Journal Pre-proof

Population connectivity of fan-shaped sponge holobionts in the deep Cantabrian Sea

Kathrin Busch, Sergi Taboada, Ana Riesgo, Vasiliki Koutsouveli, Pilar Ríos, Javier Cristobo, Andre Franke, Klaus Getzlaff, Christina Schmidt, Arne Biastoch, Ute Hentschel

PII: S0967-0637(20)30214-4

DOI: https://doi.org/10.1016/j.dsr.2020.103427

Reference: DSRI 103427

To appear in: *Deep-Sea Research Part I*

Received Date: 25 June 2020

Revised Date: 23 October 2020

Accepted Date: 23 October 2020

Please cite this article as: Busch, K., Taboada, S., Riesgo, A., Koutsouveli, V., Ríos, P., Cristobo, J., Franke, A., Getzlaff, K., Schmidt, C., Biastoch, A., Hentschel, U., Population connectivity of fan-shaped sponge holobionts in the deep Cantabrian Sea, *Deep-Sea Research Part I* (2020), doi: https://doi.org/10.1016/j.dsr.2020.103427.

Population connectivity of fan-shaped sponge holobionts in the deep Cantabrian Sea

Kathrin Busch[1], Sergi Taboada[2,3,4], Ana Riesgo[3,4], Vasiliki Koutsouveli[3], Pilar Ríos[4,5], Javier Cristobo[4,6], Andre Franke[7], Klaus Getzlaff[1], Christina Schmidt[1], Arne Biastoch[1,8], Ute Hentschel[1,8*]

[1] GEOMAR Helmholtz Centre for Ocean Research Kiel, Düsternbrooker Weg 20, 24105 Kiel, Germany

[2] Universidad Autónoma de Madrid, Departamento de Biología (Zoología), Cantoblanco 28049, Madrid, Spain

[3] The Natural History Museum of London, Cromwell Road, London SW7 5BD, United Kingdom

[4] Universidad de Alcalá, Departamento de Ciencias de la Vida, Apdo. 20, 28805, Alcalá de Henares, Spain

[5] Instituto Español de Oceanografía. Centro Oceanográfico de Santander C/ Promontorio San Martín s/n, 39004 Santander, Spain

[6] Instituto Español de Oceanografía. Centro Oceanográfico de Gijón, C/ Príncipe de Asturias 70 bis, 33212 Gijón, Asturias, Spain

[7] Christian-Albrechts University of Kiel, Institute of Clinical Molecular Biology (IKMB), Rosalind-Franklin-Straße 12, 24105 Kiel, Germany

[8] Christian-Albrechts University of Kiel, Christian-Albrechts-Platz 4, 24118 Kiel, Germany

*__Corresponding author__: uhentschel@geomar.de

Key words: Porifera, population genetics, single-nucleotide polymorphisms (SNPs), amplicon sequencing, Lagrangian modelling, ocean sensing

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This study was funded by the European Union's Horizon 2020 research and innovation program under Grant Agreement No. 679849 (the SponGES project). This document reflects only the authors' view and the Executive Agency for Small and Medium-sized Enterprises (EASME) is not responsible for any use that may be made of the information it contains. The model computations were performed at the North German Supercomputing Alliance (HLRN) supported by the cooperative program 'RACE-Regional Atlantic Circulation and Global Change' (BMBF Grant 03F0824C).

Highlights

- Fan-shaped sponges display panmixia at three locations in the Cantabrian Sea. Subtle sponge population genetic and pronounced microbial differences were observed between a canyon and bank location lying < 100 km apart.
- Lagrangian modelling reveals variable inter-annual connectivity via ocean currents between the sampling regions.
- Interdisciplinary approaches, such as the here presented combination of sponge taxonomy, genetics, microbiology, and particle tracking modelling, can help to improve understanding about connectivity in the deep-sea. This is particularly crucial for key organisms like sponges and the ecosystems they form.

Abbreviations

TaP clade = *Topsentia*-and-*Petromica* clade

Abstract

Connectivity is a fundamental process driving the persistence of marine populations and their adaptation potential in response to environmental change. In this study, we analysed the population genetics of two morphologically highly similar deep-sea sponge clades (*Phakellia hirondellei* and the '*Topsentia*-and-*Petromica*' clade, (hereafter refered to as 'TaP clade')) at three locations in the Cantabrian Sea and simultaneously assessed the corresponding host microbiome by *16S* rRNA gene sequencing. A virtual particle tracking approach (Lagrangian modelling) was applied to assess oceanographic connectivity in the study area. We observed overall genetic uniformity for both sponge clades. Notably, subtle genetic differences were observed for sponges of the TaP clade and also their microbiomes between a canyon and bank location, < 100 km apart and with the same depth range. The Lagrangian model output suggests a strong retention of larvae in the study area with variable inter-annual connectivity via currents between the three sampling regions. We conclude that geologic features (canyons) and the prevailing ocean currents may dictate sponge holobiont connectivity and that differentiation can emerge even on small spatial scales.

1. Introduction

The ocean is the largest interconnected habitat on planet Earth (Ramirez-Llodra *et al.*, 2011). At the same time, it is a highly dynamic system showing natural fluctuations, but also discrete responses to human impacts. Connectivity studies are important for our understanding regarding the resilience of ecosystems to changing oceanic conditions, as well as for an evaluation of population vulnerability to environmental change and anthropogenic stressors (Fox *et al.,* 2016; Kenchington *et al.,* 2019). Understanding population- and ecosystem-connectivity in the ocean is crucial for the design of appropriate conservation measures, in particular for the design of marine protected area (MPA) networks (White *et al.,* 2014; Gallego *et al.,* 2017; Andrello *et al.,* 2017; Kenchington *et al.,* 2019). Many marine benthic organisms have pelagic larvae that ensure the maintenance of genetic diversity and allow for a colonization of new habitats (e.g. review by Levin, 2006; Cowen and Sponaugle, 2009). Larval dispersal is difficult to observe directly in the field and in particular in the deep-sea. Therefore, indirect methods such as molecular markers and virtual particle tracking are commonly applied to analyse genetic connectivity between organisms (e.g. review by Cowen and Sponaugle, 2009; Baltazar-Soares *et al.*, 2014; Breusing *et al.*, 2016)). To conduct virtual particle tracking, biophysical models based on a Lagrangian approach (i.e. individual particle tracking in space and time; Cowen and Sponaugle, 2009) are powerful tools. As the term 'biophysical' implies, in these approaches biological parameters (e.g. time point of larval release and pelagic larval duration) are integrated into a physical framework (i.e. ocean current models) to acquire large ensembles of passive drift trajectories, which represent larval dispersal (Breusing *et al.*, 2016). Information about the biology of the studied organism, including its reproductive cycle, is crucial to fine-tune respective models, but this kind of information is generally not available for the deep ocean. Along these lines, the number of population genetic studies addressing connectivity is still very small (Taylor and Roterman, 2017). Previous studies suggest that the offspring of sessile benthic invertebrates is typically transported 10-100 km distance away from the parents (Palumbi, 2004). However, identification and exploration of biophysical barriers, which may inhibit gene flow, is fundamental to understanding population connectivity. While depth is commonly known to be a main factor related to species and population structure (e.g. Van Soest and de Voogd, 2015; Taylor and Roterman, 2017; Indraningrat *et al.,* 2019), the barriers separating regions within a similar depth range are less understood.

The geology of the seafloor is relevant for the formation and maintenance of biophysical barriers on different spatial scales. On large spatial scales, massive features such

as shelf breaks, trenches and ridges direct major ocean currents. On regional scales, the seafloor may lead to up- or downwelling, or circulatory water flows. Seamounts and banks are prominent elevated geologic features on the regional scale. These settings are often habitats for diverse macroorganisms (Morato *et al.*, 2010) and microorganisms (Busch *et al.*, 2020a), which frequently are endemic to that location (De Forges *et al.*, 2000). Submarine canyons are prominent plunging geologic structures which represent crucial connections between productive shelf waters and the deep open ocean (Martín *et al.*, 2006) but may also potentially lead to retention of organic matter (Hickey *et al.*, 1986).

Around the northern shores of Spain, the continental shelf is comparably narrow (Gómez-Ballesteros *et al.*, 2014) (**Figure 1**) and streaked with numerous submarine canyons. The Avilés Canyon System, which is located in front of the Asturian coast, is one of the largest submarine canyons in Europe (Rumín-Caparrós *et al.*, 2016). Overall, the seafloor on the Cantrabrian shelf has a high structural complexity, with elevated topography (e.g. banks) occurring in close proximity to plunging geologic features (e.g. canyons). Along these lines, the Le Danois Bank (Spanish name: *El Cachucho*) is found within a few kilometers distance next to the Avilés Canyon System (**Figure 1**). Le Danois Bank is a designated Marine Protected Area (MPA) and the Avilés Canyon System is a designated Special Area of Conservation (SAC), as both areas harbor a great diversity of benthic organisms (Sánchez *et al.*, 2008; Sánchez *et al.*, 2014).

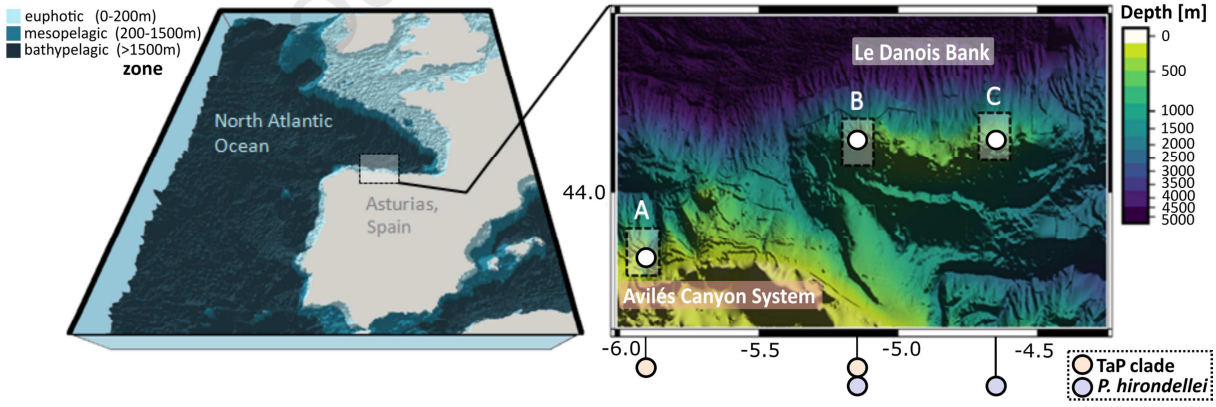

Figure 1 Combined bathymetric maps of the study area showing an overview of the study area in the Asturias region on the Spanish continental margin and a close-up of the three focus regions at the Avilés Canyon System (A: 'Canyon') and the Le Danois Bank (B: 'West-Bank' and C: 'East-Bank'). The colored circles below the map indicate the spatial distribution of the fan-shaped sponge samples included in this study.

Sponges are prominent and abundant benthic organisms in the deep North Atlantic Ocean (Klitgaard *et al.*, 1997; Klitgaard and Tendal, 2004; Murillo *et al.*, 2012). Dense aggregations of sponges in the deep-sea are commonly refered to as 'sponge grounds'.

Sponge grounds are often characterized as Vulnerable Marine Ecosystems (FAO 2009; FAO 2016) as they can be affected by anthropogenic activities such as bottom trawling (Kazanidis *et al.*, 2019; Busch *et al.*, 2020b) and oil and gas exploitation (Vad *et al.*, 2018). These threats stand opposed to the high ecological value of deep-sea sponge grounds. The ecological value of deep-sea sponge grounds includes enhancement of local biodiversity (Beazley *et al.*, 2013), formation of biomass hot-spots, and representation of a crucial role in biogeochemical cylcing (Cathalot *et al.*, 2015).

Sponges live in close association with dense and diverse symbiotic microbial consortia (Easson and Thacker, 2014; Hentschel *et al.*, 2012; Thomas *et al.*, 2016). The microbiomes of sponges are represented by diverse bacterial and archaeal clades with > 63 prokaryotic phyla having been identified so far (Thomas *et al.*, 2016; Moitinho-Silva *et al.*, 2017a). Microbial symbionts were shown to contribute to the metabolism of the sponge, i.e., via carbon and nitrogen cycling as well as vitamin production and defense (reviewed in Pita *et al.*, 2018). Sponges and their associated microbial communities (hereafter termed 'holobionts') are considered fine-tuned entities where both partners are tailored to function together. At the population-level, previous studies found sponge host population genetics and location to be related with the structure of the associated microbiomes at variable degrees, depending on the level of gene flow between the hosts (Griffiths *et al.*, 2019; Díez-Vives *et al.*, 2020; Easson *et al.*, 2020).

The present study aimed to explore the host genetic and microbial connectivity of sponge holobionts on a small geographic gradient < 100 km. We focused on fan-shaped sponges that are commonly found at the Avilés Canyon System and the Le Danois Bank of the Cantabrian Sea (**Figure 1**). These sponges of the '*Topsentia*-and-*Petromica*' clade (hereafter referred to as 'TaP clade') and *Phakellia hirondellei* are morphologically very similar even to the expert eye, but belong to different sponge orders. We sought to address whether genetic diversification can be resolved in deep-sea sponge holobionts sampled less than 100 km apart and what effect the geological setting (canyon versus bank) might have on population structure. In addition, oceanographic connectivity in the Cantabrian Sea was assessed by a virtual larvae tracking approach. Given that the continental shelf is comparably narrow in the study area, we were particularly interested to evaluate whether virtual sponge larvae would travel along the bathymetry. The presented findings contribute - in the long run - to a better management and preservation of vulnerable marine ecosystems.

2. Material and Methods

177

178 **2.1 Field work**

179 42 fan-shaped sponge specimens were collected with a rock dredge onboard *RV Ángeles*
180 *Alvariño* from the area of the Avilés Canyon System (~ 43.9 °N; 6.0 °W) and the Le Danois
181 Bank (~ 44.1 °N; 5.0 °W) in June 2017 during the expedition *SponGES0617*. In addition, 15
182 full depth CTD casts were performed at the respective and additional sampling stations. The
183 focus of this study was on three sampling regions: region A ('Canyon'), region B ('West-
184 Bank'), and region C ('East-Bank'), (**Figure 1**). Sampling at all three sampling regions was
185 conducted within a similar depth range (region A: 695 m, region B: 653 m, and region C: 541
186 m). Bathymetric data of the study area were extracted from the Bathymetry Data Portal of the
187 European Marine Observation and Data Network (EMODnet). Samples of different sponge
188 species were collected in the same grab. The sample details, corresponding (meta-) data and
189 raw sequence NCBI-accession numbers for all methodological approaches of all seawater and
190 sponge samples presented in this study were deposited in the PANGAEA data repository
191 (https://doi.org/10.1594/PANGAEA.923271).

192 After collection, sponge specimens were rinsed with fresh seawater and photographed.
193 Tissue subsamples were taken for microbial work and instantly flash-frozen. Sponge
194 fragments of each specimen (1-3 cm^3) were preserved in 96 % EtOH for molecular analysis
195 (see below), and immediately kept at -20 ºC; EtOH was replaced once after one day
196 preservation. Small sponge fragments were also incubated in sodium hypochlorite solution
197 (Panreac 5 % w/v technical grade) and kept at room temperature until the end of the cruise for
198 spicule analysis (see below). The remainder of the specimens not used for molecular and
199 morphological analyses was preserved in 80 % EtOH and kept at room temperature in the
200 reference collection at the Oceanographic Centre of Gijón (IEO, Spain). Seawater samples (2
201 L) for microbial community analyses were derived from each CTD casts' bottom depth (~ 5
202 m above seafloor), filtered onto PVDF filter membranes (pore size = 0.22 µm and ø = 47 mm)
203 and stored at -80 °C until further processing.

204

205 **2.2 Sponge morphological analyses**

206 For spicule preparation, soft sponge tissue was digested with sodium hypochlorite solution
207 (Panreac 5 % w/v technical grad.). The remaining spicules were soaked 3 x 1 h with distilled
208 water in a water bath at room temperature which was removed by repeated centrifugation (1
209 min at 12000 rpm). Finally, absolute ethanol was added. A few drops of ethanol and spicules
210 were placed on a slide which was then flamed for dehydration. A few drops of Durcupan™

ACM embedding mixture for microscopy (Sigma-Aldrich) were placed on the coverslip which was then kept in the oven at 36 ºC for 24 h before visualisation of spicules was performed using a microscope Nikon Eclipse 50i.

2.3 Sponge barcoding and phylogenetic analyses

DNA of all sponge specimens was extracted using the DNeasy Power Soil Kit (Qiagen) with approximately 0.25 g sponge tissue. The quality and quantity of resulting extracts were checked with a NanoDrop spectrophotometer. Fragments of the *18S* rDNA (*18S*) gene were amplified for all sponges using the primer pair SP18aF-SP18gR (SP18aF: 5´-CCTGCCAGTAGTCATATGCTT-3´; SP18gR: 5´-CCTTGTTACGACTTTTACTTCCTC-3´; Redmond *et al.*, 2013). The polymerase chain reaction (PCR) program for *18S* was 95 ºC/2 min - (95 ºC/2 min - 57 ºC/45 sec - 72 ºC/30 sec) x 30 cycles - 72 ºC/3 min. For the amplification of a fragment of the cytochrome *c* oxidase subunit I (*COI*), we used two sets of primers: Pcant-COIF (5'-TTTGCAGGGATGATCGGAAC-3') and Pcant-COIR (5'-CCCGGGGCCCTCATATTTAA-3') to obtain a fragment of 708 base pairs (for more details see Taboada *et al.* (*in review*)) for *P. hirondellei* and LCO1490 and HCO2198 (Folmer *et al.*, 1994) to obtain 685 bp. The PCR program for *COI* was 94 ºC/5 min - (94 ºC/30 s - 58 ºC/30 s - 72 ºC/30 s) x 35 cycles - 72 ºC/10 min. Amplification of both *18S* and *COI* was performed in 12.5 µL reactions, using 10.5 µL of VWR Red Taq DNA Polymerase 1.1x Master Mix (VWR International bvba/sprl, Belgium), 0.5 µL of the forward and reverse primers, and 1 µL of DNA template. PCR products were verified by gel electrophoresis on 1 % agarose. Sequencing was conducted with the primers mentioned above, using (dideoxy chain termination / cycle sequencing) on an ABI 3730XL DNA Analyser (Applied Biosystems, USA) by Eurofins Genomics for *18S* and at the Molecular Core Labs (Sequencing Facility) of the Natural History Museum of London (NHM) for *COI.*

Overlapping sequence fragments of *18S* and *COI* were assembled and trimmed separately into consensus sequences using the software Geneious v.10.1.3 (Kearse *et al.*, 2012). Occasionally only forward or reverse sequences were used due to the poor quality of one of the fragments (for all *18S* and three *COI* sequences). Consensus sequences were then checked for contamination using BLAST (Altschul *et al.*, 1990), aligned with MAFFT v.7.309 (Katoh and Standley, 2013) using sequences of closely related species, and used to construct a phylogenetic hypothesis with Maximum Likelihood using RAxML 8.1.22 (Stamatakis, 2014). The evolutionary model was selected using jModelTest (Darriba *et al.*, 2012), resulting in GAMMAGTR (GTR+G) for both markers. Phylogenetic analyses for *18S*

and *COI* were run separately ten times, with 100 replicates for bootstrap recovery. Our taxonomic distinction between species is then based on *18S* and *COI* phylogenetic analyses (see below) and spicule confirmation. This procedure, of combining classical sponge taxonomy with molecular sequencing, left us with 21 unambigiously identified sponge specimens for the present study, which are comprised of 11 TaP clade individuals and 10 *P. hirondellei* individuals.

2.4 Sponge reproductive state analyses

An assessment of the sponge reproductive state was conducted as biological ground-truthing for the biophysical particle tracking model. Our main aim of this analysis was to derive a reasonable timepoint of larval release into the water column. For the assessment, small pieces (5 mm^3) of five TaP clade individuals and three *P. hirondellei* were fixed in a solution of 2.5 % glutaraldehyde, 0.4M PBS and 0.34 NaCl onboard and stored at 4 ºC (modified after Koutsouveli *et al.*, 2018). In the home laboratory, samples were rinsed three times with buffer (0.4 M PBS-0.6 M NaCl) and then postfixed in 2 % osmium tetroxide for two hours at 4 ºC. After rinsing the sample pieces three additional times (with an incubation time of 15 min per washing step, at 4 °C), partial dehydration was conducted with an ascending ethanol series (2x 30 %- 1x 50 %- 1x 70 %). In the next step, sponge pieces were submerged in 5 % hydrofluoric acid overnight to remove any silica remnants from their skeleton. Subsequently, samples were washed carefully (8x 70 % EtOH) and dehydration was continued with an increasing gradient of ethanol (1x 90 %– 2x 100 %). Once dehydrated, samples were embedded in LR-White (LWR) resin, using a gradient of LRW/EtOH (1x 70/30– 1x 50/50– 1x 30/70– 2x 100). Afterwards, samples were transferred into fresh LWR and polymerized inside embedding capsules in an oven (57 °C) for two days. Blocks of LWR were cut using a Reichert Ultracut ultramicrotome equipped with a diamond knife (DIATOME, Switzerland) at 0.5–2 µm (semi-thin sections) and 70–90 nm (ultrathin sections) respectively. The semi-thin sections were stained with toluidine blue or Richardson solution (latter prepared as described in (Mulisch and Welsch, 2015)), while the ultra-thin sections for transmission electron microscopy (TEM) were contrasted with uranyl acetate and Reynold's lead citrate. Ultra-thin sections were visualised with a Hitachi Transmission Electron Microscope TEM (H-7650) and a Tecnai G2 Spirit Bio Twin TEM (FEI Company) at 90 kV. Semi-thin sections were visualised with an Axio Observer.Z1 microscope (Zeiss, Germany).

2.5 Sponge population genetic analysis

ddRADseq libraries were performed for a total of 21 individuals: 11 TaP clade specimens and 10 *P. hirondellei*. Library preparation was conducted following (Peterson *et al.*, 2012) with the following modifications by (Combosch *et al.*, 2017). Double-stranded genomic DNA (500 ng) was digested using the high-fidelity restriction enzymes SbfI and EcoRI (New England Biolabs). Resulting digested fragments were cleaned by manual pipetting using Agencourt AMPure beads (1.5X volume ratio; Beckham Coulter), and were subsequently quantified with a Qubit dsDNA HS assay (Life Technologies). In the next step, resulting fragments were ligated to custom-made P1 and P2 adapters containing sample-specific barcodes and primer annealing sites. Barcoded individuals were pooled into libraries, cleaned by manual pipetting using AMPure beads (1.5X volume ratio), and size-selected (range sizes 200–400 bp) using a Blue Pippin Prep (Sage Science). Each library was PCR-amplified using Phusion polymerase (Thermo Scientific) and a different set of PCR primers with barcodes in order to create multiplex libraries. The PCR program used was 98 ºC/30 s – (98 ºC/10 s – 65 ºC/30 s – 72 ºC/1.5 min) x 12 cycles – 72 ºC/10 min. Resulting PCR products were cleaned by manual pipetting using Agencourt AMPure beads (1.5X volume ratio), quantified with a Qubit dsDNA HS assay, and quality-checked on a Tapestation 2200 (Agilent Technologies). Libraries were pooled normalizing their concentration, and pooled together with RNA-seq libraries in the same flow cell. Libraries pair-end sequenced (150 bp) were run on an Illumina HiSeq 4000 (Illumina) at Macrogen Inc. (South Korea).

Quality filtering of reads and locus assembly was conducted with the Stacks pipeline, v2.4.1 (Catchen *et al.*, 2013). RAD-tags (DNA fragments with the two appropriate restriction enzyme cut sites that were selected, amplified, and sequenced) were processed using *process_radtags*, where raw reads were quality-trimmed to remove low quality reads, reads with uncalled bases, and reads without a complete barcode or restriction cut site. The *process_radtags* rescue feature (-r) was used to recover minimally diverged barcodes and RAD-tags (--barcode_dist 3; --adapter_mm 2). The *process_radtags* trimming feature (-t) was used to trim remaining reads to 120 bp, in order to increase confidence in single-nucleotide polymorphism (SNP) calling. After performing these filtering steps in *process_radtags*, we retained a total of 74,442,388 reads from the initial 102,673,760 raw reads (72.5 % of the total reads retained), with an average of 3,544,876 reads per sample.

Preliminary tests were carried out following (Jeffries *et al.*, 2016) and (Paris *et al.*, 2017) in order to identify optimal Stacks parameters (including m, M, and n) for our dataset. Briefly, tests were carried out for 5 sets of 3 randomly chosen individuals and, for each test, all non-test parameters were kept as default. The Stacks *populations* module was run to filter

data with r = 0.8 for each test, and the number of assembled loci, number of polymorphic loci, number of SNPs, and coverage was compared between the tests. Final parameter values were as follows: *ustacks*: M = 2, m = 3; *cstacks*: n = 1. Mean locus coverage among all samples was 47.1±13.0, ranging from 15.8 to 99.5.

The Stacks *populations* module was used to conduct a first filtering of the data, retaining those SNPs present in at least 70 % of the individuals (r = 0.7), and just retaining the first SNP from each RAD-tag using "--write_single_SNP" in order to reduce the linkage disequilibrium among loci. A subsequent more accurate filtering was performed using the *adegenet* R package (Jombart, 2008; R Development Core Team, 2008; Jombart and Ahmed, 2011), assessing SNP distributions across individual samples and sampling regions, and testing different filtering thresholds in order to maximise the number of retained SNPs and minimise missing data. This approach provides significant help in defining final thresholds in comparison with the Stacks *populations* module. The resulting assessment resulted in no further filtering of samples.

Given the known presence of symbiotic bacteria in all sponges, the resulting set of sequences containing variable SNPs obtained after running *populations* were filtered for bacterial hits. This was done using *–blastn* comparing the aforementioned set of sequences against a nr database for bacteria extracted from NCBI (accessioned on the 17/06/2018), using a e-value of 1e-6 or lower. This filtering of bacteria resulted in 0 being removed in both the *P. hirondellei* and TaP clade datasets. That said, only polymorphic SNPs were used.

We calculated population genetic diversity and demographic statistics separately for TaP clade specimens and *P. hirondellei* by grouping the samples per sampling region. Expected (*He*) and observed (*Ho*) heterozygosities, and inbreeding coefficients (F_{IS}) were calculated per each sponge clade per sampling region using Genodive v.3.02 (Meirmans and Van Tienderen, 2004). To assess global inbreeding within sampling regions and differentiation among them, we also calculated global inbreeding coefficient F_{IS} in Genodive.

We assessed the population structure for each sponge clade separately using three different methods: STRUCTURE v.2.3 (Pritchard *et al.,* 2000); the function *snapclust* in the *adegenet* R package (Beugin *et al.,* 2018); and the discriminant analysis of principal components (DAPC) as implemented in the *adegenet* R package (Jombart *et al.*, 2010). We ran STRUCTURE with 50,000 MCMC iterations using the admixture model, with a burn-in of 20,000 iterations, setting the putative *K* from 1 to 4 with 10 replicates for each run. We used STRUCTURE HARVESTER (Earl and vonHoldt, 2012) and CLUMPP v.1.1.2 (Jakobsson and Rosenberg, 2007) to determine the most likely number of clusters and to

average each individual's membership coefficient across the *K* value replicates, respectively. We used the Akaike Information Criterion (AIC) to identify the optimal number of clusters in *snapclust.choose.k*, and then initial memberships for *snapclust* were chosen using the k-means algorithm (pop.ini = "kmeans"), allowing a maximum *K* (number of clusters) of 12 (max = 12), and a n.start.kmeans of 100 (n.start.kmeans= 100). For the DAPC analysis we grouped samples by sampling region, and retained the number of principal components analysis (PCA) axes and eigen values using the cross-validation *xvalDapc* function from the *adegenet* R package. Finally, Pairwise F_{ST} values were estimated to measure the differentiation between pairs of sampling regions using Genodive v.3.02 (Meirmans and Van Tienderen, 2004) with 20,000 permutations.

2.6 Sponge-associated microbial community analyses

DNA extraction of all fan-shaped sponges was performed using the DNeasy Power Soil Kit (Qiagen) with approximately 0.25 g sponge tissue or half a seawater filter (for 46 seawater samples in total) used as input material. The concentration of resulting DNA extracts was checked with a NanoDrop spectrophotometer and their quality assessed by gel electrophoresis after a polymerase chain reaction (PCR) with the universal 16 rRNA gene primers 27F and 1492R. After the quality check, a one-step PCR (30x cycles: 98 °C/30 s - 98 °C/9 s - 55 °C/1 min - 72 °C/1.30 min - 72 °C/10 min - hold at 4°C) was conducted on the extracts to amplify the V3 to V4 variable regions of the *16S* rRNA gene (using the primer pair 341F 5'-CCTACGGGAGGCAGCAG-3' (Muyzer *et al.*, 1993) & 806R 5'-GGACTACHVGGGTWTCTAAT-3' (Caporaso *et al.*, 2011) in a dual-barcoding approach (Kozich *et al.*, 2013)). The amplicon libraries were quality checked by gel electrophoresis, normalised with the SequalPrep Normalization Plate Kit (ThermoFisher Scientific) and pooled equimolarily. Sequencing was performed on a MiSeq platform (MiSeqFGx, Illumina) using v3 chemistry and in subsequent demultiplexing 0 mismatches were allowed in the barcode sequences.

Sequences were processed using QIIME2 (version 2018.11, (Bolyen *et al.*, 2018), similar to the methods described in Busch *et al.,* (2020c). Briefly, the DADA2 algorithm (Callahan et al. 2016) was applied on forward reads (truncated to 270nt) to generate Amplicon Sequence Variants (ASVs), which in turn were used to calculate phylogenetic ASV trees with the FastTree2 plugin. Classification of representative ASV sequences (on the genus, family, order, class and phylum level) was performed by a primer-specific trained Naive Bayes taxonomic classifier, using the Silva 132 99% OTUs *16S* database (Quast *et al.*, 2013).

Statistical analyses of the microbial community were performed on all 21 sponge individuals, including eight TaP clade individuals from region A ('Canyon'), six individuals from region B ('West-Bank') (comprised of three TaP clade individuals + three *P. hirondellei* individuals), and seven *P. hirondellei* specimens from region C ('East-Bank'). Further, nine seawater samples were included in the statistical analyses (three samples from each region). Alpha diversity indices (among others Shannon index), as well as beta diversity metrices (among others weighted UniFrac distances) were calculated. Non-metric multidimensional scaling (nMDS) was performed on weighted UniFrac distances to evaluate sample separation in ordination space. To assess if groups of samples differed significantly from each other, permutational multivariate analyses of variance (PERMANOVA) were conducted in a pairwise manner on weighted UniFrac distances. Phyla differing significantly between sample groups were determined and ranked using the Linear Discriminant Analysis Effect Size (LEfSe) algorithm (Segata *et al.*, 2011). We applied a significance level of $\alpha = 0.05$ for all statistical analyses throughout this study. Plotting was performed with R (version 3.0.2, (R Development Core Team, 2008)) and further fine-tuning of visualisations conducted using Inkscape (version 0.92.3; (Harrington and Team, 2005)) and/or GIMP (version 2.8).

2.7 Larval Dispersal Modeling and Oceanographic Observations

Before running the virtual particle tracking model simulations, the following steps were performed (i) validation of technical input parameters (number of released particles), (ii) further validation of biological input parameters (timepoint of larval release), (iii) comparison of model output (TS data) with *in situ* measurements.

2.7.1 Validations with biological and physical oceanographic observations

To further validate the timepoint of larval release into the water column (as determined by the sponge reproductive state analyses), we conducted an assessment of pelagic productivity. This seemed reasonable as several previous studies have described a strong link between pulses of primary productivity and the larval release of benthic organisms (e.g. Highfield *et al.*, 2010). We assessed productivity by two means: (a) based on *in situ* biological data of bacterioplankton and chlorophyll-a concentrations (snapshot in time, relying on sampling in June 2017); and (b) based on remote sensing (assessment of temporal varitions within the year 2017). The latter was performed with the help of publicly available resources, as we derived monthly chlorophyll-a concentrations of the year 2017 from satellite data accessed via the GlobColor web portal (provided by the European Space Agency, ESA). *In situ* productivity

data was retrieved by flow cytometry using a Beckman Coulter Gallios machine. Samples for flow cytometry were derived in triplicates from bottom depths of all 15 performed CTD casts. Subsamples (4 mL seawater) were fixed with 200 µL glutaraldehyde (GDA, 25 %). Sample-containing tubes were stored in vertical position at -20 °C until processing. For measurement of bacterial cell numbers, the samples were thawed, prefiltered (syringe filters with 50 µm pore size), and sample aliquots of 400 µL were mixed with 10 µL of a SYBR Green stock solution (10,000x concentrate in DMSO). Fluoresbrite fluorescence microspheres with a diameter of 1 µm were added, followed by a product incubation time of 15 min. For measurements of phytoplankton concentrations, no staining was performed, but detection of autofluorescence conducted instead.

To validate our physical model output with *in situ* physical measurements (of TS data), and to get a better understanding of prevailing water mass properties, we performed *in situ* sensing using a Seabird 37 CTD sensor system. The bathymetry map was created with QGIS (version 3.4.4; (QGIS Development Team, 2017)). CTD data and satellite data were visualised with Python (version 3.7.3).

2.7.2 Biophysical modelling

We applied a Lagrangian modelling framework using the particle tracking toolbox Parcels (version 2.0.0.beta2, Delandmeter and Van Sebille, 2019) with three-dimensional velocity data from an Atlantic model based on the Nucleus for European Modelling of the Ocean (NEMO v3.6, Madec, 2016) system. VIKING20X is a successor of the Ocean General Circulation Model (OGCM) VIKING20 (Böning *et al.*, 2016) and combines an eddy-rich (1/20°) grid of the whole Atlantic (69 °N - 33.5 °S), two-way nested into a global 1/4° ocean-sea-ice configuration (ORCA025), forced by the JRA55-do atmospheric forcing data set of the past decades (Tsujino *et al.*, 2018). In this study we used the experiment VIKING20X.L46-KKG36107B. VIKING20X has 46 z-levels in the vertical, for which the layer thickness increases from 6 m at the surface to 250 m in the deep ocean. At the depths of the simulated regions (i.e. 628 m), the layer thickness is 94 m. The bathymetry in the model is based on the 2-minute ETOPO2 bathymetric database and represented by partial steps. Simulated velocities are stored as daily three-dimensional averages (to allow releases on an everyday basis to assure equal integration lengths). The VIKING20 model itself has been intensively validated in previous studies (e.g. Breckenfelder *et al.*, 2017) and used for dispersal studies (Baltazar-Soares *et al.*, 2014; Breusing *et al.*, 2016). To further demonstrate the reliability of the model simulations in our area of interest, we compared *in situ*

temperature and salinity data derived from the 15 priorly mentioned full ocean depth CTD casts in June 2017, with modelled data from VIKING20X (respective data are shown in the Supplementary Material and discussed further in the Results section 3.4). Lagrangian simulations were performed on ~ 90,000 particles in total (1,000 particles per release region and release day), which were released from three release regions (covering four model grid boxes each) during the assumed natural spawning period in June (1st of June - 30th of June). The coordinates of the three virtual release regions were covering the actual sampling locations in 2017 at region A ('Canyon'), region B ('West-Bank') and region C ('East-Bank') (consider the Supplementary Material for further insights into the representation of regions in model topography as well as model velocities). The number of particles was determined based on the saturation plateau of computed rarefraction curves (respective data are shown in the Supplementary Material and discussed further in the Results section 3.4). In our study, release positions of sponge larvae were close to the seafloor (in detail a larval release was performed in the vertical centre of the last gridbox above the bottom). After release, virtual larvae were allowed to drift passively at any depth with the three-dimensional time-varying ocean velocities, and no assumptions about mortality were made at any time of the simulations. A typical pelagic larval duration of 14 days was assumed and dispersal probabilities were computed as described previously (van Sebille *et al.*, 2018). To consider temporal fluctuations in current patterns, we modeled releases and dispersal over a period of 10 years (between 2009 and 2018). In-depth analyses on seasonality were performed for one year, in which case we chose the year of the conducted expedition (i.e. 2017). For plotting larval trajectories and probabilities of dispersal, all particle positions deeper than 2000 m (determined by a depth tracer) were flagged particularly. Those particles which, on drifting day 14 (i.e. the last day of the typical pelagic larval duration phase) were found in areas where the seafloor lies below 2000 m, were considered as 'lost' to the open ocean. In this study we present '100 % -areas' and '95 %- areas' of larval drift. In the latter case, the area covered by the larvae drift refers to the area including 95 % of all observed particle positions within 14 days when positions are binned in boxes of 5 km x 5 km onto a geographic grid. Due to the chosen bin size the number of particle positions considered to calculate this area is, however, up to 0.6 % lower than 95 %. The connectivity matrix shows the probability that a particle is present at least once in one of the three regions during the last seven days of its drift relative to the number of all particles released. All oceanographic modelling and visualising was performed within Python (version 3.7.3) running (inside Jupyter notebooks; Kluyver *et al.*, 2016) on an Unix system.

3. Results

3.1 Sponge taxonomy

Phakellia hirondellei and the TaP clade individuals are morphologically very similar in their external appearance and can be difficult to distinguish even to the expert eye. Although similar, the spicule content of both lineages differed slightly: TaP clade individuals possessed two types of strongyles and two sizes of styles, and *P. hirondellei* contained one type of strongyles, two sizes of oxeas and one size of styles (**Supplementary Table S1**). Only two individuals within *P. hirondellei* showed small differences in the spicule content (lack of strongyles, **Supplementary Table S1**). The molecular markers *18S* and *COI* were then used to investigate their phylogenetic placement and to confirm their taxonomic assignment in combination with traditional spicule analyses. In the *COI* analysis, the ten individuals identified based on the spicule content as *P. hirondellei* (Topsent, 1890; Bubarida order) were grouped in a monophyletic clade showing relatively moderate support (**Supplementary Material S1B**), with *Phakellia robusta* as sister clade (**Supplementary Material S1B**). However, these same individuals were grouped together in a clade with *P. robusta* based on the phylogenetic analysis of *18S* (**Supplementary Material S1A**) because of the low resolution power of this marker for sponge phylogeny at species level. Despite multiple sequencing attempts, generation of clean host sequences turned out to be difficult for the two sponge groups. Working with one single sequence direction for *18S*, we observed a few nucleotide mismatches within species, which could not be properly checked, but most likely do not reflect a true variation of the marker (as concluded based on the combination of the four different approaches used to decipher sponge taxonomy). The spicule analysis for all these ten individuals (**Supplementary Material S1C**) supported the assignation to *P. hirondellei*, despite small differences in the spicule content of two individuals. Within the same clade as *P. hirondellei* and *P. robusta*, we found other species which are assigned to either the order Axinellida or Bubarida (**Supplementary Material S1B).** The other 11 sponge individuals were classified as belonging to the TaP clade. 18S sequencing suggested that this clade included *Topsentia* and *Petromica* species (**Supplementary Material S1A).** In the *COI* analysis, the robustly supported TaP clade appeared as sister to another clade containing *Petromica*, *Cymbastela,* and *Ciocalypta* (**Supplementary Material S1B**). The remaining 21 sponges could not be unambigiously taxonomically classified, which is why we discarded the respective samples from further analyses of connectivity.

3.2 Sponge microbial community composition

The microbiomes of TaP clade specimens and *P. hirondellei* clustered apart from seawater in a non-metric multidimensional scaling plot of weighted UniFrac distances (**Figure 2A**). In addition, microbiomes of TaP clade individuals and *P. hirondellei* fell into two distinct clusters based on weighted UniFrac distances. Alpha diversity indices were highest in seawater, intermediate in TaP clade specimens and lowest in *P. hirondellei* (**Figure 2A**, small insert). Further, the microbial community composition was significantly different on ASV-level between TaP clade specimens and *P. hirondellei*, and also different from seawater (PERMANOVA, p=0.001, **Table 1**). Notably, the seawater reference microbiomes did not differ between the three sampling regions in terms of diversity and community composition (**Supplementary Material S2A-B**).

Table 1 Overview of pairwise-comparisons (PERMANOVAs) based on weighted UniFrac matrixes of microbial communities between sponge clades (TaP clade and *P. hirondellei*) and sample types (TaP clade, *P. hirondellei*, seawater).

Group 1	Group 2	Sample size	Per-mutations	pseudo-F	p-value
TaP clade	*Phakellia hirondellei*	21	999	107.04	0.001
TaP clade	Seawater	20	999	26.56	0.001
Phakellia hirondellei	Seawater	19	999	94.13	0.001

In terms of bacterial abundance and composition, the microbiomes of the TaP clade specimens and *P. hirondellei* were characteristic of low microbial abundance (LMA) sponges (according to Hentschel *et al.*, 2006; Gloeckner *et al.*, 2014; Moitinho-Silva *et al.*, 2017b) with dominant Gammaproteobacteria and Bacteroidetes (**Figure 2B**). However, the relative abundances of these microbial phyla differed between TaP clade specimens and *P. hirondellei*. While the phylum Nitrospirae was significantly enriched in *P. hirondellei*, TaP clade individuals showed a significant enrichment of Alphaproteobacteria and Bacteroidetes (**Supplementary Material S3**).

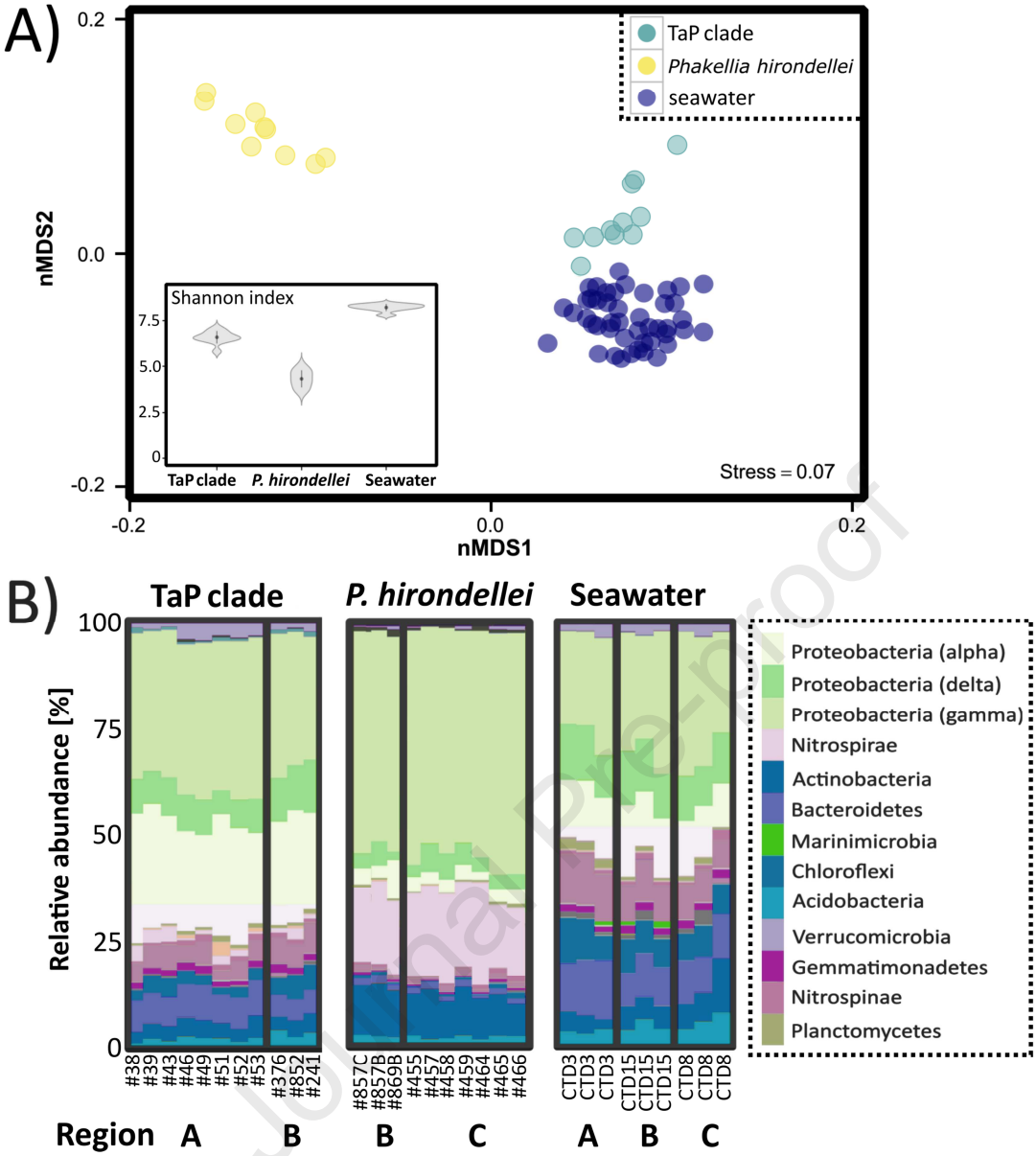

Figure 2 Microbial community composition and richness of different sample types. A) Non-metric multidimensional scaling plot on weighted UniFrac distances. Each marker is one microbial community, with colors indicating sample type (the TaP clade, *P. hirondellei* and seawater). Data of 21 sponge individuals and seawater samples from 15 stations are presented. Small plot inside Fig 3A presents Shannon indices of TaP clade individuals, *P.hirondellei* and seawater. Dots show means and whiskers represent standard deviations. In Fig.3B the microbial community composition is shown on the phylum level across all three sampling sites (A:'Canyon', B:'West-Bank', C:'East-Bank') for the TaP clade, *P. hirondellei* and seawater.

In a nMDS approach, we observed a substructuring of TaP clade microbiomes between region A ('Canyon') and region B ('West-Bank') (**Figure 3A**), which was statistically significant (PERMANOVA, p=0.011, **Table 2**). *Phakellia hirondellei* microbiomes showed no significant spatial differences between region B ('West-Bank') and region C ('East-Bank') (**Figure 3B**). In terms of alpha diversity indices (Shannon index) the TaP clade and *P. hirondellei* did not show any spatial differences between sampling region **(Figures 3A-B**, small inserts).

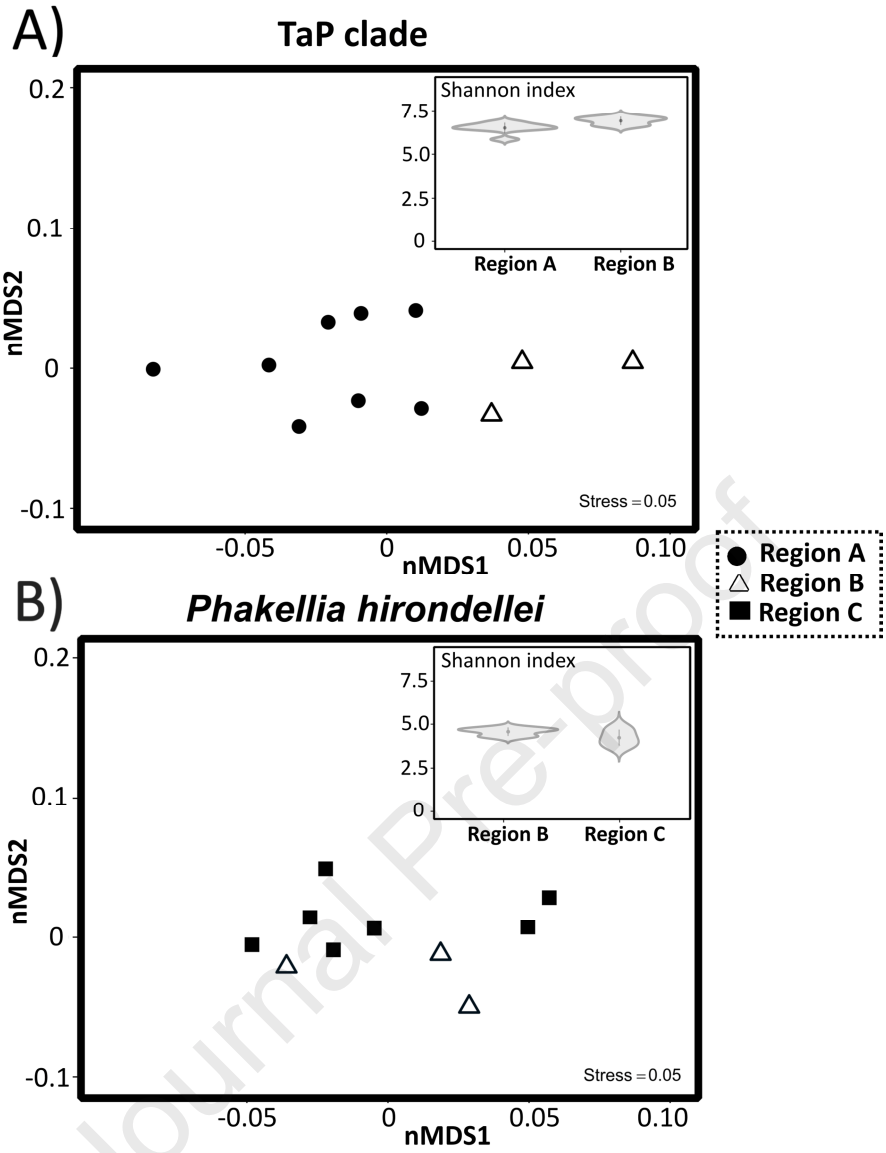

Figure 3 Microbial community composition of the TaP clade and *P. hirondellei* across the three sampling regions (A:'Canyon', B:'West-Bank', C:'East-Bank'). Non-metric multidimensional scaling (Figs 3 A-B) is based on weighted UniFrac distances (ASV-level) and symbols indicate the respective sampling regions. The violin plots presented within Figs. 3A-B show mean (dot) and standard deviation (whiskers) of Shannon indices for the two sponge clades across sites.

Table 2 Overview of pairwise-comparisons (PERMANOVAs) based on weighted UniFrac matrixes of microbial communities between sampling regions (A:'Canyon', B:'West-Bank', C:'East-Bank') for the TaP clade and *P. hirondellei*.

Group 1	Group 2	Sample size	Per-mutations	pseudo-F	p-value
TaP clade (RegionA)	TaP clade (RegionB)	11	999	3.59	0.011
Phakellia hirondellei (RegionB)	*Phakellia hirondellei* (RegionC)	10	999	1.61	0.144

3.3 Sponge population genetics

For population structure and connectivity (SNPs) analyses a total of 3,408 and 2,119 SNPs were obtained for the *P. hirondellei* and TaP clade datasets, respectively. Population genetic statistics for the two sponge clades are given in **Table 3**. Overall, the expected heterozygosity (H_e), generally considered as a measure of genetic diversity, was slightly lower for TaP clade specimens (0.369) than for *P. hirondellei* (0.400). Within sampling regions, H_e ranged from 0.366 (A:'Canyon') to 0.381 (B:'West-Bank') for TaP clade specimens, and from 0.394 (C:'East-Bank') to 0.408 (B:'West-Bank') for *P. hirondellei* (**Table 3**). Observed heterozygosity (H_O) was again lower for the TaP clade (0.535) than for *P. hirondellei* (0.601). Within sampling regions, H_O ranged from 0.532 (A:'Canyon') to 0.557 (B:'West-Bank') for TaP clade specimens, and from 0.597 (B:'West-Bank') to 0.603 (C:'East-Bank') for *P. hirondellei* (**Table 3**). All values for the inbreeding coefficient (F_{IS}) were negative and significant for the two sponge clades and the different regions analyzed, indicating an excess of observed heterozygotes (**Table 3**).

Table 3 Population genetic statistics for the TaP clade and *P. hirondellei*, grouping samples per sampling region. The abbreviations are as follows: n (number of samples), Ho ('observed heterozygosity'), He ('expected heterozyogosity') (= genetic diversity), Fis ('inbreeding coefficient'). Significant values are depicted in bold.

Sponge clade / Sampling region	n	H_O	He	F_{IS}
TaP clade				
A: 'Canyon'	8	0.532	0.366	**-0.454**
B: 'West-Bank'	3	0.557	0.381	**-0.463**
Total	11	0.535	0.369	**-0.452**
Phakellia hirondellei				
B: 'West-Bank'	3	0.597	0.408	**-0.465**
C: 'East-Bank'	7	0.603	0.394	**-0.531**
Total	10	0.601	0.400	**-0.504**

The results of the *STRUCTURE* and *adegenet* analyses revealed no significant genetic structure among the samples for the two sponge clades from the different regions (**Figure 4A-B**). This was also corroborated by the F_{ST} values for the pairwise comparisons between regions, which was 0.023 for the comparison between TaP clade regions (*p*-value = 0.06) and 0.015 for the comparison between *P. hirondellei* regions (*p*-value = 0.169). Interestingly, the DAPC analysis for which samples were grouped per sampling region, did detect differences across sampling regions for the TaP clade, which showed a subtle separation between regions A ('Canyon') and B ('West-Bank'), (**Figure 4C**; note that DAPC analyses are generally more sensitive than STRUCTURE and adegenet analyses). In contrast, for *P. hirondellei*, we did

not observe this differentiation, but samples of the canyon region (region A) were not available for this species (**Figure 4D**).

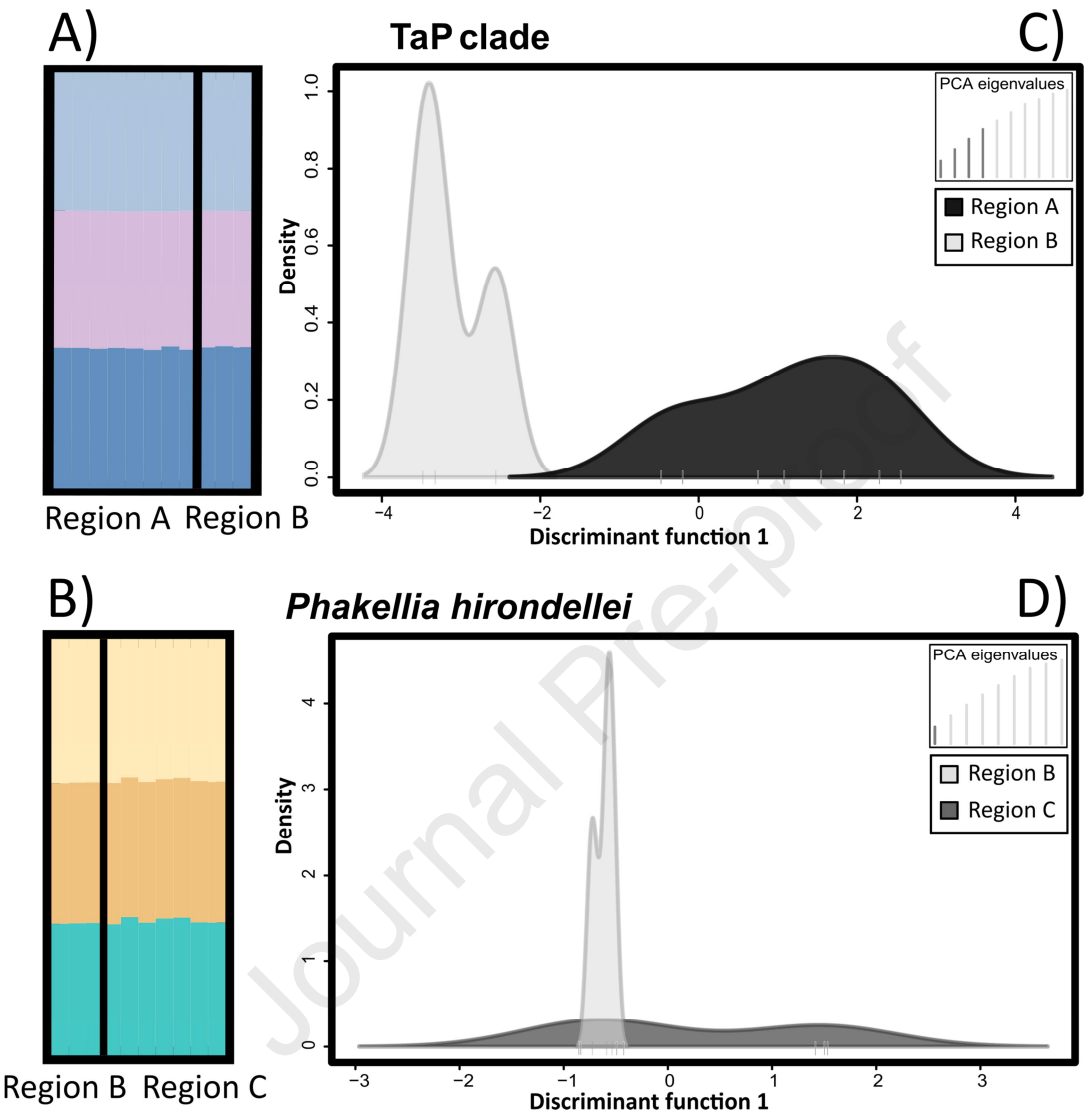

Figure 4 Sponge population genetics based on individual genotype assignments of TaP clade and *P. hirondellei* individuals to clusters (K) as inferred by STRUCTURE (Fig. 4A-B). Spatial structure of the two sponge clades across the three sampling regions (A:'Canyon', B:'West-Bank', C:'East-Bank'), as inferred by DAPC (Figs. 4C-D).

3.4 Validation of oceanographic modelling setup

In order to link population genetic data with oceanographic connectivity, VIKING20X was chosen as our model to simulate drifting of sponge larvae. The optimal number of particles to be released was determined as 1000 particles per grid box and day based on the saturation plateau of computed rarefraction curves (**Supplementary Material S4**). When determining the optimal model starting date, we observed that current direction and velocities seemed to be governed by mesoscale processes over the whole year (**Supplementary Material S5**). We

chose June as the starting month to release virtual sponge larvae into the water column; and due to strong temporal variability we performed releases over 10 different years. The month of June was chosen because it is two months after the main phytoplankton bloom in 2017 (**Supplementary Material S6**) and it is during a period when ripe reproductive stages were found in both, TaP clade and *P. hirondellei* individuals. In this month, the *in situ* chlorophyll a concentrations were comparably low at all three locations (**Supplementary Material S7A-B**) and corresponded well to the satellite data. Both, TaP clade members and *P. hirondellei* were found to contain mature gametes in mid-June, when they were collected. In TaP clade individuals, vitellogenic oocytes of ~ 40-60 μm were found close to the canals, indicating that they were relatively mature and almost ready to be released in the water column (**Supplementary Material S8A**). The oocytes were full of nutrients of proteinaceous origin (**Supplementary Material S8B-C**) and heterogenous yolk, a mix of lipid and protein (**Supplementary Material S8B, D**). In *P. hirondellei,* similar vitellogenic oocytes were observed (~ 50 μm), probably mature and ready to be released in the water column (**Supplementary Material S8E**). In contrast to TaP clade individuals, in *P. hirondellei* the nutrient reserves within the oocytes seemed to be homogeneous yolk comprised only by protein platelets (≤ 500 nm) (**Supplementary Material S8E-F**).

The three release and sampling regions (A ('Canyon'), B ('West-Bank'), C ('East-Bank')), lay within a similar depth range (**Supplementary Material S7**) and sampling at all three regions occurred within the same watermass (**Supplementary Material S9A**). *In situ* temperature and salinity measurements fitted overall well to the data extracted from VIKING20X in terms of their vertical profile (**Supplementary Material S9A**) and their rough characteristics at the bottom depths of the CTD sampling sites (**Supplementary Material S9B**). However, we noted a slight shift to lower salinities in the VIKING20X data in comparison to the *in situ* measurements. On a different note, VIKING20X does not resolve hydrodynamics inside canyons and further does not simulate important small-scale properties (such as tides), which could lead to an overestimation of connectivity between the individual regions. In our experiments we released virtual larvae at the 'opening' of the canyon (but at the depth level that corresponds to the *in situ* sampling depths). For our similuations we considered particles drifting on day 14 in regions where the seafloor is below 2,000 m water depth, as 'lost' to the open ocean. This configuration rests upon the evidence-based assumption that fan-shaped sponge abundance in the study area is very low below 2,000 m water depth: With the aid of a photogrammetic sledge (ROTV Politolana), fan-shaped sponges at the study area were recorded particularly within the mesopelagic zone (Sánchez *et*

al., 2017), especially in a depth range between 425–1300 m, with highest densities between 450–550 m.

3.5 Larvae drift model

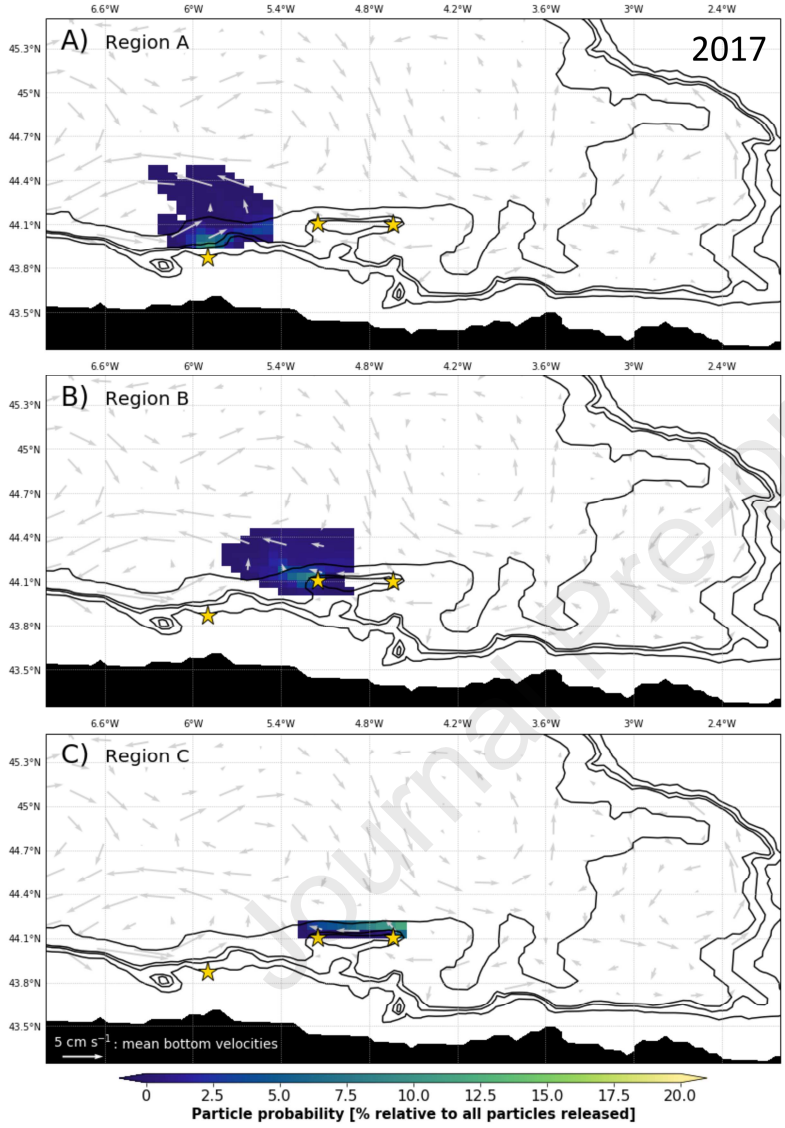

Figure 5 Output of the Lagrangian simulations showing particle probabilities in space (in % relative to all particles released) after a 14 days drifting time and a daily release of larvae over a period of 30 days in June 2017 (which was determined as suitable model start point during model validation procedure) for the three regions (A:'Canyon', B:'West-Bank', C:'East-Bank'). Bathymetry contours are shown by black lines and the Spanish coastline is indicated by the black area. Mean monthly current velocities for the analysed period are indicated by grey arrows.

For all three regions a considerable fraction of the released particles stayed within their release region for all analysed years and inter-regional connectivity was low. Passive-particle drift trajectories showed a preferential retention of simulated larvae within the study area (for a particular year see **Figure 5** and **Supplementary Material S10**; for multi-year data see **Supplementary Material S11**). Most virtual particles travelled along the bathymetry and

were strongly influenced by the mesoscale variability, i.e. the existence and direction of ocean eddies. Across all years and sampling regions on drifting day 14, on average 44.2 ± SE 5.7 % of all particles were found in areas where the seafloor is located below 2,000 m, but 0 % of the particles were actually drifting at water depths below 2,000 m. We observed strong inter-annual variations (between 2009 and 2018) in the geographic location of eddies (**Supplementary Material S12**), as well as in the average larval drift distances, and in the 100 % -areas and 95 %- areas of larval drift (**Supplementary Material S13**). These inter-annual variations were observed at all three sampling regions. To assess inter-annual patterns of connectivity, we computed a connectivity matrix (**Figure 6**). Throughout the ten analysed years region A ('Canyon') and region C ('East-Bank') were never connected. In less than half of the analysed years, region B ('West-Bank') received particles from the other two regions. Only in two years (2009 and 2011) particles from region B were transported to another region (only region C ('East-Bank')).

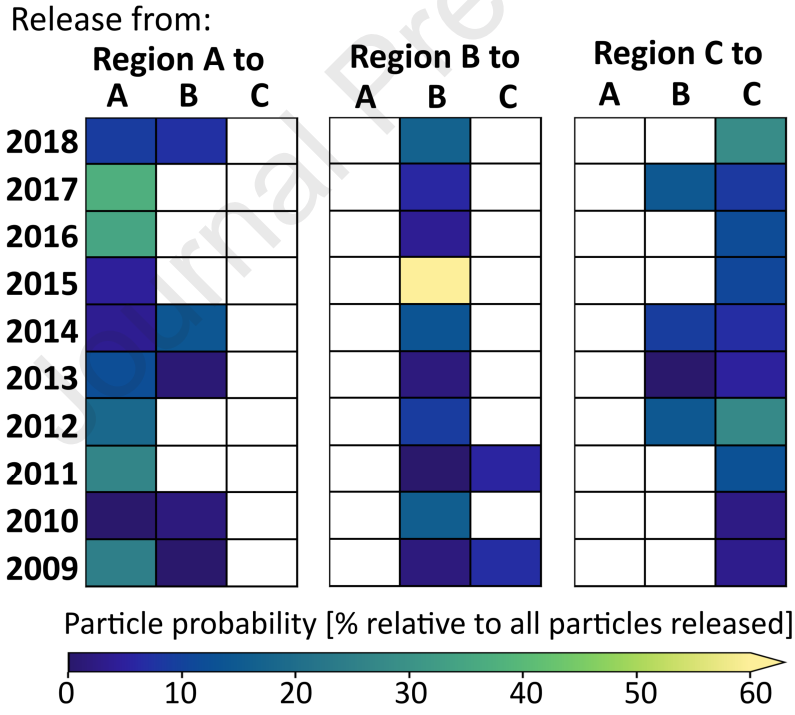

Figure 6 Connectivity matrix showing the probability that a particle is present at least once in one of the three regions (A:'Canyon', B:'West-Bank', C:'East-Bank') during the last seven days of its drift relative to the number of all particles released. The inter-annual variablilty across the ten analysed years (2009-2018) is illustrated.

4. Discussion

Almost two-thirds of the Earth are considered deep-sea, which is also the least explored habitat on Earth (Costello *et al.*, 2010; Ramirez-Llodra *et al.*, 2010; Ramirez-Llodra *et al.*, 2011). Designing population genetic studies in such understudied regions can be challenging.

In the present study half of all originally collected sponge specimens could not be used for population genetic evaluation, because the taxonomic assignment remained ambigious despite considerable efforts using both molecular (*COI, 18S, 16S*) and classical taxonomic markers (spicules). Further, even for specimens identified as belonging to the same species, it can be difficult to assign a taxonomic name. Indeed, sponge systematics and phylogenic reconstructions continue to be difficult using traditional spicule analyses and/or molecular techniques (Erpenbeck *et al.*, 2006; Philippe *et al.*, 2011). While in our study individuals belonging to the TaP clade were identified as belonging to the same species based on *COI*, *18S*, and spicule analyses, their assignment to a described order remains wanting. In order to be able to provide an official taxonomic name, a complete review of the orders to which *Topsentia* and *Petromica* (and other groups) belong will need to be accomplished.

Although the encountered difficulties in taxonomic identification led to a considerable reduction of our sampled dataset, we emphasize that a critical re-evaluation of sponge taxonomy needs to be at the basis of every population genetic analysis, as morphologically very similar specimens (i.e. specimens with a similar external body shape) may belong to different sponge orders. The microbial community composition based on *16S* amplicon data was identified as a helpful tool to support sponge taxonomy since the microbiomes were host-specific. Indeed, for many organisms phylogenetic relatedness of the host can strongly affect the associated microbial community composition (reviewed in Moran *et al.*, 2019). For sponges, species-specific prokaryotic communities are commonly found (e.g. Easson and Thacker, 2014; Thomas *et al.*, 2016; Steinert *et al.*, 2017). Besides host-specific microbiomes, both, TaP clade individuals and *P. hirondellei* showed microbial signatures typical for Low Microbial Abundance (LMA) sponges, as described in (Moitinho-Silva *et al.*, 2017b), which differed in terms of relative abundances of microbial phyla. Results of our study further imply that on a subspecies level (genotypes obtained using SNPs), the sponge microbiome seems to be a sensitive parameter for evaluating emerging taxonomic structures of sponge holobionts, and particularly so when the ASV – level is considered for microbial analyses.

Choosing the appropriate spatial scale to identify potential genetic structures of deep-sea sponges can be difficult since knowledge on population genetics/genomics of deep-sea organisms is sparse. In order to cover a maximum geographic range, the population genetic pattens the fan-shaped sponges *P. robusta* and *P. hirondellei* are examined over a few thousand kilometers in a complementary paper by Taboada *et al.* (*in review*), while the present study focussed on distances of < 100 km. For deep-sea organisms, generally low genetic structuring is reported when organisms occur at similar depths (McClain and Hardy,

2010; Taylor and Roterman, 2017), while other studies highlight the importance of biophysical barriers with respect to inhibition of gene flow. A recent meta-analysis on meiofaunal biodiversity covering 21 canyons together with reference stations on the adjacent slopes, observed a high dissimilarity in taxonomic composition between canyons and slopes (Bianchelli and Danovaro, 2019). At a population genetic level, the red gorgonian *Paramuricea clavata* showed a strong genetic differentiation among populations in different submarine canyons in the Mediterranean Sea, presumably due to limited effective dispersal (Pérez-Portela *et al.*, 2016). Similarly, another recent study found genetic differentiation among cold-water coral populations across sites in the near neighbourhood (Boavida *et al.*, 2019).

In the present study, we mostly observed genetically well-connected sponge populations for both, the TaP clade, and *P. hirondellei*. This result was consistent across the STRUCTURE analysis (**Figure 4 A-B**), *adegenet* analysis, and Fst values. Interestingly, individuals belonging to the TaP clade showed a subtle separation between sampling regions A ('Canyon') and B ('West-Bank') in the DAPC analysis (**Figure 4 C**), despite the analysis detected an overall single panmictic population. Similarly, we observed a significant difference of TaP clade microbiomes between sampling regions A ('Canyon') and B ('West-Bank') in terms of weighted UniFrac distances. Seawater microbiomes did however not differ between the sampling regions. This supports our oceanographic *in situ* observations (TS data) that sampling at all three sampling regions occurred within the same watermass and under similar environmental conditions. We thus conclude that slight but distinct substructure in host population genetics and microbial community composition occur concomitantly. We further conclude that although gene flow generally occurs between both examined geologic features (the Avilés Canyon System and the Le Danois Bank), the exchange may be low enough to promote subtle genetic differentiation. It is quite remarkable that slight differentiation can even be found on scales below 100 km distance, and interestingly on both, the host and microbiome level. For *P. hirondellei*, we did not observe a regional differentiation, but samples of the canyon region (region A) were not available for this species. Our results are overall consistent with a previous lack in strong genetic structuring in deep-sea organisms occurring at similar depths (McClain and Hardy, 2010; Taylor and Roterman, 2017).

Our performed simulations of Lagrangian connectivity suggest that virtual sponge larvae stay in the study area and ocean currents generally allow local gene flow events, despite inter-annual variations in the connectivity between the three regions. For sessile

759 invertebrates, 10–100 km distance is the typical range within which the majority of off-spring
760 is transported away from its parents (Palumbi, 2004). Particle trajectories calculated within
761 our study reveal a similar range for TaP clade individuals and *P. hirondellei*. Larval drifting
762 occurred in swirled patterns and exhibited a variable strength across different years, as the
763 major circulation patterns seemed to be governed by mesoscale processes. Overall, ocean
764 currents seem to promote virtual larvae to drift along the bathymetry in the study area. Our
765 results thus suggest that, although the continental shelf is comparably narrow around the
766 northern shores of Spain, ocean circulation may prevent the loss of sponge larvae towards the
767 open ocean. Whether this will still be the case under future climate scenarios remains to be
768 tested in future studies. Interestingly, in a recent study by Fox *et al.*, (2016) trajectories of
769 cold-water coral larvae were found to be strongly correlated to the North Atlantic Oscillation.
770 The authors of that study conclude that the connectivity of the existing MPA network is
771 vulnerable to atmospheric-driven changes in ocean circulation.

772 Besides the need of further studies on the variability of the physical settings, also
773 further biological data is urgently needed to improve understanding of connectivity in the
774 deep-sea, especially for key organisms like sponges and the ecosystems they form. As current
775 understanding of deep-sea larval behavior is extremely limited (in general) to non-existant
776 (for deep-sea sponges), our particle trajectory calculations are based on the assumption of
777 passive dispersal only. However, recent studies have shown that active larval movements can
778 have an order of magnitude impact on dispersal (Gary *et al.*, 2020). It will therefore be crucial
779 for future studies to gather more information on the ecology and behaviour of deep-sea
780 sponge larvae.

781 In summary, we show that canyons can promote slight genetic divergence of sponges
782 and also of their microbiomes, although more data would be needed to provide further
783 statistical support for this observation. We further conclude that the specific prevailing ocean
784 currents of the Cantabrian Sea dictate sponge host genetics by promoting a settlement in
785 suitable areas within kilometers from their parent sponges and consequently allowing gene
786 flow. However oceanographic connectivity showed a distinct variability between the
787 simulated years. While genetic uniformity would occur in times of connectivity between the
788 Avilés Canyon and the Le Danois Bank, genetic differentiation might happen in years of
789 disconnection, when the deep-water currents promote drifting of larvae in different directions.
790 These findings contribute to a better understanding of connectivity, genetic and microbial
791 variability of deep-sea organisms. Our results imply that a design of ecologically coherent
792 networks of MPAs is no trivial task for deep-sea sponge holobionts. We conclude for our

study site (on scales < 100 km) that (i) physical connectivity between close-by sites via ocean currents has a strong temporal component, and that (ii) differences between close-by sites are stronger reflected in the sponge-associated microbial community than in the concomitant sponge-host genetics. These points reveal that it is important for a design of ecologically coherent networks of MPAs, to consider the naturally existing variability of connectivity (e.g. physical data from multiple years) and to consider the occurring microbial diversity in addition to the macrobial diversity.

6. Acknowledgements

We dedicate this study to Hans Tore Rapp, friend, mentor and esteemed colleague whose outstanding expertise in deep-sea sponge taxonomy will be dearly missed. We greatly acknowledge the crew and scientific party of RV Ángeles Alvariño cruise SponGES0617 for their valuable support at sea. As well as members of the IEO Gijón for logistical support during a guest stay of KB before the cruise. We appreciate Andrea Hethke's and Ina Clefsen's, as well as Thomas Hansen's assistance in the laboratory after the cruise while generating the microbial amplicon and flow cytometry data. We thank Willi Rath for support in technical issues with the modelling part, and Lara Schmittmann and Ina Clefsen for technical support with *18S* and *COI* sequencing. Paco Cárdenas provided valuable input with regard to the sponge taxonomic identification and on a first draft of the manuscript, which was very much appreciated. We further thank two anonymous reviewers for their valuable input during the review process.

7. Author contributions

UH, AR, AB, KB designed the study. KB, AR, ST, VK, PR, JC participated in sampling. PR, JC, AR conducted sponge taxonomic analysis. ST, AR performed sponge population genetic analyses, while KB, UH performed microbial analyses. CS, KB, AB designed and performed Lagrangian modelling. KG provided VIKING20X data. VK and KB did microscopy of sponge tissue. KB performed assessment of environmental conditions *in situ* and remotely. AF was involved in sequencing of microbial samples. KB and UH wrote the manuscript, and the other authors reviewed and edited it.

8. References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**:403-410.

827 Andrello, M., Guilhaumon, F., Albouy, A., Parravicini, V., Scholtens, J., Verley, P., Barange,
828 M., et al. (2017) Global mismatch between fishing dependency and larval supply from
829 marine reserves. *Nat. Commun.* **8**:16039.
830 Baltazar-Soares, M., Biastoch, A., Harrod, C., Hanel, R., Marohn, L., Prigge, E., et al. (2014)
831 Recruitment collapse and population structure of the european eel shaped by local ocean
832 current dynamics. *Curr Biol* **24**: 104–108.
833 Beazley, L., Kenchington, E., Murillo, F.J., del Mar Sacau, M. (2013) Deep-sea sponge
834 grounds enhance diversity and abundance of epibenthic megafauna in the Northwest
835 Atlantic. *ICES Journal of Marine Science* **70**:1471-1490.
836 Beugin, M.P., Gayet, T., Pontier, D., Devillard, S., and Jombart, T. (2018) A fast likelihood
837 solution to the genetic clustering problem. *Methods Ecol Evol* **9**: 1006–1016.
838 Bianchelli, S. and Danovaro, R. (2019) Meiofaunal biodiversity in submarine canyons of the
839 Mediterranean Sea: A meta-analysis. *Prog Oceanogr* **170**: 69–80.
840 Boavida, J., Becheler, R., Choquet, M., Frank, N., Taviani, M., Bourillet,J.-F., Meistertzheim,
841 A.-L., Grehan, A., Savini, A., Arnaud-Haond,S. (2019) Out of the Mediterranean? Post-
842 glacial colonization pathways varied among cold-water coral species. *Journal of*
843 *Biogeography* **46**:915–931.
844 Bolyen, E., Rideout, J.R., Dillon, M.R. et al. (2019) Reproducible, interactive, scalable and
845 extensible microbiome data science using QIIME 2. *Nat Biotechnol* **37**: 852–857.
846 Böning, C.W., Behrens, E., Biastoch, A., Getzlaff, K., and Bamber, J.L. (2016) Emerging
847 impact of Greenland meltwater on deepwater formation in the North Atlantic Ocean. *Nat*
848 *Geosci* **9**: 523–527.
849 Breckenfelder, T., Rhein, M., Roessler, A., Böning, C.W., Biastoch, A., Behrens, E., and
850 Mertens, C. (2017) Flow paths and variability of the North Atlantic Current: A
851 comparison of observations and a high-resolution model. *J Geophys Res Ocean* **122**:
852 2686–2708.
853 Breusing, C., Biastoch, A., Drews, A., Metaxas, A., Jollivet, D., Vrijenhoek, R.C., et al.
854 (2016) Biophysical and Population Genetic Models Predict the Presence of “Phantom”
855 Stepping Stones Connecting Mid-Atlantic Ridge Vent Ecosystems. *Curr Biol* **26**: 2257–
856 2267.
857 Busch,K., Hanz, U., Mienis, F., Mueller, B., Franke, A., Roberts, E.M., Rapp, H.T.,
858 Hentschel, U. (2020a) On giant shoulders: how a seamount affects the microbial
859 community composition of seawater and sponges. *Biogeosciences* **17**:3471–3486.
860 Busch, K., Beazley, L., Kenchington, E., Whoriskey, F, Slaby, B.M., Hentschel, U. (2020b)

Microbial diversity of the glass sponge *Vazella pourtalesii* in response to anthropogenic activities. *Conservation Genetics*, doi:10.1007/s10592-020-01305-2.

Busch, K., Wurz, E., Rapp, H.T., Bayer, K., Franke, A., Hentschel, U. (2020c) Chloroflexi dominate the deep-sea golf ball sponges *Craniella zetlandica* and *Craniella infrequens* throughout different life stages. Front. Mar. Sci. **7**:674.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Knight, R., Lozupone, C.A., et al. (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci* **108**: 4516–4522.

Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., and Cresko, W.A. (2013) Stacks: An analysis tool set for population genomics. *Mol Ecol* **22**: 3124–3140.

Cathalot, C., van Oevelen, D., Cox, T. J. S., Kutti, T., Lavaleye, M., Duineveld, G., et al. (2015) Cold-water coral reefs and adjacent sponge grounds: hot spots of benthic respiration and organic carbon cycling in the deep sea. *Front. Mar. Sci.* **2**:37.

Combosch, D.J., Lemer, S., Ward, P.D., Landman, N.H., and Giribet, G. (2017) Genomic signatures of evolution in Nautilus—An endangered living fossil. *Mol Ecol* **26**: 5923–5938.

Costello, M.J., Cheung, A., and De Hauwere, N. (2010) Surface area and the seabed area, volume, depth, slope, and topographic variation for the world's seas, oceans, and countries. *Environ Sci Technol* **44**: 8821–8828.

Cowen, R.K. and Sponaugle, S. (2009) Larval Dispersal and Marine Population Connectivity. *Ann Rev Mar Sci* **1**: 443–466.

Darriba, D., Taboada, G.L., Doallo, R. and Posada, D. (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature methods* **9(8)**:772-772.

Delandmeter, P. and Van Sebille, E. (2019) The Parcels v2.0 Lagrangian framework: New field interpolation schemes. *Geosci Model Dev* **12**: 3571–3584.

Earl, D.A. and vonHoldt, B.M. (2012) STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* **4**: 359–361.

Easson, C.G. and Thacker, R.W. (2014) Phylogenetic signal in the community structure of host-specific microbiomes of tropical marine sponges. *Front Microbiol* **5**: 532.

Erpenbeck, D., Breeuwer, J.A.J., Parra-Velandia, F.J., and Van Soest, R.W.M. (2006) Speculation with spiculation? - Three independent gene fragments and biochemical characters versus morphology in demosponge higher classification. *Mol Phylogenet Evol* **38**: 293–305.

De Forges, B.R., Koslow, J.A., and Poore, G.C.B. (2000) Diversity and endemism of the benthic seamount fauna. *Nature* **405**: 26–29.

Díez-Vives, C., Taboada, S., Leiva, C., Busch, K., Hentschel, U., Riesgo, A. (in press) On the way to specificity - Microbiome reflects sponge genetic cluster primarily in highly structured populations. *Molecular Ecology.*

Easson, C. G., Chaves-Fonnegra, A., Thacker, R. W., Lopez, J. V. (2020) Host population genetics and biogeography structure the microbiome of the sponge Cliona delitrix. *Ecology and Evolution* **205(22)**:3505–14.

FAO (2009) The FAO international guidelines for the management of deep-sea fisheries in the High Seas. *FAO Fisheries and Aquaculture Department*, Rome.

FAO (2016) Vulnerable Marine Ecosystems: Processes and practices in the High Seas. In: Thompson, A., Sanders, J., Tandstad, M., Carocci, F., Fuller, S. (eds), *FAO Fisheries and Aquaculture Technical Paper* 595, Rome.

Folmer, O., Black, M., Hoeh, W., Lutz, R., and Vrijenhoek, R. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol* **3(5)**: 294–299.

Fox, A.D., Henry, L.-A., Corne, D.W., Roberts, J.M. (2016) Sensitivity of marine protected area network connectivity to atmospheric variability. *Royal Society Open Science* **3**:160494.

Gallego, A., Gibb, F.M., Tullet, D., Wright, P.J. (2017) Bio-physical connectivity patterns of benthic marine species used in the designation of Scottish nature conservation marine protected areas. *ICES J. Mar. Sci.* **74**:1797–1811.

Gary S.F., Fox A.D., Biastoch A., Roberts J.M., Cunningham S.A. (2020) Larval behaviour, dispersal and population connectivity in the deep sea. *Scientific Reports* **10**:10675.

Gloeckner, V., Wehrl, M., Moitinho-Silva, L., Schupp, P., Pawlik, J.R., Lindquist, N.L., et al. (2014) The HMA-LMA Dichotomy Revisited : an Electron Microscopical Survey of 56 Sponge Species. *Biol Bull* **227**: 78–88.

Gómez-Ballesteros, M., Druet, M., Muñoz, A., Arrese, B., Rivera, J., Sánchez, F., et al. (2014) Geomorphology of the Avilés Canyon System, Cantabrian Sea (Bay of Biscay). *Deep Res Part II Top Stud Oceanogr* **106**: 99–117.

Griffiths, S. M., Antwis, R. E., Lenzi, L., Lucaci, A., Behringer, D. C., Butler IV, M. J., Preziosi, R. F. (2019) Host genetics and geography influence microbiome composition in the sponge Ircinia campana. *Journal of Animal Ecology* **88(11)**:1684–1695.

Harrington, B. and Team, and the D. (2005) Inkscape. Available online at:

http://www.inkscape.org/.

Hentschel, U., Usher, K.M., and Taylor, M.W. (2006) Marine sponges as microbial fermenters. *FEMS Microbiol Ecol* **55**: 167–177.

Hentschel, U., Piel, J., Degnan, S.M., Taylor, M.W. (2012) Genomic insights into the marine sponge microbiome. *Nat Rev Microbiol* **10**: 641-654.

Hickey, B., Baker, E., and Kachel, N. (1986) Suspended particle movement in and around Quinault submarine canyon. *Mar Geol* **71**: 35–83.

Highfield, J.M., Eloire, D., Conway, D. V. P., Lindeque, P. K., Attrill, M. J., Somerfield, P. J. (2010) Seasonal dynamics of meroplankton assemblages at station L4. *Journal of Plankton Research* **32(5)**:681–691.

Indraningrat, A. A. G., Micheller, S., Runderkamp, M., Sauerland,I., Becking, L. E., Smidt, H., Sipkema, D. (2019) Cultivation of sponge-associated bacteria from *Agelas sventres* and *Xestospongia muta c*ollected from different depths. Mar. Drugs 17:1–17.

Jakobsson, M. and Rosenberg, N.A. (2007) CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**: 1801–1806.

Jeffries, D.L., Copp, G.H., Handley, L.L., Håkan Olsén, K., Sayer, C.D., and Hänfling, B. (2016) Comparing RADseq and microsatellites to infer complex phylogeographic patterns, an empirical perspective in the Crucian carp, Carassius carassius, L. *Mol Ecol* **25**: 2997–3018.

Jombart, T. (2008) Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**: 1403–1405.

Jombart, T. and Ahmed, I. (2011) adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics* **27**: 3070–3071.

Jombart, T., Devillard, S., and Balloux, F. (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet* **11**:94.

Katoh, K. and Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol Biol Evol* **30**:772–780.

Kazanidis, G., Vad, J., Henry, L.-A., Neat, F., Berx, B., Georgoulas, K., Roberts, J.M. (2019) Distribution of deep-sea sponge aggregations in an area of multisectoral activities and changing oceanic conditions. *Front. Mar. Sci.* **6**:163.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012) Geneious Basic: An integrated and extendable desktop software platform for the

963 organization and analysis of sequence data. *Bioinformatics* **28**:1647-1649.
964 Kenchington, E., Wang, Z., Lirette, C., Murillo, F.J., Guijarro, J., Yashayaev, I., Maldonado,
965 M. (2019) Connectivity modelling of areas closed to protect vulnerable marine
966 ecosystems in the northwest Atlantic. *Deep-Sea Res. Part I* **143**:85-103.
967 Klitgaard, A.B. and Tendal, O.S. (2004) Distribution and species composition of mass
968 occurrences of large-sized sponges in the northeast Atlantic. *Prog Oceanogr* **61**: 57–98.
969 Klitgaard, A.B., Tendal, O.S., and Westerberg, H. (1997) Mass occurrences of large sponges
970 (Porifera) in Faroe Island (NE Atlantic) Shelf and slope areas: Characteristics,
971 distribution and possible causes. In, Hawkins, L.E. and Hutchinson, S. (eds), *Responses*
972 *of Marine Organisms to Their Environment*. Southampton, England: Proceedings of the
973 30th European Marine Biology Symposium, pp. 129-142.
974 Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B.E., Bussonnier, M., Frederic, J., et al.
975 (2016) Jupyter Notebooks - a publishing format for reproducible computational
976 workflows. In, Loizides, F. and Schmidt, B. (eds), *Positioning and Power in Academic*
977 *Publishing: Players, Agents and Agendas*. Amsterdam, The Netherlands: IOS Press, pp.
978 87–90.
979 Koutsouveli, V., Taboada, S., Moles, J., Cristobo, J., Ríos, P., Bertran, A., Solà, J., Avila, C.,
980 Riesgo, A. (2018) Insights into the reproduction of some Antarctic dendroceratid,
981 poecilosclerid, and haplosclerid demosponges. *PLoS ONE* **13(2)**: e0192267.
982 Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K., and Schloss, P.D. (2013)
983 Development of a dual-index sequencing strategy and curation pipeline for analyzing
984 amplicon sequence data on the miseq illumina sequencing platform. *Appl Environ*
985 *Microbiol* **79**: 5112–5120.
986 Levin, L.A. (2006) Recent progress in understanding larval dispersal: New directions and
987 digressions. *Integr Comp Biol* **46**: 282–297.
988 Madec, G. (2016) NEMO ocean engine. *Note Du Pôle Modélisation, Inst Pierre-Simon*
989 *Laplace* **406**.
990 Martín, J., Palanques, A., and Puig, P. (2006) Composition and variability of downward
991 particulate matter fluxes in the Palamós submarine canyon (NW Mediterranean). *J Mar*
992 *Syst* **60**: 75–97.
993 McClain, C.R. and Hardy, S.M. (2010) The dynamics of biogeographic ranges in the deep
994 sea. *Proc R Soc B Biol Sci* **277**: 3533–3546.
995 Meirmans, P.G. and Van Tienderen, P.H. (2004) GENOTYPE and GENODIVE: Two
996 programs for the analysis of genetic diversity of asexual organisms. *Mol Ecol Notes* **4**:

792–794.

Moitinho-Silva, L., Nielsen, S., Amir, A., Gonzalez, A., Ackermann, G.L., Cerrano, C., Astudillo-García, C., Easson, C., Sipkema, D., Liu, F., Steinert, G., Kotoulas, G., McCormack, G.P., Feng, G., Bell, J.J., Vicente, J., Björk, J.R., Montoya, J.M., Olson, J.B., Reveillaud, J., Steindler, L., Pineda, M.-C., Marra, M.V., Ilan, M., Taylor, M.W., Polymenakou, P., Erwin, P.M., Schupp, P.J., Simister, R.L., Knight, R., Thacker, R.W., Costa, R., Hill, R.T., Lopez-Legentil, S., Dailianis, T., Ravasi, T., Hentschel, U., Li, Z., Webster, N.S., Thomas, T. (2017a) The sponge microbiome project. *Gigascience* **6**: gix077.

Moitinho-Silva, L., Steinert, G., Nielsen, S., Hardoim, C.C.P., Wu, Y.C., McCormack, G.P., et al. (2017b) Predicting the HMA-LMA status in marine sponges by machine learning. *Front Microbiol* **8**: 1–14.

Moran, N.A., Ochman, H., and Hammer, T.J. (2019) Evolutionary and Ecological Consequences of Gut Microbial Communities. *Annu Rev Ecol Evol Syst* **50**: 451–475.

Morato, T., Hoyle, S.D., Allain, V., and Nicol, S.J. (2010) Seamounts are hotspots of pelagic biodiversity in the open ocean. *Proc Natl Acad Sci USA* **107**: 9707–9711.

Morrow, C. and Cárdenas, P. (2015) Proposal for a revised classification of the Demospongiae (Porifera). *Front Zool* **12**: 1–27.

Mulisch, M. and Welsch, U. (2015) Romeis Mikroskopische Technik Aufl. 19. Springer Spektrum, Heidelberg, Germany.

Murillo, F.J., Muñoz, P.D., Cristobo, J., Ríos, P., González, C., Kenchington, E., and Serrano, A. (2012) Deep-sea sponge grounds of the Flemish Cap, Flemish Pass and the Grand Banks of Newfoundland (Northwest Atlantic Ocean): Distribution and species composition. *Mar Biol Res* **8**: 842–854.

Muyzer, G., Waal, E.C.D.E., and Uitierlinden, A.G. (1993) Profiling of Complex Microbial Populations by Denaturing Gradient Gel Electrophoresis Analysis of Polymerase Chain Reaction-Amplified Genes Coding for 16S rRNA GERARD. *Appl Environ Microbiol* **59**: 1–6.

Palumbi, S.R. (2004) Marine reserves and Ocean neighborhoods: The Spatial Scale of Marine Populations and Their Management. *Annu Rev Environ Resour* **29**: 31–68.

Paris, J.R., Stevens, J.R., and Catchen, J.M. (2017) Lost in parameter space: A road map for Stacks. Meth Ecol Evol **8**:1360-1373.

Pérez-Portela, R., Cerro-Gálvez, E., Taboada, S., Tidu, C., Campillo-Campbell, C., Mora, J., and Riesgo, A. (2016) Lonely populations in the deep: genetic structure of red

gorgonians at the heads of submarine canyons in the north-western Mediterranean Sea. *Coral Reefs* **35**: 1013–1026.

Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., and Hoekstra, H.E. (2012) Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One* **7**: e37135.

Philippe, H., Brinkmann, H., Lavrov, D. V., Littlewood, D.T.J., Manuel, M., Wörheide, G., and Baurain, D. (2011) Resolving difficult phylogenetic questions: Why more sequences are not enough. *PLoS Biol* **9**: e1000602.

Pita, L., Rix, L., Slaby, B.M., Franke, A., and Hentschel, U. (2018) The sponge holobiont in a changing ocean: from microbes to ecosystems. *Microbiome* **6**: 46.

Pritchard, J.K., Stephens, M., and Donnelly, P. (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.

QGIS Development Team (2017) Geographic Information System. Open Source Geospatial Foundation Project. Available online at: http://qgis.osgeo.org.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013) The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res* **41**: 590–596.

R Development Core Team (2008) R: A language and environment for statistical computing.

Ramirez-Llodra, E., Clark, M.R., Smith, C.R., Tyler, P.A., Rowden, A.A., Bergstad, O.A., et al. (2011) Man and the Last Great Wilderness: Human Impact on the Deep Sea. *PLoS One* **6**: e22588.

Ramirez-Llodra,E., Brandt, A., Danovaro, R., De Mol, B., Escobar, E., German, C.R., Levin, L.A., Martinez Arbizu, P., Menot, L., Buhl-Mortensen, P., Narayanaswamy, B.E., Smith, C.R., Tittensor, D.P., Tyler, P.A., Vanreusel, A., Vecchione, M. (2010) Deep, diverse and definitely different: unique attributes of the world's largest ecosystem. *Biogeosciences* **7(9)**:2851-2899.

Redmond, N.E., Morrow, C.C., Thacker, R.W., Diaz, M.C., Boury-Esnault, N., Cárdenas, P., et al. (2013) Phylogeny and systematics of demospongiae in light of new small-subunit ribosomal DNA (18S) sequences. *Integr Comp Biol* **53**: 388–415.

Rumín-Caparrós, A., Sanchez-Vidal, A., González-Pola, C., Lastras, G., Calafat, A., and Canals, M. (2016) Particle fluxes and their drivers in the Avilés submarine canyon and adjacent slope, central Cantabrian margin, Bay of Biscay. *Prog Oceanogr* **144**: 39–61.

Sánchez, F., Gómez-Ballesteros, M., González-Pola, C., and Punzón, A. (2014) Sistema de Cañones Submarinos de Avilés, Madrid, Spain.

1065 Sánchez, F., Rodríguez Basalo, A., García-Alegre, A., and Gómez-Ballesteros, M. (2017)
1066 Hard-bottom bathyal habitats and keystone epibenthic species on Le Danois Bank
1067 (Cantabrian Sea). *J Sea Res* **130**: 134–153.
1068 Sánchez, F., Serrano, A., Parra, S., Ballesteros, M., and Cartes, J.E. (2008) Habitat
1069 characteristics as determinant of the structure and spatial distribution of epibenthic and
1070 demersal communities of Le Danois Bank (Cantabrian Sea, N. Spain). *J Mar Syst* **72**:
1071 64–86.
1072 Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of
1073 large phylogenies. *Bioinformatics* **30(9)**:1312-1313.
1074 Taboada, S., Ríos, P., Mitchell, A., Cranston, A., Busch, K., Tonzo, V., Cárdenas, P.,
1075 Sánchez, F., Leiva, C., Koutsouveli, V., Cristobo, J., Xavier, J., Hentschel, U., Rapp,
1076 H.T., Morrow, C., Drewery, J., Romero, P., Arias, M.B., Riesgo, A. (*in review*) Genetic
1077 diversity, gene flow and hybridization in fan-shaped sponges *Phakellia* spp. in the North-
1078 East Atlantic deep sea. submitted to *Molecular Ecology.*
1079 Topsent, E. (1890). Notice préliminaire sur les spongiaires recueillis durant les campagnes de
1080 l'Hirondelle. Bulletin de la Société zoologique de France. **15**: 26-32, 65-71.
1081 van Sebille, E., Griffies, S.M., Abernathey, R., Adams, T.P., Berloff, P., Biastoch, A., et al.
1082 (2018) Lagrangian ocean analysis: Fundamentals and practices. *Ocean Model* **121**: 49–
1083 75.
1084 Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W.S., and
1085 Huttenhower, C. (2011) Metagenomic biomarker discovery and explanation. *Genome*
1086 *Biol* **12**: R60.
1087 Van Soest and de Voogd (2015) Sponge species composition of north-east Atlantic cold-water
1088 coral reefs compared in a bathyal to inshore gradient. Journal of the Marine Biological
1089 Association of the United Kingdom 95(7):1461-1474.
1090 Steinert, G., Rohde, S., Janussen, D., Blaurock, C., and Schupp, P.J. (2017) Host-specific
1091 assembly of sponge-associated prokaryotes at high taxonomic ranks. *Sci Rep* **7**: 1–9.
1092 Steinert, G., Wemheuer, B., Janussen, D., Erpenbeck, D., Daniel, R., Simon, M., et al. (2019)
1093 Prokaryotic diversity and community patterns in antarctic continental shelf sponges.
1094 *Front Mar Sci* **6**: 1–15.
1095 Taboada, S., Riesgo, A., Wiklund, H., Paterson, G.L.J., Koutsouveli, V., Santodomingo, N., et
1096 al. (2018) Implications of population connectivity studies for the design of marine
1097 protected areas in the deep sea: An example of a demosponge from the Clarion-
1098 Clipperton Zone. *Mol Ecol* **27**: 4657–4679.

1099 Taylor, M.L. and Roterman, C.N. (2017) Invertebrate population genetics across Earth's
1100 largest habitat: The deep-sea floor. *Mol Ecol* **26**: 4872–4896.
1101 Thomas, T., Moitinho-Silva, L., Lurgi, M., Björk, J.R., Easson, C., Astudillo-García, C., et al.
1102 (2016) Diversity, structure and convergent evolution of the global sponge microbiome.
1103 *Nat Commun* **7**: 11870.
1104 Tsujino, H., Urakawa, S., Nakano, H., Small, R.J., Kim, W.M., Yeager, S.G., et al. (2018)
1105 JRA-55 based surface dataset for driving ocean–sea-ice models (JRA55-do). *Ocean*
1106 *Model* **130**: 79–139.
1107 Vad, J., Kazanidis, G., Henry, L.-A., Jones, D.O.B., Tendal, O.S., Christiansen, S., Henry,
1108 T.B., Roberts, J.M. (2018) Potential impacts of offshore oil and gas activities on deepsea
1109 sponges and the habitats they form. *Adv. Mar. Biol.* **79**:33-60.
1110 White, J.W., Schroeger, J., Drake, P.T., Edwards, C.A. (2014) The value of larval
1111 connectivity information in the static optimization of marine reserve design. *Conserv.*
1112 *Lett.* **7**:533–544.
1113
1114
1115
1116

Highlights

- Fan-shaped sponges display panmixia at three locations in the Cantabrian Sea. Subtle sponge population genetic and pronounced microbial differences were observed between a canyon and bank location lying < 100 km apart.
- Lagrangian modelling reveals variable inter-annual connectivity via ocean currents between the sampling regions.
- Interdisciplinary approaches, such as the here presented combination of sponge taxonomy, genetics, microbiology, and particle tracking modelling, can help to improve understanding about connectivity in the deep-sea. This is particularly crucial for key organisms like sponges and the ecosystems they form.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: