A framework for in situ molecular characterization of coral holobionts using nanopore sequencing

Quentin Carradec, Julie Poulain, Emilie Boissin, Benjamin Hume, Christian Voolstra, Maren Ziegler, Stefan Engelen, Corinne Cruaud, Serge Planes, Patrick Wincker

OPEN

A framework for in situ molecular characterization of coral holobionts using nanopore sequencing

Quentin Carradec[1,2,9]✉, Julie Poulain[1,2,9], Emilie Boissin[3,4], Benjamin C. C. Hume[5], Christian R. Voolstra[5,6], Maren Ziegler[7], Stefan Engelen[8], Corinne Cruaud[8], Serge Planes[2,3,4] & Patrick Wincker[1,2]✉

Molecular characterization of the coral host and the microbial assemblages associated with it (referred to as the coral holobiont) is currently undertaken via marker gene sequencing. This requires bulky instruments and controlled laboratory conditions which are impractical for environmental experiments in remote areas. Recent advances in sequencing technologies now permit rapid sequencing in the field; however, development of specific protocols and pipelines for the effective processing of complex microbial systems are currently lacking. Here, we used a combination of 3 marker genes targeting the coral animal host, its symbiotic alga, and the associated bacterial microbiome to characterize 60 coral colonies collected and processed in situ, during the _Tara_ Pacific expedition. We used Oxford Nanopore Technologies to sequence marker gene amplicons and developed bioinformatics pipelines to analyze nanopore reads on a laptop, obtaining results in less than 24 h. Reef scale network analysis of coral-associated bacteria reveals broadly distributed taxa, as well as host-specific associations. Protocols and tools used in this work may be applicable for rapid coral holobiont surveys, immediate adaptation of sampling strategy in the field, and to make informed and timely decisions in the context of the current challenges affecting coral reefs worldwide.

Coral reefs are threatened worldwide by global environmental changes and local anthropogenic pressures[1]. Studying coral reef ecosystems today is essential to understand what their stressors are and how their biodiversity will be affected in the coming years. Coral holobionts are ecological units organized from an anthozoan cnidarian animal host (the coral) and its obligate photosynthetic dinoflagellate endosymbionts of the family Symbiodiniaceae[2]. These dinoflagellates are the subject of numerous scientific studies given their role in coral sensitivity and resilience to bleaching events[3]. However, an array of other organisms (Bacteria, Archaea, Fungi, Protists, and viruses) living within or around the coral colony may be as important as the dinoflagellate for coral holobiont health[4,5]. For example, symbiotic cyanobacteria have been shown to be a source of nitrogen for the scleractinian coral _Montipora_ sp.[6]. Conversely, several strains of _Vibrio_ are putative causative agents of coral bleaching[7]. Development of high-throughput sequencing techniques have been fundamental in enabling our characterization of the taxonomic diversity that makes up the coral holobiont.

Morphological identification of coral species is challenging for the non-specialist and even for coral researchers, microscopy equipment is often required to distinguish between closely related species[8]. For the past 2 decades, genetic markers have been used to help the identification of coral species or establish population structures

[1]Génomique Métabolique, Genoscope, Institut François Jacob, CEA, CNRS, Univ Evry, Université Paris-Saclay, 91057 Evry, France. [2]Research Federation for the Study of Global Ocean Systems Ecology and Evolution, R2022/Tara Oceans GO-SEE, 3 rue Michel-Ange, 75016 Paris, France. [3]PSL Research University: EPHE-UPVD-CNRS, USR 3278 CRIOBE, Université de Perpignan, Perpignan Cedex, France. [4]Laboratoire d'Excellence "CORAIL", 52 Avenue Paul Alduy, 66860 Perpignan Cedex, France. [5]Red Sea Research Center, Division of Biological and Environmental Science and Engineering (BESE), King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia. [6]Department of Biology, University of Konstanz, 78457 Konstanz, Germany. [7]Department of Animal Ecology & Systematics, Justus Liebig University, Giessen, Germany. [8]Genoscope, Institut de Biologie François-Jacob, Commissariat à l'Energie Atomique (CEA), Université Paris-Saclay, Evry, France. [9]These authors contributed equally: Quentin Carradec and Julie Poulain. ✉email: qcarrade@genoscope.cns.fr; pwincker@genoscope.cns.fr

of these intricate taxa. However, due to slow nucleotide substitution rates for mitochondrial genes within anthozoans[9], common genetic markers such as the cytochrome c oxidase subunit I or the cytochrome b are not discriminant enough to identify corals at the species level[10,11]. Nuclear markers like the 18S or the 5.8S rRNA[12,13] are generally more variable between species although multi-locus and microsatellite analysis are often required to precisely determine species boundaries[14].

The most commonly used marker gene to study the diversity of coral symbionts in the Symbiodiniaceae diversity is the Internal Transcribed Spacer 2 (ITS2) region of the rDNA[15]. Whereas only 22 Symbiodiniaceae species are formally described, a large ITS2 sequence diversity recently grouped in seven genera suggest a much larger number of species[2]. ITS2 is a multi-copy marker that can resolve Symbiodiniaceae at the level of taxa and strains, but the high intragenomic diversity of the ITS2 sequence poses analytical challenges in distinguishing intra- from inter-genomic diversity[16]. Different strategies have been developed to solve this problem including an analytical framework ("SymPortal"). This tool makes explicit use of the intragenomic diversity by employing the resolution of next-generation-sequencing approaches to determine ITS2 type profiles of putative Symbiodiniaceae taxa based on consistent co-occurrence of defining intragenomic ITS2 variants[17]. The large diversity of Symbiodiniaceae symbionts and numerous different microhabitats across coral reefs make the global comprehension of environmental and biological drivers of host-symbiont specificity challenging[18–20].

The 16S rRNA gene is the most commonly used marker to assess the bacterial diversity in coral holobionts[21]. Previous quantification of coral-associated bacterial assemblage richness have identified up to 100,000 distinct OTUs dominated by gamma- and alpha-proteobacteria[22]. Among them, *Endozoicomonas* is probably the most abundant and widely distributed bacterial genus[23,24]. For some coral species, the bacterial composition is highly variable according to coral reef site, environmental conditions, or seasons[25], whereas for other corals, bacterial compositions is rather fixed and less variable[26]. As such, comprehensive sampling efforts are needed to discover all possible associations and to identify taxa of putative functional importance[27].

The recent development of the MinION portable sequencer by Oxford Nanopore Technologies (ONT) allows real time long read sequencing and is practicable in the field. Barcoding experiments have been realized in various environments for the molecular identification of rare species[28–30]. However, very few studies have realize metabarcoding experiments on complex samples due to the lack of analysis pipelines able to work with long sequences and to manage the high error rate of ONT sequencing for the taxonomic identification[31].

Here, we investigate the diversity of microbial assemblages associated with corals sampled around an isolated reef in north eastern Papua New Guinea (Kimbe Bay). Without access to bulky sequencing instruments and with limited laboratory facilities, we used the MinION device onboard the research vessel *Tara* to evaluate coral host identity as well as Symbiodiniaceae and bacterial composition of sampled coral specimens in the field. We simultaneously used three rRNA marker genes to characterize the coral holobiont: the full-length 18S rRNA for the coral host, the ITS2 for Symbiodiniaceae diversity, and the 16S rRNA for the bacterial assemblages. In total, 55 samples of scleractinian corals from 13 different genera and 5 samples of the fire coral *Millepora* (Hydrozoa) were assessed.

Materials and methods

Coral and Hydrozoa sampling and DNA extraction.   60 coral samples were collected between 5 and 20 m depth by removing a fragment on four different reefs of the Kimbe Bay in Papua New Guinea (New Britain island) (Fig. 1). Various families of scleractinian corals as well as the hydrozoan genus *Millepora* were collected (Supplementary Fig. S1). Coral fragments were placed in zip-lock bags under water then stored in 2 ml RNAse/DNAse free tubes containing Lysing Matrix A (MP Biomedical) and 1.5 ml of DNA/RNA shield preservative buffer (Zymo Research, Irvine, California, USA). Samples were then placed in a Terralyzer Instrument (Zymo Research) and grinded for 2 min. Grinded samples (500 µl) were incubated 30 min at 55 °C with 75 µl of protein-ase K. Then, 1.5 ml of Lysis Buffer (ZR-DuetDNA/RNA Miniprep Plus, Zymo Research) was added for 15 min at room temperature (around 35 °C). Coral DNA was then extracted using the ZR-DuetDNA/RNA Miniprep Plus Kit following the manufacturer's instructions. DNA was quantified on a Qubit dsDNA HS Assays (Thermo Fisher Scientific, Waltham, USA). High resolution coral pictures were archived at the European Bioinformatics Institute.

Full-length marker gene amplification.   The ITS2, 18S and 16S rRNA sequences were targeted for amplicon generation. For each primer, the ONT tail and one unique barcode (out of twelve) were added. Full-length 18S rRNA primers: TTTCTGTTGGTGCTGATATTGC-Barcode-AACCTGGTTGATCC TGCCAGT for the forward and ACTTGCCTGTCGCTCTATCTTC-Barcode- TGATCCTTCTGCAGGTTCACCTAC for the reverse primer[32]. Bacterial-specific 16S rRNA primers: 27F (TTTCTGTTGGTGCTGATATTGC-Barcode-AGAGTTTGATCMTGG CTCAG) and 1492R (ACTTGCCTGTCGCTCTATCTTC -Barcode- TACGGYTAC CTTGTTA CGACTT)[33]. ITS2 primers SYM_VAR_5.8S: TTTCTGTTGGTGCTGATATTGC-Barcode-GAA TTGCAGAACTCCGTGAACC for the forward and SYM_VAR_REV: ACTTGCCTGTCGCTCTATCTTCT-Barcode-CGGGGTTCWCTTGTYTGACTTCATGC for the reverse[16,34,35]. PCRs were performed on board using 25 ng of DNA from each coral sample with the Advantage 2 kit (Takara Clontech) and a final primer concentration of 0.5 µM in a final reaction volume of 50 µl. PCR conditions were optimized to be able to generate the three amplicons using the same thermocycling program: initial denaturing at 95 °C for 1 min, 30 cycles each at 95 °C for 30 s, 55 °C for 30 s, and 68 °C for 1 m, followed by a final extension step at 68 °C for 10 min. Twelve barcodes were available per primer set and allowed us to process DNAs from 12 samples in parallel on one 96 well PCR plate.

PCR products for each sample were pooled by targeted gene in equimolar ratios then each pool was 1:10 diluted and subsequently cleaned with 1 volume of AMPure XP (Beckman Coulter, Brea, California, USA) before
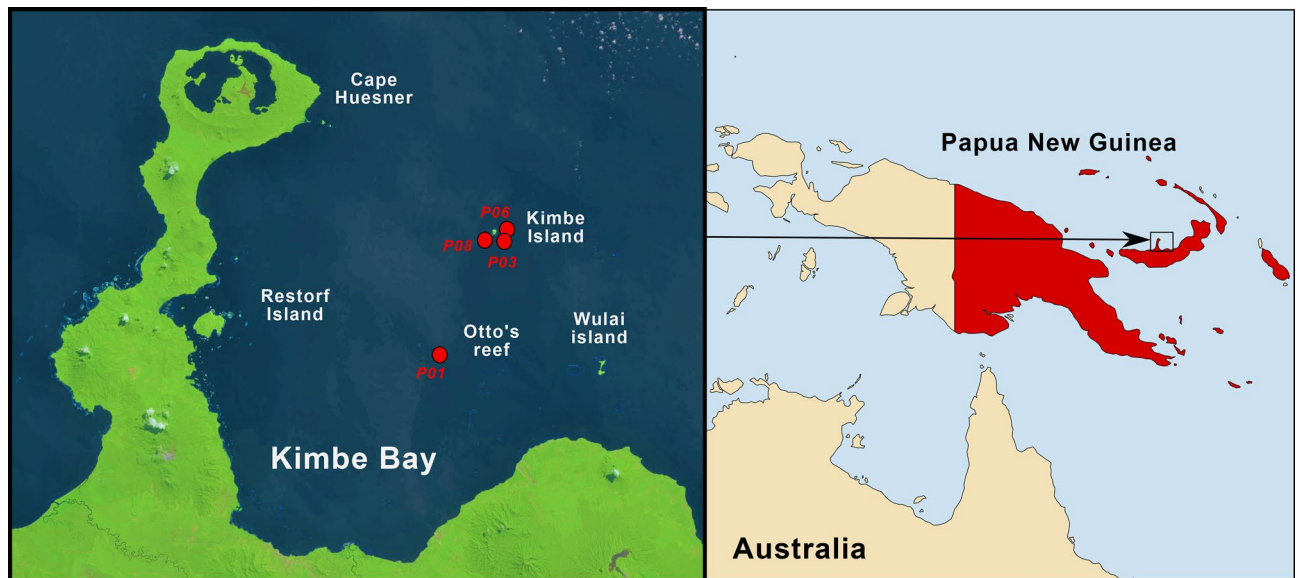
Figure 1. Map of the sampling sites. The four sampling sites are indicated by red dots on the LANDSAT satellite image of Kimbe Bay (left). The first site (P01) is close to Otto's Reef, the three other sites (P03, P06 and P08) are around the Kimbe Island. (Right panel is generated with R version 4.0.2 and the maps package, https://cran.r-project.org/bin/windows/base/old/4.0.2/).

quantification. 2.5 ng of ITS2-PCRs-pool, 250 ng of 16S-PCRs-pool and 250 ng of 18S-PCRs-pool were mixed to constitute the input of the ONT sequencing library. The smaller amount of ITS2-PCRs-pool was aimed at compensating for amplification bias, as PCR products for shorter sequences would result in a higher number of molecules sequenced.

ONT library preparation and sequencing. Sequencing libraries were prepared for R7.9 flow cells run (FLO-MAP107) on MinION device using the Low Input by PCR Sequencing Kit SQK-LWP001 according to the four-primers PCR protocol from ONT. This ONT kit is no longer available and has been replaced by the ligation sequencing kit SQK-LSK109, which is compatible with the recent flow cells R9.4 and R9.5. Previous steps are compatible with the current version of the sequencing kit, but second multiplex PCR, described below, that adds the Whole Genome Primers (WGP from the SQK-LWP001 kit) is no longer required for the SQK-LSK109 kit. We mixed 500 ng of the pool of amplicons with 1.5 µl of WGP, 3.75 µl of dNTP at 10 mM, 1 µl of Advantage 2 polymerase (Takara Clonetech, USA), 5 µl of Advantage 2 buffer in a final reaction volume of 50 µl. The following PCR conditions were used: initial denaturing at 95 °C for 3 min, 15 cycles at 95 °C for 30 s, 56 °C for 30 s, and 68 °C for 1 m, followed by a final extension step at 68 °C for 10 min. PCR products were cleaned with 0.8 volume of Ampure XP and finally eluted in 20 µl of 10 mM Tris–HCl pH8 with 50 mM NaCl. 100 ng of PCR products were diluted in 10 µl then mixed with 1 µl of Rapid 1D sequencing adapters (RPD from the SQK-LWP001 kit) and 1 µl of Ligase T4 Blunt (New England Biolabs, Ipswich, MA, USA) to increase the efficiency of ligation. Pre-sequencing mix of the SQK-LWP001 kit was incubated 10 min at 25 °C and left on ice until ready to load. The flow cell was primed with a mix of 500 µl of RBF and 500 µl of nuclease-free water according to the manufacturer's recommendations. Eleven µl of pre-sequencing mix was combined with 30.5 µl of Running Buffer with Fuel Mix Buffer, 7 µl of water and 26.5 µl of Library Loading Beads then loaded to the flow cell.

The sequencing run was performed for 12 to 24 h. Bases were called during the MinION run with the MinKnow software (v. 1.7.14). The demultiplexing and adaptor trimming were done with porechop tool (https://github.com/rrwick/Porechop) with the option discard_middle. Sequences are archived at the European Bioinformatics Institute.

Read mapping, consensus reconstruction, and species identification. Three specific databases were used to identify each set of nanopore reads. The coral 18S rRNA database contains full-length 18S rRNA for 140 scleratinia and 2 *Millepora*[12]. Five 18S rRNA sequences of Symbiodiniaceae (gi:176088, 12247076, 1224077, 148734588 and 12247080) were added in the database in order to detect and remove undesirable Symbiodiniaceae 18S sequences. The ITS2 database contains 432 sequences of Symbiodiniaceae[18] and finally the Greengenes database (v.13.5, https://greengenes.lbl.gov) was used to detect 16S rRNA reads. All nanopore reads were mapped to each database with minimap2 (v 2.0-r191) with the pre-set options "map-ont"[36]. For coral identification, the reference sequence that had the most sequences mapped to it was the only one retained in each sample. Then a second round of mapping (same parameters) was done on the selected reference in order to aggregate reads potentially mis-assigned during the first round of mapping. Two Maximum Likelihood phylogenetic trees with newly reconstructed sequences and the coral 18S rRNA database were further reconstructed, one with all sequences except Acroporidae and the other with Acroporidae only. For Symbiodiniaceae and the
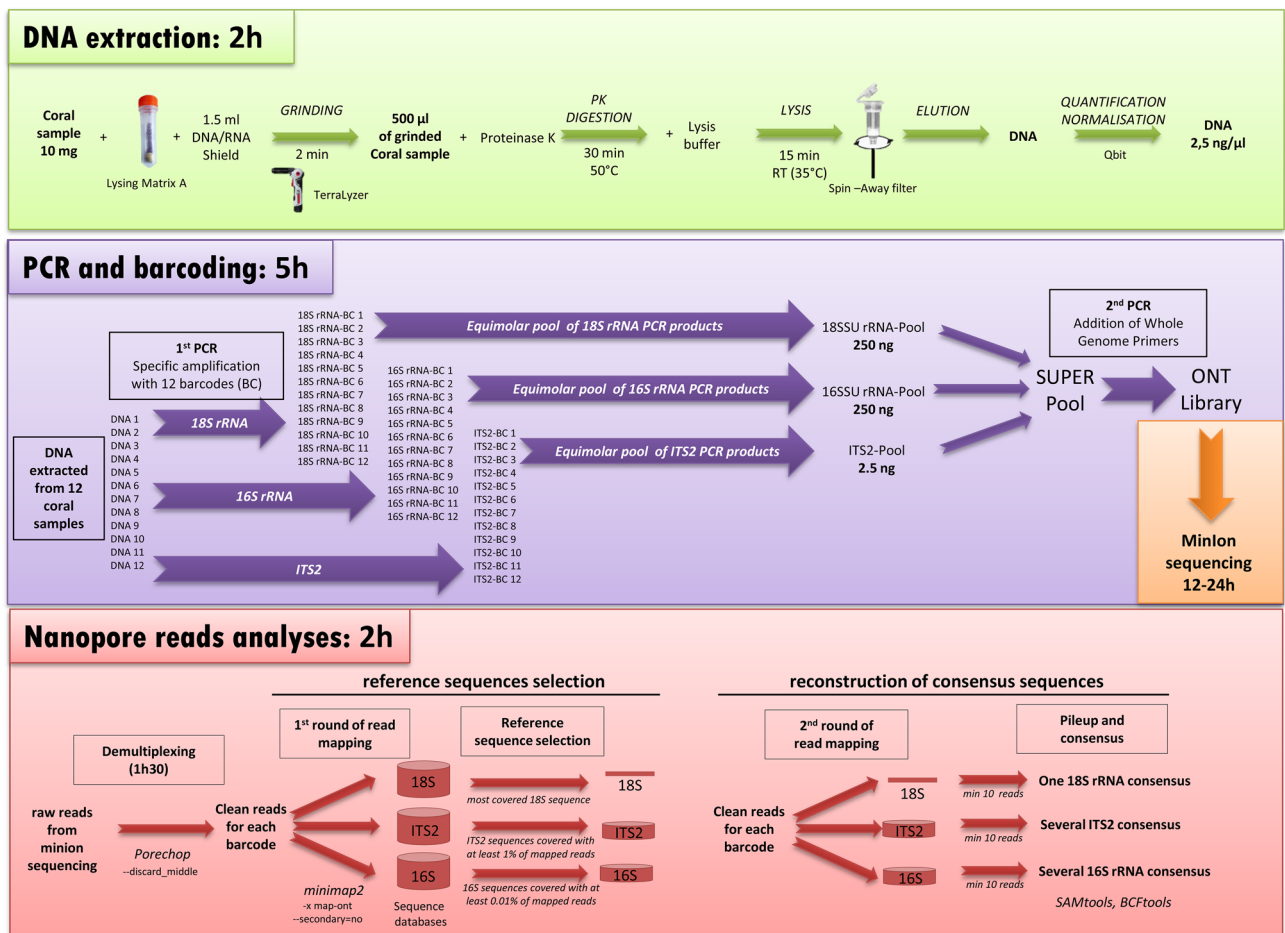
3

Figure 2. Pipeline for in situ analysis of coral holobionts. The approximate times for each step correspond to the processing of 12 coral samples. Protocols and tools parameters are detailed in the method section. (Figure generated with Microsoft Office 2016 Professional Power Point https://www.microsoft.com/).

bacterial community the same strategy was used except that all references covered with a minimum of 1% (for ITS2 sequences) or 0.01% (for 16S rRNA sequences) of all reads mapped were kept for the second round of mapping. In addition for the bacterial community, reads mapped on eukaryote chloroplastic 16S were removed after the first mapping. SAMtools and BCFtools were used to reconstruct consensus sequences for each reference sequence covered with more than 10 nanopore reads with the following programs and options: mpileup -B -a -Q 0 –u; bcftools call -c -ploidy 1;vcfutils.pl vcf2fastq. The specific command line execution and system requirements are detailed in the Supplementary Method S1. This analysis was undertaken on a laptop, without internet access and on board of the research vessel *Tara*, at sea.

Illumina sequencing and analysis of Symbiodiniaceae amplicons. The same PCR pools obtained onboard *Tara* were, later in laboratory, used for the sequencing of ITS2 amplicons on Illumina instruments in order to compare with ONT sequencing. 100 ng were directly end-repaired, A-tailed and ligated to Illumina adapters on a Biomek FX Lab Auto Workstation (Beckman Coulter). Then, library amplification was performed using Kapa Hifi HotStart NGS library Amplification kit and purified with AMPure XP (1 volume). Libraries concentrations were normalized to 10 nM by addition of Tris–Cl 10 mM, pH 8.5 and then applied to cluster generation according to the Illumina Cbot User Guide (Part # 15006165). Libraries were sequenced on a MiSeq instrument with paired end (2 × 300 bp). The taxonomic assignation of Illumina reads was done with the "SymPortal" pipeline[17]. In addition, novel ITS2 sequences were matched against the ITS2 database with BLAST (v2.6.0) then assigned to their closest match.

Results

Coral sample collection and ONT sequencing of marker genes. A total of 60 coral colonies were collected at 4 sites in Kimbe bay, Papua New Guinea during a 12-day research trip in December 2017 onboard the *Tara* vessel as part of the *Tara* Pacific expedition (Fig. 1). In order to identify Symbiodiniaceae and the bacterial community associated to each coral, we extracted DNA from each colony and sequenced 3 different marker genes on the Oxford Nanopore MinION device (Fig. 2 and "Methods"). For each colony, the full-length 18S rRNA sequence (1.8 kb) was PCR-amplified with primers designed for coral identification[32], the ITS2 region

(~ 250 bp) was amplified with primers specific for the Symbiodiniaceae family[16], and the full-length 16S rRNA (1.3 kb) with bacteria-specific primers[33]. 12 unique identifier sequences (barcodes) were added to the 5′ends of each primer pair. This strategy allowed us to sequence the holobiont of 12 coral colonies in a single MinION run lasting 24 h (Fig. 2).

After 5 sequencing runs over a period of 10 days, we obtained a total of 2,019,607 reads (2.0 Gb) from 60 coral colonies and the unique barcodes were recognized for 32% of the sequences on average (Supplementary Table S1). For 48 samples, more than one thousand demultiplexed sequences were obtained for each barcode, which is seemingly sufficient to identify the coral as well as dominant Symbiodiniaceae and bacterial taxa. For the 12 remaining samples (11 sequenced during the last run), the number of reads was too low to identify the coral and bacterial community. In this last run, most of the nanopore channels became inactive within a few seconds of run resulting in a very low output. This phenomenon is most probably due to the conservation of the flow cell for more than 2 months before the run which is longer than ONT specifications. However, given the number of sequencing reads obtained was sufficient to study the ITS2 amplicon, we decided to keep this MinION run for downstream analysis. The identification of 18S rRNA, ITS2 and 16S rRNA sequences was done by the mapping of nanopore reads against specific databases for each marker gene (see the Method section and Supplementary Method S1).

Identification of corals, Symbiodiniaceae, and bacterial communities.

Corals were identified by both their morphological traits (Supplementary Fig. S1) and the analysis of 18S rRNA sequences. A full-length 18S rRNA consensus was obtained for 50 corals covering 10 scleratinia families and the fire coral *Millepora*. The coral identification with the 18S rRNA was limited at the genus level for non-acroporid corals because the 18S rRNA sequence is not discriminant at species level for these genera[12] (Supplementary Table S2). A phylogenetic tree was reconstructed confirming the taxonomic identification of these corals (Supplementary Fig. S2). The DNA extraction and/or the PCR amplification failed for 10 coral colonies. In these cases, we based our identification on the morphological traits only.

In order to identify Symbiodiniaceae diversity in each coral colony, nanopore reads were mapped against a database of 432 *Symbiodinium* ITS2 sequences[18]. Several Symbiodiniaceae species may coexist in a coral colony so we can expect multiple ITS2 sequences per sample. As such, for each sample, ITS2 sequences covered with at least 1% of all nanopore reads aligned were conserved. In 53 of the sampled coral colonies we succeeded in identifying at least 1 ITS2 sequence. A similar method was used to characterize the bacterial community from the full-length 16S rRNA (see Methods).

In order to get a broad overview of the coral holobionts analyzed in this study, we constructed a force-directed graph (Fig. 3). This network showed that Symbiodiniaceae taxonomy correlates to coral host at the family level, whereas bacterial specificity to its host is dependent on the bacterial taxa. Some bacterial taxa are detected in almost all coral hosts sampled (in the centre of Fig. 3), while others appear host-specific.

Symbiodiniaceae sequence diversity in coral holobionts.

Symbiodiniaceae ITS2 sequences from taxa in the genus *Durusdinium* (Clade D) were dominant in 22 coral colonies including 4 Fungiidae colonies and all *Diploastrea, Galaxea*, and *Pocillopora* (Fig. 3 and Fig. 4a). A total of 8 different ITS2 sequences were reconstructed from these samples. The D1 sequence was the most abundant representing between 40 and 60% of reads in each coral colony. The D2, D3, D4, D10, and D17 sequences were also recovered in all of these colonies in lower proportions suggesting non-specific alignment of ONT reads. To test this hypothesis, we re-sequenced the same PCR amplicons using Illumina technologies, which have a reduced sequencing error rate (Fig. 4b). Sequencing results confirmed the dominance of the D1 sequence in these 22 coral colonies, but also revealed distinct patterns of less abundant, i.e. minor, ITS2 sequences: D4 in Fungiidae, *Diploastrea* and *Galaxea* colonies, and D2 in *Pocillopora* colonies.

These differences between ONT and Illumina reveal a limitation for the identification of minor ITS2 sequences of the genus *Durusdinium* with the current error rate of ONT sequences.

Several sequences representing *Cladocopium* taxa (Clade C) were detected in different coral families (Figs. 3 and 4a). The C3 and C40 sequences (1 insertion and 1 substitution between them) were present in all *Acropora* colonies. Their co-presence is confirmed with the Illumina sequencing (Fig. 4b and Supplementary Fig. S3). Three *Cladocopium* sequences were also recovered from the *Millepora* colonies: C57, C66, and C74 in consistent proportion (9% C74, 72%, C66, 19% C57) (Fig. 4a). C57 and C74 variants (respectively 1 and 2 substitutions with C66) were not observed with Illumina sequencing suggesting that their detection is also due to the error rate of ONT sequences. We detected C15 sequences in all *Porites* colonies and C60, a C15-derived sequence (2 substitutions) confirmed by Illumina sequencing, in 1 colony (Fig. 4a). Finally, 7 coral colonies from 6 different genera present a large diversity of *Cladocopium* ITS2 variants, however several samples of the same coral species would have been necessary to substantiate these associations (Fig. 4a). Sequences from the *Cladocopium* and *Durusdinium* genera were found co-occurring in 4 colonies: one *Plathygyra* with the C40 and D1, one *Symphyllia* colony with C3, C40 and D1, one *Pavona* colony with the C116 and the D6 and one *Pocillopora* with C116 and D1 ITS2 sequences (Fig. 4). The co-occurrence of these two genera in a coral colony were previously reported in Papua New Guinea[37].

On average of all samples, 4.9 unique ITS2 sequences per sample were observed with ONT sequencing versus 2.9 with Illumina sequencing (Supplementary Table S3). This higher richness of ITS2 sequences with ONT reflect a miss-assignation of several reads likely due to the sequencing error rate of this technology.

Coral-associated bacterial assemblages.

We retrieved 1637 unique bacterial 16S sequences belonging to 77 orders of bacteria. Among these sequences, 175 are detected in 2 or more coral colonies. These Bacteria
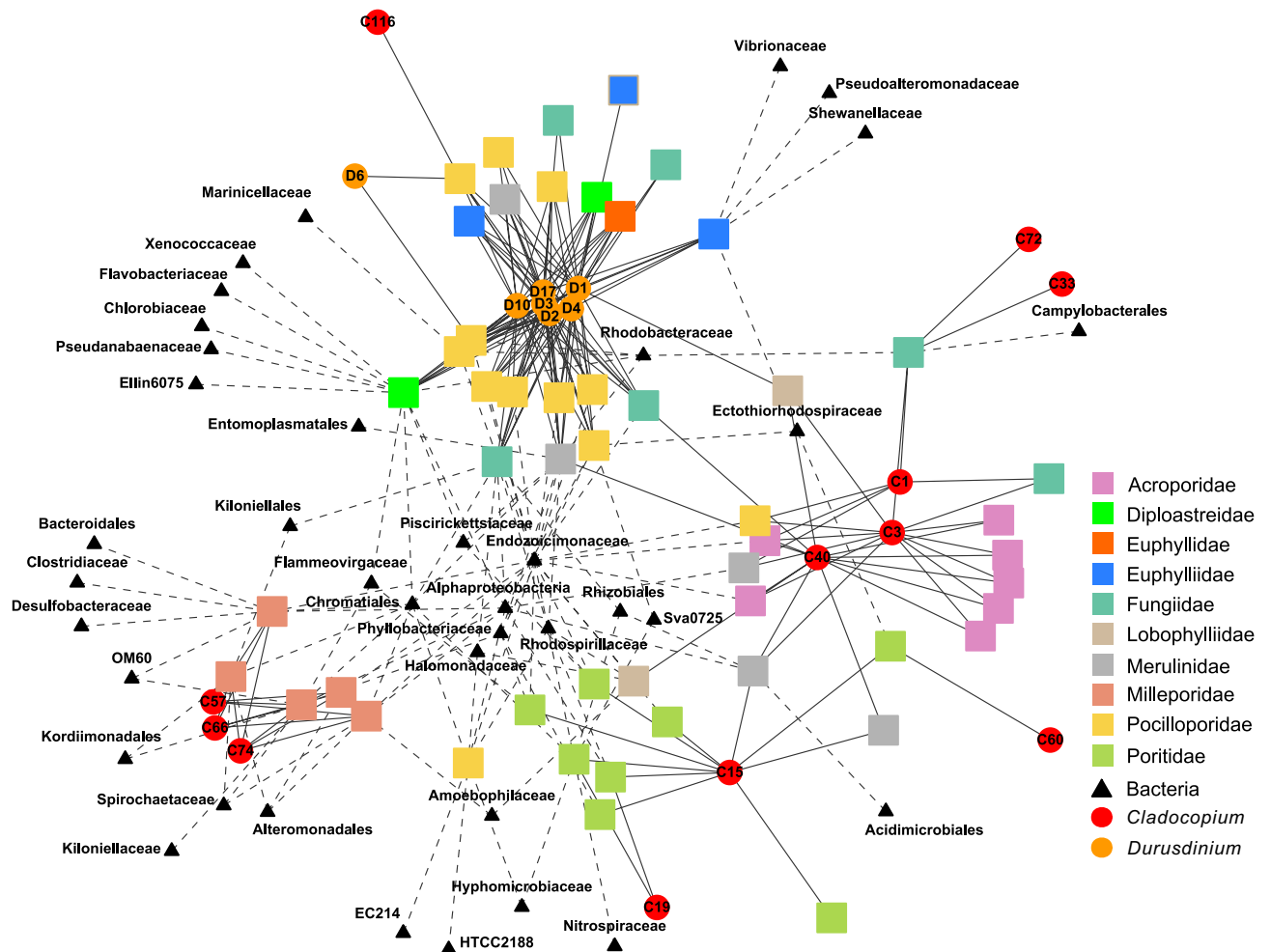
Figure 3. Coral holobiont network. Network representation of a force directed graph of Symbiodiniaceae and bacterial families living within or around each coral colony. Coral colonies are represented by a square and coloured according to their taxonomy (family level). Each node connected to a coral colony represents an organism living within or around the colony. Symbiodiniaceae taxa identified with nanopore sequencing are represented by a circle coloured by taxonomic origin; red for *Cladocopium* (Clade C) and orange for *Durusdinium* (Clade D). ITS2 sequences are indicated inside each circle. Each bacterial sequence is represented by a triangle and its family of origin is indicated below. Symbiodiniaceae ITS2 sequences and Bacteria 16S rRNA sequences covered with less than 5% and 1% of all mapped reads respectively are not represented in this figure. (Figure generated with R version 3.5.0, https://cran.r-project.org/bin/windows/base/old/3.5.0/).

were identified in 31 coral colonies (Supplementary Table S2 and S4). Of the coral colonies sampled, *Oceanospirillales* was the most common bacterial order (detected in 22 coral colonies), followed by *Rhizobiales* (12), uncharacterized alphaproteobacteria (12) and *Chromatiales* (11) (Supplementary Tables S2 and S4). These marine bacteria are commonly detected in a large diversity of corals colonies and are known to be associated with coral tissue and mucus. For instance, *Endozoicomonadaceae* (*Oceanospirillales*) are known to be abundant in healthy mucus and absent or in very low abundance in diseased corals[25,38–41]. Their presence in a large number of corals (20 colonies including *Poritidae, Pocilloporidae, Acroporidae*, and *Milleporidae*) may indicate the health of this coral reef, although the functional role of *Endozoicomonas* remains elusive[23,42]. Among the 186 unique full-length 16S sequences belonging to the *Endozoicomonas* family, several are shared between different coral hosts (maximum of 12 samples) showing that the family is large, commonly found in corals, but also display a pattern of fine-scale genetic differentiation with host (Supplementary Table S4) as recently shown[43,44]. In addition, *Ectothiorhodospiraceae* family and *Kordiimonadales* order detected in several samples were so far not commonly identified within corals. From our data, we argue that they may represent important families, either for this particular reef or in a broader context, awaiting further studies.

At the same time, we found opportunistic bacterial taxa known to be associated with corals under stress. For instance, a bacterium belonging to the *Vibrionaceae* family was detected in one colony of *Galaxea* (*Euphylliidae*). This family of bacteria has been described to be prevalent in diseased corals[45,46]. Furthermore, two *Alteromonadales* species were detected in this colony. *Shewanellaceae*, already described in *Favia* corals[47], and *Pseudoalteromonadaceae* describe in corals affected by sedimentation and local sewage[48] suggests that this colony may be under stress.
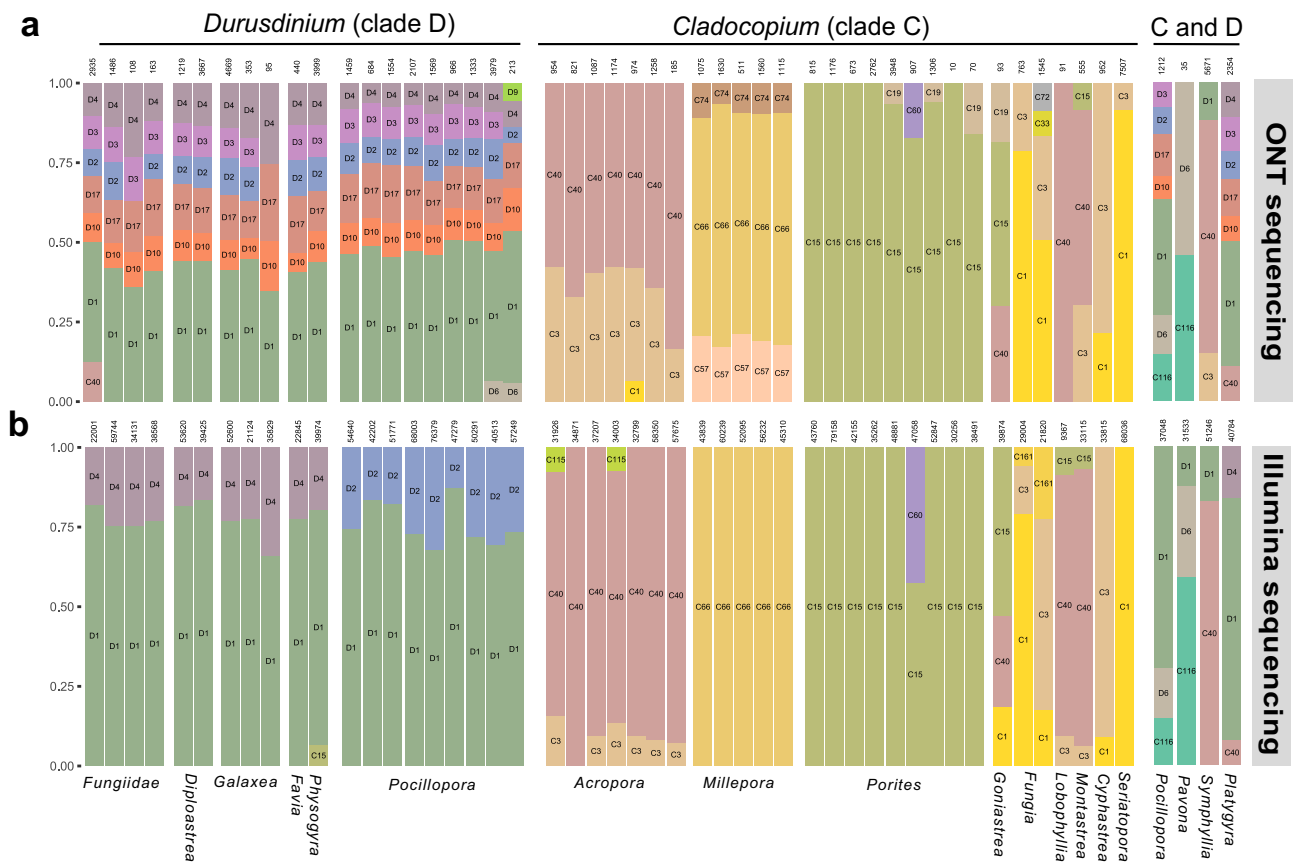
Figure 4. Symbiodiniaceae ITS2 diversity in coral colonies sequenced with ONT or Illumina technology. Each bar represents the proportion of Symbiodiniaceae ITS2 sequences with a relative abundance above 5% detected in a coral colony. The abundances of ITS2 sequences are grouped by Symbiodiniaceae subtype and are represented by a specific colour. The total number of reads is indicated on top of each bar. Corals colonized by Symbiodiniaceae of *Cladocopium* genus are on the left panel, *Durusdinium* genus on the middle panel and by both genera on the right panel. (**a**) PCR amplicons sequenced with ONT. (**b**) Same PCR amplicons sequenced with Illumina technology (Figure generated with R version 3.5.0, https://cran.r-project.org/bin/windows/base/old/3.5.0/).

An endosymbiotic bacterium of *Millepora* belonging to *Spirochaetaceae* family was detected in four out of five coral colonies and a *Kordiimonadales* bacterium in two colonies, these bacterial families were already described in healthy as well as sick tissues of *Millepora*[49,50]. The Gammaproteobacteria *Congregibacter* (OM60, *Alteromonadales*) was also detected in two colonies; this photosynthetic bacterium is found abundant in coastal ecosystems, but was never reported in association with *Millepora*[51]. Three other bacterial families (*Clostridiaceae*, *Desulfobacteraceae* and *Bacteroidales*) were detected in one sample of *Millepora* (P03-C051) suggesting a diseased colony (Fig. 2 and Supplementary Table S2)[52,53]. The aspect of this *Millepora* colony with the presence of parasites, the strong space competition with other coral species and the presence of dead parts colonized by algae corroborate this hypothesis based on the bacterial composition (Supplementary Fig. S1).

Discussion

Recent studies have shown the efficiency of ONT sequencing for rapid species identification in the field in remote locations[28,29]. However, this technology has so far not been applied to characterize complex ecosystems or holobionts. In this study, the MinION device was used to describe the diversity of one of the most diverse ecological units on earth: the coral holobiont[54].

A discriminant marker gene for all corals is still lacking, but the full-length 18S rRNA sequence used in this study was sufficient to identify corals at the species level for acroporids and at the genus level for most of the other corals. Coral identification with 18S rRNA sequencing in the field may open coral studies to non-specialist in contrary to the morphological identification that requires taxonomic expertise. Moreover, the identification in less than 24 h on MinION device could significantly evolve sampling methods for corals. Given the ongoing uncertainties with coral taxonomy designations at the species level, it should be noted that even a designation of corals to the genus level, may be considered a big step forward with regard to diversity assessment to aid conservation efforts.

In addition to the identification of the coral host species, we characterized the Symbiodiniaceae community and recovered sequences from the genera *Cladocopium* and *Durusdinium*. The *Cladocopium* symbiont genus

is the most diverse of the Symbiodiniaceae family in the Arabian Seas, the Indo-Pacific, and the Atlantic-Caribbean[20,55,56]. This diversity is mainly driven by rapid host specialization, but also by specific environmental conditions[57], suggesting that the diversity observed in this study may be best explained by local adaptation to the specific host genotypes and environmental condition of this remote reef. If correct, this suggests sampling more coral reefs is essential to get a complete view of all possible Symbiodiniaceae symbionts for a coral species. Sequences belonging to the *Symbiodinium* genus (Clade A) were not detected in any coral colony sampled in Kimbe Bay, corroborating previous observations in PNG[37].

Regarding the coral-associated bacterial assemblages, we observed highly distinct phyla. Among those, we found common marine bacteria associated to coral mucus as well as more specific endosymbiotic relationships. Although the sequencing depth is insufficient to detect rare bacterial taxa, our results support that bacterial community composition assessed with the current technology may be used as an indicator of coral health.

The small size of the MinION sequencer coupled to a simple laptop is particularly convenient for use on a research vessel where the work space and electricity use are extremely limited. At present, this device is the only one able to execute a complete sequencing run under these conditions. A − 20 °C freezer is sufficient to conserve reagents for DNA extraction, ONT library preparation, and MinION sequencing, allowing molecular experimentations during long-term expeditions in distant islands in total autonomy. In addition, several new developments ongoing by ONT to lyophilize sequencing reagents will be useful improvements to enable room temperature storage of reagents for several months.

The bioinformatic pipeline developed in this article can be run on a simple laptop without internet access, so long as reference databases are prepared and downloaded before the expedition. The method described in this study will significantly improve the capability of local surveys and enable researchers in remote locations to make informed and effective sampling and experimental decisions while in the field. This approach also carries an element of capacity building in remote areas, as it allows local users to have full access over the data generated and analysed. In remote places, this method could be performed to rapidly evaluate the state of numerous coral colonies on a reef in terms of holobiont diversity and then orient further sampling in the most interesting features like the presence of unexpected bacteria or Symbiodiniaceae or rare coral species. In addition, ONT allows the sequencing of full marker genes instead of short amplicon regions, often insufficient for resolving taxonomies to the species level. Long-read sequencing have already shown their efficiency on coral holobiont characterization[33].

This study has also revealed several limitations of ONT sequencing for diversity analysis of complex samples. We have observed that the combination of low sequencing depth and high error rate limits the detection and the precise assignment of rare sequences. The low sequencing depth in this study was due to the difficult conditions under which the sequencing was conducted, in particular the high temperature (∼ 32 °C). The problem of low sequencing depth is amplified by the difficulty in identifying the correct barcode at each side of the amplicon due to the high error rate. Taken together, the number of de-multiplexed reads is sometimes the limiting factor for the diversity analysis but the current error rate is by far the main challenge to assign correctly each read to a specific taxon. We aimed to address this problem through the generation of consensus sequences from reads aligned to each sequence of the database. Despite this analysis pipeline, very similar sequences diverging only by a single nucleotide were assigned incorrectly as was the case for Symbiodiniaceae delineation and minor ITS2 variants. The difficulty to detect accurately lower abundance ITS2 sequences will limit the assignment of taxa to the level of genus or basal lineage.

These limitations are a significant problem for these metabarcoding analyses with ONT, however, the expected improvements in error rates will impact the detection of rare taxa by diminishing the number of reads necessary to confidently build a consensus, which is presently high. This would lead to delineation of a more complex microbiome. The frequent new versions of the library reagents and the MinION flow cells make the production of stable protocols and analysis pipelines difficult but also indicate that the technology is still progressing which is a very good sign for further diversity studies using ONT.

Coral holobionts require large sampling efforts to obtain an accurate and complete network of the diversity of coral-associated Symbiodiniaceae and bacteria present in a reef. The here-presented study conducted in Papua New Guinea demonstrates the efficacy of the ONT method for a rapid survey of a large number of coral colonies which could help future evaluation of threatened coral reefs and has applications for monitoring -and capacity building.

Data availability
High resolution coral pictures and sequenced reads are accessible at the European Bioinformatics Institute repository under the BioProject PRJEB32905. All accession numbers are indicated in Supplementary Table S1.

References
1. Hughes, T. P. *et al.* Global warming and recurrent mass bleaching of corals. *Nature* **543**, 373–377. https://doi.org/10.1038/nature21707 (2017).
2. LaJeunesse, T. C. *et al.* Systematic revision of symbiodiniaceae highlights the antiquity and diversity of coral endosymbionts. *Curr. Biol.* **28**, 2570–2580. https://doi.org/10.1016/j.cub.2018.07.008 (2018).
3. Bourne, D. G., Morrow, K. M. & Webster, N. S. Insights into the coral microbiome: underpinning the health and resilience of reef ecosystems. *Annu. Rev. Microbiol.* **70**, 317–340. https://doi.org/10.1146/annurev-micro-102215-095440 (2016).
4. Peixoto, R. S., Rosado, P. M., Leite, D. C. D., Rosado, A. S. & Bourne, D. G. Beneficial microorganisms for corals (BMC): proposed mechanisms for coral health and resilience. *Front. Microbiol.* https://doi.org/10.3389/Fmicb.2017.00341 (2017).
5. Reshef, L., Koren, O., Loya, Y., Zilber-Rosenberg, I. & Rosenberg, E. The coral probiotic hypothesis. *Environ. Microbiol.* **8**, 2068–2073. https://doi.org/10.1111/j.1462-2920.2006.01148.x (2006).

6. Lesser, M. P. *et al.* Nitrogen fixation by symbiotic cyanobacteria provides a source of nitrogen for the scleractinian coral Montastraea cavernosa. *Mar. Ecol. Prog. Ser.* **346**, 143–152. https://doi.org/10.3354/meps07008 (2007).

7. Ben-Haim, Y. *et al. Vibrio coralliilyticus* sp. nov., a temperature-dependent pathogen of the coral *Pocillopora damicornis. Int. J. System. Evol. Microbiol.* **53**, 309–315. https://doi.org/10.1099/ijs.0.02402-0 (2003).

8. Johnston, E. C. *et al.* A genomic glance through the fog of plasticity and diversification in Pocillopora. *Sci. Rep.* https://doi.org/10.1038/S41598-017-06085-3 (2017).

9. Shearer, T. L., Van Oppen, M. J., Romano, S. L. & Worheide, G. Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria). *Mol. Ecol.* **11**, 2475–2487 (2002).

10. Hellberg, M. E. No variation and low synonymous substitution rates in coral mtDNA despite high nuclear variation. *BMC Evol. Biol.* **6**, 24. https://doi.org/10.1186/1471-2148-6-24 (2006).

11. Wares, J. P. Mitochondrial cytochrome b sequence data are not an improvement for species identification in scleractinian corals. *PeerJ* **2**, e564. https://doi.org/10.7717/peerj.564 (2014).

12. Arrigoni, R. *et al.* A new sequence data set of SSU rRNA gene for Scleractinia and its phylogenetic and ecological applications. *Mol. Ecol. Resour.* **17**, 1054–1071. https://doi.org/10.1111/1755-0998.12640 (2017).

13. Suzuki, G. & Nomura, K. Species boundaries of Astreopora corals (Scleractinia, Acroporidae) inferred by mitochondrial and nuclear molecular markers. *Zool. Sci.* **30**, 626–632. https://doi.org/10.2108/zsj.30.626 (2013).

14. Gelin, P., Postaire, B., Fauvelot, C. & Magalon, H. Reevaluating species number, distribution and endemism of the coral genus Pocillopora Lamarck, 1816 using species delimitation methods and microsatellites. *Mol. Phylogenet. Evol.* **109**, 430–446. https://doi.org/10.1016/j.ympev.2017.01.018 (2017).

15. LaJeunesse, T. C. Investigating the biodiversity, ecology, and phylogeny of endosymbiotic dinoflagellates in the genus Symbiodinium using the its region: In search of a "species" level marker. *J. Phycol.* **37**, 866–880. https://doi.org/10.1046/j.1529-8817.2001.01031.x (2001).

16. Hume, B. C. C. *et al.* An improved primer set and amplification protocol with increased specificity and sensitivity targeting the Symbiodinium ITS2 region. *PeerJ* **6**, e4816. https://doi.org/10.7717/peerj.4816 (2018).

17. Hume, B. C. C. *et al.* SymPortal: A novel analytical framework and platform for coral algal symbiont next-generation sequencing ITS2 profiling. *Mol. Ecol. Resour.* **19**, 1063–1080. https://doi.org/10.1111/1755-0998.13004 (2019).

18. Arif, C. *et al.* Assessing Symbiodinium diversity in scleractinian corals via next-generation sequencing-based genotyping of the ITS2 rDNA region. *Mol. Ecol.* **23**, 4418–4433. https://doi.org/10.1111/mec.12869 (2014).

19. Smith, E. G., Ketchum, R. N. & Burt, J. A. Host specificity of Symbiodinium variants revealed by an ITS2 metahaplotype approach. *Isme J.* **11**, 1500–1503. https://doi.org/10.1038/ismej.2016.206 (2017).

20. Ziegler, M. *et al.* Biogeography and molecular diversity of coral symbionts in the genus Symbiodinium around the Arabian Peninsula. *J. Biogeogr.* **44**, 674–686. https://doi.org/10.1111/jbi.12913 (2017).

21. Mouchka, M. E., Hewson, I. & Harvell, C. D. Coral-associated bacterial assemblages: current knowledge and the potential for climate-driven impacts. *Integr. Comp. Biol.* **50**, 662–674. https://doi.org/10.1093/icb/icq061 (2010).

22. Hernandez-Agreda, A., Leggat, W., Bongaerts, P. & Ainsworth, T. D. The microbial signature provides insight into the mechanistic basis of coral success across reef habitats. *mBio* https://doi.org/10.1128/mBio.00560-16 (2016).

23. Neave, M. J., Apprill, A., Ferrier-Pages, C. & Voolstra, C. R. Diversity and function of prevalent symbiotic marine bacteria in the genus Endozoicomonas. *Appl. Microbiol. Biotechnol.* **100**, 8315–8324. https://doi.org/10.1007/s00253-016-7777-0 (2016).

24. Hernandez-Agreda, A., Gates, R. D. & Ainsworth, T. D. Defining the Core Microbiome in Corals' Microbial Soup. *Trends Microbiol.* **25**, 125–140. https://doi.org/10.1016/j.tim.2016.11.003 (2017).

25. Roder, C., Bayer, T., Aranda, M., Kruse, M. & Voolstra, C. R. Microbiome structure of the fungid coral Ctenactis echinata aligns with environmental differences. *Mol. Ecol.* **24**, 3501–3511. https://doi.org/10.1111/mec.13251 (2015).

26. Pogoreutz, C. *et al.* Dominance of Endozoicomonas bacteria throughout coral bleaching and mortality suggests structural inflexibility of the Pocillopora verrucosa microbiome. *Ecol. Evol.* **8**, 2240–2252. https://doi.org/10.1002/ece3.3830 (2018).

27. Neave, M. J. *et al.* Differential specificity between closely related corals and abundant Endozoicomonas endosymbionts across global scales. *Isme J.* **11**, 186–200. https://doi.org/10.1038/ismej.2016.95 (2017).

28. Menegon, M. *et al.* On site DNA barcoding by nanopore sequencing. *PLoS ONE* **12**, e0184741. https://doi.org/10.1371/journal.pone.0184741 (2017).

29. Parker, J., Helmstetter, A. J., Devey, D., Wilkinson, T. & Papadopulos, A. S. T. Field-based species identification of closely-related plants using real-time nanopore sequencing. *Sci. Rep.* **7**, 8345. https://doi.org/10.1038/s41598-017-08461-5 (2017).

30. Pomerantz, A. *et al.* Real-time DNA barcoding in a rainforest using nanopore sequencing: opportunities for rapid biodiversity assessments and local capacity building. *Gigascience* https://doi.org/10.1093/gigascience/giy033 (2018).

31. Santos, A., van Aerle, R., Barrientos, L. & Martinez-Urtaza, J. Computational methods for 16S metabarcoding studies using Nanopore sequencing data. *Comput. Struct. Biotechnol. J.* **18**, 296–305. https://doi.org/10.1016/j.csbj.2020.01.005 (2020).

32. Berntson, E. A., Bayer, F. M., McArthur, A. G. & France, S. C. Phylogenetic relationships within the Octocorallia (Cnidaria:Anthozoa) based on nuclear 18S rRNA sequences. *Mar. Biol.* **138**, 235–246. https://doi.org/10.1007/s002270000457 (2001).

33. Pootakham, W. *et al.* High resolution profiling of coral-associated bacterial communities using full-length 16S rRNA sequence data from PacBio SMRT sequencing system. *Sci. Rep.* **7**, 2774. https://doi.org/10.1038/s41598-017-03139-4 (2017).

34. Hume, B. *et al.* Corals from the Persian/Arabian Gulf as models for thermotolerant reef-builders: prevalence of clade C3 Symbiodinium, host fluorescence and ex situ temperature tolerance. *Mar. Pollut. Bull.* **72**, 313–322. https://doi.org/10.1016/j.marpolbul.2012.11.032 (2013).

35. Hume, B. C. *et al. Symbiodinium thermophilum* sp. nov., a thermotolerant symbiotic alga prevalent in corals of the world's hottest sea, the Persian/Arabian Gulf. *Sci. Rep.* **5**, 8562. https://doi.org/10.1038/srep08562 (2015).

36. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100. https://doi.org/10.1093/bioinformatics/bty191 (2018).

37. Noonan, S. H. C., Fabricius, K. E. & Humphrey, C. Symbiodinium community composition in scleractinian corals is not affected by life-long exposure to elevated carbon dioxide. *PLoS ONE* https://doi.org/10.1371/journal.pone.0063985 (2013).

38. Bayer, T. *et al.* Bacteria of the genus Endozoicomonas dominate the microbiome of the Mediterranean gorgonian coral *Eunicella cavolini. Mar. Ecol. Prog. Ser.* https://doi.org/10.3354/meps10197 (2013).

39. Glasl, B., Herndl, G. J. & Frade, P. R. The microbiome of coral surface mucus has a key role in mediating holobiont health and survival upon disturbance. *Isme J.* **10**, 2280–2292. https://doi.org/10.1038/ismej.2016.9 (2016).

40. Morrow, K. M. *et al.* Natural volcanic CO2 seeps reveal future trajectories for host-microbial associations in corals and sponges. *Isme J.* **9**, 894–908. https://doi.org/10.1038/ismej.2014.188 (2015).

41. Morrow, K. M., Bromhall, K., Motti, C. A., Munn, C. B. & Bourne, D. G. Allelochemicals produced by brown macroalgae of the lobophora genus are active against coral larvae and associated bacteria, supporting pathogenic shifts to vibrio dominance. *Appl. Environ. Microb.* https://doi.org/10.1128/AEM.02391-16 (2017).

42. Neave, M. J., Michell, C. T., Apprill, A. & Voolstra, C. R. Endozoicomonas genomes reveal functional adaptation and plasticity in bacterial strains symbiotically associated with diverse marine hosts. *Sci. Rep.* **7**, 40579. https://doi.org/10.1038/srep40579 (2017).

43. Cardenas, A. *et al.* Excess labile carbon promotes the expression of virulence factors in coral reef bacterioplankton. *Isme J.* **12**, 59–76. https://doi.org/10.1038/ismej.2017.142 (2018).

44. Pollock, F. J. *et al.* Coral-associated bacteria demonstrate phylosymbiosis and cophylogeny. *Nat. Commun.* https://doi.org/10.1038/S41467-018-07275-X (2018).
45. Cardenas, A., Rodriguez, L. M., Pizarro, V., Cadavid, L. F. & Arevalo-Ferro, C. Shifts in bacterial communities of two caribbean reef-building coral species affected by white plague disease. *Isme J.* **6**, 502–512. https://doi.org/10.1038/ismej.2011.123 (2012).
46. Gajigan, A. P., Diaz, L. A. & Conaco, C. Resilience of the prokaryotic microbial community of *Acropora digitifera* to elevated temperature. *Microbiologyopen* https://doi.org/10.1002/mbo3.478 (2017).
47. Shnit-Orland, M., Sivan, A. & Kushmaro, A. *Shewanella corallii* sp. nov., a marine bacterium isolated from a Red Sea coral. *Int. J. System. Evol. Microbiol.* **60**, 2293–2297. https://doi.org/10.1099/ijs.0.015768-0 (2010).
48. Ziegler, M. *et al.* Coral microbial community dynamics in response to anthropogenic impacts near a major city in the central Red Sea. *Mar. Pollut. Bull.* **105**, 629–640. https://doi.org/10.1016/j.marpolbul.2015.12.045 (2016).
49. Paramasivam, N. *et al.* Bacterial Consortium of *Millepora dichotoma* exhibiting unusual multifocal lesion event in the gulf of Eilat Red Sea. *Microb Ecol* **65**, 50–59. https://doi.org/10.1007/s00248-012-0097-8 (2013).
50. Paramasivam, N., Ben-Dov, E., Arotsker, L. & Kushmaro, A. *Eilatimonas milleporae* gen. nov., sp. nov., a marine bacterium isolated from the hydrocoral Millepora dichotoma. *Int. J. Syst. Evol. Microbiol.* **63**, 1880–1884. https://doi.org/10.1099/ijs.0.043976-0 (2013).
51. Spring, S., Lunsdorf, H., Fuchs, B. M. & Tindall, B. J. The photosynthetic apparatus and its regulation in the aerobic Gammaproteobacterium *Congregibacter litoralis* gen. nov., sp nov. *PLoS ONE* https://doi.org/10.1371/journal.pone.0004866 (2009).
52. Roder, C. *et al.* Bacterial profiling of White Plague Disease in a comparative coral species framework. *Isme J.* **8**, 31–39. https://doi.org/10.1038/ismej.2013.127 (2014).
53. Sekar, R., Mills, D. K., Remily, E. R., Voss, J. D. & Richardson, L. L. Microbial communities in the surface mucopolysaccharide layer and the black band microbial mat of black band-diseased Siderastrea siderea. *Appl. Environ. Microbiol.* **72**, 5963–5973. https://doi.org/10.1128/AEM.00843-06 (2006).
54. Blackall, L. L., Wilson, B. & van Oppen, M. J. Coral-the world's most diverse symbiotic ecosystem. *Mol. Ecol.* **24**, 5330–5347. https://doi.org/10.1111/mec.13400 (2015).
55. LaJeunesse, T. C. "Species" radiations of symbiotic Dinoflagellates in the Atlantic and Indo-Pacific since the Miocene-Pliocene transition (vol 22, pg 570, 2005). *Mol. Biol. Evol.* **22**, 1158–1158. https://doi.org/10.1093/molbev/msi042 (2005).
56. Hume, B. C. *et al.* Ancestral genetic diversity associated with the rapid spread of stress-tolerant coral symbionts in response to Holocene climate change. *Proc. Natl. Acad. Sci. USA* **113**, 4416–4421. https://doi.org/10.1073/pnas.1601910113 (2016).
57. Thornhill, D. J., Lewis, A. M., Wham, D. C. & LaJeunesse, T. C. Host-specialist lineages dominate the adaptive radiation of reef coral endosymbionts. *Evolution* **68**, 352–367. https://doi.org/10.1111/evo.12270 (2014).

Acknowledgements

Author contributions

J.P, Q.C. and P.W. designed the study. Q.C. wrote the paper with substantial input from J.P., B.H., E.B., C.V., M.Z., S.P. and P.W. E.B collected coral samples. J.P coordinated in situ experiments and ONT sequencing assisted by Q.C. Q.C., B.H., M.Z. and E.B. performed barcoding data analyses. J.P., Q.C., C.C. and S.E. contributed to the development of protocols and tools for ONT sequencing and analysis. All authors discussed the results and commented on the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-72589-0.

Correspondence and requests for materials should be addressed to Q.C. or P.W.

Reprints and permissions information is available at www.nature.com/reprints.

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