Symbiosis, hybridization and speciation in Mediterranean octocorals (Octocorallia, Eunicellidae)

Didier Aurelle[1,2*], Anne Haguenauer[3], Marc Bally[1], Frédéric Zuberer[4], Dorian Guillemain[4], Jean-Baptiste Ledoux[5], Stéphane Sartoretto[6], Cédric Cabau[7], Rachel Lapeyre[8], Lamya Chaoui[9], Hichem Kara[9], Sarah Samadi[2], Pierre Pontarotti[10,11,12]

[1] Aix Marseille Univ, Université de Toulon, CNRS, IRD, MIO, Marseille, France
[2] Institut Systématique Evolution Biodiversité (ISYEB), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, CP 26, 75005 Paris, France.
[3] CNRS - Délégation Provence et corse, Marseille, France
[4] Aix Marseille Univ, CNRS, IRD, INRAE, OSU Inst. PYTHEAS, Marseille, France
[5] CIIMAR/CIMAR, Centro Interdisciplinar de Investigação Marinha e Ambiental, Universidade do Porto, Porto, Portugal.
[6] Ifremer, LITTORAL, 83500 La Seyne-sur-Mer, France
[7] Sigenae, GenPhySE, Université de Toulouse, INRAE, ENVT, 31326, Castanet Tolosan, France.
[8] MGX-Montpellier GenomiX, Univ. Montpellier, CNRS, INSERM, Montpellier France
[9] Laboratoire Bioressources marines. Université d'Annaba Badji Mokhtar, Annaba - Algérie.
[10] Aix Marseille Univ, MEPHI, Marseille, France.
[11] IHU Méditerranée Infection, Marseille, France.
[12] CNRS SNC5039

*Corresponding author
Correspondence: didier.aurelle@univ-amu.fr

ABSTRACT

Understanding how species can form and remain isolated in the marine environment still stimulates active researches. Here we study the differentiation and the possibility of hybridization among three temperate octocorals: *Eunicella cavolini*, *E. singularis* and *E. verrucosa*. Morphologically intermediate individuals have been observed between them. Among these three species, *E. singularis* is the only one described in mutualistic symbiosis with photosynthetic Symbiodiniaceae. The symbiosis between Symbiodiniaceae and scleractinian corals is well studied, especially in the context of the response to climate change. Nevertheless, the potential role of symbiotic interactions in speciation processes remains unknown in cnidaria. We tested here the possibility of hybridization between symbiotic and non-symbiotic *Eunicella* species. Through multivariate analyses and hybrid detection, we prove the existence of current gene flow between *E. singularis* and *E. cavolini*, with the observation of F1 and F2 hybrids, and backcrosses. Demographic inferences indicate a scenario of secondary contact between these two species. Despite current gene flow, these two species appear genetically well differentiated. Our data also suggest an intermediate abundance of Symbiodiniaceae in the hybrids of the two species. We discuss the evolution of the Symbiodiniaceae / cnidarian symbiosis in the light of our results.

Keywords: speciation, hybridization, symbiosis, transcriptome, RAD sequencing, octocoral

Introduction

As corner stones of evolutionary biology, species and speciation still raise a wealth of questions fuelled by the technological and conceptual advancements in genomics. Genomic data, either from complete or partial representation of genomes, allow testing hypotheses about species boundaries and origins. Named species are indeed hypotheses, built on available data, that can be rejected or validated through the integration of additional data and / or the use of additional criteria based on evolutionary concepts (Pante *et al.*, 2015b). Sound species delimitations are useful, among others, to better estimate species range and biodiversity patterns (Muir *et al.*, 2022; Coelho *et al.*, 2023), to avoid biases in studies of connectivity (Pante et al., 2015b), and adaptive abilities (Brener-Raffalli *et al.*, 2022). However, proposing sound species delimitation can be problematic because different delimitation criteria may bring contradictory conclusions about species boundaries (the Grey Zone of de Queiroz, 2007). This grey zone corresponds to puzzling cases such as the absence of gene flow among morphologically undifferentiated sets of organisms (i.e. cryptic species, Cahill *et al.*, 2024), or conversely, the detection of gene flow among sets of organisms recognized, based on morphological distinctiveness, as distinct species (Leroy *et al.*, 2020). Evolutionary inferences, based on genomic data, allow testing scenarios of speciation and current gene flow: this provides a better understanding on the origin and persistence of species at the light of genomic divergence (Roux *et al.*, 2016; De Jode *et al.*, 2023).

In the marine realm, the question of speciation is considered as particularly confusing. Notably, how new species can originate from species with large effective size associated to high level of gene flow is still abundantly debated in the literature (e.g. Palumbi, 1992; Mayr, 2001; Faria *et al.*, 2021). Difficulties in sampling and rearing organisms also hamper experiments to test reproductive isolation (Faria *et al.*, 2021). Important progresses in methodologies allow to better understand spatial patterns of genetic structure in marine organisms, for example through the study of oceanographic connectivity (Reynes *et al.*, 2021), clines in allele frequencies (Gagnaire *et al.*, 2015), and hybrid zones (Bierne *et al.*, 2003).

In this context, the role of symbiotic interactions in reproductive isolation remains poorly investigated. There are various examples of the involvement of microbial species in reproductive isolation, especially in insects (Brucker and Bordenstein, 2012). For marine species, microbial communities have been mainly explored in light of adaptative evolution (Rosenberg and Zilber-Rosenberg, 2018). Shallow water scleractinian corals (hexacorals) are usually associated with various species of photosynthetic zooxanthellae, in the family Symbiodiniacae (Cairns, 2007; LaJeunesse *et al.*, 2018). Changes in associated Symbiodiniaceae can impact the thermotolerance of the coral holobiont, and the possibility of adaptation facing climate change (Berkelmans and van Oppen, 2006; van Oppen & Medina,

2020). Inferences from the phylogeny of Anthozoans (hexacorals and octocorals) have shown multiple acquisitions of the symbiotic state throughout evolution (Cairns, 2007; Campoy *et al.*, 2020, Mc Fadden *et al.*, 2021). The symbiotic interactions between Anthozoans and Symbiodiniaceae provide important mutualistic benefits especially from a nutrional point of view (Furla *et al.*, 2005). These interactions require specific adaptations for the animal host, as for example protection against oxygen produced by photosynthesis (Furla *et al.*, 2005). The association with Symbiodiniaceae can also range from mutualism to parasitism (Sachs and Wilcox, 2006; Lesser *et al.*, 2013). Therefore, one can hypothesize that the fitness of hybrids could be impaired by a modification in host – symbiont interactions. The presence of Symbiodiniaceae could also be involved in genetic incompatibilities with the host genome, as previously observed with bacterial species (Bordenstein, 2003; Brucker and Bordenstein, 2012). It is therefore interesting to test the potential role of Symbiodiniaceae in reproductive isolation in Anthozoans.

Here we explore the robustness of species limits between named species of the gorgonian genus *Eunicella* (Octocorallia, Eunicellidae) documented as displaying different symbiotic relationships. In shallow conditions (above 50 m depth), three *Eunicella* species are mainly present in the Mediterranean Sea: *Eunicella cavolini* (Koch, 1887), *E. singularis* (Esper, 1971), and *E. verrucosa* (Pallas, 1766). These three species have partially overlapping ranges, and they can be observed in sympatry, as in the area of Marseille (France). *Eunicella singularis* hosts Symbiodiniaceae corresponding to the *Philozoon* genus (Forcioli *et al.*, 2011; LaJeunesse *et al.*, 2018, 2022; Porro, 2019), whereas the two other gorgonian species are devoided of these symbionts (Carpine and Grasshoff, 1975). The Symbiodiniaceae contribute to the carbon metabolism of *E. singularis*, but a non-symbiotic *aphyta* morph has already been observed (Gori *et al.*, 2012). The lack of variability in mitochondrial DNA does not allow to distinguish these three species (Calderón *et al.*, 2006), and a study using two nuclear introns suggested the possibility of hybridization between *E. singularis* and *E. verrucosa* (Aurelle *et al.*, 2017). Moreover, demographic inferences based on the transcriptome sequences of *E. cavolini* and *E. verrucosa* indicated the possibility of current gene flow between these two species (Roux *et al.*, 2016). However, these data are incomplete because neither individuals identified as *E. singularis*, nor individuals that are morphologically difficult to attribute to a named species (potential hybrids) have been analysed with transcriptomes. Here, we will go further on these topics with the following objectives: i) estimate the genomic differentiation among these three species, ii) test whether differences in the presence of symbionts prevent hybridisation, iii) examine if hybrids show a breakdown of symbiosis, and iv) infer scenarios of speciation. Studying the history of speciation is useful to test is isolation is complete (no gene flow), and how divergence happened. The main analyses presented here are based on transcriptome sequencing which is powerful for such analyses (Roux *et al.*, 2016). We

completed transcriptome data with restriction sites associated DNA sequencing (RAD-sequencing; Baird *et al.*, 2008) to include more individuals in testing species limits and hybridization. The results will be useful to better understand the evolution of these species in different environments and particularly the possible impact of hybridization in adaptation to changing environment.

Material and methods

Species distribution

Eunicella singularis and *E. cavolini* are only present in the Mediterranean Sea, whereas *E. verrucosa* is present both in the Eastern Atlantic Ocean and the Mediterranean Sea (Carpine & Grasshoff, 1975). In the Atlantic, *E. verrucosa* can be found from Ireland, West coasts of Britain to Angola (Grasshoff, 1992; Readman and Hiscock, 2017). *Eunicella verrucosa* is locally present in the North Western Mediterranean Sea with a patchy distribution, in Sardinia (Canessa et al., 2022), and in the Adriatic and Aegean Seas (Chimienti, 2020). In the Mediterranean Sea, it can be observed from shallow conditions (20-40 m) up to 200 m depth (Sartoretto and Francour, 2011; Fourt and Goujard, 2012; Chimienti, 2020). *Eunicella cavolini* is present in the Western Mediterranean, Adriatic and Aegean Seas, from 5 to 200 m depth (Sini *et al.*, 2015; Carugati *et al.*, 2022). *Eunicella singularis* is the only Mediterranean octocoral known to harbour Symbiodiniaceae (but see Bonacolta *et al.*, 2024). These Symbiodiniaceae belong to the temperate clade A (Forcioli *et al.*, 2011; Casado-Amezúa *et al.*, 2016), now corresponding to the *Philozoon* genus (LaJeunesse *et al.*, 2018, 2022). *Eunicella singularis* can be found in the Western Mediterranean and Adriatic Seas, and less frequently in the Eastern Mediterranean (Gori *et al.*, 2012). It is usually observed up top 40 m depth; deeper occurrences (up to 70 m) have been mentioned, which correspond to the *aphyta* morph, without Symbiodiniaceae (Gori et al., 2012). In the area of Marseille, these three species can be observed in sympatry, sometimes at the same depth (Sartoretto and Francour, 2011).

Sampling

For transcriptome sequencing, specimens attributed to *E. cavolini*, *E. singularis*, and *E. verrucosa* have been collected by scuba diving in the Mediterranean (for the three species), and in the Atlantic (*E. verrucosa* only; Table S1; Figure 1) in 2016. In the area of Marseille, the three species have been sampled in sympatry. Four morphologically intermediate individuals (i.e. intermediate colors and branching patterns between *E. cavolini* and

E. singularis; Figure S1) were collected near Marseille to test their hybrid status (Aurelle *et al.*, 2017). The final sampling for transcriptomics included five *E. cavolini*, eight *E. singularis*, three *E. verrucosa*, and four potential hybrids.

We completed the analysis of hybridization and species limits with a higher number of specimens sampled at different periods in the area of Marseille (Figure S2; Table S2). These additional samples have been analysed with RAD sequencing which allows to include more individuals, but generally with less loci than with transcriptomes. The final sampling for RAD sequencing included 25 specimens identified as *E. cavolini*, 23 *E. singularis*, seven *E. verrucosa*, and 12 morphologically intermediate individuals (potential hybrids).

Sampling was non-destructive, with authorizations from the local authorities, including Marine Protected Areas.

To test the genetic proximity of three *Eunicella* species studied here, we built a tree with mitochondrial MutS sequences (McFadden *et al.*, 2011), available in GenBank. The methods and sequences are detailed in supplementary Figure S3, and Table S3.

Transcriptome sequencing

Total RNA has been extracted as in Haguenauer *et al.* (2013). RNAs were sent to the LIGAN genomic platform for sequencing (Lille, France) on four flow cells of Illumina NextSeq 500 (2 x 75 bp). The transcriptomes have been assembled with the *de novo* RNA-Seq Assembly Pipeline (DRAP ; Cabau *et al.*, 2017) with Oases (Schulz *et al.*, 2012) and default parameters. We performed an individual assembly, and a meta-assembly to be used as reference. The statistics describing the assembled transcriptomes are given in Table S1.

We used the BLAT software (Kent, 2002) and the blat_parser.pl script to remove potential Symbiodiniaceae sequences in the obtained transcriptomes, with the transcriptome of the type A1 (Baumgarten *et al.*, 2013) as a reference.

RAD sequencing

DNA has been extracted with the Macherey-Nagel NucleoSpin DNA RapidLyse kit. RAD library preparation (with the PstI restriction enzyme) and sequencing (Illumina NovaSeq600 with 150 nucleotides paired-end sequencing) have been performed at the MGX platform (CNRS). The MGX platform performed control quality, demultiplexing and removal of PCR duplicates with unique molecule identifiers. Potential contaminants have been removed with kraken2 (Wood *et al.*, 2019; Lu *et al.*, 2022). RAD loci have been assembled with ipyrad (Eaton and Overcast, 2020). We tested four assembly strategies to test the robustness of the results: a *de novo* assembly, with a clustering threshold of 0.85, and assembly on a reference genome,

with each of the three available genomes: for *E. cavolini*, *E. singularis*, and *E. verrucosa* (Ledoux et al., in prep).

Analysis of the presence of Symbiodiniaceae

We analysed the presence of Symbiodiniaceae in *Eunicella* gorgonians with transcriptome data. First, we counted the number of reads corresponding to the Symbiodiniaceae transcriptome type A1 with Salmon (Patro *et al.*, 2017). Second, we used the percentage of assembled sequences (contigs) in the *Eunicella* transcriptomes corresponding to Symbiodiniaceae following the BLAT analysis. We used a Kruskal-Wallis test in R to test for differences among the four groups of samples (the three *Eunicella* species and the potential hybrids) for each metric. Additionally, we performed a blast analysis with the LSU, ITS and psbA sequences of *Philozoon* (LaJeunesse *et al.*, 2022) to try to identify the Symbiodiniaceae genera present in the different samples.
As our data pointed to the unexpected presence of Symbiodiniaceae in *E. cavolini* (see Results), we further explored this topic with preliminary data from another experiment dedicated to studying the microbiome of *E. cavolini* and *E. singularis*. This pilot study involved an analysis of microeukaryotic communities through 18S rDNA metabarcoding on two colonies of *E. cavolini*, and one *E. singularis (*Supplementary File S2).

Transcriptomes SNPs calling and filtering

We mapped the reads on the meta transcriptome filtered for Symbiodiniaceae sequences with bwa option mem (Li and Durbin, 2009). The obtained sam files were converted in bam format with samtools 1.9 (Li *et al.*, 2009), and sorted with Picard tools ('Picard Toolkit', 2019). The SNPs calling has been performed with reads2snp 2.0 with default parameters (Tsagkogeorga *et al.*, 2012; Gayral *et al.*, 2013). The obtained dataset, including variable and non variable sites, will thereafter be referred as the "all sites" dataset. We performed separate SNP calls with reads2snp for pairwise comparisons among species and without the potential hybrid samples. These three datasets have been used for demographic inferences, and will be referred as "all-CS" for the *E. cavolini* / *E. singularis* comparison, "all-CV" for the *E. cavolini* / *E. verrucosa* comparison, and "all-SV" for the *E. singularis* / *E. verrucosa* comparison.
For an analysis of genetic differentiation, we filtered the "all sites" vcf file with vcftools (Danecek *et al.*, 2011). We retained biallelic sites, without missing data, and separated by at least 1 kb: this is the "polymorphic sites" dataset. From this dataset, we built a dataset focused on the differentiation between *E. cavolini* and *E. singularis*: we excluded *E. verrucosa* samples

and we retained the first percent of the loci with the highest F_{ST} between *E. cavolini* and *E. singularis*. This last dataset will be referred as "1% SNPs" dataset. The characteristics of the different datasets are summarised in Table S4.

RAD sequencing assembly and SNPs filtering

The vcf files obtained with ipyrad were filtered with vcftools to remove SNPs with more than 50% missing data and separated by less than 1 kb. We used the R package SNPfiltR (De Raad, 2023) for the following additional filters: remove heterozygous genotypes outside from the 0.25-0.75 allele balance range, maximum genotype depth at 200, remove samples with more than 40% missing data, remove SNPs with more than 20% missing data. The four resulting datasets will be referred as "RAD_denovo", "RAD_EC", "RAD_ES" and "RAD_EV" for the *de novo* assembly, and for the assembly on the genome of *E. cavolini*, *E. singularis*, and *E. verrucosa* respectively (Table S4). As with the transcriptome dataset, for each RAD sequencing vcf file, we built a dataset without *E. verrucosa* samples and with the first percent of the loci with the highest F_{ST} between *E. cavolini* and *E. singularis* ("1% SNPs RAD datasets").

Genetic differentiation and analysis of hybrids

With transcriptomes, we analysed the genetic structure and differentiation among species with the "polymorphic sites" dataset. We used the LEA R package to estimate ancestry coefficients (Frichot *et al.*, 2014; Frichot and François, 2015). We tested K values from 1 to 10, with 10 replicates for each K. To analyse the genetic differences among individuals, we performed a Principal Component Analysis (PCA) with the R package adegenet (Jombart, 2008). The pairwise F_{ST} (Weir and Cockerham, 1984) estimated among species were computed with the R package Genepop (Rousset, 2008; Rousset *et al.*, 2020), after conversion of the vcf file with PGDSpider (Lischer & Excoffier, 2012). The distribution of F_{ST} among loci was obtained with vcftools.

The hybrid status (e.g. first generation hybrids) of morphologically intermediate individuals was analysed with the NewHybrids software (Anderson and Thompson, 2002). We used the genepopedit R package to prepare the input file from genepop format (Stanley *et al.*, 2017). Following the results of the LEA and PCA analyses, we focused here on the comparison between *E. cavolini*, *E. singularis* and potential hybrids. The NewHybrids analysis has difficulties to converge with a high number of loci compared to the number of individuals (https://github.com/eriqande/newhybrids/issues/5). We therefore used the "1% SNP" dataset for the NewHybrids analysis. As a prior, we used individuals with the lowest levels of admixture

in LEA as potential parental individuals: this corresponded to three individuals for *E. cavolini*, and six individuals for *E. singularis*. The NewHybrids analysis was repeated ten times to test the robustness of the results.

With RAD sequencing data, we performed the analysis of genetic structure and PCA as above, and we computed pairwise F_{ST} with the four datasets including all loci. Then we performed the NewHybrids analysis as previously described, and separately for the four "1% SNPs RAD datasets", with ten individuals of each parental species as priors.

Scenarios of speciation

We tested scenarios of speciation with the Demographic Inferences with Linked Selection (DILS) pipeline (Csilléry, *et al.*, 2012; Pudlo *et al.*, 2016; Fraïsse *et al.*, 2021) on transcriptome data only. Note that with the high number of loci recovered with transcriptomes, the numbers of specimens used here are adequate for robust inferences (Roux *et al.*, 2016). The DILS pipeline allows the analysis of two species scenarios only: we therefore performed separate analyses for the three two-species comparisons, with the "all-CS", "all-CV", and "all-SV" pairwise datasets. We did not include the potential hybrids in the analysis, which would have required the consideration of a separate population. The tested scenarios are presented in Figure S4 (see Fraïsse *et al.*, 2021 for details). Briefly, DILS allows testing scenario with current migration (i.e. gene flow), such as isolation / migration or secondary contact, versus scenarios of current isolation (no gene flow), such as complete or ancestral migration (gene flow among ancestral populations).

We used the same priors for all analyses, with different numbers of sequences per gene and per sample according to the dataset (Table S5). For all pairwise comparisons, we performed two DILS analyses: one with constant population sizes, and one with variable population sizes.

Results

Mitochondrial MutS

The mitochondrial MutS sequences available in GenBank confirmed the proximity of the three *Eunicella* species analysed here: all sequences were identical for these three species, as well as for three other sequences deposited in GenBank as unidentified *Eunicella* (Figure S3). The closest species to this group was *Eunicella racemosa*. All other *Eunicella* MutS sequences (*E. tricoronata* and *E. albicans*) grouped separately with *Complexum monodi*, but with low bootstrap support.

Presence of Symbiodiniaceae

The transcriptomes showed low numbers of reads counts aligning on the Symbiodiniaceae transcriptome (1868 to 58406 reads; Table S6). The proportion of contigs corresponding to Symbiodiniaceae with BLAT was also very low (between 0.00276 and 0.03686; Table S6). Significant differences were observed among species in both cases (Kruskal-Wallis test, p = 0.047 for reads counts, and p = 0.002 for the proportions of contigs). The pairwise Wilcoxon-Test showed significant differences only for the comparisons of proportions of contigs involving *E. singularis*, which was higher than in other species (Table S7; Figure 2). The mean values of reads counts and contigs for the Symbiodiniaceae in the hybrids were lower than in *E. singularis* and *E. cavolini* but higher than in *E. verrucosa*, although pairwise tests were not significant.

The blast analysis with the LSU, ITS and psbA sequences of *Philozoon* only retrieved corresponding sequences in the transcriptomes of *E. singularis*. Regarding the pilot study of 18S rDNA metabarcoding, a diversity of 92 Operational Taxonomic Units (OTUs) corresponding to Symbiodiniaceae in the Silva database was observed in *E. singularis*, with a single OTU largely dominant in abundance (Supplementary file S2). The same OTU was also observed in *E. cavolini* with a low abundance of reads, but still representing 99% of all 12 to 13 Symbiodiniaceae OTUs detected in the two analysed colonies. A Blast search in GenBank identified a subset of Symbiodiniaceae sequences related to this OTU. Phylogenetic inference based on these data indicated that this OTU was related to clade A of the Symbiodiniaceae.

Genetic differentiation and analysis of hybrids

With transcriptomes, we obtained 31 369 SNPs for the “polymorphic sites” dataset. With this dataset, the highest F_{ST} values were observed for the comparisons between *E. verrucosa* and all other samples (F_{ST} > 0.43; Table S8). The F_{ST} between *E. cavolini* and *E. singularis* was much lower (0.21), and the lowest F_{ST} values were observed for hybrids compared to these two species (F_{ST} around 0.07 in both cases). These differences corresponded to different distributions of F_{ST} over SNPs (Figure S5). For the 1% SNPs with the highest F_{ST} estimates, 52 SNPs were shared by both comparisons involving *E. cavolini* (i.e. *E. cavolini* vs *E. singularis* and *E. cavolini* vs *E. verrucosa*), 116 top 1% SNPs were shared by both comparisons involving *E. singularis*, and 1042 top 1% SNPs were shared by both comparisons involving *E. verrucosa*. The F_{ST} estimates from RAD sequencing were higher than with transcriptome but with similar patterns of relationships among species (Table S9).

The cross-entropy analysis using LEA with transcriptomes indicated a best clustering solution corresponding to K = 2 or K = 3 clusters (Figure S6). At K = 2, the first distinction was

observed between *E. verrucosa* and all other samples (Figure 3). The K = 3 analysis further separated *E. cavolini* and *E. singularis*, with morphologically intermediate individuals admixed between these two species. Conversely the individuals representative of *E. cavolini* and *E. singularis* presented low levels of admixture, apart from the *E. cavolini* of the site in Algeria (code anb), and, at a small level, two *E. singularis* individuals from Banyuls (ban). At K = 4, the two *E. cavolini* individuals from Algeria separated from their conspecifics from the northern part of the Mediterranean.

The cross-entropy analysis using LEA with RAD sequencing showed a minimum at K = 3 for the four datasets (results not shown). The barplots of coancestry coefficients were very similar for the four datasets, with a separation of the three species, and an admixture between *E. cavolini* and *E. singularis* for the morphologically intermediate individuals (Figure S7).

The PCA on transcriptome SNPs separated *E. verrucosa* from other samples on the first axis (33.2% of variance; Figure 4). The second axis (13% of variance) separated *E. cavolini* and *E. singularis*, with the potential hybrids in intermediate position between them. PCA on RAD sequencing gave very similar results (Figure S8).

The NewHybrids analysis with transcriptomes indicated that the morphologically intermediate individuals were indeed hybrids with a probability of one in all ten iterations of the analysis. One individual was a first-generation hybrid, another one was a second-generation hybrid, and the two other ones corresponded to backcrossing with *E. singularis* (Figure 3; Table 1). In the same analysis, the *E. cavolini* and *E. singularis* individuals not included as priors for parental species (see Figure 2 for the individuals used as priors), were indeed inferred as parental with a probability of one, including the *E. cavolini* individuals from Algeria. With RAD sequencing, all morphologically intermediate individuals, except one, appeared as hybrids: F1, F2 or backcrosses with *E. singularis* or *E. cavolini* (Table 1). For four individuals the hybrid status varied according to the dataset: F2 or backcross with *E. cavolini* in two cases, F1 or F2 in two cases. Parental individuals not included in the priors were well inferred as parental with NewHybrids.

Scenarios of speciation

The average pairwise net divergence estimated from DILS was 0.0018 between *E. cavolini* and *E. singularis,* and around 0.007 for the two comparisons with *E. verrucosa* (Table S8, https://zenodo.org/records/12532817/files/results_DILS_suppl_file.ods?download=1). The DILS analysis indicated the existence of current gene flow between *E. cavolini* and *E. singularis* with high probability, both with constant and variable population sizes (p = 0.87 and 0.88 respectively; Table 2). This possibility of gene flow corresponded to a scenario of secondary contact. Conversely, a model of current isolation was inferred for the comparisons

between *E. verrucosa* and each of the two other species, with a probability p ≥ 0.87: in these two cases, the inferred scenario included a period of ancestral migration, though with moderate support (p between 0.61 and 0.69). A genomic heterogeneity in effective size (i.e. variations among loci) was inferred with strong support (p ≥ 0.99) for all analyses. In the case of current gene flow (between *E. cavolini* and *E. singularis*), a genomic heterogeneity in migration rates was inferred (p ≥ 0.82). The inferred parameters for the different scenarios are presented in Supplementary Table S9. We will first present the results obtained for the constant population sizes models. The divergence time between *E. cavolini* and *E. singularis* (median 403 273 generations) was much lower than between *E. cavolini* and *E. verrucosa* (median 1 054 488 generations), and between *E. singularis* and *E. verrucosa* (median 899 098 generations). For the comparison between *E. cavolini* and *E. singularis,* the time of secondary contact was estimated after around 85% of time spent in isolation since divergence. Following secondary contact, the gene flow was similar in both directions for these two species. The duration of ancestral migration roughly corresponded to 6% and 8% of the total time since divergence for the comparison between *E. cavolini* and *E. verrucosa*, and for the comparison between *E. singularis* and *E. verrucosa*, respectively. For these last two cases, the gene flow (forward in time) during ancestral migration was higher towards *E. verrucosa* than in the opposite direction. The estimated effective sizes were of similar order for *E. cavolini* and *E. verrucosa*. Similar results were obtained for the models including variations in effective size, except for the estimate of current gene flow between *E. cavolini* and *E. singularis*: with variable population size, gene flow from *E. singularis* to *E. cavolini* was higher than in the opposite direction.

Discussion

The three named *Eunicella* species studied here have been previously described with differences in colony morphology, sclerites shape, and in the presence of photosynthetic Symbiodiniaceae. Our results demonstrate a continuum between *E. cavolini* and *E. singularis*, with morphologically intermediate individuals, current gene flow, and hybrids characterised by a reduced frequency of Symbiodiniaceae compared to *E. singularis*. On the other hand, *E. verrucosa* appears genetically isolated from these two species. We will discuss here the differences observed among markers, the outcome of hybridization, the speciation scenarios, and what can be learnt on the evolution of symbiosis.

Discordances between molecular markers

As previously observed (Aurelle *et al.*, 2017), mitochondrial DNA did not allow to discriminate the three species due to the usually slow evolution of mitochondrial DNA in octocorals (McFadden *et al.*, 2011; Muthye *et al.*, 2022). The use of transcriptome sequences first confirmed the closer proximity between *E. cavolini* and *E. singularis* than with *E. verrucosa*. This had been previously suggested with two intron sequences, but with incomplete lineage sorting (Aurelle *et al.*, 2017). The Mediterranean *Eunicella* then add a new example of the lack of power of mitochondrial DNA to discriminate genetically differentiated octocoral species, as demonstrated in other genera (Erickson *et al.*, 2021; Pante *et al.*, 2015a). The slow rate of evolution of mitochondrial DNA in octocorals has been linked to the presence of the mitochondrial locus MutS, an homolog of a bacterial gene involved in DNA repair. However, there are contradictory examples showing that the presence of this locus is not the only factor explaining the slow evolution of mitochondrial DNA in octocorals (Muthye *et al.*, 2022). More generally, as hybridization can lead to the sharing of mitochondrial DNA among species, the use of multiple independent nuclear loci is required for species discrimination in such cases.

Incomplete reproductive isolation among two named species

Inferences of genetic ancestry and hybrid status confirmed that morphologically intermediate individuals are indeed hybrids between *E. singularis* and *E. cavolini*, with the identification of F1, F2 and backcrosses with both parental lineages: first generation hybrids can then be fertile. The fact that gene flow indeed goes further than the hybrid levels is confirmed by the DILS analysis, which did not include hybrid individuals. Reproductive isolation is therefore at least partial between these lineages. The ease to find hybrids in the area studied here, as well as similar observations in other sites (S. Sartoretto, pers. com.) indicate that hybridization is not rare on an evolutionary scale.

The alternation of populations with and without hybrids would point to a mosaic hybrid zone (Bierne *et al.*, 2003), where hybrids could form in different areas and from different parental populations. As, or because, hybridization between *E. cavolini* and *E. singularis* had not been reported before, the presence of hybrids has probably been overlooked up to now. This may be the consequence of previously focusing on colonies with “typical” morphologies. The frequency of hybridization therefore remains to be studied.

Our results allow discussing the evolution of genomic divergence among these species. The persistence of genomic differentiation between these lineages in sympatry, despite current gene flow, indicates that genetic incompatibilities must exist, potentially coupled with differences in adaptation to local environments (Bierne *et al.*, 2011). A better characterization of the ecological range of parental and hybrid populations would be useful to test if local adaptation is involved in their distribution. A genome wide analysis of differentiation is also

required to investigate whether divergence between *E. cavolini* and *E. singularis*, is homogeneous along the genome (as suggested by the DILS analysis which inferred a homogeneity of gene flow), or whether genomic islands of differentiation exist (Peñalba *et al.*, 2024). We could then better understand to what stage of divergence the *E. cavolini* / *E. singularis* split corresponds: from intra-specific polymorphism to species separated by semipermeable barriers to gene flow.

One interesting question in this context is whether changes in selection regimes induced by human activities can change the outcome of hybridization (Ålund *et al.*, 2023). For example, Mediterranean octocorals are impacted by mortality events linked with climate change (Sini *et al.*, 2015; Estaque *et al.*, 2023), and the impact of these events could be different for hybrids and parental individuals. In scleractinian corals, interspecific hybridization has been reported to enhance the survival under elevated temperature conditions (Chan *et al.*, 2018) .

Regarding *E. verrucosa*, the more ancient divergence corresponded to much more loci with high F_{ST}. Among the list of the most highly differentiated loci, more overlap was also observed for the two comparisons involving *E. verrucosa* than for the other pairwise comparisons: this may indicate that few genomic areas of potential incompatibilities with *E. verrucosa* are involved in the divergence between *E. cavolini* and *E. singularis*.

Scenarios of speciations

The scenarios of speciations inferred with DILS supported the current isolation (no gene flow) of *E. verrucosa* with the two other species with high posterior probability. Conversely current gene flow was strongly supported versus isolation between *E. cavolini* and *E. singularis*. The posterior probabilities for ancestral migration (for *E. verrucosa* versus the two other species), and secondary contact (*E. cavolini* and *E. singularis*), were lower than for inferences on current gene flow. These scenarios were indeed the best ones among those tested here but they might not provide the best possible representation of the evolutionary history. Other models of evolution could be tested for better inferences, for example by including the three species and hybrids, or gene flow from unsampled taxa (Tricou *et al.*, 2022). The current isolation of *E. verrucosa* from *E. cavolini* is also at odds with previous results which showed the possibility of current gene flow between these two species despite an important divergence (Roux *et al.*, 2016). It will be useful to explore the reasons for the discrepancy between this study and the present one, which are both based on transcriptome datasets but obtained from different samples and sequencing platforms.

Eunicella verrucosa is currently widely distributed in the North Eastern Atlantic Ocean, and less frequent in the Mediterranean Sea, whereas both other species are only present in the Mediterranean Sea. The Atlantic / Mediterranean Sea transition does not seem to act as a

phylogeographic barrier for *E. verrucosa* (Macleod *et al.*, 2024). We can propose a scenario where the split between *E. verrucosa* and both other species occurred in allopatry between the Atlantic Ocean and the Mediterranean Sea, followed by the colonization of the Mediterranean Sea by *E. verrucosa*. The generation time remains unknown for the *Eunicella* species, and previous studies have shown important variation in the age at first reproduction in gorgonians, from 2 to 13 years (see references in Munro, 2004). If we use a generation time of two years for *Eunicella* species, with a median estimate of divergence time around 900 000 generations for *E. verrucosa* / *E. singularis* and 1 000 000 for *E. verrucosa* / *E. cavolini*, and based on a mutation rate set at 3.10^{-9}, this would indicate a divergence at least around 2 000 000 years (2 Ma). The divergence time between *E. cavolini* and *E. singularis* would be 2.5 times more recent, around 800 000 years, with a median time of secondary contact around 60 000 generations, corresponding to 15% of the time spent since divergence. It is difficult to infer past distributions of *E. singularis* and *E. cavolini*, but one can note that even if they are currently found in sympatry in different areas, their range do not completely overlap. For example *E. cavolini* is nearly absent at the West of the Rhone estuary on the French coast, whereas *E. singularis* is present there. The ecological range of *E. singularis* and *E. cavolini* is also not completely overlapping, as *E. cavolini* can be observed deeper than *E. singularis* (Gori *et al.*, 2012; Carugati *et al.*, 2022). Therefore one can envision an historical separation of these two species either geographically or ecologically, followed by a secondary contact where gene flow took place. In any case, additional information on generation time, mutation rate and past demographic fluctuations are required to be more precise on the history of these species.

Evolution of symbiosis

As previously discussed, we clearly demonstrated here the possibility of gene flow between symbiotic (i.e. hosting Symbiodiniaceae) and non-symbiotic octocorals. Symbiodiniaceae could nevertheless be involved in genetic incompatibilities with the genome of some cnidarian hosts, but this would require additional analysis of symbiotic status in hybrids. The methods used here did not aim at a precise quantification of Symbiodiniaceae, and one can note the low levels of sequences corresponding to these symbionts, even in *E. singularis,* which may be due to difficulties in extracting the RNA of the symbionts (but see Guzman *et al.*, 2018; Rivera-García *et al.*, 2019). Despite these limits we observed, as expected, a higher Symbiodiniaceae concentration in *E. singularis* than in *E. cavolini* and *E. verrucosa.* Interestingly, the hybrids showed a lower frequency of Symbiodiniaceae than *E. singularis*, and possibly than *E. cavolini*, though this last result remains to be confirmed. In *E. singularis*, the transmission of Symbiodiniaceae seems to occur both vertically, through ovules, and

horizontally, from the environment (Forcioli *et al.*, 2011). Both transmission modes did not restore the levels of Symbiodiniaceae in the hybrids to those of *E. singularis*. This suggests a breakdown of or failure to establish symbiosis for hybrid genotypes, which may impact the fitness of hybrids and consequently the possibility of introgression. The *aphyta* type of *E. singularis* observed in deep conditions indicates a plasticity of symbiotic status apart from hybridization. Nevertheless, the hybrids were sampled here in shallow conditions (10-20 m depth) which underlines the role of hybridization in reducing the extent of symbiosis. More precise estimates of Symbiodiniaceae abundance, and of physiological parameters such as photosynthetic and respiration rates (Ezzat *et al.*, 2013). would help understanding the role of symbionts in hybrids fitness.

Our results also question the evolution and significance of octocoral / Symbiodiniaceae symbiosis. In scleractinians, the transition between symbiotic and non-symbiotic states happened repeatedly, but mostly in the direction of the acquisition of symbiosis, with very low rates of reversal (Campoy *et al.*, 2020). This could indicate that investing in such mutualistic interactions for the cnidarian would lead to increasingly relying on autotrophy for energetic supply, making reversal to heterotrophy difficult. In octocorals, an evolutionary versatility in symbiotic state seems possible, as in various families and genera, both symbiotic and non-symbiotic species are present (Van Oppen *et al.*, 2005). In the Mediterranean Sea, all octocoral species are non-symbiotic, except for *E. singularis* (but see Bonacolta et al. 2024). The most parsimonious scenario here would be an acquisition of symbiosis in *E. singularis* during or following its divergence from *E. cavolini*. The symbiotic status of *E. singularis* nevertheless could be facultative as previously mentioned for the *aphyta* type (Gori *et al.*, 2012). Additionally, experimental physiological studies have demonstrated the nutritional plasticity of *E. singularis* which is able to use either heterotrophy or autotrophy for its metabolism (Ezzat *et al.*, 2013). Nevertheless, in natural conditions, autotrophy seems to provide an important contribution to the metabolism of *E. singularis*, and the collapse of photosynthetic capacities in too warm conditions could contribute to mortality events in this species (Coma *et al.*, 2015).

The question of symbiosis could be reversed as well: why are Symbiodiniaceae not more abundant in *E. cavolini*? This species can be observed in shallow conditions (less than 10 m depth) where there is enough light for photosynthesis, and in syntopy with *E. singularis*. The availability of preys or particulate organic matter may provide enough energy to *E. cavolini* in its habitat, but this species may have never engaged in mutualistic interaction with Symbiodiniaceae. Interestingly we observed a low rate of sequences related to Symbiodiniaceae in the transcriptomes of *E. cavolini* (and even lower, but not null in *E. verrucosa*). This could either correspond to a signal from free living Symbiodiniaceae, or to rare, transient, associations with the cnidarian. In addition, a Symbiodiniaceae OTU that is

common to *E. singularis* and *E. cavolini* was identified among the microeukaryotes associated with the two species: this OTU is related to strains observed in symbiosis with *E. singularis* and other cnidarians. Molecular markers also allowed to evidence the presence of Symbiodiniaceae in species previously supposed to be asymbiotic, as in the Mediterranean octocoral *Paramuricea clavata*, and in several Hawaiian antipatharian species (Wagner *et al.*, 2011; Bonacolta *et al.*, 2024). These results, and our observations in *Eunicella* species, obviously underline the dynamic nature of interactions between Symbiodiniaceae and cnidarians: the establishment of symbiosis may be preceded by more or less stable, and more or less mutualistic interactions. The development of effective symbiosis, with stable relationships, and higher abundance of symbiont, would require specific adaptation from both partners. We can see here that even if on a macro-evolutionary scale the acquisition of symbiosis is much more frequent than its loss, on a micro-evolutionary scale the gene flow between the *Eunicella* species considered here has not led to the full development of symbiosis in *E. cavolini*.

Conclusions and perspectives

We demonstrated the lack of genetic isolation between octocorals with contrasted levels of mutualistic interaction with Symbiodiniaceae. Understanding the evolution and adaptation of these species in heterogeneous environments should then take into account the possible impact of introgression. We also show that symbiosis is more flexible that previously envisioned in octocorals. For these species it will be useful to estimate the frequency and spatial extent of hybrid zones: does it correlate with particular environments with a coupling between endogenous and exogenous barriers to gene flow (Bierne *et al.*, 2011)? Characterizing the genomic landscape of introgression would help to look for the effects of introgression on adaptation or symbiosis for example. Indeed, even low levels of interspecific gene flow can have important consequences on the evolution of species (Arnold *et al.*, 1999). Finally, various cases of hybridization have been demonstrated in symbiotic scleractinian corals (e.g. Combosch and Vollmer, 2015): it would then be interesting to study the dynamics of symbiosis in these cases, especially when different Symbiodiniaceae strains are involved.

Acknowledgements:

We thank the ECCOREV Research Federation (FR 3098) for the financial support of part of this study (https://www.eccorev.fr/). The project leading to this publication has received funding from European FEDER Fund under project 1166-39417. The project leading to this

publication has received funding from Excellence Initiative of Aix-Marseille University - A*MIDEX, a French “Investissements d’Avenir” programme. The authors thank the UMR 8199 LIGAN-PM Genomics platform (Lille, France, especially Véronique Dhennin) which belongs to the 'Federation de Recherche' 3508 Labex EGID (European Genomics Institute for Diabetes; ANR-10-LABX-46) and was supported by the ANR Equipex 2010 session (ANR-10-EQPX-07-01; 'LIGAN-PM'). The LIGAN-PM Genomics platform (Lille, France) is also supported by the FEDER and the Region Nord-Pas-de-Calais-Picardie. JBL was supported by the strategic funding UIDB/04423/2020, UIDP/04423/2020 and 2021.00855.CEECIND through national funds provided by FCT -Fundaço para a Ciência e a Tecnologia. Camille Roux, Jonathan Romiguier and Christelle Fraïsse were of a great help for the analysis scenarios of speciation. We thank the diving servive of INSU/OSU Pytheas for fieldwork. We acknowledge the staff of the "Cluster de calcul intensif HPC" Platform of the OSU Institut Pythéas (Aix-Marseille Université, INSU-CNRS) for providing the computing facilities. We are grateful to the Genotoul bioinformatics platform Toulouse Occitanie (Bioinfo Genotoul, https://doi.org/10.15454/1.5572369328961167E12) for providing help, computing and storage resources. We thank Christophe Klopp and Marie-Stéphane Trotard for their help. We acknowledge the use of the computing cluster of MNHN (Plateforme de Calcul Intensif et Algorithmique PCIA, Muséum National d’Histoire Naturelle, Centre national de la recherche scientifique, UAR 2700 2AD, CP 26, 57 rue Cuvier, F-75231 Paris Cedex 05, France). Part of the bioinformatics analyses have been performed on the Core Cluster of the Institut Français de Bioinformatique (IFB) (ANR-11-INBS-0013). MGX acknowledges financial support from France Génomique National infrastructure, funded as part of “Investissement d’Avenir” program managed by Agence Nationale pour la Recherche (contract ANR-10-INBS-09). Part of this work has been performed during a CNRS detachment position of D. Aurelle at the ISYEB laboratory.

Data availability

The transcriptome raw sequences are available in Genbank under BioProject ID PRJNA1037721. The RAD raw sequences are available in Genbank under BioProject ID PRJNA1122331.

The scripts used in this study, the vcf files from RAD sequencing, and the results of the DILS analysis are available at https://doi.org/10.5281/zenodo.12532817 .

Conflict of interest disclosure

The authors declare that they have no conflict of interest in relation to the content of the article

592

References

Ålund M, Cenzer M, Bierne N, *et al*. Anthropogenic Change and the Process of Speciation. *Cold Spring Harbor Perspectives in Biology* 2023;**15**:a041455. https://cshperspectives.cshlp.org/content/15/12/a041455.short

Anderson E, Thompson EA. A model-based method for identifying species hybrids using multilocus genetic data. *Genetics* 2022;**160**:1217–1229. https://academic.oup.com/genetics/article-abstract/160/3/1217/6052497

Arnold ML, Bulger MR, Burke JM, *et al*. 1999. Natural hybridization: how low can you go and still be important? *Ecology* 1999;**80**:371–381. https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/0012-9658(1999)080[0371:NHHLCY]2.0.CO;2

Aurelle D, Pivotto ID, Malfant M, *et al*. Fuzzy species limits in Mediterranean gorgonians (Cnidaria, Octocorallia): inferences on speciation processes. *Zoologica Scripta* 2017;**46**:767–778. https://onlinelibrary.wiley.com/doi/abs/10.1111/zsc.12245

Baird NA, Etter PD, Atwood TS, *et al*. Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *PLoS ONE* 2008;**3**:e3376. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0003376

Baumgarten S, Bayer T, Aranda M, *et al*. Integrating microRNA and mRNA expression profiling in Symbiodinium microadriaticum, a dinoflagellate symbiont of reef-building corals. *BMC Genomics* 2013;**14**:704. https://link.springer.com/article/10.1186/1471-2164-14-704

Berkelmans R and van Oppen MJH. The role of zooxanthellae in the thermal tolerance of corals: a 'nugget of hope' for coral reefs in an era of climate change. *Proceedings of thex Royal Society B: Biological Sciences* 2006;**273**:2305–2312. https://royalsocietypublishing.org/doi/abs/10.1098/rspb.2006.3567

Bierne N, Borsa P, Daguin C, *et al*. Introgression patterns in the mosaic hybrid zone between Mytilus edulis and M. galloprovincialis. *Molecular Ecology* 2003;**12**:447–462. https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1365-294X.2003.01730.x

Bierne N, Welch J, Loire E, *et al*. The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Molecular Ecology* 2011;**20**:2044–2072. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-294X.2011.05080.x

Bonacolta AM, Miravall J, Gómez-Gras D, *et al*. Differential apicomplexan presence predicts thermal stress mortality in the Mediterranean coral Paramuricea clavata. *Environmental Microbiology* 2024;**26**:e16548. https://enviromicro-journals.onlinelibrary.wiley.com/doi/abs/10.1111/1462-2920.16548

Bordenstein S. Symbiosis And The Origin Of Species. In: *Insect Symbiosis*. CRC Press, 2003;283–303.

Brener-Raffalli K, Vidal-Dupiol J, Adjeroud M, *et al*. Gene expression plasticity and frontloading promote thermotolerance in *Pocillopora* corals. *Peer Community Journal* 2022;**2**. https://peercommunityjournal.org/articles/10.24072/pcjournal.79/

Brucker RM and Bordenstein SR. Speciation by symbiosis. *Trends in Ecology & Evolution* 2012;**27**:443. https://www.cell.com/trends/ecology-evolution/fulltext/S0169-5347(12)00076-6

Cabau C, Escudié F, Djari A, *et al*. Compacting and correcting Trinity and Oases RNA-Seq de novo assemblies. *PeerJ* 2017;**5**:e2988. https://peerj.com/articles/2988/

Cahill AE, Meglécz E, Chenuil A. Scientific history, biogeography, and biological traits predict presence of cryptic or overlooked species. *Biological Reviews* 2024;**99**:546-561. https://onlinelibrary.wiley.com/doi/abs/10.1111/brv.13034

Cairns SD. Deep-water corals: an overview with special reference to diversity and distribution of deep-water scleractinian corals. *Bulletin of marine Science* 2007;**81**:311–322. https://www.ingentaconnect.com/content/umrsmas/bullmar/2007/00000081/00000003/art00002

Calderón I, Garrabou J, Aurelle D. Evaluation of the utility of COI and ITS markers as tools for population genetic studies of temperate gorgonians. *Journal of Experimental Marine Biology and Ecology* 2006;**336**:184–197. https://www.sciencedirect.com/science/article/pii/S0022098106002498

Campoy AN, Addamo AM, Machordom A, *et al.* The Origin and Correlated Evolution of Symbiosis and Coloniality in Scleractinian Corals. *Frontiers in Marine Science* 2020;**7**. https://www.frontiersin.org/articles/10.3389/fmars.2020.00461/full

Carpine C, Grasshoff M. Les gorgonaires de la Méditerranée. *Bulletin de l'Institut Océanographique de Monaco* 1975;**71**(140).

Carugati L, Moccia D, Bramanti L, *et al.* Deep-Dwelling Populations of Mediterranean Corallium rubrum and Eunicella cavolini: Distribution, Demography, and Co-Occurrence. *Biology* 2022;**11**. https://www.mdpi.com/2079-7737/11/2/333

Casado-Amezúa P, Terrón-Sigler A, Pinzón JH, *et al.* General ecological aspects of Anthozoan-Symbiodinium interactions in the Mediterranean Sea. In: Goffredo S, Dubinsky Z, (ed). The cnidaria, past, present and future: the world of medusa and her sisters. Springer, 2016;375–386. https://link.springer.com/chapter/10.1007/978-3-319-31305-4_24

Chan WY, Peplow LM, Menéndez P, Hoffmann AA & Van Oppen MJ. 2018. Interspecific hybridization may provide novel opportunities for coral reef restoration. *Frontiers in Marine Science* 5: 160. https://www.frontiersin.org/articles/10.3389/fmars.2018.00160/full

Chimienti G. 2020. Vulnerable forests of the pink sea fan Eunicella verrucosa in the Mediterranean Sea. *Diversity* 12: 176.

Coelho M, Pearson G, Boavida J, *et al.* Not out of the Mediterranean: Atlantic populations of the gorgonian Paramuricea clavata are a separate sister species under further lineage diversification. *Ecology and Evolution* 2023;**13**. https://onlinelibrary.wiley.com/doi/abs/10.1002/ece3.9740

Coma R, Llorente-Llurba E, Serrano E, *et al.* Natural heterotrophic feeding by a temperate octocoral with symbiotic zooxanthellae: a contribution to understanding the mechanisms of die-off events. *Coral Reefs* 2015;**34**:549–560. https://link.springer.com/article/10.1007/s00338-015-1281-3

Combosch DJ, Vollmer SV. Trans-Pacific RAD-Seq population genomics confirms introgressive hybridization in Eastern Pacific Pocillopora corals. *Molecular phylogenetics and evolution* 2015;**88**:154–162. https://www.sciencedirect.com/science/article/pii/S1055790315000858

Csilléry K, François O, Blum MGB. abc: an R package for approximate Bayesian computation (ABC). *Methods in Ecology and Evolution* 2012;**3**:475–479. https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/j.2041-210X.2011.00179.x

Danecek P, Auton A, Abecasis G, *et al.* The variant call format and VCFtools. *Bioinformatics* 2011;**27**:2156–2158. https://academic.oup.com/bioinformatics/article/27/15/2156/402296

De Jode A, Le Moan A, Johannesson K,*et al.* Ten years of demographic modelling of divergence and speciation in the sea. *Evolutionary Applications* 2023;**16**:542–559. https://onlinelibrary.wiley.com/doi/abs/10.1111/eva.13428

De Queiroz K. Species Concepts and Species Delimitation. *Systematic Biology* 2007;**56**:879–886. https://academic.oup.com/sysbio/article-abstract/56/6/879/1653163
De Raad D. SNPfiltR: Interactively Filter SNP Datasets. R package version 1.0.1. https://devonderaad.github.io/SNPfiltR/, 2023
Eaton DA, Overcast I. ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics* 2020;**36**:2592–2594. https://academic.oup.com/bioinformatics/article-abstract/36/8/2592/5697088
Erickson KL, Pentico A, Quattrini AM *et al.* New approaches to species delimitation and population structure of anthozoans: Two case studies of octocorals using ultraconserved elements and exons. *Molecular Ecology Resources* 2021;**21**:78–92. https://onlinelibrary.wiley.com/doi/abs/10.1111/1755-0998.13241
Estaque T, Richaume J, Bianchimani O, *et al.* Marine heatwaves on the rise: One of the strongest ever observed mass mortality event in temperate gorgonians. *Global change biology* 2023;**29**:6159-6162. https://onlinelibrary.wiley.com/doi/10.1111/gcb.16931
Ezzat L, Merle PL, Furla P, *et al.* The Response of the Mediterranean Gorgonian Eunicella singularis to Thermal Stress Is Independent of Its Nutritional Regime. *PLoS ONE* 2013;**8**:e64370. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0064370
Faria R, Johannesson K, Stankowski S. Speciation in marine environments: Diving under the surface. *Journal of Evolutionary Biology* 2021;**34**:4–15. https://academic.oup.com/jeb/article-abstract/34/1/4/7326591
Forcioli D, Merle PL, Caligara C, *et al.* Symbiont diversity is not involved in depth acclimation in the Mediterranean sea whip Eunicella singularis. *Marine Ecology Progress Series* 2011;**439**:57–71. https://www.int-res.com/abstracts/meps/v439/p57-71/
Fourt M, Goujard A. Rapport final de la campagne MEDSEACAN (Têtes des canyons méditerranéens continentaux) novembre 2008–avril 2010. *Partenariat Agence des aires marines protégées–GIS Posidonie*: 1–218. 2012. http://paleopolis.rediris.es/benthos/TaP/Rapport_Final_MEDSEACAN.pdf
Fraïsse C, Popovic I, Mazoyer C, *et al.* DILS: Demographic inferences with linked selection by using ABC. *Molecular Ecology Resources* 2021;**21**:2629–2644. https://onlinelibrary.wiley.com/doi/abs/10.1111/1755-0998.13323
Frichot E, Mathieu F, Trouillon T, *et al.* Fast and efficient estimation of individual ancestry coefficients. *Genetics* 2014;**196**:973–983. https://academic.oup.com/genetics/article/196/4/973/5935614
Frichot E, François O. LEA: an R package for landscape and ecological association studies. *Methods in Ecology and Evolution* 2015;**6**:925–929. https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.12382
Furla P, Allemand D, Shick JM, *et al.* The symbiotic anthozoan: a physiological chimera between alga and animal. *Integrative and Comparative Biology* 2005;**45**:595–604. https://academic.oup.com/icb/article-abstract/45/4/595/636401
Gagnaire P, Broquet T, Aurelle D, *et al.* Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. *Evolutionary Applications* 2015;**8**:769–786. https://onlinelibrary.wiley.com/doi/abs/10.1111/eva.12288
Gayral P, Melo-Ferreira J, Glemin S, *et al.* Reference-free population genomics from next-generation transcriptome data and the vertebrate–invertebrate gap. *PLoS Genetics* 2013;**9**:e1003457. https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1003457

Gori A, Bramanti L, López-González P, *et al.* Characterization of the zooxanthellate and azooxanthellate morphotypes of the Mediterranean gorgonian Eunicella singularis. *Marine biology* 2012;**159**:1485–1496. https://link.springer.com/article/10.1007/s00227-012-1928-3
Grasshoff,M. Die Flachwasser-Gorgonarien von Europa und Westafrika (Cnidaria, Anthozoa). *Courier Forschunginstitut Senckenberg 1992:***149**. Frankfurt a. M.
Guzman C, Shinzato C, Lu TM *et al.* Transcriptome analysis of the reef-building octocoral, Heliopora coerulea. *Scientific Reports* 2018;**8**:8397. https://www.nature.com/articles/s41598-018-26718-5
Haguenauer A, Zuberer F, Ledoux JB *et al.* Adaptive abilities of the Mediterranean red coral Corallium rubrum in a heterogeneous and changing environment: from population to functional genetics. *Journal of Experimental Marine Biology and Ecology* 2013;**449**:349–357.
Jombart T. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 2008;**24**:1403-1405. https://www.sciencedirect.com/science/article/pii/S0022098113003493
Kent WJ. BLAT—the BLAST-like alignment tool. *Genome research* 2002;**12**:656–664. https://genome.cshlp.org/content/12/4/656.short
Krueger-Hadfield S. marmap. https://www.molecularecologist.com/2015/07/03/marmap/. 2015
LaJeunesse TC, Parkinson JE, Gabrielson PW, *et al.* Systematic revision of Symbiodiniaceae highlights the antiquity and diversity of coral endosymbionts. *Current Biology* 2018;**28**:2570–2580. https://www.cell.com/current-biology/fulltext/S0960-9822(18)30907-2
LaJeunesse TC, Wiedenmann J, Casado-Amezúa P, *et al.* Revival of Philozoon Geddes for host-specialized dinoflagellates,'zooxanthellae', in animals from coastal temperate zones of northern and southern hemispheres. *European Journal of Phycology* 2022;**57**:166–180. https://www.tandfonline.com/doi/abs/10.1080/09670262.2021.1914863
Leroy T, Louvet JM, Lalanne C, *et al*. Adaptive introgression as a driver of local adaptation to climate in European white oaks. *New Phytologist* 2020;**226**:1171–1182. https://nph.onlinelibrary.wiley.com/doi/abs/10.1111/nph.16095
Lesser MP, Stat M, Gates RD. The endosymbiotic dinoflagellates (Symbiodinium sp.) of corals are parasites and mutualists. *Coral Reefs* 2013;**32**:603–611. https://link.springer.com/article/10.1007/s00338-013-1051-z
Li H, Handsaker B, Wysoker A, *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* 2009;**25**:2078–2079. https://academic.oup.com/bioinformatics/article-abstract/25/16/2078/204688
Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 2009;**25**:1754–1760. https://academic.oup.com/bioinformatics/article/25/14/1754/225615
Lischer HE, Excoffier L. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* 2012;**28**:298–299. https://academic.oup.com/bioinformatics/article/28/2/298/198891
Lu J, Rincon N, Wood DE, *et al.* Metagenome analysis using the Kraken software suite. *Nature protocols* 2022;**17**:2815-2839. https://www.nature.com/articles/s41596-022-00738-y
Macleod KL, Jenkins TL, Witt MJ *et al.* Rare, long-distance dispersal underpins genetic connectivity in the pink sea fan, Eunicella verrucosa. *Evolutionary Applications* 2024;**17**:e13649. https://onlinelibrary.wiley.com/doi/abs/10.1111/eva.13649
Mayr E. Wu's genic view of speciation. *Journal of Evolutionary Biology* 2001;**14**:866–867. https://academic.oup.com/jeb/article-abstract/14/6/866/7322934

McFadden CS, Benayahu Y, Pante E, *et al.* Limitations of mitochondrial gene barcoding in Octocorallia. *Molecular Ecology Resources* 2011;**11**:19–31. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1755-0998.2010.02875.x

McFadden CS, Quattrini AM, Brugler MR, *et al.* Phylogenomics, origin, and diversification of Anthozoans (Phylum Cnidaria). *Systematic Biology* 2021;**70**:635-647. https://academic.oup.com/sysbio/article-abstract/70/4/635/6122449

Muir PR, Obura DO, Hoeksema BW, *et al.* Conclusions of low extinction risk for most species of reef-building corals are premature. *Nature Ecology & Evolution* 2022;**6**:357–358. https://www.nature.com/articles/s41559-022-01659-5

Munro L. Determining the reproductive cycle of Eunicella verrucosa. *Reef Research: ETR* 2004;**11**. https://www.marine-bio-images.com/RR_Eunicella_PDFS/Report_RR12Jul2004reproductive%20cycle%20pdf.pdf

Muthye V, Mackereth CD, Stewart JB *et al.* Large dataset of octocoral mitochondrial genomes provides new insights into mt-mutS evolution and function. *DNA repair* 2022;**110**:103273. https://www.sciencedirect.com/science/article/pii/S1568786422000027

van Oppen MJH, Medina M. Coral evolutionary responses to microbial symbioses. *Philosophical Transactions of the Royal Society B: Biological Sciences* 2020;**375**:20190591. https://royalsocietypublishing.org/doi/abs/10.1098/rstb.2019.0591

Palumbi SR. Marine speciation on a small planet. *Trends in Ecology & Evolution* 1992;**7**:114–118. https://www.sciencedirect.com/science/article/pii/016953479290144Z

Pante E, Puillandre N, Viricel A, *et al.* Species are hypotheses: avoid connectivity assessments based on pillars of sand. *Molecular Ecology* 2015a;**24**:525–544. https://onlinelibrary.wiley.com/doi/abs/10.1111/mec.13048

Pante E, Abdelkrim J, Viricel A, *et al.* Use of RAD sequencing for delimiting species. *Heredity* 2015b;**114**:450–459. https://www.nature.com/articles/hdy2014105

Pante E, Simon-Bouhet B. marmap: a package for importing, plotting and analyzing bathymetric and topographic data in R. *PLoS one* 2013;**8**:e73051. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0073051

Patro R, Duggal G, Love MI, *et al.* Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods* 2017;**14**:417–419. https://www.nature.com/articles/nmeth.4197

Peñalba JV, Runemark A, Meier JI, Singh P, Wogan GO, Sánchez-Guillén R, Mallet J, Rometsch SJ, Menon M, Seehausen O. The Role of Hybridization in Species Formation and Persistence. *Cold Spring Harbor perspectives in biology* 2014;**a041445**. https://cshperspectives.cshlp.org/content/early/2024/03/01/cshperspect.a041445.short

Picard Toolkit. Broad Institute. GitHub Repository. https://broadinstitute.github.io/picard/. 2019

Porro B. Diversités génétiques chez l'holobiote Anemonia viridis: des morphotypes de l'hôte à la différenciation symbiotique. (Doctoral dissertation, COMUE Université Côte d'Azur (2015-2019). 2019. https://theses.hal.science/tel-02736573

Pudlo P, Marin JM, Estoup A, *et al.* Reliable ABC model choice via random forests. *Bioinformatics* 2016;**32**:859–866. https://academic.oup.com/bioinformatics/article-abstract/32/6/859/1744513

Quilodrán CS, Ruegg K, Sendell-Price AT, *et al.* The multiple population genetic and demographic routes to islands of genomic divergence. *Methods in Ecology and Evolution* 2020;**11**:6–21. https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.13324

Readman J, Hiscock K. Eunicella verrucosa. Pink sea fan. https://www.marlin.ac.uk/species/detail/1121. 2017

Reynes L, Aurelle D, Chevalier C, *et al.* Population genomics and Lagrangian modeling shed light on dispersal events in the Mediterranean endemic Ericaria zosteroides (= Cystoseira zosteroides) (Fucales). *Frontiers in Marine Science* 2021;**8**:683528. https://www.frontiersin.org/articles/10.3389/fmars.2021.683528/full

Rivera-García L, Rivera-Vicéns RE, Veglia AJ *et al.* De novo transcriptome assembly of the digitate morphotype of Briareum asbestinum (Octocorallia: Alcyonacea) from the southwest shelf of Puerto Rico. *Marine Genomics* 2019;**47**:100676. https://www.sciencedirect.com/science/article/pii/S1874778718302393

Rosenberg E, Zilber-Rosenberg I. The hologenome concept of evolution after 10 years. *Microbiome* 2018;**6**:78. https://link.springer.com/article/10.1186/S40168-018-0457-9

Rousset F. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources* 2008;**8**:103–106. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1471-8286.2007.01931.x

Rousset F, Lopez J, Belkhir K. Package 'genepop'. *R package version* 1. 2020. https://cran.r-project.org/web/packages/genepop/index.html

Roux C, Fraïsse C, Romiguier J, *et al.* Shedding Light on the Grey Zone of Speciation along a Continuum of Genomic Divergence. *PLOS Biology* 2016;**14**:e2000234. https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.2000234

Sachs JL, Wilcox TP. A shift to parasitism in the jellyfish symbiont Symbiodinium microadriaticum. *Proceedings of the Royal Society B: Biological Sciences* 2006;**273**:425–429. https://royalsocietypublishing.org/doi/abs/10.1098/rspb.2005.3346

Sartoretto S, Francour P. Bathymetric distribution and growth rates of Eunicella verrucosa (Cnidaria: Gorgoniidae) populations along the Marseilles coast (France). *Scientia Marina* 2011;**76**:349–355. https://archimer.ifremer.fr/doc/00087/19859/

Schulz MH, Zerbino DR, Vingron M *et al.* Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics* 2012;**28**:1086–1092. https://academic.oup.com/bioinformatics/article-abstract/28/8/1086/195757

Sini M, Kipson S, Linares C, *et al.* The Yellow Gorgonian Eunicella cavolini: Demography and Disturbance Levels across the Mediterranean Sea. *PLoS ONE* 2015;**10**:e0126253. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0126253

Stanley RRE, Jeffery NW, Wringe BF, *et al.* GENEPOPEDIT: a simple and flexible tool for manipulating multilocus molecular data in R. *Molecular Ecology Resources* 2017;**17**:12–18. https://onlinelibrary.wiley.com/doi/abs/10.1111/1755-0998.12569

Tricou T, Tannier E, de Vienne DM. Ghost lineages can invalidate or even reverse findings regarding gene flow. *PLoS Biology* 2022;**20**:e3001776. https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3001776

Tsagkogeorga G, Cahais V, Galtier N. The population genomics of a fast evolver: high levels of diversity, functional constraint, and molecular adaptation in the tunicate Ciona intestinalis. *Genome biology and evolution* 2012;**4**:852–861. https://academic.oup.com/gbe/article-abstract/4/8/852/580636

Van Oppen M, Mieog JC, Sanchez C, *et al.* . Diversity of algal endosymbionts (zooxanthellae) in octocorals: the roles of geography and host relationships. *Molecular Ecology* 2005;**14**:2403–2417. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-294X.2005.02545.x

Wagner D, Pochon X, Irwin L, *et al.* Azooxanthellate? Most Hawaiian black corals contain Symbiodinium. *Proceedings of the Royal Society B: Biological Sciences* 2011;**278**:1323–1328. https://royalsocietypublishing.org/doi/abs/10.1098/rspb.2010.1681

Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. *Evolution* 1984;**38**:1358–1370. https://www.jstor.org/stable/2408641

Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Biology* 2019;**20**:257. https://link.springer.com/article/10.1186/s13059-019-1891-0

Figure 1: map of sampling sites for transcriptomes: A) general view, B) zoom on the area of Marseille. The symbols correspond to different samples: EC *E. cavolini*, ES *E. singularis*, EV *E. verrucosa*, HY potential hybrids. The three letters correspond to the codes of the sampling. The maps have been produced with the marmap R package (Pante & Simon-Bouhet, 2013) and following the tutorial of Krueger-Hadfield (2015).

Figure 2: distribution of the frequency of Symbiodiniaceae sequences in the individual transcriptomes according to the species based A) on the number of reads estimated with Salmon, and B) on the proportion of assembled sequences (contigs) with the BLAT analyses.

A) Read counts with Salmon; mean values per group: *E. cavolini*: 16508; hybrids: 10238; *E. singularis*: 26023; *E. verrucosa*: 4285. Kruskal-Wallis test of the differences among groups: chi-squared = 7.9467, df = 3, p-value = 0.047.

B) Assembled sequences with BLAT; mean values per group: *E. cavolini*: 0.0034; hybrids: 0.0029; *E. singularis*: 0.0219; *E. verrucosa*: 0.0028. Kruskal-Wallis test of the differences among groups: chi-squared = 14.352, df = 3, p-value = 0.002.

Figure 3: barplots of coancestry coefficients inferred with the LEA R package. The analysis is based on the "polymorphic sites" transcriptome dataset. The red asterisks indicate the individuals used as prior for parental status in the NewHybrids analysis. The results of the NewHybrids analysis are indicated below the hybrid individuals: F1, 1st generation; F2, 2nd generation; Sbx, backcross with *E. singularis*. The coancestry analysis is based on 31 369 SNPs, whereas the NewHybrids analysis is based on 326 SNPs showing high differentiation between *E. cavolini* and *E. singularis*.

Figure 4: principal Component Analysis based on the "polymorphic SNPs" transcriptome dataset; axis 1 represents 33.2% of the variance, axis 2 represents 13% of the variance

Table 1 : inference of hybrid status with NewHybrids for transcriptome and RAD sequencing. For transcriptomes, all probabilities were at 1 for the inferred status and for the ten replicates. For RAD sequencing, the results are given for the four datasets (i.e. different assembly strategies). If no probability is mentioned for RAD sequencing, the hybrid status was supported by a probability higher than 0.999 over the ten replicates. In the other cases, the numbers indicate the minimal probability threshold over the ten replicates for this status (and the status was coherent over the ten replicates as well, with slight variations in probability). NA indicates an individual which was removed during the filtering of SNPs because of too many missing data. The lines highlighted in grey indicate the cases where different status were inferred depending on the dataset. Bx-ES and Bx-EC indicate backcrosses with *E. singularis* and *E. cavolini* respectively; ES indicates parental *E. singularis*.

Individual - transcriptome				
EH-JPB-a	F1			
EH-MFN-a	Bx-ES			
EH-MFN-b	F2			
EH-MFN-e	Bx-ES			
Individual - RAD sequencing	***de novo***	**ref. *E. cavolini***	**ref. *E. singularis***	**ref. *E. verrucosa***
EC-X-MFNB	F2	F2	F2	F2
EC-X-MFNC	F2	NA	Bx-EC	NA
EC-X-MFND	Bx-ES	Bx-ES	Bx-ES	Bx-ES
EC-X-MFNE	Bx-EC	Bx-EC	Bx-EC	Bx-EC
EC-X-MFNF	Bx-EC	F2 > 0.95	Bx-EC > 0.92	F2
EC-X-MFNG	F2	F2	F2	F2
EC-X-MFNH	Bx-ES	Bx-ES	Bx-ES > 0.67	Bx-ES > 0.98
EC-X-MFNI	F1	F1	F1	F2
EC-X-MFNL	F1	F1 > 0.99	F1	F2
ES-X-MFNA	Bx-ES	Bx-ES	Bx-ES	Bx-ES
ES-X-MFNJ	F2	F2 > 0.96	F2	F2
ES-X-MFNK	ES	ES	ES	ES

Table 2: results of demographic inferences with DILS with transcriptome data. The columns indicate the species comparison, the model choice for population size (constant vs. variable), and the results of inferences: current gene flow (migration vs isolation); if current migration, isolation / migration (IM) vs ancestral migration (AM); if no current migration, strict isolation (SI) vs ancestral migration (AM); homogeneity (N-homo) vs heterogeneity in effective size (N-hetero) among loci; homogeneity (M-homo) vs heterogeneity (M-hetero) in gene flow among loci. The probability of each scenario is given in the same case. Homogeneity and heterogeneity indicate no variation or variation among loci respectively.

Comparison	Population size	Current gene flow	IM / SC	SI / AM	Heterogeneity effective size	Heterogeneity gene flow
cavolini / singularis	constant	Migration; 0.87	SC; 0.79	-	N-hetero; 0.99	M-homo; 0.82
cavolini / singularis	variable	Migration; 0.88	SC; 0.77	-	N-hetero; 1	M-homo; 0.87
cavolini / verrucosa	constant	Isolation; 0.90	-	AM; 0.65	N-hetero; 1	-
cavolini / verrucosa	variable	Isolation; 0.89	-	AM; 0.69	N-hetero; 1	-
singularis / verrucosa	constant	Isolation; 0.87	-	AM; 0.61	N-hetero; 1	-
singularis / verrucosa	variable	Isolation; 0.87	-	AM; 0.61	N-hetero; 1	-

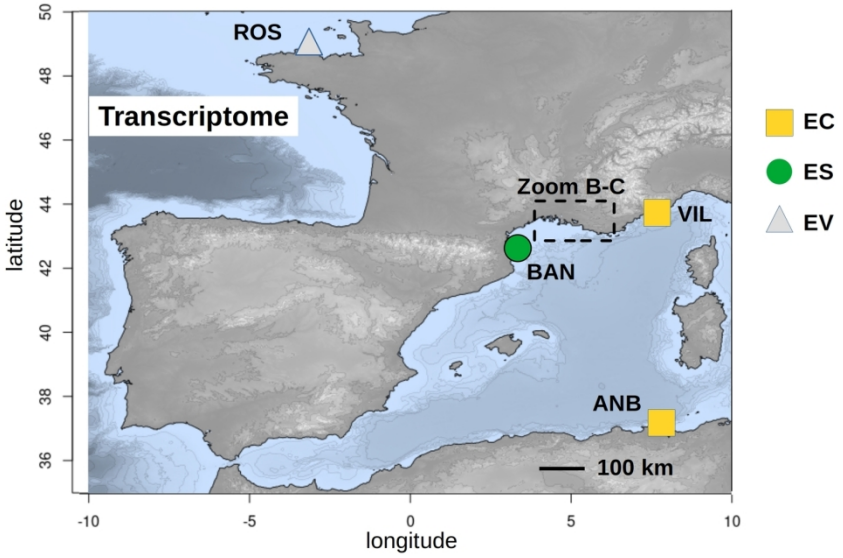


Figure 1A, sampling sites with transcriptomes

80x45mm (600 x 600 DPI)

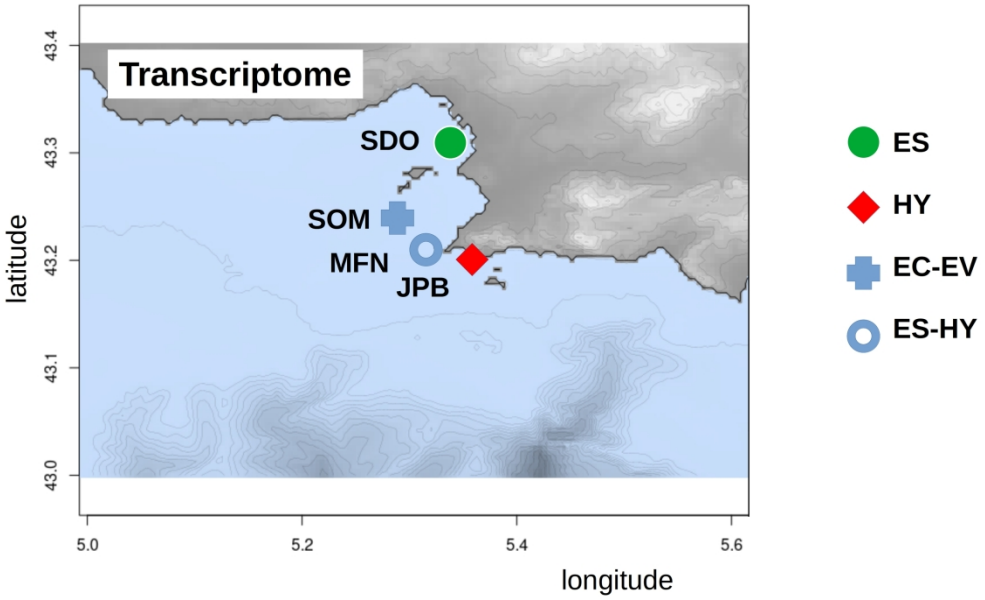


Figure 1B, sampling sites with transcriptomes

122x74mm (600 x 600 DPI)

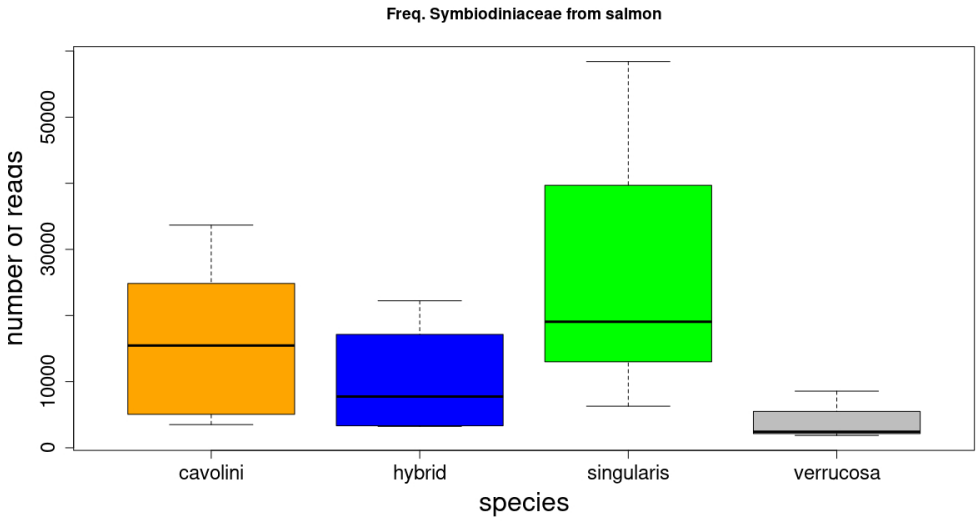


Figure 2A, read counts with Salmon

317x174mm (96 x 96 DPI)

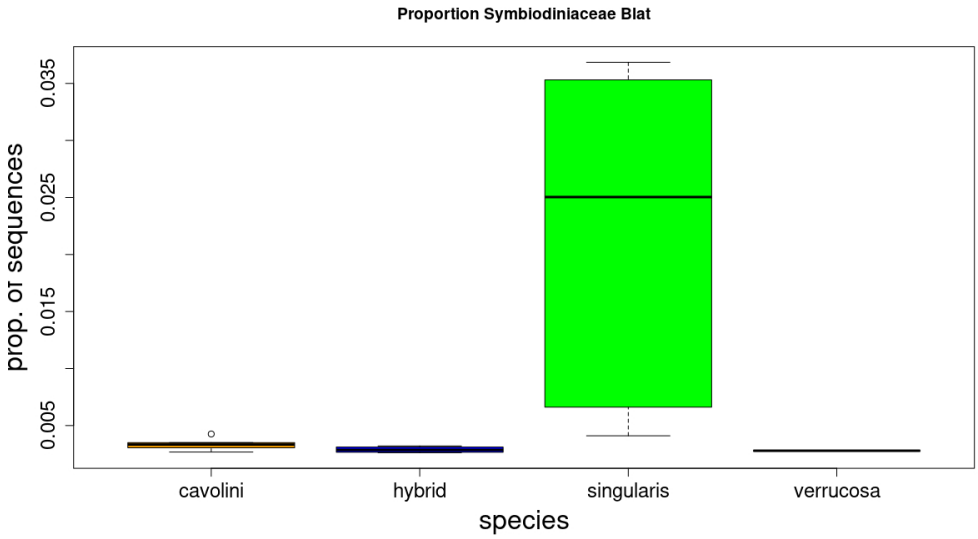


Figure 2B, assembled sequences with BLAT

317x180mm (96 x 96 DPI)

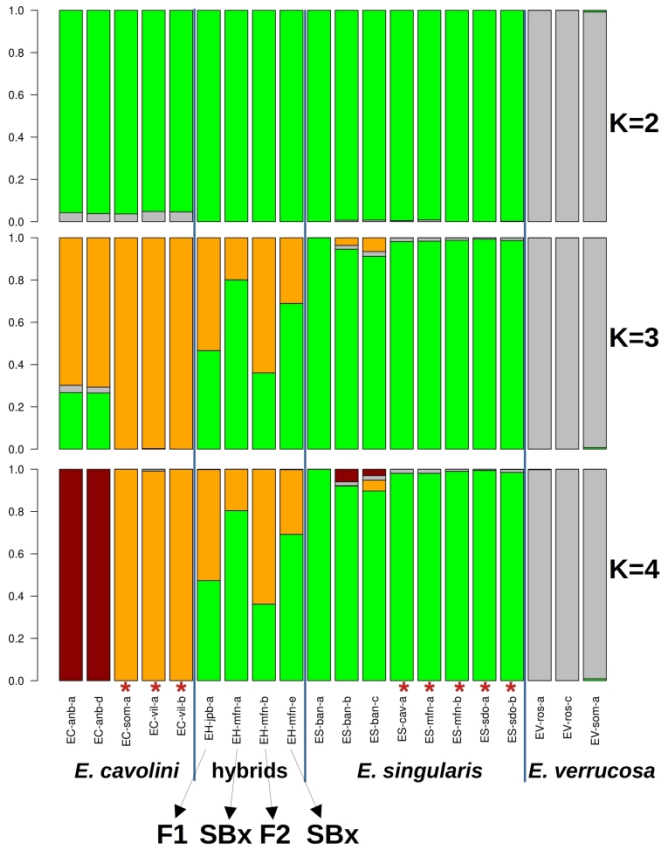


Figure 3, barplots of coancestry coefficients inferred with the LEA R package with transcriptome data

99x177mm (600 x 600 DPI)

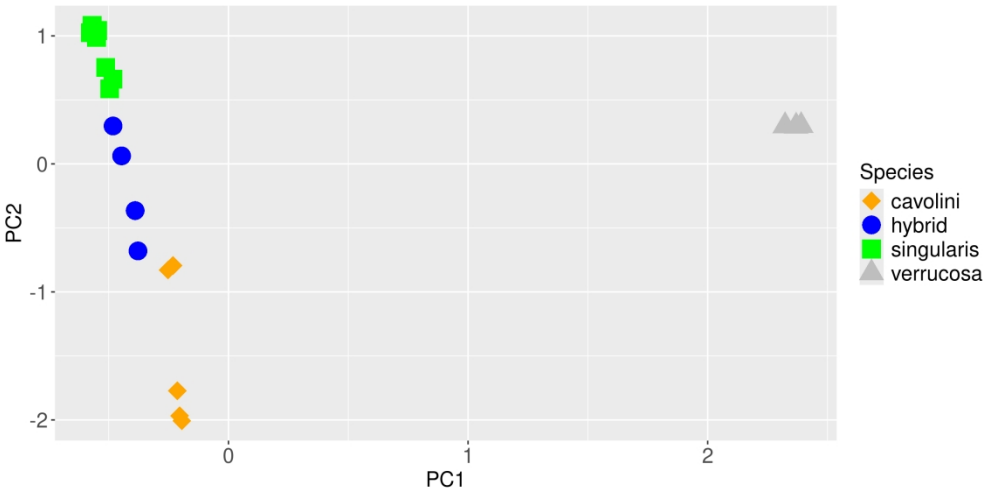


Figure 4, PCA from transcriptome data

254x127mm (300 x 300 DPI)

Symbiotic status does not preclude hybridisation in Mediterranean octocorals

Supplementary Material S1

Didier Aurelle[1,2*], Anne Haguenauer[3], Marc Bally[1], Frédéric Zuberer[4], Jean-Baptiste Ledoux[5], Stéphane Sartoretto[6], Lamya Chaoui[7], Hichem Kara[7], Sarah Samadi[2], Pierre Pontarotti[8,9,10]

[1] Aix Marseille Univ, Université de Toulon, CNRS, IRD, MIO, Marseille, France

[2] Institut Systématique Evolution Biodiversité (ISYEB), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, CP 26, 75005 Paris, France.

[3] CNRS - Délégation Provence et corse, Marseille, France

[4] Aix Marseille Univ, CNRS, IRD, INRAE, OSU Inst. PYTHEAS, Marseille, France

[5] CIIMAR/CIMAR, Centro Interdisciplinar de Investigação Marinha e Ambiental, Universidade do Porto, Porto, Portugal.

[6] Ifremer, LITTORAL, 83500 La Seyne-sur-Mer, France

[7] Laboratoire Bioressources marines. Université d'Annaba Badji Mokhtar, Annaba - Algérie.

[8] Aix Marseille Univ, MEPHI, Marseille, France.

[9] IHU Méditerranée Infection, Marseille, France.

[10] CNRS SNC5039

*Corresponding author

Correspondence: didier.aurelle@univ-amu.fr

Table S1 : list of samples (all sampled in 2016) and statistics of assembled transcriptomes for individual transcriptomes and meta-transcriptomes. The intermediate samples correspond to individuals with intermediate morphology, suspected to be hybrids before genetic analyses. For Annaba and Villefranche sur Mer, we indicate a range depth, as the precise sampling depth had not been recorded. The assembly is based on paired-ends sequencing (2 x 75 bp) and the number of raw sequences corresponds to the number of pairs. Contigs indicates the number of contigs for each assembly, with the corresponding N50 and L50. The Lg columns corresponds to the contigs length in bp, with the sum, minimum, mean, median and maximum of Lg. The last two lanes refer to the meta-transcriptome obtained from all individual transcriptomes with or without potential Symbiodiniaceae sequences. See main text for details.

morphologial identification	**sample**	**site**	**depth**	**raw sequences**	**contigs**	**N50**	**L50**	**Lg sum**	**Lg min**	**Lg mean**	**Lg median**	**Lg max**
E. cavolini	e-cavol-anb-a	Annaba, Algeria	20-30	21432997	33627	1978	7240	46041698	201	1369.19	1023	13533
E. cavolini	e-cavol-anb-d	Annaba, Algeria	20-30	20070761	33624	2002	7025	46288676	201	1376.66	1004	24422
E. cavolini	e-cavol-som-a	Marseille, France	58	22986734	43541	1757	8987	52113269	201	1196.88	840	24228
E. cavolini	e-cavol-vil-a	Villefranche sur Mer, France	20-40	31846763	36908	2056	7687	51709228	201	1401.03	1018	18381
E. cavolini	e-cavol-vil-b	Villefranche sur Mer, France	20-40	28751407	34961	2044	7470	48971519	201	1400.75	1040	19504
intermediate	e-hybri-jpb-a	Marseille, France	25	34392918	39407	2031	8098	54290945	201	1377.7	998	25557
intermediate	e-hybri-mfn-a	Marseille, France	10	44256795	40762	2081	8451	57794280	201	1417.85	1039	25573
intermediate	e-hybri-mfn-b	Marseille, France	10	34705411	39672	2046	8040	54738734	201	1379.78	981.5	16650
intermediate	e-hybri-mfn-e	Marseille, France	10	36536647	39532	2038	8090	54685655	201	1383.33	995.5	25578
E. singularis	e-singu-ban-a	Banyuls, France	10	44325669	45364	1919	9379	58576839	201	1291.26	928	28882
E. singularis	e-singu-ban-b	Banyuls, France	10	33184944	38095	1930	8114	50868966	201	1335.32	987	20211
E. singularis	e-singu-ban-c	Banyuls, France	10	46271612	43821	2023	9132	60512898	201	1380.91	1007	21714
E. singularis	e-singu-cav-a	Marseille, France	25	48947180	51120	1967	10031	65261049	201	1276.62	868	22527
E. singularis	e-singu-mfn-a	Marseille, France	10	52588076	70114	1761	13336	79649263	201	1136	739	16808
E. singularis	e-singu-mfn-b	Marseille, France	10	43713977	55035	1894	10583	67120524	201	1219.6	808	21143
E. singularis	e-singu-sdo-a	Marseille, France	30	37444166	55928	1741	10464	62326140	201	1114.4	715	16387
E. singularis	e-singu-sdo-b	Marseille, France	30	39266148	72419	1652	13837	78950323	201	1090.19	715	24245
E. verrucosa	e-verru-ros-a	Roscoff, France	20	19398629	31195	1936	6630	41727111	201	1337.62	981	16974
E. verrucosa	e-verru-ros-c	Roscoff, France	20	20495748	31526	1968	6729	42779660	201	1356.96	1005	16663
E. verrucosa	e-verru-som-a	Marseille, France	58	23332185	33133	2005	6944	45531674	201	1374.21	1005	25577
Meta transcriptome				**number of contigs**	**retained contigs**	**N50**	**L50**	**Lg sum**	**Lg min**	**Lg mean**	**Lg median**	**Lg max**
meta				891354	68386	2144	14309	102621319	201	1500,62	1098	28882
meta no Symb				300085	59697	1975	12316	80903965	201	1355,24	967	25577

Table S2 : samples used for RAD sequencing, with putative species on the basis of field identications. All sampling sites are in the area of Marseille (see Figure 1).

Sample	morphologial identification	sampling year	site	depth (m)	raw reads
ECESC37	*E. cavolini*	2020	Escu	10	4588103
ECESC38	*E. cavolini*	2020	Escu	10	7379895
ECESC39	*E. cavolini*	2020	Escu	10	7749308
ECMEL37	*E. cavolini*	2020	Mélette	20	10448058
ECPLD22	*E. cavolini*	2020	Veyron	40	9621675
ECPLD23	*E. cavolini*	2020	Veyron	40	9266239
ECPLD24	*E. cavolini*	2020	Veyron	40	4517126
ECPLD25	*E. cavolini*	2020	Veyron	40	4363166
ECPLD26	*E. cavolini*	2020	Veyron	40	5409613
ECPLD38	*E. cavolini*	2020	Veyron	40	8258381
ECPLD39	*E. cavolini*	2020	Veyron	40	5241648
ECRID13	*E. cavolini*	2020	Riou	40	5146046
ECRID14	*E. cavolini*	2020	Riou	40	4743650
ECRID15	*E. cavolini*	2020	Riou	40	4639611
ECRID16	*E. cavolini*	2020	Riou	40	11001580
ECRID17	*E. cavolini*	2020	Riou	40	10657450
ECRID19	*E. cavolini*	2020	Riou	40	6801196
ECRID20	*E. cavolini*	2020	Riou	40	4546053
ECRID21	*E. cavolini*	2020	Riou	40	6800451
ECRID22	*E. cavolini*	2020	Riou	40	5310390
ECRID23	*E. cavolini*	2020	Riou	40	4702724
ECRID25	*E. cavolini*	2020	Riou	40	5472492
ECRID27	*E. cavolini*	2020	Riou	40	8302669
ECRID28	*E. cavolini*	2020	Riou	40	9097420
ECRID29	*E. cavolini*	2020	Riou	40	8163462
ESFRO10	*E. singularis*	2022	Fromages	10-25	9983228
ESFRO11	*E. singularis*	2022	Fromages	10-25	9752609
ESFRO12	*E. singularis*	2022	Fromages	10-25	13963735
ESFRO13	*E. singularis*	2022	Fromages	10-25	7473750
ESFRO14	*E. singularis*	2022	Fromages	10-25	11379910
ESFRO15	*E. singularis*	2022	Fromages	10-25	7291156
ESFRO2	*E. singularis*	2022	Fromages	10-25	5024054
ESFRO3	*E. singularis*	2022	Fromages	10-25	6235592
ESFRO5	*E. singularis*	2022	Fromages	10-25	3539506
ESFRO6	*E. singularis*	2022	Fromages	10-25	6217585
ESFRO7	*E. singularis*	2022	Fromages	10-25	12405704
ESFRO8	*E. singularis*	2022	Fromages	10-25	14113042

ESFRO9	*E. singularis*	2022	Fromages	10-25	13734698
ESSFI1	*E. singularis*	2013	Figuier	15	3885117
ESSFI10	*E. singularis*	2013	Figuier	15	4886425
ESSFI2	*E. singularis*	2013	Figuier	15	9360315
ESSFI3	*E. singularis*	2013	Figuier	15	12083132
ESSFI4	*E. singularis*	2013	Figuier	15	5542586
ESSFI5	*E. singularis*	2013	Figuier	15	5957124
ESSFI6	*E. singularis*	2013	Figuier	15	5030751
ESSFI7	*E. singularis*	2013	Figuier	15	7657655
ESSFI8	*E. singularis*	2013	Figuier	15	4283226
ESSFI9	*E. singularis*	2013	Figuier	15	5367231
EVCSO2	*E. verrucosa*	2022	Fromages	10-25	3794532
EVCSO4	*E. verrucosa*	2022	Fromages	10-25	6731471
EVFRO1	*E. verrucosa*	2022	Fromages	10-25	7811130
EVFRO2	*E. verrucosa*	2022	Fromages	10-25	4631737
EVFRO3	*E. verrucosa*	2022	Fromages	10-25	4507318
EVFRO4	*E. verrucosa*	2022	Fromages	10-25	3661582
EVFRO5	*E. verrucosa*	2022	Fromages	10-25	5269993
EC-X-MFNB	intermediate	2013	Maïre	10	5903675
EC-X-MFNC	intermediate	2013	Maïre	10	3478362
EC-X-MFND	intermediate	2013	Maïre	10	5023674
EC-X-MFNE	intermediate	2013	Maïre	10	5102971
EC-X-MFNF	intermediate	2013	Maïre	10	3427364
EC-X-MFNG	intermediate	2013	Maïre	10	4594893
EC-X-MFNH	intermediate	2013	Maïre	10	9326619
EC-X-MFNI	intermediate	2013	Maïre	10	6228674
EC-X-MFNL	intermediate	2013	Maïre	10	6013320
ES-X-MFNA	intermediate	2013	Maïre	10	5686156
ES-X-MFNJ	intermediate	2013	Maïre	10	5972257
ES-X-MFNK	intermediate	2013	Maïre	10	5383938

Table S3: list of mitochondrial MutS sequences used for the phylogenetic reconstruction with the corresponding Genbank accession numbers. The location and voucher code are indicated when available.

Accession number	Genus	species	location	voucher
KP036906	*Complexum*	*monodi*	*Congo*	*CSM-SEN3*
NC_035666	*Eunicella*	*albicans*	-	*SNSB-BSPG 2015 XXXI GW1815*
JQ397290	*Eunicella*	*cavolini*	*Isola d'Elba*	-
JQ397291	*Eunicella*	*cavolini*	*Isola d'Elba*	-
JQ397292	*Eunicella*	*cavolini*	*Isola d'Elba*	-
NC_035667	*Eunicella*	*cavolinii*	-	*SNSB-BSPG 2015 XXXI GW4597*
KX051577	*Eunicella*	*racemosa*	*Atlantic - Morocco*	*BEIM-26*
JQ397293	*Eunicella*	*singularis*	*Cap de Creus*	-
JQ397294	*Eunicella*	*singularis*	*Cap de Creus*	-
KX051571	*Eunicella*	*singularis*	*Cap de Creus*	*BEIM-11*
KX051572	*Eunicella*	*singularis*	*Cap de Creus*	*BEIM-13*
JQ397307	*Eunicella*	*sp.*	-	-
JQ397308	*Eunicella*	*sp.*	-	-
JQ397311	*Eunicella*	*sp.*	-	-
JX203795	*Eunicella*	*tricoronata*	-	*RMNH Coel.40814*
NC_062012	*Eunicella*	*tricoronata*	-	-
JQ397300	*Eunicella*	*verrucosa*	*Tarragona*	-
JQ397302	*Eunicella*	*verrucosa*	*Tarragona*	-
JQ397305	*Eunicella*	*verrucosa*	*Tarragona*	-
JQ397306	*Eunicella*	*verrucosa*	*Tarragona*	-
NC_073494	*Eunicella*	*verrucosa*	*United Kingdom: England, Lyme Bay, East, Tennants Reef*	-
KX904973	*Swiftia*	*pacifica*	-	-
KX905018	*Swiftia*	*simplex*	-	-

Table S4 : summary of the different datasets; for transcriptomes, the first four datasets include variable and non variable sites (all sites), while the "polymorphic sites" and the "1% SNPs" datasets only consider SNPs, i.e. variable sites. For the "all" datasets we indicate the number of contigs and the number of sites retained from reads2snp. See main text for details

dataset	samples	sites / assembly	number of individuals	number of contigs / SNPs	analyses
Transcriptomes					
all sites	all	all from reads2snp	20	61500 contigs / 101516577 sites	build SNPs datasets
all-CS	*cavolini* / *singularis*	all from reads2snp	20	61947 contigs / 101515803 sites	speciation scenarios with DILS
all-CV	*cavolini* / *verrucosa*	all from reads2snp	20	59702 contigs / 100704015 sites	speciation scenarios with DILS
all-SV	*singularis* / *verrucosa*	all from reads2snp	20	61373 contigs / 101444729 sites	speciation scenarios with DILS
polymorphic sites	all	polymorphic sites ; no missing data	20	31369 SNPs	F_{ST}, LEA, PCA
1 % SNPs	without *verrucosa*	polymorphic sites ; no missing data ; 1 % highest F_{ST} *cavolini* / *singularis*	20	326 SNPs	NewHybrids
RAD sequencing					
RAD_denovo	all	all, *de novo* assembly	67	16362 SNPs	F_{ST}, LEA
RAD_EC	all	all, assembly on *E. cavolini* genome	65	12952 SNPs	F_{ST}, LEA
RAD_ES	all	all, assembly on *E. singularis*	67	13342 SNPs	F_{ST}, LEA

		genome			
RAD_EV	all	all, assembly on *E. verrucosa* genome	65	29061 SNPs	F_{ST}, LEA, PCA
RAD_denovo 1%	without *verrucosa*	1 % highest F_{ST} *cavolini* / *singularis*	67	163 SNPs	NewHybrids
RAD_EC 1%	without *verrucosa*	1 % highest F_{ST} *cavolini* / *singularis*	65	130 SNPs	NewHybrids
RAD_ES 1%	without *verrucosa*	1 % highest F_{ST} *cavolini* / *singularis*	67	133 SNPs	NewHybrids
RAD_EV 1%	without *verrucosa*	1 % highest F_{ST} *cavolini* / *singularis*	65	290 SNPs	NewHybrids

Table S5: parameters used in the DILS analyses: Max_NA : maximum proportion of missing data ; Lmin : minimum sequence length per gene ; nMin : minimum number of sequences per gene and per species ; jSFS : use of joint Site Frequency Spectrum as an additional set of summary statistics ; constant / variable : consider constant or variable population size ; minimum and maximum values for the following priors : Tsplit : time of split, Ne : population size, M : migration rate. All other priors were kept at default values. For all analyses we used the option for coding regions, we didn't use any outgroup, we used the bimodal model for barriers, and the "normal" computation mode. The last column indicates the code used to describe the corresponding analysis in the text. The ranges of prior were chosen after preliminary analyses where we analysed the goodness of fit of the data to the models and priors. We used a mutation rate of 3.10^{-9}.

dataset	max_NA	Lmin	nMin	jSFS	Tsplit	Ne	M
all-CS	0.1	30	10	yes	100 – 2 000 000	100 – 2 000 000	0-30
all-CV	0.1	30	6	yes	100 – 2 000 000	100 – 2 000 000	0-30
all-SV	0.1	30	6	yes	100 – 2 000 000	100 – 2 000 000	0-30

Table S6: frequency of Symbiodiniaceae sequences in the individual transcriptomes on the basis i) of the proportion of raw reads mapped on the Symbiodiniaceae transcriptome, and ii) on the proportion of contigs in individual transcriptomes following the BLAT analysis. “meta” indicate the meta-transcriptome assembly based on all samples. See Table S2 for the codes of samples.

Sample	Species	Raw reads	Transcriptome
e-cavol-anb-a	*E. cavolini*	0,0171	0.00305
e-cavol-anb-d	*E. cavolini*	0,0087	0.00268
e-cavol-som-a	*E. cavolini*	0,0087	0.00426
e-cavol-vil-a	*E. cavolini*	0,0184	0.00350
e-cavol-vil-b	*E. cavolini*	0,0255	0.00333
e-hybri-jpb-a	hybrid	0,0123	0.00262
e-hybri-mfn-a	hybrid	0,0076	0.00321
e-hybri-mfn-b	hybrid	0,0079	0.00270
e-hybri-mfn-e	hybrid	0,0162	0.00302
e-singu-ban-a	*E. singularis*	0,0192	0.00675
e-singu-ban-b	*E. singularis*	0,0140	0.00410
e-singu-ban-c	*E. singularis*	0,0080	0.00647
e-singu-cav-a	*E. singularis*	0,0261	0.02263
e-singu-mfn-a	*E. singularis*	0,0233	0.03419
e-singu-mfn-b	*E. singularis*	0,0129	0.02745
e-singu-sdo-a	*E. singularis*	0,0158	0.03644
e-singu-sdo-b	*E. singularis*	0,0207	0.03686
e-verru-ros-a	*E. verrucosa*	0,0075	0.00276
e-verru-ros-c	*E. verrucosa*	0,0082	0.00282
e-verru-som-a	*E. verrucosa*	0,0098	0.00279
meta			0.01393

Table S7: p-values of the Pairwise-Wilcoxon test on the frequency of Symbiodiniaceae. A) on the basis of read counts with Salmon; B) on the proportion of assembled sequences with the BLAT analysis

A)

	E. cavolini	hybrids	*E. singularis*
hybrids	0.69		
E. singularis	0.69	0.36	
E. verrucosa	0.57	0.69	0.15

B)

	E. cavolini	hybrids	*E. singularis*
hybrids	0.571		
E. singularis	0.019	0.020	
E. verrucosa	0.571	1	0.048

Table S8: transcriptome data; above diagonal: average net divergence estimated from DILS for the "all" pairwise datasets (the hybrids were not included in the DILS analysis); below diagonal pairwise F_{ST} estimated from variable sites only ("polymorphic SNPs" dataset; see main text and Table S3 for details)

	E. cavolini	hybrids	*E. singularis*	*E. verrucosa*
E. cavolini	-	-	0.0018	0.0067
hybrids	0.069	-	-	-
E. singularis	0.207	0.073	-	0.0070
E. verrucosa	0.432	0.456	0.529	-

Table S9: RAD sequencing data; pairwise F_{ST} estimated from all SNPs after filtering, for the four assembly strategies

de novo assembly

	E. cavolini	hybrids	*E. singularis*	*E. verrucosa*
E. cavolini	-			
hybrids	0.156	-		
E. singularis	0.380	0.122	-	
E. verrucosa	0.587	0.574	0.658	-

assembly on the genome of *E. cavolini*

	E. cavolini	hybrids	*E. singularis*	*E. verrucosa*
E. cavolini	-			
hybrids	0.155	-		
E. singularis	0.361	0.113	-	
E. verrucosa	0.563	0.553	0.630	-

assembly on the genome of *E. singularis*

	E. cavolini	hybrids	*E. singularis*	*E. verrucosa*
E. cavolini	-			
hybrids	0.141	-		
E. singularis	0.352	0.114	-	
E. verrucosa	0.563	0.554	0.634	-

assembly on the genome of *E. verrucosa*

	E. cavolini	hybrids	*E. singularis*	*E. verrucosa*
E. cavolini	-			
hybrids	0.126	-		
E. singularis	0.294	0.093	-	
E. verrucosa	0.514	0.515	0.568	-

Table S9: estimated parameters for the different evolutionary scenarios for the three pairwise comparisons. We present here the results of estimations for the optimized posterior with the random forests approach implemented in DILS. For each parameter we present the highest posterior density (HPD), with the median, and the lower and higher 2.5 % limits. Models : SC : secondary contact ; AM : ancestral migration. Parameters : N : effective size ; $founders_X$: number of founder individuals in species X ; T_{split} : time of split at which the ancestral population subdivides in two populations ; T_{SC} : time of secondary contact ; T_{AM} : time of the end of gene flow for ancestral migration ; T_{dem_X} : time of demographic event for species X ; M_{XY} : introgression rate from Y to X . For all parameters, the subscripts indicate the species : A for ancestral, C for *E. cavolini*, S for *E. singularis,* and V for *E. verrucosa*. Times are given in generations, migration in numbers of migrants per generation.

A) comparison *E. cavolini* / *E. singularis*

	HPD 0.025	HPD median	HPD 0.0975
constant size, SC			
N_C	545985	633894	733842
N_S	168290	192199	225073
N_A	537403	581831	632310
T_{split}	336413	403273	476196
T_{SC}	51536	62039	71760
M_{CS}	12	15	17
M_{SC}	12	15	18
variable size, SC			
N_C	531986	665965	875780
N_S	185826	222258	276889
N_A	515018	578861	640504
$founders_C$	0	1	1
$founders_S$	0	1	1
T_{dem_C}	250520	339056	418963
$T_{demS\_}$	245400	350132	454320
T_{split}	330907	434060	542765
T_{SC}	40560	57552	75405
M_{CS}	14	19	24
M_{SC}	8	12	16

B) comparison *E. cavolini* / *E. verrucosa*

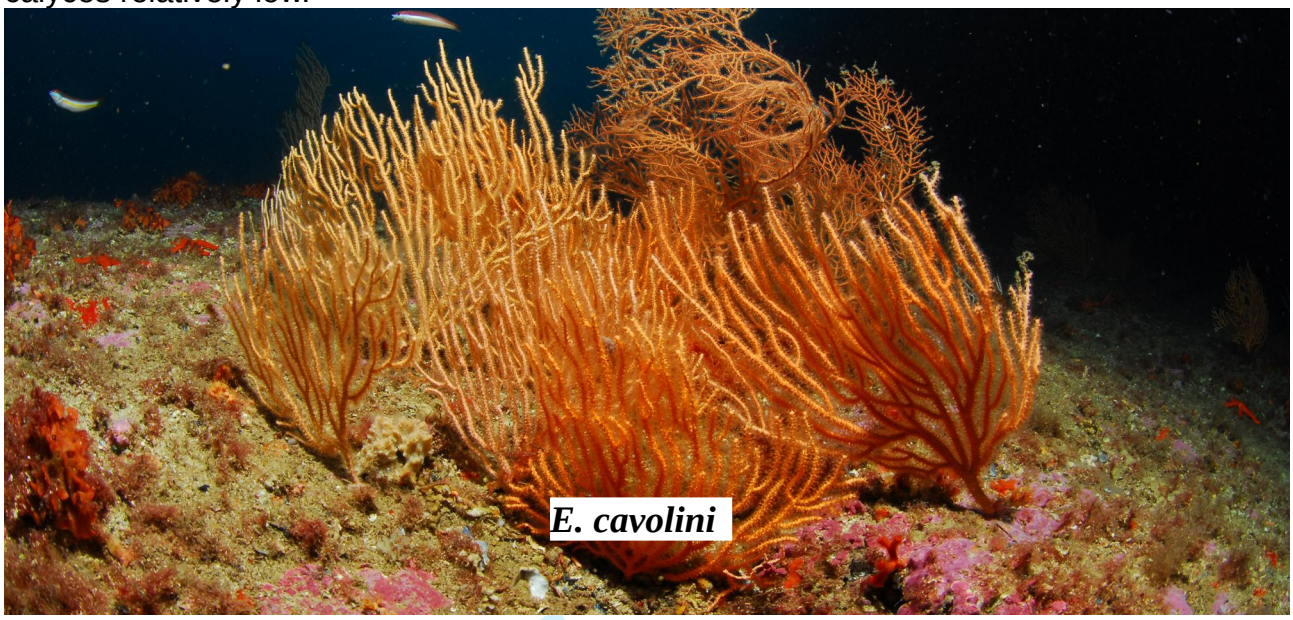
	HPD 0.025	HPD median	HPD 0.0975
constant size, AM			
N_C	630969	744556	875220
N_V	648850	755298	920095
N_A	698501	784512	879664
T_{split}	909392	1054488	1225792
T_{AM}	840920	991118	1147073
M_{CV}	4	6	7
M_{VC}	9	12	14
variable size, AM			
N_C	777526	1099410	1694348
N_V	871210	1230360	1803231
N_A	692366	793880	930000
$founders_C$	0	1	1
$founders_V$	0	0	1
T_{dem_C}	237960	369260	496633
T_{dem_V}	335620	509096	679348
T_{split}	819714	1051517	1367074
T_{AM}	782210	930590	1104120
M_{CV}	7	12	16
M_{VC}	11	22	31

C) comparison *E. singularis* / *E. verrucosa*

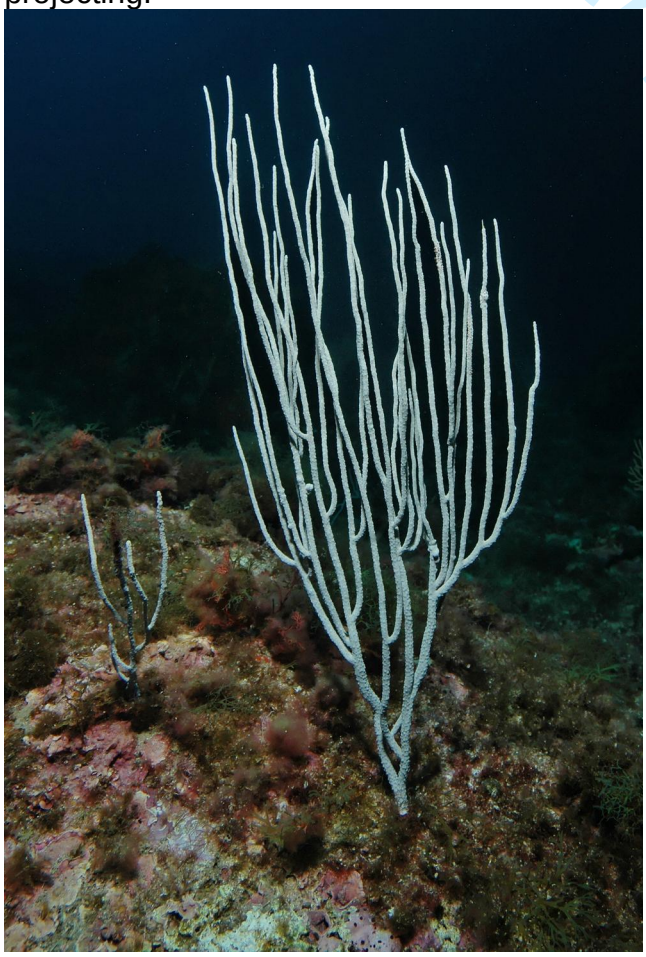
	HPD 0.025	HPD median	HPD 0.0975
constant size, AM			
N_S	263390	298162	336536
N_V	490519	592796	715930
N_A	632004	708517	790246
T_{split}	741840	899098	1091610
T_{AM}	698891	811827	934655
M_{SV}	10	14	17
M_{VS}	21	27	33
variable size, AM			
N_S	281023	386388	494606
N_V	856542	1165039	1566087
N_A	592428	697054	797828
$founders_C$	0	0	0
$founders_V$	0	0	1
T_{dem_S}	166517	273546	374076
T_{dem_V}	226988	360174	493360
T_{split}	713634	926756	1207281
T_{AM}	454059	659458	858558
M_{SV}	3	4	6
M_{VS}	1	1	2

Figure S1: examples of morphological diversity in *Eunicella* species in the area of Marseille. See Carpine and Grasshoff (1975) for details.

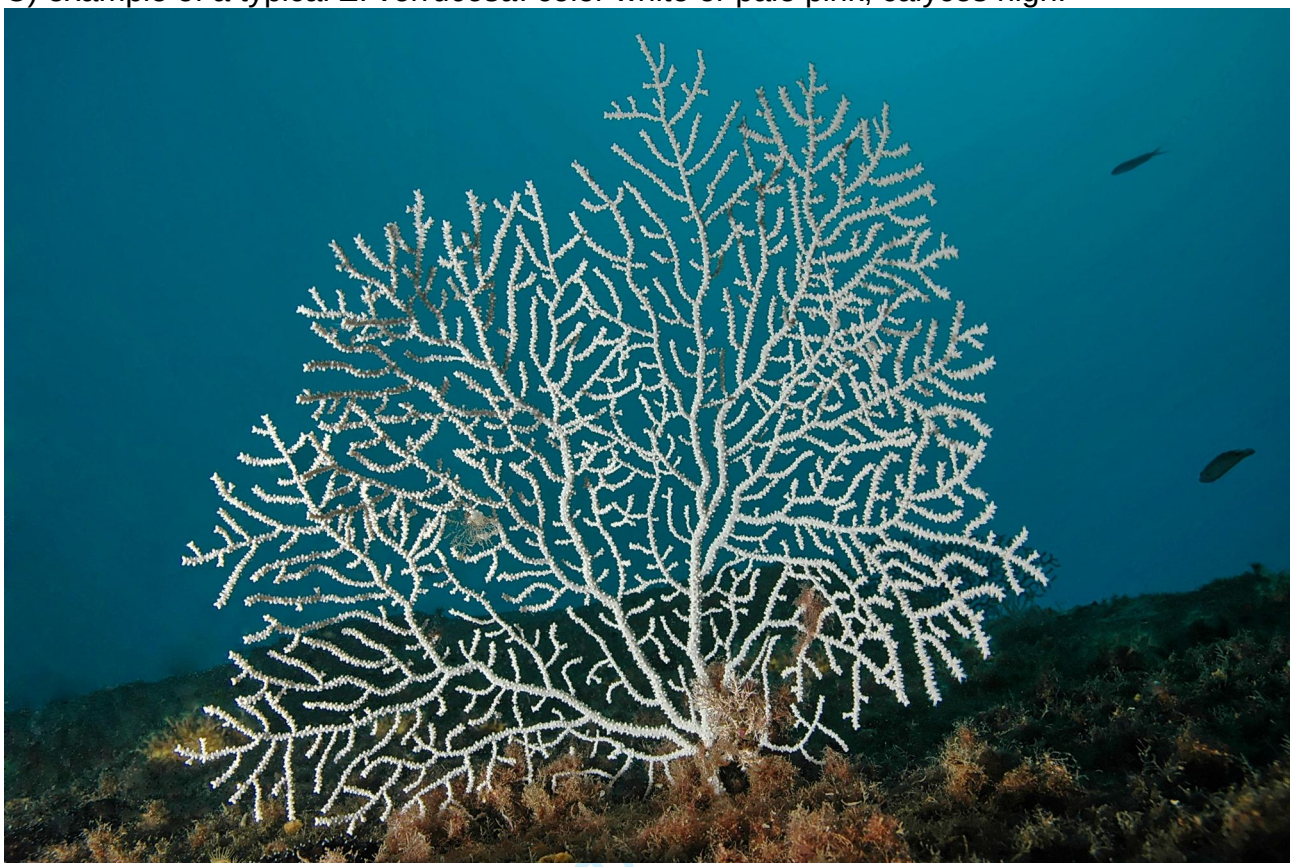
A) example of typical *E. cavolini* colonies (in the foreground): color yellow - orange, calyces relatively low.



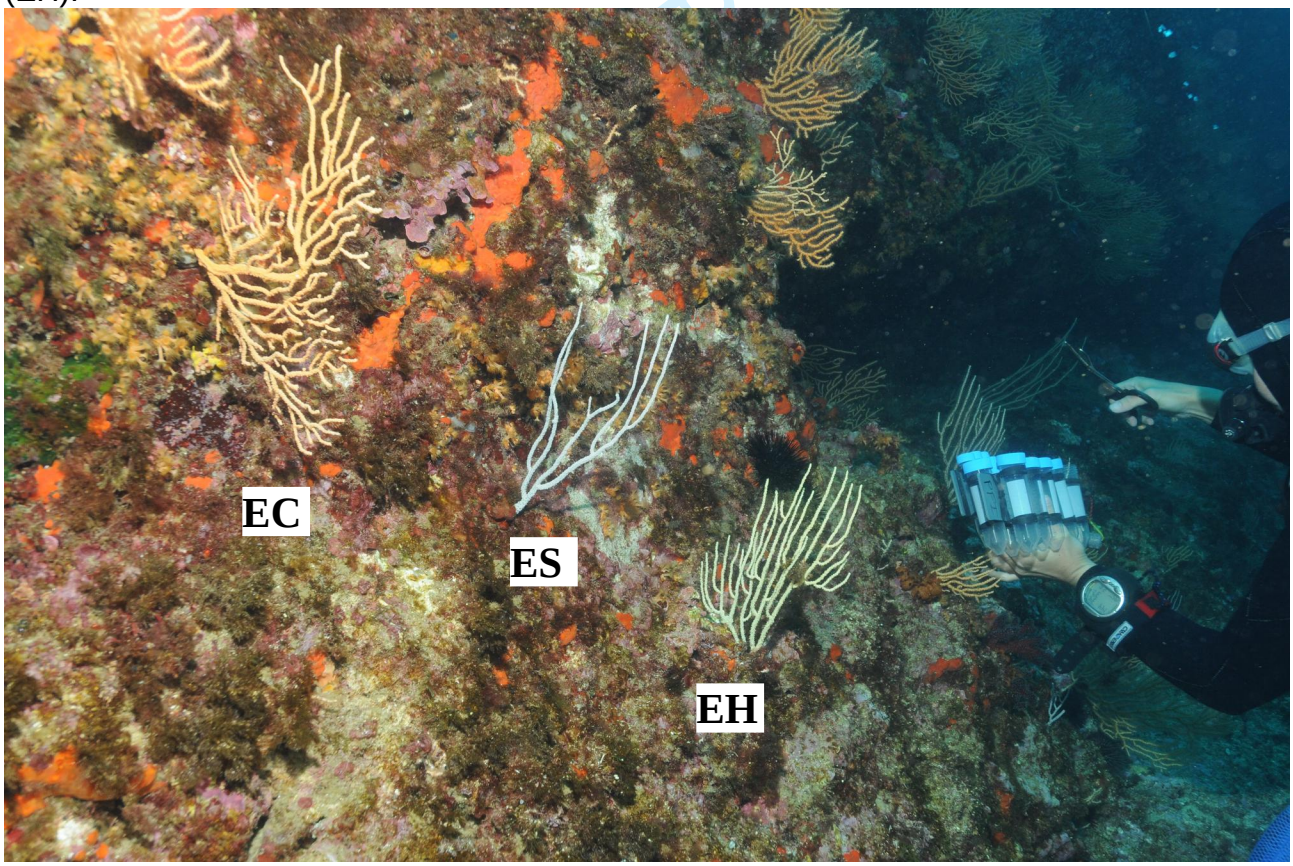
B) example of a typical *E. singularis*: color white, long terminal branches, calyces not projecting.



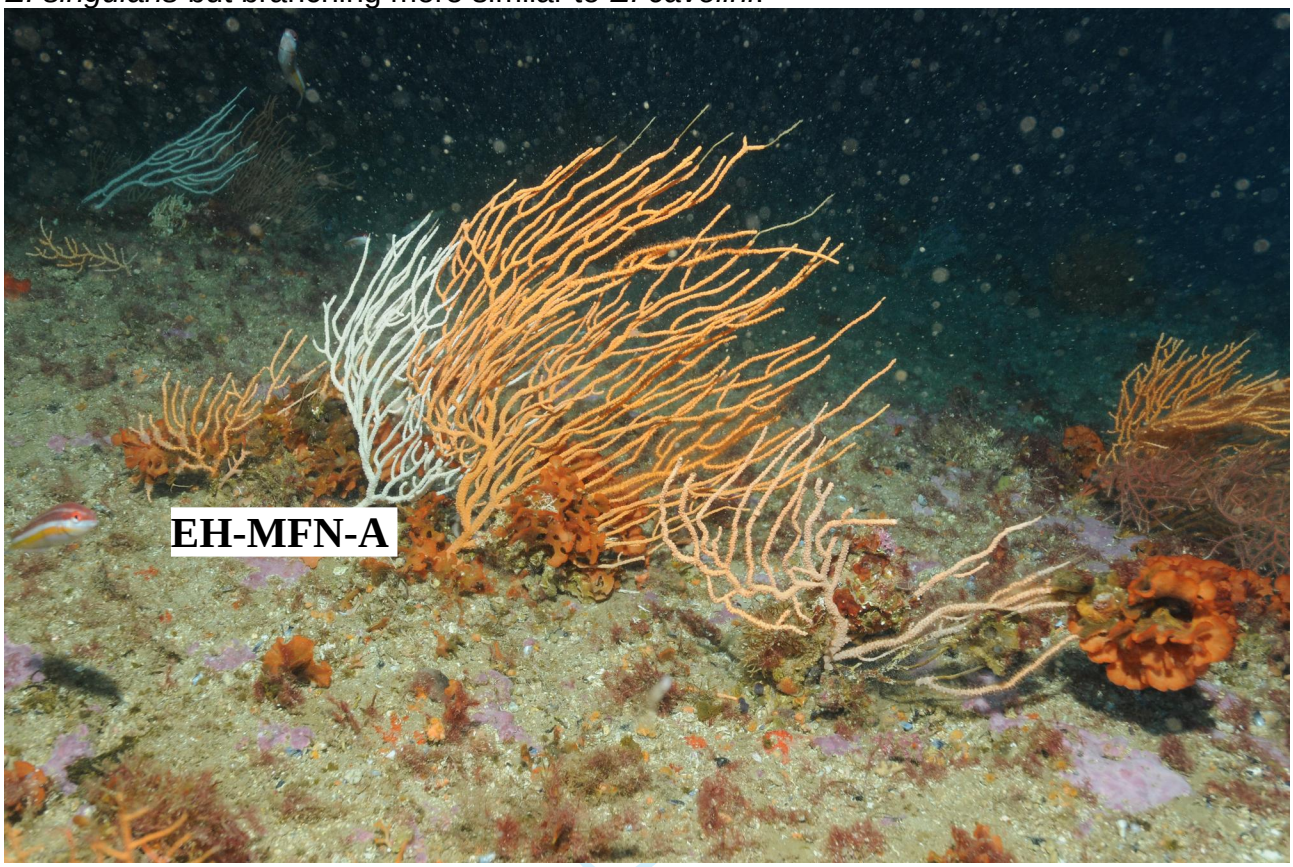
C) example of a typical *E. verrucosa*: color white or pale pink, calyces high.



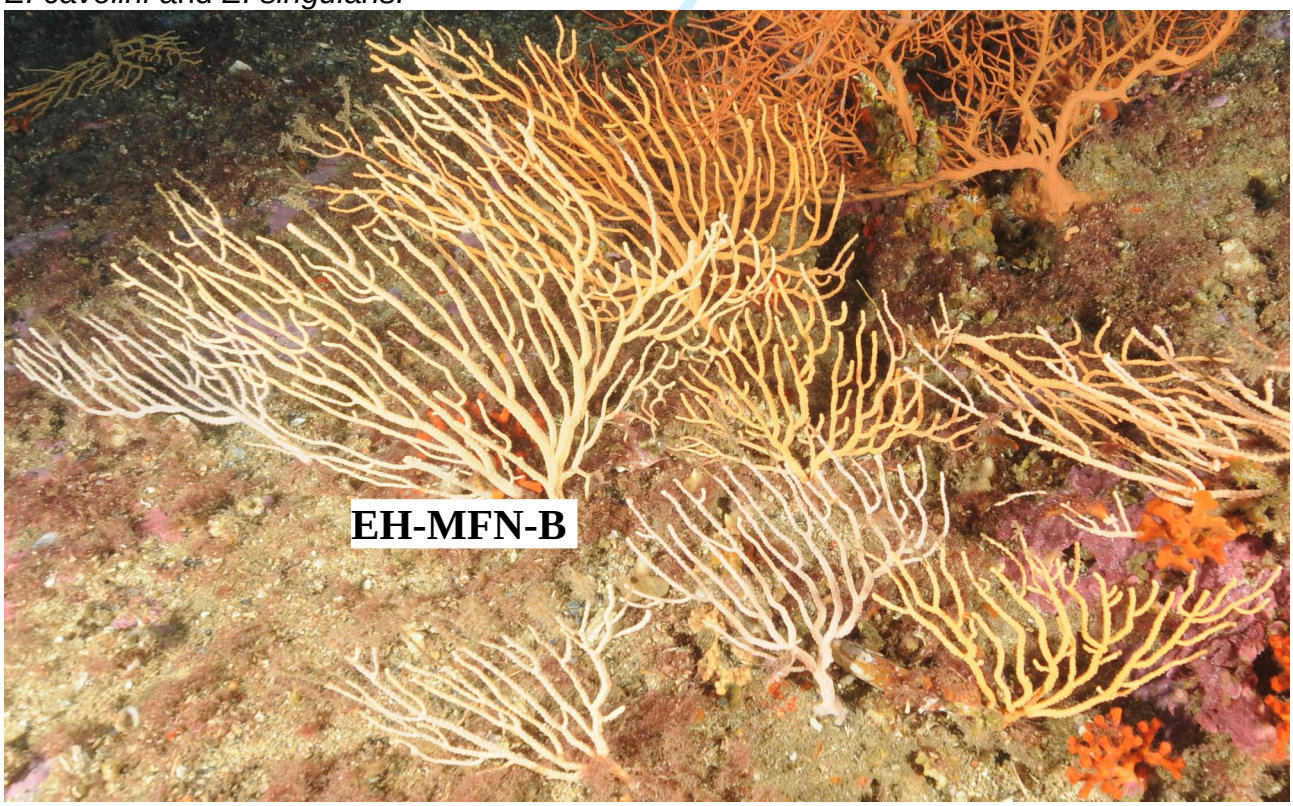
D) sampling with the presence of *E. cavolini* (EC), *E. singularis* (ES) and a potential hybrid (EH).



E) morphology of the colony EH-MFN-A (white, in the background) with white color as *E. singularis* but branching more similar to *E. cavolini*.



F) morphology of the colony EH-MFN-B with intermediate branching and color between *E. cavolini* and *E. singularis.*



G) morphology of the colony EH-MFN-E with intermediate branching and color between *E. cavolini* and *E. singularis.*

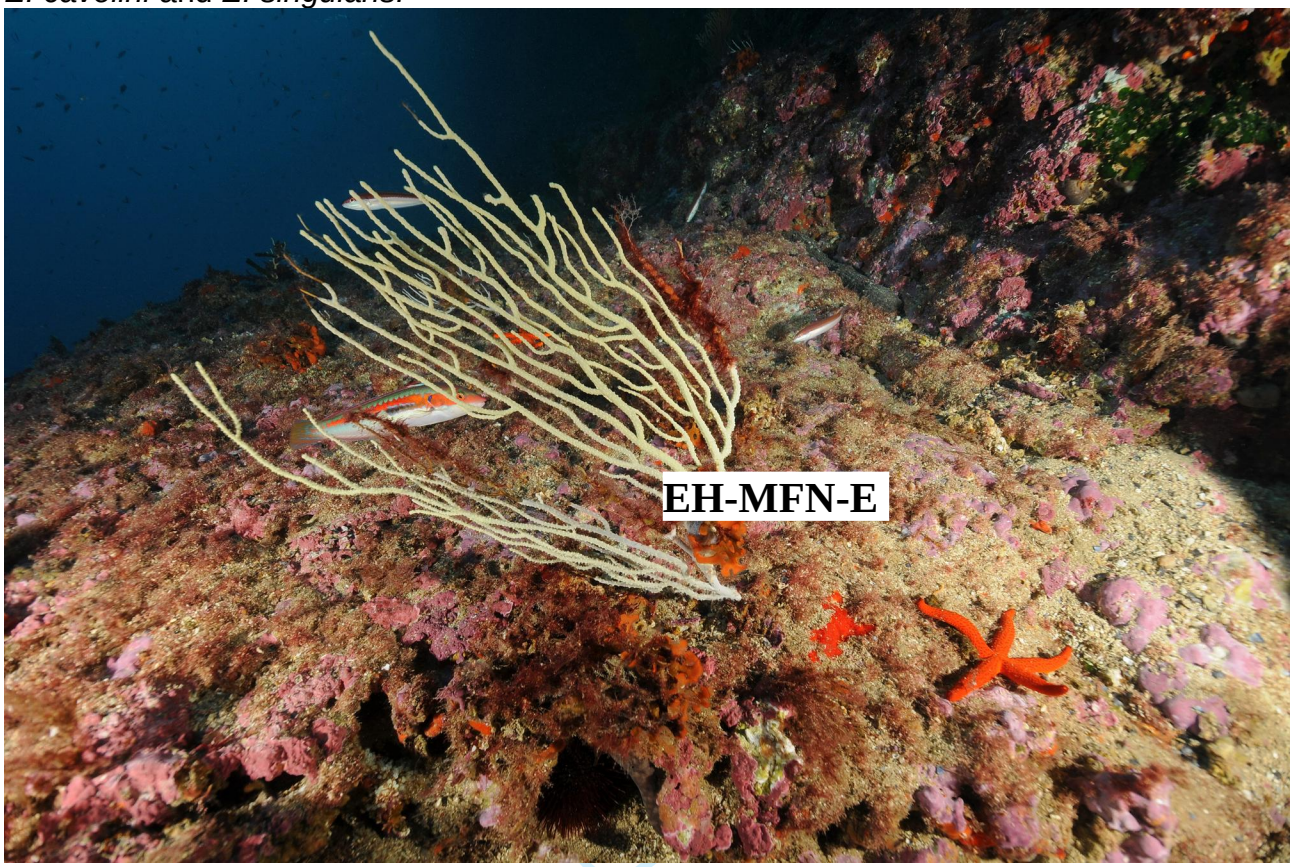


Figure S2: map of sampling sites for RAD sequencing in the area of Marseille. The symbols present the different samples: EC E. cavolini, ES E. singularis, EV E. verrucosa, HY hybrids. The three letters correspond to the codes of the sampling. The maps have been produced with the marmap R package (Pante & Simon-Bouhet, 2013) and following the tutorial of Krueger-Hadfield (2015).

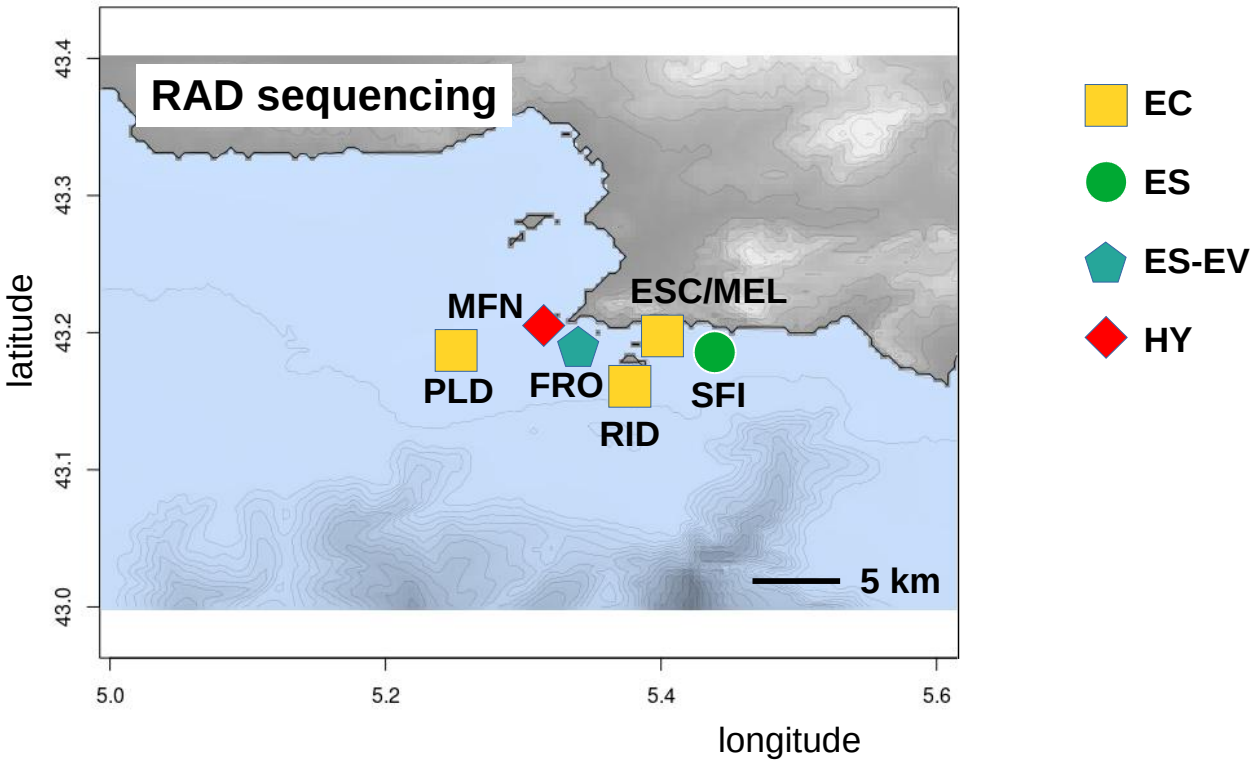


Figure S3: phylogenetic relationships among *Eunicella* species. The phylogenetic reconstruction has been performed with mitochondrial MutS sequences obtained from Genbank, with a search focused on *Eunicella* species. Sequences from the *Complexum* and *Swiftia* genera have been retained on the basis of a Blast search with the MutS sequence of *E. cavolini*, and according to the current systematics of octocorals (McFadden et al., 2022). The sequences corresponding to our three focal species come from previous studies and do not correspond to specimens sampled for the present study. The sequences have been edited with ugene (Okonechnikov et al., 2012). The phylogenetic reconstructions have been performed with the Maximum-Likelihood (ML) approach of IQ-TREE 2.1.1 (Nguyen et al., 2015). We used the ModelFinder option (Kalyaanamoorthy et al., 2017), and robustness was evaluated with 1000 ultrafast bootstraps (Hoang et al., 2018). The tree has been visualized with FigTree 1.4.4 (Rambaut, 2006) and was rooted with *Swiftia simplex* as outgroup. The numbers to the left of the nodes indicate the percentages of bootsraps. The Genbank accession numbers are listed in table S1.

References:

Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, *35*(2), 518–522.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587–589. https://doi.org/10.1038/nmeth.4285

McFadden, C. S., van Ofwegen, L. P., & Quattrini, A. M. (2022). Revisionary systematics of Octocorallia (Cnidaria: Anthozoa) guided by phylogenomics. *Bulletin of the Society of Systematic Biologists*, *1*(3).

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Molecular Biology and Evolution*, *32*(1), 268–274. https://doi.org/10.1093/molbev/msu300

Okonechnikov, K., Golosova, O., Fursov, M., & Ugene Team. (2012). Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics*, *28*(8), 1166–1167.

Rambaut, A. (2006). *FigTREE v1.4*. University of Edinburgh. http://tree.bio.ed.ac.uk/software/figtree/

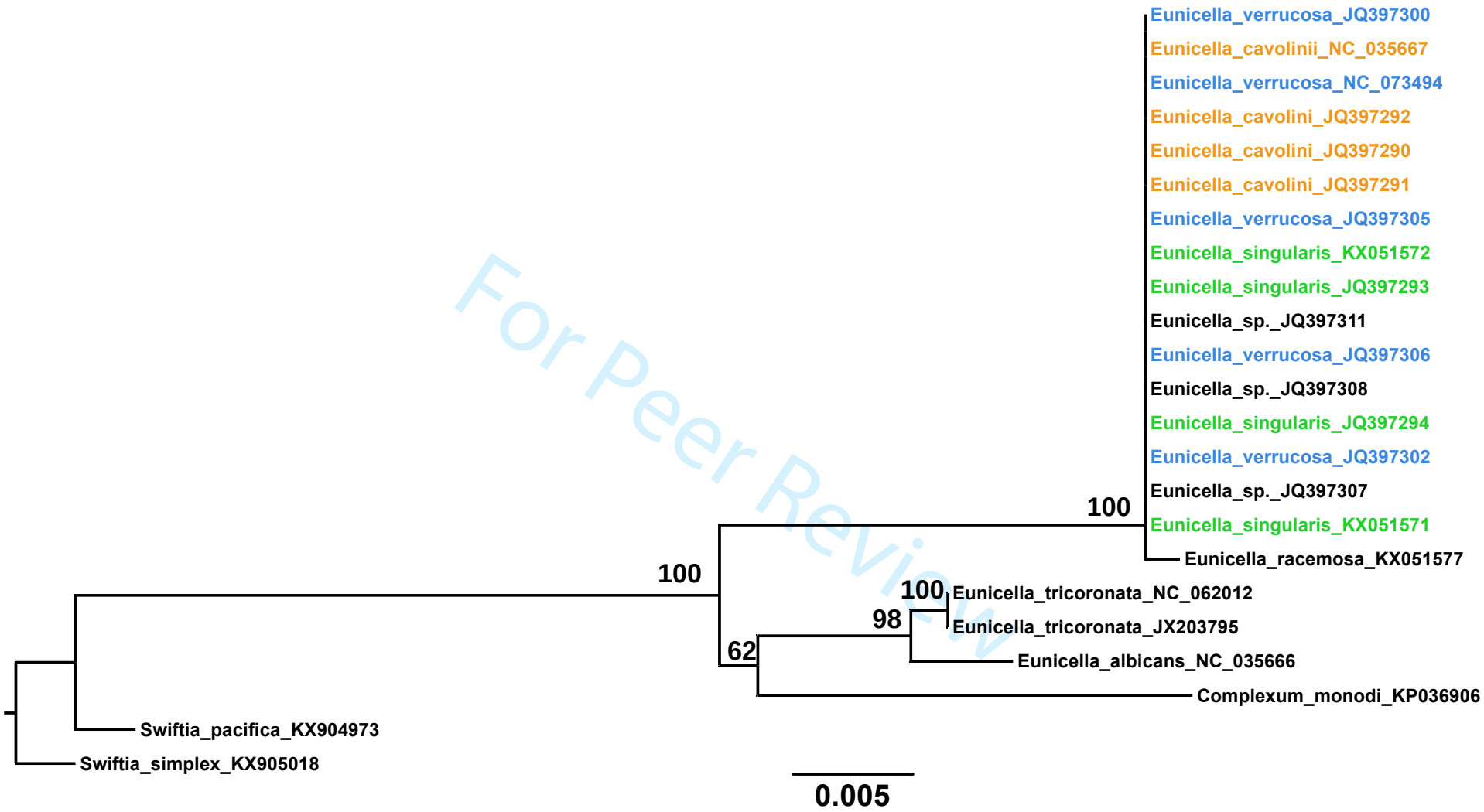
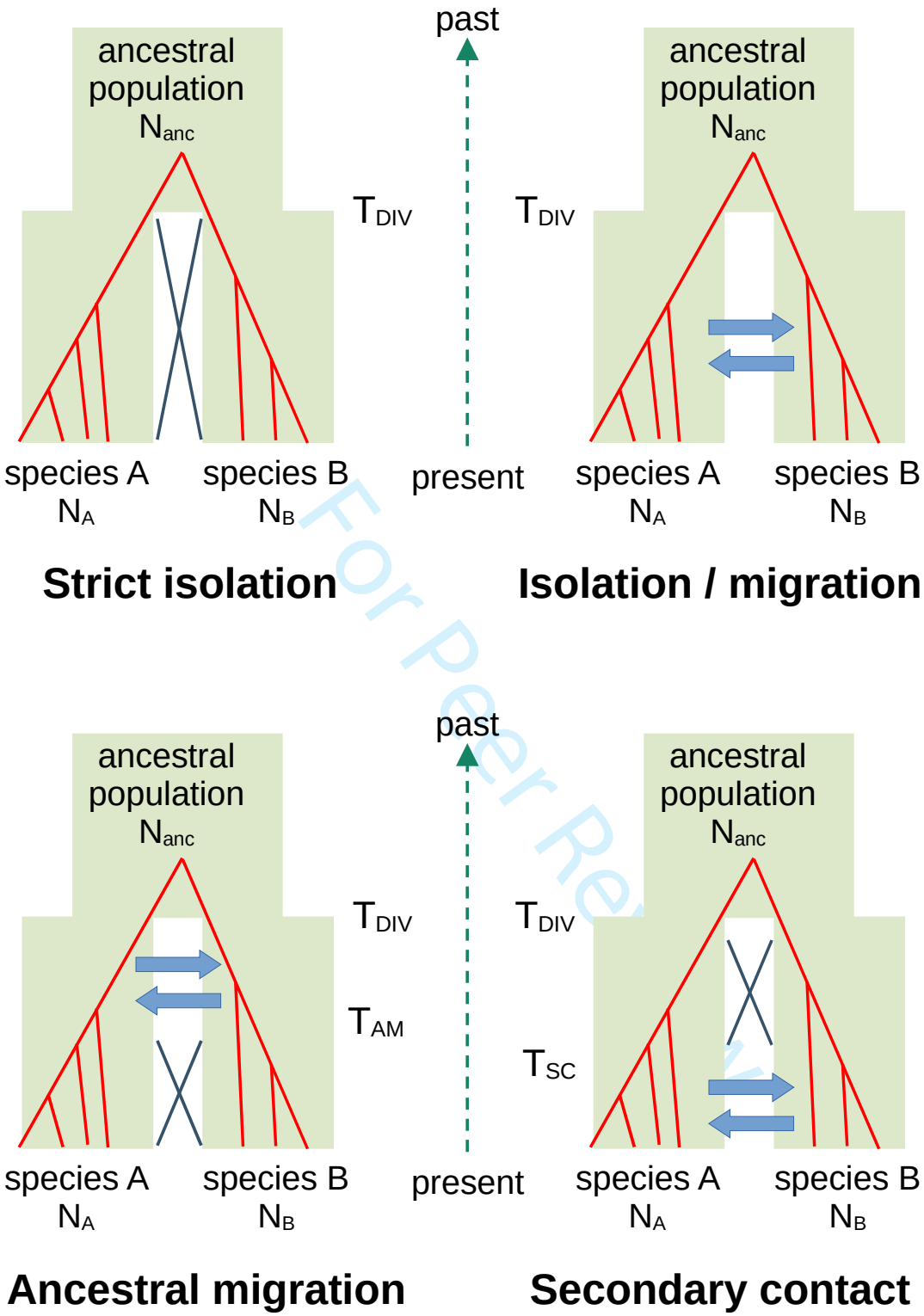


Figure S4: schematic representation of the four speciation scenarios tested with DILS (adapted from Roux *et al.*, 2016). The green background indicates the species history, and the red lines show an example of the history of one gene. All scenarios correspond to a divergence from an ancestral population but differ by the possibility and timing of migration (i.e. gene flow) versus isolation (no gene flow). Among these four scenarios, only isolation / migration and secondary contact imply current gene flow among species. In the strict isolation scenario, there is no gene flow after divergence, whereas in the ancestral migration scenario, divergence is followed by a period of gene flow, then by isolation. For simplicity we did not add here the possibility of variation in effective sizes (NA and NB). DILS also allowed to test genomic heterogeneity in gene flow, which corresponds to differences in probability of gene flow among loci.



Isolation (no gene flow)

Gene flow

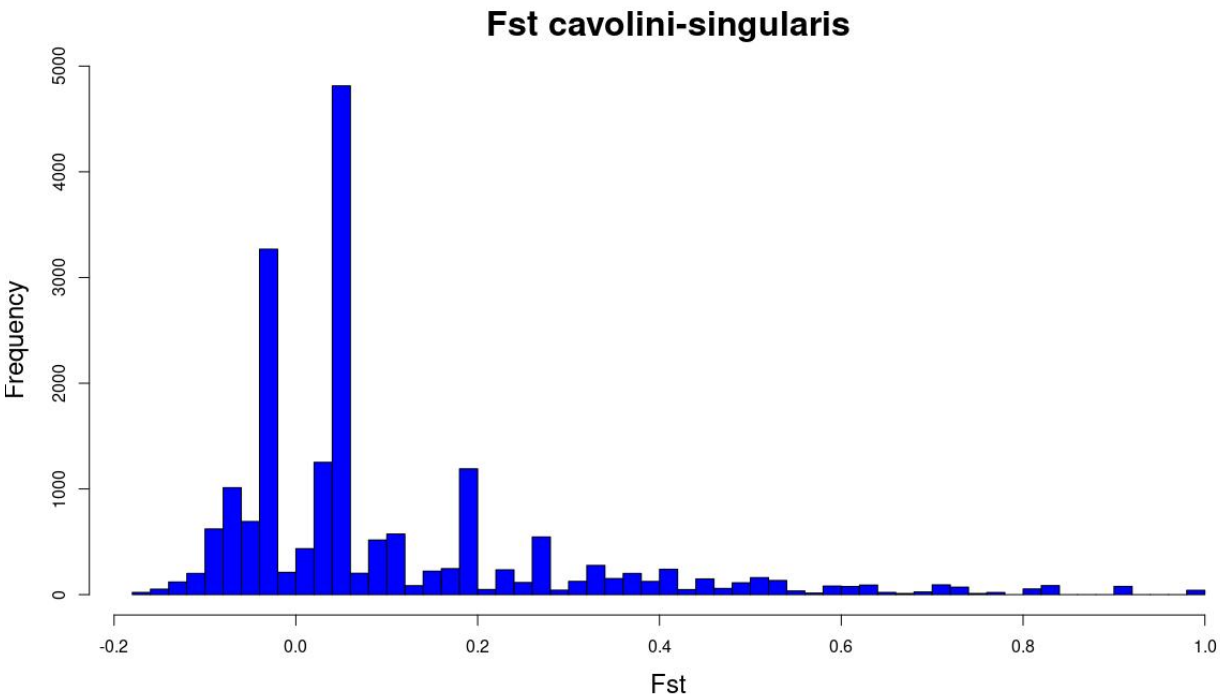
T_{DIV} divergence time

T_{AM} end of ancestral migration

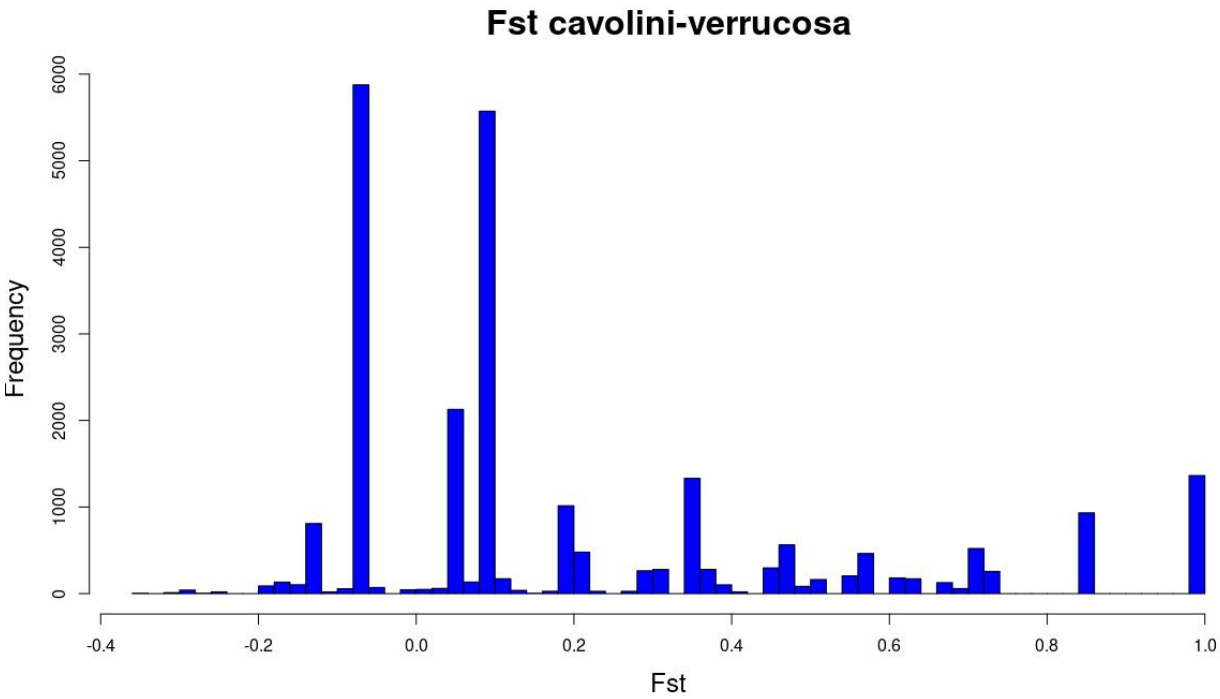
T_{SC} time of secondary contact

Figure S5: distribution of F_{ST} estimates over loci, for the pairwise comparisons among the three species, with the exclusion of potential hybrids.

A) comparison between *E. cavolini* and *E. singularis*



B) comparison between *E. cavolini* and *E. verrucosa*



C) comparison between *E.singularis* and *E. verrucosa*

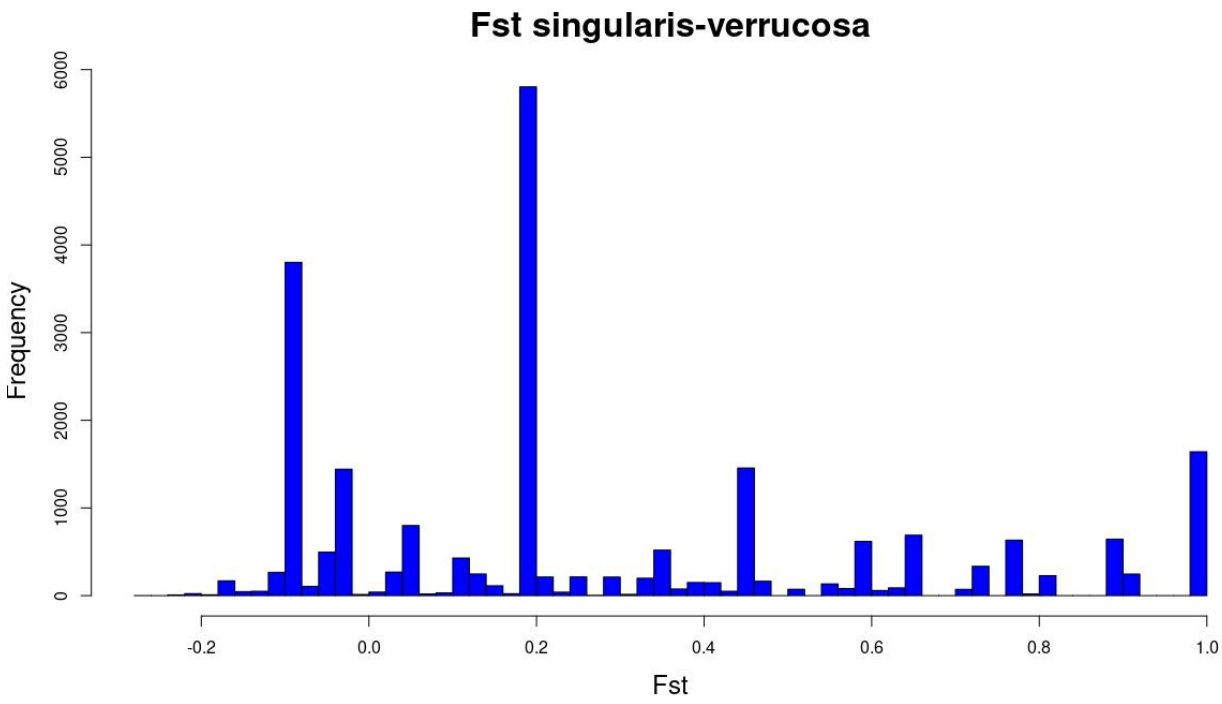


Figure S6: result of the cross-entropy analysis with the LEA R package

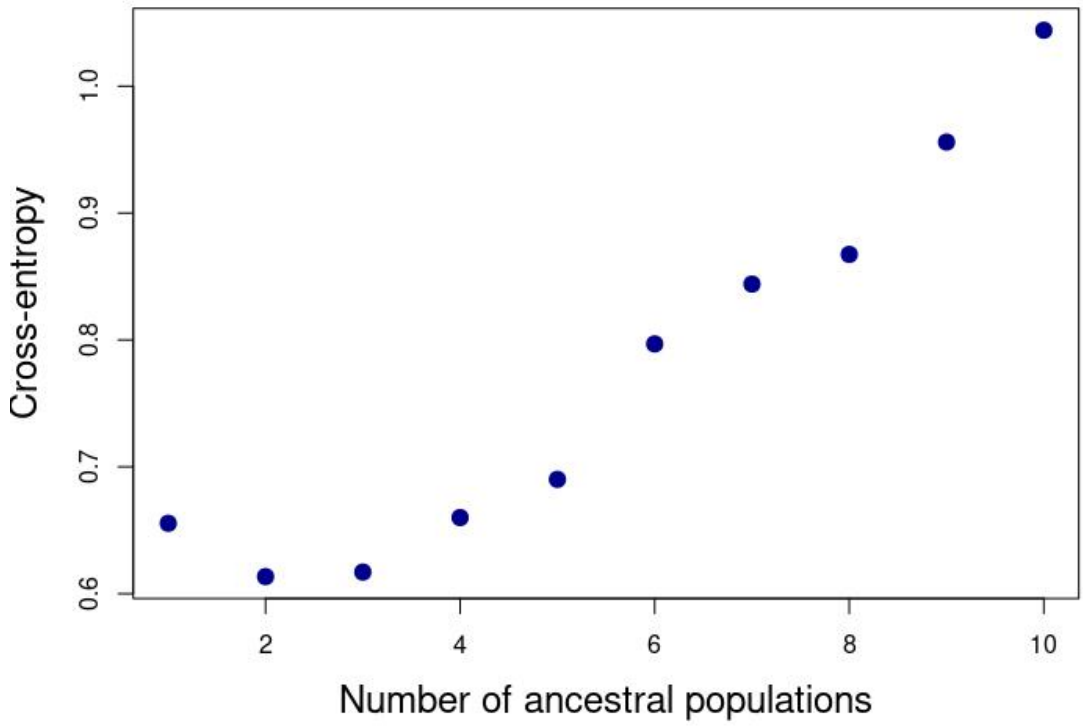


Figure S7: barplots of coancestry coefficients inferred with the LEA R package for K = 3 with RAD sequencing with the four assembly strategies. The red asterisks indicate the individuals used as prior for parental status in the newhybrids analysis. For clarity reasons, the results of the newhybrids analysis are not indicated here but they can be found in Table 1.

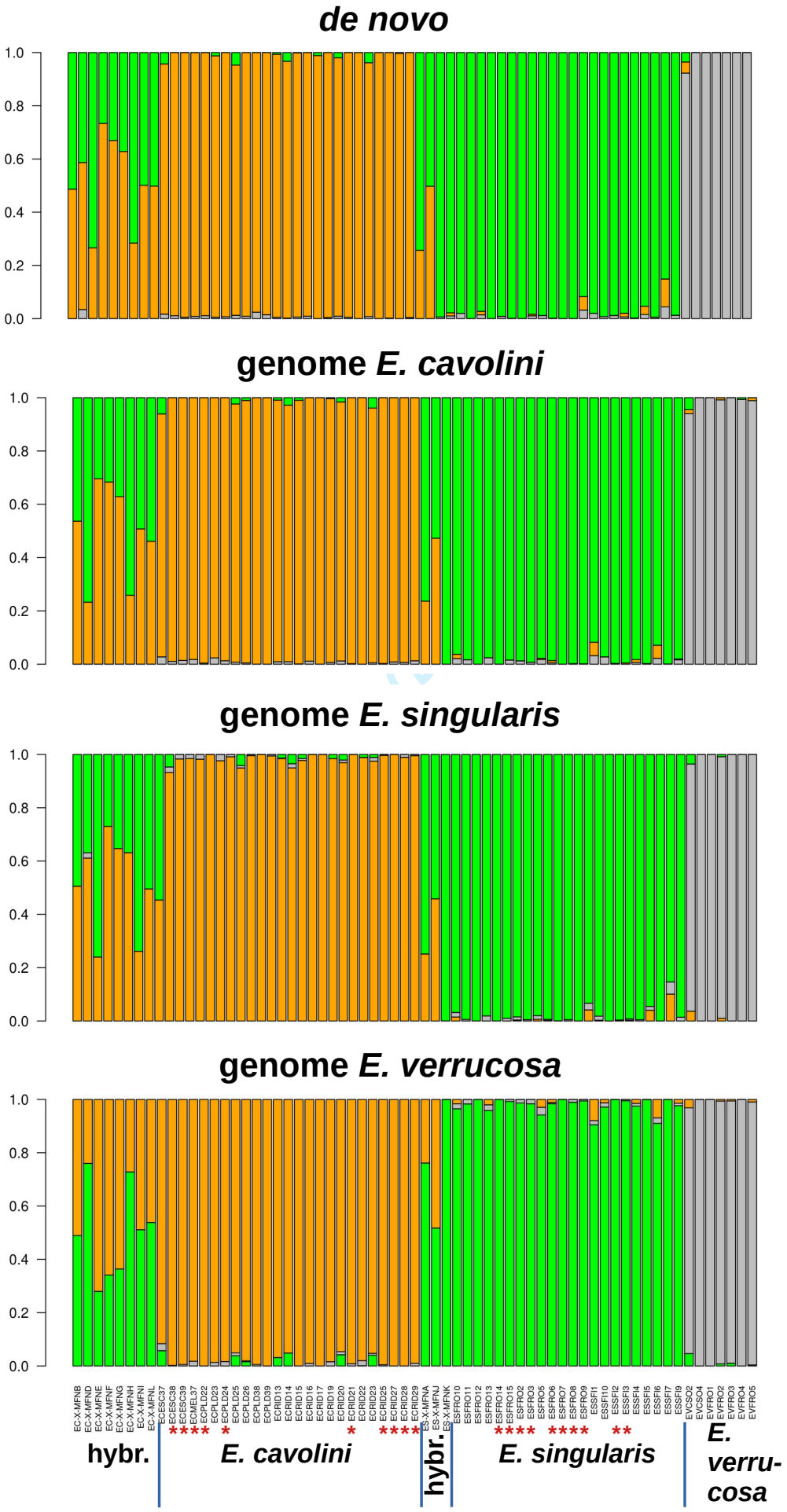
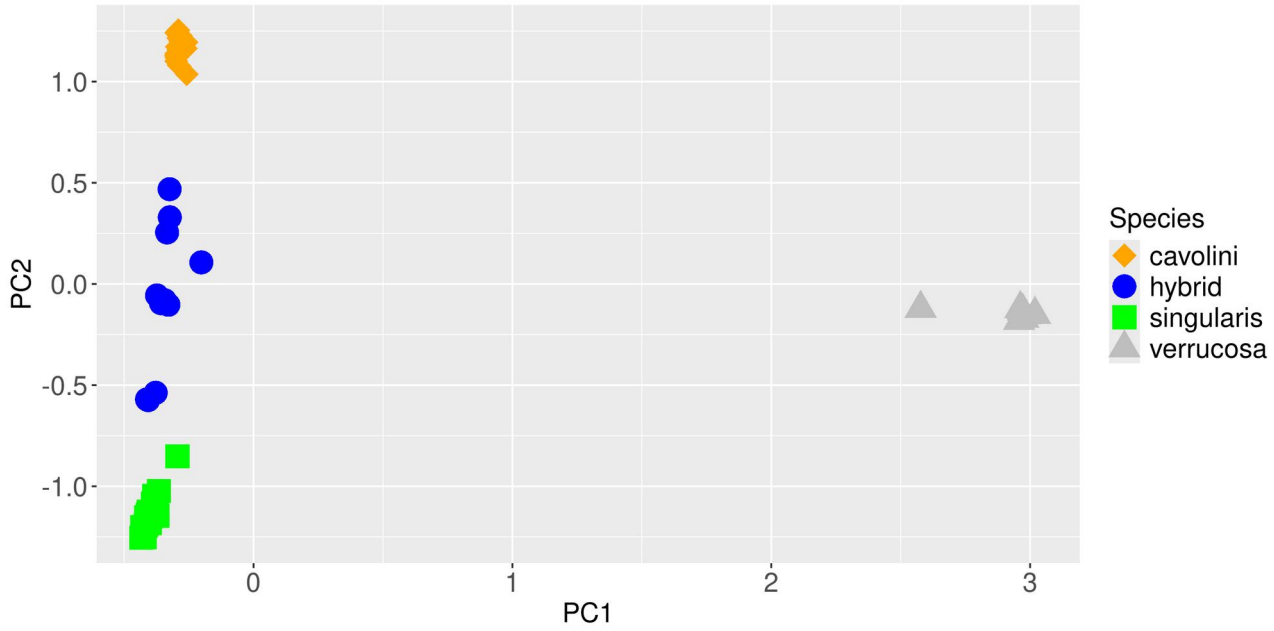
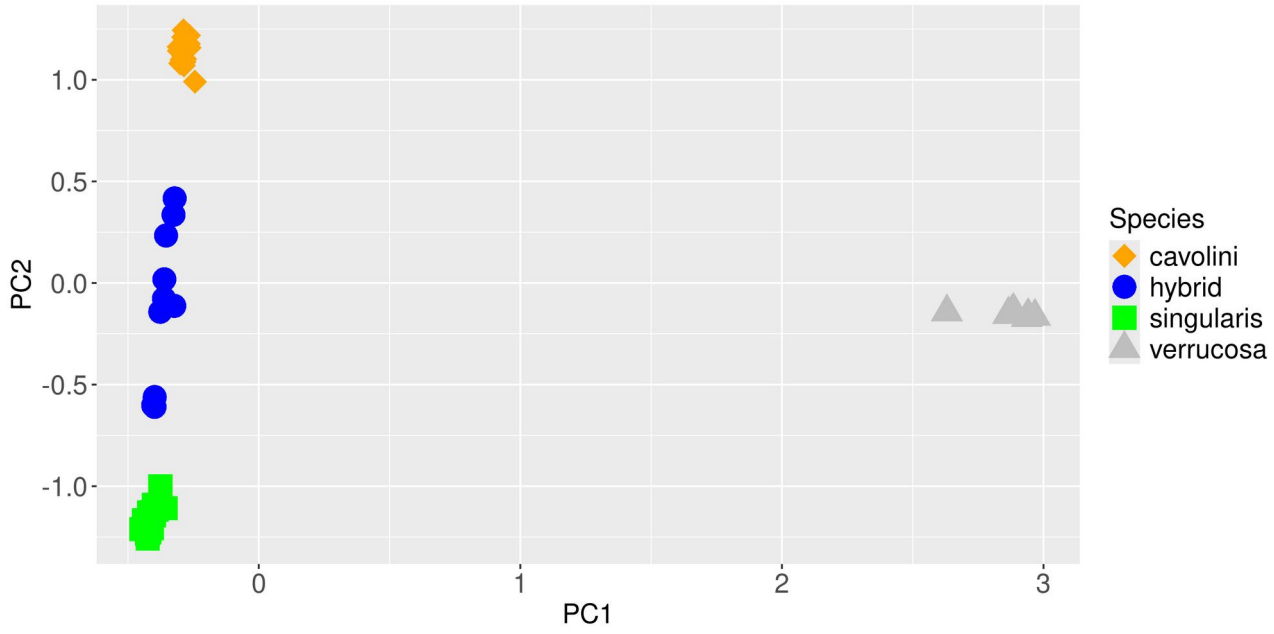

de novo
genome E. cavolini
genome E. singularis
genome E. verrucosa
1.0
0.8
0.6
0.4
0.2
0.0
hybr.
E. cavolini
hybr.
E. singularis
E. verru-cosa

Figure S8: Principal Component Analysis based A) on the RAD_denovo dataset; B) on the RAD_EC dataset; C) on the RAD_ES dataset; D) on the RAD_EV dataset.

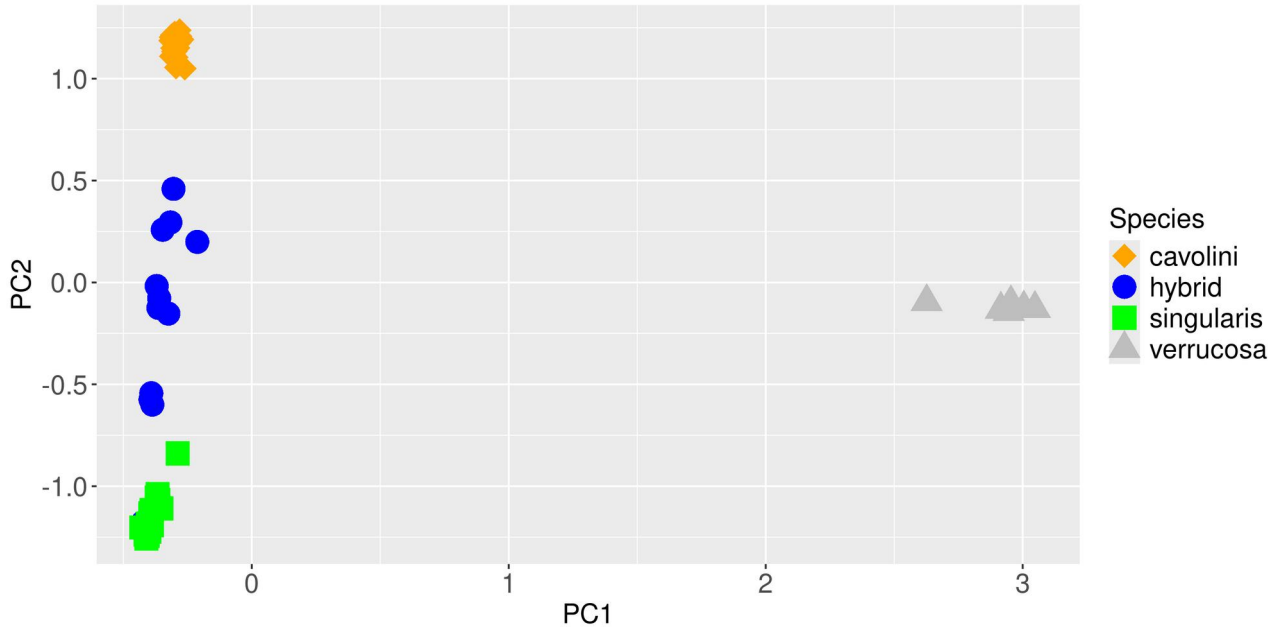
A) RAD sequencing, de novo assembly; axis 1 represents 21% of the variance, axis 2 represents 10.2% of the variance



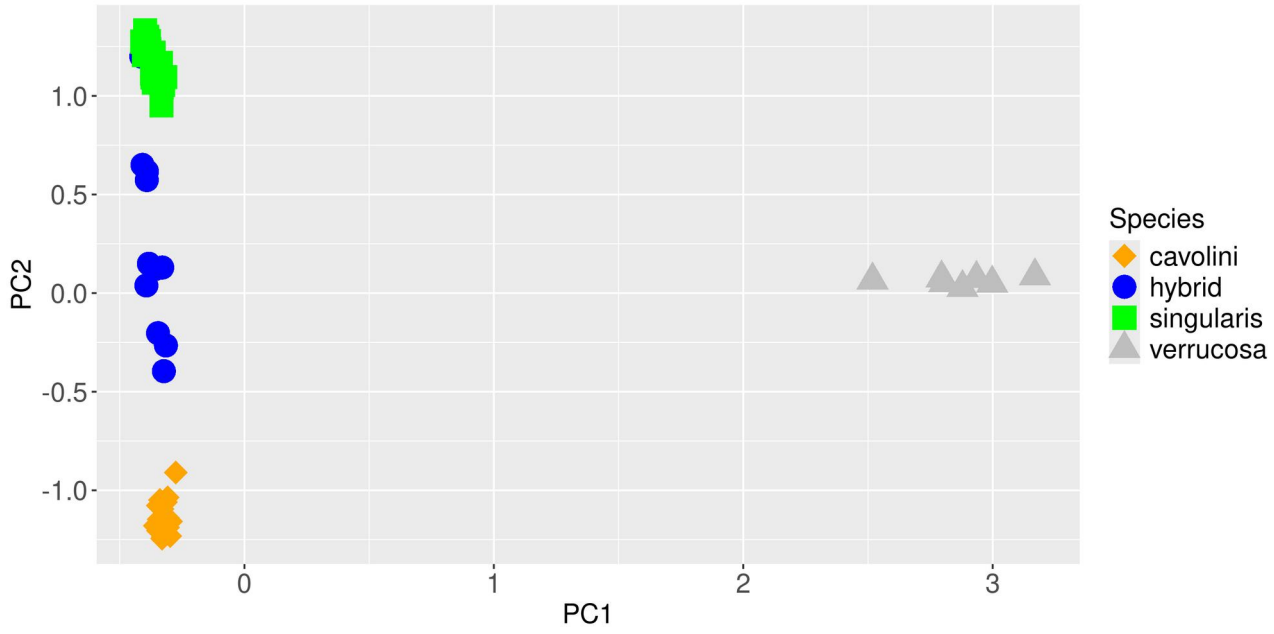
B) RAD sequencing, assembly on *E. cavolini* genome; axis 1 represents 18.7% of the variance, axis 2 represents 10.2% of the variance



C) RAD sequencing, assembly on *E. singularis* genome; axis 1 represents 18.8% of the variance, axis 2 represents 9.8% of the variance



D) RAD sequencing, assembly on *E. verrucosa* genome; axis 1 represents 14.1% of the variance, axis 2 represents 7.1% of the variance



Symbiotic status does not preclude hybridisation in Mediterranean octocorals

Supplementary Material S2: 18S rDNA metabarcoding

Didier Aurelle[1,2*], Anne Haguenauer[3], Marc Bally[1], Frédéric Zuberer[4], Jean-Baptiste Ledoux[5], Stéphane Sartoretto[6], Lamya Chaoui[7], Hichem Kara[7], Sarah Samadi[2], Pierre Pontarotti[8,9,10]

[1] Aix Marseille Univ, Université de Toulon, CNRS, IRD, MIO, Marseille, France
[2] Institut Systématique Evolution Biodiversité (ISYEB), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, CP 26, 75005 Paris, France.
[3] CNRS - Délégation Provence et corse, Marseille, France
[4] Aix Marseille Univ, CNRS, IRD, INRAE, OSU Inst. PYTHEAS, Marseille, France
[5] CIIMAR/CIMAR, Centro Interdisciplinar de Investigação Marinha e Ambiental, Universidade do Porto, Porto, Portugal.
[6] Ifremer, LITTORAL, 83500 La Seyne-sur-Mer, France
[7] Laboratoire Bioressources marines. Université d'Annaba Badji Mokhtar, Annaba - Algérie.
[8] Aix Marseille Univ, MEPHI, Marseille, France.
[9] IHU Méditerranée Infection, Marseille, France.
[10] CNRS SNC5039

*Corresponding author
Correspondence: didier.aurelle@univ-amu.fr

Objectives of the study

This section describes a preliminary sequencing test carried out to analyse the microeukaryotic community associated with gorgonians of the genus *Eunicella*.

Given that gorgonian host DNA accounts for the vast majority of DNA extracted from colonies, the detection of microeukaryotic diversity is challenging as their less abundant sequences are severely disadvantaged by PCR, which favours amplification of dominant matrices, here the ribosomal DNA (rDNA) of the host. To circumvent this problem, we tested a strategy that relies on the use of a blocking primer complementary to the gorgonian rDNA sequence to reduce the proportion of host amplicons. This approach has been reported, for example, in previous studies on coral-associated protists (Clerissi et al., 2018) and krill stomach contents (Vestheim & Jarman, 2008).

Methods

PCR amplification and metabarcoding of 18S rDNA

Using the 18S rDNA gorgonian sequences available in GenBank and one *Eunicella cavolini* sequence determined in the laboratory, we confirmed that the 18SV4 blocking primer of Clerissi et al. (2018) (5'-TCTTGATTAATGAAAACATTCTTGGC-3' modified with a C3 spacer at the 3' end) initially designed for scleractinian corals was also complementary to the octocorallia sequences.

We therefore tested the efficiency of amplification of microeukaryotes on DNA samples obtained from one colony of *E. singularis* and two colonies of *E. cavolini* sampled in Marseille Bay, using the blocking primer 18SV4 in combination with the primer pair 18SV4-F (5'-CCAGCASCYGCGGTAATTCC-3') and 18SV4-R (5'-ACTTTCGTTCTTGATYRA-3') (Stoek et al., 2010) targeting a fragment of approximately 420 base pairs in the V4 variable region of the 18S rRNA gene.

Gorgonian DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen) and PCR reactions were performed according to the conditions of Clerissi et al. (2018), except that different concentration ratios between the blocking primer and the 18SV4 primers (1.5:1, 3:1, 5:1 and 10:1) were tested to optimise the proportion of microeukaryotic amplicons.

Sequencing of the final library of pooled amplicons was performed at the Génome Québec Centre of Expertise and Services (Montréal, Canada) on Illumina MiSeq platform using 2 x 250 bp v2 chemistry and following the manufacturer's guidelines.

Analysis of sequencing data

The FROGS pipeline v4.0 (Escudié et al., 2018) implemented in a Galaxy instance at GenoToul bioinformatics facility (Toulouse, France; https://bioinfo.genotoul.fr/) was used to align reads, remove chimera sequences, define Operational Taxonomic Units (OTUs), and to assign taxonomy based on the Silva 138.1 18S reference database (Quast et al., 2012).

For the phylogenetic analysis, the sequences were edited with ugene (Okonechnikov et al., 2012). The phylogenetic reconstructions were performed on a 376 bp alignment with the Maximum-Likelihood (ML) approach of IQ-TREE 2.1.1 (Nguyen et al., 2015). We used the ModelFinder option (Kalyaanamoorthy et al., 2017), and robustness was evaluated with 1000 ultrafast bootstraps (Hoang et al., 2018). The tree has been visualized with FigTree 1.4.4 (Rambaut, 2006) and rooted at mid-point.

Results

Identification of Symbiodiniaceae sequences in *E. singularis* and *E. cavolini*

Depending on the primer ratio used, the proportion of non-cnidarian 18S sequences reached up to 37.6% and 23.2% of the total sequences for *E. singularis* and *E. cavolini*, respectively. For both species, inhibition of host 18S rDNA gene amplification was most effective with the highest concentration of blocking primer.

In *E. singularis*, we identified 92 OTUs belonging to the family Symbiodiniaceae, in good agreement with the intracolonial diversity of zooxanthellae genotypes previously reported in this host species (Forcioli et al., 2011). Among these OTUs, a single OTU (OTU_7) was highly dominant and contributed up to 45.6% of the total Symbiodiniaceae abundance in the studied colony.

A small number of Symbiodiniaceae OTUs were detected in *E. cavolini* colonies (12 to 13 OTUs depending on the colony). Between the two colonies analysed, the proportion of Symbiodiniaceae sequences varied considerably, accounting for 0.21% to 2,3% of the non-cnidarian sequences when the blocking primer concentration was the highest. However, in both cases OTU_7 was the most abundant, representing up to 99% of all Symbiodiniaceae sequences. The sequence of this OTU_7 (381 bp in length) has been submitted to GenBank under reference SUB14400021.

Phylogenetic analysis

The 18S rDNA OTU_7 shared between *E. cavolini* and *E. singularis* was used for a Blast search in GenBank. In the list of Blast hits, we retained a subset of sequences corresponding to different levels of identity and to different clades of Symbiodiniaceae for phylogenetic reconstruction. The phylogenetic inference on these data produced a tree (Fig. S2.1) that allowed sequences belonging to clades A, B, C and D to be distinguished (most previously defined clades were recovered with more than 90% bootstrap support). According to the tree topology, the putative Symbiodiniaceae species associated with OTU_7 is sister to a clade containing *Symbiodinium microadriaticum*, *S.*

pilosum and other symbiotic *Symbiodinium* species belonging to Clade A (now corresponding to the *Philozoon* genus; Lajeunesse et al., 2022). The closest sequence to OTU_7 is from a dinoflagellate isolated from a South China Sea reef (accession MZ621018; to be released upon publication).

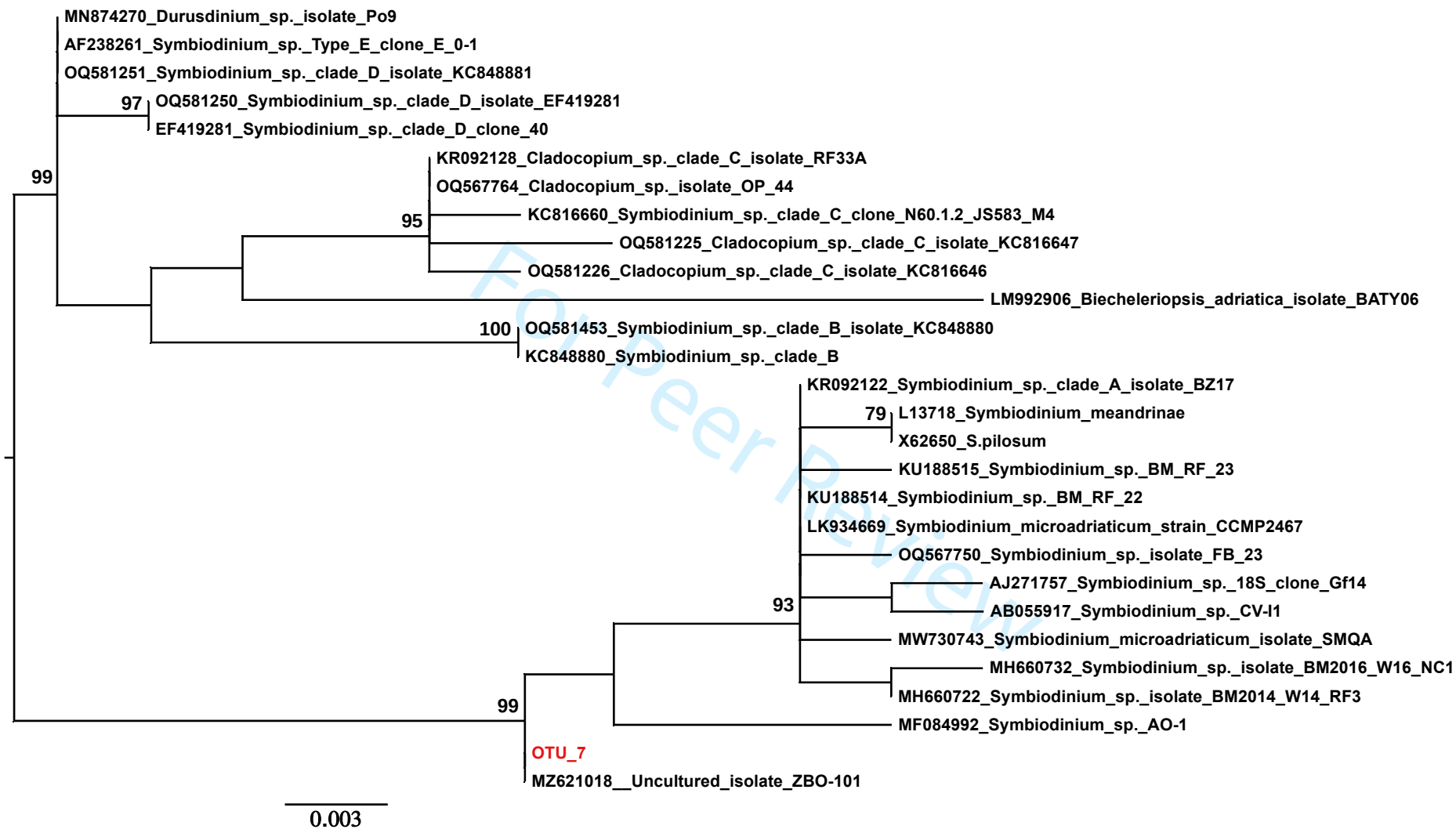


Figure S2.1. Maximum-likelihood phylogeny of Symbiodiniaceae based on the variable V4 region of the 18S rRNA gene, illustrating the relationship of OTU_7 sequence (in red) with Symbiodiniaceae spp. belonging to Clade A. The numbers to the left of the nodes indicate the percentages of bootstraps, for values superior to 75%. The first part of each sequence name corresponds to the accession number in GenBank.

References

Clerissi, C., Brunet, S., Vidal-Dupiol, J., Adjeroud, M., Lepage, P., Guillou, L., ... & Toulza, E. (2018). Protists within corals: the hidden diversity. *Frontiers in Microbiology*, *9*, 2043. https://doi.org/10.3389/fmicb.2018.02043

Escudié, F., Auer, L., Bernard, M., Mariadassou, M., Cauquil, L., Vidal, K., ... & Pascal, G. (2018). FROGS: find, rapidly, OTUs with galaxy solution. *Bioinformatics*, *34*(8), 1287-1294. https://doi.org/10.1093/bioinformatics/btx791

Forcioli, D., Merle, P. L., Caligara, C., Ciosi, M., Muti, C., Francour, P., ... & Allemand, D. (2011). Symbiont diversity is not involved in depth acclimation in the Mediterranean sea whip Eunicella singularis. *Marine Ecology Progress Series*, *439*, 57-71. https://doi.org/10.3354/meps09314

Hoang, D. T., Chernomor, O., Von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, *35*(2), 518–522.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587–589. https://doi.org/10.1038/nmeth.4285

LaJeunesse, T. C., Wiedenmann, J., Casado-Amezúa, P., D'ambra, I., Turnham, K. E., Nitschke, M. R., Oakley, C. A., Goffredo, S., Spano, C. A., & Cubillos, V. M. (2022). Revival of Philozoon Geddes for host-specialized dinoflagellates,'zooxanthellae', in animals from coastal temperate zones of northern and southern hemispheres. *European Journal of Phycology*, *57*(2), 166–180. https://doi.org/10.1080/09670262.2021.1914863

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Molecular Biology and Evolution*, *32*(1), 268–274. https://doi.org/10.1093/molbev/msu300

Okonechnikov, K., Golosova, O., Fursov, M., & Ugene Team. (2012). Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics*, *28*(8), 1166–1167. https://doi.org/10.1093/bioinformatics/bts091

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., ... & Glöckner, F. O. (2012). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic acids research*, *41*(D1), D590-D596. https://doi.org/10.1093/nar/gks1219

Rambaut, A. (2006). *FigTREE v1.4*. University of Edinburgh. http://tree.bio.ed.ac.uk/software/figtree/

Stoeck, T., Bass, D., Nebel, M., Christen, R., Jones, M. D., Breiner, H. W., & Richards, T. A. (2010). Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine

anoxic water. *Molecular ecology*, *19*, 21-31. https://doi.org/10.1111/j.1365-294X.2009.04480.x

Vestheim, H., & Jarman, S. N. (2008). Blocking primers to enhance PCR amplification of rare sequences in mixed samples–a case study on prey DNA in Antarctic krill stomachs. *Frontiers in zoology*, *5*, 1-11. https://doi.org/10.1186/1742-9994-5-12