

Bacterial epibiont communities of panmictic Antarctic krill are spatially structured

Laurence J. Clarke^{1,2,3*}, Léonie Suter¹, Rob King¹, Andrew Bissett⁴, Sophie Bestley², Bruce E. Deagle^{1,4}

¹ Australian Antarctic Division, Channel Highway, Kingston, Tasmania 7050, Australia

² Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Tasmania 7001, Australia

³ Antarctic Climate & Ecosystems Cooperative Research Centre, University of Tasmania, Hobart, Tasmania 7001, Australia

⁴ Commonwealth Scientific and Industrial Research Organisation, Hobart, Tasmania 7004, Australia

*Corresponding author: Laurence J. Clarke, Australian Antarctic Division, Kingston, Tasmania 7050, Australia. Ph: +61 (3) 6232 3117

Email: laurence.clarke@utas.edu.au

Short title: Krill microbiomes are spatially structured

Abstract

Antarctic krill (*Euphausia superba*) are amongst the most abundant animals on Earth, with a circumpolar distribution in the Southern Ocean. Genetic and genomic studies have failed to detect any population structure for the species, suggesting a single panmictic population. However, the hyper-abundance of krill slows the rate of genetic differentiation, masking potential underlying structure. Here we use high-throughput sequencing of bacterial 16S rRNA genes to show that krill bacterial epibiont communities exhibit spatial structuring, driven mainly by distance rather than environmental factors, especially for strongly krill-associated bacteria. Estimating the ecological processes driving bacterial community turnover indicated this was driven by bacterial dispersal limitation increasing with geographic distance. Furthermore, divergent epibiont communities generated from a single krill swarm split between aquarium tanks under near identical conditions suggests physical isolation in itself can cause krill-associated bacterial communities to diverge. Our findings show that Antarctic krill-associated bacterial communities are geographically structured, in direct contrast with the lack of structure observed for krill genetic and genomic data.

Introduction

Antarctic krill (*Euphausia superba*) is a keystone species of Southern Ocean food webs (Croxall *et al.* 1999), with an estimated biomass of 379 million tonnes (Atkinson *et al.* 2009). The commercial Antarctic krill fishery catch is the largest in the Southern Ocean and has been expanding in recent years (Nicol & Foster 2016). However, management of the krill fishery is hampered by lack of knowledge regarding krill population structure and the number of effective stocks.

Krill have a circumpolar, but non-uniform, distribution, with highest densities in the South Atlantic near the Antarctic Peninsula and Scotia Arc (Siegel & Watkins 2016). The presence of persistent areas of high krill density has led to suggestions that there may be between three and six separate krill stocks in the Southern Ocean (Latogurski 1979; Mackintosh 1973), but this has not been supported by genetic

and genomic studies over the last four decades (Deagle *et al.* 2015). The current consensus is that the genetic and genomic data are consistent with a single panmictic circumpolar population, however, any demographic structure may be obscured by the hyper-abundance of Antarctic krill (Jarman & Deagle 2016). Genetic drift in neutral markers is inversely related to population size, and the vast number of individual krill would slow genetic differentiation between subpopulations, even in the absence of homogenising gene flow (Deagle *et al.* 2015).

All eukaryotes carry a characteristic bacterial microbiome. Bacterial populations often exhibit more genetic structuring than those of higher eukaryotes, as large bacterial populations accumulate diversity that tends to create barriers to recombination between lineages (Yang *et al.* 2019). Spatial structuring of microbiomes has been demonstrated in several marine invertebrates (Cregeen 2016; Zwirgmaier *et al.* 2015), and can arise as a result of either environmental (including host) heterogeneity combined with selection, and/or dispersal limitation combined with ecological drift (Stegen *et al.* 2015). In a recent study, we showed that four krill swarms separated by 260-1270 km each supported distinct microbiomes, with the strongest differences observed in bacteria associated with the exoskeleton (Clarke *et al.* 2019). In this study, we expand the number of swarms and the number of krill sampled per swarm, and examine a greater range of geographic distances to test whether krill-associated bacteria show spatial structuring, and whether the observed structure is more closely related to geographic or environmental distance. In supporting experiments, we investigate whether divergent krill microbiomes can be generated in aquaria under near identical conditions. Our results demonstrate krill microbiomes in different areas represent distinct communities, and may also be functionally different; these different microbe-host associations in different spatial areas may be meaningful as conservation units.

Results

We analysed the bacterial communities on 155 krill moults collected from krill from 13 swarms (n=11-12 moults per swarm) separated by 4-3481 km (Fig. 1a). Following filtering and quality control, this dataset included 4.49 million Illumina paired-end sequencing reads (10,787 to 87,547 reads per sample) representing 3,808 zero-radius OTUs (zOTUS). Molt bacterial communities were dominated by Gammaproteobacteria (mean relative abundance 53%), followed by Bacteroidia (27%) and Alphaproteobacteria (7%). Swarm membership (trawl) explained half the variance in molt microbiome composition based on weighted UniFrac distance, and more than 40% for unweighted UniFrac distance (weighted UniFrac: $F_{12,142}=11.85$, $R^2=0.500$, $P<0.001$, unweighted UniFrac: $F_{12,142}=8.18$, $R^2=0.409$, $P<0.001$). Ecological (unweighted UniFrac) distance between molt bacterial communities increased with increasing geographic distance between trawls (Fig. 2a). An UPGMA tree based on unweighted UniFrac distance showed bacterial communities clustered by geographic region (Fig. 1a and b), with 3-5 swarms sharing similar microbiomes near each of Mawson, Casey and Dumont d'Urville stations in East Antarctica. A single swarm near BANZARE Bank north-east of Mawson station (K29) supported a distinct microbiome, though most closely related to the swarms sampled off Mawson (Fig. 1b). The spatial structuring of krill microbiomes from the vicinity of Casey and Mawson is in direct contrast to the lack of structure observed in SNP markers for krill collected from the same two regions (Deagle *et al.* 2015) (Fig. 3).

Although we have previously demonstrated that krill molt bacterial communities are distinct from seawater bacterial communities (Clarke *et al.* 2019), molt bacterial communities are an open system that can exchange bacteria with the surrounding seawater. To investigate if variation in seawater bacteria could be driving the observed geographic structuring, we repeated the above analyses using only strongly krill-associated bacteria (present in $\geq 50\%$ of samples) to reduce the potential influence of local environmental bacteria. This reduced the dataset from 3,808 to 267 zOTUs but, as these were the more common bacteria, still represented 73.3% of reads. The entire molt and strongly krill-associated bacterial communities were taxonomically very similar (see Supporting Information Text)

and strongly krill-associated bacterial communities clustered by region as per Fig. 1b (Fig. S1). Swarm membership explained approximately half the variance in strongly krill-associated bacterial community composition (weighted UniFrac: $R^2=0.469$, unweighted UniFrac: $R^2=0.529$, $P<0.001$ for both). Unweighted UniFrac distance between strongly krill-associated bacterial communities also increased with geographic distance between trawls (Fig. S2a).

Identity of dominant core zOTUs differs between regions

The core microbiome (zOTUs present in 85% or more of the moult samples) comprised only 21 zOTUs representing 30.9% of reads. The core microbiome was dominated by zOTUs from the families Colwelliaceae (13 zOTUs), Rhodobacteraceae (4 zOTUs, 2 *Sulfitobacter*) and Flavobacteriaceae (3 zOTUs, 2 *Polaribacter*). Most of the Colwelliaceae zOTUs (11/13) had 97-99% identity to two closely related (99% identity) *Colwellia* species isolated from Antarctic sea-ice assemblages (Bowman *et al.* 1998) (Table S1). Other core zOTUs were highly similar (99-100% pairwise identity) to bacterial isolates from polar marine invertebrates ($n=3$ zOTUs) and to Antarctic sea ice bacteria ($n=3$ zOTUs), indicating these high nutrient surface environments in polar oceans are commonly colonised by similar bacterial assemblages. Interestingly, despite being present in the majority of samples and representing a high proportion of the total reads, the identity of dominant core zOTUs varied between geographic regions (Fig. 4). Four core zOTUs (three *Colwellia* and one *Sulfitobacter*) were identified as having significantly different relative abundances between regions (LEfSe, LDA > 3, $\alpha < 0.01$). Inferred metagenomic traits also suggest krill moult bacterial communities from different regions are functionally distinct (Fig. S3 and 4).

Environmental and spatial effects

We used linear mixed-effects (LME) modelling to test whether spatial (geographic distance) and/or environmental variables (ocean temperature, salinity, sea-ice melt and productivity) explained variation in krill microbial community composition between trawl sites. Geographic distance (LME

model coefficient estimate: 0.065 ± 0.003 [0.060–0.070 95%CI]) and salinity (0.009 ± 0.003 [0.003–0.014 95%CI]) were the only significant predictors for moult community composition ($z = 23.01$, $p < 0.001$ and $z = 3.19$, $p = 0.0014$, respectively, Table S4), based on model-averaging over the top-ranked models ($n=8$) using $\Delta AICc \leq 10$. In contrast, geographic distance (0.066 ± 0.003 [0.059–0.073 95%CI]) was the only significant predictor for strongly krill-associated bacterial communities ($z = 18.8$, $p < 0.001$, Table S6), with weaker evidence for an effect of salinity depending on the random-effects specification (custom covariance structure: z -value=1.48, p -value=0.14, Table S6; crossed effects z -value=2.21, p -value=0.027, Table S10). Results were very consistent overall using the two different random-effects specifications (Table S3-10). All top-ranking (i.e., lower AICc) models included geographic distance (rather than only environmental variables) for both the entire moult community and strongly krill-associated bacteria.

Krill microbiome differentiation driven by dispersal limitation over large geographic distances

We used a combination of null models to estimate the contribution of homogeneous or variable selection, homogenising dispersal and dispersal limitation to krill microbiome community turnover (Stegen *et al.* 2013; Stegen *et al.* 2015). In our data, bacterial communities on krill with increased geographic separation showed increased dispersal limitation and decreased influence of homogenising dispersal, homogeneous selection and undominated processes. Undominated processes were the major influence for moult communities sampled from krill within a trawl (52%, Fig. 2b). Comparing samples between trawls within a cluster (e.g., Mawson, Casey and Dumont d’Urville), the influence of dispersal limitation increased from 9 to 29%, with decreased influence of homogenising dispersal and undominated processes. Dispersal limitation was the main factor when comparing communities between clusters (57%), with a further decrease in the influence of undominated processes (18%). The contribution of homogeneous selection and homogenising dispersal also decreased from 11 to 2% and 11 to 0.3%, respectively, with increasing geographic

separation. Strongly krill-associated bacteria showed similar patterns of increasing dispersal limitation with increasing geographic distance (Fig. S2b).

Krill in identical aquarium conditions develop divergent microbiomes

To investigate the process of microbiome differentiation, krill from a single swarm off Casey station (T04, see Fig. 1a) were split across four 200 L aquarium tanks, with approximately 400 krill per tank, and with consistent conditions for seven months prior to sampling. Krill from each tank developed significantly different epibiont bacterial communities, with twice the proportion of variation explained by tank of origin when relative abundance was taken into account, compared to presence/absence of taxa (weighted UniFrac (WU): $F_{3,44}=12.09$, $R^2=0.452$, $P<0.001$, unweighted UniFrac (UU): $F_{3,44}=4.57$, $R^2=0.237$, $P<0.001$, Fig. 5a and b). Bacterial communities in water within the four tanks were clearly distinct from krill communities (Fig. S5, WU: $F_{1,54}=46.56$, $R^2=0.463$, $P<0.001$, UU: $F_{1,54}=11.91$, $R^2=0.181$, $P<0.001$), but were also significantly different between tanks (WU: $F_{3,4}=2.87$, $R^2=0.682$, $P=0.015$, UU: $F_{3,4}=1.65$, $R^2=0.553$, $P=0.024$). However, this was largely driven by tank 2A, as tank water bacterial communities were not significantly different when this tank was excluded (2B, C and D, WU: $F_{2,3}=2.78$, $R^2=0.650$, $P=0.13$, UU: $F_{2,3}=1.56$, $R^2=0.510$, $P=0.067$), whereas epibiont communities were distinct (WU: $F_{2,33}=11.50$, $R^2=0.411$, UU: $F_{2,33}=4.91$, $R^2=0.229$, $P<0.001$ for both).

Comparing the ecological processes contributing to community turnover within versus between tanks, the contribution of homogenising dispersal decreased from 17 to less than 1%, whereas dispersal limitation increased from 9 to 20% (Fig. 5c). Interestingly, the influence of variable selection (reflecting different environments) decreased from 7% within tanks to 1% between tanks, supporting the notion that the tanks represented highly similar environments.

Discussion

We show that Antarctic krill-associated bacterial communities are geographically structured, in direct contrast with the lack of structure observed for krill genetic and genomic data (Deagle *et al.* 2015). Individual taxa and communities within the human microbiome can be used as markers of migration and ancestry (Dominguez-Bello & Blaser 2011; Henne *et al.* 2014), allowing resolution of migration events previously indistinguishable with human genetic markers (Moodley *et al.* 2009; Wirth *et al.* 2004). The potential for members of the human microbiome to accurately trace human migrations is dependent on the degree to which they are vertically transmitted (Dominguez-Bello & Blaser 2011). In contrast, structuring of krill-associated bacterial communities is driven by homogenising dispersal within a krill swarm (horizontal transmission), but increasing dispersal limitation with increasing geographic distance (Fig. 2). Our findings support the conclusion from a study of bat fur microbiomes (Kolodny *et al.* 2018) that for species experiencing high levels of homogenising dispersal between individuals, the colony or swarm rather than the individual may be the meaningful biological unit on which selection influences the microbiome. Krill microbiome dynamics can thus be understood best in light of metacommunity theory, being influenced by both processes within the community (e.g., competition) and dispersal between communities (Burns *et al.* 2017; Miller *et al.* 2018).

The significant effect of spatial distance on both moult and strongly krill-associated bacterial communities, but limited evidence for environmental influence on the latter suggests different environments are not essential for the development of distinct bacterial communities. This is further supported by development of divergent microbiomes in aquarium krill populations under near-identical conditions. The contribution of variable selection, reflecting community differences arising due to selection in different environments, showed a modest increase comparing bacterial communities within a swarm versus those in separate bacterial geographic clusters (17 to 21% for entire moults communities, 23 to 26% for strongly krill-associated bacteria). In contrast, the contribution of dispersal limitation showed a much larger increase from 9 to 57% (Fig. 2b). If

environmental factors were driving variation in krill microbiomes, we would expect swarms in different water masses to have distinct bacterial communities, as the Southern Ocean fronts separate waters with distinct physical and chemical properties. However, the Casey and Dumont d'Urville clusters include swarms north and south of one or more oceanographic fronts (Fig. 1a). Similarly, LEfSe analysis showed that swarms north or south of the southern boundary of the Antarctic Circumpolar Current were not enriched in particular zOTUs (Table S11). Rather, clusters were mainly distributed west to east rather than north to south, consistent with meso-scale circulation in the region. In particular, Mawson (and BANZARE Bank) clusters correspond to the Prydz Bay gyre, and the Casey cluster corresponds to recirculation within the Australian-Antarctic basin (Bindoff *et al.* 2000). A series of cyclonic eddies off the Adelie Coast that terminates near 130° E may link krill swarms in the Dumont d'Urville cluster, whilst isolating them from sites further west (Aoki *et al.* 2007). However, the influence of circulation patterns, environmental factors and variable selection on krill microbiomes warrants further investigation. The difference observed for aquarium populations between distance measures including or excluding relative abundance suggests ecological drift may be driving divergence in isolated populations in the absence of environmental selection. Weighted UniFrac distance (including relative abundance) explained twice the proportion of variation between tanks compared to unweighted UniFrac distance (a presence/absence metric), suggesting changes in relative abundance precede changes in community membership.

Ecological drift in krill epibiont communities is likely exacerbated by the regular moulting of the exoskeleton, which occurs every 12-30 days during summer (Reiss 2016). Characterising bacteria from moulted exoskeletons means samples represent bacterial communities at the same stage of the moult cycle. However, this was not the case for the aquarium samples. The occurrence of ciliate epibionts decreases from 66% in pre-moult individuals to 0% in post-moult krill (Tarling & Cuzin-Roudy 2008), with bacterial epibionts likely to experience a similar, but potentially less extreme, drop in abundance. Subsequent colonisation may then occur essentially at random by whichever species gets there first,

following the “competitive lottery model” (Burke *et al.* 2011; Hubbell 2001), although the presence of protozoan epibionts could also influence bacterial moult community composition. The existence of a core microbiome is consistent with the need for species to have similar ecologies to compete for the same vacant niche. Given krill swarms can exceed densities of 1000 individuals m⁻³ (Nowacek *et al.* 2011; Tarling *et al.* 2009), colonisation is most likely to occur from epibionts on neighbouring krill, hence the greater influence of homogenising dispersal for bacterial communities within a swarm (Fig. 2b). Homogenising dispersal should have less influence in smaller or less dense swarms, leading to faster rates of drift. Although bacteria on model marine (chitin) particles exhibit rapid and reproducible succession patterns (Datta *et al.* 2016), further aquarium studies are required to better understand krill epibiont colonisation dynamics.

The fishery for Antarctic krill is the largest by tonnage in the Southern Ocean (Nicol & Foster 2016). The krill fishery is managed using regional catch limits, but whether there is a single circumpolar krill stock or if there are distinct populations within each region is unclear (Jarman & Deagle 2016). We have shown that krill microbiomes in different regions represent distinct communities, indicating a non-homogeneous mixture throughout their range. Microbiomes may also be functionally different and furthermore affect host fitness. It seems plausible that different microbe-host associations in different regions may be meaningful as conservation units. Current spatial management of the krill fishery is focussed within the southwest Atlantic, where small-scale management units have been established. Our results suggest that the Indian Sector of the Southern Ocean contains at least two distinct spatial clusters of krill bacterial communities, reflecting different microbe-host associations, near Casey and Dumont d’Urville which both lie within a single large management division. A better understanding of how microbiomes vary at different temporal scales, including throughout the moult cycle, and both seasonal and interannual variation, will be necessary to understand how to apply microbe-host associations to inform regional krill fishery management.

There are many ways population connectivity can be measured; from ancient linkages through phylogeography, to more recent linkages with population genetics, and contemporary movement through animal tracking. Several studies have now demonstrated links between host population genetic structure and host-associated microbiome composition in marine species (Díez-Vives *et al.* 2020; Easson *et al.* 2020; Fietz *et al.* 2018). Our study does not explicitly test whether there is a direct association between krill moult microbiomes and krill population dynamics; indeed this is not possible given that we have little independent information about krill population dynamics. However, we consider that some link between krill microbiomes and population connectivity is a reasonable hypothesis given the apparent lack of an environmental driver and observed isolation by distance. Future studies should explore and ideally test whether there is an association between krill microbiomes and krill population dynamics, especially in the key krill fishery areas of the Antarctic Peninsula and Scotia Arc. Supporting aquarium experiments that investigate horizontal transmission rates between krill with distinct microbiomes, and between krill and the water column, would also assist establishing the scope of inference.

The potential for long-distance dispersal in many marine taxa leads to high population connectivity, making it difficult to resolve population structure with classic genetic approaches (Kelley *et al.* 2016). An improved view of host-microbiome structuring could expand our understanding of system connectivity in the marine environment. Studying the structure of the bacterial communities that effectively hitch a ride with their host may allow previously unseen linkages and barriers to dispersal to be identified.

Materials and Methods

Sample collection

Samples were collected on board the *RSV Aurora Australis* during voyage 3 between 31 January and 19 February 2016 (Kerguelen Axis voyage), and voyage 2 between 16 December 2016 and 16 January

2017 (Totten Glacier voyage). Antarctic krill were sampled from 13 swarms across the Indian sector of the Southern Ocean (Fig. 1a) using targeted trawls with a Rectangular Mid-water Trawl 8+1 (RMT-8+1 metre square) net. Trawls were made on acoustically identified targets at depths between the surface and 60-70 m. The pairwise distance between trawls ranged from 4 to 3481 km.

In order to isolate moults, live krill were transferred immediately after capture to 250 mL jars (one krill per jar) which were ventilated with small holes to allow seawater exchange as per Virtue et al. (2010). The jars were incubated in a large (1600 L) flow-through seawater tank close to ambient ocean temperature (approx. 1 °C) with no additional food provided (Kawaguchi *et al.* 2006). Jars were inspected for moults at 12 hour intervals, with the first 12 animals to moult from each trawl sampled for microbial community profiles (all collected within 48 hours). Moults were removed from the jar, rinsed with 0.22 µm-filtered seawater and stored separately in liquid nitrogen before being stored at -86 °C on return to Australia.

Aquarium samples

Krill from a single swarm caught off Casey station mid-December 2016 (T04 in Fig. 1a) were split across four 200 L krill aquaria tanks at the end of January 2017, with approximately 400 krill per tank. Krill in each tank experienced identical conditions for seven months, including light environment, food (phytoplankton), with the same filtered and UV-sterilised water supply for each tank (Kawaguchi *et al.* 2010; King *et al.* 2003). In September 2017, 12 krill per tank were individually netted and both lateral surfaces swabbed with Epicentre Catch-All™ sample collection swabs for approximately 10 seconds. Krill were then sexed and staged. Swabs were transferred to sterile Eppendorf tubes and stored at -86 °C. Nets were treated with 1% bleach then 0.5% sodium thiosulphate and rinsed with reverse-osmosis and deionised water between individuals. Tank water bacterial communities were sampled by filtering 1 to 1.6 L onto 0.22 µm Sterivex™ filters, with two samples per tank. Filters were transferred to tubes prior to extraction.

DNA extraction, PCR amplification and high-throughput sequencing

Swabs (n=48) were sent to the Australian Genome Research Facility (AGRF, Adelaide, Australia; <http://www.agrf.org.au>) on dry ice, and moults (n=155), Sterivex filters (n=8) and extraction controls (1 ml ethanol) were sent in ethanol on dry ice. DNA was extracted from all samples using the QIAGEN DNeasy PowerLyzer PowerSoil kit. DNA concentrations were quantified using a NanoDrop ND-8000 Spectrophotometer (ThermoFisher Scientific). PCR amplification, amplicon purification and high-throughput sequencing of bacterial 16S V1-3 rRNA (primers 27F: AGAGTTTGATCMTGGCTCAG, Lane 1991; and 519R: GWATTACCGCGGCKGCTG, Lane *et al.* 1985) were carried out at the Ramaciotti Centre for Genomics (Sydney, Australia) on an Illumina MiSeq following the Australian Marine Microbes protocol (Brown *et al.* 2018).

Data analysis

DNA sequence processing and taxonomic assignment followed the Australian Marine Microbial Biodiversity Initiative workflow (Brown *et al.* 2018), with data presented as amplicon sequence variants, or zero-radius operational taxonomic units (zOTUs, Edgar 2016), to maximise potential phylogenetic resolution. In brief, paired-end reads were merged, short sequences (<400 base pair, bp) and sequences containing N's or homopolymer runs >8 bp were removed. Sequences were de-replicated and those with <4 representatives removed. Chimeras were removed and zOTUs identified using the UNOISE3 algorithm (Edgar 2016). Quality-filtered sequences (including those <4 representatives) were mapped to the zOTUs to create a sample-by-read abundance matrix. Taxonomy was assigned to each zOTU using the RDP Bayesian classifier (Wang *et al.* 2007) based on the SILVA v132 database (Yilmaz *et al.* 2014) and 60% probability cut-off. Lastly, zOTUs present in only one sample or with less than 10 reads across the dataset were removed. The number of reads per sample was 10,787-87,547 for moults, 13,083-72,384 reads for swabs, and 5,107-10,215 for the Sterivex filters, with the number of observed zOTUs approaching saturation at the minimum read depth for

each sample type (Fig. S6). DNA sequence data for this study can be found in the NCBI database under BioProject ID: PRJNA505226. The zOTU table and mapping file are available on the Dryad data repository (Clarke *et al.* 2020).

Differences in entire moult and strongly krill-associated (present in $\geq 50\%$ of samples) bacterial community composition between swarms were explored using weighted and unweighted UniFrac distances (Lozupone & Knight 2005) in QIIME v1.8.0 (Caporaso *et al.* 2010) (beta_diversity_through_plots.py) based on a rarefied zOTU table (10,000 reads or 5,000 reads for entire moult and strongly krill-associated bacteria, respectively), with strength and significance of swarm assessed using the Adonis method (Anderson 2001) (compare_categories.py, 999 permutations). The same method was used to characterise differences between krill microbiomes from separate aquarium tanks, with the zOTU table rarefied to 5,000 reads for analyses including the Sterivex filters. The phylogenetic tree used for the UniFrac analysis and to estimate the contribution of selection and dispersal processes to microbiome turnover (see below) was generated in QIIME using PyNAST (Caporaso *et al.* 2009) to align sequences against the Greengenes (v13_8) core set (McDonald *et al.* 2012), then filtering the alignment (removing 0.0005% most variable positions and those that were $>80\%$ gaps) and building the tree using FastTree 2.1.3 (Price *et al.* 2010).

We defined core moult microbiome membership as zOTUs present in 85% or more moult samples ($>131/155$ samples) using QIIME (compute_core_microbiome.py). Ecological dissimilarity between krill moult bacterial communities was visualised by an UPGMA tree based on unweighted UniFrac distance created using the jackknifed_beta_diversity.py workflow in QIIME, with subsampling repeated 100 times. The Linear Discriminant Analysis (LDA) Effect Size (LEfSe, Segata *et al.* 2011) method was used to highlight zOTUs that showed different abundances between bacterial geographic clusters (Fig. 1). Default settings were used except the LDA threshold was increased to 3.0 and the α -value reduced to 0.01 to highlight the most significant taxa discriminating between clusters.

We tested whether spatial (geographic distance) and/or environmental variables explained variation in krill microbial community composition between trawl sites using a linear mixed-effects (LME) modelling approach (R package 'lme4', Bates *et al.* 2015). Analyses were performed separately for both the entire moult community and strongly krill-associated bacteria. The response variable was calculated as the mean unweighted UniFrac distance across all pairwise krill samples for each pair of trawl sites (see Fig. 2a). We included a random-effects term for the non-independent error structure of pairwise datasets (Clarke *et al.* 2002; Row *et al.* 2017) by setting up the covariance structure such that a proportion (ρ_i) of the total variance (σ^2) is due to the correlation between data points that share a common site. Thus, the covariance for n pairwise data points that share a common site is $\rho_i\sigma^2$ and zero for those that do not (Clarke *et al.* 2002; Van Strien *et al.* 2012). These results were compared against LME models that more straightforwardly specified the two trawl sites as crossed random effects.

Predictor variables considered as fixed-effects included geographic distance (km) between trawl sites and a suite of environmental variables obtained from both ship-based underway (sea surface temperature [°C] and salinity [on the practical salinity scale, PSS]) and satellite-derived data (time [weeks] since sea-ice melt, surface chlorophyll-*a* [mg m^{-3}]). Full descriptions of these environmental data are given in the Supporting Information. Euclidean distances were calculated for all environmental variable pairs between sites. We centred and scaled all predictor variables and ensured that the maximum correlation between predictor variables (all Pearson correlations <0.66) and variance inflation factors (all VIFs ≤ 1.52) were acceptably low prior to model fitting (Zuur *et al.* 2010).

We adopted a multi-model inference and model averaging approach (Burnham & Anderson 2002; Burnham *et al.* 2011) using the R package 'MuMIn' (Barton 2019) to generate a set of models with all combinations of fixed-effects. We calculated the marginal (R_m^2) and conditional (R_c^2) coefficient of

determination (Nakagawa & Schielzeth 2013) for all models, and for model evaluation used the second-order Akaike Information Criterion (AICc) suitable for small sample sizes, ranking the models via Akaike weights. For model averaging we included all models with a $\Delta AICc \leq 10$ from the top-ranked model. We calculated the 95% confidence interval (CI) of the model-averaged regression coefficients (full average) and considered those which did not straddle zero to have a significant effect on the response under investigation. Results are presented in text as LME coefficient estimate \pm SE [95% CI].

We used the procedures described by Stegen *et al.* (2013; 2015) to estimate the contribution of selection and dispersal processes to krill microbiome assembly. Variable selection generates divergent communities due to distinct selective environments, whereas homogeneous selection generates convergent communities due to their presence in similar selective environments. Similarly, divergent communities arising primarily due to low rates of dispersal, which causes communities to drift apart, is referred to as 'dispersal limitation', whereas high dispersal preventing drift is referred to as 'homogenising dispersal'. The fraction of community turnover where neither selection nor dispersal is the primary driver of community turnover (e.g., due to the combination of moderate dispersal and weak selection) is referred to as 'undominated' (called 'drift' in Stegen *et al.* 2013; Stegen *et al.* 2015).

Firstly, we computed the between-community mean nearest taxon distance (β MNTD), the mean phylogenetic distance between each zOTU in one community and its closest relative in a second community. A null-model distribution of this parameter was generated by randomly shuffling zOTUs across the tips of the phylogeny (999 permutations). β -nearest taxon indices (β NTI) were calculated as the difference between the observed β MNTD and the mean of the null distribution, expressed in units of standard deviations; β NTI values < -2 or $> +2$ were deemed significant deviations and indicative of homogeneous selection or variable selection, respectively. Where pairwise comparisons showed $|\beta$ NTI| < 2 , we used the Bray-Curtis-based Raup-Crick metric (Chase *et al.* 2011; Stegen *et al.* 2013) to compare observed and expected turnover without using phylogenetic information to infer the

contribution of homogenising dispersal, dispersal limitation, or undominated processes (moderate dispersal and weak selection, referred to as ‘drift’ in Stegen *et al.* 2013). For pairwise comparisons where $|\beta_{NTI}| < 2$, $RC_{bray} < -0.95$ represented homogenising dispersal, $RC_{bray} > +0.95$ represented dispersal limitation, and $RC_{bray} < |0.95|$ suggested no single ecological process dominated compositional turnover. We compared the relative contribution of each process for pairwise comparisons of krill bacterial communities within a swarm, between swarms within a bacterial geographical cluster (see Fig. 1), and between swarms in separate clusters; and within and between tanks for the aquarium populations.

Data Accessibility

DNA sequence data for this study can be found in the NCBI database under BioProject ID: PRJNA505226.

Acknowledgements

We thank the Science Technical Support Team of the Australian Antarctic Division (AAD) and the master and crew of the *Aurora Australis* for their considerable efforts to make this work happen. Jessica Melvin (U. Tasmania) collected samples during the Totten Glacier Voyage. Michael Sumner (AAD) helped extract the satellite-derived environmental data. Chris Burridge (U. Tasmania) gave advice on mixed-effect modelling. Stephen Nicol provided valuable discussions and comments on an earlier version of this manuscript. The manuscript was improved by insights provided by two anonymous reviewers. Molecular work was funded through the Australian Antarctic Science Program (AAS-4015 and 4313) and a Bioplatforms Australia Industry Access Voucher. LS was supported by an Early Postdoc Mobility fellowship from the Swiss National Science Foundation (P2EZP3_162241). SB is supported under an Australia Research Council Discovery Early Career Researcher Award (DE180100828). This work was a contribution to the Australian Antarctic Science Kerguelen Axis

project (AAS-4344) and was supported by the Australian Government's Business Cooperative Research Centres Programme through the Antarctic Climate and Ecosystems Cooperative Research Centre.

Funding statement

This work is supported by the Australian Antarctic Science Program (AAS-4015 and 4313 to BED), a Bioplatforms Industry Access Voucher (to BED and LJC), an Australia Research Council Discovery Early Career Researcher Award (DE180100828 to SB), and an Early Postdoc Mobility fellowship from the Swiss National Science Foundation (P2EZP3_162241 to LS). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

- Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. *Austral Ecol.* **26**, 32-46.
- Aoki S, Fukai D, Hirawake T, *et al.* (2007) A series of cyclonic eddies in the Antarctic Divergence off Adélie Coast. *J. Geophys. Res.* **112**, C05019.
- Atkinson A, Siegel V, Pakhomov EA, Jessopp MJ, Loeb V (2009) A re-appraisal of the total biomass and annual production of Antarctic krill. *Deep-Sea Res. Pt. 1 Oceanogr. Res. Pap.* **56**, 727-740.
- Avena CV, Parfrey LW, Leff JW, *et al.* (2016) Deconstructing the bat skin microbiome: influences of the host and the environment. *Front. Microbiol.* **7**, 1753.
- Barton K (2019) *MuMIn: Multi-Model Inference. R package version 1.43.6.* <https://CRAN.R-project.org/package=MuumIn>
- Bates D, Mächler M, Bolker B, Walker S (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* **67**, 1-48.
- Bindoff NL, Rosenberg MA, Warner MJ (2000) On the circulation and water masses over the Antarctic continental slope and rise between 80 and 150°E. *Deep-Sea Res. Pt. II. Top. Stud. Oceanogr.* **47**, 2299-2326.
- Bowman JP, Gosink JJ, McCammon SA, *et al.* (1998) *Colwellia demingiae* sp. nov., *Colwellia hornerae* sp. nov., *Colwellia rossensis* sp. nov. and *Colwellia psychrotropica* sp. nov.: psychrophilic Antarctic species with the ability to synthesize docosahexaenoic acid (22:6 ω 3). *Int. J. Syst. Bacteriol.* **48**, 1171-1180.
- Brown MV, van de Kamp J, Ostrowski M, *et al.* (2018) Systematic, continental scale temporal monitoring of marine pelagic microbiota by the Australian Marine Microbial Biodiversity Initiative. *Sci. Data.* **5**, 180130.
- Burke C, Steinberg P, Rusch D, Kjelleberg S, Thomas T (2011) Bacterial community assembly based on functional genes rather than species. *Proc. Natl. Acad. Sci. USA* **108**, 14288-14293.
- Burnham KP, Anderson DR (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretical Approach.*, 2nd edn. Springer-Verlag, New York.
- Burnham KP, Anderson DR, Huyvaert KP (2011) AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behavioral Ecology and Sociobiology* **65**, 23-35.
- Burns AR, Miller E, Agarwal M, *et al.* (2017) Interhost dispersal alters microbiome assembly and can overwhelm host innate immunity in an experimental zebrafish model. *Proc. Natl. Acad. Sci. USA* **114**, 11181-11186.
- Caporaso JG, Bittinger K, Bushman FD, *et al.* (2009) PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* **26**, 266-267.
- Caporaso JG, Kuczynski J, Stombaugh J, *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7**, 335-336.
- Chase JM, Kraft NJB, Smith KG, Vellend M, Inouye BD (2011) Using null models to disentangle variation in community dissimilarity from variation in α -diversity. *Ecosphere* **2**, art24.
- Clarke LJ, Suter L, King R, *et al.* (2020) Data for manuscript "Bacterial epibiont communities of panmictic Antarctic krill are spatially structured". *Dryad, dataset*, DOI: .
- Clarke LJ, Suter L, King R, Bissett A, Deagle BE (2019) Antarctic krill are reservoirs for distinct Southern Ocean microbial communities. *Front. Microbiol.* **9**, 3226.
- Clarke RT, Rothery P, Raybould AF (2002) Confidence limits for regression relationships between distance matrices: Estimating gene flow with distance. *Journal of Agricultural, Biological, and Environmental Statistics* **7**, 361-372.
- Cregeen SJJ (2016) *Microbiota of dominant Atlantic copepods: Pleuromamma sp. as a host to a betaproteobacterial symbiont* Doctoral thesis, University of Southampton.
- Croxall JP, Reid K, Prince PA (1999) Diet, provisioning and productivity responses of marine predators to differences in availability of Antarctic krill. *Mar. Ecol. Prog. Ser.* **177**, 115-131.

- Datta MS, Sliwerska E, Gore J, Polz MF, Cordero OX (2016) Microbial interactions lead to rapid micro-scale successions on model marine particles. *Nat Commun* **7**, 11965.
- Deagle BE, Faux C, Kawaguchi S, Meyer B, Jarman SN (2015) Antarctic krill population genomics: apparent panmixia, but genome complexity and large population size muddy the water. *Mol. Ecol.* **24**, 4943-4959.
- Díez-Vives C, Taboada S, Leiva C, *et al.* (2020) On the way to specificity - Microbiome reflects sponge genetic cluster primarily in highly structured populations. *Mol. Ecol.*, DOI: 10.1111/mec.15635.
- Dominguez-Bello MG, Blaser MJ (2011) The human microbiota as a marker for migrations of individuals and populations. *Annu. Rev. Anthropol.* **40**, 451-474.
- Easson CG, Chaves-Fonnegra A, Thacker RW, Lopez JV (2020) Host population genetics and biogeography structure the microbiome of the sponge *Cliona delitrix*. *Ecol. Evol.* **10**, 2007-2020.
- Edgar RC (2016) UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. *bioRxiv*, 081257.
- Fietz K, Hintze COR, Skovrind M, *et al.* (2018) Mind the gut: genomic insights to population divergence and gut microbial composition of two marine keystone species. *Microbiome* **6**, 82.
- Henne K, Li J, Stoneking M, *et al.* (2014) Global analysis of saliva as a source of bacterial genes for insights into human population structure and migration studies. *BMC Evolutionary Biology* **14**, 190.
- Hubbell SP (2001) *The unified neutral theory of biodiversity and biogeography*. Princeton University Press, Princeton.
- Jarman SN, Deagle BE (2016) Genetics of Antarctic Krill. In: *Biology and Ecology of Antarctic Krill*. (ed. Siegel V), pp. 247-277. Springer, Cham.
- Kawaguchi S, Candy SG, King R, Naganobu M, Nicol S (2006) Modelling growth of Antarctic krill. I. Growth trends with sex, length, season, and region. *Mar. Ecol. Prog. Ser.* **306**, 1-15.
- Kawaguchi S, King R, Meijers R, *et al.* (2010) An experimental aquarium for observing the schooling behaviour of Antarctic krill (*Euphausia superba*). *Deep-Sea Res. Pt. II. Top. Stud. Oceanogr.* **57**, 683-692.
- Kelley JL, Brown AP, Therkildsen NO, Foote AD (2016) The life aquatic: advances in marine vertebrate genomics. *Nat. Rev. Genet.* **17**, 523-534.
- King R, Nicol S, Cramp P, Swadling KM (2003) Krill maintenance and experimentation at the Australian Antarctic Division. *Mar. Freshw. Behav. Physiol.* **36**, 271-283.
- Kolodny O, Weinberg M, Reshef L, *et al.* (2018) Coordinated change at the colony level in fruit bat fur microbiomes through time. *Nat. Ecol. Evol.* **3**, 116-124.
- Lane DJ (1991) 16S/23S rRNA sequencing. In: *Nucleic acid techniques in bacterial systematics* (eds. Stachebrandt E, Goodfellow M), pp. 115-175. John Wiley and Sons, Chichester, NY.
- Lane DJ, Pace B, Olsen GJ, *et al.* (1985) Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proc. Natl. Acad. Sci. USA* **82**, 6955-6959.
- Lankau EW, Hong PY, Mackie RI (2012) Ecological drift and local exposures drive enteric bacterial community differences within species of Galapagos iguanas. *Mol. Ecol.* **21**, 1779-1788.
- Latogurski VI (1979) Recognizing of independent population of Antarctic krill. *Rybnoe Khoziaistvo* **1979**, 12-14 (in Russian).
- Linnenbrink M, Wang J, Hardouin EA, *et al.* (2013) The role of biogeography in shaping diversity of the intestinal microbiota in house mice. *Mol. Ecol.* **22**, 1904-1916.
- Lozupone C, Knight R (2005) UniFrac: a new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* **71**, 8228-8235.
- Mackintosh NA (1973) Distribution of post-larval krill in the Antarctic. *Discov. Rep.* **36**, 95-156.
- McDonald D, Price MN, Goodrich J, *et al.* (2012) An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J.* **6**, 610-618.

536 Miller ET, Svanback R, Bohannan BJM (2018) Microbiomes as metacommunities: understanding host-
537 associated microbes through metacommunity ecology. *Trends Ecol. Evol.* **33**, 926-935.

538 Moodley Y, Linz B, Yamaoka Y, *et al.* (2009) The peopling of the Pacific from a bacterial perspective.
539 *Science* **323**, 527-530.

540 Nakagawa S, Schielzeth H (2013) A general and simple method for obtaining R² from generalized linear
541 mixed-effects models. *Methods in Ecology and Evolution* **4**, 133-142.

542 Nicol S, Foster J (2016) The fishery for Antarctic Krill: Its current status and management regime. In:
543 *Biology and Ecology of Antarctic Krill*. (ed. Siegel V), pp. 387-421. Springer, Cham.

544 Nowacek DP, Friedlaender AS, Halpin PN, *et al.* (2011) Super-aggregations of krill and humpback
545 whales in Wilhelmina Bay, Antarctic Peninsula. *PLoS One* **6**, e19173.

546 Orsi AH, Whitworth T, Nowlin WD (1995) On the meridional extent and fronts of the Antarctic
547 Circumpolar Current. *Deep-Sea Res. Pt. 1 Oceanogr. Res. Pap.* **42**, 641-673.

548 Peijnenburg KT, Goetze E (2013) High evolutionary potential of marine zooplankton. *Ecol. Evol.* **3**,
549 2765-2781.

550 Price MN, Dehal PS, Arkin AP (2010) FastTree 2 – Approximately Maximum-Likelihood Trees for Large
551 Alignments. *PLOS ONE* **5**, e9490.

552 Reiss CS (2016) Age, growth, mortality, and recruitment of Antarctic Krill, *Euphausia superba*. In:
553 *Biology and Ecology of Antarctic Krill*. (ed. V. S), pp. 101-144. Springer, Cham.

554 Row JR, Knick ST, Oyler-McCance SJ, Loughheed SC, Fedy BC (2017) Developing approaches for linear
555 mixed modeling in landscape genetics through landscape-directed dispersal simulations. *Ecol.*
556 *Evol.* **7**, 3751-3761.

557 Segata N, Izard J, Waldron L, *et al.* (2011) Metagenomic biomarker discovery and explanation. *Genome*
558 *Biol.* **12**, R60.

559 Siegel V, Watkins JL (2016) Distribution, biomass and demography of Antarctic Krill, *Euphausia*
560 *superba*. In: *Biology and Ecology of Antarctic Krill* (ed. Siegel V), pp. 21-100. Springer, Cham.

561 Stegen JC, Lin X, Fredrickson JK, *et al.* (2013) Quantifying community assembly processes and
562 identifying features that impose them. *ISME J.* **7**, 2069-2079.

563 Stegen JC, Lin X, Fredrickson JK, Konopka AE (2015) Estimating and mapping ecological processes
564 influencing microbial community assembly. *Front. Microbiol.* **6**, 370.

565 Tarling GA, Cuzin-Roudy J (2008) External parasite infestation depends on moult-frequency and age in
566 Antarctic krill (*Euphausia superba*). *Polar Biol.* **31**, 121-130.

567 Tarling GA, Klevjer T, Fielding S, *et al.* (2009) Variability and predictability of Antarctic krill swarm
568 structure. *Deep-Sea Res. Pt. 1 Oceanogr. Res. Pap.* **56**, 1994-2012.

569 Thompson KF, Patel S, Baker CS, Constantine R, Millar CD (2016) Bucking the trend: genetic analysis
570 reveals high diversity, large population size and low differentiation in a deep ocean cetacean.
571 *Heredity* **116**, 277-285.

572 Van Strien MJ, Keller D, Holderegger R (2012) A new analytical approach to landscape genetic
573 modelling: least-cost transect analysis and linear mixed models. *Mol Ecol* **21**, 4010-4023.

574 Virtue P, Kawaguchi S, McIvor J, *et al.* (2010) Krill growth and condition in Western Indian Ocean sector
575 of the Southern Ocean 30-80°E in austral summer 2006. *Deep-Sea Res. Pt. II. Top. Stud.*
576 *Oceanogr.* **57**, 948-955.

577 Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA
578 sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**, 5261-5267.

579 Ward RD, Woodward M, Skibinski DOF (1994) A comparison of genetic diversity levels in marine,
580 freshwater, and anadromous fishes. *J. Fish Biol.* **44**, 213-232.

581 Wirth T, Wang X, Linz B, *et al.* (2004) Distinguishing human ethnic groups by means of sequences from
582 *Helicobacter pylori*: lessons from Ladakh. *Proc. Natl. Acad. Sci. USA* **101**, 4746-4751.

583 Yang C, Cui Y, Didelot X, Yang R, Falush D (2019) Why panmictic bacteria are rare. *bioRxiv*, 385336.

584 Yilmaz P, Parfrey LW, Yarza P, *et al.* (2014) The SILVA and "All-species Living Tree Project (LTP)"
585 taxonomic frameworks. *Nucleic Acids Res.* **42**, D643-D648.

586 Zuur AF, Ieno EN, Elphick CS (2010) A protocol for data exploration to avoid common statistical
587 problems. *Methods in Ecology and Evolution* **1**, 3-14.
588 Zwirgmaier K, Reid WD, Heywood J, *et al.* (2015) Linking regional variation of epibiotic bacterial
589 diversity and trophic ecology in a new species of Kiwaidae (Decapoda, Anomura) from East
590 Scotia Ridge (Antarctica) hydrothermal vents. *MicrobiologyOpen* **4**, 136-150.

591

592 **Author contributions**

593 LC, LS, RK, and BD conceived the experiments and collected the samples. LC, AB and SB performed the
594 data analysis. LC wrote the manuscript. All authors contributed to discussing results and the final
595 version of the manuscript.

596

597 **Competing Interests**

598 The authors declare no competing interests.

599

Figures

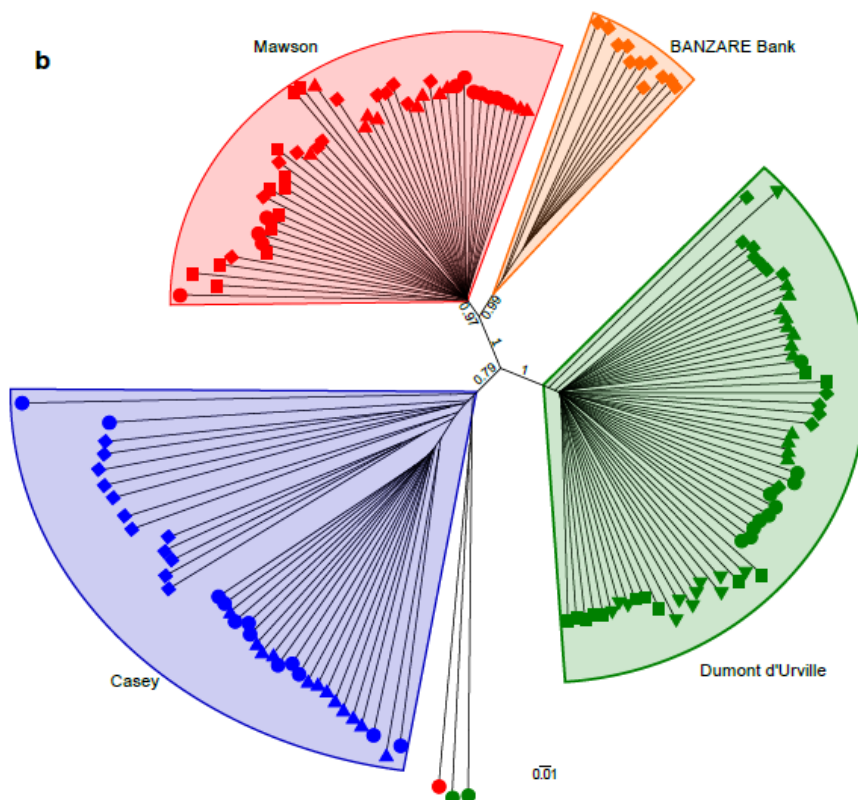
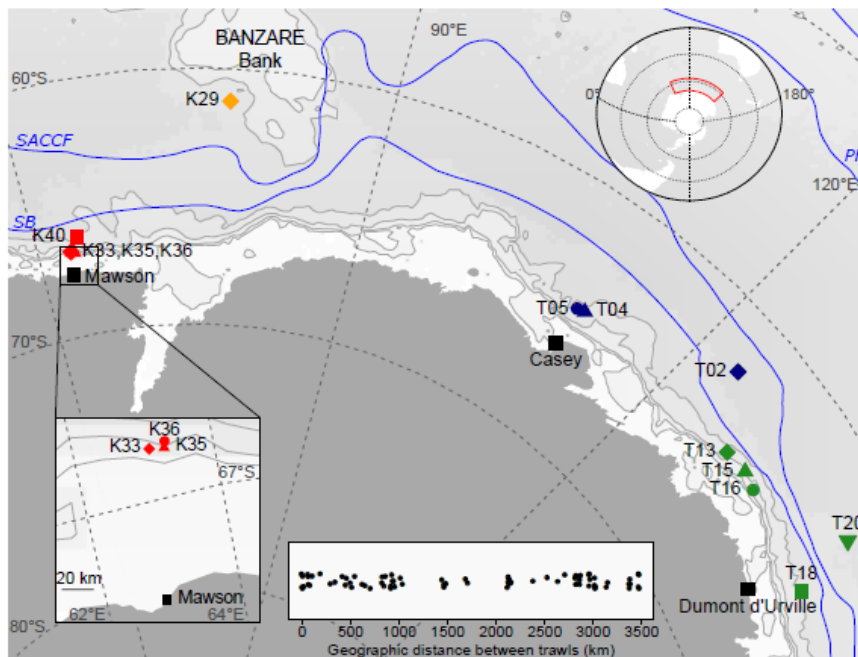


Fig. 1. Spatial structuring of krill moult bacterial communities. **a** Map of Southern Ocean trawl locations off the East Antarctic coast used to study krill-associated bacteria. Trawls are coloured according to clustering in **(b)**. The inset shows three trawls near Mawson station separated by 4-11 km. Mean locations of the principal fronts (following Orsi *et al.* 1995) are shown as blue lines. PF – Polar Front, SACCF – southern Antarctic Circumpolar Current front, SB – Southern Boundary of ACC. **b** UPGMA clustering of krill moult bacterial communities based on unweighted UniFrac distance. Trawls within each region are differentiated by symbols. Jack-knife support is shown for central branches.

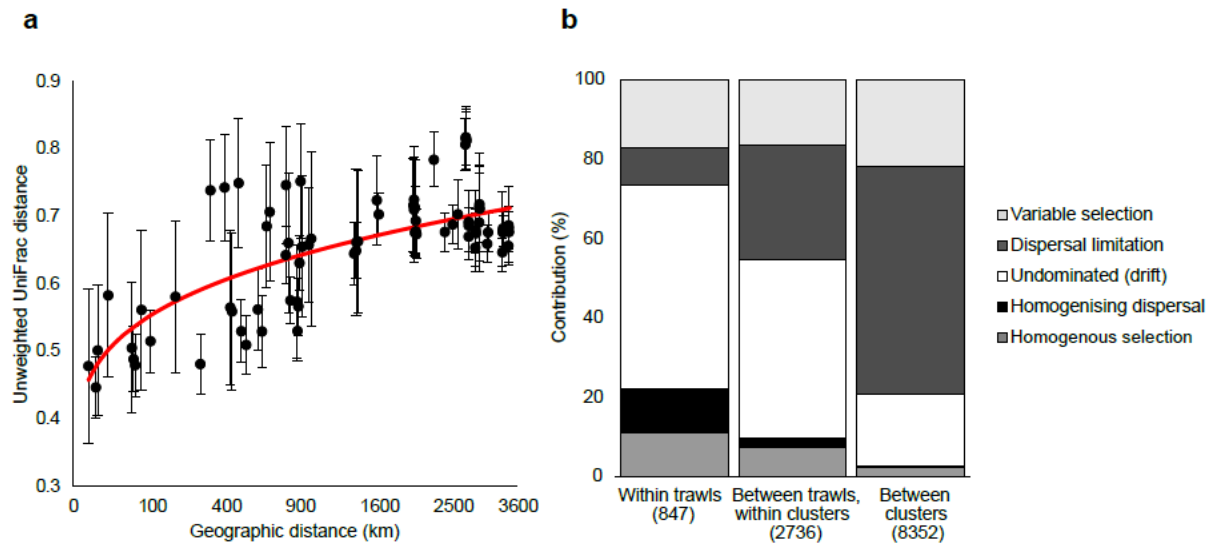


Fig. 2. Krill microbiome differentiation driven by dispersal limitation over large geographic distances. **a** Geographic vs. ecological (unweighted UniFrac) distance for krill moult bacterial communities from separate trawls. The x-axis has been square-root transformed for clarity. Equation of the fitted line: $y = 1/(2.354 - 0.116 * \log(x))$, adjusted $R^2 = 0.432$, $p < 2 \times 10^{-16}$. Values are means \pm SD unweighted UniFrac distance for all samples from each pair of trawls. **b** Contribution of ecological processes to krill moult bacterial community assembly at different geographical scales. The number of pairwise comparisons for each category is shown in brackets.

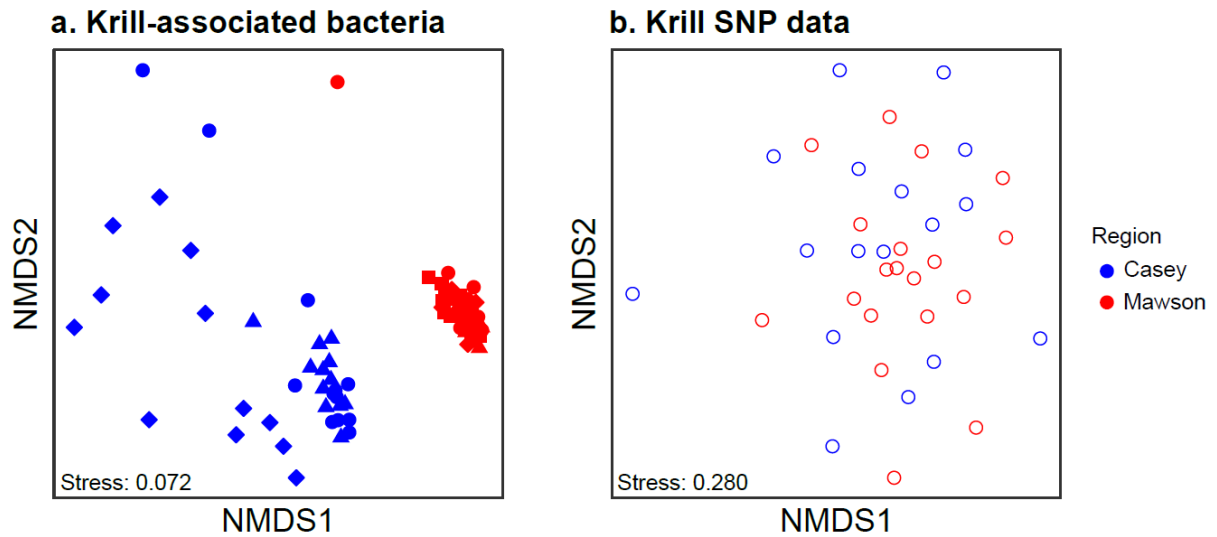


Fig. 3. Spatial structuring of krill microbiomes contrasted with lack of structure observed in SNP markers. Non-metric multidimensional scaling (nMDS) plots for krill from the vicinity of Mawson and Casey stations based on either (a) moult bacterial communities using unweighted UniFrac distance, or (b) Single Nucleotide Polymorphism (SNP) markers using Bray-Curtis dissimilarity of allelic sequence counts (data from Deagle *et al.* 2015). Krill-associated bacterial community data is from 3 and 5 trawls from the Casey and Mawson regions, respectively, coloured by bacterial geographic cluster as per Fig. 1.

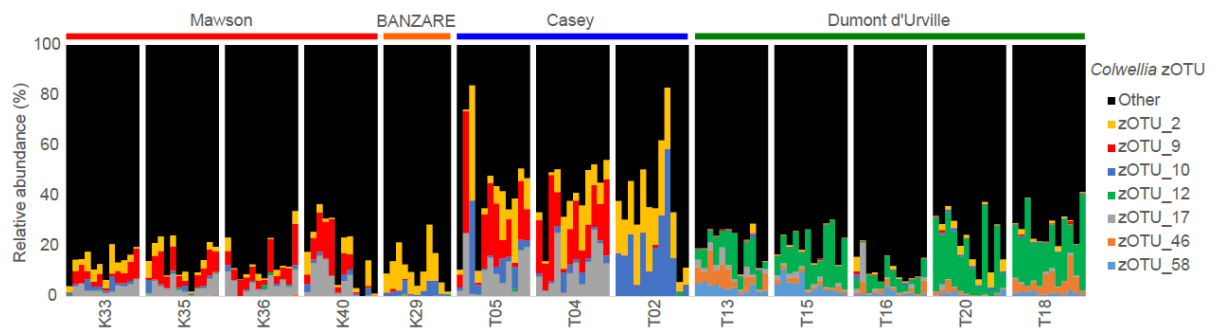


Fig. 4. Relative abundance of dominant *Colwellia* zOTUs in krill moults shown relative to all other zOTUs, highlighting variability between swarms within a single genus. Samples are grouped by trawl and arranged west to east. Each of the seven zOTUs are members of the core microbiome (present in 85% or more moult samples).

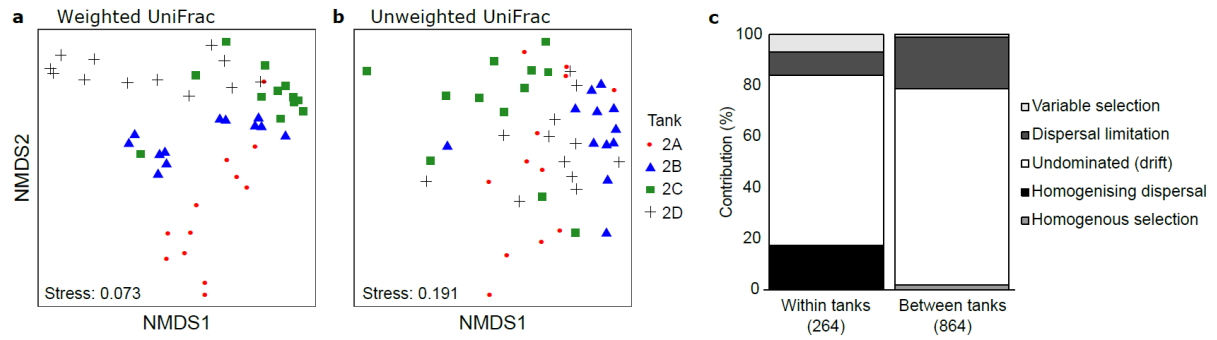


Fig. 5. Differences between krill bacterial communities originating from a single swarm split across four aquarium tanks for seven months. Non-metric multidimensional scaling (nMDS) plot based on weighted (a) or unweighted UniFrac distance (b). c Contribution of ecological processes to krill epibiont bacterial community assembly within versus between aquarium tanks. The number of pairwise comparisons for each category is shown in brackets.