Supplementary Materials

Highly structured prokaryote communities exist within the skeleton of coral colonies

Vanessa R. Marcelino, Madeleine J.H. van Oppen and Heroen Verbruggen

Sampling design

Coral skeletons were collected at three sites: three colonies of *Porites lutea* were sampled at Paradise beach (Western Australia), two colonies of *Porites lobata* were sampled at a reef slope in Coral Bay (~ 5 km from Paradise beach, Western Australia), and another three colonies of *P. lobata* were sampled in Heron Island (Queensland, Euclidean distance ~ 3900 km from Paradise beach). GPS coordinates are given in Supplementary table S1. These coral species were chosen because they form large bolder structures where a ~ 2 m transect can be done along a homogeneous surface. Twelve samples were collected from each colony using hammer and chisel, except for one smaller colony where only the first 6 samples were collected (90 samples in total, Supplementary table S1). To avoid environmental effects, samples were collected following the surface of the colony and at the same depth in the water column. Each sample consisted of a piece of skeleton of ~ 0.25 cm^3, situated at 0.5 – 1.5 cm (i.e. 1 cm long piece) below the upper surface of the coral's tissue. All samples were physically similar, and environmental differences are not known to occur at the same depth into the skeletons. The skeleton samples analysed here were underneath a visibly healthy layer of coral living tissue and mucus, and therefore were not in direct contact with the surrounding seawater. It is still unknown how microbes populate coral skeletons, it is possible that lesions (e.g. parrot fish bites) and basal parts of the colony not covered in living tissue constitute major entry points for these microbes.

Samples were collected according to a geometric progression design (Webster & Boag 1992): the distances between successive samples were: 0.4 cm, 1.2 cm, 4 cm, 11 cm, 33 cm, 100 cm, 33 cm, 11 cm, 4 cm, 1.2 cm and 0.4 cm (Supplementary figure S1, Supplementary table S2). The three first and the three last samples from each colony (distances 0.4cm – 4cm) were collected as a single coral fragment, with the exact positions of the samples carved in the coral tissue. Pliers and a Dremel tool were used in the field laboratory to separate samples across smaller distances and obtain the ~ 0.25 cm^3 fragments. Samples were stored in RNAlater (samples collected in Western Australia in 2013) or 100% ethanol (samples collected in Queensland in 2015). We assessed that there were no differences in DDR patterns between sampling sites and preservation methods (Supplementary figure S8).

Library preparation

The DNA isolation and amplification followed previously described protocols (Marcelino & Verbruggen 2016), with the addition of the amplification and sequencing of the Internal transcribed spacer (ITS) region for corals (White *et al.* 1990). The ITS region was amplified with Kapa Taq (Kapa biosystems) following the manufacturer's instructions for the PCR reaction mixture.

The PCR conditions consisted of an initial denaturation step at 94°C for 2 min, followed by 26 cycles of denaturation (94°C for 30s), annealing (45 s at 51°C for the first 6 cycles and 55°C for the remaining 20 cycles) and extension (72°C for 60 s), 20 cycles of denaturation (94°C for 30 s), and a final extension step at 72°C for 7 min. Libraries were quantified with the Quant-iT PicoGreen dsDNA assay kit (Invitrogen), pooled according to their DNA quantity, and sequenced using the Illumina MiSeq platform (2×300 bp paired end reads).

Biodiversity assessment

The initial parsing and quality filtering of the reads was carried out as described in Marcelino and Verbruggen (2016). After quality control, a total of 2,131,374 16S rRNA gene reads, 2,039,715 UPA reads, 1,296,183 *tuf*A reads and 218,315 ITS reads were retrieved and used to cluster Operational Taxonomic Units (OTUs). Sequences were clustered into OTUs using UPARSE (Edgar 2013). A similarity threshold of 98% was set for the *tuf*A and ITS markers, and 97% for the 16S rRNA gene and UPA markers. A taxonomy was assigned to the OTUs using the Naïve Bayesian Classifier (RDP) implemented in QIIME v.1.9.1 (Wang *et al.* 2007; Caporaso *et al.* 2010). The Greengenes v.13.8 dataset (DeSantis *et al.* 2006) was used to classify the 16S rRNA gene sequences, and custom-made reference datasets were used for *tuf*A and UPA (Marcelino & Verbruggen 2017). A known issue with amplicon sequencing is the incorrect assignment of reads to incorrect samples – known as cross-talk or tag-jumping (Schnell *et al.* 2015; Edgar 2016). To address this issue and reduce the risk of false-positives, OTUs with less than 50 reads across all samples and OTUs from samples where they were present with 50 or less reads were removed from the analysis. Chloroplast sequences were excluded from the 16S rRNA gene dataset and bacterial sequences were excluded from the *tuf*A dataset. After this quality control, 2.9% of the 16S rRNA gene sequence reads were Archaea, 0.2% were Cyanobacteria and the remaining 96.9% were other bacterial taxa. For the UPA marker, 0.2% of the reads were Cyanobacteria and the remaining were eukaryotic algae. For the *tuf*A marker, 100% of the reads used in the analyses were eukaryotic green algae. To investigate whether the sequencing effort was deep enough to represent the community, rarefaction curves of the number of observed OTUs per number of reads were constructed by randomly subsampling the reads in QIIME (Supplementary figure S6). To correct for different sequencing depth among samples, a rarefaction threshold was set for each marker where the curve reaches an asymptote – 2500 reads for 16S rRNA gene, UPA and ITS, and 1000 reads for *tuf*A marker. Samples containing less reads than this threshold were excluded from the analyses. After quality control and rarefaction, the remaining number of OTUs were 1,331 in the 16S rRNA gene, 52 in the *tuf*A and 370 in the UPA datasets. Sørensen index and UniFrac distances (Lozupone *et al.* 2011) were calculated for each OTU pair in QIIME.

Distance decay relationship and species accumulation curve

Analyses were carried out separately for the Sørensen, UniFrac and Bray-Curtis distances. These distances represent of community dissimilarity (*d*), and were converted to community similarity

using the formula S = 1- *d*, where S is the community similarity. The rate of the decay in community similarity with distance was calculated for each marker separately as the slope of the linear least squares regression on the relationship between pairwise spatial distance and similarity (Nekola & White 1999; Martiny *et al.* 2011). In cases where pairwise similarity was 0 (i.e. no OTUs in common), it was replaced by the lowest nonzero community similarity observed in the similarity matrix (Martiny *et al.* 2011). The community similarity and spatial distances were $\log_{10}$-transformed prior DDR analyses. The significance of the slope was tested using $\log_{10}$-transformed community dissimilarities and distances with Mantel tests (9,990 permutations), using the *vegan* R package (Oksanen *et al.* 2007).

To investigate the degree of patchiness within colonies, species accumulation curves were calculated as a function of the number of samples included as well as linear distances. This was done for each colony, and using 100 permutations for sample-based curves.

Collection site effects on the Distance Decay Relationship

To investigate whether samples from the three coral reefs had different β-diversities, we coloured the samples in the distance decay curve according to their location. The DDR patterns were similar irrespective of sampling site (Supplementary Figure S7).

Similarity index effect on the Distance Decay Relationship

To test the effects of different similarity metrics on the results, we performed the DDR analysis using Sørensen, Bray-Curtis and weighted UniFrac distances. Since most studies using DDR are based on the Sørensen similarity index, the results based on Sørensen similarities are more comparable to other studies found in the literature. The Sørensen index is a presence-absence metric while Bray-Curtis takes into consideration the relative abundances of the OTUs. The weighted UniFrac metric takes into consideration the relative abundances and the phylogenetic distances among organisms (see Lozupone *et al.* 2011). The resulting slopes were slightly higher when using Bray-Curtis similarity and less step when using UniFrac distances for all groups of organisms. The intra-colony variation was significant for prokaryotes (16S rRNA dataset) and insignificant for algae regardless of the distance metric used (Supplementary table S3).

The reduced rates of species turnover observed when using UniFrac may be explained by its' use of abundance and phylogenetic relatedness data. The weighted UniFrac distance is known to assign a disproportional weight to highly abundant OTUs (Chen *et al.* 2012), and it is possible that the rate of turnover of the most abundant species is slow. Another (non-exclusive) possibility is the existence of competitive exclusion of closely related microorganisms in small patches, as it has been demonstrated for some bacterial strains (Perez-Gutierrez *et al.* 2013; Cordero & Datta 2016). Assuming that closely-related OTUs perform similar ecological functions, the relatively shallow slope observed with UniFrac (compared to the very high slope observed with other metrics) suggests functional redundancy exist despite taxonomic turnover.

Coral and colony identification

Massive coral boulder structures were considered individual colonies when they were not interconnected with each other by coral living tissue. Supplementary figure S1, for example, shows three colonies of *P. lutea.*

Coral species identifications were based on a combination of Internal Transcribed Spacer (ITS) sequences and corallite morphology. Samples for which the ITS region was sequenced are indicated in Supplementary table S1. The 5 most common OTUs (with relative abundances greater than 75%) were aligned with reference ITS sequences of previously identified *Porites* species (Forsman *et al.* 2009; Hellberg *et al.* 2016) using MAFFT v.7.222 (Katoh & Standley 2013). A maximum likelihood tree was constructed with RAxML v.8.2.6 using a GTR + Gamma model (Stamatakis 2006) (Supplementary figure S8). The morphology of the taxa phylogenetically close to the OTUs (Veron *et al.* 2013) were contrasted with pictures and skeleton vouchers of the samples to obtain species level classification.

Although coral colonies can hybridize, all colonies analysed here had only one or two closely-related host ITS sequences, and by unifying ITS sequencing with morphological features we were able to obtain unambiguous species-level identifications. There is no reason to believe that the host genetic background will gradually change with distance within the same colony, and therefore it is unlikely to influence the rates of species turnover observed here.

Distance-Decay relationships within individual coral colonies

To assess the generality of our results across colonies, we calculated the rate of species turnover and the significance of the distance-decay relationship for individual coral colonies. The analyses were based on the Sørensen similarity and were performed as described above. Note however that for colonies with less than 7 samples analysed (i.e. samples retained after rarefaction) it was not possible to perform 9990 Mantel permutations and the maximum number of permutations allowed were used (Supplementary table S4). The results show that different colonies have different rates of species turnover, nevertheless, a significant decay in bacterial community similarity with distance was observed for all but one colony (P8) where only 6 samples were collated and analysed (Supplementary figure S2 A, Supplementary table S4). The algal community showed a more variable pattern, some colonies showed a significant correlation between community similarity and distance, but the slope values (and therefore rates of species turnover) were much shallower than the ones observed for the prokaryotic community (Supplementary figure S2 B-C, Supplementary table S4).

Generality of the results and guidelines for future studies

The results show a high rate of bacterial species turnover and a slow rate of algal species turnover in eight coral colonies, two different coral species and in different locations (~ 3900 km

apart), suggesting that this is a recurring pattern in large massive corals. Branching corals forming monospecific reefs (e.g. *Acropora* spp. and *Porites rus*) might feature an even higher β-diversity due to their structural complexity and size, but that remains to be investigated. We cannot extrapolate our conclusions to smaller reef organisms like oysters, these studies would require an even finer-scale sampling strategy.

Beta-diversity is expected to be higher in environmentally heterogeneous areas. The low β-diversity of endolithic algae observed here may be due to the environmentally homogeneous sampling design and their ability to bore through limestone, and we expect that they can feature higher species turnover across environmental gradients.

The species accumulation curves obtained here help to design sampling strategies in future studies, but it is important to keep in mind that exact number of samples that should be collected per colony naturally depends on the coral species, size of colonies and available resources. For example, if the study focuses on similarly large coral species, at least three samples from different parts of the colony are necessary to capture over 50% of the 16S rRNA gene OTUs, while one sample is enough to observe 50% of green algal OTUs (Supplementary figures S3 and S4). Our prokaryotic OTU accumulation curves did not reach an asymptote, suggesting that more samples would retrieve more species diversity. Thus, if the goal would be to detect the core microbiome of a coral species, then the researchers should aim for the maximum number of samples per colony that is feasible within their timeframe and budget.

Supplementary References:

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **7**, 335-336.

Chen J, Bittinger K, Charlson ES, Hoffmann C, Lewis J, *et al.* (2012) Associating microbiome composition with environmental covariates using generalized UniFrac distances. *Bioinformatics* **28**, 2106-2113.

Cordero OX, Datta MS (2016) Microbial interactions and community assembly at microscales. *Current Opinion in Microbiology* **31**, 227-234.

DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, *et al.* (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology* **72**, 5069-5072.

Edgar RC (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods* **10**, 996-998.

Edgar RC (2016) UNCROSS: Filtering of high-frequency cross-talk in 16S amplicon reads. *bioRxiv*.

Forsman ZH, Barshis DJ, Hunter CL, Toonen RJ (2009) Shape-shifting corals: molecular markers show morphology is evolutionarily plastic in Porites. *BMC Evolutionary Biology* **9**, 45.

Hellberg ME, Prada C, Tan MH, Forsman ZH, Baums IB (2016) Getting a grip at the edge: recolonization and introgression in eastern Pacific Porites corals. *Journal of Biogeography*.

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**, 772-780.

Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R (2011) UniFrac: an effective distance metric for microbial community comparison. *ISME Journal* **5**, 169-172.

Marcelino VR, Verbruggen H (2016) Multi-marker metabarcoding of coral skeletons reveals a rich microbiome and diverse evolutionary origins of endolithic algae. *Scientific Reports* **6**, 31508.

Martiny JB, Eisen JA, Penn K, Allison SD, Horner-Devine MC (2011) Drivers of bacterial beta-diversity depend on spatial scale. *Proceedings of the National Academy of Sciences, USA* **108**, 7850-7854.

Nekola JC, White PS (1999) The distance decay of similarity in biogeography and ecology. *Journal of Biogeography* **26**, 867-878.

Oksanen J, Kindt R, Legendre P, O'Hara B, Stevens MHH, Oksanen MJ, Suggests M (2007) The vegan package. *Community ecology package* **10**.

Perez-Gutierrez RA, Lopez-Ramirez V, Islas A, Alcaraz LD, Hernandez-Gonzalez I, *et al.* (2013) Antagonism influences assembly of a Bacillus guild in a local community and is depicted as a food-chain network. *ISME Journal* **7**, 487-497.

Schnell IB, Bohmann K, Gilbert MT (2015) Tag jumps illuminated--reducing sequence-to-sample misidentifications in metabarcoding studies. *Mol Ecol Resour* **15**, 1289-1303.

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688-2690.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naïve Bayesian Classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology* **73**, 5261-5267.

Webster R, Boag B (1992) Geostatistical analysis of cyst nematodes in soil. *Journal of Soil Science* **43**, 583-595.

White TJ, Bruns T, Lee SJWT, Taylor JW (1990) Amplification and direct sequencing of fungal ribosomal RNA Genes for phylogenetics. In: *PCR protocols: a guide to methods and applications*, pp. 315-322.

Supplementary tables

Highly structured prokaryote communities exist within the skeleton of coral colonies

Vanessa R. Marcelino, Madeleine J.H. van Oppen and Heroen Verbruggen

Supplementary table S1. Coral skeleton samples and geographical coordinates analysed in this study. Please refer to the excel file.

Supplementary table S2: Distances between samples, between and within colonies. Please refer to the excel file.

Supplementary table S3: Slopes of the distance-decay relationships and results of Mantel analyses to test the significance of the correlation between distance and community similarity. Analyses were based on Sørensen similarity distance matrices (presence-absence data) and on weighted UniFrac distance matrices (taking into consideration phylogenetic relatedness between species and their relative abundance).

	16S	**UPA**	***tuf*A**
Sørensen			
DDR slope	-0.1494	-0.0301	-0.0612
Mantel r	0.7059	0.2147	0.3029
Mantel *P*	0.0001*	0.1072	0.0543
Bray-Curtis			
DDR slope	-0.1766	-0.0705	-0.0995
Mantel r	0.698	0.1289	0.1102
Mantel *P*	0.0002*	0.2247	0.2611
UniFrac			
DDR slope	-0.0222	-0.0042	-0.0109
Mantel r	0.4252	0.0644	0.1286
Mantel *P*	0.0128*	0.3428	0.2345

Supplementary table S4. Slopes of the distance-decay relationships and results of Mantel tests for individual colonies. Analyses were based on Sørensen similarity distance matrices. N = number of samples that remained after OTU rarefaction and that were used in this test. Perm = number of permutations used to assess the significance of the Mantel test, this number is limited by the number of samples. *P*-values ≤ 0.05 suggest a significant correlation between distance and community similarity and are indicated with an asterisk.

Prokaryotes - 16S RNA gene					
Colony	**N**	**Perm**	**DDR Slope**	**Mantel *r***	***P***
P1	8	9990	-0.2748	0.8133	0.0027*
P2	7	5039	-0.1009	0.5731	0.0488*
P3	8	9990	-0.1787	0.8421	0.0003*
P4	8	9990	-0.1230	0.7388	0.0006*
P5	9	9990	-0.1009	0.7181	0.0002*
P6	11	9990	-0.0691	0.5288	0.0008*
P7	12	9990	-0.2338	0.8569	0.0001*
P8	6	719	-0.0423	0.3889	0.2000

Algae - UPA					
Colony	**N**	**Perm**	**DDR Slope**	**Mantel *r***	***P***
P1	5	119	-0.0422	0.5889	0.1000
P2	7	5039	-0.0114	0.0619	0.3901
P3	11	9990	-0.1029	0.5658	0.0004*
P4	10	9990	-0.0669	0.5535	0.0004*
P5	8	9990	-0.0419	0.4482	0.0070*
P6	11	9990	0.0030	-0.0089	0.4990
P7	12	9990	-0.0173	0.1080	0.1510
P8	6	719	0.0135	-0.1310	0.5792

Eukaryotic green algae - *tuf*A					
Colony	**N**	**Perm**	**DDR Slope**	**Mantel *r***	***P***
P1	6	719	-0.0757	0.2491	0.1333
P2	5	119	-0.0804	0.1447	0.2667
P3	10	9990	-0.0656	0.3035	0.0330*
P4	8	9990	-0.1638	0.5123	0.0005*
P5	7	5039	-0.0310	0.2694	0.1587
P6	11	9990	-0.0307	0.2985	0.0215*
P7	12	9990	-0.0554	0.4341	0.0015*
P8	6	719	0.0106	-0.0138	0.4750

Supplementary table S5. The strength of the distance-decay relationship (DDR-slope) for endolithic bacteria (16S rDNA) within colonies of *Porites* spp. in comparison with other organisms. The DDR slope of phototrophs was not significantly different from zero and therefore it is not shown. The taxa-area exponents (z-values) calculated from the DDR slope following the method of Harte et al (1999) or calculated independently based on area values (in Zinger et al 2014) are shown. It is important to note that z-values and DDR slope values can vary substantially according to spatial scale (Martiny et al 2011), sequencing approach (Terrat et al 2015) and methodologies (Zinger et al 2014). The z-values reviewed in Horner-Devine et al (2004) represent an average across several studies, therefore scale and sequencing technologies are variable. HTS = High-throughput sequencing.

Organism	DDR slope	z-value	Scale	Sequencing approach	Reference
Plants	-	0.228	-	-	Reviewed in Horner-Devine et al 2004
Birds	-	0.149	-	-	Reviewed in Horner-Devine et al 2004
Butterflies	-	0.101	-	-	Reviewed in Horner-Devine et al 2004
Earthworms	-	0.092	-	-	Reviewed in Horner-Devine et al 2004
Ants	-	0.088	-	-	Reviewed in Horner-Devine et al 2004
Diatoms	-	0.066	-	-	Reviewed in Horner-Devine et al 2004
Ciliates	-	0.060	-	-	Reviewed in Horner-Devine et al 2004
Salt marsh bacteria (97% OTUs)	-0.039	0.020	3cm - 300m	Cloning and sequencing	Horner Devine et al 2004
Lake sediment bacteria	-0.018	0.009	1cm - 1400m	Fingerprint (T-RFLP)	Barreto et al 2014
Lake sediment Archaea	-0.031	0.016	1cm - 1400m	Fingerprint (T-RFLP)	Barreto et al 2014
Desert soil fungi	-0.147	0.074	1m - 100km	Fingerprint (ARISA)	Green et al 2004
Rainforest soil diazotrophs	-	0.060	1-200m	HTS (MiSeq)	Tu et al 2016
Temperate forest soil diazotrophs	-	0.108	1-200m	HTS (MiSeq)	Tu et al 2016
Soil bacteria and Archaea (97% OTUs)	-0.063	0.032	-	HTS (454)	Terrat et al 2015
Seawater (free-living) bacteria	-0.04	0.44	Global	HTS (454)	Zinger et al 2014
Marine sediments bacteria	-0.07	0.51	Global	HTS (454)	Zinger et al 2014
Deep sea waters bacteria	-0.02	0.33	Global	HTS (454)	Zinger et al 2014
Marine surface waters bacteria	-0.03	0.29	Global	HTS (454)	Zinger et al 2014
Coastal waters bacteria	-0.05	0.36	Global	HTS (454)	Zinger et al 2014
Deep sea sediments bacteria	-0.05	0.46	Global	HTS (454)	Zinger et al 2014
Coastal sediments bacteria	-0.08	0.40	Global	HTS (454)	Zinger et al 2014
Endolithic bacteria (97% OTUs)	**-0.149**	**0.075**	**Intra-colony: 0.4cm - 199cm**	**HTS (MiSeq)**	**This study**

Table S5 references:

Barreto DP, Conrad R, Klose M, Claus P, Enrich-Prast A (2014) Distance-decay and taxa-area relationships for bacteria, archaea and methanogenic Archaea in a tropical lake sediment. *PLoS One* **9**, e110128.

Green JL, Holmes AJ, Westoby M, Oliver I, Briscoe D, *et al.* (2004) Spatial scaling of microbial eukaryote diversity. *Nature* **432**, 747-750.

Horner-Devine MC, Lage M, Hughes JB, Bohannan BJ (2004) A taxa-area relationship for bacteria. *Nature* **432**, 750-753.

Martiny JB, Eisen JA, Penn K, Allison SD, Horner-Devine MC (2011) Drivers of bacterial beta-diversity depend on spatial scale. *Proceedings of the National Academy of Sciences, USA* **108**, 7850-7854.

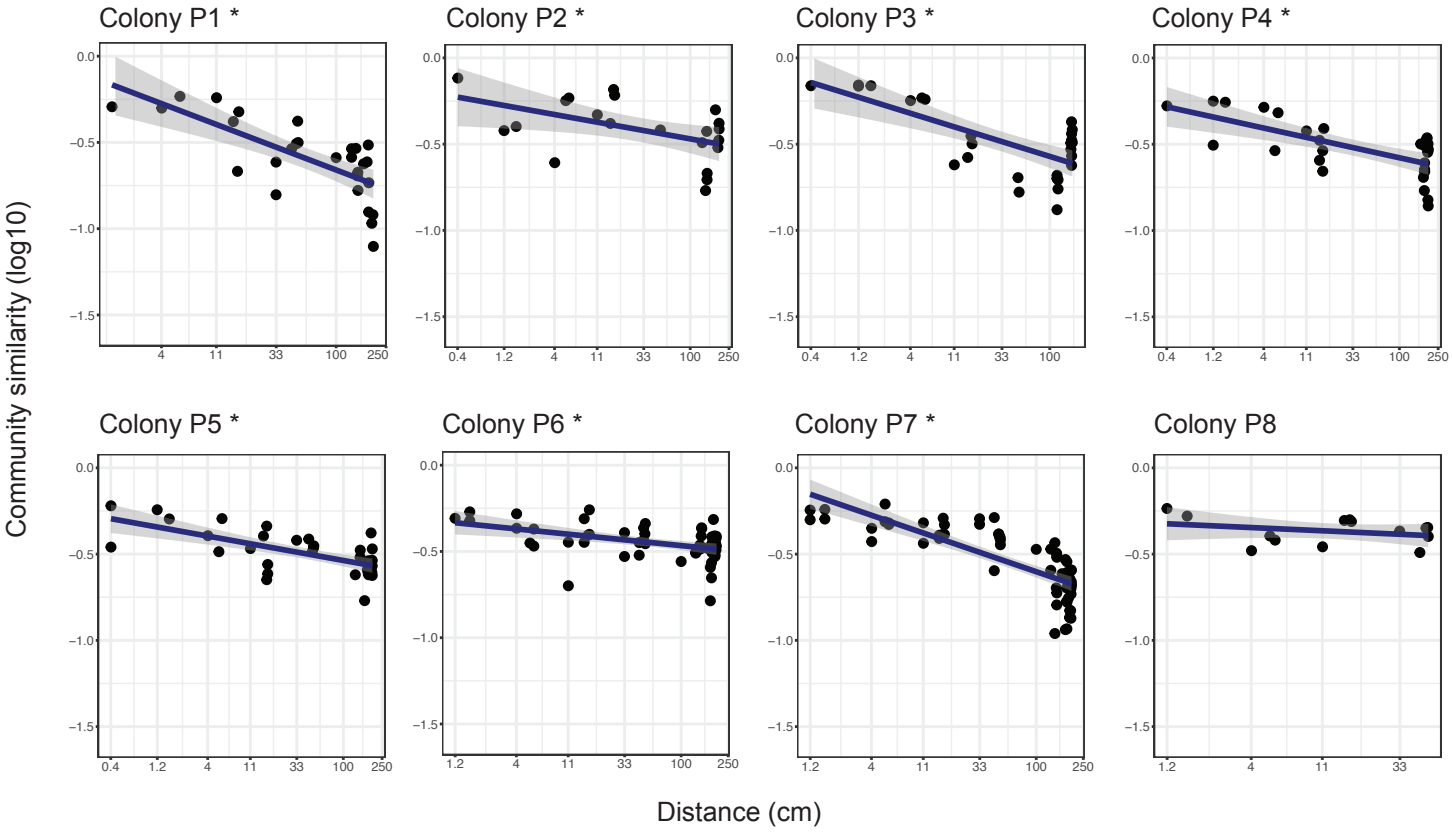
Terrat S, Dequiedt S, Horrigue W, Lelievre M, Cruaud C, *et al.* (2015) Improving soil bacterial taxa-area relationships assessment using DNA meta-barcoding. *Heredity* **114**, 468-475.

Tu Q, Deng Y, Yan Q, Shen L, Lin L, *et al.* (2016) Biogeographic patterns of soil diazotrophic communities across six forests in the North America. *Molecular Ecology* **25**, 2937-2948.

Zinger L, Boetius A, Ramette A (2014) Bacterial taxa-area and distance-decay relationships in marine environments. *Molecular Ecology* **23**, 954-964.

Supplementary Figure S1: Massive *Porites lutea* colonies in Western Australia and sampling design. The red line illustrate the transect along which samples were collected, the varying distances between samples is shown.

A) 16S rRNA gene - Prokaryotes



B) UPA - Algae

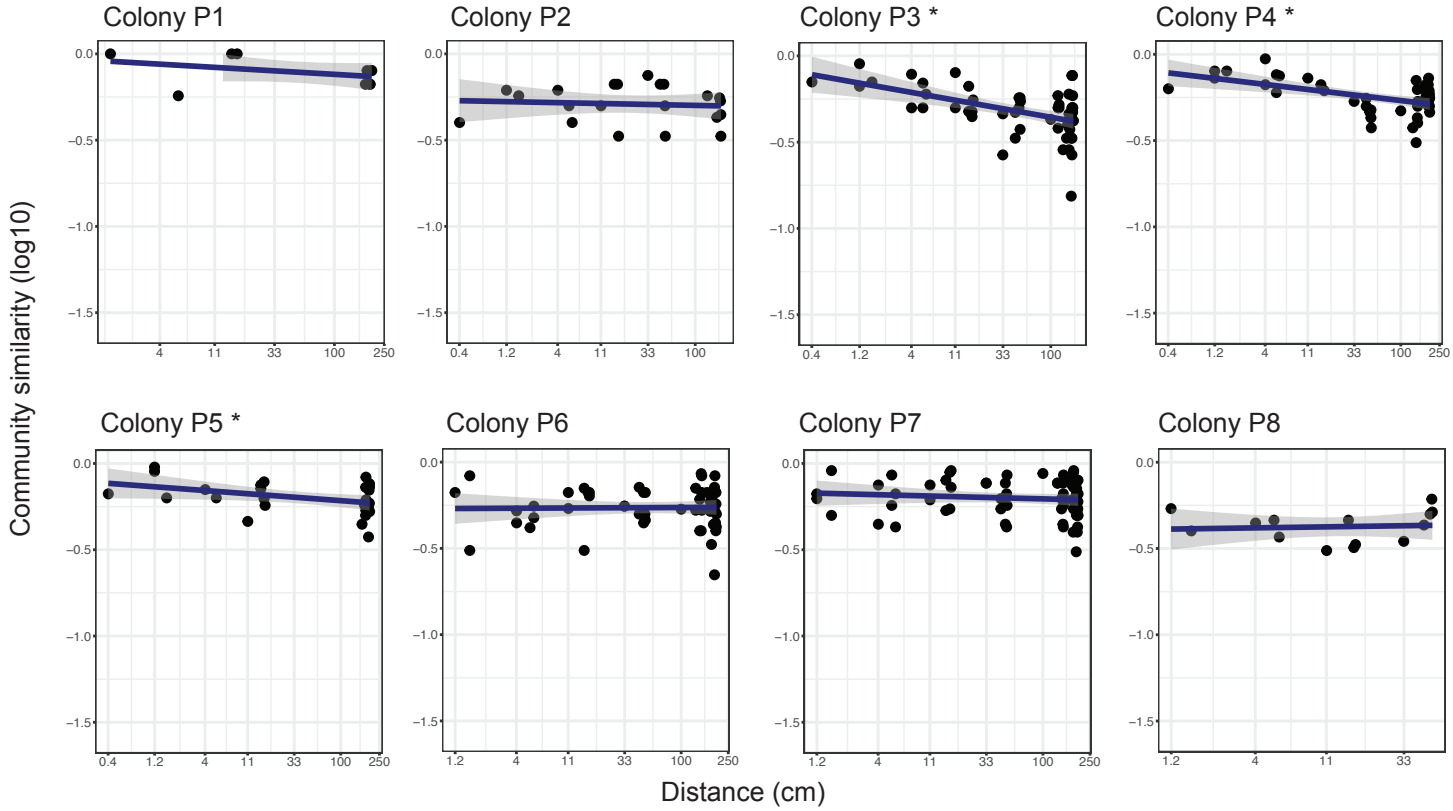
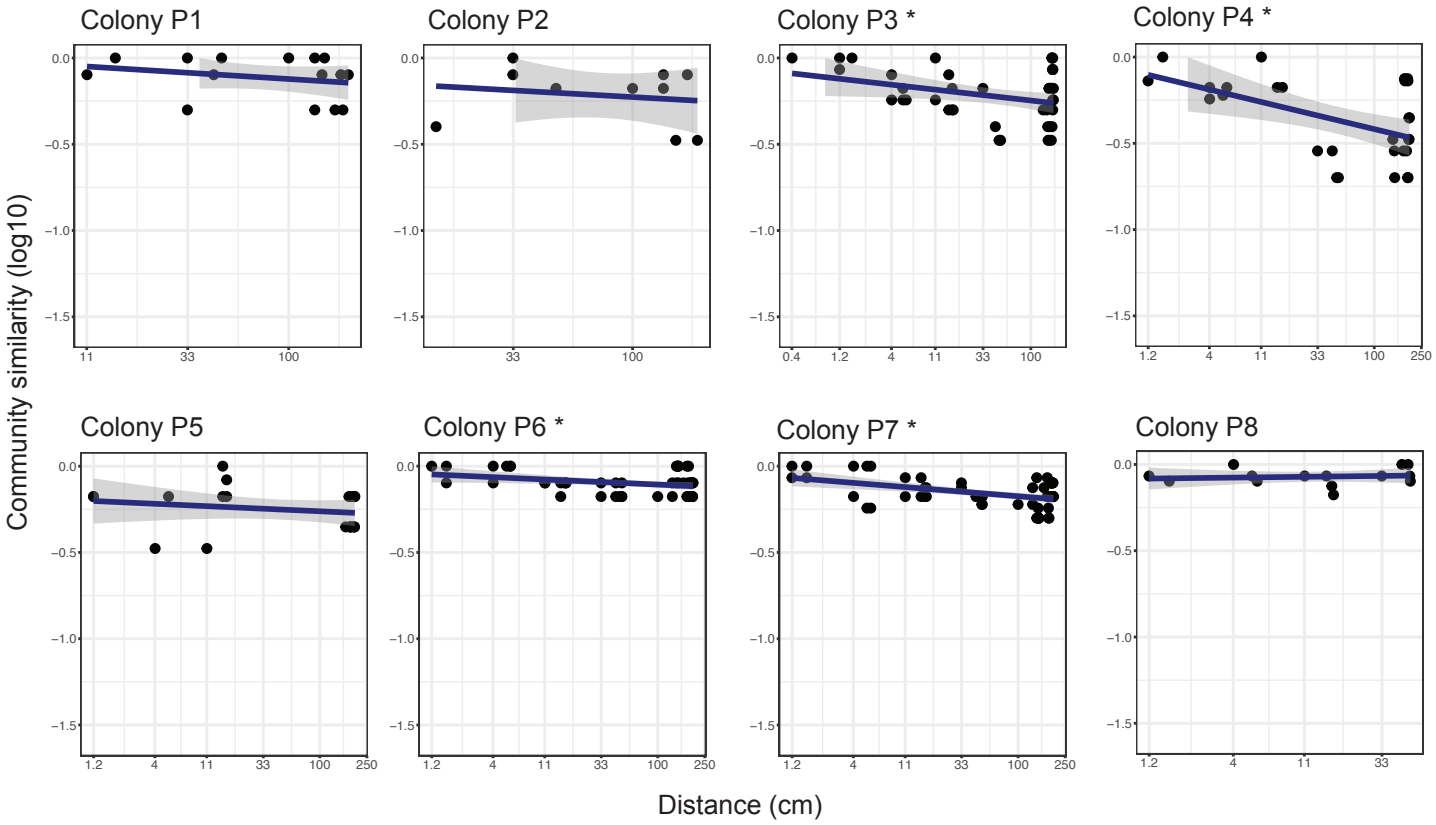
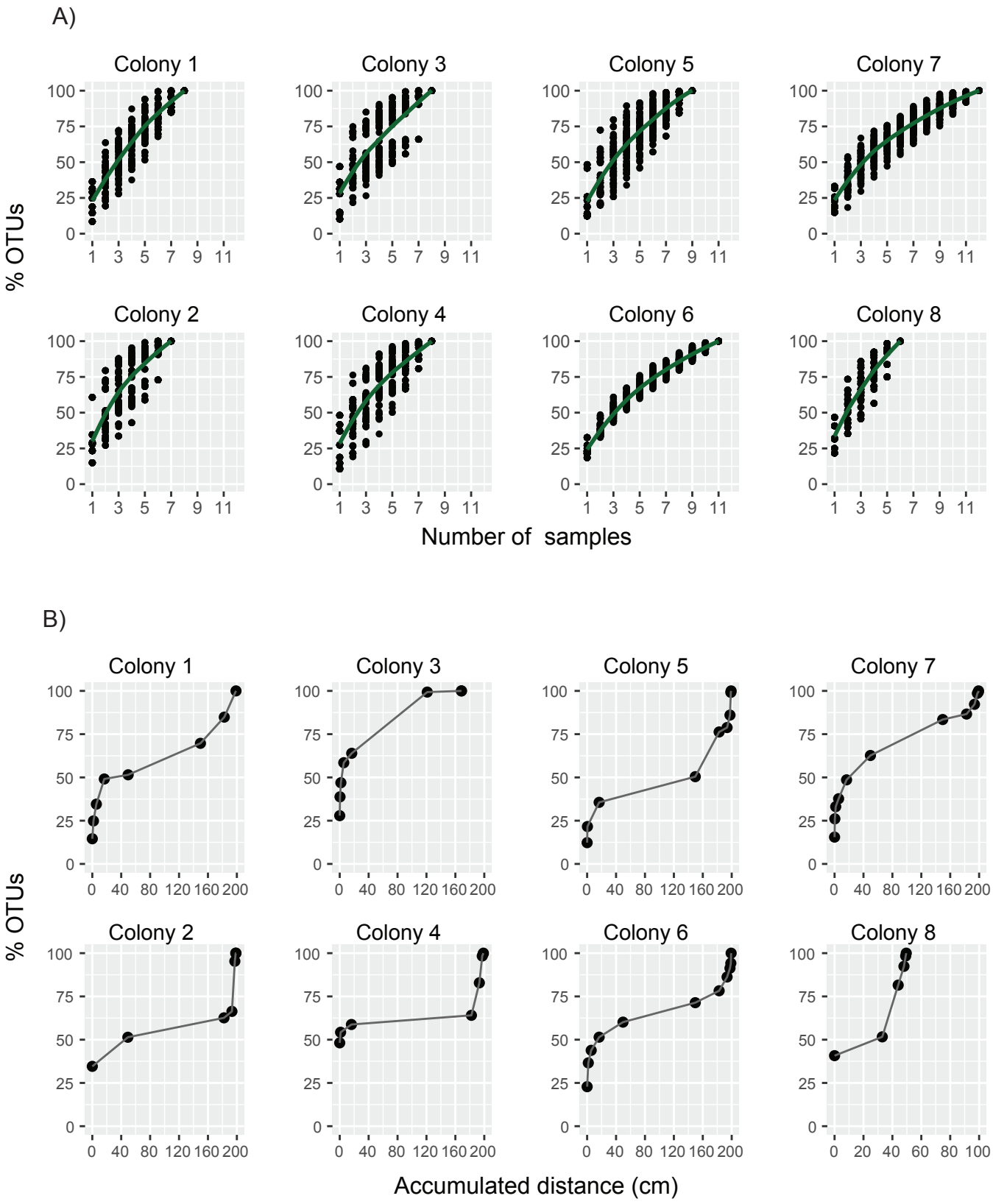


Fig. continues

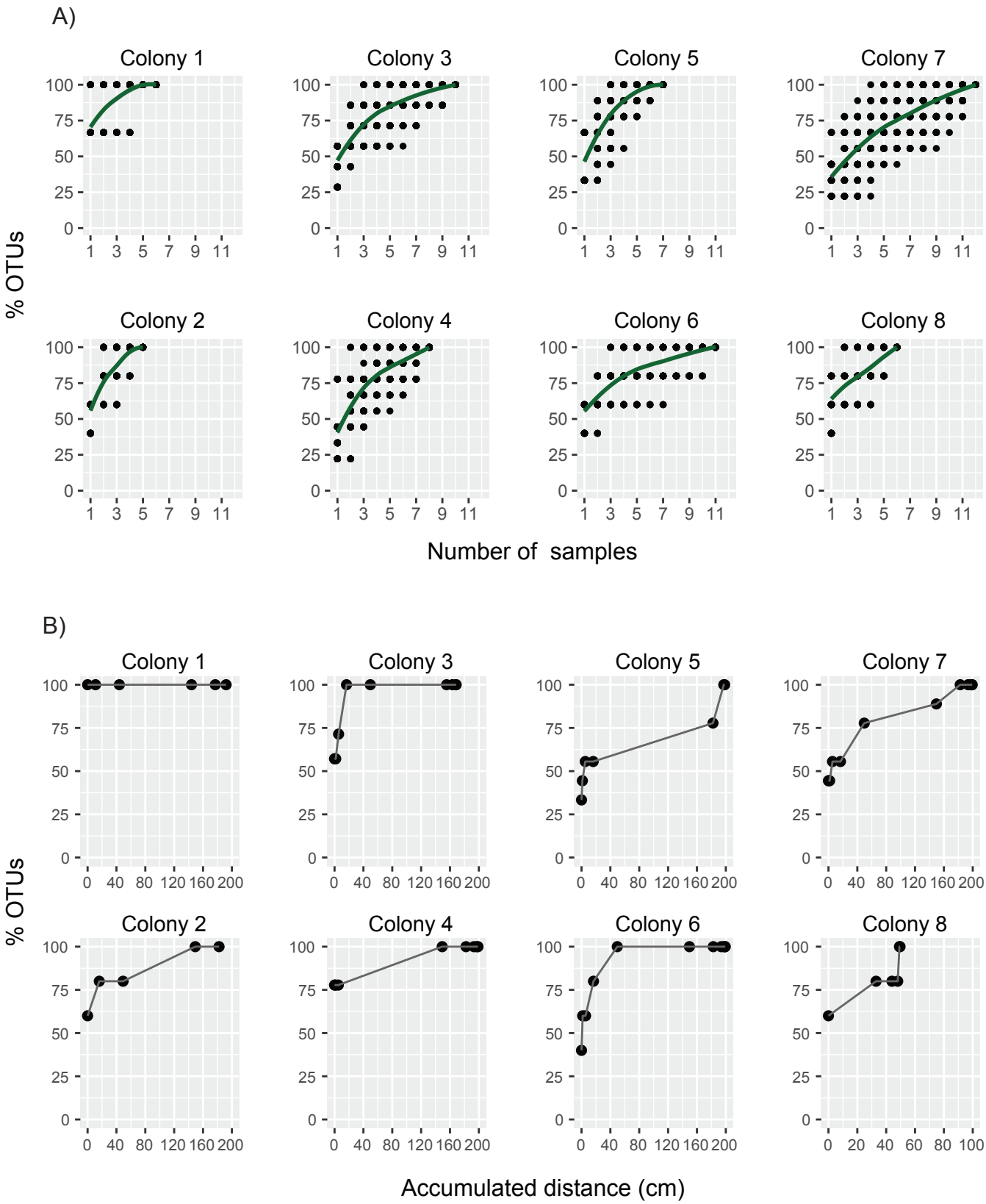
C) *tufA* - Eukaryotic green algae



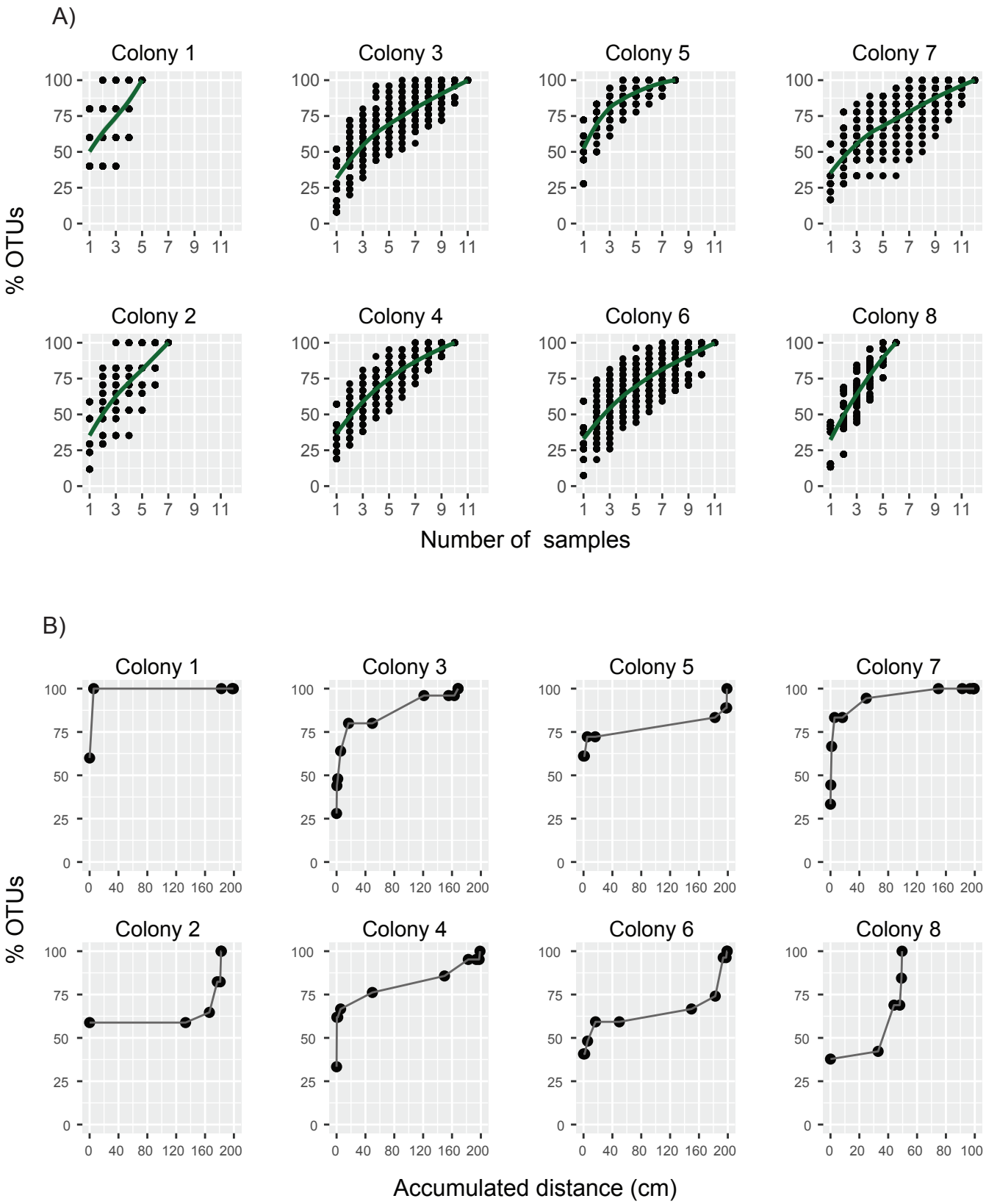
Supplementary figure S2: Distance-decay relationships for the endolithic communities in individual Porites colonies. Blue lines indicate the linear regression between (log_{10} transformed) geographical distance and (log_{10} transformed) Sørensen community similarity. Colonies where a significant decay was observed are indicated with an asterisk (see Supplementary table S4 for slope values, Mantel r and significance values. A – C) Distance-decay relationships obtained with the 16S rRNA gene, UPA and tufA markers respectively.



Supplementary figure S3. Prokaryotic species accumulation curves, based on the 16S rRNA gene marker. A) Percentage of the total number of OTUs observed in each colony that is recovered with increasing skeleton samples. These accumulation curves were obtained by randomizing the samples and storing the recovered OTU percentage 100 times. B) Percentage of the total number of OTUs observed in each colony that is recovered with increasing distance between samples.

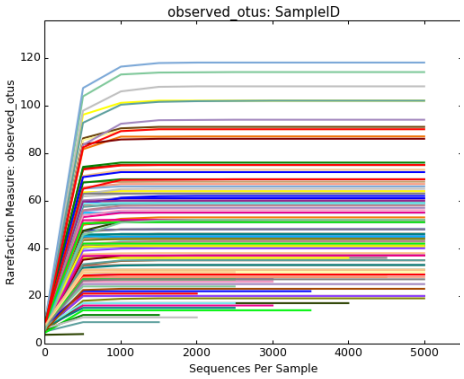


Supplementary figure S4. Species accumulation curves of eukaryotic green algae based on the *tuf*A marker. A) Percentage of the total number of OTUs observed in each colony that is recovered with increasing skeleton samples.These accumulation curves were obtained by randomizing the samples and storing the recovered OTU percentage 100 times. B) Percentage of the total number of OTUs observed in each colony that is recovered with increasing distance between samples.

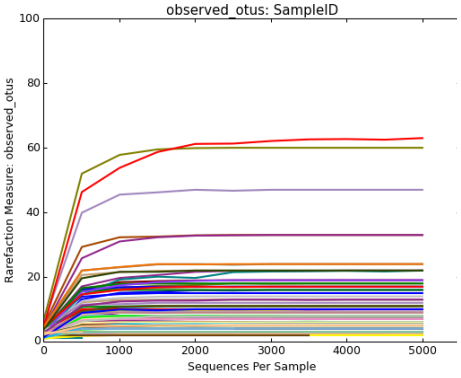


Supplementary figure S5. Species accumulation curves of photosynthetic eukaryotes and cyanobacteria based on the UPA marker. A) Percentage of the total number of OTUs observed in each colony that is recovered with increasing skeleton samples.These accumulation curves were obtained by randomizing the samples and storing the recovered OTU percentage 100 times. B) Percentage of the total number of OTUs observed in each colony that is recovered with increasing distance between samples.

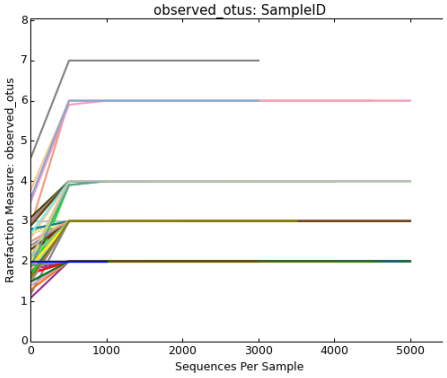
16S rRNA gene



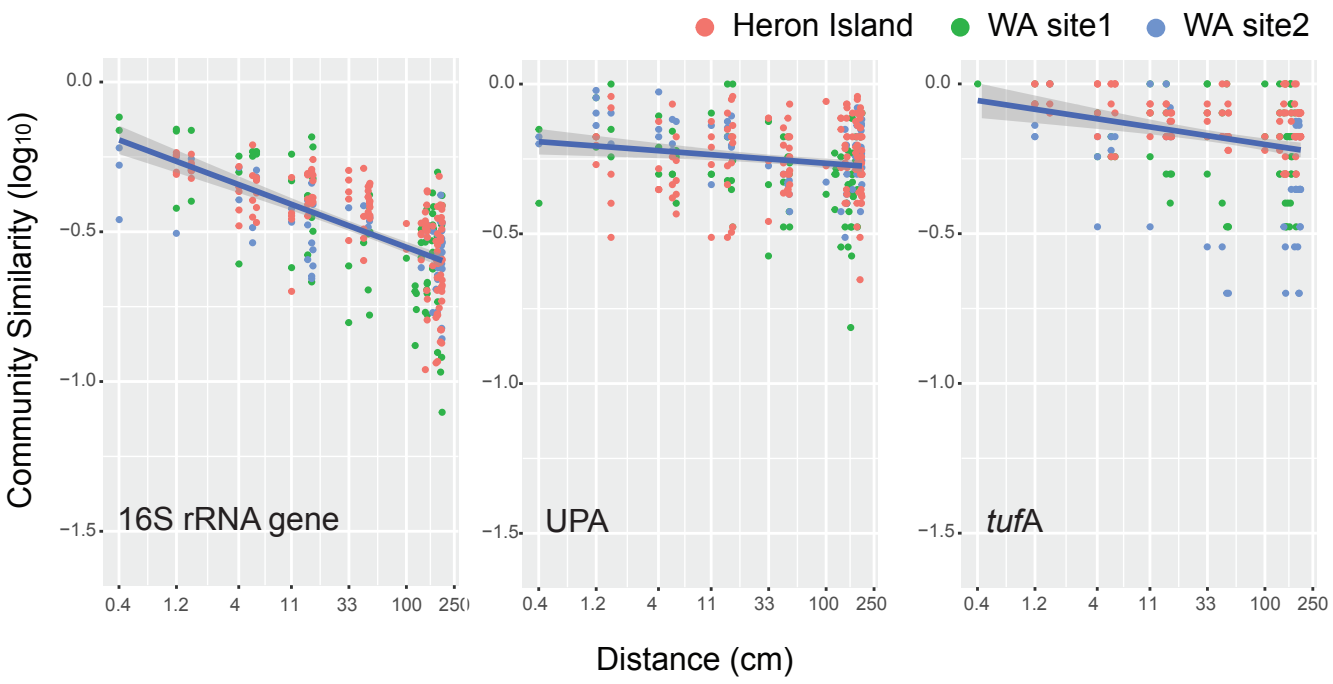
UPA



*tuf*A



Supplementary Figure S6: Alpha rarefaction curves showing the number of OTUs per number of sequences for the different skeleton samples (coloured lines).



Supplementary figure S7: Distance-decay relationships for the prokaryotic (16S rRNA gene) and phototrophic (UPA and *tuf*A) communities found in coral skeletons, highlighting the sites where corals were surveyed.

Supplementary figure S8: Maximum Likelihood tree of ITS sequences. Reference sequences from previously identified corals are in black and the OTUs retrieved from the corals analysed in this study are in red. High bootstrap support values (>80) in internal nodes are indicated. Bootstrap values from terminal nodes and smaller than 80 are ommited.