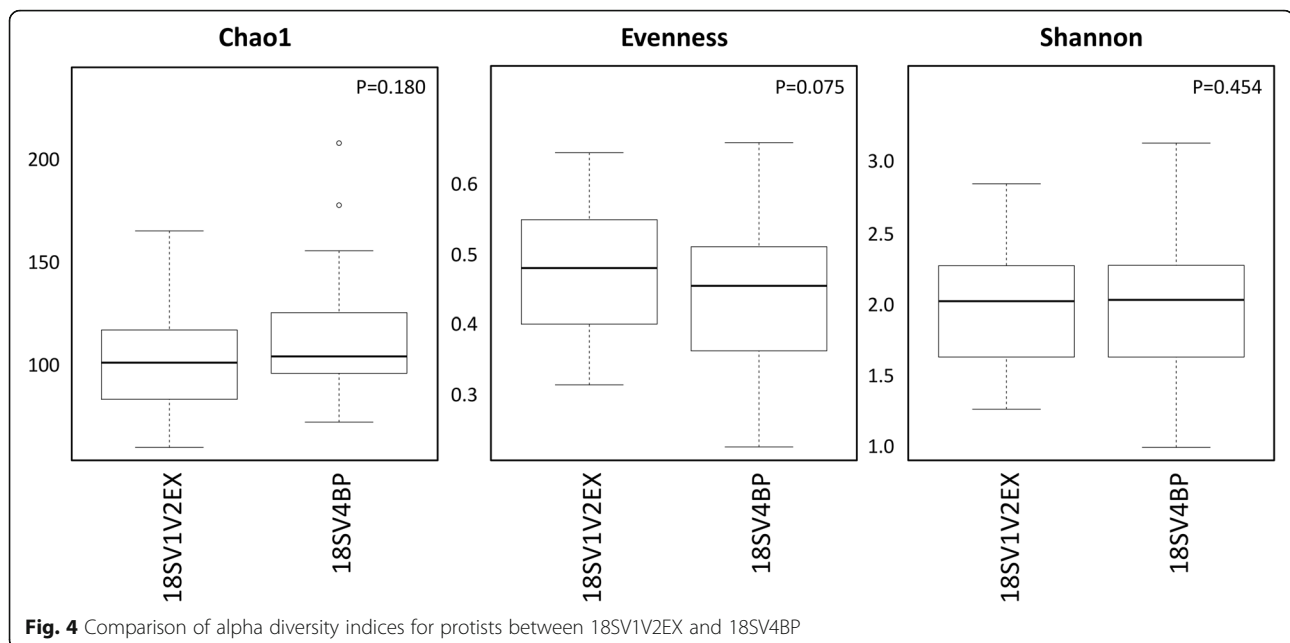


Fig. 4 Comparison of alpha diversity indices for protists between 18SV1V2EX and 18SV4BP

Discussion

Advantages and disadvantages of blocking and excluding primers

The UNonMet primers were previously designed to exclude most metazoan sequences [16]. Our in silico analyses validated the estimated high performances of this primer set [17]. UNonMet primers had potentially two main limitations (Table 4). First, expected amplicon size was not compatible with current sequencing technology for metabarcoding. Secondly, while in silico analyses suggested that this primer set performed well, it was proposed that excluding primers might preclude as well the amplification of unexpected taxa, in particular from taxa not described so far [18].

Blocking primers are highly specific and prevent the amplification of a small range of sequences [18]. As a consequence, we decided to design blocking primer (18SV4BP) to study oyster microbiota, and to compare it in vivo (i.e., through metabarcoding analyses) to a primer set that was expected to exclude *C. gigas* sequences (18SV1V2EX) [20]. In silico analyses highlighted that the blocking primer was expected to prevent amplification of most Ostreoida and *Crassostrea* DNA. However, various efficiencies were observed for the different samples when we performed PCR. On average, oyster sequences still represented 83% (from 56 to 99%). Such variations were already observed in previous studies based on blocking primers [18–20], and they might be mainly related to the ratio between host and total DNA. In contrast, most host sequences were not amplified when using the excluding 18SV1V2EX primer set (<1%). However, according to in silico analyses, we found that many non-metazoans (e.g.,

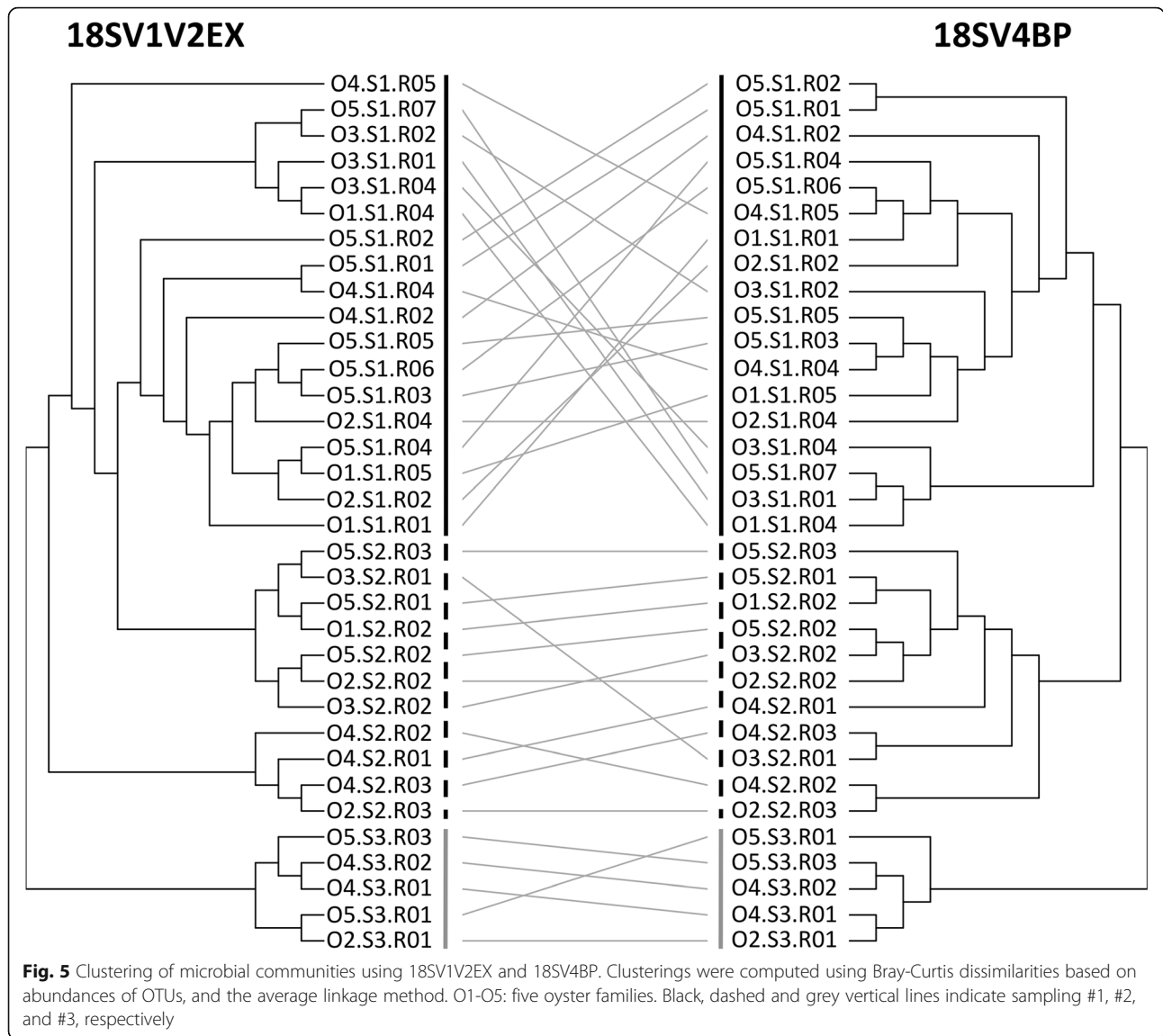
Archaeplastida, Excavata, and Picozoa) might not be amplified in comparison to 18SV4BP.

Altogether, these analyses revealed advantages and disadvantages for the three primer sets (Table 4). While blocking primers are less expected to exclude other sequences than the targeted ones [18], they did not remove all host DNA. In contrast, while excluding primers might possibly exclude unexpected taxa, the metabarcoding diversity was similar to blocking primers.

Protist assemblages within oyster microbiota

Both blocking (18SV4BP) and excluding (18SV1V2EX) primers showed similar patterns of protist diversity using metabarcoding sequencing. No significant differences were found for alpha diversity indices, and protist compositions were significantly correlated between both primer sets at each taxonomic rank (from OTU to phylum). The major difference between protist assemblages was found at the genus level. In particular, *Codonellopsis* was dominant for 18SV1V2EX, but absent from 18SV4BP dataset. Surprisingly, this genus was already identified in *C. gigas* using morphology [29]. Because both 18SV1V2EX and 18SV4 are degenerated primer sets (Table 1), they might amplify differently some protist groups. Differences between marker regions were already observed in previous metabarcoding studies. For example, different protist assemblages were obtained between 18SV4 and 18SV9 marker regions when they were used to study coastal phytoplankton [22].

Alveolata, Stramenopiles and Archaeplastida were the main phyla within oyster microbiota. Other



studies highlighted that these groups were highly abundant in seawater [30, 31]. Thus, it was difficult to conclude if these phyla were resident within oyster microbiota, or represent only environmental protists obtained through filtration activity [32]. However, potential oyster pathogens were identified in this study, such as the *Alexandrium* and *Pseudoperkinsus* genera. In addition, another dominant genus was already found to affect oysters, such as *Gymnodinium* that might cause oyster tissue injuries [33]. Lastly, BLASTn searches revealed that the *Trichodina* genus was similar to an isolate already identified within another bivalve species (*Mizuhopecten yessoensis*). Overall, these results highlighted the potential of these primer sets to study the whole eukaryotic microbes within oyster microbiota, but also diseases caused by protists.

Conclusions

To conclude, we developed a blocking primer to study eukaryotic microbes within oyster microbiota, and we compared it to excluding primers. We found that the three primer sets had advantages and disadvantages, but they offered the possibility of targeting a compartment that was rarely described so far in most known microbiota. As a consequence, these primers are promising tools to better understand oyster homeostasis and disease development, such as the Pacific Oyster Mortality Syndrome (POMS) targeting juveniles.

Methods

Blocking and excluding primers

Blocking primers were designed for 18SV4 primer sets (Table 1) in order to reduce the proportion of host sequences. First, we downloaded the non-redundant (99%)

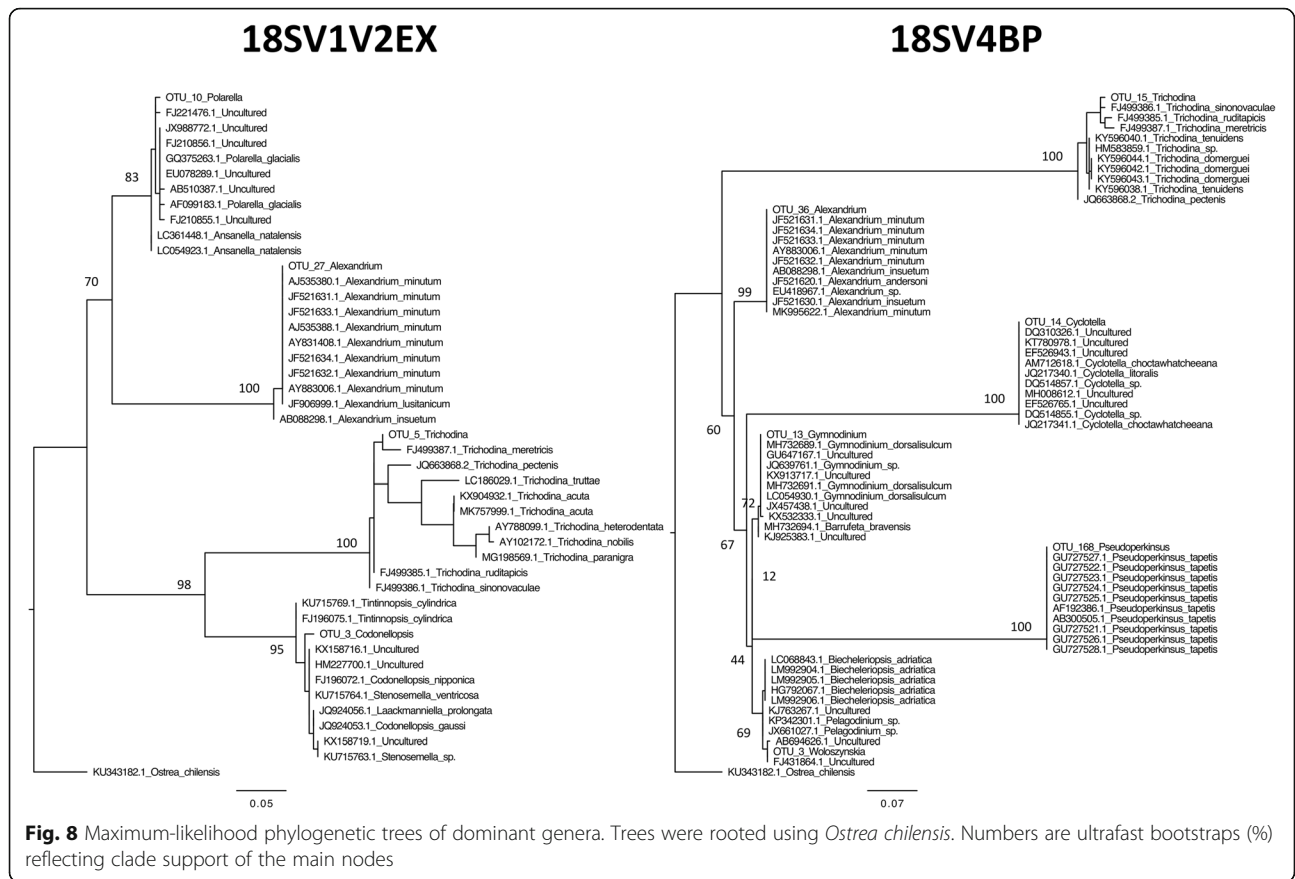


Table 3 Host annotations for BLASTn searches of dominant protist genera (NCBI)

Accession number	Description	Host	Marker
JX661027.1	<i>Pelagodinium</i> sp.	<i>Acanthochiasma</i> sp. (Radiolaria)	18SV4BP
AB300505.1	<i>Pseudoperkinsus tapetis</i>	<i>Adipicola pacifica</i> (Mussel)	
GU727525.1		<i>Crassostrea gigas</i> (Oyster)	
GU727528.1			
GU727524.1		<i>Katharina tunicata</i> (Chiton)	
GU727527.1		<i>Leptosynapta clarki</i> (Sea cucumber)	
GU727522.1		<i>Mytilus</i> sp. (Mussel)	
GU727523.1			
GU727526.1		<i>Phascolosoma agassizii</i> (Peanut worm)	
GU727521.1		<i>Venerupis philippinarum</i> (Clam)	
KY596042.1	<i>Trichodina domerguei</i>	<i>Gasterosteus aculeatus</i> (Teleostei)	
KY596043.1			
KY596044.1			
KY596038.1	<i>Trichodina tenuidens</i>		
KY596040.1			
JQ663868.1	<i>Trichodina pecten</i>	<i>Mizuhopecten yessoensis</i> (Bivalvia)	18SV4BP/ 18SV1V2EX
HM583859.1	<i>Trichodina</i> sp.	<i>Salmo salar</i> (Teleostei)	18SV4BP

Table 4 Advantages and disadvantages between blocking and excluding primers to study protists within oyster microbiota

	18SV4BP	UNonMet	18SV1V2EX
Advantages	BP specifically targets host sequences	No host sequences are expected to be amplified Protist diversity is expected to be well represented	No host sequences were amplified High nucleotide polymorphism compared to 18SV4BP
Disadvantages	Host sequences were not completely removed Low nucleotide polymorphism compared to 18SV1V2EX	Excluding primers might not only exclude metazoans, but also other unexpected taxa Expected amplicon size is too large for current Illumina MiSeq technology	

primer) of *Crassostrea* sequences with the non-metazoan database using MAFFT (default parameters) [36]. According to previous studies, we designed several blocking primers having less than 30 bp, with 10 bp overlapping the reverse primer, and having a T_m similar to the targeted primer set [18, 19]. The best candidate was finally identified using specificity tests against metazoan and non-metazoan databases (i.e., targeting only oysters and no non-metazoans). Lastly, this primer was synthesized and modified at the 3'-end with a Spacer C3 CPG (3 hydrocarbons) [18].

In addition, we compared the specificity of 18SV4BP to two already published excluding primer sets: the UNonMet primers [16], and a primer set that targeted the variable loops V1 and V2 of the 18S rRNA gene [20]. To estimate the specificity of the three primer sets, we used the non-redundant (99%) Silva SSU database (release 128, September 2016). First, we randomly selected a sequence that matched with the three primer sets, and we used it as a query to identify sequences that matched with each amplicon in the Silva SSU database using BLASTn [37] (evaluate $< 10^{-5}$). Secondly, BLASTn subjects were then aligned with the query sequence using MAFFT (default parameters) [36] to find sequences having the complete amplicon regions. Thirdly, we compared sequence annotations (metazoa and non-metazoa) between sequences that matched or not with the different primer sets.

Biological material

Five biparental oyster families (hereafter named O1-O5) of *Crassostrea gigas* were used in this study (Additional file 1: Table S1). They were produced at the hatchery (Argenton, France) using a methodology that allowed the production of pathogen-free juveniles (please see reference [4] for more details). Individuals of each oyster family were either kept at the hatchery (please see reference [38] for more details) (sampling #1, Additional file 1: Table S1), or placed in the natural environment (Atlantic Ocean, latitude: 48.335263; longitude: 4.317922) for 5 days in April 2016 (sampling #2) or in July 2016

(sampling #3). Oysters were sampled, flash frozen in liquid nitrogen, and stored at -80°C .

DNA extraction, PCR and sequencing

DNA extractions from frozen oysters were done using the DNA from tissue Macherey-Nagel kit (ref. 740, 952.250) according to the manufacturer's protocol (please see reference [38] for more details).

The rRNA genes were amplified and sequenced using the 18S variable V1V2 and V4 loops for eukaryotic communities (Table 1) [22, 39]. PCR reactions were carried in a 25 μl volume with final concentrations of 0.4 μM of each PCR primers, 0.02 U of the Qiagen Hotstar Taq DNA Polymerase, 0.2 mM of the dNTP mix and 1xTaq buffer. In order to reduce amplification of *C. gigas* amplicons for 18SV4, the blocking primer was added to the PCR mix at a final concentration of 1.2 μM (Table 1). PCR cycling included an initial incubation of 15 min at 96°C followed by 35 cycles of 96°C for 30 s, 52°C for 30 s and 72°C for 1 min, with a final 10 min incubation at 72°C . Paired-end sequencing (250 bp) was done at the McGill University (Génome Québec Innovation Centre, Montréal, Canada) on the MiSeq system (Illumina, v2 chemistry) according to the manufacturer's protocol. Raw sequence data are available in the SRA database (accession number PRJNA579900).

Sequence analyses

We used the FROGS pipeline (Find Rapidly OTU with Galaxy Solution) implemented into a galaxy instance (<http://sigenae-workbench.toulouse.inra.fr/galaxy/>) to define Operational Taxonomic Units (OTUs), and compute taxonomic annotations [40] (please see reference [38] for more details). We filtered the dataset for singletons and we annotated OTUs using Blast+ against the Protist Ribosomal Reference database (PR²) [41]. Rarefaction curves of species richness were computed using the {phyloseq} R package and the ggrare function. The rarefy_even_depth function was used to subsample dataset to 5160 reads per sample using. We did not compare low coverage samples (< 5160 sequences). The estimate_richness function was used to compute alpha diversity

metrics (Observed, Chao1 and Shannon). Pielou's measure of species evenness was obtained using the diversity function {vegan}. In order to compare length and nucleotide diversity of 18SV1V2EX and 18SV4BP amplicons, protist sequences from subsampled dataset were aligned for each region using MAFFT, and alignments were trimmed at each extremity. Then, nucleotide diversity of OTU sequences was computed using the nuc.div function {pegas}. The tax_glom function was used to obtain abundances at different taxonomic ranks (from genus to phylum). Multi-affiliations were not considered for these taxonomic ranks. Then, Bray-Curtis dissimilarities were computed at each taxonomic rank to study differences between samples for protist compositions (beta diversity) (vegdist function, {vegan}).

Phylogenetic reconstructions

We performed BLASTn searches of the dominant protist genera against the non-redundant nucleotide collection of NCBI. We kept the 10 first hits for each query (coverage and identity > 90%) to compute phylogenetic reconstructions. Sequences were aligned using MAFFT [36], and trimmed at each extremity. Poorly aligned and highly variable regions of the alignment were automatically removed using GBlocks [42], and maximum likelihood (ML) trees were computed with IQ-TREE v1.3.8 using the best model (selected with the Bayesian information criterion) [43], and validated via an ultrafast bootstrap procedure with 1000 replicates [44].

Statistical analyses

All statistical analyses were done using R v3.3.1 [45].

Hierarchical clusterings (average linkages (hclust {stats})) were computed to describe composition of microbial communities between samples using Bray-Curtis dissimilarities (vegdist {vegan}). Clusterings of 18SV1V2EX and 18SV4BP were plotted face to face using the tanglegram function {dentextend}. Heatmaps of dominant protist genera were computed using relative abundances and the heatmap.2 function ({gplots}).

We performed Student's t-test (t.test {stats}) or non-parametric Wilcoxon test (wilcox.test {stats}) (when normality was rejected with the Shapiro-Wilk test, (shapiro.test {stats})) to compare (i) amplicon sizes, (ii) abundances of total sequences, (iii) abundances of high taxonomic ranks (oyster, Embryophyceae, protists, others), and (iii) alpha diversity metrics (Chao1, evenness and Shannon) between 18SV1V2EX and 18SV4BP. Mantel test (mantel {vegan}) was used to compare 18SV1V2EX and 18SV4BP dissimilarities (Bray-Curtis index) at each taxonomic rank. Permutational multivariate analysis of variance (PERMANOVA, adonis2 {vegan}) was used to investigate the variation of the different OTUs under the constraint of environmental conditions (S1-S3) and oyster families (O1-O5).

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12866-020-01860-1>.

Additional file 1: Table S1. Microbiota samples

Additional file 2: Figure S1. Sequences of high taxonomic ranks (oysters, Embryophyceae, protists) within oyster sample

Additional file 3: Table S2. Number of sequences. Values correspond to 18SV1V2EX/18SV4BP.

Additional file 4: Figure S2. Rarefaction analyses.

Additional file 5: Table S3. OTU annotations and abundances for 18SV1V2EX samples.

Additional file 6: Table S4. OTU annotations and abundances for 18SV4BP samples.

Additional file 7: Figure S3. Comparison of amplicon sizes between 18SV1V2EX and 18SV4BP

Additional file 8: Table S5. Metadata and alpha diversity indices. Alpha diversity values correspond to 18SV1V2EX/18SV4BP. NA: not analysed (protist sequences < 5160).

Abbreviations

18SV1V2: Variable loops V1 and V2 of the 18S rRNA gene; 18SV1V2EX: Sequences obtained using the excluding primer set targeting 18SV1V2; 18SV4: Variable loop V4 of the 18S rRNA gene; 18SV4BP: Sequences obtained using the blocking primer targeting the 18SV4 primer set; OshV-1: *Ostreid herpesvirus 1*; OTU: Operational taxonomic units; POMS: Pacific oyster mortality syndrome; PR²: Protist ribosomal reference database; SSU rRNA: Small subunit ribosomal ribonucleic acid; UNonMet: Universal non-metazoa excluding primers

Acknowledgements

We thank IHPE members for stimulating discussions. We are grateful to Bruno Petton from the Ifremer station of Argenton for the generation and maintaining of all animals used in this study. We thank Lucie Subirana from the Oceanological Observatory of Banyuls-sur-Mer for fruitful discussions at the beginning of the project. We thank the genotoul bioinformatics platform Toulouse Midi-Pyrenees and Sigenea group for providing help and computing resources thanks to Galaxy instance <http://sigenea-workbench.toulouse.inra.fr>. We are also grateful to Sébastien Brunet and Pierre Lepage from the McGill University and Genome Quebec Innovation Center for technical assistances.

Authors' contributions

CC, LG, JME and ET were involved in the study concept and design. CC and ET were involved in data acquisition and analysis. CC and ET drafted the manuscript and all authors contributed to critical revisions and approved the final manuscript.

Funding

CC benefited of post-doctoral fellowships from CNRS and IFREMER. This work was supported by the DHOF program of the UMR5244/IHPE (<http://ihpe.univ-perp.fr/en/ihpe-transversal-holobiont/>). This project has received funding from the European Union's Horizon 2020 Research and innovation programme under grant agreement N° 678589 (VIVALDI project). This study is set within the framework of the "Laboratoires d'Excellences (LABEX)" TULIP (ANR-10- LABX-41) and CeMEB (ANR-10-LABX-04-01).

Availability of data and materials

The sequences can be accessed at the Sequence Read Archive repository (<https://www.ncbi.nlm.nih.gov/sra/>) with BioProject ID PRJNA579900. Until then, the sequences are available from the corresponding author upon reasonable request.

Ethics approval and consent to participate

Not applicable. There is no European ethical framework nor regulation for experimentation on invertebrates (except cephalopods).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹IHPE, Univ. Montpellier, CNRS, Ifremer, Univ. Perpignan Via Domitia, Perpignan, France. ²PSL Université Paris: EPHE-UPVD-CNRS, USR 3278 CRIOBE, Université de Perpignan, 52 Avenue Paul Alduy, 66860 Perpignan Cedex, France. ³Sorbonne Université, CNRS, UMR7144 Adaptation et Diversité en Milieu Marin, Ecology of Marine Plankton (ECOMAP), Station Biologique de Roscoff SBR, Roscoff, France. ⁴IHPE, Univ. Montpellier, CNRS, Ifremer, Univ. Perpignan Via Domitia, Montpellier, France.

Received: 21 November 2019 Accepted: 16 June 2020

Published online: 03 July 2020

References

- Barbosa Solomieu V, Renault T, Travers M-A. Mass mortality in bivalves and the intricate case of the Pacific oyster, *Crassostrea gigas*. *J Invertebr Pathol*. 2015;131:2–10.
- Pernet F, Lupo C, Bacher C, Whittington RJ. Infectious diseases in oyster aquaculture require a new integrated approach. *Philos Trans R Soc B Biol Sci*. 2016;371:20150213.
- Petton B, Bruto M, James A, Labreuche Y, Alunno-Bruscia M, Le Roux F. *Crassostrea gigas* mortality in France: the usual suspect, a herpes virus, may not be the killer in this polymicrobial opportunistic disease. *Front Microbiol*. 2015;6:686.
- de Lorgeril J, Lucasson A, Petton B, Toulza E, Montagnani C, Clerissi C, et al. Immune-suppression by OshV-1 viral infection causes fatal bacteraemia in Pacific oysters. *Nat Commun*. 2018;9:4215.
- Azéma P, Lamy J-B, Boudry P, Renault T, Travers M-A, Dégremont L. Genetic parameters of resistance to *Vibrio aestuarianus*, and OshV-1 infections in the Pacific oyster, *Crassostrea gigas*, at three different life stages. *Genet Sel Evol*. 2017;49:23.
- Desriac F, Le Chevalier P, Brillet B, Leguerinel I, Thuillier B, Paillard C, et al. Exploring the hologenome concept in marine bivalvia: haemolymph microbiota as a pertinent source of probiotics for aquaculture. *FEMS Microbiol Lett*. 2014;350:107–16.
- Becker CD, Pauley GB. An ovarian parasite (*Protista incertae sedis*) from the Pacific oyster, *Crassostrea gigas*. *J Invertebr Pathol*. 1968;12:425–37.
- Comps M, Park MS, Desportes I. Fine structure of *Marteilioides chungmuensis* n.g., n.sp., parasite of the oocytes of the oyster *Crassostrea gigas*. *Aquaculture*. 1987;67:264–5.
- Grijalva-Chon JM, Castro-Longoria R, Enríquez-Espinoza TL, Maeda-Martínez AN, Mendoza-Cano F. Molecular evidence of the protozoan parasite *Marteilia refringens* in *Crassostrea gigas* and *Crassostrea corteziensis* from the Gulf of California. *Lat Am J Aquat Res*. 2015;43:776–80.
- Arzeta-Pino L, Acosta A, Sarmiento ME, Rojas-Contreras M, Rodríguez-Jaramillo C, Vázquez-Juárez R. Herpes virus OshV-1 and the protist *Perkinsus marinus* modify the expression of the Down syndrome cell adhesion molecule gene in gill and mantle of *Crassostrea* spp. *Aquat Res*. 2018;49:3638–46.
- Tun K, Itoh N, Shimizu Y, Yamanoi H, Yoshinaga T, Ogawa K. Pathogenicity of the protozoan parasite *Marteilioides chungmuensis* in the Pacific oyster *Crassostrea gigas*. *Int J Parasitol*. 2008;38:211–7.
- Berthe FJ, Le Roux F, Adlard RD, Figueras A. Marteiliosis in molluscs: a review. *Aquat Living Resour*. 2004;17:433–48.
- Fernández Robledo JA, Vasta GR, Record NR. Protozoan parasites of bivalve molluscs: literature follows culture. *PLoS One*. 2014;9:e100872.
- Katrec J, Soudant P, Payton L, Tran D, Miner P, Lambert C, et al. Bioactive extracellular compounds produced by the dinoflagellate *Alexandrium minutum* are highly detrimental for oysters. *Aquat Toxicol*. 2018;199:188–98.
- Anderson DM, Cembella AD, Hallegraeff GM. Progress in understanding harmful algal blooms: paradigm shifts and new technologies for research, monitoring, and management. *Annu Rev Mar Sci*. 2012;4:143–76.
- BOWER SM, CARNEGIE RB, GOH B, JONES SR, LOWE GJ, MAK MW. Preferential PCR amplification of parasitic protistan small subunit rDNA from metazoan tissues. *J Eukaryot Microbiol*. 2004;51:325–32.
- del Campo J, Pons MJ, Herranz M, Wakeman KC, del Valle J, Vermeij MJA, et al. Validation of a universal set of primers to study animal-associated microeukaryotic communities. *Environ Microbiol*. 2019;21:3855–3861.
- Vestheim H, Jarman SN. Blocking primers to enhance PCR amplification of rare sequences in mixed samples - a case study on prey DNA in Antarctic krill stomachs. *Front Zool*. 2008;5:12.
- Leray M, Agudelo N, Mills SC, Meyer CP. Effectiveness of annealing blocking primers versus restriction enzymes for characterization of generalist diets: unexpected prey revealed in the gut contents of two coral reef fish species. *PLoS One*. 2013;8:e58076.
- Clerissi C, Brunet S, Vidal-Dupiol J, Adjeroud M, Lepage P, Guillou L, et al. Protists within corals: the hidden diversity. *Front Microbiol*. 2018;9:2043.
- Tan S, Liu H. Unravel the hidden protistan diversity: application of blocking primers to suppress PCR amplification of metazoan DNA. *Appl Microbiol Biotechnol*. 2018;102:389–401.
- Stoeck T, Bass D, Nebel M, Christen R, Jones MDM, Breiner HW, et al. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol Ecol*. 2010;19:21–31.
- Tragin M, Zingone A, Vault D. Comparison of coastal phytoplankton composition estimated from the V4 and V9 regions of the 18S rRNA gene with a focus on photosynthetic groups and especially Chlorophyta. *Environ Microbiol*. 2018;20:506–20.
- Massana R, del Campo J, Sieracki ME, Audic S, Logares R. Exploring the uncultured microeukaryote majority in the oceans: reevaluation of ribogroups within stramenopiles. *ISME J*. 2014;8:854–66.
- Hu SK, Liu Z, Lie AAY, Countway PD, Kim DY, Jones AC, et al. Estimating protistan diversity using high-throughput sequencing. *J Eukaryot Microbiol*. 2015;62:688–93.
- Decelle J, Romac S, Sasaki E, Not F, Mahé F. Intracellular diversity of the V4 and V9 regions of the 18S rRNA in marine protists (radiolarians) assessed by high-throughput sequencing. *PLoS One*. 2014;9:e104297.
- Giner CR, Forn I, Romac S, Logares R, de Vargas C, Massana R. Environmental sequencing provides reasonable estimates of the relative abundance of specific picoeukaryotes. *Appl Environ Microbiol*. 2016;82:4757–66.
- Piredda R, Tomasino MP, D'Erchia AM, Manzari C, Pesole G, Montresor M, et al. Diversity and temporal patterns of planktonic protist assemblages at a mediterranean long term ecological research site. *FEMS Microbiol Ecol*. 2017;93:fw200.
- Kamiyama T. Microzooplankton as a food source for the Pacific oyster *Crassostrea gigas*: seasonal variation in gut contents and food availability. *Fish Sci*. 2011;77:961–74.
- de Vargas C, Audic S, Henry N, Decelle J, Mahé F, Logares R, et al. Eukaryotic plankton diversity in the sunlit ocean. *Science*. 2015;348:1261605.
- Not F, Latasa M, Marie D, Cariou T, Vault D, Simon N. A single species, *Micromonas pusilla* (Prasinophyceae), dominates the eukaryotic picoplankton in the western English Channel. *Appl Environ Microbiol*. 2004;70:4064–72.
- Theis KR, Dheilly NM, Klassen JL, Brucker RM, Baines JF, TCG B, et al. Getting the hologenome concept right: an eco-evolutionary framework for hosts and their microbiomes. *mSystems*. 2016;1:e00028–16.
- García-Lagunas N, de Jesús R-GR, Hernández-Saavedra NY. Changes in gene expression and histological injuries as a result of exposure of *Crassostrea gigas* to the toxic dinoflagellate *Gymnodinium catenatum*. *J Molluscan Stud*. 2016;82:193–200.
- Quast C, Pruesse E, Yilmaz P, Gerken J, Schaefer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*. 2013;41:D590–6.
- Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, et al. The SILVA and "all-species living Tree project (LTP)" taxonomic frameworks. *Nucleic Acids Res*. 2014;42:D643–8.
- Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res*. 2002;30:3059–66.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol*. 1990;215:403–10.
- Clerissi C, de Lorgeril J, Petton B, Lucasson A, Escoubas J-M, Gueguen Y, et al. Microbiota composition and evenness predict survival rate of oysters confronted to Pacific oyster mortality syndrome. *Front Microbiol*. 2020;11:311.

39. Wuyts J, Van de Peer Y, Winkelmans T, De Wachter R. The European database on small subunit ribosomal RNA. *Nucleic Acids Res.* 2002;30:183–5.
40. Escudie F, Auer L, Bernard M, Cauquil L, Vidal K, Maman S, et al. FROGS: find rapidly OTU with galaxy solution. Montpellier, France: The environmental genomics Conference; 2015.
41. Guillou L, Bachar D, Audic S, Bass D, Berney C, Bittner L, et al. The Protist ribosomal reference database (PR2): a catalog of unicellular eukaryote small sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* 2012; 41:D597–604.
42. Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol.* 2000;17:540–52.
43. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol.* 2015;32:268–74.
44. Minh BQ, Nguyen MAT, von Haeseler A. Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol.* 2013;30:1188–95.
45. R: a language and environment for statistical computing. R development Core team. Vienna, Austria: R Foundation for Statistical Computing; 2008. <http://www.R-project.org>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

