# Temporal, environmental, and biological drivers of the mucosal microbiome in a wild marine fish, *Scomber japonicus*

Jeremiah J Minich[1], Semar Petrus[2], Julius D Michael[2], Todd P Michael[1,2], Rob Knight[3,4,5], and Eric E. Allen[1,3,6*]

[1] Marine Biology Research Division, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA, USA

[2] J. Craig Venter Institute, La Jolla, CA, USA

[3] Center for Microbiome Innovation, University of California San Diego, La Jolla, CA, USA

[4] Department of Pediatrics, University of California San Diego, La Jolla, CA, USA

[5] Department of Computer Science and Engineering, University of California San Diego, La Jolla, CA, USA

[6] Division of Biological Sciences, University of California, San Diego, La Jolla, CA, USA

* To whom correspondence should be addressed: Eric E. Allen <eallen@ucsd.edu>

**Running Title**: Mackerel microbiome

**Keywords**: fish microbiome, mackerel, microbial ecology, microbial biogeography, *Scomber japonicus*

# Abstract

Changing ocean conditions driven by anthropogenic activity may have a negative impact on fisheries by increasing stress and disease with the mucosal microbiome as a potentially important intermediate role. To understand how environment and host biology drives mucosal microbiomes in a marine fish, we surveyed five body sites (gill, skin, digesta, GI, and pyloric caeca) from 229 Pacific chub mackerel, *Scomber japonicus*, collected across 38 time points spanning one year from the Scripps Institution of Oceanography Pier, making this the largest and longest wild marine fish microbiome survey. Mucosal sites had unique communities significantly different from the surrounding sea water and sediment communities with over 10 times more diversity than sea water alone. Although, external surfaces such as skin and gill were more similar to sea water, digesta was similar to sediment. Both alpha and beta diversity of the skin and gill was explained by environmental and biological factors, especially sea surface temperature, chlorophyll a, and fish age, consistent with an exposure gradient relationship. We verified that seasonal microbial changes were not confounded by migrations of chub mackerel sub-populations by nanopore sequencing a 14 769 bp region of the 16 568 bp mitochondria. A cosmopolitan pathogen, *Photobacterium damselae*, was prevalent across multiple body sites all year, but highest in the skin, GI, and digesta between June and September. Our study evaluates the extent which the environment and host biology drives mucosal microbial ecology, establishing a baseline for long term monitoring surveys for linking environment stressors to mucosal health of wild marine fish.


# Introduction

Pacific chub mackerel, *Scomber japonicus* (Houtuyn 1782), is an economically and ecologically important, cosmopolitan, marine coastal-pelagic fish found in the temperate and tropical waters of the Pacific, Atlantic, and Indian Oceans [1, 2]. *S. japonicus* is currently the fifth largest commercial fishery (purse-seine) in the world [3], processed for human consumption and animal food. Historically in the US, *S. japonicus* was a prominent commercial fishery, but has been on the decline since the 1980s due to a collapse in spawning and fishery stock biomass leading to the last US mackerel canary closing in 1992 [4]. The boom and busts of the fishery have been attributed to large scale environmental factors such as Pacific decadal oscillation, North Pacific gyre oscillation, sea surface temperature, sea level, upwelling, and chlorophyll *a* [5–8]. Juveniles grow quickly reaching 50% of total growth by the first 1.5 years of life with larval growth highest in warmer water (16.8 - 22.1$^{o}$C) [9]. Larvae eat copepods and zooplankton [3], while juveniles and adults consume primarily small fish and pelagic crustaceans [2]. *S. japonicus* are an important prey item for marine mammals, sea birds, and higher trophic fish such as tunas and sharks [2]. In the eastern North Pacific, *S. japonicus* migrate North in the summer and South in the winter [10] with seasonal offshore migrations occurring from March to May. Climate change and warming oceans likely have contributed to stocks shifting to more northerly migrations [4]. Modeling has shown that nearly 90% of the *S. japonicus* catch was explained by temperature (28-29.4 $^{o}$C), salinity 33.6-34.2 psu, and chlorophyll *a* of 0.15-0.5 mg/m$^3$ [8] whereas survival rates to one year recruits was highly associated with low plankton biomass [11]. *S. japonicus* are ecologically and commercially important while occupying broad environmental gradients. This combined with their relative ease of collection make them an excellent model organism to study the environmental and biological drivers of microbiome diversity in a marine vertebrate.

69

70 The primary mucosal environments of fish include the gills, skin, and throughout the
71 gastrointestinal tract, all of which are important to fish health. Disease resistance in the host is
72 promoted in the mucus through continual epithelial shedding and immune cell regulation [12,
73 13]. The mucus is an important physical barrier to the environment and is generally thought to be
74 colonized with a unique microbiome [14]. The skin and gut both have mucosal associated
75 lymphoid tissues which produce IgT+ B cells protecting the host from invasion of mucosal
76 microbiota [15]. The establishment of microbiomes on mucosal sites is a function of exposure
77 and successful colonization. Mucosal environments within a fish, have varying levels of
78 environmental exposure such as habitat (sea water, sediment, kelp forests), with gills and skin
79 having different exposure rates than the gastrointestinal tract. The GI tract, however will have
80 varying exposure to nutrients during digestion of food. Successful colonization within mucosal
81 sites will further be driven by variables regulated by the host which can include different
82 physiological conditions of the host, thus microbial communities are likely to be reflected.
83 Various protective enzymes related to the innate immune response including lysozymes,
84 proteases, phosphatases, esterases, and sialic acid can be differentially abundant in the mucus
85 depending on the host fish exposure to environmental microbes [16].

86

87 To understand the full microbiome potential of a given host, it is important to evaluate the
88 variability longitudinally throughout an entire season (year), and then to continue sampling
89 throughout consistent periods for multiple years. Including long term biological monitoring of
90 commercially and ecologically important marine fish to complement the 100 years of sea water
91 data taken from the SIO pier will be important for understanding marine ecosystem dynamics.

92 Although most ecological studies since 2004 are less than a year and have sampling frequencies
93 of 1 month or greater [17], we have designed our study to include 38 sampling frequencies
94 across 1 year. Previous work looking at seasonal or temporal microbiome changes in the marine
95 environment has focused on free-living pelagic seawater microbes [18]. Gilbert *et al.* found that
96 day length described over 65% of microbial community diversity with richness highest during
97 winter months in the North Atlantic [19]. Very few time series datasets spanning an entire year
98 exist for analyzing the host-associated microbiome. Within humans, most seasonally-active
99 microbes in the gut are associated with populations spending more time outdoors suggesting that
100 seasonal variance in the environment has a greater influence on those with higher environmental
101 exposure [20]. In freshwater fish, lower microbial diversity and altered composition in the gut
102 was associated with warmer summer months in tilapia reared in earthen ponds [21]. In salmon
103 however, no seasonal variations of gut microbiota composition were detected although alpha
104 diversity was highest during warm water months [22]. To date, no systematic analysis of the
105 temporal variability in wild fish microbiomes has been done previously.
106
107 The purpose of this study was to quantify the effects of the environmental and biological drivers
108 across five unique mucosal body sites in a marine fish over a longitudinal time course spanning
109 one year. From Jan 28 2017 to Jan 26 2018, 229 pacific chub mackerel, *Scomber japonicus*, were
110 collected off the SIO pier across 38 sampling events. Mucosal microbiome communities were
111 sampled from five body sites including gill, skin, digesta, gastrointestinal tract, and pyloric caeca
112 within each fish. Regional coastal sampling of sea water and marine sediment microbial
113 communities were collected to compare mucosal communities to potential environmental
114 sources. Microbiome processing was performed using the Earth Microbiome protocol using the

16S rRNA gene V4 region. Water conditions (salinity, temperature, pressure, and chlorophyll a) and fish biometrics (length, mass, condition factor, age) were collected and compared to mucosal microbiomes to determine significant ecological drivers. We evaluate both alpha diversity measures (Shannon entropy and Faith's Phylogenetic diversity) and beta diversity (weighted and unweighted UniFrac) to assess these changes. In addition, we calculate microbial gamma diversity across body sites and time to understand effects of sampling effort on capturing true host-microbiome diversity. Our results show that mucosal communities across body sites are highly differentiated in a single species of fish and that seasonal environmental drivers partially account for this differentiation.

# Materials and Methods

### Sample collection *S. japonicus* time series

From Jan 28 2017 to Jan 26 2018, 1-8 *Scomber japonicus* specimens were collected across 38 sampling events from the end of the SIO Pier (32.867, -117.257). Sea surface water samples were collected from each sampling event and immediately stored on dry ice. Environmental conditions at time of sampling, including sea water temperature, salinity, pressure, and chlorophyll *a* concentration, were collected using publicly available data from the Southern California Coastal Ocean Observing System (SCCOOS) SIO pier shore station data archive (http://www.sccoos.org) (Fig 1a). Fishing occurred at or near sunset with exact times recorded in the metadata (see Qiita Study ID 11721 for full metadata). Fish were caught using hook and line with a Sabiki rig, immediately euthanized upon landing using accepted protocols according to

AVMA guidelines and stored on dry ice. Individual fish were wrapped in aluminum foil and handled with gloves prior to storage on dry ice to minimize contamination and then stored long term at -80ºC for up to 6 months prior to dissection. Upon processing, frozen fish were weighed and measured, along with calculation of Fulton's condition factor, which is a proxy for fish health [23, 24]. Age was estimated using fish length as derived from the most recent Pacific chub mackerel stock assessment [4] (Fig 1b-c) [25] where otoliths were compared to 25 fish individuals per catch across multiple years (1962-2008). Specifically, the von-Bertalanffy equation was used with two separate growth coefficients: $LA = L\infty (1 - e^{-k(A-to)})$ where LA=length at age, L∞=theoretical maximum length of fish, k = growth coefficient, to = theoretical age when length is 0 mm. After 30 minutes of thawing the fish, a cotton swab (Puritan, Cat #806-wc) was swiped back and forth five times along the left gill and then put directly into a 2 ml Mo Bio PowerSoil (Mo Bio, Cat # 12888) bead beating tube. The skin was also swabbed in a 3 cm x 3 cm area on the left side behind the gill and above the pectoral fin (Fig 1d). After carefully dissecting the fish with a new razor blade, the last 3 cm of GI tract was cleared and the digesta sampled. This same distal portion of GI tract was cut and also sampled. Lastly, an approximate 50 mg sample of pyloric caeca was sampled from the fish and placed in a tube. The tubes were then stored at -80ºC until DNA extraction commenced. For additional environmental controls, surface seawater and sediment samples were collected across two time points (Dec 8 2017 and Jan 12 2018) at 30 coastal locations, approximately 200 m offshore, spanning 10 km throughout San Diego including soft bottom, reef, river mouth, and bay areas.

**Microbiome processing**

Samples were processed using the standard 16S rRNA gene Earth Microbiome Protocol (EMP),

with only slight modifications (http://wwwearthmicrobiome.org). Specifically, genomic DNA was extracted using a hybrid approach where lysis is performed in 2 ml bead beating tubes and then cleanup performed using the KingFisher robot to reduce well to well contamination (Minich et al, in prep). The initial cell lysis steps were performed in single-tube reactions (instead of 96-well plate format) followed by transfer to plates for the standard magnetic bead cleanup on the KingFisher robots using the Mo Bio PowerMag kit (Mo Bio, Cat # 27000-4-KF) which has improved limits of detection for low biomass samples [26]. The EMP extraction procedure includes modifications including the use of RNaseA during lysis and a 10 minute incubation at 65ºC prior to bead beating. All sample batches had positive and negative controls included with each extraction set so that sample exclusions based on read counts could be calculated [26]. Extracted gDNA was then PCR amplified using the EMP 16S V4 515f/806rB bar-coded primers [27, 28]. The miniaturized PCR method, which generates libraries at a 58% cost reduction of $1.42 per sample, was used for all samples that included the use of the Echo-550 instrument to do triplicate 5 ul PCR reactions [29]. Amplicons were quantified using a pico green assay, and then 2 ul of each sample was equally pooled into a single tube. This final pool was then cleaned up to remove dNTPs and primer dimers using the QIAquick PCR purification kit (Qiagen, Cat# 28106). Final pools contained up to 768 samples which were then sequenced on an Illumina MiSeq using a 2x150 bp strategy (300-cycles v2 kit, Illumina, San Diego, CA). Bioinformatic processing of samples occurred using Qiita [30] and QIIME 1.9.1 or QIIME 2.0 [31], with the first 150 bp read trimmed to 150 bp and processed through deblur [32], a *de novo* sOTU picking method. A phylogenetic tree of the 16S sOTU single-sequence tags was created using SEPP (SATé-Enabled Phylogenetic Placement) [33]. Rarefaction levels were empirically determined by calculating the read counts at which 90% of the reads from the DNA extraction positive

184 controls map back to the positive controls [26].

185

186 **Summary microbiome statistics**

187 Alpha, beta, and gamma diversity of microbial communities was measured [34]. Alpha diversity
188 was calculated using measures of Shannon [35] and Faith's Phylogenetic Diversity [36] while
189 beta diversity was calculated using weighted and unweighted UniFrac [37, 38] distance and
190 visualized in Emperor [39]. Alpha and beta diversity statistical significance was tested using
191 Kruskal-Wallis test [40]. Taxonomies were classified in Silva [41] using the Greengenes and
192 RDP databases [42] using the following parameters: minimum identity with query sequence
193 (0.95), number of neighbors per query sequence (10), greengenes- reference NR database, search
194 kmer-candidates (1000), lca-quorum (0.8), search-kmer-length (10), search-kmer-mm (0),
195 search-no-fast, reject sequences below 70%.

196

197 **Statistical analysis of environmental and biological drivers of fish mucosal microbiomes**

198 To evaluate the extent to which the environment and biology of the fish influences the microbial
199 communities of the various body sites, both alpha diversity and beta diversity were analyzed.
200 Only samples which had environmental values for all water conditions (temperature, salinity,
201 pressure, and chlorophyll a) and biological conditions (weight, fork length, condition factor, and
202 age) were included. Thus, some samples had to be excluded due to temporary failure of the
203 chlorophyll *a* fluorometer instrument on the SIO pier (August 4 2017). Alpha diversity measures
204 for each body site were independently verified and tested to ensure they met the assumptions for
205 the General Linear Model (GLM). Specifically, to test for normally distributed residuals, sets
206 were analyzed using the R package Library(car) [43] and run through the Shapiro-Wilk

Normality test [44]. To evaluate and test for homoscedasticity, the non-constant error variance test (ncvTest) commonly known as the Breusch-Pagan test was used [45]. To meet GLM criteria, the Faith's Phylogenetic Diversity samples were log-transformed. Both Shannon and Faith's PD were then processed through the GLM in R while controlling for collinearity of variables. Individual $R^2$ values and P-values for each environmental and biological variable are reported along with total $R^2$, F-statistic, and P-values for all variables. Gill samples using Shannon diversity and GI samples using Faith's PD were excluded from analysis due to not meeting required assumptions of the GLM. To evaluate the effects of environmental and biological variables on beta diversity, we assessed both unweighted and weighted UniFrac for each body site independently using Adonis, a non-parametric analysis of variation method, [46] in QIIME 1.9.1 [31] and Calypso [47].

**Validation of *Photobacterium damselae* sOTU phylogeny**

To validate the taxonomy assignments of five *Photobacterium* sOTUs in our dataset, we performed multiple sequence alignment (Neighbor-Joining) of a 148 bp region of the 16S rRNA gene v4 region from eight other strains. Specifically, we used default settings (nucleotide scoring 200 PAM / k=2, Gap opening penalty 1.53, offset value = 0, 'nzero' where Ns have no effect on alignment score) in the MAFFT alignment tool [48, 49]. The phylogenetic tree was visualized using Phylo.io [50]. The comparison bacteria strains included: two pathogenic *Photobacterium spp.* isolates (*P. damselae* ATCC 33539T, Genbank X74700.1; *P. damselae*, Genbank D25308.1), four non-pathogenic *Photobacterium spp*. (*P. leiognathi* strain ATCC 25521, Genbank NR_115541.1; *P. angustum* ATCC 25915T, Genbank X74685.1; *P. phosphoreum* strain ATCC 11040, Genbank NR_115205.1; *P. rosenbergii* strain CC1, Genbank

NR_042343.1) and two outgroup *Vibrio* species (*Vibrio pelagius* strain ATCC 25916, Genbank NR_119059.1; *Aliivibrio fischeri* strain ATCC 7744, Genbank NR_115204.1) which were identified from various studies [51–53].

**Population genetics of *S. japonicus***

We developed a high-throughput two fragment, mitochondrial amplicon workflow for the Oxford Nanopore long read sequencer. A total of 96 gDNA skin mucus samples, spanning 5 collection months (Aug 27 2017 – Jan 26 2018) were amplified in 192 separate 10 ul PCR reactions using 1 ul gDNA, 5 ul NEB Long Amp mastermix (NEB, Cat# M0287S), and 3.4 ul molecular grade water and one of two primer combinations. The first mtDNA fragment (96 PCR reactions) used 0.4 ul of 10uM forward primer (SJ_F1_655: 5'-TTT CTG TTG GTG CTG ATA TTG | CAA ACC TCA CCC TCC CTT GTT-3') and 0.4 ul of 10 uM reverse primer (SJ_R1_7653: 5'- ACT TGC CTG TCG CTC TAT CTT | CAC CAC TAT TCG GTG GTC TGC-3'). The second fragment (96 PCR reactions) used 0.4 ul of 10uM forward primer (SJ_F2_7425: 5'-TTT CTG TTG GTG CTG ATA TTG | CTC CCT GCC GTC ATT CTT ATC) and 0.4 ul of 10 uM reverse primer (SJ_R2_15424: 5'-ACT TGC CTG TCG CTC TAT CTT | CGA CGA CTA CGT CTG CGA CAA). All primers have ONT adaptor regions on the first 27 bases as indicated by '|'. All PCR reactions followed the following protocol: 94 ºC 3 minutes, 25 cycles of 94 ºC 30s, 60 ºC 30s, 65 ºC 8:20, a final extension of 68 ºC 10 minutes followed by storage at 4 ºC. Following the first PCR, a second 5 ul PCR reaction was conducted for each of the 96 samples by combining 2.5 ul NEB Long Amp mastermix, 0.1 ul of a unique barcode (Oxford Nanopore PCR Barcode kit 01-96, batch DK601001 brown box) and finally pooling 1.2 ul of each PCR product (first + second fragment). Barcodes were transferred to the PCR reaction

plate using the acoustic liquid handler Labcyte Echo 550. The same PCR reaction was used. A final 2 ul of sample was pulled from all samples and processed through the QiaQuick PCR purification kit (Qiagen, Cat#28106) and run on a 1% agarose gel to confirm size. The pool was then run on a used MinION using the 1D PCR barcoding protocol (SQK-LSK109). Samples were demultiplexed, uploaded to Galaxy [54], and aligned against the *Scomber japonicus* reference mitochondria genome (NC_013723) using LASTZ aligner (Galaxy version 1.3.2)[55] using defaults. Consensus sequences were visualized, calculated, and exported using the quick consensus mode in Integrated Genome Viewer [56]. Samples with either less than an average of 10x coverage or samples with more than 20 ambiguous basepairs (Ns) were excluded from the analysis (n=2, BC52, BC82). A phylogenetic tree of all 91 *S. japonicus* samples along with three reference *S. japonicus* mitochondrial genomes from NCBI (AB488405.1, NC_013723.1, AB102724.1), and two outgroup species *Scomber colias* (NC_013724.1) and *Scomber australasicus* (AB102725.1) was generated using MAFFT [49] (NJ conserved sites 12388, Jukes-Cantor substitution model, bootstrap 100) and visualized with Phylo.io using default parameters [50].

All microbiome data is publicly available through Qiita (sample ID 11721, prep ID 4638), EBI (ERP109537), and NCBI (BioProject PRJEB27458).

# Results

**Microbial diversity associated with a marine pelagic fish across body sites over one year**

From January 2017 to January 2018, 229 wild *Scomber japonicus* were collected from the SIO pier across 38 sampling events at approximately five fish per week, although actual takes varied due to weather and other constraints. Sea water temperature, salinity, pressure, and chlorophyll *a* were recorded using the SCOOS online database (Figure 1a). Fork length and mass were recorded and approximate age of the fish determined from length (Figure 1b). The condition factor of the fish was positively associated with older fish ($P<0.0001$, $R^2=0.307$) (Figure 1c). Along with paired sea water samples, mucosal microbiome samples were sequenced from the gill, skin, digesta, GI, and pyloric caeca of each fish (Figure 1d).

A total of 612 samples resulting in 18 857 sOTUs, processed with the miniaturized PCR method, passed the sample exclusion criteria. Sample exclusion criteria was based on the KatharoSeq method where the read counts from DNA extraction positive controls of varying cell counts was compared to compositional read out and the read count at which 90% of the reads mapped appropriately was chosen as the rarefaction depth, which was 1 362 reads (Supplementary Figure 1). Alpha diversity measured by Faith's PD, was significantly different when compared across mackerel body sites and sea water (Kruskal-Wallis, $P<0.0001$, KW statistic 87.48) (Figure 2a). Gill, skin, and digesta samples had higher diversity than the GI and pyloric caeca samples while gill and digesta had higher diversity than sea water (Figure 2a). Beta diversity indicates the gill and skin mucosal samples were most similar to sea water while digesta, GI, and pyloric caeca were uniquely clustered and also had higher within body site variability (Figure 2b). Some of the digesta samples also appeared to cluster more closely to sediment samples. When tested, skin followed by gill samples were most similar to sea water samples whereas the digesta samples were most similar to sediment (Figure 2c). To understand sample size requirements for capturing

298 novel microbial diversity in fish, we compared accumulation of microbial richness over the one-
299 year sampling period across all sample types. Overall microbial richness in the gill, skin, GI,
300 pyloric caeca, and sea water appeared to level off after only a couple months (20-50 samples)
301 whereas digesta samples continued to increase perhaps requiring another few years of data
302 collection to approach saturation. For comparison, we included gill, skin, and digesta samples
303 from 14 other local San Diego species of fish (right of the dotted line) (Figure 2d). While digesta
304 diversity increased with the addition of the first new species it followed a similar trend while gill
305 and skin samples did not increase much suggesting an overall conservation of microbes in other
306 species of fish. Lastly, total Gamma diversity or richness was calculated for all samples in this
307 study showing that sediment samples had the most microbial diversity followed by mackerel
308 digesta and mackerel gill. The total unique microbial diversity in a single species of fish, *S.*
309 *japonicus*, was 8.8 fold more than sea water (9 172 vs. 1 039 sOTUs) (Figure 2d), demonstrating
310 the potential for microbial discovery within and upon fish hosts in the ocean.
311
312 **Environmental and biological drivers of the *S. japonicus* mucosal community**
313 We next quantified the combined and specific effects of four environmental variables including
314 chlorophyll *a* concentration, sea water temperature, salinity, and pressure along with four
315 biological variables including fish age, fork-length, mass, and condition factor on the fish-
316 associated mucosal microbiomes. Alpha diversity measures were assessed using the General
317 Linear Model (GLM). For alpha diversity measures of Shannon diversity, skin mucus was
318 significantly influenced by the factors ($P<0.001$, $R^2 = 0.38$, F-stat 6.595), with chlorophyll *a*
319 having a negative association and temperature a positive association ($P<0.0001$, $P=0.0004$),
320 respectively. Gill samples were not assessed (grey line Table 1) because the Shannon diversity

321 did not meet the assumptions of normally distributed residuals (Shapiro test P <0.05) and was
322 not homoscedastic (Breusch-Pagan P<0.05) (Table 1). For the alpha diversity measure of Faith's
323 phylogenetic diversity (PD), which takes into account phylogenetic diversity with richness, all
324 data was log transformed to meet the assumptions of the GLM. The gastrointestinal tract samples
325 however, still did not meet the assumptions as the residuals were not normally distributed
326 (Shapiro test P<0.05), thus were excluded from analysis (grey line Table 1). Both gill, skin, and
327 pyloric caeca Faith's PD were significantly influenced by the measured factors (gill: P<0.0001,
328 $R^2$=0.33, F-state=7.042; skin: P=0.00039, $R^2$=0.26, F-stat=4.239; pyloric caeca: P=0.00891,
329 $R^2$=0.22, F-stat=2.972). The gill sample diversity was negatively associated with Chlorophyll a
330 concentration (P=0.00549). Skin was negatively associated with Chlorophyll *a* concentration
331 (P=0.00182) and age (P=0.00811), while positively associated with temperature (P=0.04434).
332 The pyloric caeca was positively associated with age (P=0.04787) and temperature (P=0.00305)
333 while negatively associated with salinity (P=0.04921).
334
335 The extent to which environmental and biological variables explain microbial diversity was also
336 assessed for Beta diversity including both unweighted UniFrac and weighted UniFrac distances.
337 The Adonis permutational multivariate statistical analysis was used to test overall significance
338 along with variance explanation by factor. Unweighted UniFrac distance measures showed that
339 gill, skin, and digesta samples were influenced by measured factors (Adonis, P<0.0001, $R^2$=0.12,
340 $R^2$=0.15, $R^2$=0.09). The gill was primarily driven by Chlorophyll *a* concentration and age while
341 skin was influenced mostly by Chlorophyll *a*, age, and fork-length. For weighted UniFrac
342 distances, both gill (P<0.0001, $R^2$=0.14) and skin (P=0.001, $R^2$=0.20) were significantly
343 influenced by factors with age being the most significant driver. In summary, the skin mucosal

344 microbiome was significantly influenced by environmental and biological factors in each of the
345 four measures across alpha and beta diversity while gill was significant in each of the three
346 measures. The environmental variables of Chlorophyll *a* followed by temperature had the most
347 frequent influences on microbial communities across body sites while age was the most frequent
348 biological factor (Table 1).
349
350 **Population structure of *S. japonicus***
351 *S. japonicus* are thought to have three spawning populations along the Pacific coast of North
352 America[4], which suggest that our environmental and biological associations could be explained
353 in part by population dynamics over the year. To estimate the changes in *S. japonicus* population
354 structure over our time course we sequenced two fragments of mitochondrial DNA directly from
355 skin mucus gDNA for a total length of 14 769 bp for 93 fish samples landed between Aug 27
356 2017 and Jan 26 2018. Two samples were removed from the analysis due to having lower
357 coverage (less than 10x) or more than 20 Ns in the consensus sequence (Supplemental Figure
358 2a). The majority of samples (93%) had at least 100x coverage of the mitochondria target region
359 (Supplemental Figure 2b). Based on near full-length mitochondria data, no population structure
360 was observed, consistent with our sampling of one population of *S. japonicus* over the course of
361 the study (Supp Figure 3).
362
363
364 **Candidate pathogen and probiotic associations**
365 Microbial diversity was unique within the various *S. japonicus* body sites and environment (sea
366 water and sediment) with the top 25 most abundant genera comprising the majority of reads

(Figure 3). *Rhodobacteraceae* were found in all environments particularly the seawater with lower levels in the gill, skin, digesta, and GI body sites. The gill was dominated by several sOTUs within the *Shewanella* genera along with microbes from the *Rickettsiales* and *Polynucleobacter* genera. Sea water microbes including *Rhodobacter* and *Pelagibacter* were also present on the gill in lower numbers. Various sOTUs of *Photobacterium* were highly abundant across the skin, digesta, GI, and pyloric caeca. *Enterovibrio*, another fish pathogen, was abundant on the fish skin and pyloric caeca. Digesta samples were comprised of many seawater dwelling Cyanobacteria including *Synecochoccus* but was also high in the family *Pirellulaceae* (Planctomycetes) and sporadically high amounts of the genus *Clostridium*. *Mycoplasma* (Tenericutes) was a dominant genus in the GI and pyloric caeca. Various *Bacillus spp.* and *Lactobacillus spp*. were found to be present across multiple body sites (Supplemental Figure 4). As expected, seawater contained many common groups including *Synechococcus*, *Rhodobacteraceae*, *Pelagibacter*, and *Flavobacteriaceae* while the sediment had consistently higher levels of *Pirellulaceae* (Figure 4).

There were five highly abundant *Photobacterium* sOTUs in the dataset which prompted further phylogenetic evaluation to elucidate species level assignments. The full 16S rRNA genes of two pathogenic isolates of *P. damselae*, four other Photobacterium species including *P. angustum*, *P. phosphoreum*, *P. leiognathi*, and *P. rosenbergii* and two vibrio species as outgroups were aligned with the five *Photobacterium* unique 150 bp sOTUs. The phylogenetic tree (Figure 4a) of the five Photobacterium sOTUs in this dataset with the known *Photobacterium* strains is able to identify and resolved the taxonomic assignments. The *Photobacterium damselae* sOTU identified in our dataset was 100% identical to the v4 region of 16S rRNA from the two

pathogenic strains while distinct from the other *Photobacterium spp*. This *P. damselae* sOTU was identified across various body sites of fish, but was most prevalent in the GI, skin, and digesta samples (present across 22.2%, 16.1%, and 14.7% of samples, respectively) (Figure 4b). For the GI, skin, and digesta samples which had *P. damselae* present, the single *P. damselae* sOTU made up 5.88%, 6.99%, and 5.32% of the total microbial composition, respectively. Further, the temporal enrichment and prevalence of this *P. damselae* sOTU was highest between June and September.

# Discussion

Our study evaluated how the mucosal microbial community of a wild marine fish species is influenced according to environmental and biological variance as experienced over the course of an annual season in coastal temperate waters. Body sites had unique microbial signatures that were uniquely influenced by environmental and biological measures. Alpha diversity was highest in the gill, skin, and digesta communities as compared to the gastrointestinal tract and pyloric caeca. Beta diversity measures demonstrated that fish mucosal sites were primarily driven by body site location and were further unique to the surrounding environment. An exposure gradient was observed with skin and gill surfaces being more similar to the water column while the digesta community more similar to the sediment. Further, the environmental and biological variables best explained variation in the skin and gill microbiomes as opposed to the internal body sites (digesta, GI, pyloric caeca). Lastly, an important fish pathogen, *Photobacterium damselae* was observed in high prevalence on GI, skin, and digesta communities and was associated with the summer months which exhibit higher temperatures and low nutrients. This is

413 the first comprehensive microbiome study of a marine fish that evaluates multiple body sites
414 from a large sample size over an extended time series.
415
416 Regardless of environmental conditions, the mackerel mucosal body site was the strongest driver
417 of microbiome diversification, with each site associated with a specific gradient of
418 environmental exposure. The gill and skin communities were most similar to the seawater
419 whereas the gastrointestinal samples were more divergent. This environmental gradient which
420 distinguishes host-associated gut microbes from free-living microbes, was first described in
421 mammalian vertebrates [57]. Environmental exposure gradients have also been shown to
422 influence gut or skin microbiomes in amphibians, fish [58, 59], and other vertebrates [60, 61],
423 whereas our study is the first comprehensive community assessment that explicitly tests this
424 comparison with multiple mucosal body sites in fish. Marine fish differ from other vertebrates in
425 that their microbial exposure rates are greatly elevated compared to terrestrial or freshwater
426 species. Seawater can harbor as many as 1 million cells per ml [62], while coastal sediments can
427 be two-orders of magnitude higher at 100 million cells per $cm^3$ [63]. Gill microbial communities
428 may be supported physically by complex morphological structure of laminae and filaments and
429 chemically through gas exchange, ion transport, and waste excretion. Age, phylogeny, diet, and
430 individual have been implicated as influencing the gill microbial community in tropical fish, with
431 *Shewanella* taxa being dominant [64]. Skin microbiomes of marine tropical fish have been also
432 shown to be driven by phylogeny and diet [65]. Digesta and GI samples in *S. japonicus* were the
433 most variable suggesting that either niche differentiation is more static in the gill and skin
434 environments or that microbial turnover is lower. Few studies, however, have evaluated these
435 body sites in temperate marine fish. Discovery of novel microbial lineages and metabolic activity

436 should focus on fish mucosal associated environments, specifically the gill, skin, and digesta
437 communities that had the highest levels of phylogenetic diversity in our dataset. Sediment
438 samples had the highest diversity, yet were most similar, thus having the lowest inter-individual
439 variability.
440
441 The environmental and biological variables most explained the skin and gill microbiomes as
442 compared to the internal GI communities. Within the environmental variables, Chlorophyll *a*
443 concentration followed by temperature and salinity were the strongest drivers while age was the
444 most pronounced of the biological metrics. Chlorophyll *a* concentration is a general indicator of
445 primary production and microbial growth or proliferation in the water column. As phytoplankton
446 blooms occur in the ocean due to nutrient enrichments through upwelling, bacterial communities
447 in the water column also change thus altering exposure to fish and other marine animals. While
448 many studies have examined the effects of harmful algal blooms on marine organisms [66], few
449 have quantified the extent of these exposures in the wild. Temperature has been shown to
450 influence marine macroalgae [67] and oyster hemolymph microbiomes [68]. Salinity was one of
451 the first major abiotic conditions shown to drive microbiomes in free-living freshwater versus
452 marine environments [69] and has also been shown to influence fish microbiomes
453 deterministically [70]. Fish gill parasite load has been shown to be positively associated with fish
454 age, season, eutrophic water conditions [71], and temperature [72]. This may be explained by
455 increased biofouling activity or biofilm formation over time on the gills or could be a response to
456 parasite persistence. Unfortunately, we did not measure parasite abundances on the gill, but this
457 would be an important area of research to examine the impact of parasite load on microbiome
458 diversity, or vice versa. Understanding the effects of age on the microbiome was first

459 demonstrated in African turquoise killifish where it was shown that microbiomes from older fish
460 were associated with inflammation in the gut which could be rescued by fecal microbiome
461 transplants from younger fish [73]. It has been suggested that during host ageing, gut
462 communities of vertebrates may shift from commensal to pathogenic leading to increased
463 inflammation and overall dysbiosis [74]. Our results indicate that microbial communities from
464 other body sites may also be influenced by ageing or development of fish and is deserving of
465 additional research. Additional host-associated explanatory variables not measured in our study,
466 include diet or trophic level and host genotype. However, our assessment did determine at least
467 based on mitochondria DNA, that the genetic population of mackerel was heterogeneous which
468 further verifies the importance of environment and fish development stage on driving microbial
469 communities.
470
471 Along with being most influenced by environmental and biological factors, overall the skin and
472 gill communities were more similar phylogenetically to the sea water. Interestingly, of all
473 mucosal sites, digesta samples were most similar to the sediment. This could be explained if the
474 fish were feeding on benthic organisms [75] such as crustaceans buried in the sand. Although not
475 quantified, we did find various types of crustaceans in the stomachs of the larger fish along with
476 occasional gritty material which appeared to be sediment [76]. It is also possible that wave
477 turbulence in nearshore environments where the fish were caught could also cause fish to be
478 exposed to higher sediment levels through resuspension [77]. Since sediments are often
479 repositories of decaying organic matter including anthropogenic contaminants, it is important to
480 consider the negative health implications on a fish population as well as the potential human

481 impacts associated with recreational fishing that occurs in near-shore locations, such as piers,
482 and consumption of these fish.
483
484 Various potential pathogenic and beneficial microbes were persistently abundant across seasons
485 which has important implications for climate change and aquaculture. *Photobacterium damselae*
486 was present in skin and gut communities and in relatively high abundance compared to other
487 microbes. High abundance relative to other microbes and prevalence across fish replicates could
488 have important negative health implications as this is an important globally distributed [53] fish
489 [78] pathogen causing bacterial septicemia that has also been implicated as an important genera
490 for co-evolution in marine hosts [79]. If *P. damselae* is an important host-associated microbe,
491 understanding the conditions by which it becomes pathogenic will be important in modeling
492 fishery impacts. This pathogen has caused financial losses in marine fish farms across numerous
493 species including yellowtail, gilthead seabream, and seabass [80, 81] and is thought to be
494 transferred through water [82] to other species even infecting humans [83]. Chub mackerel are a
495 very important forage fish consumed by many higher trophic level fish including tunas, billfish,
496 and jacks which could have implications for trophic transfer of pathogens warranting future
497 studies. Further, this microbe was most prevalent and abundant during the summer months
498 suggesting it could be associated with high water temperatures and low nutrients. Extending this
499 time series for another 3-10 years will be crucial to continue monitoring. While time series
500 datasets exist for marine free-living microbial communities, few exist for marine vertebrates.
501 Evaluating the extent by which exposure to marine pathogens influences disease is important for
502 estimating impacts to fisheries. Further, as marine aquaculture activities continue to expand in
503 coastal waters, farm monitoring through the host-microbiome could be an important tool for

504 preventing disease outbreaks and massive losses. Experimental mesocosm studies could also be
505 useful to model this marine vertebrate pathogen. Examples in other vertebrates of wide spread
506 prevalence of opportunistic pathogens includes 20-80% carriage rates of *Staphylococcus aureus*
507 in humans [84].
508
509 Some novel candidate symbiotic interactions were discovered when evaluating microbial
510 ecology across the various mucosal sites. In the gills, several *Shewanella spp.* were highly
511 prevalent in mackerel which is consistent with tropical fish microbiome studies [64] suggesting a
512 potential symbiotic role. Some *Shewanella spp*. are common marine genera responsible for
513 eicosapentaenoic acid (20:5n-3, an omega-3 polyunsaturated fatty acid) production [85, 86] and
514 have been documented in freshwater fish [87]. Both Bacillus and Lactobacillus strains make up
515 close to 50% of the microbial taxa in commercially available probiotics for aquaculture [88, 89].
516 Therefore, mucus from wild marine fish could provide a novel source of probiotics for use in the
517 aquaculture industry.
518
519 In our study, we have for the first time evaluated the full mucosal microbiome of a single marine
520 fish species. This design makes it the longest running wild marine fish microbiome study
521 encompassing five body sites within a single species of fish caught over one year. The Pacific
522 chub mackerel microbiome was primarily differentiated by mucosal body site. Environmental
523 conditions and host biology primarily drives the skin and gill mucosal microbiomes with
524 Chlorophyll a concentration, age, and temperature having the broadest effects. Our results
525 provide the foundation to understanding natural microbiome variation of a wild marine fish

526 which is economically important and provides a basis for asking how questions about how
527 climate change may impact the marine fish microbiome in a positive or negative way.
528

## 529 Conflict of interest

530 The authors declare there is no conflict of interest.
531

## 532 Acknowledgements

533

534 This work was supported by grants NSF (OCE-1313747) and NIEHS (P01-ES021921) to E.E.A.
535 We thank the Center for Microbiome Innovation at UC San Diego for support through the
536 Microbial Sciences Graduate Research Fellowship to J.J.M. We thank Karen Minich for graphic
537 design assistance on Figure 1. We thank Shane Poplawski and Rachael Gominsky from JCVI for
538 advice and assistance on using Oxford Nanopore sequencing.
539

## 540 References

541

542 1. Hernández JJC, Ortega ATS, Nations F and AO of the U. Synopsis of Biological Data on the
543 Chub Mackerel (Scomber Japonicus Houttuyn, 1782). 2000. Food & Agriculture Org.
544 2. COLLETTE B. Scombrids of the world. An annotated and illustrated catalogue of tunas,
545 mackerels, bonitos and related species known to date. *FAO Fish Synop* 1983; **125**: 1–137.
546 3. FAO. The State of World Fisheries and Aquaculture 2016 (SOFIA): Contributing to food
547 security and nutrition for all. 2016. FAO, Rome, Italy.

4. Crone PR. PACIFIC MACKEREL (Scomber japonicus) STOCK ASSESSMENT FOR USA MANAGEMENT IN THE 2009-10 FISHING YEAR.

5. MACCALL AD. Recent increased abundance and potential productivity of Pacific mackerel (Scomber japonicus). *Calif Coop Ocean Fish Invest Rep* 1985; **26**: 119–129.

6. Yatsu A, Watanabe T, Ishida M, Sugisaki H, Jacobson LD. Environmental effects on recruitment and productivity of Japanese sardine Sardinops melanostictus and chub mackerel Scomber japonicus with recommendations for management. *Fish Oceanogr* 2005; **14**: 263–278.

7. Yatsu A, Mitani T, Watanabe C, Nishida H, Kawabata A, Matsuda H. Current stock status and management of chub mackerel, Scomber japonicus, along the Pacific coast of Japan. *Fish Sci* 2002; **68**: 93–96.

8. Chen X, Li G, Feng B, Tian S. Habitat suitability index of Chub mackerel (<Emphasis Type="Italic">Scomber japonicus</Emphasis>) from July to September in the East China Sea. *J Oceanogr* 2009; **65**: 93–102.

9. HUNTER JR. Early life history of Pacific mackerel, Scamber japonicus. *Fish Bull* 1980; **78**: 89–101.

10. Fry DH, Roedel PM. Fish Bulletin No. 73. Tagging Experiments on the Pacific Mackerel (Pneumatophorus diego). 1949.

11. Sinclair M, Tremblay MJ, Bernal P. El Niño Events and Variability in a Pacific Mackerel (Scomber japonicus) Survival Index: Support for Hjort's Second Hypothesis. *Can J Fish Aquat Sci* 1985; **42**: 602–608.

12. Ouwerkerk JP, de Vos WM, Belzer C. Glycobiome: Bacteria and mucus at the epithelial interface. *Best Pract Res Clin Gastroenterol* 2013; **27**: 25–38.

13. Beck BH, Peatman E. Mucosal Health in Aquaculture. 2015. Academic Press.

14. Gomez D, Sunyer JO, Salinas I. The mucosal immune system of fish: The evolution of tolerating commensals while fighting pathogens. *Fish Shellfish Immunol* 2013; **35**: 1729–1739.

15. Xu Z, Parra D, Gómez D, Salinas I, Zhang Y-A, von Gersdorff Jørgensen L, et al. Teleost skin, an ancient mucosal surface that elicits gut-like immune responses. *Proc Natl Acad Sci U S A* 2013; **110**: 13097–13102.

16. Nigam AK, Kumari U, Mittal S, Mittal AK. Comparative analysis of innate immune parameters of the skin mucous secretions from certain freshwater teleosts, inhabiting different ecological niches. *Fish Physiol Biochem* 2012; **38**: 1245–1256.

17. Estes L, Elsen PR, Treuer T, Ahmed L, Caylor K, Chang J, et al. The spatial and temporal domains of modern ecology. *Nat Ecol Evol* 2018; **2**: 819–826.

18. Fuhrman JA, Cram JA, Needham DM. Marine microbial community dynamics and their ecological interpretation. *Nat Rev Microbiol* 2015; **13**: 133–146.

19. Gilbert JA, Steele JA, Caporaso JG, Steinbrück L, Reeder J, Temperton B, et al. Defining seasonal marine microbial community dynamics. *ISME J* 2012; **6**: 298–308.

20. Smits SA, Leach J, Sonnenburg ED, Gonzalez CG, Lichtman JS, Reid G, et al. Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania. *Science* 2017; **357**: 802–806.

21. Al-Harbi AH, Naim Uddin M. Seasonal variation in the intestinal bacterial flora of hybrid tilapia (Oreochromis niloticus×Oreochromis aureus) cultured in earthen ponds in Saudi Arabia. *Aquaculture* 2004; **229**: 37–44.

22. Karlsen C, Ottem KF, Brevik ØJ, Davey M, Sørum H, Winther-Larsen HC. The environmental and host-associated bacterial microbiota of Arctic seawater-farmed Atlantic salmon with ulcerative disorders. *J Fish Dis* 2017; **40**: 1645–1663.

23. FULTON T. The rate of growth of fishes. *Twenty-Second Annu Rep* 1904; 141–241.

24. Froese R. Cube law, condition factor and weight–length relationships: history, meta-analysis and recommendations. *J Appl Ichthyol* ; **22**: 241–253.

25. Methot RD, Wetzel CR. Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. *Fish Res* 2013; **142**: 86–99.

26. Minich JJ, Zhu Q, Janssen S, Hendrickson R, Amir A, Vetter R, et al. KatharoSeq Enables High-Throughput Microbiome Analysis from Low-Biomass Samples. *mSystems* 2018; **3**: e00218-17.

27. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* 2012; **6**: 1621–1624.

28. Walters W, Hyde ER, Berg-Lyons D, Ackermann G, Humphrey G, Parada A, et al. Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. *mSystems* 2016; **1**: e00009-15.

29. Minich JJ, Humphrey G, Benitez RAS, Sanders J, Swafford A, Allen EE, et al. High-Throughput Miniaturized 16S rRNA Amplicon Library Preparation Reduces Costs while Preserving Microbiome Integrity. *mSystems* 2018; **3**: e00166-18.

30. Gonzalez A, Navas-Molina JA, Kosciolek T, McDonald D, Vázquez-Baeza Y, Ackermann G, et al. Qiita: rapid, web-enabled microbiome meta-analysis. *Nat Methods* 2018; **15**: 796–798.

31. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010; **7**: 335–336.

32. Amir A, McDonald D, Navas-Molina JA, Kopylova E, Morton JT, Zech Xu Z, et al. Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. *mSystems* 2017; **2**.

33. Janssen S, McDonald D, Gonzalez A, Navas-Molina JA, Jiang L, Xu ZZ, et al. Phylogenetic Placement of Exact Amplicon Sequences Improves Associations with Clinical Information. *mSystems* 2018; **3**: e00021-18.

34. Whittaker RH. Evolution and Measurement of Species Diversity. *Taxon* 1972; **21**: 213–251.

35. Shannon CE. A Mathematical Theory of Communication. *SIGMOBILE Mob Comput Commun Rev* 2001; **5**: 3–55.

36. Faith DP. Conservation evaluation and phylogenetic diversity. *Biol Conserv* 1992; **61**: 1–10.

37. Lozupone C, Knight R. UniFrac: a New Phylogenetic Method for Comparing Microbial Communities. *Appl Environ Microbiol* 2005; **71**: 8228–8235.

38. Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R. UniFrac: an effective distance metric for microbial community comparison. *The ISME Journal*. https://www.nature.com/articles/ismej2010133. Accessed 22 May 2018.

632 39. Vázquez-Baeza Y, Pirrung M, Gonzalez A, Knight R. EMPeror: a tool for visualizing high-
633 throughput microbial community data. *GigaScience* 2013; **2**: 16.
634 40. Kruskal WH, Wallis WA. Use of Ranks in One-Criterion Variance Analysis. *J Am Stat Assoc*
635 1952; **47**: 583–621.
636 41. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA
637 gene database project: improved data processing and web-based tools. *Nucleic Acids Res*
638 2013; **41**: D590–D596.
639 42. McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, et al. An improved
640 Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of
641 bacteria and archaea. *ISME J* 2012; **6**: 610–618.
642 43. Fox J. Applied Regression Analysis and Generalized Linear Models. 2015. SAGE
643 Publications.
644 44. Shapiro SS, Wilk MB. An Analysis of Variance Test for Normality (Complete Samples).
645 *Biometrika* 1965; **52**: 591–611.
646 45. Breusch TS, Pagan AR. A Simple Test for Heteroscedasticity and Random Coefficient
647 Variation. *Econometrica* 1979; **47**: 1287–1294.
648 46. Anderson MJ. A new method for non-parametric multivariate analysis of variance. *Austral*
649 *Ecol* 2001; **26**: 32–46.
650 47. Zakrzewski M, Proietti C, Ellis JJ, Hasan S, Brion M-J, Berger B, et al. Calypso: a user-friendly
651 web-server for mining and visualizing microbiome–environment interactions.
652 *Bioinformatics* 2017; **33**: 782–783.

48. Kuraku S, Zmasek CM, Nishimura O, Katoh K. aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. *Nucleic Acids Res* 2013; **41**: W22-28.

49. Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* .

50. Robinson O, Dylus D, Dessimoz C. Phylo.io: Interactive Viewing and Comparison of Large Phylogenetic Trees on the Web. *Mol Biol Evol* 2016; **33**: 2163–2166.

51. RUIMY R, BREITTMAYER V, ELBAZE P, LAFAY B, BOUSSEMART O, GAUTHIER M, et al. Phylogenetic Analysis and Assessment of the Genera Vibrio, Photobacterium, Aeromonas, and Plesiomonas Deduced from Small-Subunit rRNA Sequences. *Int J Syst Evol Microbiol* 1994; **44**: 416–426.

52. Osorio CR, Collins MD, Toranzo AE, Barja JL, Romalde JL. 16S rRNA Gene Sequence Analysis of Photobacterium damselae and Nested PCR Method for Rapid Detection of the Causative Agent of Fish Pasteurellosis. *Appl Environ Microbiol* 1999; **65**: 2942–2946.

53. Terceti MS, Ogut H, Osorio CR. Photobacterium damselae subsp. damselae, an Emerging Fish Pathogen in the Black Sea: Evidence of a Multiclonal Origin. *Appl Env Microbiol* 2016; **82**: 3736–3745.

54. Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P, et al. Galaxy: A platform for interactive large-scale genome analysis. *Genome Res* 2005; **15**: 1451–1455.

55. Harris RS. Improved Pairwise Alignment of Genomic Dna. 2007. PhD Thesis, Pennsylvania State University.

674 56. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-
675 performance genomics data visualization and exploration. *Brief Bioinform* 2013; **14**: 178–
676 192.
677 57. Ley RE, Lozupone CA, Hamady M, Knight R, Gordon JI. Worlds within worlds: evolution of
678 the vertebrate gut microbiota. *Nat Rev Microbiol* 2008; **6**: 776–788.
679 58. Sullam KE, Essinger SD, Lozupone CA, O'connor MP, Rosen GL, Knight R, et al.
680 Environmental and ecological factors that shape the gut bacterial communities of fish: a
681 meta-analysis. *Mol Ecol* 2012; **21**: 3363–3378.
682 59. Wong S, Rawls JF. Intestinal microbiota composition in fishes is influenced by host ecology
683 and environment. *Mol Ecol* ; **21**: 3100–3102.
684 60. Bevins CL, Salzman NH. The potter's wheel: the host's role in sculpting its microbiota. *Cell*
685 *Mol Life Sci* 2011; **68**: 3675.
686 61. Colston TJ, Jackson CR. Microbiome evolution along divergent branches of the vertebrate
687 tree of life: what is known and unknown. *Mol Ecol* ; **25**: 3776–3800.
688 62. Azam F, Fenchel T, Field JG, Gray JS, Meyer-Reil LA, Thingstad F. The Ecological Role of
689 Water-Column Microbes in the Sea. *Mar Ecol Prog Ser* 1983; **10**: 257–263.
690 63. Kallmeyer J, Pockalny R, Adhikari RR, Smith DC, D'Hondt S. Global distribution of microbial
691 abundance and biomass in subseafloor sediment. *Proc Natl Acad Sci* 2012; **109**: 16213–
692 16216.
693 64. Pratte ZA, Besson M, Hollman RD, Stewart FJ. The Gills of Reef Fish Support a Distinct
694 Microbiome Influenced by Host-Specific Factors. *Appl Environ Microbiol* 2018; **84**.

65. Chiarello M, Auguet J-C, Bettarel Y, Bouvier C, Claverie T, Graham NAJ, et al. Skin microbiome of coral reef fish is highly variable and driven by host phylogeny and diet. *Microbiome* 2018; **6**: 147.

66. Anderson DM, Glibert PM, Burkholder JM. Harmful algal blooms and eutrophication: Nutrient sources, composition, and consequences. *Estuaries* 2002; **25**: 704–726.

67. Minich JJ, Morris MM, Brown M, Doane M, Edwards MS, Michael TP, et al. Elevated temperature drives kelp microbiome dysbiosis, while elevated carbon dioxide induces water microbiome disruption. *PloS One* 2018; **13**: e0192772.

68. Lokmer A, Mathias Wegner K. Hemolymph microbiome of Pacific oysters in response to temperature, temperature stress and infection. *ISME J* 2015; **9**: 670–682.

69. Lozupone CA, Knight R. Global patterns in bacterial diversity. *Proc Natl Acad Sci* 2007; **104**: 11436–11440.

70. Schmidt VT, Smith KF, Melvin DW, Amaral-Zettler LA. Community assembly of a euryhaline fish microbiome during salinity acclimation. *Mol Ecol* 2015; **24**: 2537–2550.

71. Hanek G, Fernando CH. The role of season, habitat, host age, and sex on gill parasites of Lepomis gibbosus (L.). *Can J Zool* 1978; **56**: 1247–1250.

72. Rohde K, Hayward C, Heap M. Aspects of the ecology of metazoan ectoparasites of marine fishes. *Int J Parasitol* 1995; **25**: 945–970.

73. Smith P, Willemsen D, Popkes M, Metge F, Gandiwa E, Reichard M, et al. Regulation of life span by the gut microbiota in the short-lived African turquoise killifish. *eLife* 2017; **6**.

74. Seidel J, Valenzano DR. The role of the gut microbiome during host ageing. *F1000Research* 2018; **7**.

717 75. Weiss R. Feeding behaviour and formation of fish concentrations in the chub mackerel
718 (Scomber colias) in the Northwest African fishing grounds. *ICES CM* 1974; **15**: 6.
719 76. Falk V. Sediments as food of the chub mackerel (Scomber colias Gmelin) off Northwest
720 Africa. *ICES CM* 1967; 5.
721 77. Reise K. Sediment mediated species interactions in coastal waters. *J Sea Res* 2002; **48**:
722 127–141.
723 78. Romalde JL. Photobacterium damselae subsp. piscicida: an integrated view of a bacterial
724 fish pathogen. *Int Microbiol Off J Span Soc Microbiol* 2002; **5**: 3–9.
725 79. Urbanczyk H, Ast JC, Dunlap PV. Phylogeny, genomics, and symbiosis of Photobacterium.
726 *FEMS Microbiol Rev* 2011; **35**: 324–342.
727 80. Magariños B, Toranzo AE, Romalde JL. Phenotypic and pathobiological characteristics of
728 Pasteurella piscicida. *Annu Rev Fish Dis* 1996; **6**: 41–64.
729 81. Romalde JL, Magariños B. Immunization with bacterial antigens: pasteurellosis. *Dev Biol*
730 *Stand* 1997; **90**: 167–177.
731 82. Fouz B, Toranzo AE, Milán M, Amaro C. Evidence that water transmits the disease caused
732 by the fish pathogen Photobacterium damselae subsp. damselae. *J Appl Microbiol* 2000;
733 **88**: 531–535.
734 83. Rivas AJ, Lemos ML, Osorio CR. Photobacterium damselae subsp. damselae, a bacterium
735 pathogenic for marine animals and humans. *Front Microbiol* 2013; **4**.
736 84. Kluytmans J, van Belkum A, Verbrugh H. Nasal carriage of Staphylococcus aureus:
737 epidemiology, underlying mechanisms, and associated risks. *Clin Microbiol Rev* 1997; **10**:
738 505–520.

85. Yazawa K. Production of eicosapentaenoic acid from marine bacteria. *Lipids* 1996; **31**: S297–S300.

86. Nogi Y, Kato C, Horikoshi K. Taxonomic studies of deep-sea barophilic Shewanella strains and description of Shewanella violacea sp. nov. *Arch Microbiol* 1998; **170**: 331–338.

87. Dailey FE, McGraw JE, Jensen BJ, Bishop SS, Lokken JP, Dorff KJ, et al. The Microbiota of Freshwater Fish and Freshwater Niches Contain Omega-3 Fatty Acid-Producing Shewanella Species. *Appl Environ Microbiol* 2015; **82**: 218–231.

88. Gatesoupe FJ. The use of probiotics in aquaculture. *Aquaculture* 1999; **180**: 147–165.

89. Martínez Cruz P, Ibáñez AL, Monroy Hermosillo OA, Ramírez Saad HC. Use of probiotics in aquaculture. *ISRN Microbiol* 2012; **2012**: 916845.

# Figure Legends

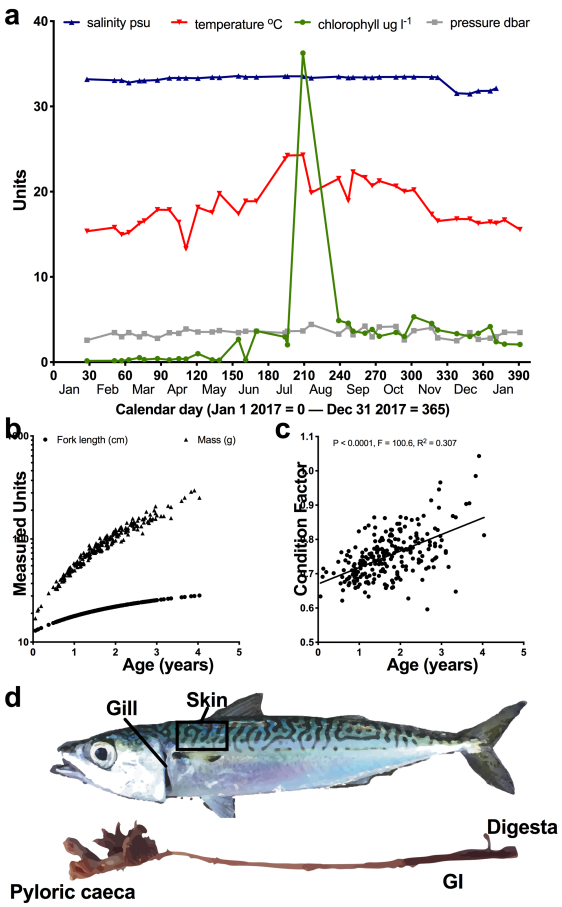



Figure 1

Figure 1. Environmental sampling design. Throughout 2017 and into early 2018, 229 Pacific chub mackerel, *S. japonicus*, were caught across 38 sampling events from the SIO pier. (a) Pier sea water measurements of temperature, salinity, pressure, and chlorophyll a were collected using the scoos.org database. (b) Ages of *S. japonicus* were inferred from fish lengths. (c) Condition factor was calculated for each fish based on length and mass. (d) Mucosal microbiome samples were collected from five body sites including gill, skin, pyloric caeca, GI biopsy, and fecal or digesta material removed from the lower GI.

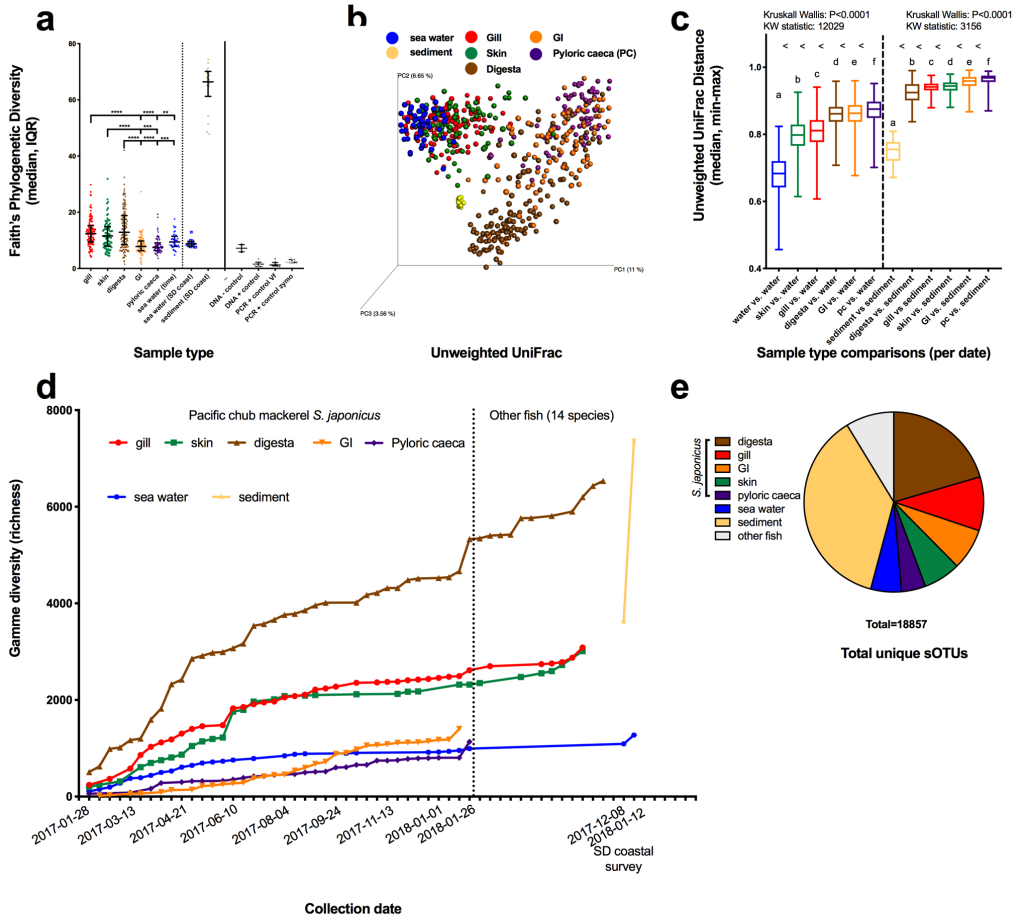



Figure 2

Figure 2. Microbial diversity of coastal environmental controls and *S. japonicus* mucosal microbiome. (a) Alpha diversity was calculated using Faith's Phylogenetic diversity metric in Qiime 1.9.1 with the median and interquartile range displayed. (b) PCoA plot of beta diversity as calculated using Unweighted UniFrac with a rooted phylogenetic tree inserted using the SEPP method in Qiita and Qiime1.9.1. (c) Distances of mucosal microbial communities (gill, skin, digesta, GI, and pyloric caeca) compared to sea water and sediment samples using non-parametric Kruskal-Wallis test. (d) Accumulation of total microbial diversity across chronological sampling events within fish (*S. japonicus* and 14 other species) mucosal sites, water samples, and sediment. (e) Proportion of unique microbial diversity (sOTUs) contributed by body site or environment to the whole dataset.

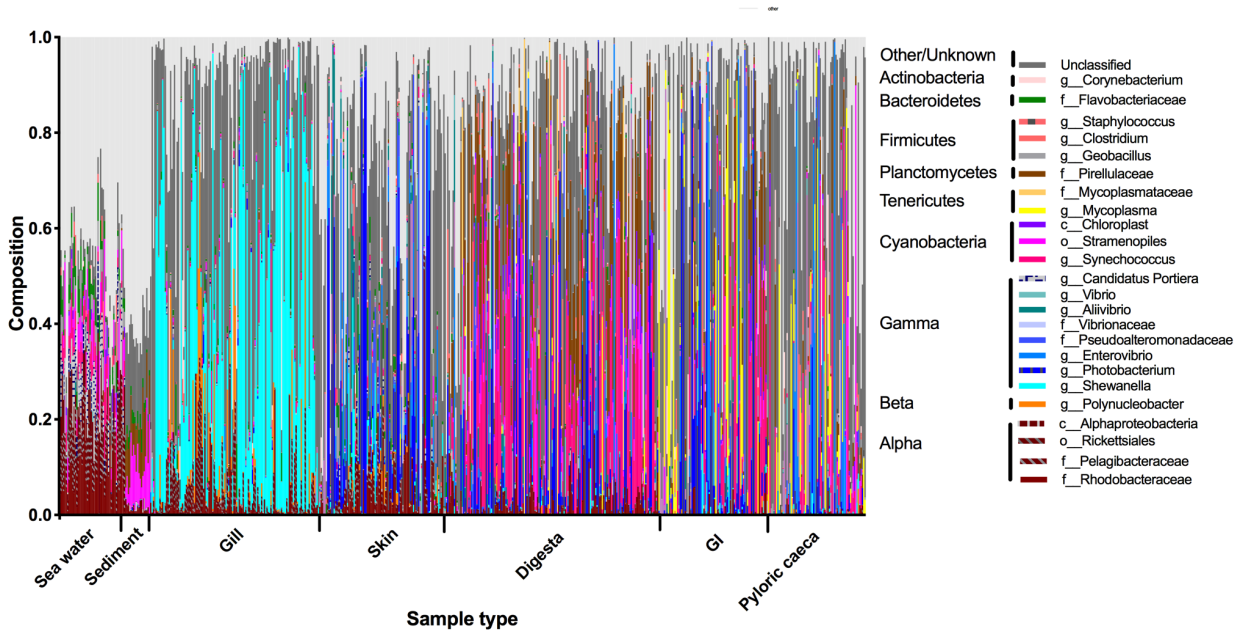
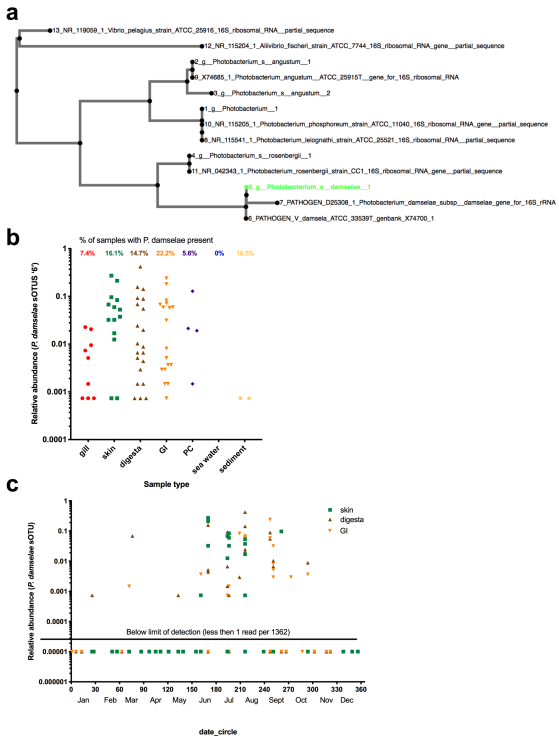



Figure 3

Figure 3. Top 25 most abundant genera across body sites.




788 Figure 4

789 Figure 4. Prevalence of marine vertebrate pathogen, *Photobacterium damselae* on *S. japonicus*
790 body sites throughout the sampling effort. (a) Validation of *P. damselae* 148 bp v4 region by
791 way of phylogenetic comparison to two known pathogen isolates and and non-pathogenic strains.
792 (b) Total relative abundance of *P. damselae* sOTU across five body sites and environments for
793 successfully sequenced samples. Total prevalence or percent of samples with *P. damelsae*
794 present is also calculated for each sample type and displayed at top of graph (7.4% gill samples,
795 16.1% skin, 14.7% digesta, 22.2% GI, 5.6% pyloric caeca, 0% water, 10.5% sediment). (c)
796 Proportion of microbial community comprised of *P. damselae* sOTU across the most prevalent
797 body sites (skin, digesta, GI) over the sampling effort of 1 year. Relative abundance is calculated
798 as number of *P. damselae* sOTU reads divided by 1362, the rarefraction number. Any samples
799 with 0 *P. damselae* reads are considered under the detection limit and are displayed as equal to
800 0.00001 relative abundance in order to visualize on the log scale.

Table 1. . Quantification of environmental and biological variables on fish mucosal microbiomes as measured by (a) alpha diversity with Generalized Linear Model and (b) beta diversity with multivariate statistics (Adonis).

| | | | | | | Host (biometrics) | | | | Environmental (water conditions) | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | Age (years) | FL mm | Mass kg | KNA | Chl a µg $L^{-1}$ | Press dbar | Sal psu | Temp °C |
| a | Body site | Alpha | Adj. $R^2$ | F-stat | P value | P value | | | | | | | |
| | gill* | Shan | 0.13 | 2.89 | 0.006 | | | | | | | | |
| | skin | Shan | 0.38 | 6.60 | <0.0001 | | | | | <0.001 (-) | | | <0.001 (+) |
| | digesta | Shan | -0.01 | 0.84 | 0.567 | | | | | | | | |
| | GI | Shan | -0.01 | 0.94 | 0.494 | | | | | | | | |
| | PC | Shan | 0.05 | 1.35 | 0.244 | 0.027 (+) | 0.026 (+) | | | | | | |
| | gill | PD | 0.33 | 7.04 | <0.0001 | | | | | 0.005 (-) | | | |
| | skin | PD | 0.26 | 4.24 | <0.0001 | 0.008 (-) | | | | 0.002 (-) | | | 0.044 (+) |
| | digesta | PD | 0.02 | 1.29 | 0.258 | | | | | | 0.036 (-) | | |
| | GI** | PD | -0.01 | 0.91 | 0.514 | | | | | | | | |
| | PC | PD | 0.22 | 2.97 | 0.009 | 0.048 (+) | | | | | | 0.049. (-) | 0.003 (+) |
| b | Body site | Beta | Total $R^2$ | | P value | P value | | | | | | | |
| | gill | uUniF | 0.12 | | <0.0001 | <0.001 | | | | <0.001 | | | |
| | skin | uUniF | 0.15 | | <0.0001 | <0.001 | <0.001 | | | <0.001 | | | |
| | digesta | uUniF | 0.09 | | <0.0001 | | | | | | | | |

| | | | | | |
|---|---|---|---|---|---|
| GI | u UniF | 0.14 | 0.099 | | |
| PC | u UniF | 0.15 | 0.32 | | |
| gill | w UniF | 0.14 | <0.0001 | <0.001 | |
| skin | w UniF | 0.20 | 0.001 | <0.001 | |
| digesta | w UniF | 0.10 | 0.038 | | |
| GI | w UniF | 0.13 | 0.38 | | |
| PC | w UniF | 0.14 | 0.49 | | |

* Shannon diversity of gill sample results excluded from analysis because residuals are non-normal & homoscedasticity

** Faiths Phylogenetic Diversity for GI sample results excluded from analysis because residuals are non normal

Alpha: Shan = Shannon, PD = Faiths Phylogenetic diversity;

Beta: u UniF = unweighted UniFrac, w UniF = weighted UniFrac

P value: (-) indicates a negative association, (+) indicates a positive association

Transformation:Faiths PD is log transformed

Host biometrics: FL = fork length in mm, K = fultons condition factor

Environmental factors: Press = pressure, Sal = salinity, Temp = temperature

Shapiro: (Residual normality > 0.05)

Breusch-Pagan: (homoscedasticity > 0.05)

801