



# The many faced symbiotic snakelocks anemone (*Anemonia viridis*, Anthozoa): host and symbiont genetic differentiation among colour morphs

Barbara Porro<sup>1,2</sup> · Cédric Mallien<sup>1</sup> · Benjamin C. C. Hume<sup>1,3</sup> · Alexis Pey<sup>4</sup> · Emilie Aubin<sup>1,5</sup> · Richard Christen<sup>1</sup> · Christian R. Voolstra<sup>1,6</sup> · Paola Furla<sup>1,2</sup> · Didier Forcioli<sup>1,2</sup>

Received: 24 December 2018 / Revised: 30 July 2019 / Accepted: 15 August 2019 / Published online: 16 September 2019  
© The Author(s), under exclusive licence to The Genetics Society 2019

## Abstract

How can we explain morphological variations in a holobiont? The genetic determinism of phenotypes is not always obvious and could be circumstantial in complex organisms. In symbiotic cnidarians, it is known that morphology or colour can misrepresent a complex genetic and symbiotic diversity. *Anemonia viridis* is a symbiotic sea anemone from temperate seas. This species displays different colour morphs based on pigment content and lives in a wide geographical range. Here, we investigated whether colour morph differentiation correlated with host genetic diversity or associated symbiotic genetic diversity by using RAD sequencing and symbiotic dinoflagellate typing of 140 sea anemones from the English Channel and the Mediterranean Sea. We did not observe genetic differentiation among colour morphs of *A. viridis* at the animal host or symbiont level, rejecting the hypothesis that *A. viridis* colour morphs correspond to species level differences. Interestingly, we however identified at least four independent animal host genetic lineages in *A. viridis* that differed in their associated symbiont populations. In conclusion, although the functional role of the different morphotypes of *A. viridis* remains to be determined, our approach provides new insights on the existence of cryptic species within *A. viridis*.

## Introduction

The most straightforward way to recognize a species is from its morphological description. However, there are limits to

---

These authors contributed equally: Barbara Porro, Cédric Mallien

---

These authors contributed equally: Paola Furla, Didier Forcioli

---

**Supplementary information** The online version of this article (<https://doi.org/10.1038/s41437-019-0266-3>) contains supplementary material, which is available to authorized users.

- ✉ Barbara Porro  
barbara.porro@hotmail.com
- ✉ Didier Forcioli  
forcioli@unice.fr

<sup>1</sup> UPMC Univ Paris 06, Univ Antilles, Univ Nice Sophia Antipolis, CNRS, Symbiose Marine, Evolution Paris Seine—Institut de Biologie Paris Seine (EPS—IBPS), Sorbonne Universités, 75005 Paris, France

<sup>2</sup> Institute for Research on Cancer and Aging, Nice (IRCAN), Université Côte d'Azur, CNRS, INSERM, 06107 Nice, France

the applicability of this. Indeed, in several cases, different species have no distinguishing phenotypic features, leading to the concept of cryptic species (see for review Struck et al. 2018; Fišer et al. 2018). Many studies have shown that cryptic species are quite common in most of the animal phyla (Pfenninger and Schwenk 2007; Pérez-Ponce de León and Poulin 2016). In the marine realm, the number of described cryptic species has increased in the last years, thanks to new and diversified molecular approaches (see for review Pante Puillandre et al. 2015), but Appeltans et al. (2012) estimated that many species remain to be discovered.

<sup>3</sup> Red Sea Research Center, Division of Biological and Environmental Science and Engineering, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia

<sup>4</sup> THALASSA Marine research & Environmental awareness, 17 rue Gutenberg, 06000 Nice, France

<sup>5</sup> Laboratoire Génome et Développement des Plantes—UMR CNRS/UPVD 5096, Université de Perpignan, 66860 Perpignan, France

<sup>6</sup> Department of Biology, University of Konstanz, 78457 Konstanz, Germany

The forces explaining the existence of cryptic species in some groups are not clear, but could be the result of low standing genetic variation or environmental/developmental constraints (see for review Struck et al. 2018).

On the other hand, morphological polymorphisms could be due to phenotypic plasticity and response to a given environment (Price et al. 2003). Many species transplant experiments have shown that the same genotype could generate different phenotypes depending on the environmental conditions (Merilä and Hendry 2013). Those morphological variations among individuals could even be due to environmentally stabilized allelic polymorphisms, as in the well-known example of the peppered moth *Biston betularia* (Creed et al. 1980), rather than to a loss of interfecundity among morphs.

In cnidarians, morphological variation is often a poor indicator of species divergence. Corals are especially well known to demonstrate extraordinary variation in ecomorphology that may or may not correspond well with the underlying genetic diversity (Veron and Pichon 1976; Flot et al. 2011; Schmidt-Roach et al. 2013). In addition, Schmidt-Roach et al. (2013) showed that depending on the molecular marker used, the ecomorphs could be either grouped together or sub-divided. In sea anemones also, morphological traits are poor predictors of species delimitation (Rodríguez et al. 2014). In the species *Actinia equina* (Douek et al. 2002; Pereira et al. 2014) and *Phymanthus crucifer* (González-Muñoz et al. 2015), phenotypic diversity was not strictly linked to genetic differences. Moreover, in the genus *Anthothoe*, Spano et al. (2018) recently identified the existence of several cryptic species not revealed by the characters used in morphological classification. In addition, many species of cnidarians live in symbiosis with intracellular photosynthetic dinoflagellates belonging to the family Symbiodiniaceae (LaJeunesse et al. 2018) that play an important trophic role (Muscatine et al. 1991) and could also influence the cnidarian phenotype. Indeed, it has been shown that in the coral *Madracis pharensis*, a correlation between the symbiont *Breviolum* sp. (formerly clade B) present in the host cells and the colour morph of the host exists (Frade et al. 2008). The authors suggested that this diversity in colour could be a functional response to light, thus demonstrating the existence of putative strong links between host phenotype and symbiont population.

The sea anemone *Anemonia viridis* (Forskål 1775) is a conspicuous member of benthic communities throughout the Mediterranean Sea and from the Azores to the North Sea. It is among the largest sea anemones in the region and can form extensive aggregations of clonemates that can carpet the benthos. It thus plays an important role in shaping benthic community structure, as a dominant benthic predator in temperate habitats. *A. viridis* displays five colour morphs, characterised by the expression pattern of genes

coding for fluorescent and non-fluorescent proteins (FPs): the three most frequent of these morphs are var. *rustica* (no detectable fluorescence and no pink tentacle apex), var. *smaragdina* (green FPs expression and pink tentacle apices) and var. *rufescens* (green and red-orange FPs expression and pink tentacle apices) (Wiedenmann et al. 1999; Wiedenmann et al. 2000). The taxonomic nature of these morphs has been debated for a long time and the functional role of the fluorescence is not elucidated in *A. viridis* yet. Recently, De Brauwer et al. (2018) suggested to use the fluorescence properties of coral reef fishes as a survey tool for cryptic marine species. Therefore, the hypothesis of the existence of a correlation between genetic differentiation and fluorescence among *A. viridis* morphs could be relevant. In the past, based on allozyme differentiation and presumed different reproduction strategies, some authors already raised the non-fluorescent var. *rustica* to species level, (Bulnheim and Sauer 1984; Sauer 1986; Sauer et al. 1986). Furthermore, Wiedenmann et al. (1999) highlighted a differential bathymetric distribution of the var. *rustica* and var. *smaragdina* in the first 2 meters of depth, raising the question of the adaptive nature of the morphs. In our previous work (Mallien et al. 2017), using exon-primed intron-crossing (EPIC) polymorphism analysis with relatively few EPIC loci ( $n = 5$ ) for a relatively low number of individuals ( $n = 34$ ), we did not detect any clear phylogenetic split among the morphs. However, our conclusions may have lacked the resolution needed to detect incompletely separated lineages. This last scenario will correspond to what De Queiroz (2007) defined as the “grey zone of speciation”, where the level of population interconnectivity is decreasing leading to the build-up of independent genetic pools. This scenario could not be a priori excluded for the *A. viridis* colour morphs, but to evaluate such an ongoing split between emergent morph lineages, it is necessary to use more powerful molecular investigations than the ones we used previously. Consequently, here we have used a RAD sequencing (RADseq) approach in order to generate several tens of thousands of genetic markers in a larger collection of individuals, a sampling scheme fit to uncover population level diversity patterns, species delimitation, and their phylogenetic relationships (Emerson et al. 2010; Pante et al. 2015; Ree and Hipp 2015). Moreover, previous studies investigating the nature of *A. viridis* morphs did not systematically consider *A. viridis* as a holobiont. *A. viridis*' gastrodermal tissue harbours millions of dinoflagellate cells (Muscatine et al. 1998; Suggett et al. 2012; Zamoum and Furla 2012; Ventura et al. 2016) belonging to the family Symbiodiniaceae (LaJeunesse et al. 2018) that live in a close trophic relationship (Davy et al. 1996) and that are vertically transmitted (Schäfer 1984). The Symbiodiniaceae associated to *A. viridis* belong to the temperate clade A (LaJeunesse et al. 2018; or *A' sensu* Savage et al. 2002;

Visram et al. 2006) presenting not only an intra-clade genetic diversity partially structured by host species but also an intra-host genetic diversity (Visram et al. 2006; Forcioli et al. 2011; Casado-Amezúa et al. 2014). Among *A. viridis* morphs, no study provided a clear view of the symbiont genetic diversity distribution that could however be a driver of the morphological differentiation of *A. viridis* morphs.

In the present study, we tested for the existence of genetic divergence among the three most frequent *A. viridis* morphotypes var. *smaragdina*, *rustica* and *rufescens* using high throughput host genotyping by RADseq. In addition, we also tested for correlation between host morph and the genetic diversity of symbiont populations. This diversity was assessed using the rDNA internal transcribed spacer 2 (ITS2) sequence variation, a known standard for Symbiodiniaceae diversity assessment, through a targeted sequencing approach and using two independent methods for the identification of the ITS2 variants: SymPortal (Hume et al. 2019) and an ad hoc pipeline (D&S).

## Materials and methods

### Sampling

We sampled 177 individuals of *Anemonia viridis* belonging to the three most frequent colour morphs (25 var. *rufescens*, 74 var. *rustica* and 78 var. *smaragdina*) sampled from five locations (three in the Mediterranean Sea and two in the English Channel) between 2011 and 2014. To identify the morphotypes, we detected their respective diagnostic fluorescence patterns using a fluorescence torch (Tekna T6, FireDiveGear, The Netherlands). In addition, four individuals of *Paranemonia cinerea* were used as an outgroup. The sampling sites are detailed in Fig. 1. For each specimen, a dozen tentacles were cut, fixed as soon as possible in 70% ethanol and preserved at  $-80^{\circ}\text{C}$  until DNA extraction.

### DNA extraction

All DNA extractions were performed as in Mallien et al. (2017) using a modified “salting out” protocol (Miller et al. 1988). However, to test for Symbiodiniaceae DNA influence on RADseq efficiency, we isolated symbiont-free epidermal tissue of a tentacle from a subset of 63 samples following Richier et al. (2003) and performed DNA extraction on this tissue fraction following the same procedure. This allowed us to compare the yield in total RADseq loci among symbiont-rich and symbiont-poor samples to infer the impact of Symbiodiniaceae DNA contamination. Extracted DNA concentrations were determined using Quant-it Picogreen kit (Invitrogen).

## Host RADseq and genetic analyses

### Library preparation

For this study, libraries were prepared following Etter et al. (2011) with these modifications: (i) DNA extracts were classified depending on DNA degradation (measured by electrophoresis of the total DNA) and symbiont contamination level (measured by PCR using Symbiodiniaceae cp23S specific primers from Santos et al. (2003)) and split into nine libraries containing between 32 and 44 individuals; (ii) libraries were constructed using 0.6–1 µg of DNA per individual digested by PstI-HF enzyme (New England Biolabs), and ligated to sample specific RADseq adaptors; (iii) each pooled library was sheared for 60 s using the S220 Focused-ultrasonicator (Covaris); (iv) after ligation of the second RADseq adaptor, a final amplification of 16 cycles was carried out for each library. Sequencing was performed on an Illumina HiSeq 2000 (100-bp, single read format) at MGX (Montpellier, France).

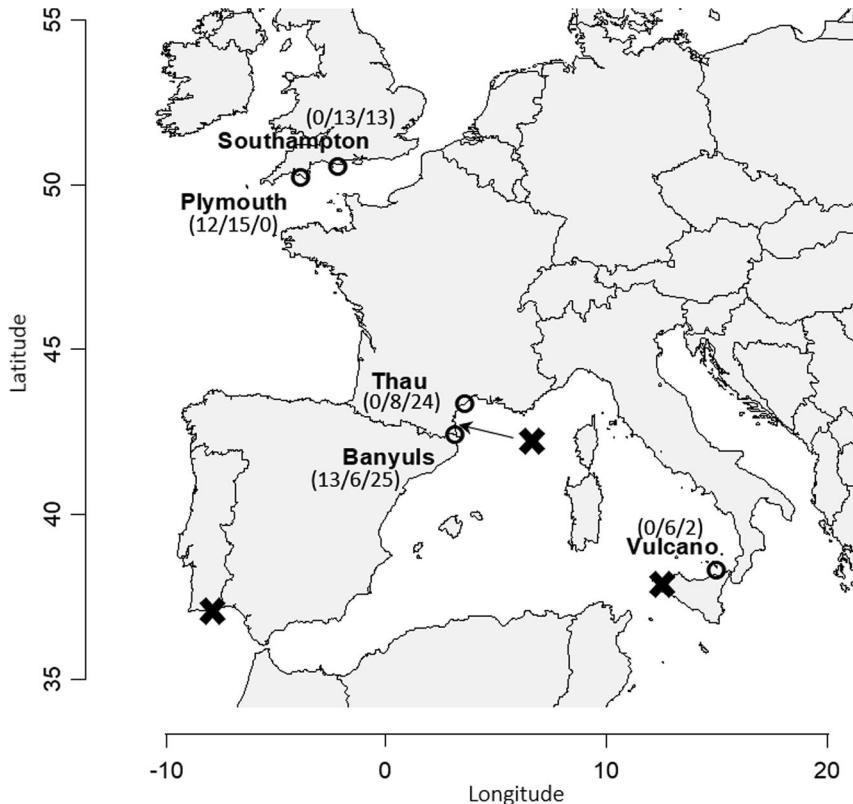
### De novo RADseq loci identification

Raw sequence reads were demultiplexed, filtered and clustered with iPyRAD v0.3.41 (Eaton 2014) in the 181 initial individuals. Briefly, restriction sites and adaptors were trimmed out and raw reads were regrouped under individual names using their sample specific barcode. Base mismatches were not allowed in those barcodes. The clustering threshold was set to 90%. A minimum of five individuals were needed to validate a locus and the maximum proportion of shared polymorphic sites within a locus was set to 50%. Based on locus count in all individuals, some individuals turned out almost unamplified and were consequently discarded from the subsequent analyses.

A second run of iPyRAD was then performed with the same parameters, excluding individuals with low locus count (less than 4000) and using a clustering threshold of 90%. This led us to a dataset of 137 *A. viridis* individuals (25 from var. *rufescens*, 48 from var. *rustica* and 64 from var. *smaragdina*) and three *P. cinerea* individuals (see Fig. 1 for their geographical origin). The iPyRAD loci were further filtered using VCFTools v0.1.15 (Danecek et al. 2011) to keep only bi-allelic loci present in at least 70% of all individuals. This allowed us to define a first total dataset, N140, of 140 individuals (137 *A. viridis* and 3 *P. cinerea* anemones), containing a total of 45,519 SNPs.

We verified the distribution of the missing data within this 140 host individuals’ dataset by computing the proportion of shared loci between all the pairs of individuals in R using the package RADami (Hipp et al. 2014, <https://github.com/cran/RADami>). This analysis led to the selection of a subset of 85 individuals (N85 dataset) with an

**Fig. 1** Sampling sites of the N140 dataset in the Mediterranean Sea (Vulcano: 38°21'14"N 14°59'22"E; Thau lagoon: 43°23'32.82"N 3°36'1.43"E; Banyuls: 42°28'50"N 03°07'50"E) and in the English Channel (Plymouth: 50°16'40"N 03°54'49"W; Southampton: 50°36'38.16"N 02°8'3.252"W). The brackets indicate the number of sampled individuals per site per morph (var. *rufescens*/var. *rustica*/var. *smaragdina*, respectively). The black crosses mark the three sampling sites for the three *P. cinerea* individuals



average of 5% of missing data. The distribution of missing data was represented with a heatmap using R (Fig. S1).

### Host phylogeny

An individual-based tree was built for the N140 dataset using RAxML v8.2.4 (Stamatakis 2014), from the phylip output of iPyRAD, using a GTR + $\Gamma$  nucleotide substitution model and bootstrap support was estimated from 100 replicates. Mega v6.06 (Tamura et al. 2013) was used to visualize the output of RAxML.

### Clonality test

Considering the propensity for clonal reproduction of *A. viridis* (Wiedenmann et al. 1999; Wiedenmann et al. 2000; Mallien et al. 2017), we determined the putative presence of clonal lineages within host samples. Since all the individual RADseq genotypes we obtained were different, the classical way of identifying clonal genets through the detection of repeated multilocus genotypes was not applicable. Instead, using the R package RClone (Bailleul et al. 2016), we computed between individuals pairwise genetic distances keeping only SNPs present in all individuals in the N85 dataset (396 SNPs). These distances were computed at the level where recombination would be expected to occur that is within each of the secondary hypothesis species that we

obtained after a first run of SVDquartets on the N140 dataset (see below). The observed distributions of pairwise genetic distances thus obtained were compared each to a simulated distribution of expected genetic distances under the assumption of purely sexual reproduction (with or without inbreeding) computed with RClone. This allowed us to define for each SVD clade a threshold pairwise distance value that we used to affect individuals to clonal lineages (MLL) using the function mlg.filter from the R package poppr (Kamvar et al. 2014). Keeping only one ramet per genet, this led us to the definition of a 109 individuals dataset (N109). To test for an effect of clonality on species delimitation, SVDquartets, snmf and BFD\* analyses were performed as well on this non-clonal dataset (see below).

### Host species delimitation

To construct species hypotheses within *A. viridis*, we first performed a taxonomic partitioning from the N140 and the N109 datasets with SVDquartets software from Chifman and Kubatko (2014) implemented in PAUP\* v4.0a152 (Swofford 1998) using default parameters. This non-Bayesian approach infers the relationship among quartets of taxa under a coalescent model and uses this information to build the species tree. *P. cinerea* individuals were identified as outgroups. We performed a 500 bootstrap replicate

analysis for each dataset. We considered the strongest monophyletic clades identifying at least the expected split between Mediterranean and English Channel individuals as possible species hypothesis to test. To validate these hypotheses, we also performed a quantitative genetic clustering by computing individual ancestry coefficients in the N109 dataset using the snmf function from the LEA R package (Frichot and François 2015). This function allows for the identification of population structure in a way similar to STRUCTURE (Pritchard et al. 2000) and ADMIXTURE (Alexander et al. 2009) analysis, but in a more efficient way (Frichot et al. 2014). First, the optimal number of clusters (identified as ancestral populations, K) from which the individuals in the dataset could be traced back was determined using an entropy criterion. The admixture coefficient among these K clusters was then estimated for each individual, using sparse non-negative matrix factorization (snmf) (Frichot and François 2015). The result was visualized using a STRUCTURE-like representation showing, for each individual, the contribution of each cluster to its multilocus genotype.

We then compared our primary species hypothesis (the three morphs as species, irrespective of the geographical origins of the individuals) to the secondary species hypothesis obtained with SVDquartets and snmf. We also compared the SVD clades/snmf clusters to another, more geographically based, primary species hypothesis in which the two non *rufescens* Mediterranean SVD clades/snmf clusters were lumped together. These species delimitation hypotheses were tested using BFD\* (Leaché et al. 2014). In order to manage computing times, these analyses were performed using only a subset of non-clonal individuals representing most of the genetic diversity from the N85 dataset (see Table S1) and a set of 1000 filtered bi-allelic SNPs (no missing data, minor allele frequency 0.1). The BFD\* analyses were performed in BEAST2 (Bouckaert et al. 2014) with 48 path sampler sets of 100,000 MCMC repetitions, a 10% burnin to sample in 1,000,000 MCMC iterations of SNAPP (Bryant et al. 2012). The number of SNAPP iterations was chosen to ensure proper convergence of the species tree inference. The parameters for the BFD\* analysis were chosen following the recommendations of the 2018 version of the BFD\* tutorial (<https://github.com/BEAST2-Dev/beast-docs/releases/download/v1.0/BFD-tutorial-2018.zip>), that is mutation rates priors set to 1, coalescent rate sampled from the data and a fixed prior value of the speciation rate Lambda estimated as stated in the document table.

The Wright's F-statistics among and within populations, along with basic diversity indices, were computed using the R package hierfstat (Goudet 2005, version 0.04-22. <https://CRAN.R-project.org/package=hierfstat>). These statistics were calculated on the N109 dataset for the SVD clades/snmf cluster.

## Temperate clade A Symbiodiniaceae ITS2 and genetic analyses

### Amplification and sequencing of the rDNA ITS2

To infer the genetic diversity of the temperate clade A Symbiodiniaceae populations, the nuclear ribosomal DNA ITS2 region was amplified by PCR using the primers from Stat et al. (2009): the forward itsD (5'-GTGAATTGCAGA ACTCCGTG-3') and the reverse its2rev2 (5'-CCTCCG CTTACTTATATGCTT-3').

PCR amplifications of the ITS2 were performed by Access Array (FLUIDIGM) PCR and the amplicons were sequenced by MiSeq Illumina Technology (2 × 250 bp paired-end) at the Brain and Spine Institute (ICM, Paris, France).

### Genotyping of the symbiont populations

The aim here was to build a catalogue of all Symbiodiniaceae ITS2 sequences out of the sequencing background noise for each host sea anemone. To do so, we tested two different ITS2 genotyping pipelines: one based on the depth (i.e. number of reads) and the sharing of the variants among individuals (identified as D&S pipeline in the rest of the article), and another based on minimum entropy decomposition (MED) algorithm (Eren et al. 2015) and referred to as the SymPortal pipeline in the remainder of this article (Hume et al. 2019; Hume 2019).

The D&S pipeline consisted in assembling forward and reverse reads using OBITools scripts (Boyer et al. 2016). We then aligned the assembled reads to a temperate A reference ITS2 sequence obtained in the laboratory, using BWA MEM v0.7.12-5 (Li 2013). To separate the real ITS2 alleles from sequencing errors in the reads, which mapped to the reference, we used a three-step protocol (Table S2). For the first step, we clustered the reads across all samples at a threshold of 80% of identity among the reads, and we conserved only clusters that contained more than 5% of the total number of reads to exclude the rarest sequences, most likely sequencing errors. For the second step, we pooled all the sequences that passed the first step and applied a second clustering identity threshold of 100% to identify unique sequences. Finally, a last step was necessary to determine which unique sequence was a valid ITS2 variant and not background sequencing noise. For this, we considered (i) the number of hosts harbouring a given unique sequence (with a minimum of two individuals), and (ii) the median number of reads per host for this unique sequence, to set up a validation threshold as follows:

$$Q_{50}(\text{reads})_i \times N_i \geq 50$$

where  $Q_{50}(\text{reads})_i$  is the median number of reads for the putative ITS2 allele  $i$  and  $N_i$  the number of hosts in which it was found. This dataset was identified as the T50 dataset.

For the second pipeline, the MED algorithm implemented in SymPortal (Hume et al. 2019; Hume 2019) was used. In the MED algorithm (Eren et al. 2015), the most variable positions in the sequences (those with the highest Shannon's entropy value forming an entropy peak) were used to split the dataset and decompose it into more homogeneous groups within the symbiont population of each animal host. At each step of decomposition, if there are still entropy peaks in the entropy profile of the dataset, the dataset is sub-split depending on the identified variable position, otherwise the algorithm stops. The number of homogeneous groups (or MED nodes) obtained thus corresponded to the biologically informative ITS2 sequence variants within each animal host (see Hume et al. 2019 for a more thorough description of SymPortal). We used SymPortal up to this step to obtain a dataset comparable with the D&S pipeline previously described. This dataset was identified as the SP dataset.

## ITS2 diversity

To visualise the diversity of the ITS2 variants, we built an ITS2 median joining network (Bandelt et al. 1995, 1999) for each pipeline, using SplitsTree (Huson and Bryant 2006). PopART v1.7 (Leigh and Bryant 2015) was used to calculate the nucleotide diversity ( $\pi$ ) among the sequences obtained by the both pipelines.

## Measure of the differentiation among symbiont populations

Each sea anemone hosts up to 1 million *Symbiodinium* cells per cm<sup>2</sup> of tissue (Suggett et al. 2012), and therefore constitutes a pooled sample of symbionts. Because we could not estimate properly the number of symbionts per anemone in our samples, we considered the presence (1) or the absence (0) of each ITS2 variants rather than its number of reads in each individual. Consequently, the genotype of each symbiont population was coded as for binary haploids.

We tested the differences in the composition of symbiont populations among sea anemones depending on (i) the morph of the host, or (ii) the host genetic lineages by PERMANOVA using Jaccard's distance (based on presence/absence of variants) over 9999 permutations with the R package vegan (Oksanen et al. 2016). When these factors had an effect on the distribution of the genetic diversity of symbionts, we then computed pairwise PERMANOVAs between each level of the given factor with an FDR correction to avoid false positives.

## Results

### Host species delimitation

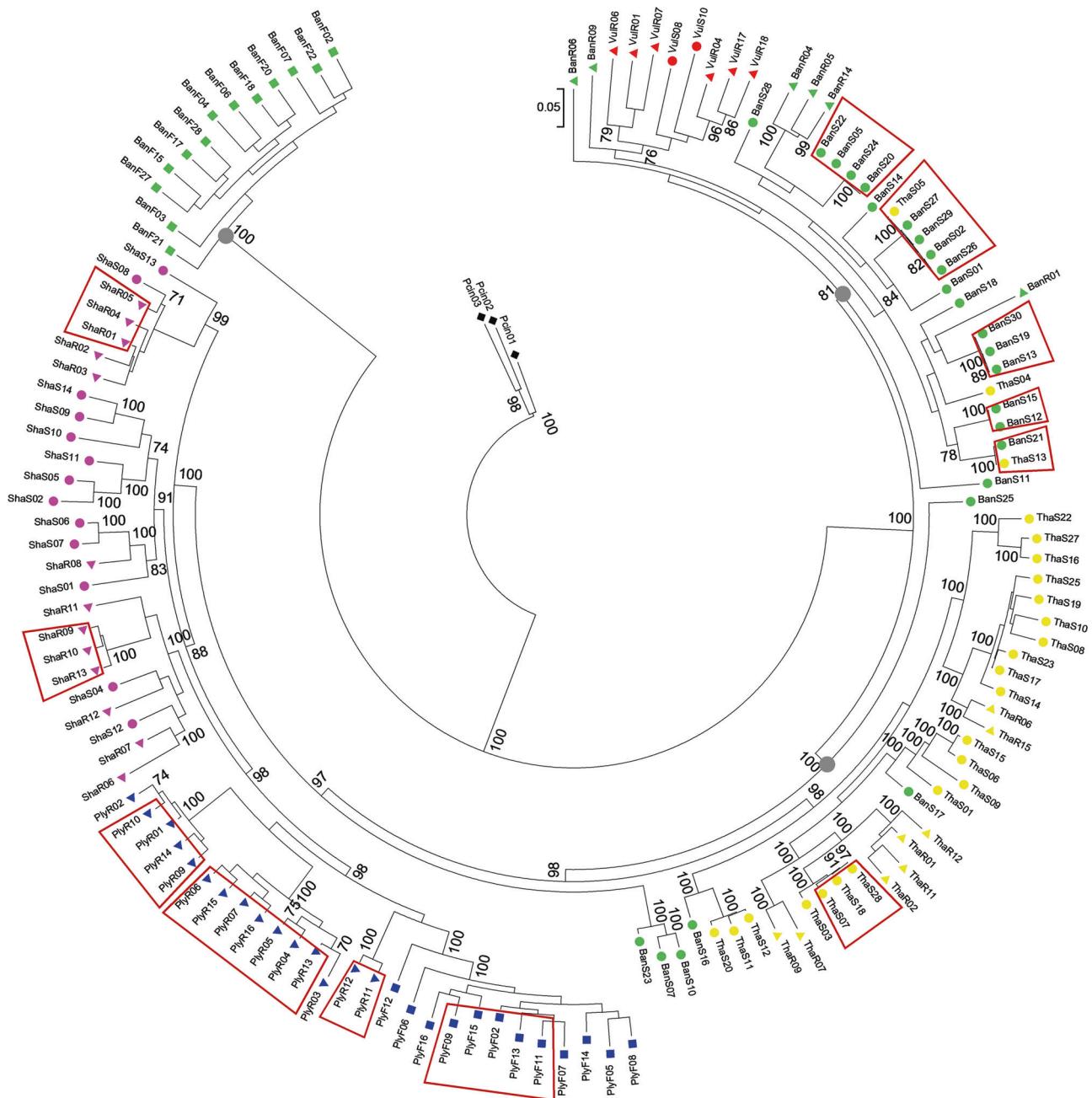
The analysis of sequences obtained by RADseq, as described in section "Materials and methods", allowed the identification of an average of  $33\,407 \pm 1162$  SE RADseq loci (with a minimum number of SNPs of 4214 up to 45,447 for the maximal value) per animal host individual.

As there is still no reference genome either for *Anemonia viridis* or for the Symbiodiniaceae temperate clade A, there was no efficient way to filter symbiont contamination from the obtained RADseq loci. However, the symbiont-poor samples obtained from the epidermal DNA extracts displayed significantly more RADseq loci ( $N = 65$ ,  $57,696.26 \pm 25,693.28$  loci) than the symbiont-rich samples obtained from both the epidermal and the gastrodermal DNA extracts ( $N = 116$ ,  $29,720.77 \pm 30,948.38$  loci) (Student's  $t$  test  $p$ -value =  $9.58 \cdot 10^{-10}$ ), even if they contained less DNA per extract ( $N = 65$ ,  $56.52 \pm 28.65$  ng/ $\mu$ L vs  $N = 116$ ,  $122.30 \pm 56.62$  ng/ $\mu$ L, Student's  $t$  test  $p$ -value =  $5.42 \cdot 10^{-20}$ ).

After the different filtering steps, these RAD loci provided 45,519 bi-allelic SNPs shared by at least 70% of the 140 selected individuals. The missing data were not structured by host populations but depended on the quality of DNA extracts (Fig. S1). On the full dataset, we computed a global  $F_{st}$  and a  $F_{is}$  of 0.38 and 0.37, respectively.

To measure the actual genetic differentiation among host individuals, we computed a RAxML individual tree on the N140 dataset (Fig. 2). The outgroup *P. cinerea*, which was well differentiated from *A. viridis*, was used to root the tree. We identified three monophyletic clades in *A. viridis* (grey circles on the Fig. 2): the first clear split separated the sea anemones var. *rufescens* from Banyuls from the other individuals, then a Mediterranean clade with sea anemones mostly from Banyuls and Vulcano was separated from a clade with the English Channel and Thau sea anemones (each with bootstrap values higher than 80%).

Some of the individuals did cluster very closely in this RAxML tree. We therefore tested for the occurrence of clonal lineages in the N85 dataset that was the most relevant to detect extremely close RAD genotypes, as it contained very few missing data. RClone (Bailleul et al. 2016) simulations allowed us to identify individuals that, although genetically different, were too genetically close to be issued from sexual recombination (Fig. S2). Several clonal lineages (MLLs) were thus identified: five (of respectively 2, 2, 3, 4 and 5 individuals) were found among non var. *rufescens* individuals from the Banyuls sampling site, one of three individuals in Thau, two of three individuals each in Southampton and four in Plymouth (of 2, 4, 5 and 7 individuals). The var. *rufescens* individuals were very likely the



**Fig. 2** Animal host RAD differentiation at the individual level. Maximum likelihood phylogeny of the N140 dataset using RAxML with GTR + $\Gamma$  model of evolution and 100 bootstrap replicates. For each individual, the origin is coded by colour (red: Vulcano, orange: Thau, green: Banyuls, blue: Plymouth, pink: Southampton and black: *P.*

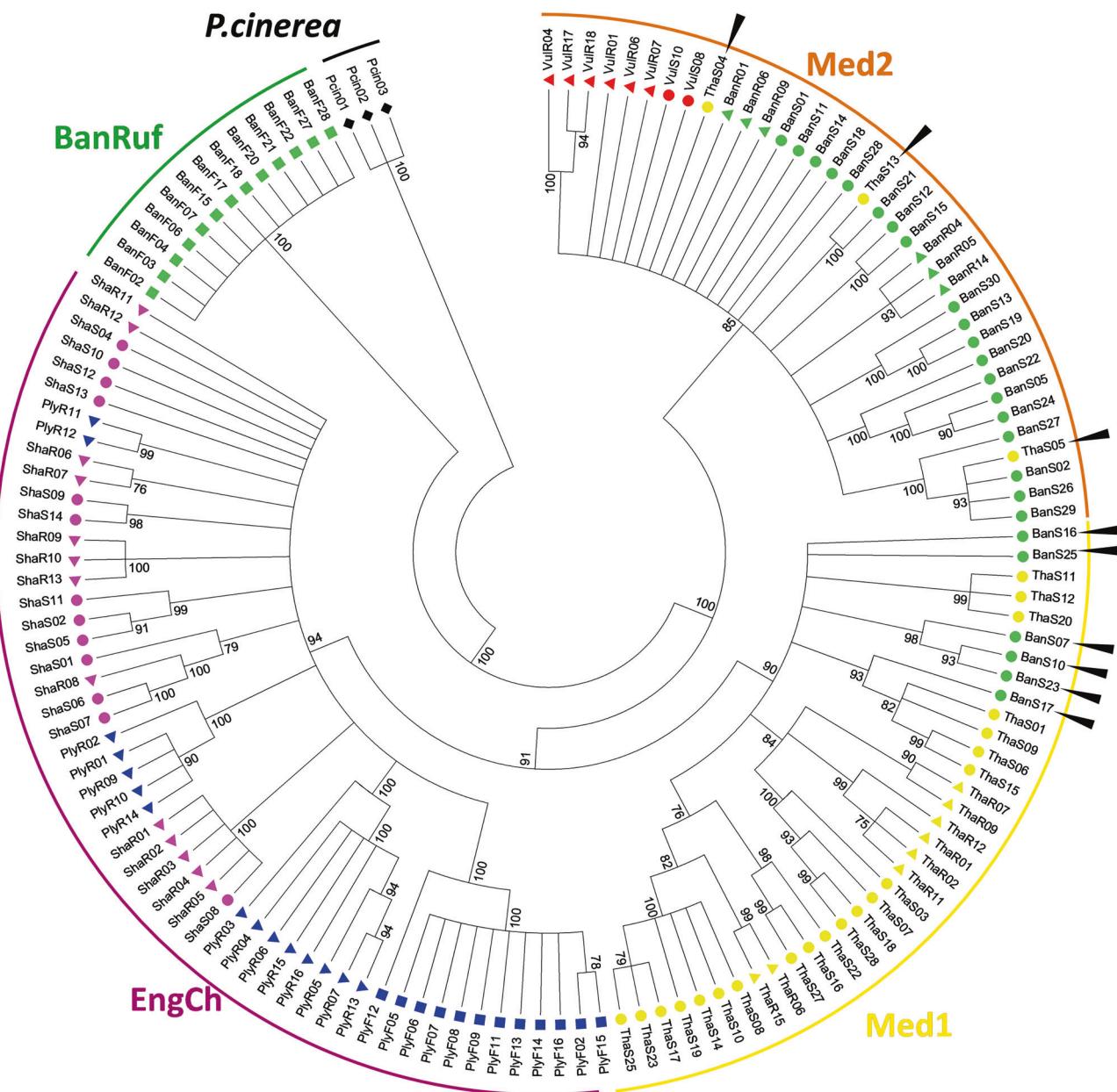
*cinerea* outgroup) and its morph by symbol (square: *A. viridis* var. *rufescens*, triangle: *A. viridis* var. *rustica*, and circle: *A. viridis* var. *smaragdina*). The outgroup *P. cinerea* is marked by a diamond symbol. The individuals falling in the same MLL are boxed in red

products of sexual reproduction (Fig. S2). Keeping only one individual per clonal genetic lineage (i.e. one ramet per genet), we thus obtained the N109 dataset on which we checked the species hypotheses formulated on the N140 dataset.

To formulate secondary species hypotheses from these data, we used SVDquartets to produce a coalescent taxonomic partitioning from the N140 dataset (Fig. 3) and N109

dataset (Fig. S3) and snmf clustering on the N109 dataset (Fig. 4).

For the SVDquartets analyses, as in the RAxML tree, the *P. cinerea* individuals were separated from *A. viridis* and formed an independent clade. Then the first split among *A. viridis* separated with high support the var. *rufescens* anemones from Banyuls and all the other anemones, forming the BanRuf clade. These individuals from Banyuls were in



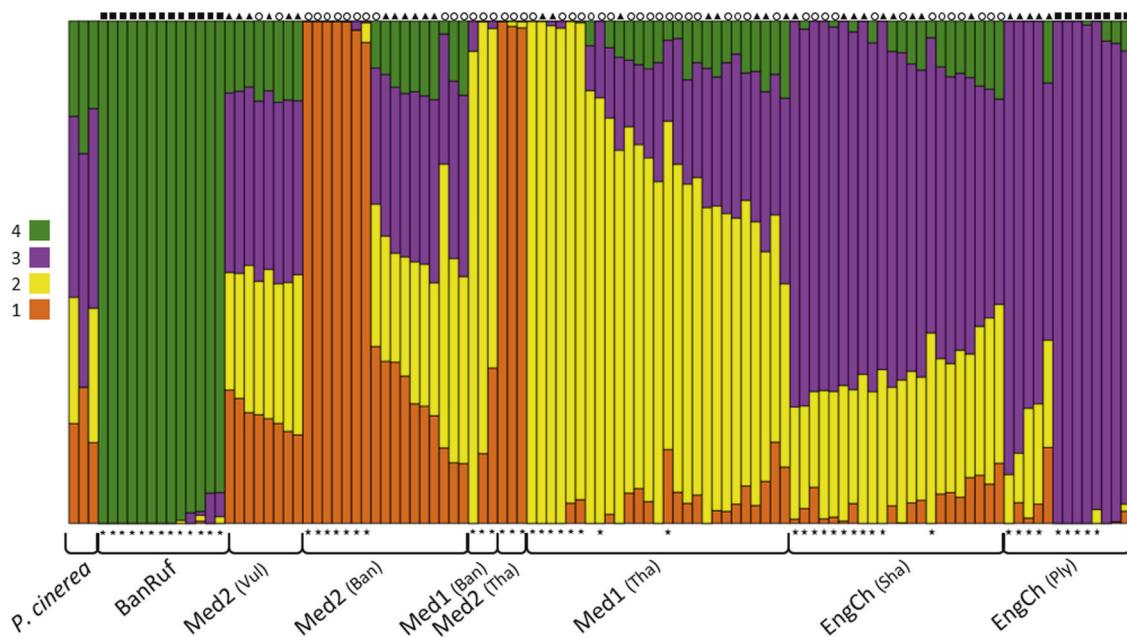
**Fig. 3** Non-Bayesian taxonomic partitioning among *A. viridis* morphs (SVDquartets analysis on the N140 dataset) at the individual level. Support values at the nodes are shown only if equal or superior to 70%. The coloured bars correspond to the five identified clades (i) *P.*

*cinerea* (black), (ii) BanRuf (green), (iii) English Channel (EngCh, purple), (iv) Med1 (yellow) and (v) Med2 (orange). The black arrowheads point to the sea anemones from Banyuls attributed to the Med1 clade and to the Thau sea anemones attributed to the Med2 clade

fact as divergent as the outgroup (100% of the quartets including var. *rufescens* anemones from Banyuls supported this split). In addition, two clades were identified, one grouping the other Banyuls individuals and the Vulcano ones, that we named the Med2 clade, and the second grouping individuals from the English Channel and Thau. This last clade was further split in two by SVDquartets separating English Channel populations (EngCh clade) from Thau (Med1 clade). The BanRuf, Med1 and Med2 SVD clades can be found in sympatry: the three have been sampled in Banyuls (six var. *smaragdina* Banyuls

individuals belonged to the Med1 clade), and the Med1 and Med2 clades have been sampled in Thau (three var. *smaragdina* Thau individuals were attributed to the Med2 clade). These few rare individuals that did not belong to the locally more frequent clade are identified by arrowheads in Fig. 3. These SVD clades were retained in the analysis of the MLL pruned N109 dataset (Fig. S3).

To validate these SVD clades as secondary species hypotheses, we also performed a quantitative genetic clustering by snmf (Frichot and François 2015). We determined  $K = 4$  as the optimal number of genetic clusters in this



**Fig. 4** snmf clustering on the N109 dataset: the probabilities of attribution to the most probable  $K = 4$  clusters or ancestral populations are displayed in a bar plot for each individual. The morph of *A. viridis* sea anemones is coded by the shape above the diagram (triangle: var. *rustica*, circle: var. *smaragdina*, and square: var. *rufescens*). The

individuals kept in the N85 dataset are identified by black stars below the diagram. The sampling sites are detailed in brackets (Vul: Vulcano, Ban: Banyuls, Tha: Thau, Sha: Southampton, and Ply: Plymouth). The four clusters are identified by the colours of the corresponding SVD clades

dataset, based on the minimum value of cross entropy (Fig. S4). For the individuals with less missing data (marked with stars in Fig. 4), the snmf clustering was in complete accordance with the SVD clades. On the other hand, most of the individuals with too much missing data had similar probabilities of attribution to all four clusters (Fig. 4), showing the impact of low data quality on this analysis.

In individuals with high quality data, snmf clustering can be used to measure introgression (Frichot and François 2015). The a priori most likely introgressed individuals, i.e. the var. *smaragdina* from Thau lagoon belonging to the Med2 clade/cluster and the ones from Banyuls belonging to the Med1 clade/cluster (identified by arrowheads in Fig. 3 and identified as Med2(Tha) and Med1(Ban), respectively, on the Fig. 4), had coancestry patterns similar to those of anemones from the same cluster. However, whereas no admixture was detected in the Med2(Tha) individuals, the Med1(Ban) individuals were weakly introgressed by Med2 genetic background (Fig. 4). Likewise, the BanRuf clade corresponded to a homogenous cluster different from the other clusters, and notably from the others found in Banyuls.

The SVD clades/snmf clusters are the best secondary species hypothesis as shown by the Bayes factor comparisons following BFD\* computations (Table S3). Splitting the individuals according to their SVD clades/snmf clusters is decisively better than splitting them either by their

morphs or by their sampling region. Hence, we identified five clear independent genetic lineages in the dataset: (i) the outgroup *P. cinerea*, and among *A. viridis* samples (ii) the var. *rufescens* individuals from Banyuls (BanRuf), (iii) English Channel populations (EngCh) and two main Mediterranean groups (iv) one with mostly anemones from Thau lagoon (Med1), and (v) one with mostly sea anemones from Banyuls and Vulcano (Med2). These lineages harboured the same level of diversity and all displayed a marked deficit in heterozygosity (Table S4).

### Influence of the symbiont on morphs differentiation

The analysis of the ITS2 sequences, following the D&S pipeline, allowed the identification of 92 ITS2 sequence variants (T50 dataset) from the populations of *A. viridis* symbionts (Fig. S5.A, Table S2). Most of the sampled populations harboured more than half of this allelic diversity: between 44 and 91 alleles were present within each host genetic lineage (Table S4, T50 dataset). Similarly, each anemone harboured on average  $32.4 \pm 1.6$  ITS2 variants. The most abundant variant was harboured by 95% of the animal hosts (Fig. S5.A). There was a low divergence among the ITS2 sequences, 90% of them were one substitution away from the most frequent one and the nucleotide diversity was  $\pi = 0.0117$ . *P. cinerea* harboured less variants than *A. viridis* (12 alleles from three individuals,

**Table 1** PERMANOVA analysis of the differentiation and partitioning of the temperate clade A T50-ITS2 diversity in *A. viridis*

	<i>df</i>	Sums of squares	<i>F</i>	<i>R</i> <sup>2</sup>	<i>P</i> (> <i>F</i> )
Among morphs	2	0.779	1.2732	0.02325	0.0764
Residual	107	29.974		0.89478	
Total	109	33.499		1	
Among host genetic lineage	3	2.329	2.6416	<b>0.06952</b>	<b>0.0001</b>
Among morphs within host genetic lineage	4	1.196	1.0175	<b>0.05117</b>	<b>0.0031</b>
Residual	102	29.974		0.89478	
Total	109	33.499		1	

The values in bold are statistically significant with *p*-value < 0.05

Table S4) but all the sequence variants found in *P. cinerea* were also found in *A. viridis* (Fig. S5.A).

There was no statistically significant differentiation of the symbiont populations among the three morphs (Table 1). As host independent genetic lineages were identified (Fig. 3 and Table S3), we tested the differentiation in symbiont populations and the ITS2 composition of these populations among genetic lineages and among morphs nested within these lineages. These analyses revealed an influence of the genetic lineage on the symbiont populations (Table 1). The pairwise comparisons highlighted a difference among EngCh and Mediterranean lineages, and also between Med1 and Med2 (Table 2). A weak difference between BanRuf and Med1 was detected (Table 2). Within these lineages, we detected a differentiation among morphs in symbiont content (Table 1). However, the morphs were not systematically differentiated within all lineages but only within the Med2 lineage (PERMANOVA comparisons Table S5).

To validate this image of symbiont differentiation, we also used the MED algorithm implemented in SymPortal to identify the ITS2 sequence variants by an alternative method (SP dataset). We obtained 93 ITS2 variants in 106 animal hosts. Thirty nine of these variants were common with the T50 dataset (grey circles on Fig. S5). The different populations harboured 4–23 variants out of the 93 identified (Table S4, SP dataset), with an average of  $4.17 \pm 0.2$  variants per sea anemone. The most frequent variant was harboured by 97% of the anemones and the ITS2 sequences were less divergent than the sequences obtained with D&S pipeline ( $\pi = 0.0019$ ). *P. cinerea* was also less diverse than *A. viridis* and again did not harbour any unique ITS2 variant.

In agreement with the SP dataset, the var. *smaragdina* and *rustica* shared the same symbiont ITS2 diversity (Table S7). By contrast, the var. *rufescens* individuals were differentiated for their symbiont content (Table S7), which led

**Table 2** PERMANOVA analysis of the pairwise differentiation and partitioning of the temperate clade A T50-ITS2 diversity among *A. viridis* genetic lineages

	<i>F</i>	<i>R</i> <sup>2</sup>	Adjusted <i>p</i> -value
BanRuf vs EngCh	2.738076	0.04742236	<b>0.003</b>
BanRuf vs Med1	1.844418	0.06870769	<b>0.0156</b>
BanRuf vs Med2	1.013963	0.03071315	0.439
EngCh vs Med1	2.166855	0.02844879	<b>0.0003</b>
EngCh vs Med2	4.364921	0.0511325	<b>0.0003</b>
Med1 vs Med2	2.225183	0.04180696	<b>0.0066</b>

The values in bold are statistically significant with *p*-value < 0.05. The *p*-value was adjusted with an FDR correction

to an overall statistically significant differentiation among morphs (Table S6). It should however be noted that all the var. *rufescens* anemones did not share a common symbiont pool, as the English Channel var. *rufescens* individuals were more differentiated from Mediterranean var. *rufescens* individuals than from individuals belonging to the other morphs (Table S8). Indeed, the composition of symbiont populations was correlated to the host genetic lineages independently of the morphs (Table S6). The pairwise analysis of the SP-ITS2 variants among *A. viridis* lineages showed that the EngCh lineage was different from the Mediterranean ones and that Med1 and Med2 were different from each other (Table S9). No difference between BanRuf and Med1 lineages has been detected with the SP-ITS2 variants. Moreover, no effect of the morphs was detected within lineages with the SP dataset (Table S6).

## Discussion

### NGS data validation

#### Symbiont contamination among host RADseq loci

One of the major concerns when working on symbiotic organism is separating the effects due to each of the partners. This is all the more true when using reduced representation sequencing to measure the genetic diversities of the host and symbionts separately without reference genomes. However, in symbiotic hexacorals, as the sequencing depth is usually tailored for the host, it results in an under sampling of the bigger symbiont genome, and therefore, an even modest filtering on missing data yields a very low Symbiodiniaceae contamination in the final set of loci (Bellis et al. 2016; Titus and Daly 2018). Moreover, the size of *A. viridis* tentacles allowed us to efficiently dissect the symbiont-free epidermis from the symbiont rigged gastrodermal cell layer in fresh or properly preserved samples. In the epidermis fraction, Symbiodiniaceae contamination at

the protein level was estimated at a mere 3% (Richier et al. 2003). This probably explains the fact that the mean number of RAD loci per individual we obtained after filtering on missing data was significantly lower in symbiont-rich DNA extracts (from whole tentacles), although they contained more DNA than the symbiont-poor DNA extracts (from epidermis only). The lower yield of the symbiont-rich extracts probably reflects the contamination dispersion effect as described by Titus and Daly (2018).

Moreover, in the subsequent analysis, symbiont-rich and symbiont-poor samples did finally belong to the same lineages. In addition, the BanRuf lineage that was the most divergent on host RADseq (Fig. 2) did not contain a divergent ITS2 symbiont population (Tables 2 and S9). Although we cannot exclude that a few of the loci among the tens of thousands per individual we obtained could be Symbiodiniaceae loci, we are therefore quite confident that the RAD genotypes reflect the host's genetic diversity and not the symbiont's.

### Pipelines for the identification of symbiont ITS2 sequence variants

To detect intra-cladal/generic diversity, we used sequencing of ITS2 amplicons, now a standard method for genotyping Symbiodiniaceae populations (Quigley et al. 2014; Arif et al. 2014; Batovska et al. 2017; Bonthond et al. 2018). An operational taxonomic units (OTU) approach is commonly used to filter and quality check the huge quantity of raw data obtained by NGS. OTUs are identified by a clustering approach of the variant sequences, at for example a 97% similarity threshold (Arif et al. 2014). However, such a 97% cut-off would miss most of the intra-cladal/generic Symbiodiniaceae diversity. To overcome that issue, we used two different pipelines generating catalogues of ITS2 sequence variants: one based on sequencing depth and shared occurrence of variants among individuals, D&S, that we developed (generating T50 dataset) and the other based on the MED algorithm applied within animal hosts, implemented in SymPortal (Hume et al. 2019) (generating the SP dataset). Through the D&S and SP pipelines, we identified 92 and 93 ITS2 sequence variants, most of which were identical between both pipelines. In fact, the two datasets differed mainly in two ways: they did not detect the same rare sequence variants (Fig. S5) and the SymPortal pipeline conserved less sequences per animal host (Table S4). These differences are due to the different rationale of the two pipelines: to identify ITS2 sequence variants, one works over the whole dataset (D&S pipeline), whereas the other works within individuals (SymPortal pipeline). Cunning et al. (2017), using an OTU-based approach, did also obtain different images of the same raw diversity whether they applied a 97% identity threshold over their whole dataset or within individuals.

Obviously, a large part, or even the majority, of this detected genetic diversity probably corresponds to multi-copy intragenomic ITS2 variants, considering the number of sequence variants we obtained within each host individual (Sampayo et al. 2009; Smith et al. 2017). However, these intragenomic variants can be used to better define the Symbiodiniaceae lineages harboured by the hosts, this is even the very goal of the SymPortal pipeline (Hume et al. 2019). But, in order to be able to use this information, SymPortal relies on the assumption that a majority of hosts harbours a single dominant genetic strain per symbiont genus. Considering that in the Mediterranean Sea, the previously observed genetic diversity corresponds to within clade/genera diversity (Forcioli et al. 2011) it seems that this basic assumption may not be respected. The dearth of Symbiodiniaceae temperate clade A ITS2 sequences in SymPortal database so far precludes a formal test of this hypothesis. Further analyses could however be done using single copy markers (as psbA sequences LaJeunesse and Thornhill 2011). By consequence, we chose to use the observed ITS2 sequence diversity as a fingerprint of the symbiont population content, without further tentative to sort intragenomic variation from among strains differentiation.

### Absence of genetic differentiation among the morphs of *Anemonia viridis*

In the present study we investigated whether the different *A. viridis* morphs were genetically differentiated and therefore could represent adaptations to different environmental conditions. This hypothesis is supported by the previous observations of Wiedenmann et al. (1999) who identified a differential bathymetric distribution for some of the recognized *A. viridis* morphs (var. *rustica* and *smaragdina*). Previously, using five EPIC markers, we showed that these morphs were not differentiated species (Mallien et al. 2017). However, we could have missed a low level of differentiation if the morphs were still in the speciation grey zone (De Queiroz 2007). In the present work, using high throughput RADseq, we could demonstrate that *A. viridis* var. *smaragdina* and *rustica* are definitely not monophyletic groups, independent of the origin of the samples (Mediterranean Sea or English Channel). However, independent genetic lineages were found in our dataset (Figs. 3 and S3), which were not correlated to morphology but, on the whole, rather correlated to the geographic origin of the samples. In addition, the clustering pattern analysis identified several groups of individuals mixing different morphs together and demonstrating admixture, and hence interfecundity between morphs (Fig. 4). Concerning var. *rufescens*, our previous study suggested that only this morph was more differentiated (Mallien et al. 2017). The present analysis showed

that this morph was not a monophyletic independent lineage because the var. *rufescens* individuals belonged to at least two different genetic lineages (BanRuf in the Mediterranean Sea and EngCh in the English Channel, Figs. 3 and 4). This discordance between morphological differentiation and genetic polymorphism was also found in another sea anemones, e.g. *Actinia equina*, which displays three different morphotypes (brown, red and green), out of which at least two have been shown by phylogenetic analysis to share two genetic units (Pereira et al. 2014). In the study models for developmental and functional studies belonging to the genus *Exaiptasia*, Grajales and Rodríguez (2016) identified a new cryptic species, *E. brasiliensis*, morphologically indistinguishable and living in sympatry with *E. pallida*. Spano et al. (2018) using RADseq, inferred the existence of a “wide-spread complex of sea anemones” in the genus *Anthothoe*, not correlated with classical morphological features but here also structured by geographical distribution.

### A similar symbiotic dinoflagellate composition among the morphs

Having demonstrated that *A. viridis* var. *smaragdina* and *rustica* belong to the same genetic pool, we tested the correlation between the symbiotic composition and the morphotypes mainly differentiated by FP and related pigments expression (Wiedenmann et al. 1999; Wiedenmann et al. 2000). These proteins may have photoprotective functions (Salih et al. 2000; Gittins et al. 2015) or they may enhance symbiont photosynthesis (Quick et al. 2018), suggesting a possible correlation between morph fluorescence pattern and the symbiotic composition of the morphs, irrespective of genetic differentiation. To assess the genetic diversity of the *A. viridis* symbiont populations, that all belong to the unique temperate clade A (Savage et al. 2002; Visram et al. 2006; Forcioli et al. 2011; Casado-Amezúa et al. 2014), we genotyped the pooled Symbiodiniaceae from each single host and compared the obtained symbiotic compositions with one another in order to detect any differentiation among the different host morphs or lineages.

In *A. viridis*, the distribution of the ITS2 sequence variants did not correlate with the three host morphs. If a differentiation between var. *smaragdina* and *rustica* was indeed detected with the T50 dataset (Table S5), it was restricted to within the Med2 lineage, thus suggesting a sampling artefact. Such a differentiation was not found with the SP dataset and if we only used the individuals present in the SP datasets in the D&S pipeline, this morph differentiation within Med2 lineage disappears (result not shown). This absence of correlation between colour morphs and symbiont populations was not in accordance with the previous observations in the coral *Madracis pharensis*

(Frade et al. 2008), where a relation between the symbiont type harboured by the coral colonies and the colour of the colonies was found. As such the generalisation that the cnidarian FP pattern is a response to Symbiodiniaceae diversity may be excluded. We also could hypothesise that the FP expressed by *A. viridis* has a less important role in the photoprotection of Symbiodiniaceae than in other symbiotic cnidarians (Salih et al. 2000; Smith et al. 2013; Gittins et al. 2015).

### What is the origin of morph differentiation in *A. viridis*?

The question of the nature of the *A. viridis* morphs still remains. Even if the morphs are not divergent host lineages nor due to morph specific symbiont assemblages, we still cannot exclude that morphs are genetically determined. Very few genes could be involved in such morphological variation, and even a unique gene could be responsible for a critical effect on the phenotype, as in maize (Dorweiler et al. 1993). Even in *Littorina saxatilis*, a marine gastropod harbouring different ecomorphs along the foreshore and known as a classical example of ecological speciation, only 4% of SNPs were identified as outliers between the morphologically differentiated species (Kess et al. 2018). Given this, a genome wide association study could be powerful enough to detect SNPs correlated to the morph pattern (Korte and Farlow 2013) in *A. viridis* if the morphs are due to allelic differences at a few major genes. Unfortunately, the RADseq loci generated in our study covered only a small part of the genome (less than 1% considering the number of RADseq loci obtained and an estimated genome size of around 480 Mb) and would not be enough for this analysis.

Obviously, the morphological variations within *A. viridis* could also be a plastic phenotypic response of the organism. In corals, the expression profile of FPs can vary during the different development stages (Roth et al. 2013). However, to the best of our knowledge, there was no report of this kind of variation in *A. viridis* neither in the literature nor during monitored embryonic development (Zamoum et al., personal communication). Nevertheless, *A. viridis* morphs could still be the consequence of a local plastic response at the larval settlement stage, followed by developmental canalization (Zamer and Mangum 1979; Kawecki and Ebert 2004).

### Cryptic species in the snakelocks anemone

The RADseq was resolute enough to detect differentiation among different independent genetic groups but not among the studied morphs. This approach enabled us to identify four cryptic species (BanRuf, Med1, Med2 and EngCh) in *A. viridis*, which were supported by three lines of analysis:

SVDquartets (Fig. 3), *snmf* (Fig. 4) and BFD\* (Table S3). Three of these species globally coincided with the geographical origin of the sea anemones: an English lineage (EngCh), a Mediterranean Sea lineage (Med2) and a mostly Mediterranean lagoon lineage (Med1). Our data also suggest further structuring within these lineages, as all displayed relatively strong  $F_{IS}$  values probably diagnostic of Wahlund effects (Table S4).

### Vicariance between the English Channel and Mediterranean Sea

During the last glacial period, the English Channel coasts were not suitable habitats for *A. viridis* (Bianchi and Morri 2000). The English populations of *A. viridis* are therefore the result of a post-glacial recolonization. The differences observed between EngCh and the Mediterranean lineages could then be the result of a vicariance process, in line with what was described for the lagoon cockle *Cerastoderma glaucum* (Sromek et al. 2019) or *Cymodocea nodosa* (Alberto et al. 2008). This cleavage was also supported by the distribution of the symbiont assemblages (Tables 2 and S9). A similar pattern of distribution of the symbiotic genetic diversity was observed in another sea anemone species, *Aiptasia couchii*, with a temperate clade A “Atlantic strain” in the English Channel and a “Mediterranean strain” on Portuguese and Spanish Atlantic and Mediterranean coasts (Grajales et al. 2016). However, the lack of sampling points between the North-Western Mediterranean and English Channel in our datasets does not allow for a formal test of this scenario.

### Cryptic species within the Western Mediterranean basin?

In our study, the Med1 lineage individuals were, for their vast majority, sampled in the Thau lagoon. This lineage could therefore correspond to some differentiation due to ecological conditions (Thau lagoon vs Mediterranean open sea), as found in other lagoon species (Lemaire et al. 2000; Chaoui et al. 2012). This was reinforced by the differentiation of symbiont populations between Med1 and Med2, which could be linked to a functional intra-cladal diversity (Howells et al. 2012), or as the result of a strict vertical transmission of the symbionts in *A. viridis* (Schäfer 1984). Nevertheless, we also found individuals belonging to the Med1 lineage at Banyuls (i.e. not in a lagoon) and individuals belonging to the Med2 lineage in the Thau lagoon. Furthermore, we do not know the extent of the geographical distribution of the species in the Western Mediterranean Sea.

At Banyuls, we found three species in sympatry (BanRuf, Med1 and Med2) with or without introgression among them (Fig. 4). We know however that the BanRuf lineage could not be a “sampling site effect” because the var-

*rufescens* individuals in the study of Mallien et al. (2017) came from another more northern site (Collioure, France, about 7 km away from Banyuls) and were already differentiated from the others *A. viridis* individuals. Nevertheless, as Banyuls was the more genetically diverse site, we hence hypothesise that the French–Spanish border could harbour a secondary contact zone between a Western Mediterranean genetic pool (which contains the Med2 and Med1 lineages) and another highly divergent one (BanRuf) from the Northern coast of Africa or from the Atlantic Ocean. Again, a formal test of this hypothesis would require a more thorough sampling of the geographical zone than the one we performed for this study, which was aimed at clarifying the taxonomic status of *A. viridis* colour morphs.

To conclude, *A. viridis*, as it is described today, appears as a complex of more or less introgressed cryptic species sharing morphotypes and harbouring differentiated symbiont populations. These different lineages, from the English Channel to the Mediterranean Sea, do live under different environmental conditions. Therefore, if there is any adaptive genetic differentiation to be found in *A. viridis*, it should be looked for among the genetic lineages rather than among the morphs.

### Data archiving

Data available from the Dryad Digital Repository: <https://doi.org/10.5061/dryad.2j6t1mb>.

**Acknowledgements** This work was supported by the ANR Bioadapt research program “AdaCni” (ANR-12-ADAP-0016) as a Ph.D. grant to CM and by a doctoral fellowship from the French Ministère de l’Enseignement supérieur et de la Recherche to BP. We thank J. Wiedenmann, P. Ventura, A. Haguenauer, R. Rodolpho-Metalpa, IMBE and M. Wanguemert for their contribution to the sampling effort. We also thank P.-A. Gagnaire and L. Sromek for the RADseq individual tags; P. Barby from IPMC who gave us the authorization to use the CovarisB sonicator and V. Magnone and M.-J. Arguel for their help during its utilization.

### Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

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

### References

- Alberto F, Massa S, Manent P, Diaz-Almela E, Arnaud-Haond S, Duarte CM et al. (2008) Genetic differentiation and secondary contact zone in the seagrass *Cymodocea nodosa* across the Mediterranean-Atlantic transition region. *J Biogeogr* 35:1279–1294
- Alexander DH, Novembre J, Lange K (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Res* 19:1655–1664

- Appeltans W, Ahyong ST, Anderson G, Angel MV, Artois T, Bailly N et al. (2012) The magnitude of global marine species diversity. *Curr Biol* 22:2189–2202
- Arif C, Daniels C, Bayer T, Banguera-Hinestrosa E, Barbrook A, Howe CJ et al. (2014) Assessing *Symbiodinium* diversity in scleractinian corals via next-generation sequencing-based genotyping of the ITS2 rDNA region. *Mol Ecol* 23:4418–4433
- Bailleul D, Stoeckel S, Arnaud-Haond S (2016) RClone: a package to identify multilocus clonal lineages and handle clonal data sets in r. *Methods Ecol Evol* 7:966–970
- Bandelt HJ, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16:37–48
- Bandelt HJ, Peter F, Bryan CS, Richards MB (1995) Mitochondrial portraits of human populations using median networks. *Genetics* 141:743–753
- Batovska J, Cogan NOI, Lynch SE, Blacket MJ (2017) Using next-generation sequencing for DNA barcoding: capturing allelic variation in ITS2. *G3: Genes, Genomes, Genet* 7:19–29
- Bellis ES, Howe DK, Denver DR (2016) Genome-wide polymorphism and signatures of selection in the symbiotic sea anemone *Aiptasia*. *BMC Genomics* 17:160
- Bianchi CN, Morri C (2000) Marine biodiversity of the Mediterranean Sea: situation, problems and prospects for future research. *Mar Pollut Bull* 40:367–376
- Bonthond G, Merselis DG, Dougan KE, Graff T, Todd W, Fourqurean JW et al. (2018) Inter-domain microbial diversity within the coral holobiont *Siderastrea siderea* from two depth habitats. *PeerJ* 6:e4323
- Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D et al. (2014) BEAST 2: a software platform for Bayesian evolutionary analysis. *PLOS Comput. Biol* 10:e1003537
- Boyer F, Mercier C, Bonin A, Le Bras Y, Taberlet P, Coissac E (2016) Obitools: a unix -inspired software package for DNA metabarcoding. *Mol Ecol Resour* 16:176–182
- Bryant D, Bouckaert R, Felsenstein J, Rosenberg NA, RoyChoudhury A (2012) Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Mol Biol Evol* 29:1917–1932
- Bulnheim HP, Sauer KP (1984) *Anemonia sulcata*-two species? Genetical and ecological evidence. *Verhandlungen der Dtsch Zoologische Ges* 77:264
- Casado-Amezúa P, Machordom A, Bernardo J, González-Wangüemert M (2014) New insights into the genetic diversity of zooxanthellae in Mediterranean anthozoans. *Symbiosis* 63:41–46
- Chaoui L, Gagnaire P-A, Guinand B, Quignard J-P, Tsigenopoulos C, Kara MH et al. (2012) Microsatellite length variation in candidate genes correlates with habitat in the gilthead sea bream *Sparus aurata*. *Mol Ecol* 21:5497–5511
- Chifman J, Kubatko L (2014) Quartet inference from SNP data under the coalescent model. *Bioinformatics* 30:3317–3324
- Creed ER, Lees DR, Bulmer MG (1980) Pre-adult viability differences of melanic *Biston betularia* (L.) (Lepidoptera). *Biol J Linn Soc* 13:251–262
- Cunning R, Gates RD, Edmunds PJ (2017) Using high-throughput sequencing of ITS2 to describe *Symbiodinium* metacommunities in St. John, US Virgin Islands. *PeerJ* 5:e3472
- Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA et al. (2011) The variant call format and VCFtools. *Bioinformatics* 27:2156–2158
- Davy SK, Lucas IAN, Turner JR (1996) Carbon budgets in temperate anthozoan-dinoflagellate symbioses. *Mar Biol* 126:773–783
- De Brauwer M, Hobbs J-PA, Ambo-Rappe R, Jompa J, Harvey ES, McIlwain JL (2018) Biofluorescence as a survey tool for cryptic marine species. *Conserv Biol* 32:706–715
- De Queiroz K (2007) Species concepts and species delimitation. *Syst Biol* 56:879–886
- Dorweiler J, Stec A, Kermicle J, Doebley J (1993) Teosinte glume architecture 1: a genetic locus controlling a key step in maize evolution. *Science* 262:233–235
- Douek J, Barki Y, Gateño D, Rinkevich B (2002) Possible cryptic speciation within the sea anemone *Actinia equina* complex detected by AFLP markers. *Zool J Linn Soc* 136:315–320
- Eaton DAR (2014) PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics* 30:1844–1849
- Emerson KJ, Merz CR, Catchen JM, Hohenlohe PA, Cresko WA, Bradshaw WE et al. (2010) Resolving postglacial phylogeography using high-throughput sequencing. *PNAS* 107:16196–16200
- Eren AM, Morrison HG, Lescault PJ, Reveillaud J, Vineis JH, Sogin ML (2015) Minimum entropy decomposition: unsupervised oligotyping for sensitive partitioning of high-throughput marker gene sequences. *ISME J* 9:968–979
- Etter PD, Preston JL, Bassham S, Cresko WA, Johnson EA (2011) Local de novo assembly of RAD paired-end contigs using short sequencing reads. *PLoS ONE* 6:e18561
- Fišer C, Robinson CT, Malard F (2018) Cryptic species as a window into the paradigm shift of the species concept. *Mol Ecol* 27:613–635
- Flot J-F, Blanchot J, Charpy L, Cruaud C, Licuanan WY, Nakano Y et al. (2011) Incongruence between morphotypes and genetically delimited species in the coral genus *Stylophora*: phenotypic plasticity, morphological convergence, morphological stasis or interspecific hybridization? *BMC Ecol* 11:22
- Forcioli D, Merle P, Caligara C, Ciosi M, Muti C, Francour P et al. (2011) Symbiont diversity is not involved in depth acclimation in the Mediterranean Sea whip *Eunicella singularis*. *Mar Ecol Prog Ser* 439:57–71
- Forskal, P (1775) *Descriptiones Animalium Avium, Amphibiorum, Piscium, Insectorum, Vermium; quae in itinere orientali observavit*. Copenhagen: Mölleri, Havniae, p. 164
- Frade PR, Englebert N, Faria J, Visser PM, Bak RPM (2008) Distribution and photobiology of *Symbiodinium* types in different light environments for three colour morphs of the coral *Madracis pharensis*: is there more to it than total irradiance? *Coral Reefs* 27:913–925
- Frichot E, François O (2015) LEA: an R package for landscape and ecological association studies. *Methods Ecol Evol* 6:925–929
- Frichot E, Mathieu F, Trouillon T, Bouchard G, François O (2014) Fast and efficient estimation of individual ancestry coefficients. *Genetics* 196:973–983
- Gittins JR, D'Angelo C, Oswald F, Edwards RJ, Wiedenmann J (2015) Fluorescent protein-mediated colour polymorphism in reef corals: multicopy genes extend the adaptation/acclimatization potential to variable light environments. *Mol Ecol* 24:453–465
- González-Muñoz R, Simões N, Mascaró M, Tello-Musi JL, Brugler MR, Rodríguez E (2015) Morphological and molecular variability of the sea anemone *Phymanthus crucifer* (Cnidaria, Anthozoa, Actiniaria, Actinoidea). *J Mar Biol Assoc U.K.* 95:69–79
- Goudet J (2005). hierfstat, a package for r to compute and test hierarchical F-statistics. *Molecular Ecology Notes* 5:184–186
- Grajales A, Rodriguez E (2016) Elucidating the evolutionary relationships of the Aiptasiidae, a widespread cnidarian-dinoflagellate model system (Cnidaria: Anthozoa: Actiniaria: Metridioidea). *Mol Phylogenetics Evol* 94:252–263
- Grajales A, Rodríguez E, Thornhill DJ (2016) Patterns of *Symbiodinium* spp. associations within the family Aiptasiidae, a monophyletic lineage of symbiotic of sea anemones (Cnidaria, Actiniaria). *Coral Reefs* 35:345–355
- Hipp AL, Eaton DAR, Cavender-Bares J, Fitzek E, Nipper R et al. (2014) A Framework Phylogeny of the American Oak Clade Based on Sequenced RAD Data. *PLoS ONE* 9:e93975. <https://doi.org/10.1371/journal.pone.0093975>

- Howells EJ, Beltran VH, Larsen NW, Bay LK, Willis BL, van Oppen MJH (2012) Coral thermal tolerance shaped by local adaptation of photosymbionts. *Nat Clim Change* 2:116–120
- Hume BCC, Smith EG, Ziegler M, Warrington HJM, Burt JA, LaJeunesse TC et al. (2019) SymPortal: a novel analytical framework and platform for coral algal symbiont next-generation sequencing ITS2 profiling. *Mol Ecol Resour* 19(4):1063–1080
- Hume BCC (2019) didillysquat/SymPortal\_framework: v0.2.7. Zenodo
- Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* 23:254–267
- Kamvar ZN, Tabima JF, Grünwald NJ (2014) Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* 2:e281
- Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecol Lett* 7:1225–1241
- Kess T, Galindo J, Boulding EG (2018) Genomic divergence between Spanish *Littorina saxatilis* ecotypes unravels limited admixture and extensive parallelism associated with population history. *Ecol Evol* 8:8311–8327
- Korte A, Farlow A (2013) The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods* 9:29
- LaJeunesse TC, Parkinson JE, Gabrielson PW, Jeong HJ, Reimer JD, Voolstra CR et al. (2018) Systematic revision of symbiodiniaceae highlights the antiquity and diversity of coral endosymbionts. *Curr Biol* 28:2570–2580.e6
- LaJeunesse TC, Thornhill DJ (2011) Improved resolution of reef-coral endosymbiont (*Symbiodinium*) species diversity, ecology and evolution through pdbA non-coding region genotyping. *PLoS ONE* 6:e29013
- Leaché AD, Fujita MK, Minin VN, Bouckaert RR (2014) Species delimitation using genome-wide SNP data. *Syst Biol* 63:534–542
- Leigh JW, Bryant D (2015) Popart: full-feature software for haplotype network construction. *Methods Ecol Evol* 6:1110–1116
- Lemaire C, Allegrucci G, Naciri M, Bahri-Sfar L, Kara H, Bonhomme F (2000) Do discrepancies between microsatellite and allozyme variation reveal differential selection between sea and lagoon in the sea bass (*Dicentrarchus labrax*)? *Mol Ecol* 9:457–467
- Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997v2 [q-bio.GN]
- Mallien C, Porro B, Zamoum T, Olivier C, Wiedenmann J, Furla P et al. (2017) Conspicuous morphological differentiation without speciation in *Anemonia viridis* (Cnidaria, Actiniaria). *Syst Biodivers* 0:1–16
- Merilä J, Hendry AP (2013) Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evolut Appl* 7:1–14
- Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 16:1215
- Muscanté L, Ferrier-Pagès C, Blackburn A, Gates RD, Baghdasarian G, Allemand D (1998) Cell-specific density of symbiotic dinoflagellates in tropical anthozoans. *Coral Reefs* 17:329–337
- Muscanté L, Grossman D, Doino J (1991) Release of symbiotic algae by tropical sea anemones and corals after cold shock. *Mar Ecol Prog Ser* 77:233–243
- Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D et al. (2019). vegan: Community Ecology Package. R package version 2.5-5. <https://CRAN.R-project.org/package=vegan>
- Pante E, Abdelkrim J, Viricel A, Gey D, France S, Boisselier M et al. (2015) Use of RAD sequencing for delimiting species. *Heredity* 114:450–459
- Pante E, Puillandre N, Viricel A, Arnaud-Haond S, Aurelle D, Castelin M et al. (2015) Species are hypotheses: avoid connectivity assessments based on pillars of sand. *Mol Ecol* 24:525–544
- Pereira AM, Brito C, Sanches J, Sousa-Santos C, Robalo JI (2014) Absence of consistent genetic differentiation among several morphs of *Actinia* (Actiniaria: Actiniidae) occurring in the Portuguese coast. *Zootaxa* 3893:595–600
- Pérez-Ponce de León G, Poulin R (2016) Taxonomic distribution of cryptic diversity among metazoans: not so homogeneous after all. *Biol Lett* 12:20160371
- Pfenninger M, Schwenk K (2007) Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. *BMC Evolut Biol* 7:121
- Price TD, Qvarnström A, Irwin DE (2003) The role of phenotypic plasticity in driving genetic evolution. *Proc R Soc Lond Ser B: Biol Sci* 270:1433–1440
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- Quick C, D'Angelo C and Wiedenmann J (2018) Trade-offs associated with photoprotective green fluorescent protein expression as potential drivers of balancing selection for color polymorphism in reef corals. *Front Mar Sci* 5:1–11. <https://doi.org/10.3389/fmars.2018.00011>
- Quigley KM, Davies SW, Kenkel CD, Willis BL, Matz MV, Bay LK (2014) Deep-sequencing method for quantifying background abundances of symbiodinium types: exploring the rare symbiodinium biosphere in reef-building corals. *PLoS ONE* 9:e94297
- Ree RH, Hipp AL (2015) Inferring phylogenetic history from restriction site associated DNA (RADseq). In: Hörandl E, Appelhans MS (eds) Next-generation sequencing in plant systematics'. Chapter 6, <https://doi.org/10.14630/000007>. Koeltz botanical books, Oberreifenberg © International Association for Plant Taxonomy (IAPT), pp 181–204
- Richier S, Merle P-L, Furla P, Pigozzi D, Sola F, Allemand D (2003) Characterization of superoxide dismutases in anoxia- and hyperoxia-tolerant symbiotic cnidarians. *Biochimica et Biophysica Acta* 1621:84–91
- Rodríguez E, Barbeitos MS, Brugler MR, Crowley LM, Grajales A, Gusmão L et al. (2014) Hidden among sea anemones: the first comprehensive phylogenetic reconstruction of the order Actiniaria (Cnidaria, Anthozoa, Hexacorallia) reveals a novel group of hexacorals. *PLoS ONE* 9:e96998
- Roth MS, Fan T-Y, Deheyn DD (2013) Life history changes in coral fluorescence and the effects of light intensity on larval physiology and settlement in *Seriatopora hystrix*. *PLoS ONE* 8:e59476
- Salih A, Larkum A, Cox G, Kühl M, Hoegh-Guldberg O (2000) Fluorescent pigments in corals are photoprotective. *Nature* 408:850–853
- Sampayo EM, Dove S, LaJeunesse TC (2009) Cohesive molecular genetic data delineate species diversity in the dinoflagellate genus *Symbiodinium*. *Mol Ecol* 18:500–519
- Santos SR, Gutierrez-Rodriguez C, Coffroth MA (2003) Phylogenetic identification of symbiotic dinoflagellates via length heteroplasmy in domain V of chloroplast large subunit (cp23S)-ribosomal DNA sequences. *Mar Biotechnol* 5:130–140
- Sauer KP (1986) Strategien zeitlicher und räumlicher Einnischung. *Verhandlungen der Deutschen Zoologischen Gesellschaft*, 79:11–30
- Sauer KP, Müller M, Weber M (1986) Alloimmune memory for glycoprotein recognition molecules in sea anemones competing for space. *Mar Biol* 92:73–79
- Savage AM, Goodson MS, Visram S, Trapido-Rosenthal H, Wiedenmann J, Douglas AE (2002) Molecular diversity of symbiotic algae at the latitudinal margins of their distribution: dinoflagellates of the genus *Symbiodinium* in corals and sea anemones. *Mar Ecol Prog Ser* 244:17–26
- Schäfer W (1984) Fortpflanzung und Entwicklung von *Anemonia sulcata* (Anthozoa, Actiniaria). I: Fortpflanzungszyklus und

- Struktur der Oocyten vor und nach der Besamung. Helgoländer Meeresuntersuchungen 38:135–148
- Schmidt-Roach S, Lundgren P, Miller KJ, Gerlach G, Noreen AME, Andreakis N (2013) Assessing hidden species diversity in the coral *Pocillopora damicornis* from Eastern Australia. *Coral Reefs* 32:161–172
- Smith EG, D'Angelo C, Salih A, Wiedenmann J (2013) Screening by coral green fluorescent protein (GFP)-like chromoproteins supports a role in photoprotection of zooxanthellae. *Coral Reefs* 32:463–474
- Smith EG, Ketchum RN, Burt JA (2017) Host specificity of *Symbiodinium* variants revealed by an ITS2 metahaplotype approach. *ISME J* 11:1500–1503
- Spano CA, Häussermann V, Acuña FH, Griffiths C, Seeb LW, Gomez-Uchida D (2018) Hierarchical biogeographical processes largely explain the genomic divergence pattern in a species complex of sea anemones (Metridioidea: Sagartiidae: Anthothoe). *Mol Phylogenetics Evol* 127:217–228
- Sromek L, Forcioli D, Lasota R, Furla P, Wolowicz M (2019) Next-generation phylogeography of the cockle *Cerastoderma glaucum*: highly heterogeneous genetic differentiation in a lagoon species. *Ecol Evol* 9:4667–4682
- Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313
- Stat M, Pochon X, Cowie R, Gates R (2009) Specificity in communities of *Symbiodinium* in corals from Johnston Atoll. *Mar Ecol Prog Ser* 386:83–96
- Struck TH, Feder JL, Bendiksby M, Birkeland S, Cerca J, Gusarov VI et al. (2018) Finding evolutionary processes hidden in cryptic species. *Trends Ecol Evol* 33:153–163
- Suggett DJ, Hall-Spencer JM, Rodolfo-Metalpa R, Boatman TG, Payton R, Tye Pettay D et al. (2012) Sea anemones may thrive in a high CO<sub>2</sub> world. *Glob Change Biol* 18:3015–3025
- Swofford DL (1998) Phylogenetic analysis using parsimony (paup). Sinauer Associates, Sunderland, MA
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30:2725–2729
- Titus BM, Daly M (2018) Reduced representation sequencing for symbiotic anthozoans: are reference genomes necessary to eliminate endosymbiont contamination and make robust phyogeographic inference? *bioRxiv* 440289. <https://doi.org/10.1101/440289>
- Ventura P, Jarrold MD, Merle P-L, Barnay-Verdier S, Zamoum T, Rodolfo-Metalpa R et al. (2016) Resilience to ocean acidification: decreased carbonic anhydrase activity in sea anemones under high pCO<sub>2</sub> conditions. *Mar Ecol Prog Ser* 559:257–263
- Veron J, Pichon M (1976) Scleractinia of eastern Australia. Part I: families Thamnasteriidae, Astrocoeniidae, Pocilloporidae. Australian Government Publishing Service, Canberra. p 208.
- Visram S, Wiedenmann J, Douglas AE (2006) Molecular diversity of symbiotic algae of the genus *Symbiodinium* (Zooxanthellae) in cnidarians of the Mediterranean Sea. *J Mar Biol Assoc UK* 86:1281–1283
- Wiedenmann J, Elke C, Spindler KD, Funke W (2000) Cracks in the β-can: fluorescent proteins from *Anemonia sulcata* (Anthozoa, Actinaria). *Proc Natl Acad Sci USA* 97:14091–14096
- Wiedenmann J, Kraus P, Funke W, Vogel W (2000) The relationship between different morphs of *Anemonia* aff. *sulcata* evaluated by DNA fingerprinting (Anthozoa, Actinaria). *Ophelia* 52:57–64
- Wiedenmann J, Röcker C, Funke W (1999) The morphs of *Anemonia* aff. *sulcata* (Cnidaria, Anthozoa) in particular consideration of the ectodermal pigments. In: Pfadenhauer J (ed.) Verhandlungen der Gesellschaft für Ökologie, vol 29. Spektrum Akademischer Verlag, Heidelberg Berlin, pp 497–503
- Zamer WE, Mangum CP (1979) Irreversible nongenetic temperature adaptation of oxygen uptake in clones of the sea anemone *Haliplanella luciae* (verrill). *Biol Bull* 157:536–547
- Zamoum T, Furla P (2012) Symbiodinium isolation by NaOH treatment. *J Exp Biol* 215:3875–3880