The coupling of carbon and iron cycles in the Southern Ocean through microbial metabolism

Pavla Debeljak

Sorbonne Université

École Doctorale des Sciences de l'Environnement d'Ile-de-France

Laboratoire d'Océanographie Microbienne (UMR 7621)

The coupling of carbon and iron cycles in the Southern Ocean through microbial metabolism

Par Pavla Debeljak

Thèse de doctorat de Sciences de l'environnement
Dirigée par Ingrid Obernosterer et Gerhard J. Herndl
Présentée et soutenue publiquement le 18 octobre 2019

Devant un jury composé de :

Maria Vila Costa, CSIC Barcelona, Spain, Rapporteur

David Kirchman, University of Delaware, USA, Rapporteur

Loïs Magnien, IUEM, Brest, France, Examinateur

Geraldine Sarthou, LEMAR, Brest, France, Examinateur

Gwanael Piganeau, BIOM, Banyuls-sur-Mer, France, Examinateur

Ingrid Obernosterer, LOMIC, Banyuls-sur-Mer, France, Directrice de thèse

Gerhard J. Herndl, University of Vienna, Vienna, Austria, Co-Directeur de thèse

Ce qui est étrange avec le voyage, c'est qu'on ne comprend qu'après
– et encore pas toujours – ce qu'on est allé chercher.

~ Emmanuel Lepage

Za mamu it tatu, učenje se ipak izplati

Table of content

Résumé

L'océan Austral est caractérisé par l'acronyme anglais (HNLC) signifiant qu'il contient des concentrations en nutriments élevées, mais que la biomasse phytoplanctonique y est faible. Cette situation paradoxale est due à la limitation du phytoplancton par le fer (Fe). L'augmentation de la biomasse phytoplanctonique en réponse à l'apport de Fe et, par conséquent, une plus grande absorption de dioxyde de carbone (CO_2), a été démontrée dans plusieurs expériences de fertilisation artificielle et dans des régions naturellement fertilisées de l'océan Austral. Cependant, l'impact de Fe sur les procaryotes marins, acteurs clés du cycle du carbone marin, reste mal compris. En plus du Fe, les concentrations de carbone organique dissous (COD) sont faibles dans les eaux de surface de l'océan Austral, ce qui entraîne une double limitation pour l'activité et de la croissance des procaryotes hétérotrophes. Pour mieux comprendre le rôle des différents taxons procaryotes dans les cycles de Fe et de C, plusieurs aspects ont été pris en compte dans la thèse actuelle. À l'aide d'une approche ciblant des gènes spécifiques de données métatranscriptomiques, nous avons étudié sept voies métaboliques liées au Fe (incorporation, stockage et métabolisme du carbone cellulaire) dans des communautés microbiennes présent dans des eaux type HNLC et celles naturellement fertilisées en Fe au large des iles Kerguelen (Océan Austral) (Chapitre 1). Pour répondre à la question de savoir si les particules marines représentent une source de métaux traces, en particulier du Fe, pour différents membres de la communauté des procaryotes attachés à ou dans la sphère des particules, nous avons effectué une expérience, combinant des approches classiques et métatranscriptomiques (Chapitre 2). Dans le chapitre 3, nous utilisons la protéomique pour une étude comparative des protéines de transporteurs en combinaison avec la méta-génomique et la meta-transcriptomique, pour décrire les modèles d'utilisation du carbone et organique et du Fe par la communauté microbienne pendant deux saisons à des sites contrastés dans le Océan Austral. Les résultats présentés dans cette thèse fournissent de nouvelles perspectives sur le lien entre les divers taxons microbiens et leur rôle dans le cycle du Fe- et du C dans l'océan Austral.

Summary

The Southern Ocean is known as a High Nutrient Low Chlorophyll (HNLC) region where major nutrients are present at high concentrations, but phytoplankton biomass remains low. This paradoxical situation is due to iron (Fe) limitation of phytoplankton growth. The increase in phytoplankton biomass in response to Fe input and the consequently higher uptake of carbon dioxide (CO_2) were demonstrated in several artificial fertilization experiments and in naturally fertilized regions of the Southern Ocean. However, the impact of Fe on marine prokaryotes, key players in the marine carbon cycle, remains poorly understood. In addition to Fe, concentrations of dissolved organic carbon (DOC) are low in surface waters of the Southern Ocean, leading to a double constraint of microbial heterotrophic activity and growth. To better understand the role of different prokaryotic taxa in the Fe- and C-cycles, several aspects were considered in the present thesis. Using a gene-specific approach from metatranscriptomics data, we investigated seven Fe-related metabolic pathways (uptake, storage and cellular carbon metabolism) in microbial communities from HNLC and naturally Fe-fertilized waters in the Kerguelen region (Southern Ocean)(Chapter 1). To address the question of whether marine particles represent a source of trace metals, in particular Fe, to different members of the prokaryotic community attached to or in the particle sphere, we performed an on-board experiment, combining classical and metatranscriptomics analyses (Chapter 2). In Chapter 3, we use comparative proteomics of transporter proteins in combination with meta-genomics and –transcriptomics, to describe uptake patterns of Fe and organic carbon substrates by the microbial community during two seasons at contrasting sites in the Southern Ocean. The results presented in this thesis provide novel insights on the link between diverse microbial taxa and their role in the Fe- and C-cycling in the Southern Ocean.

Introduction

The Southern Ocean and iron limitation in the context of the biological pump

The Southern Ocean covers 20 % of the world's ocean and interconnects the three oceanic basins by circumpolar currents as discovered through the voyage of Captain James Cook in 1770. This interconnection defines the functioning of the Southern Ocean and is dependent on water masses which can be distinguished by physical as well as chemical properties (Orsi *et al.*, 1995; Sokolov and Rintoul, 2002, 2009; Wilkins, *et al.*, 2013). The cold, sinking Antarctic surface waters mix with deep waters from the Atlantic and Indian Ocean resulting in nutrient-rich upper and lower circumpolar deep waters (Fig.1). Upwelling of these circumpolar deep waters is a source of nutrients to the surface waters of lower latitudes fueling 75 % of global primary production north of 30°S (Sarmiento *et al.*, 2004; Wilkins, *et al.*, 2013). The deepest water mass, the Antarctic Bottom Water, formed in the Ross and Weddell Sea, is pushed northwards through sinking and drives together with the North Atlantic Deep Water the global thermohaline circulation. The warmer waters of the Sub Antarctic Zone (SAZ) are separated from the cold Antarctic Zone (AAZ) waters by the Polar Front. The AAZ is characterized by a remaining temperature minimum layer of 2°C and can further be divided in the Permanently Open Ocean Zone (POOZ) and the Seasonal Ice Zone (SIZ). The Polar Front Zone (PFZ) is marked by drastic water temperature changes over a short distance and is seen as the most pronounced biogeographic boundary of Antarctic waters where significant downwelling occurs (Chiba *et al.*, 2001; Hunt *et al.*, 2001; Weber and Deutsch, 2010; Wilkins, *et al.*, 2013). Recent adjustments with innovative research tools such as temperature and salinity profiles from Argo floats have challenged the idea of the PFZ as a continuous boundary (Pauthenet *et al.*, 2018). Enclosed by the PFZ to the north and the Antarctic Zone to the south, westerlies generate the eastward flowing Antarctic Circumpolar Current (ACC) that circulates surface waters around the Antarctic continent and is crucial to the seasonal sea surface ice formation and loss (Rintoul *et al.*, 2001).

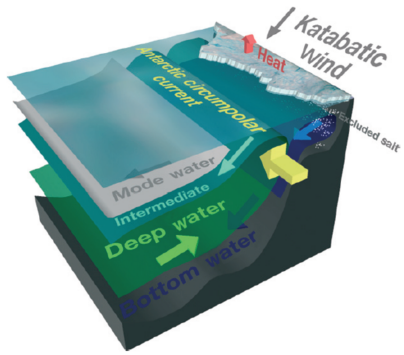


Figure 1. *Schematic view of Antarctic water masses from Rintoul et al. (2001)*

Because of the importance its for the global ocean circulation the ACC was called "great mix-master of the world ocean" by Wally Broeker, an oceanographer at the Columbia University, an expression

repeatedly used ever since (Broeker, 1991). These southernmost waters are responsible for the exchange of heat, nutrients, salt and gases between the basins of the Pacific, Atlantic and Indian Ocean and are increasingly important due to their uptake of anthropogenic CO_2. Over the last century, uptake of 40 % of global CO_2 has occurred south of 40°S describing this region as a significant CO_2 sink of the ocean (Sabine *et al.*, 2004; Fletcher *et al.*, 2006; Khatiwala *et al.*, 2009; Frölicher *et al.*, 2014).

The remoteness to any coast and the unbounded water masses make the Southern Ocean the largest High Nutrient Low Chlorophyll (HNLC) area where macronutrients are recurrently present at high concentrations yet phytoplankton biomass remains low. Together with the eastern subarctic Pacific and eastern equatorial Pacific these three HNLC regions represent 30 % of the global ocean (Minas and Minas, 1992; Pitchford, 1999; Wilkins, *et al.*, 2013).

Phytoplankton require a wide range of other elements for growth, such as nitrogen and phosphorus (Falkowski *et al.*, 1998; Sarmiento and Gruber, 2006; Falkowski and Raven, 2013). These macronutrients are generally present in the higher μmol L^{-1} range in surface ocean waters, while micronutrients, for example trace metals such as iron, zinc or cobalt are generally available in the lower nmol L^{-1} range (Dugdale and Wilkerson, 1991). Within certain limits, inorganic carbon and nutrients are incorporated into new cellular material in a fixed ratio (Redfield, 1958), even though the plasticity of the C:P and C:N ratios has recently been described (Sarmiento and Gruber, 2006; Galbraith and Martiny, 2015). Following Liebig's law of the minimum, biomass production is limited by the nutrient species in shortest supply. Nitrogen and phosphorus availability limit productivity throughout the surface low-latitudes as observed in the Atlantic Ocean and Mediterranean Sea (Krom *et al.*, 1991, 2004; Mills *et al.*, 2008), however, in HNLC regions these nutrients are never depleted (Ruud, 1930; Pitchford, 1999; Tagliabue *et al.*, 2017). This anomaly of high phosphorus and nitrogen concentrations apparently not utilized by primary producers in the Southern Ocean was coined "Antarctic paradox". In the 1930s oceanographers proposed that this might be due to the limitation of the trace metal iron as well as the lack of light for photosynthesis in high latitudes (Ruud, 1930; Gran, 1931; Hart, 1934; Martin and Fitzwater, 1988; Tagliabue *et al.*, 2017).

Early laboratory studies showed that iron enrichment stimulated the growth of phytoplankton indicating that iron could indeed be limiting (Cooper and Gelston, 1935). Additionally, iron was confirmed as a critical phytoplankton micronutrient necessary as co-factor for protein assembly, the electron transport chain and nitrogen assimilation (Holm *et al.*, 1996; Takeda, 1998).

Once trace metal clean sampling methods were developed that could measure in pico- and nanomolar dissolved iron concentrations, iron limitation of phytoplankton growth was recognized as a specific feature of the Southern Ocean (Martin, Fitzwater, *et al.*, 1990; Martin, Gordon, *et al.*, 1990; Bruland *et al.*, 1991; de Baar *et al.*, 2005). Hence, in the year 1990 John Martin postulated the "Iron Hypothesis", a theory on higher input of iron by dust during glacial periods in the Southern Ocean which enhanced the drawdown of atmospheric CO_2 by efficient utilization of macronutrients (Martin, 1990). Since then, models (Sarmiento and Orr, 1991) as well as multiple artificial (Seeds, Eisenex, IronEx,

Soiree, SoFex; de Baar *et al.*, 2005) and a few natural (KEOPS, CROZEX, Blain *et al.*, 2007; Pollard *et al.*, 2009) iron fertilization experiments have studied the impact of iron input in Southern Ocean waters. Martin's hypothesis has been widely accepted and fueled research in iron biogeochemistry (Boyd *et al.*, 2007 review and references therein; Tagliabue *et al.*, 2017). Furthermore, the launch of the international research project GEOTRACES to study trace elements and isotopes has increased the quality and consistency of trace metal data (SCOR Working Group, 2007).

The assimilation of carbon, nitrogen and phosphorus into biomass through the conversion to particulate matter form the basis of the food web in the ocean. This primary production is performed through photosynthetic activity of autotrophic microorganisms, such as phytoplankton and cyanobacteria. These enzymatic activities used by microorganisms to transform energy and matter (Falkowski *et al.*, 2008) contain or depend partly on metals. These metals might control the growth of organisms as well as the cycling of major nutrients (Morel and Price, 2003). In the case of organic material, remineralization (or decomposition) processes, such as prokaryotic respiration, convert the organic carbon to carbon dioxide while the non-remineralized compounds are sequestrated in the ocean's interior or transferred to higher trophic levels (Azam, 1983; Cho and Azam, 1988; Azam and Malfatti, 2007). The biological pump is an efficient mechanism fixing CO_2 by primary producers and transporting organic carbon into the ocean's interior, resulting in sequestration of carbon. Understanding factors involved in the production and remineralization of carbon is of main interest to predict the factors influencing the biological pump efficiency (Herndl and Reinthaler, 2013). The efficiency of the biological pump is measured by the ability of phytoplankton to maintain low nutrient concentrations in surface waters. Without the biological pump operating, the distribution of most chemicals in the ocean would be as uniform as that of salinity. A common method is to estimate primary production fueled by nitrate and ammonium as these nutrients have different sources that are related to the remineralization of sinking material (Sarmiento and Gruber, 2006).

In the Southern Ocean, the low light supply coupled with low iron concentration restricts photosynthesis. Consequently, the supply of nutrients as well as the export of organic matter is low, defining the Southern Ocean biological pump as currently inefficient (Sarmiento and Orr, 1991; Volk and Hoffert 2013). Therefore, artificial iron fertilization studies aimed at the potential decrease of atmospheric CO_2 by increasing the biological pump efficiency. However, results showed variable consequences on carbon export and the objective to decrease atmospheric CO_2 on a large scale remains speculative even though the discussion is re-appearing in recent literature (Chisholm *et al.*, 2001; Zeebe and Archer, 2005; Emerson, 2019).

Iron in the ocean

Even though iron is the 4th most abundant element in the Earth's crust its concentration in seawater remains low. The horizontal and vertical distribution of iron as well as other trace metals generally decreases from coastal to offshore environments and increases from the euphotic to deeper waters (Johnson *et al.*, 1997; Blain and Tagliabue, 2016; Tagliabue *et al.*, 2017). Iron in the ocean can be divided into primarily particulate iron and the dissolved fraction (Fig. 2). Particulate iron, defined as > 0.2 µm, is a dynamic pool defined by its sources which can be lithogenic (silicates, aluminosilicates and oxyhydroxides), detrital and biogenic and shows high spatial variability (Johnson *et al.*, 1997; Blain and Tagliabue, 2016). The dissolved iron fraction (operationally defined as < 0.2 µm) consists largely of colloidal hydrolysis species such as $Fe(OH)_3$ which are rapidly scavenged by adsorption onto particulate material, keeping dissolved iron concentrations low. Dissolved inorganic iron is further divided into its reduced ferrous (Fe^{2+}) or oxidized ferric (Fe^{3+}) state (Fig. 2, Katoh *et al.*, 2001; Cartron *et al.*, 2006). In oxygenated seawater, the thermodynamically favored oxidation state is Fe^{3+} (Kuma *et al.*, 1998). The dissolved iron concentration in waters below 500 m remains relatively stable (Liu and Millero, 2002; Blain and Tagliabue, 2016). This is in part explained by the existence of high affinity organic ligands which specifically bind Fe^{3+} (Johnson *et al.*, 1997; Tortell *et al.*, 1999). The occurrence of organic ligands increases iron solubility and binding > 99.9 % of dissolved iron (Kuma *et al.*, 1998). Ultraviolet (UV) irradiation can decrease the relative proportion of dissolved Fe through photochemical reduction (Fig 2., Kuma *et al.*, 1996; Tortell *et al.*, 1999; Liu and Millero, 2002).

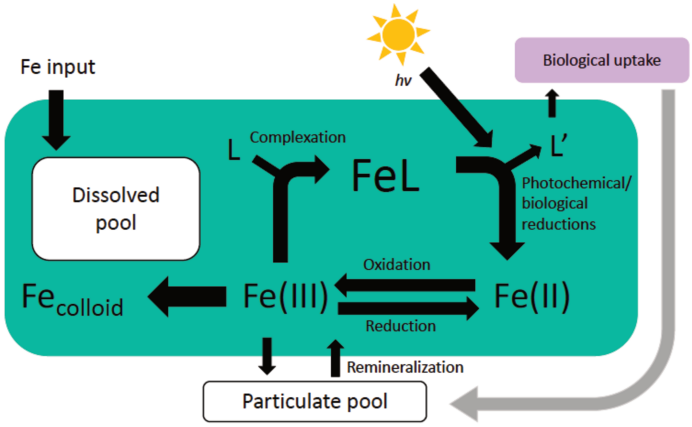


Figure 2. *Forms of iron in seawater. Modified from Yoshiko Kondo.*

This organic complexation results in low inorganic Fe concentrations of 0.2 – 0.3 nM in surface waters (Rue and Bruland, 1995; Tortell *et al.*, 1999). Organic ligands are present in the seawater through prokaryotic release, viral lysis, or grazing. Prokaryotes can produce siderophores, molecules that chelate dissolved Fe, and release them into the environment to acquire Fe (Winkelmann, 1991; Martinez *et al.*, 2003; Hider and Kong, 2010). Laboratory cultures have shown that marine heterotrophic and phototrophic bacteria produce siderophores when grown under low Fe levels typical for oceanic environments (Trick, 1989; Wilhelm and Trick, 1994). Siderophore production has been observed *in*

situ under photochemical reactions in iron replete conditions and chemically diverse ligands have been identified (Barbeau *et al.*, 2001, 2003; Butler, 2005; Vraspir and Butler, 2009).

A recent illustration of the current ocean iron cycle depicts the iron-limited Southern Ocean in strong contrast to the macronutrient (N, P)-limited low-latitude regions (Fig. 3, Tagliabue *et al.*, 2017). Dust input is the main source of iron at low latitudes while in the Southern Ocean input from deep waters and resuspension from coastal areas are the major sources (Duce and Tindale, 1991; Coale *et al.*, 1996; Johnson *et al.*, 1997; Blain *et al.*, 2007). High variability of iron uptake and biological cycling, along with the production of excess iron-binding ligands is observed in the Southern Ocean, whereas the majority of nitrogen fixation occurs at low latitudes. More recently evidence of nitrogen fixing genes in metagenomes from the Southern Ocean, however, are challenging this classical view (Delmont *et al.*, 2018).

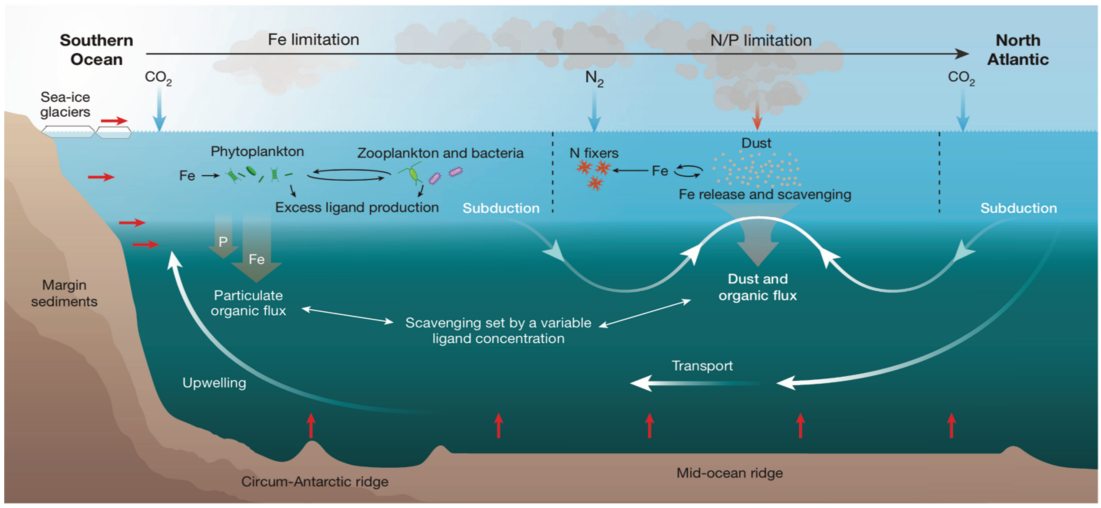


Figure 3. *Illustration of the ocean iron cycle from Tagliabue et al. 2017.*

A feature that could be of more importance than previously assumed is the subduction of iron-binding ligand produced at high latitudes that can impact deeper water masses at low latitudes (Tagliabue *et al.*, 2017). Hydrothermal dissolved iron has also been proposed as an important source of iron in the southern hemisphere by a modelling approach (Tagliabue *et al.*, 2010). Recently, data from profiling floats confirm that hydrothermal vents trigger phytoplankton blooms in the Southern Ocean by mixing of upwelled hydrothermal dissolved iron (Ardyna *et al.*, 2019).

Heterotrophic prokaryotes in the microbial loop

Functional classifications of organisms are usually based on their trophic level, their position within an ecosystem with regard to the flow of energy and material. The primary distinction is between autotrophs, which can create organic matter directly from inorganic nutrients using solar or chemical energy, and heterotrophs, gaining both energy and nutrition only from pre-existing organic matter (Kirchman, 2000; Sarmiento and Gruber, 2006) .

Prokaryotes are the most abundant living component of the biosphere and consequently also in the ocean, amounting to ~10^{29} cells in the oceanic water column (Whitman *et al.*, 1998). Planktonic

prokaryotes are both phylogenetically and metabolically diverse (Venter *et al.*, 2004; DeLong *et al.*, 2006) and play dominant roles in mediating a variety of different biogeochemical cycles (Azam and Malfatti, 2007).

Heterotrophic prokaryotes are crucial components of marine food webs and have key roles in controlling carbon fluxes in the oceans (Hedges, 1992; Hansell *et al.*, 2009; Kirchman, 2010). These prokaryotes are part of the microbial loop including the uptake and incorporation of DOM while respiring CO_2 and the release of DOM during growth and decay. The microbial loop is driven by DOM, which would otherwise have been lost from the food web. Up to half of primary production is taken up by heterotrophic prokaryotes, utilized and respired to the food web through prokaryotic grazing or viral lysis. The microbial loop is embedded in the classic grazing food chain of phytoplankton, zooplankton and top-predators and is responsible for carbon, nutrient and energy fluxes towards higher trophic levels (Azam, 1983; Pomeroy *et al.*, 2007; Gasol and Kirchman, 2018).

The effect on iron-limitation in the Southern Ocean has been extensively studied for phytoplankton and especially diatoms (Marchetti and Cassar, 2009; Smetacek *et al.*, 2012; Strzepek *et al.*, 2012; Assmy *et al.*, 2013; Quéguiner, 2013; Lasbleiz *et al.*, 2016). Research on heterotrophic prokaryotes in iron- and nutrient-limited conditions has indicated that they are physiologically able to adapt to rapid environmental changes (Kirchman, 1990, 1994; Holm *et al.*, 1996; Tortell *et al.*, 1999; Andrews *et al.*, 2003; Almaas *et al.*, 2005). In the case of iron this was first shown by Tortell (1996) when multiple heterotrophic strains had different Fe:C ratios depending on the Fe-concentration in the experiment. The ability to adjust metabolic functioning was reported for the TCA cycle, glycolysis and the respiratory chain for iron-limited conditions (Oexle *et al.*, 1999; Shakoury-Elizeh *et al.*, 2010; Fourquez *et al.*, 2014; Koedooder *et al.*, 2018).

Carbon in the ocean and co-limitation

The ocean stock of dissolved organic matter (DOM) equals roughly the amount of CO_2 in the atmosphere and represents the largest, most heterogenous and dynamic reservoir of reduced carbon on Earth (Mopper *et al.*, 1991; Jiao *et al.*, 2010; Bauer *et al.*, 2013; Moran *et al.*, 2016). DOM is highly complex, consisting of thousands of different organic molecules (Kim *et al.*, 2003; Hansman *et al.*, 2015). The chemical composition and spatial and temporal dynamics of the marine carbon pool as well as the relationship between marine microorganisms and the molecules of the DOM pool are still poorly understood (Moran *et al.*, 2016). DOM is released as a by-product of metabolically active microbes, but may also be released for nutrient acquisition and communication (in the forms of metal-binding ligands and quorum sensing chemicals, respectively) (Gram *et al.*, 2002; Hansell *et al.*, 2009; Vraspir and Butler, 2009), predation or viral lysis (Nagata, 2000).

There are major gaps, however, in our understanding of the biogeochemical significance of the released compounds (Flynn *et al.*, 2008). Although the largest fraction of DOM remains uncharacterized, phytoplankton exudates consist to a considerable extent of labile compounds such as

carbohydrates (mono-, and polysaccharides), proteins and amino acids (Obernosterer and Herndl, 1995; Myklestad, 2000; Granum *et al.*, 2002). Hence, extracellular release of DOM by phytoplankton supports a major fraction of the labile carbon flux in the surface ocean thereby fueling prokaryotic production (Fogg, 1983; Baines and Pace, 1991; Morán *et al.*, 2013). Labile DOM is defined as being consumed by microbes within hours to days of production but typically only accounts for a fraction of the prokaryotic C- and N-demand (Fuhrman, 1987). However, the majority of the oceanic DOM pool appears to be recalcitrant to semi-labile with lifetimes from months to millennia (Hansell, 2013). The composition of DOM released by heterotrophic prokaryotes is even less known than for phytoplankton, however, they do produce and secrete organic compounds for similar purposes (Heissenberger and Herndl, 1994; Stoderegger and Herndl, 1998; Hansell *et al.*, 2009; Kujawinski *et al.*, 2009).

Novel techniques based on electrospray ionization coupled to mass spectrometry allow for a more detailed characterization of the low molecular weight (LMW, < 1000 Da) DOM pool, which represents the major component of DOM and were hardly accessible for detailed molecular characterization before (Kujawinski, 2011). These recent methodological advances gave new insights into the vast complexity of DOM secreted by phytoplankton and heterotrophic prokaryotes (Becker *et al.*, 2014; Romano *et al.*, 2014; Wienhausen *et al.*, 2017).

An interesting feature that characterizes surface waters of the Southern Ocean is the low concentration of dissolved organic carbon (DOC, roughly 50 µM; Hansell *et al.*, 2009). Organic carbon availability also affects the ocean's biogeochemistry of other nutrients such as Fe. Glycolysis, the citric acid cycle as well as processes related to the respiratory chain rely on Fe-containing enzymes (Kirchman, 2000). Fe and organic carbon represent therefore both limiting elements for heterotrophic prokaryotes in Southern Ocean surface waters (Church *et al.*, 2000; Obernosterer *et al.*, 2015; Baltar *et al.*, 2018). Nutrient co-limitation has been proposed in this context as the condition when two or more nutrients, both necessary for growth, are limited, however, the usage of the term varies (Arrigo, 2005; Saito *et al.*, 2008; Moore *et al.*, 2013). Several nutrients can be depleted to equally limiting amounts which is why both need to be added to observe a growth response. Substitution of one limiting nutrient with the same macromolecule through biological processes was also observed (Sunda and Huntsman, 1995). Another scenario is the dependence to take up low concentrations of one nutrient on the availability of another nutrient (Saito *et al.*, 2008; Moore *et al.*, 2013). The response to the addition of one nutrient might also vary between members of the microbial community (Arrigo, 2005; Moore *et al.*, 2013). Within nutrient addition experiments it is challenging to differentiate these types of co-limitation from each other. The response could possibly be of secondary nature due to addition of a nutrient depleted to levels where it is close to co-limiting (Moore *et al.*, 2008, 2013).

Co-limitation for Fe, silicate and light in three Southern Ocean diatom species has been observed (Hoffmann *et al.*, 2007, 2008). The co-limitation of Fe and light on two different diatom species resulted in C and N per cell decrease with increasing Fe limitation (Bucciarelli *et al.*, 2010). A similar effect was described for phytoplankton with waters from the Ross Sea to which Fe and Vitamin B_{12} were added,

which is largely produced by bacteria and archaea (Bertrand *et al.*, 2015). For the diatom *Phaeodactylum tricornutum,* carbonate availability can limit iron-uptake which was only recently discovered (McQuaid *et al.*, 2018).

Field data from the Atlantic and Pacific Ocean have shown that the nitrogen fixing cyanobacterium *Trichodesmium* is Fe and phosphorous co-limited (Sañudo-Wilhelmy *et al.*, 2001; Sohm *et al.*, 2011; Chappell *et al.*, 2012) and laboratory studies on Fe/P co-limited cells have observed cellular responses in terms of cell size reduction and growth rates with implications on biogeochemical cycles (Walworth *et al.*, 2016). Addition experiments with water from HNLC and Fe-fertilized stations in the Southern Ocean led to a significant increase in heterotrophic prokaryotic production when organic carbon (in the form of glucose) and Fe were added (Obernosterer *et al.*, 2015).

Functional analysis of microbial communities through 'omics' approaches

Classical microbiological analyses depend to a large extent on the cultivation of organisms in pure culture. Through these approaches most of the diversity in ocean and soil ecosystems remained undiscovered as the cultivation and isolation of many microorganisms are difficult or impossible (Rappé and Giovannoni, 2003; Daniel, 2005; Rinke *et al.*, 2013). The discovery of the abundant SAR11 in the ocean through amplification of 16S rRNA gene clone libraries by Giovannoni (1990a), opened new avenues for microbial ecology and the analysis of microbial diversity (Giovannoni *et al.*, 1990). A variety of 16S rRNA gene approaches became popular such as Fluorescence *in situ* hybridization (FISH) and DNA fingerprinting such as terminal restriction fragment length polymorphism (TRFLP) or denaturing gradient gel electrophoresis (DGGE) (Giovanonni and Rappé 2000; Furhman and Hagström 2008; Gasol and Kirchman, 2018). These techniques revealed changes in archaeal and bacterial community composition with biotic and abiotic factors and along spatial and temporal scales (Simon and Daniel, 2011). Sanger sequencing was a slow and expensive method and was replaced around the year 2005 by the so-called 'next-generation sequencing methodologies' (also called high-throughput sequencing) which enables massive parallel sequencing at comparatively low costs (Goodwin *et al.*, 2016). The first published study using next generation 454 pyrosequencing on 16S rRNA (Sogin *et al.*, 2006) shifted the field of microbial oceanography (Fig. 4). It became clear that the diversity of archaea as well as bacteria was heavily underestimated with cultivation-based or fingerprinting approaches. Since then, sequencing machines have further developed into the commonly used Illumina technology yielding several million sequencing reads per sample (Koboldt *et al.*, 2013; Vollmers *et al.*, 2017). However, these modern approaches have a large drawback to classical Sanger sequencing in the base pair length of obtained sequences (several 100 bp to about 1000 bp for Sanger). This affects the taxonomic resolution, especially for the 16S rRNA gene where multiple regions of variable length could be used (Logares *et al.*, 2012; Goodwin *et al.*, 2016). However, it is now even possible to sequence the complete 16S rRNA fragment (Singer *et al.*, 2016). While the analysis of 16S rRNA as the standard marker gene of prokaryotic taxonomy provided enormous information on community composition, its

sequencing depends on amplification approaches for the chosen region (Apprill *et al.*, 2015; Brown *et al.*, 2015; Parada *et al.*, 2016; Walters *et al.*, 2016). Additionally, many microorganisms have more than one copy of the rRNA gene in their genome introducing high biases for the subsequent conclusions on diversity from 16S rRNA (Stoddard *et al.*, 2015).

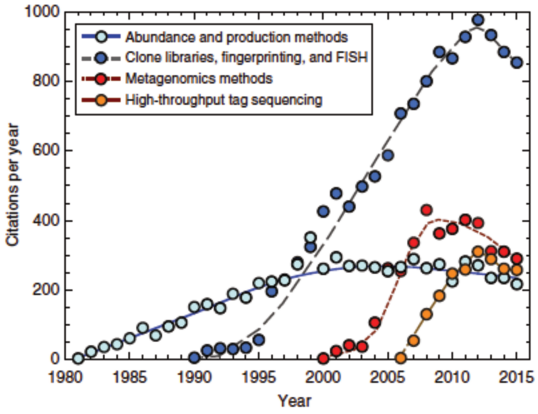


***Figure 4**. Changes in molecular techniques from 1980 to 2015 in terms of citations. Figure from Microbial Ecology of the Ocean, Third Edition, Wiley library, 2018, Gasol & Kirchman*

New approaches with novel technologies aim for the sequencing of individual DNA molecules, avoiding the specific amplification process. Now, even longer sequencing reads can be obtained in this way (Oxford Nanopore, PacBio RS Pacific Biosystems) and it is only a question of time when a high-throughput long-read sequencer will appear on the market and further advance the field. With these methods sequencing of DNA and RNA molecules and thus observing the actual fingerprint of microbial activity at a specific timepoint became possible (Gilbert *et al.*, 2008; Poretsky *et al.*, 2009; Klingenberg and Meinicke, 2017). The term omics has been widely established, originating from the Greek description for a totality of a whole, and in the field of genetics, metagenomics describes the genomic potential of the whole microbial community (Eisen, 2007; Simon and Daniel, 2009).

Through the advances in bioinformatic research the reconstruction of assembled contigs (consensus sequences of DNA) from metagenomic reads into draft genomes has become possible (Sharon and Banfield, 2013). In recent metagenomic approaches the raw sequencing reads are assembled into longer contigs. The assembly is a crucial step and multiple assemblers exist for metagenomic datasets. Once the contigs are assembled, they are grouped by nucleotide composition and coverage data in multiple samples, a process known as metagenomic binning. These draft genomes are also called metagenome-assembled-genomes (MAGs). The number of MAGs assembled from metagenomic projects is increasing (Bowers *et al.*, 2017; Parks *et al.*, 2018; Stewart *et al.*, 2019). This field of analysis is developing rapidly and several terms are not standardized yet. The same holds true for the standardization of workflows which is becoming more important (Quince *et al.*, 2017).

The approaches used in the following chapters are visually summarized in Figure 5. From sampling on board to extraction of nucleic acids for the sequencing of DNA and messenger RNA (mRNA, rRNA depleted) as well as proteins for mass spectrometry analysis. Metatranscriptomics is the study of RNA expressed by a microbial community present in an environmental sample determined by DNA sequencing of reverse transcribed cDNA generated from RNA. In the past decade several studies have shed light on expression profiles of microorganisms in the ocean (Poretsky *et al.*, 2005, 2009; Frias-Lopez *et al.*, 2008; Gilbert *et al.*, 2008; Satinsky *et al.*, 2010; Gifford *et al.*, 2011; Beier *et al.*, 2015). Metaproteomics is the study of the total community proteins with protein identifications using mass spectrometry to determine the mass of peptides derived from extracted proteins. The few existing studies from ocean metaproteomics have provided information on organic matter utilization, trace metal availability and have identified new compounds used as energy sources for microorganisms (Sowell *et al.*, 2009; Morris *et al.*, 2010; Williams *et al.*, 2012; Saito *et al.*, 2014, 2015, 2017, 2019; Bergauer *et al.*, 2018). The major succession from metagenomics is the possibility to know which genes were transcribed, answering which part of the community was active, and with protein signatures even which of these were translated. Prokaryotic half-time of mRNA is limited (up to 5 minutes) and about an order of magnitude less abundant than genes as well as four orders of magnitude less abundant than proteins (Ingraham *et al.*, 1983; Bernstein *et al.*, 2002; Hambraeus *et al.*, 2003; Selinger *et al.*, 2003; Taniguchi *et al.*, 2010; Moran *et al.*, 2013). Prokaryotic proteins reside longer in the oceanic system (up to 20 hours) than mRNA. Thus, there is a temporal decoupling that needs to be considered as proteins persist in a prokaryotic cell long after the mRNA that encoded them (Koch and Levy, 1955; Borek *et al.*, 1958; Mandelstam, 1958; Moran *et al.*, 2013).

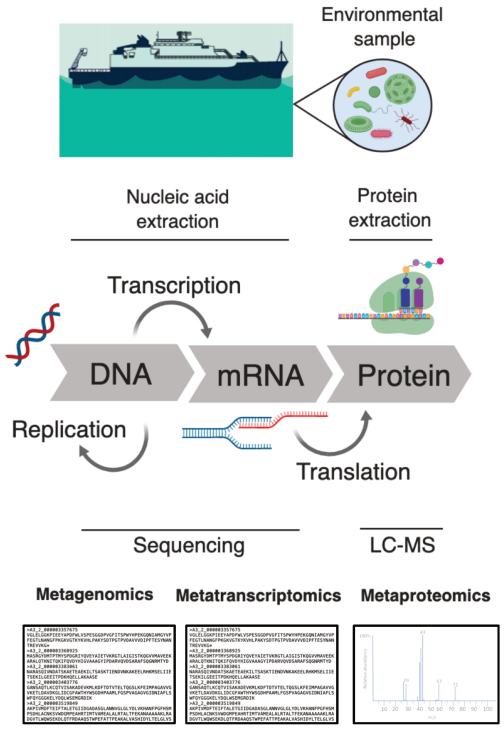


Figure 5. *Workflow of from sampling to analysis of omics analysis used throughout the chapters of this thesis.*

In oceanography, metagenomics has been the objective of several global ocean sequencing projects. The earliest was the Global Ocean sampling by Craig J. Venter resulting in the first description of the largely unknown diversity identifying over 1.2 million previously unknown genes (Fig. 6, Venter *et al.*, 2004; Yooseph *et al.*, 2007). The Spanish Malaspina expedition was an interdisciplinary research project to assess the impact of global change on the oceans and explored their biodiversity with special focus on bathypelagic waters, the least explored biosphere on Earth. The recently published Malaspina Deep-Sea Gene Collection consists of 71% novel genes of patchy distribution challenging the view of

a uniform dark ocean ecosystem (Duarte, 2015; Pernice *et al.*, 2016; Morán *et al.*, 2017; Acinas *et al.*, 2019).

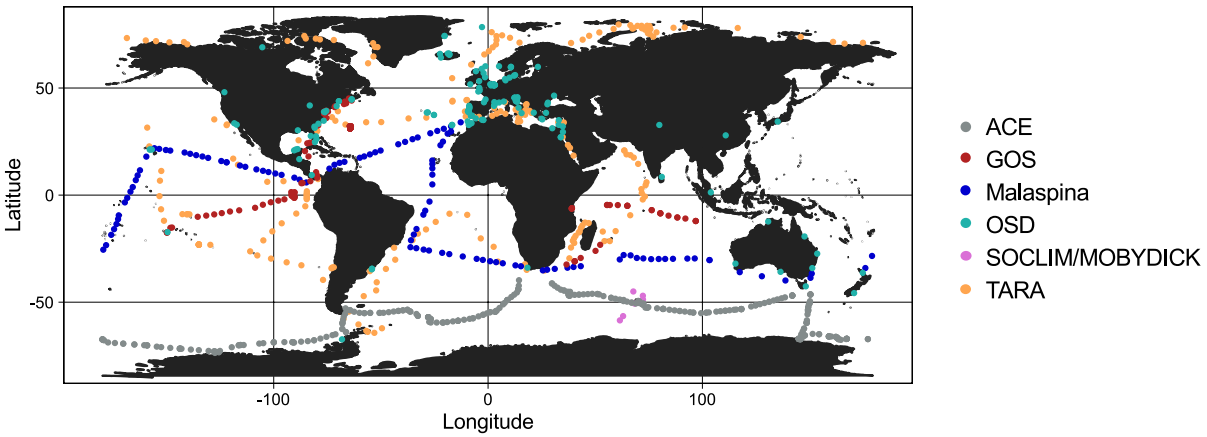


***Figure 6**. Metagenomic sequencing project of the past two decades as well as location of sampling site of this thesis. ACE – Antarctic Circumnavigation Expedition, GOS – Global Ocean Sampling, Malaspina – Expedition, OSD – Ocean Sampling Day, Tara – Oceans Sampling, SOCLIM/MOBYDICK - Data from this thesis*

The TARA Oceans project, a European initiative, aimed at describing the ocean microbiome by sequencing ocean plankton of 8 different size classes in a three-year voyage around the world. Through deep sequencing efforts an ocean microbial reference gene catalog with >40 million nonredundant, mostly novel sequences from viruses, prokaryotes, and picoeukaryotes was created (Brum *et al.*, 2015; Sunagawa *et al.*, 2015; Vargas *et al.*, 2015; Guidi *et al.*, 2016; Roux *et al.*, 2016; Carradec *et al.*, 2018). In 2012, the first ocean sampling day (OSD) was launched as the world's first simultaneous mega sequencing campaign. The idea behind it is to analyze marine microbial community composition and functional traits on a single day which is possible through sampling at coastal sites (Kopf *et al.*, 2015). Another study that is of importance in this field and certainly for the work of this thesis is the Antarctic Circumnavigation Expedition (ACE) which was carried out by the Swiss Polar Institute from December 2016 to March 2017. It is a project of broad spectrum from biology to climatology to oceanography, combining three legs from Cape Town to Hobart, Hobart to Punta Arenas and Punta Arenas to Cape Town circumnavigating 33,463 km. While first publications on climatology had high impact on our understanding of past climate events (Hasenfratz *et al.*, 2019) there are no results published yet on the project profiling the Southern Ocean microbial community.

This 'era of big data' has recently been discussed in terms of conclusions on global microbial diversity (Spang and Offre, 2019). Several studies in the past years have proposed new phyla for bacteria and archaea and a new taxonomy has been introduced based on genome phylogeny (Castelle and Banfield, 2018; Parks *et al.*, 2018). Among the challenges are the improvement of computational infrastructure for the coherent analysis of these massive data and a standardized database for the identification of sequences (Koonin, 2018). Additionally, it is not yet clear how the rare biosphere can be recovered from metagenomics data and a large amount of sequencing reads (up to 70 % in large metagenomic data) have been named 'the dark side of metagenomes' as they cannot be identified with

current database search approaches (Spang and Offre, 2019). However, from a computational point of view the coming years promise further developments and standardization in the 'omics' field. Hopefully, this will result in a 'microbial map of the ocean' as proposed by Santoro (2019), containing information on microorganisms and their genes in the ocean basins alongside with macro- and micronutrient concentrations.

Table 1. *Meta- 'omics' studies on prokaryotes from Antarctic and Southern Ocean waters.*

Analysis	**Reference**	**Sampling location**	**Study focus**
Metagenomic	Grzymski *et al.*, 2012	Antarctic coastal site at Palmer station	Assessment of winter and summer bacterioplankton communities
Metaproteomic	Williams *et al.*, 2012	Antarctic coastal site at Palmer station	
Metagenomic & Metaproteomic	Williams *et al.*, 2013	Coastal East Antarctica, Australian Sector	Algal organic matter utilization by *Flavobacteria*
Metagenomic	Brown *et al.*, 2012	Southern Ocean & Antarctic marine samples as well as under the Ross Sea ice shelf	SAR11 global biogeography
Metagenomic	Wilkins *et al.*, 2013	Transect in Australian sector from Hobart to Mertz Glacier, Antarctica	Biogeographic partitioning of Southern Ocean microorganisms
Metatranscriptomic	Bertrand *et al.*, 2015	Sea ice edge in McMurdo Sound of the Ross Sea	Co-limitation of micronutrients in phytoplankton – bacteria interactions
Metagenomic	Sunagawa *et al.*, 2015	Tara Oceans sampling, Station TARA_84 & TARA_85, Drake passage	Massive sequencing effort for ecology and evolution of marine plankton
Metagenomic & Metatranscriptomic	Caputi *et al.*, 2019	Tara Oceans sampling, (to be noted: stations TARA_84 and TARA_85 were excluded from cooccurrence analysis and considered as outliers)	Community level response of plankton to iron availability in the open ocean

While these environmental databases have contributed to our knowledge on metabolic potential from metagenomics studies, we can see in Fig. 6 that most of these campaigns did not sample waters of the Southern Ocean. In terms of metagenomics or coupled omics studies there are only a handful covering Antarctic waters (Table 1).

The case of the Kerguelen plateau

In the year 1772, Yves de Kerguelen-Tramarec was sent by King Louis XV in exploration of previously reported rich and fertile lands. He discovered, but due to poor weather conditions did not set foot on the sub-Antarctic islands today named Kerguelen and nowadays part of the French Southern and Antarctic Lands (Amiral de Brossard, 1970). Together with the Heard and MacDonald Islands, they represent emerged parts of the mostly submerged Kerguelen plateau which constitutes a microcontinent. The plateau was built from basaltic lava due to volcanic eruptions associated with the Kerguelen volcanic hotspot about 130 million years ago (Olierook *et al.*, 2016; Direen *et al.*, 2017).

Even though as described above, the Southern Ocean is the largest HNLC site in the ocean, close to the island plateau the conditions regularly lead to the development of massive, diatom-dominated phytoplankton blooms that can be observed from satellite images (Blain *et al.*, 2007; Mongin *et al.*, 2008; Quéguiner, 2013). This phenomenon has been described as the island mass effect, a localized increase in phytoplankton biomass near islands (Doty, & Oguri, 1956; Gove *et al.*, 2016) and has been crucial in understanding Darwin's paradox (Darwin, 1842) on the occurrence of coral reefs in the middle of abyssal descents. This form of natural fertilization occurs through several mechanisms that enhance nearshore nutrient concentrations. Current and bathymetry interactions that can drive vertical transport of water masses via upwelling, downwelling and eddy formation, and internal waves have been described as well as island-associated input, such as submarine groundwater discharge and outflow from rivers, which can mobilize and transport sediment and other terrigenous material rich in nutrients (Gove *et al.*, 2016).

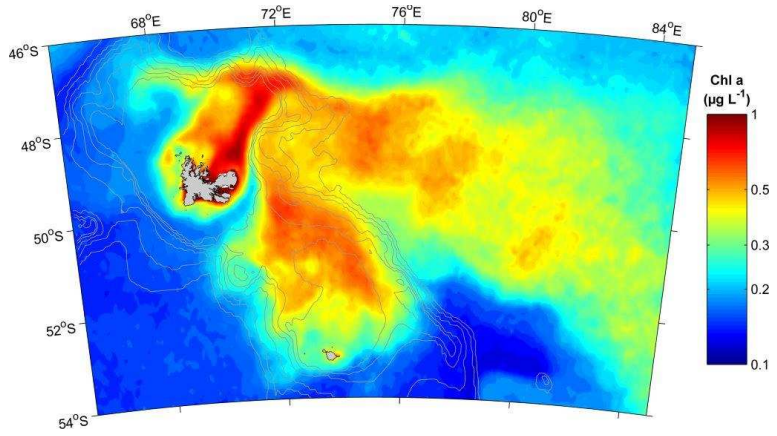


Figure 7. *Climatology of the Kerguelen plateau over a period of ten years. Chlorophyll map courtesy of Prof. Stéphane Blain.*

In the case of the Kerguelen island, vertical mixing supplying iron and major nutrients to surface waters from iron-rich deep waters is most probably due to internal wave activity (Blain *et al.*, 2007). The Kerguelen bloom consists of a plume northeast of the island and a larger area southeast of the island; the latter bloom is constrained by the bathymetry of the plateau (Fig. 7; Blain *et al.*, 2007; Jouandet *et al.*, 2008). The blooms are re-occurring every year, with one large bloom in early summer (mid-December to mid-January). Occasionally, a second bloom is observed in spring (November). The

research cruises to the Kerguelen plateau (KEOPS 1, Jan – Feb 2005; KEOPS 2, Oct – Nov 2011) have contributed enormously to describe the ecological and biogeochemical traits during different bloom stages of this system.

During KEOPS 1, the mechanism of natural iron fertilization was clearly demonstrated (Blain *et al.*, 2007). KEOPS 2 focused on the sources of new and regenerated iron and investigated the ecosystem structure of different phytoplankton blooms in the naturally iron-fertilized region east of Kerguelen Island (Bowie *et al.*, 2015; Quéroué *et al.*, 2015; van der Merwe *et al.*, 2015). Stations that have been studied during the last decade are the two contrasting sites A3 (50°38S – 72°02 E) and KERFIX (50°40S – 68°25 E) (Fig. 8).

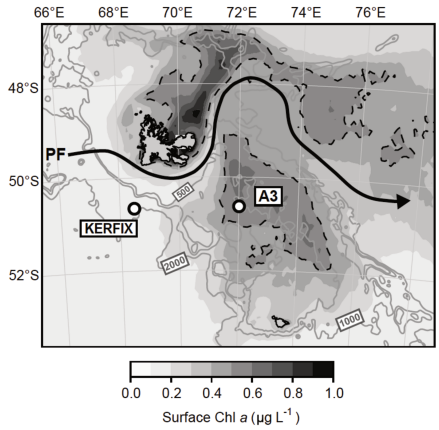


Figure 8. *Chlorophyll and bathymetry map of Kerguelen island and stations KERFIX and A3. Map courtesy Dr. Mathieu Rembauville.*

KERFIX (Kerguelen fixed station) was a five-year observation program that ran from 1991 to 1995 (Jeandel *et al.*, 1998). The KERFIX station is located on the southwestern side of the Kerguelen Plateau. A key objective of the sampling was to describe the factors responsible for the low primary production in a region of the Antarctic Zone (AAZ) characterized by high macronutrient concentrations and low phytoplankton biomass of the HNLC site (Table 2, Fiala *et al.*, 1998; Kopczyńska *et al.*, 1998).

Station A3, on the other hand, is a productive and shallow (540 m) station on the Kerguelen plateau located south of the Polar Front in the Permanently Open Ocean Zone (POOZ) (Table 2).
The so-called "natural laboratory" off Kerguelen Island offers an excellent opportunity to significantly advance the understanding on how natural iron fertilization affects the biogeochemistry of the Southern Ocean.

Table 2. *Dissolved iron (dFe), Chlorophyll a and dissolved organic carbon (DOC) at Station A3 and KERFIX obtained from oceanographic cruises. NA – not available.*

	A3 (on plateau)			**KERFIX (off plateau)**		
Cruise	**dFe (nmol L^{-1})**	**Chl *a* (µg L^{-1})**	**DOC (µM)**	**dFe (nmol L^{-1})**	**Chl *a* (µg L^{-1})**	**DOC (µM)**
KEOPS 1 (January – February 2005)[a]	0.13	1.6 ± 0.5	NA	0.09*	0.2 ± 0.1*	NA
KEOPS 2 (November 2011)[b]	0.16 ± 0.03	2.03 ± 0.34	51.3 ± 1.5	0.13 ± 0.05**	0.25 ± 0.08**	47.8 ± 0.4**
SOCLIM (October 2016)[c]	NA	1.64	52.0	NA	0.32	51.0
MOBYDICK (March 2018)[d]	NA	0.39 ± 0.14	52.5	NA	0.17 ± 0.03	50.4

*Reference site C11 for KEOPS 1 cruise.

**Reference site R-2 for KEOPS 2 cruise. Off-plateau station in vicinity of KERFIX (50°35S - 66°71E)

[a] Lefèvre *et al.*, 2008

[b] Lasbleiz *et al.*, 2014; Quéroué *et al.*, 2015; Tremblay *et al.*, 2015

[c] Liu *et al.*, 2019

[d] unpublished

The Kerguelen area has been extensively studied in terms of microbial diversity and ecology and food web structure as well as DOM composition and sources. The first molecular techniques applied were 16S rRNA clone libraries for the analysis of the KEOPS 1 project. During the bloom, the dominant operational taxonomic units (OTUs) were *Roseobacter*, SAR92 and the Flavobacteria-Bacteroidetes and Cytophaga (FCB) cluster, now coined Bacteroidetes, whereas in the HNLC region, SAR11, *Roseobacter* and *Polaribacter* dominated (West *et al.*, 2008). In a second approach oligonucleotide probes were designed to target five OTUs at a narrow phylogenetic level of 16S rRNA showing that the mentioned groups contributed substantially to bulk abundance and leucine incorporation during the spring phytoplankton bloom in the naturally fertilized region off Kerguelen (Obernosterer *et al.*, 2011). Preferences in the degradation of different DOM were identified for bacterial taxa associated with contrasting blooms. Diatom-derived DOM was identified as an important factor influencing prokaryotic communities in the Southern Ocean in early spring (Landa *et al.*, 2016, 2018). Using co-occurrence analysis, the connection between diatoms and heterotrophic prokaryotes was identified leading to distinct associations (Liu *et al.*, 2019). Through 16S rRNA analysis, the high contribution of *Pelagibacterales*, the order of the most abundant SAR11 in the ocean, to total prokaryotic abundance and activity (micro-autoradiography combined with CARD-FISH) was documented in spring phytoplankton blooms off the Kerguelen Island (Dinasquet *et al.*, 2019).

Microbial uptake of Fe was measured during the spring blooms showing that SAR11 as well as (FCB) were actively taking up Fe (Fourquez *et al.*, 2015, 2016). The gene expression response of SAR11 was analyzed with specific qPCR primers for the isocitrate lyase gene, which induces the glyoxylate shunt, indicating that the HNLC community of SAR11 actively uses this mechanism (Beier *et al.*, 2015).

Chapter overview

The objective of the doctoral thesis was to perform a detailed study on gene expression of heterotrophic prokaryotes with regard to the low organic carbon and iron availability in the study region. The thesis is structured along 3 chapters described below.

Chapter 1

Microbial iron metabolism as revealed by gene expression profiles in contrasted Southern Ocean regimes

Pavla Debeljak, Eve Toulza, Sara Beier, Stéphane Blain and Ingrid Obernosterer
Published in Environmental Microbiology (2019) 21(7), 2360–2374

The aim of this manuscript was to describe gene expression of targeted genes connected to iron uptake, intracellular machinery and low iron adaptations for different prokaryotic communities at three stations close to the Kerguelen plateau.

Author contributions:
ET, IO, SBE and SBL designed the research. SBE extracted RNA and verified the quality of the extracts. ET and PD designed databases for iron metabolism specific genes. PD did the bioinformatic analysis and visualized the data. SBE advised with the normalization approach. PD, IO and SBL wrote the manuscript and all co-authors revised the draft version and approved the final version of the manuscript.

Chapter 2

Marine particles as source of trace metals for prokaryotes: A metatranscriptomic view

Pavla Debeljak, Stéphane Blain, Pier van der Merwe, Andrew R. Bowie, Ingrid Obernosterer,

The aim of this project was to investigate trace metal bioavailability of marine particles for prokaryotic communities from the Southern Ocean, though geochemical particle description, prokaryotic activity and gene expression.

Author contributions:
This thesis chapter is part of the Heard Earth Ocean Biosphere Interactions Project (PI Mike Coffin, IMAS, Hobart, Australia). PD performed all molecular analyses, including the nucleic acid extraction

and the synthesis of mRNA standards as well as the design of all bioinformatic procedures resulting in data analysis and visualization for amplicon sequencing and metatranscriptomic data. Contextual data was provided by SB, Jocelyn Carparros and Philippe Catala. IO and SB designed the research and carried out the experiment in collaboration with Pier van der Merve and Andy Bowie (UTAS-IMAS, Hobart, Australia) during an oceanographic cruise aboard the Australian R/V Investigator (January to February 2016).

Chapter 3

Deciphering organic matter utilization through prokaryotic membrane transporter expression at two Southern Ocean sites in early spring and late summer

Pavla Debeljak, Barbara Bayer, Ying Sun, Ingrid Obernosterer & Gerhard J. Herndl

The aim of this project was to focus on organic matter availability through membrane transporter by investigating and comparing the meta-genomes, -transcriptomes and -proteomes of the free-living microbial communities present in early spring (October 2016) and late summer (March 2018) at station A3 and KERFIX.

Author contribution:
PD, BB, IO and GJH designed the research and PD conducted all on-board sampling for both cruises. PD extracted DNA and RNA and organized sequencing. PD and BB extracted all protein samples and organized MS measurements at the Vienna metabolomics center. PD designed all bioinformatic procedures and conducted the analysis. YS curated metagenome assembled genomes from the co-assembly and calculated phylogenetic similarity.

References

Acinas, S.G., Sánchez, P., Salazar, G., Cornejo-Castillo, F.M., Sebastián, M., Logares, R., et al. (2019) Metabolic Architecture of the Deep Ocean Microbiome. *bioRxiv*.

Almaas, E., Oltvai, Z.N., and Barabási, A.-L. (2005) The Activity Reaction Core and Plasticity of Metabolic Networks. *PLOS Computational Biology* **1**: e68.

Andrews, S.C., Robinson, A.K., and Rodríguez-Quiñones, F. (2003) Bacterial iron homeostasis. *FEMS Microbiology Reviews* **27**: 215–237.

Apprill, A., McNally, S., Parsons, R., and Weber, L. (2015) Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquatic Microbial Ecology* **75**: 129–137.

Ardyna, M., Lacour, L., Sergi, S., d'Ovidio, F., Sallée, J.-B., Rembauville, M., et al. (2019) Hydrothermal vents trigger massive phytoplankton blooms in the Southern Ocean. *Nature Communications* **10**: 2451.

Arrigo, K.R. (2005) Marine microorganisms and global nutrient cycles. *Nature* **437**: 349.

Assmy, P., Smetacek, V., Montresor, M., Klaas, C., Henjes, J., Strass, V.H., et al. (2013) Thick-shelled, grazer-protected diatoms decouple ocean carbon and silicon cycles in the iron-limited Antarctic Circumpolar Current. *PNAS* **110**: 20633–20638.

Azam (1983) The ecological role of water-column microbes in the sea. *The Ecological Role of Water-Column Microbes in the Sea* **10**: 257–263.

Azam, F. and Malfatti, F. (2007) Microbial structuring of marine ecosystems. *Nature Reviews Microbiology* **5**: 782–791.

de Baar, H.J.W., Boyd, P.W., Coale, K.H., Landry, M.R., Tsuda, A., Assmy, P., et al. (2005) Synthesis of iron fertilization experiments: From the Iron Age in the Age of Enlightenment. *Journal of Geophysical Research: Oceans* **110**:.

Baines, S.B. and Pace, M.L. (1991) The production of dissolved organic matter by phytoplankton and its importance to bacteria: Patterns across marine and freshwater systems. *Limnology and Oceanography* **36**: 1078–1090.

Baltar, F., Gutiérrez-Rodríguez, A., Meyer, M., Skudelny, I., Sander, S., Thomson, B., et al. (2018) Specific Effect of Trace Metals on Marine Heterotrophic Microbial Activity and Diversity: Key Role of Iron and Zinc and Hydrocarbon-Degrading Bacteria. *Front Microbiol* **9**:.

Barbeau, K., Rue, E.L., Bruland, K.W., and Butler, A. (2001) Photochemical cycling of iron in the surface ocean mediated by microbial iron(iii)-binding ligands. *Nature* **413**: 409.

Barbeau, K., Rue, E.L., Trick, C.G., Bruland, K.W., and Butler, A. (2003) Photochemical reactivity of siderophores produced by marine heterotrophic bacteria and cyanobacteria based on characteristic Fe(III) binding groups. *Limnology and Oceanography* **48**: 1069–1078.

Bauer, J.E., Cai, W.-J., Raymond, P.A., Bianchi, T.S., Hopkinson, C.S., and Regnier, P.A.G. (2013) The changing carbon cycle of the coastal ocean. *Nature* **504**: 61–70.

Becker, J.W., Berube, P.M., Follett, C.L., Waterbury, J.B., Chisholm, S.W., DeLong, E.F., and Repeta, D.J. (2014) Closely related phytoplankton species produce similar suites of dissolved organic matter. *Front Microbiol* **5**:.

Beier, S., Rivers, A.R., nn Moran, M.A., and Obernosterer, I. (2015) The transcriptional response of prokaryotes to phytoplankton-derived dissolved organic matter in seawater. *Environmental microbiology* **17**: 3466–3480.

Bergauer, K., Fernandez-Guerra, A., Garcia, J.A.L., Sprenger, R.R., Stepanauskas, R., Pachiadaki, M.G., et al. (2018) Organic matter processing by microbial communities throughout the Atlantic water column as revealed by metaproteomics. *PNAS* **115**: E400–E408.

Bernstein, J.A., Khodursky, A.B., Lin, P.-H., Lin-Chao, S., and Cohen, S.N. (2002) Global analysis of mRNA decay and abundance in Escherichia coli at single-gene resolution using two-color fluorescent DNA microarrays. *Proc Natl Acad Sci USA* **99**: 9697–9702.

Bertrand, E.M., McCrow, J.P., Moustafa, A., Zheng, H., McQuaid, J.B., Delmont, T.O., et al. (2015) Phytoplankton-bacterial interactions mediate micronutrient colimitation at the coastal Antarctic sea ice edge. *Proceedings of the National Academy of Sciences of the United States of America* **112**: 9938–43.

Blain, S., Queguiner, B., Armand, L.K., Belviso, S., and Bomb, B. (2007) Effect of natural iron fertilization on carbon sequestration in the Southern Ocean. *Nature* **446**: 1070–1074.

Blain, S. and Tagliabue, A. (2016) Iron Cycle in Oceans, 1st ed. Wiley-ISTE.

Borek, E., Ponticorvo, L., and Rittenberg, D. (1958) PROTEIN TURNOVER IN MICRO-ORGANISMS*. *Proc Natl Acad Sci U S A* **44**: 369–374.
Bowers, R.M., Kyrpides, N.C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T.B.K., et al. (2017) Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nature Biotechnology* **35**: 725–731.
Bowie, A.R., van der Merwe, P., Quéroué, F., Trull, T., Fourquez, M., Planchon, F., et al. (2015) Iron budgets for three distinct biogeochemical sites around the Kerguelen Archipelago (Southern Ocean) during the natural fertilisation study, KEOPS-2. *Biogeosciences* **12**: 4421–4445.
Boyd, P.W., Jickells, T., Law, C.S., Blain, S., Boyle, E.A., Buesseler, K.O., et al. (2007) Mesoscale Iron Enrichment Experiments 1993-2005: Synthesis and Future Directions. *Science* **315**: 612–617.
Broeker, W. (1991) The Great Ocean Conveyor. *oceanog* **4**: 79–89.
Brown, C.T., Hug, L.A., Thomas, B.C., Sharon, I., Castelle, C.J., Singh, A., et al. (2015) Unusual biology across a group comprising more than 15% of domain Bacteria. *Nature* **523**: 208–211.
Brown, M.V., Lauro, F.M., DeMaere, M.Z., Muir, L., Wilkins, D., Thomas, T., et al. (2012) Global biogeography of SAR11 marine bacteria. *Mol Syst Biol* **8**: 595.
Bruland, K.W., Donat, J.R., and Hutchins, D.A. (1991) Interactive influences of bioactive trace metals on biological production in oceanic waters. *Limnology and Oceanography* **36**: 1555–1577.
Brum, J.R., Ignacio-Espinoza, J.C., Roux, S., Doulcier, G., Acinas, S.G., Alberti, A., et al. (2015) Patterns and ecological drivers of ocean viral communities. *Science* **348**: 1261498.
Bucciarelli, E., Pondaven, P., and Sarthou, G. (2010) Effects of an iron-light co-limitation on the elemental composition (Si, C, N) of the marine diatoms <I>Thalassiosira oceanica</I> and <I>Ditylum brightwellii</I>. *Biogeosciences* **7**: 657–669.
Butler, A. (2005) Marine Siderophores and Microbial Iron Mobilization. *Biometals* **18**: 369–374.
Caputi, L., Carradec, Q., Eveillard, D., Kirilovsky, A., Pelletier, E., Karlusich, J.J.P., et al. (2019) Community-Level Responses to Iron Availability in Open Ocean Plankton Ecosystems. *Global Biogeochemical Cycles*.
Carradec, Q., Pelletier, E., Da Silva, C., Alberti, A., Seeleuthner, Y., Blanc-Mathieu, R., et al. (2018) A global ocean atlas of eukaryotic genes. *Nature Communications* **9**: 373.
Cartron, M.L., Maddocks, S., Gillingham, P., Craven, C.J., and Andrews, S.C. (2006) Feo – Transport of Ferrous Iron into Bacteria. *Biometals* **19**: 143–157.
Castelle, C.J. and Banfield, J.F. (2018) Major New Microbial Groups Expand Diversity and Alter our Understanding of the Tree of Life. *Cell* **172**: 1181–1197.
Chappell, P.D., Moffett, J.W., Hynes, A.M., and Webb, E.A. (2012) Molecular evidence of iron limitation and availability in the global diazotroph Trichodesmium. *ISME J* **6**: 1728–1739.
Chiba, S., Ishimaru, T., Hosie, G.W., and Fukuchi, M. (2001) Spatio-temporal variability of zooplankton community structure off east Antarctica (90 to 160°E). *Marine Ecology Progress Series* **216**: 95–108.
Chisholm, S.W., Falkowski, P.G., and Cullen, J.J. (2001) Dis-Crediting Ocean Fertilization. *Science* **294**: 309–310.
Cho, B.C. and Azam, F. (1988) Major role of bacteria in biogeochemical fluxes in the ocean's interior. *Nature* **332**: 441.
Church, M.J., Hutchins, D.A., and Ducklow, H.W. (2000) Limitation of Bacterial Growth by Dissolved Organic Matter and Iron in the Southern Ocean. *Appl Environ Microbiol* **66**: 455–466.
Coale, K.H., Johnson, K.S., Fitzwater, S.E., Gordon, R.M., Tanner, S., Chavez, F.P., et al. (1996) A massive phytoplankton bloom induced by an ecosystem-scale iron fertilization experiment in the equatorial Pacific Ocean. *Nature* **383**: 495.
Cooper Leslie Hugh Norman and Atkins William Ringrose Gelston (1935) Iron in the sea and in marine plankton. *Proceedings of the Royal Society of London Series B - Biological Sciences* **118**: 419–438.
Daniel, R. (2005) The metagenomics of soil. *Nature Reviews Microbiology* **3**: 470.
Darwin, C. (1842) The Structure and Distribution of Coral Reefs: Being the First Part of the Geology of the Voyage of the Beagle, Under the Command of Capt. Fitzroy, R.N. During the Years 1832 to 1836, Smith, Elder and Company.
de, B.A. (1970) Kerguelen le découvreue et ses îles., France-Empire, 1970.
Delmont, T.O., Quince, C., Shaiber, A., Esen, Ö.C., Lee, S.T., Rappé, M.S., et al. (2018) Nitrogen-fixing populations of Planctomycetes and Proteobacteria are abundant in surface ocean metagenomes. *Nature Microbiology* **3**: 804.

DeLong, E.F., Preston, C.M., Mincer, T., Rich, V., Hallam, S.J., Frigaard, N.-U., et al. (2006) Community Genomics Among Stratified Microbial Assemblages in the Ocean's Interior. *Science* **311**: 496–503.
Dinasquet, J., Landa, M., and Obernosterer, I. (2019) High contribution of Pelagibacterales to bacterial community composition and activity in spring blooms off Kerguelen Island (Southern Ocean). *bioRxiv* 633925.
Direen, N.G., Cohen, B.E., Maas, R., Frey, F.A., Whittaker, J.M., Coffin, M.F., et al. (2017) Naturaliste Plateau: constraints on the timing and evolution of the Kerguelen Large Igneous Province and its role in Gondwana breakup. *Australian Journal of Earth Sciences* **64**: 851–869.
Doty, M. S., & Oguri, M. (1956) The island mass effect. *Journal du Conseil* **22**: 33–37.
Duarte, C.M. (2015) Seafaring in the 21St Century: The Malaspina 2010 Circumnavigation Expedition. *Limnology and Oceanography Bulletin*.
Duce, R.A. and Tindale, N.W. (1991) Atmospheric transport of iron and its deposition in the ocean. *Limnology and Oceanography* **36**: 1715–1726.
Dugdale, R.C. and Wilkerson, F.P. (1991) Low specific nitrate uptake rate: A common feature of high-nutrient, low-chlorophyll marine ecosystems. *Limnology and Oceanography* **36**: 1678–1688.
Eisen, J.A. (2007) Environmental Shotgun Sequencing: Its Potential and Challenges for Studying the Hidden World of Microbes. *PLOS Biology* **5**: e82.
Emerson, D. (2019) Biogenic Iron Dust: A Novel Approach to Ocean Iron Fertilization as a Means of Large Scale Removal of Carbon Dioxide From the Atmosphere. *Front Mar Sci* **6**:.
Falkowski, P.G., Barber, R.T., and Smetacek, V. (1998) Biogeochemical Controls and Feedbacks on Ocean Primary Production. *Science* **281**: 200–206.
Falkowski, P.G., Fenchel, T., and Delong, E.F. (2008) The microbial engines that drive Earth's biogeochemical cycles. *Science* **320**: 1034–1039.
Falkowski, P.G. and Raven, J.A. (2013) Aquatic Photosynthesis: Second Edition, Princeton University Press.
Fiala, M., Kopczynska, E.E., Jeandel, C., Oriol, L., and Vetion, G. (1998) Seasonal and interannual variability of size-fractionated phytoplankton biomass and community structure at station Kerfix, off the Kerguelen Islands, Antarctica. *J Plankton Res* **20**: 1341–1356.
Fletcher, S.E.M., Gruber, N., Jacobson, A.R., Doney, S.C., Dutkiewicz, S., Gerber, M., et al. (2006) Inverse estimates of anthropogenic CO2 uptake, transport, and storage by the ocean. *Global Biogeochemical Cycles* **20**:.
Flynn, K.J., Clark, D.R., and Xue, Y. (2008) Modeling the Release of Dissolved Organic Matter by Phytoplankton1. *Journal of Phycology* **44**: 1171–1187.
Fogg, G.E. (1983) ecological significance of extracellular products of phytoplankton photosynthesis. *Botanica marina*.
Fourquez, M., Beier, S., Jongmans, E., Hunter, R., and Obernosterer, I. (2016) Uptake of Leucine, Chitin, and Iron by Prokaryotic Groups during Spring Phytoplankton Blooms Induced by Natural Iron Fertilization off Kerguelen Island (Southern Ocean). *Frontiers in Marine Science* **3**:.
Fourquez, M., Devez, A., Schaumann, A., Guéneuguès, A., Jouenne, T., Obernosterer, I., and Blain, S. (2014) Effects of iron limitation on growth and carbon metabolism in oceanic and coastal heterotrophic bacteria. *Limnology and Oceanography* **59**: 349–360.
Fourquez, M., Obernosterer, I., Davies, D.M., Trull, T.W., and Blain, S. (2015) Microbial iron uptake in the naturally fertilized waters in the vicinity of the Kerguelen Islands: phytoplankton–bacteria interactions. *Biogeosciences* **12**: 1893–1906.
Frias-Lopez, J., Shi, Y., Tyson, G.W., Coleman, M.L., Schuster, S.C., Chisholm, S.W., and DeLong, E.F. (2008) Microbial community gene expression in ocean surface waters. *Proc Natl Acad Sci USA* **105**: 3805.
Frölicher, T.L., Sarmiento, J.L., Paynter, D.J., Dunne, J.P., Krasting, J.P., and Winton, M. (2014) Dominance of the Southern Ocean in Anthropogenic Carbon and Heat Uptake in CMIP5 Models. *J Climate* **28**: 862–886.
Fuhrman, J. (1987) Close coupling between release and uptake of dissolved free amino acids in seawater studied by an isotope dilution approach. *Marine Ecology Progress Series* 8.
Galbraith, E.D. and Martiny, A.C. (2015) A simple nutrient-dependence mechanism for predicting the stoichiometry of marine ecosystems. *PNAS* **112**: 8199–8204.
Gasol, J.M. and Kirchman, D.L. (2018) Microbial Ecology of the Oceans, John Wiley & Sons.
Gifford, S.M., Sharma, S., Rinta-Kanto, J.M., and Moran, M.A. (2011) Quantitative analysis of a deeply sequenced marine microbial metatranscriptome. *The ISME journal* **5**: 461–472.

Gilbert, J.A., Field, D., Huang, Y., Edwards, R., Li, W., Gilna, P., and Joint, I. (2008) Detection of Large Numbers of Novel Sequences in the Metatranscriptomes of Complex Marine Microbial Communities. *PLOS ONE* **3**: e3042.
Giovannoni, S.J., Britschgi, T.B., Moyer, C.L., and Field, K.G. (1990) Genetic diversity in Sargasso Sea bacterioplankton. *Nature* **345**: 60.
Goodwin, S., McPherson, J.D., and McCombie, W.R. (2016) Coming of age: ten years of next-generation sequencing technologies. *Nature Reviews Genetics* **17**: 333–351.
Gove, J.M., McManus, M.A., Neuheimer, A.B., Polovina, J.J., Drazen, J.C., Smith, C.R., et al. (2016) Near-island biological hotspots in barren ocean basins. *Nature Communications* **7**: 10581.
Gram, L., Grossart, H.-P., Schlingloff, A., and Kiørboe, T. (2002) Possible Quorum Sensing in Marine Snow Bacteria: Production of Acylated Homoserine Lactones by Roseobacter Strains Isolated from Marine Snow. *Appl Environ Microbiol* **68**: 4111–4116.
Gran, H.H. (1931) On the conditions for the production of plankton in the sea. *Conseil Perm Internat pour I'Explor de la Mer Rapp et Proces-Verb* **75**: 37–46.
Granum, E., Kirkvold, S., and Myklestad, S.M. (2002) Cellular and extracellular production of carbohydrates and amino acids by the marine diatom Skeletonema costatum: diel variations and effects of N depletion. *Marine Ecology Progress Series* **242**: 83–94.
Grzymski, J.J., Riesenfeld, C.S., Williams, T.J., Dussaq, A.M., Ducklow, H., Erickson, M., et al. (2012) A metagenomic assessment of winter and summer bacterioplankton from Antarctica Peninsula coastal surface waters. *The ISME Journal* **6**:.
Guidi, L., Chaffron, S., Bittner, L., Eveillard, D., Larhlimi, A., Roux, S., et al. (2016) Plankton networks driving carbon export in the oligotrophic ocean. *Nature* **532**: 465–470.
Hambraeus, G., von Wachenfeldt, C., and Hederstedt, L. (2003) Genome-wide survey of mRNA half-lives in Bacillus subtilis identifies extremely stable mRNAs. *Mol Genet Genomics* **269**: 706–714.
Hansell, D., Carlson, C., Repeta, D., and Schlitzer, R. (2009) Dissolved organic matter in the ocean. *Oceanography* **22**: 202–211.
Hansell, D.A. (2013) Recalcitrant Dissolved Organic Carbon Fractions. *Annual Review of Marine Science* **5**: 421–445.
Hansman, R.L., Dittmar, T., and Herndl, G.J. (2015) Conservation of dissolved organic matter molecular composition during mixing of the deep water masses of the northeast Atlantic Ocean. *Marine Chemistry* **177**: 288–297.
Hart, T.J. (1934) On the Phytoplankton of the South-west Atlantic and the Bellingshausen Sea, 1929-31, University Press.
Hasenfratz, A.P., Jaccard, S.L., Martínez-García, A., Sigman, D.M., Hodell, D.A., Vance, D., et al. (2019) The residence time of Southern Ocean surface waters and the 100,000-year ice age cycle. *Science* **363**: 1080–1084.
Hedges, J.I. (1992) Global biogeochemical cycles: progress and problems. *Marine Chemistry* **39**: 67–93.
Heissenberger, A. and Herndl, G. (1994) Formation of high molecular weight material by free-living marine bacteria. *Marine Ecology Progress Series* **111**: 129–135.
Herndl, G.J. and Reinthaler, T. (2013) Microbial control of the dark end of the biological pump. *Nature Geoscience* **6**: 718–724.
Hider, R.C. and Kong, X. (2010) Chemistry and biology of siderophores. *Nat Prod Rep* **27**: 637–657.
Hoffmann, L.J., Peeken, I., and Lochte, K. (2007) Co-limitation by iron, silicate, and light of three Southern Ocean diatom species. *Biogeosciences Discussions* **4**: 209–247.
Hoffmann, L.J., Peeken, I., and Lochte, K. (2008) Iron, silicate, and light co-limitation of three Southern Ocean diatom species. *Polar Biol* **31**: 1067–1080.
Holm, R.H., Kennepohl, P., and Solomon, E.I. (1996) Structural and Functional Aspects of Metal Sites in Biology. *Chem Rev* **96**: 2239–2314.
Hunt, B.P.V., Pakhomov, E.A., and McQuaid, C.D. (2001) Short-term variation and long-term changes in the oceanographic environment and zooplankton community in the vicinity of a sub-Antarctic archipelago. *Marine Biology* **138**: 369–381.
Ingraham, J.L., Maaløe, O., and Neidhardt, F.C. (1983) Growth of the bacterial cell, Sinauer Associates.
Jeandel, C., Ruiz-Pino, D., Gjata, E., Poisson, A., Brunet, C., Charriaud, E., et al. (1998) KERFIX, a time-series station in the Southern Ocean: a presentation. *Journal of Marine Systems* **17**: 555–569.
Jiao, N., Herndl, G.J., Hansell, D.A., Benner, R., Kattner, G., Wilhelm, S.W., et al. (2010) Microbial production of recalcitrant dissolved organic matter: long-term carbon storage in the global ocean. *Nature Reviews Microbiology* **8**: 593–599.

Johnson, K.S., Gordon, R.M., and Coale, K.H. (1997) What controls dissolved iron concentrations in the world ocean? *Marine Chemistry* **57**: 137–161.
Jouandet, M.P., Blain, S., Metzl, N., Brunet, C., Trull, T.W., and Obernosterer, I. (2008) A seasonal carbon budget for a naturally iron-fertilized bloom over the Kerguelen Plateau in the Southern Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography* **55**: 856–867.
Katoh, H., Hagino, N., Grossman, A.R., and Ogawa, T. (2001) Genes Essential to Iron Transport in the Cyanobacterium Synechocystis sp. Strain PCC 6803. *Journal of Bacteriology* **183**: 2779–2784.
Khatiwala, S., Primeau, F., and Hall, T. (2009) Reconstruction of the history of anthropogenic CO_2 concentrations in the ocean. *Nature* **462**: 346–349.
Kim, S., Simpson, A.J., Kujawinski, E.B., Freitas, M.A., and Hatcher, P.G. (2003) High resolution electrospray ionization mass spectrometry and 2D solution NMR for the analysis of DOM extracted by C18 solid phase disk. *Organic Geochemistry* **34**: 1325–1335.
Kirchman, D. (1990) Limitation of bacterial growth by dissolved organic matter in the subarctic Pacific. *Marine Ecology Progress Series* **62**: 47–54.
Kirchman, D.L. (2000) Microbial ecology of the oceans, Wiley-Liss.
Kirchman, D.L. (2010) Microbial Ecology of the Oceans, John Wiley & Sons.
Kirchman, D.L. (1994) The uptake of inorganic nutrients by heterotrophic bacteria. *Microb Ecol* **28**: 255–271.
Klingenberg, H. and Meinicke, P. (2017) How to normalize metatranscriptomic count data for differential expression analysis. *PeerJ* **5**: e3859.
Koboldt, D.C., Steinberg, K.M., Larson, D.E., Wilson, R.K., and Mardis, E.R. (2013) The next-generation sequencing revolution and its impact on genomics. *Cell* **155**: 27–38.
Koch, A.L. and Levy, H.R. (1955) Protein turnover in growing cultures of Escherichia coli. *J Biol Chem* **217**: 947–957.
Koedooder, C., Gueneugues, A., Van Geersdaele, R., Verge, V., Bouget, F.-Y., Labreuche, Y., et al. (2018) The Role of the Glyoxylate Shunt in the Acclimation to Iron Limitation in Marine Heterotrophic Bacteria. *Frontiers in Marine Science* **5**:.
Koonin, E.V. (2018) Environmental microbiology and metagenomics: the Brave New World is here, what's next? *Environ Microbiol* **20**: 4210–4212.
Kopczyńska, E.E., Fiala, M., and Jeandel, C. (1998) Annual and interannual variability in phytoplankton at a permanent station off Kerguelen Islands, Southern Ocean. *Polar Biol* **20**: 342–351.
Kopf, A., Bicak, M., Kottmann, R., Schnetzer, J., Kostadinov, I., Lehmann, K., et al. (2015) The ocean sampling day consortium. *GigaScience* **4**: 27.
Krom, M.D., Herut, B., and Mantoura, R.F.C. (2004) Nutrient budget for the Eastern Mediterranean: Implications for phosphorus limitation. *Limnology and Oceanography* **49**: 1582–1592.
Krom, M.D., Kress, N., Brenner, S., and Gordon, L.I. (1991) Phosphorus limitation of primary productivity in the eastern Mediterranean Sea. *Limnology and Oceanography* **36**: 424–432.
Kujawinski, E.B. (2011) The Impact of Microbial Metabolism on Marine Dissolved Organic Matter. *Annual Review of Marine Science* **3**: 567–599.
Kujawinski, E.B., Longnecker, K., Blough, N.V., Vecchio, R.D., Finlay, L., Kitner, J.B., and Giovannoni, S.J. (2009) Identification of possible source markers in marine dissolved organic matter using ultrahigh resolution mass spectrometry. *Geochimica et Cosmochimica Acta* **73**: 4384–4399.
Kuma, K., Katsumoto, A., Kawakami, H., Takatori, F., and Matsunaga, K. (1998) Spatial variability of Fe(III) hydroxide solubility in the water column of the northern North Pacific Ocean. *Deep Sea Research Part I: Oceanographic Research Papers* **45**: 91–113.
Kuma, K., Nishioka, J., and Matsunaga, K. (1996) Controls on iron(III) hydroxide solubility in seawater: The influence of pH and natural organic chelators. *Limnology and Oceanography* **41**: 396–407.
Landa, M., Blain, S., Christaki, U., Monchy, S., and Obernosterer, I. (2016) Shifts in bacterial community composition associated with increased carbon cycling in a mosaic of phytoplankton blooms. *The ISME journal* **10**: 39–50.
Landa, M., Blain, S., Harmand, J., Monchy, S., Rapaport, A., and Obernosterer, I. (2018) Major changes in the composition of a Southern Ocean bacterial community in response to diatom-derived dissolved organic matter. *FEMS Microbiol Ecol* **94**:.
Lasbleiz, M., Leblanc, K., Armand, L.K., Christaki, U., Georges, C., Obernosterer, I., and Quéguiner, B. (2016) Composition of diatom communities and their contribution to plankton biomass in the naturally iron-fertilized region of Kerguelen in the Southern Ocean. *FEMS Microbiology Ecology* **92**: fiw171.

Lasbleiz, M., Leblanc, K., Blain, S., Ras, J., Cornet-Barthaux, V., Hélias Nunige, S., and Queguiner, B. (2014) Pigments, elemental composition (C, N, P, and Si), and stoichiometry of particulate matter in the naturally iron fertilized region of Kerguelen in the Southern Ocean. *Biogeosciences* **11**: 5931–5955.

Lefèvre, D., Guigue, C., and Obernosterer, I. (2008) The metabolic balance at two contrasting sites in the Southern Ocean: The iron-fertilized Kerguelen area and HNLC waters. *Deep Sea Research Part II: Topical Studies in Oceanography* **55**: 766–776.

Liu, X. and Millero, F.J. (2002) The solubility of iron in seawater. *Marine Chemistry* **77**: 43–54.

Liu, Y., Debeljak, P., Rembauville, M., Blain, S., and Obernosterer, I. (2019) Diatoms shape the biogeography of heterotrophic prokaryotes in early spring in the Southern Ocean. *Environmental Microbiology* **21**: 1452–1465.

Logares, R., Haverkamp, T.H.A., Kumar, S., Lanzén, A., Nederbragt, A.J., Quince, C., and Kauserud, H. (2012) Environmental microbiology through the lens of high-throughput DNA sequencing: Synopsis of current platforms and bioinformatics approaches. *Journal of Microbiological Methods* **91**: 106–113.

Mandelstam, J. (1958) Turnover of protein in growing and non-growing populations of Escherichia coli. *Biochem J* **69**: 110–119.

Marchetti, A. and Cassar, N. (2009) Diatom elemental and morphological changes in response to iron limitation: a brief review with potential paleoceanographic applications. *Geobiology* **7**: 419–431.

Martin, J.H. (1990) Glacial-interglacial CO2 change: The Iron Hypothesis. *Paleoceanography* **5**: 1–13.

Martin, J.H. and Fitzwater, S.E. (1988) Iron deficiency limits phytoplankton growth in the north-east Pacific subarctic. *Nature* **331**: 341.

Martin, J.H., Fitzwater, S.E., and Gordon, R.M. (1990) Iron deficiency limits phytoplankton growth in Antarctic waters. *Global Biogeochemical Cycles* **4**: 5–12.

Martin, J.H., Gordon, R.M., and Fitzwater, S.E. (1990) Iron in Antarctic waters. *Nature* **345**: 156.

Martinez, J.S., Carter-Franklin, J.N., Mann, E.L., Martin, J.D., Haygood, M.G., and Butler, A. (2003) Structure and membrane affinity of a suite of amphiphilic siderophores produced by a marine bacterium. *PNAS* **100**: 3754–3759.

McQuaid, J.B., Kustka, A.B., Oborník, M., Horák, A., McCrow, J.P., Karas, B.J., et al. (2018) Carbonate-sensitive phytotransferrin controls high-affinity iron uptake in diatoms. *Nature* **555**: 534–537.

van der Merwe, P., Bowie, A.R., Quéroué, F., Armand, L., Blain, S., Chever, F., et al. (2015) Sourcing the iron in the naturally fertilised bloom around the Kerguelen Plateau: particulate trace metal dynamics. *Biogeosciences* **12**: 739–755.

Mills, M.M., Moore, C.M., Langlois, R., Milne, A., Achterberg, E., Nachtigall, K., et al. (2008) Nitrogen and phosphorus co-limitation of bacterial productivity and growth in the oligotrophic subtropical North Atlantic. *Limnology and Oceanography* **53**: 824–834.

Minas, H. and Minas, M. (1992) Net community production in high nutrient-low chlorophyll waters of the tropical and antarctic oceans - grazing vs iron hypothesis. *Oceanologica Acta* **15**: 145–162.

Mongin, M., Molina, E., and Trull, T.W. (2008) Seasonality and scale of the Kerguelen plateau phytoplankton bloom: A remote sensing and modeling analysis of the influence of natural iron fertilization in the Southern Ocean. *Deep-Sea Research Part II* **55**: 880–892.

Moore, C.M., Mills, M.M., Arrigo, K.R., Berman-Frank, I., Bopp, L., Boyd, P.W., et al. (2013) Processes and patterns of oceanic nutrient limitation. *Nature Geoscience* **6**: 701–710.

Moore, C.M., Mills, M.M., Langlois, R., Milne, A., Achterberg, E.P., Roche, J.L., and Geider, R.J. (2008) Relative influence of nitrogen and phosphorous availability on phytoplankton physiology and productivity in the oligotrophic sub-tropical North Atlantic Ocean. *Limnology and Oceanography* **53**: 291–305.

Mopper, K., Zhou, X., Kieber, R.J., Kieber, D.J., Sikorski, R.J., and Jones, R.D. (1991) Photochemical degradation of dissolved organic carbon and its impact on the oceanic carbon cycle. *Nature* **353**: 60.

Moran, M.A., Kujawinski, E.B., Stubbins, A., Fatland, R., Aluwihare, L.I., Buchan, A., et al. (2016) Deciphering ocean carbon in a changing world. *PNAS* **113**: 3143–3151.

Moran, M.A., Satinsky, B., Gifford, S.M., Luo, H., Rivers, A., Chan, L., et al. (2013) Sizing up metatranscriptomics. *The ISME Journal* **7**: 237–43.

Morán, X.A.G., Ducklow, H.W., and Erickson, M. (2013) Carbon fluxes through estuarine bacteria reflect coupling with phytoplankton. **489**: 75–85.

Morán, X.A.G., Gasol, J.M., Pernice, M.C., Mangot, J.-F., Massana, R., Lara, E., et al. (2017) Temperature regulation of marine heterotrophic prokaryotes increases latitudinally as a breach between bottom-up and top-down controls. *Global Change Biology* **23**: 3956–3964.
Morel, F.M.M. and Price, N.M. (2003) The biogeochemical cycles of trace metals in the oceans. *Science* **300**: 944–947.
Morris, R.M., Nunn, B.L., Frazar, C., Goodlett, D.R., Ting, Y.S., and Rocap, G. (2010) Comparative metaproteomics reveals ocean-scale shifts in microbial nutrient utilization and energy transduction. *ISME J* **4**: 673–685.
Myklestad, S.M. (2000) Dissolved Organic Carbon from Phytoplankton. In, Wangersky,P.J. (ed), *Marine Chemistry*, The Handbook of Environmental Chemistry. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 111–148.
Nagata, T. (2000) Production mechanisms of dissolved organic matter. *Wiley Series in Ecological and Applied Microbiology*.
Obernosterer, I., Catala, P., Lebaron, P., and West, N.J. (2011) Distinct bacterial groups contribute to carbon cycling during a naturally iron fertilized phytoplankton bloom in the Southern Ocean. *Limnology and Oceanography* **56**: 2391–2401.
Obernosterer, I., Fourquez, M., and Blain, S. (2015) Fe and C co-limitation of heterotrophic bacteria in the naturally fertilized region off the Kerguelen Islands. *Biogeosciences* **12**: 1983–1992.
Obernosterer, I. and Herndl, G. (1995) Phytoplankton extracellular release and bacterial growth: dependence on the inorganic N: P ratio. *Marine Ecology Progress Series* **116**: 247–257.
Oexle, H., Gnaiger, E., and Weiss, G. (1999) Iron-dependent changes in cellular energy metabolism: influence on citric acid cycle and oxidative phosphorylation. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1413**: 99–107.
Olierook, H.K.H., Jourdan, F., Merle, R.E., Timms, N.E., Kusznir, N., and Muhling, J.R. (2016) Bunbury Basalt: Gondwana breakup products or earliest vestiges of the Kerguelen mantle plume? *Earth and Planetary Science Letters* **440**: 20–32.
Orsi, A.H., Whitworth, T., and Nowlin, W.D. (1995) On the meridional extent and fronts of the Antarctic Circumpolar Current. *Deep Sea Research Part I: Oceanographic Research Papers* **42**: 641–673.
Parada, A.E., Needham, D.M., and Fuhrman, J.A. (2016) Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental Microbiology* **18**: 1403–1414.
Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarshewski, A., Chaumeil, P.-A., and Hugenholtz, P. (2018) A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nature Biotechnology* **36**: 996–1004.
Pauthenet, E., Roquet, F., Madec, G., Guinet, C., Hindell, M., McMahon, C.R., et al. (2018) Seasonal Meandering of the Polar Front Upstream of the Kerguelen Plateau. *Geophysical Research Letters* **45**: 9774–9781.
Pernice, M.C., Giner, C.R., Logares, R., Perera-Bel, J., Acinas, S.G., Duarte, C.M., et al. (2016) Large variability of bathypelagic microbial eukaryotic communities across the world's oceans. *The ISME Journal* **10**: 945–958.
Pitchford, J. (1999) Iron limitation, grazing pressure and oceanic high nutrient-low chlorophyll (HNLC) regions. *Journal of Plankton Research* **21**: 525–547.
Pollard, R.T., Salter, I., Sanders, R.J., Lucas, M.I., Moore, C.M., Mills, R.A., et al. (2009) Southern Ocean deep-water carbon export enhanced by natural iron fertilization. *Nature* **457**: 577–580.
POMEROY, L.R., leB. WILLIAMS, P.J., AZAM, F., and HOBBIE, J.E. (2007) The Microbial Loop. *Oceanography* **20**: 28–33.
Poretsky, R.S., Bano, N., Buchan, A., LeCleir, G., Kleikemper, J., Pickering, M., et al. (2005) Analysis of Microbial Gene Transcripts in Environmental Samples. *Appl Environ Microbiol* **71**: 4121–4126.
Poretsky, R.S., Hewson, I., Sun, S., Allen, A.E., Zehr, J.P., and Moran, M.A. (2009) Comparative day/night metatranscriptomic analysis of microbial communities in the North Pacific subtropical gyre. *Environmental Microbiology* **11**: 1358–1375.
Quéguiner, B. (2013) Iron fertilization and the structure of planktonic communities in high nutrient regions of the Southern Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography* **90**: 43–54.
Quéroué, F., Sarthou, G., Planquette, H.F., Bucciarelli, E., Chever, F., van der Merwe, P., et al. (2015) High variability in dissolved iron concentrations in the vicinity of the Kerguelen Islands (Southern Ocean). *Biogeosciences* **12**: 3869–3883.
Quince, C., Walker, A.W., Simpson, J.T., Loman, N.J., and Segata, N. (2017) Shotgun metagenomics, from sampling to analysis. *Nature Biotechnology* **35**: 833–844.

R. Rintoul, S., W. Hughes, C., and Olbers, D. (2001) Chapter 4.6 The antarctic circumpolar current system. In, Siedler,G., Church,J., and Gould,J. (eds), *International Geophysics*, Ocean Circulation and Climate. Academic Press, pp. 271–XXXVI.

Rappé, M.S. and Giovannoni, S.J. (2003) The Uncultured Microbial Majority. *Annual Review of Microbiology* **57**: 369–394.

REDFIELD, A.C. (1958) THE BIOLOGICAL CONTROL OF CHEMICAL FACTORS IN THE ENVIRONMENT. *American Scientist* **46**: 230A–221.

Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N.N., Anderson, I.J., Cheng, J.-F., et al. (2013) Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **499**: 431–437.

Romano, S., Dittmar, T., Bondarev, V., Weber, R.J.M., Viant, M.R., and Schulz-Vogt, H.N. (2014) Exo-Metabolome of Pseudovibrio sp. FO-BEG1 Analyzed by Ultra-High Resolution Mass Spectrometry and the Effect of Phosphate Limitation. *PLOS ONE* **9**: e96038.

Roux, S., Brum, J.R., Dutilh, B.E., Sunagawa, S., Duhaime, M.B., Loy, A., et al. (2016) Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. *Nature* **537**: 689–693.

Rue, E.L. and Bruland, K.W. (1995) Complexation of iron(III) by natural organic ligands in the Central North Pacific as determined by a new competitive ligand equilibration/adsorptive cathodic stripping voltammetric method. *Marine Chemistry* **50**: 117–138.

Ruud, J.T. (1930) Nitrates and Phosphates in the Southern Seas. *ICES J Mar Sci* **5**: 347–360.

Sabine, C.L., Feely, R.A., Gruber, N., Key, R.M., Lee, K., Bullister, J.L., et al. (2004) The Oceanic Sink for Anthropogenic CO2. *Science* **305**: 367–371.

Saito, M.A., Bertrand, E.M., Duffy, M.E., Gaylord, D.A., Held, N.A., Hervey, W.J., et al. (2019) Progress and Challenges in Ocean Metaproteomics and Proposed Best Practices for Data Sharing. *Journal of Proteome Research* **18**: 1461–1476.

Saito, M.A., Breier, C., Jakuba, M., McIlvin, M., and Moran, D. (2017) Envisioning a Chemical Metaproteomics Capability for Biochemical Research and Diagnosis of Global Ocean Microbiomes, National Academies Press (US).

Saito, M.A., Dorsk, A., Post, A.F., McIlvin, M.R., Rappé, M.S., Ditullio, G.R., and Moran, D.M. (2015) Needles in the blue sea: sub-species specificity in targeted protein biomarker analyses within the vast oceanic microbial metaproteome. *Proteomics* **15**: 3521–3531.

Saito, M.A., Goepfert, T.J., and Ritt, J.T. (2008) Some thoughts on the concept of colimitation: Three definitions and the importance of bioavailability. *Limnology and Oceanography* **53**: 276–290.

Saito, M.A., McIlvin, M.R., Moran, D.M., Goepfert, T.J., DiTullio, G.R., Post, A.F., and Lamborg, C.H. (2014) Multiple nutrient stresses at intersecting Pacific Ocean biomes detected by protein biomarkers. *Science* **345**: 1173–1177.

Sañudo-Wilhelmy, S.A., Kustka, A.B., Gobler, C.J., Hutchins, D.A., Yang, M., Lwiza, K., et al. (2001) Phosphorus limitation of nitrogen fixation by Trichodesmium in the central Atlantic Ocean. *Nature* **411**: 66–69.

Sarmiento, J.L. and Gruber, N. (2006) Ocean Biogeochemical Dynamics, Princeton University Press.

Sarmiento, J.L., Gruber, N., Brzezinski, M.A., and Dunne, J.P. (2004) High-latitude controls of thermocline nutrients and low latitude biological productivity. *Nature* **427**: 56.

Sarmiento, J.L. and Orr, J.C. (1991) Three-dimensional simulations of the impact of Southern Ocean nutrient depletion on atmospheric CO2 and ocean chemistry. *Limnology and Oceanography* **36**: 1928–1950.

Satinsky, B.M., Zielinski, B.L., Doherty, M., Smith, C.B., Sharma, S., Paul, J.H., et al. (2010) The Amazon continuum dataset: quantitative metagenomic and metatranscriptomic inventories of the Amazon River plume, June 2010.

Selinger, D.W., Saxena, R.M., Cheung, K.J., Church, G.M., and Rosenow, C. (2003) Global RNA half-life analysis in Escherichia coli reveals positional patterns of transcript degradation. *Genome Res* **13**: 216–223.

Shakoury-Elizeh, M., Protchenko, O., Berger, A., Cox, J., Gable, K., Dunn, T.M., et al. (2010) Metabolic response to iron deficiency in Saccharomyces cerevisiae. *J Biol Chem* **285**: 14823–14833.

Sharon, I. and Banfield, J.F. (2013) Genomes from Metagenomics. *Science* **342**: 1057–1058.

Simon, C. and Daniel, R. (2009) Achievements and new knowledge unraveled by metagenomic approaches. *Appl Microbiol Biotechnol* **85**: 265–276.

Simon, C. and Daniel, R. (2011) Metagenomic Analyses: Past and Future Trends. *Appl Environ Microbiol* **77**: 1153–1161.

Singer, E., Bushnell, B., Coleman-Derr, D., Bowman, B., Bowers, R.M., Levy, A., et al. (2016) High-resolution phylogenetic microbial community profiling. *The ISME Journal* **10**: 2020–2032.

Smetacek, V., Klaas, C., Strass, V.H., Assmy, P., Montresor, M., Cisewski, B., et al. (2012) Deep carbon export from a Southern Ocean iron-fertilized diatom bloom. *Nature* **487**: 313–319.
Sogin, M.L., Morrison, H.G., Huber, J.A., Welch, D.M., Huse, S.M., Neal, P.R., et al. (2006) Microbial diversity in the deep sea and the underexplored "rare biosphere." *PNAS* **103**: 12115–12120.
Sohm, J.A., Webb, E.A., and Capone, D.G. (2011) Emerging patterns of marine nitrogen fixation. *Nat Rev Microbiol* **9**: 499–508.
Sokolov, S. and Rintoul, S.R. (2009) Circumpolar structure and distribution of the Antarctic Circumpolar Current fronts: 1. Mean circumpolar paths. *Journal of Geophysical Research: Oceans*.
Sokolov, S. and Rintoul, S.R. (2002) Structure of Southern Ocean fronts at 140°E. *Journal of Marine Systems* **37**: 151–184.
Sowell, S.M., Wilhelm, L.J., Norbeck, A.D., Lipton, M.S., Nicora, C.D., Barofsky, D.F., et al. (2009) Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *ISME J* **3**: 93–105.
Spang, A. and Offre, P. (2019) Towards a systematic understanding of differences between archaeal and bacterial diversity. *Environmental Microbiology Reports* **11**: 9–12.
Stewart, R.D., Auffret, M.D., Snelling, T.J., Roehe, R., and Watson, M. (2019) MAGpy: a reproducible pipeline for the downstream analysis of metagenome-assembled genomes (MAGs). *Bioinformatics* **35**: 2150–2152.
Stoddard, S.F., Smith, B.J., Hein, R., Roller, B.R.K., and Schmidt, T.M. (2015) rrnDB: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. *Nucleic Acids Research* **43**: D593–D598.
Stoderegger, K. and Herndl, G.J. (1998) Production and release of bacterial capsular material and its subsequent utilization by marine bacterioplankton. *Limnology and Oceanography* **43**: 877–884.
Strzepek, R.F., Hunter, K.A., Frew, R.D., Harrison, P.J., and Boyd, P.W. (2012) Iron-light interactions differ in Southern Ocean phytoplankton. *Limnology and Oceanography* **57**: 1182–1200.
Sunagawa, S., Coelho, L.P., Chaffron, S., Kultima, J.R., Labadie, K., Salazar, G., et al. (2015) Structure and function of the global ocean microbiome. *Science* **348**: 1261359.
Sunda, W.G. and Huntsman, S.A. (1995) Iron uptake and growth limitation in oceanic and coastal phytoplankton. *Marine Chemistry* **50**: 189–206.
Tagliabue, A., Bopp, L., Dutay, J.-C., Bowie, A.R., Chever, F., Jean-Baptiste, P., et al. (2010) Hydrothermal contribution to the oceanic dissolved iron inventory. *Nature Geoscience* **3**: 252–256.
Tagliabue, A., Bowie, A.R., Philip, W., Buck, K.N., Johnson, K.S., and Saito, M.A. (2017) The integral role of iron in ocean biogeochemistry. *Nature* **543**: In Press.
Takeda, S. (1998) Influence of iron availability on nutrient consumption ratio of diatoms in oceanic waters. *Nature* **393**: 774.
Taniguchi, Y., Choi, P.J., Li, G.-W., Chen, H., Babu, M., Hearn, J., et al. (2010) Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. *Science* **329**: 533–538.
Tortell, P.D., Maldonado, M.T., Granger, J., and Price, N.M. (1999) Marine bacteria and biogeochemical cycling of iron in the oceans. *FEMS Microbiology Ecology* **29**: 1–11.
Tortell, P.D., Maldonado, M.T., and Price, N.M. (1996) The role of heterotrophic bacteria in iron-limited ocean ecosystems. *Nature* **383**: 330–332.
Tremblay, L., Caparros, J., LEBLANC, K., and Obernosterer, I. (2015) Origin and fate of particulate and dissolved organic matter in a naturally iron-fertilized region of the Southern Ocean. *Biogeosciences* **12**: 607–621.
Trick, C.G. (1989) Hydroxamate-siderophore production and utilization by marine eubacteria. *Current Microbiology* **18**: 375–378.
Vargas, C. de, Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., et al. (2015) Eukaryotic plankton diversity in the sunlit ocean. *Science* **348**: 1261605.
Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., et al. (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**: 66–74.
Volk, T. and Hoffert, M.I. (2013) Ocean Carbon Pumps: Analysis of Relative Strengths and Efficiencies in Ocean-Driven Atmospheric CO2 Changes. In, *The Carbon Cycle and Atmospheric CO2: Natural Variations Archean to Present*. American Geophysical Union (AGU), pp. 99–110.
Vollmers, J., Wiegand, S., and Kaster, A.-K. (2017) Comparing and Evaluating Metagenome Assembly Tools from a Microbiologist's Perspective - Not Only Size Matters! *PLOS ONE* **12**: e0169662.
Vraspir, J.M. and Butler, A. (2009) Chemistry of marine ligands and siderophores. *Annual review of marine science* **1**: 43–63.

Walters, W., Hyde, E.R., Berg-Lyons, D., Ackermann, G., Humphrey, G., Parada, A., et al. (2016) Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. *mSystems* **1**: e00009-15.
Walworth, N.G., Fu, F.-X., Webb, E.A., Saito, M.A., Moran, D., Mcllvin, M.R., et al. (2016) Mechanisms of increased Trichodesmium fitness under iron and phosphorus co-limitation in the present and future ocean. *Nat Commun* **7**:.
Weber, T.S. and Deutsch, C. (2010) Ocean nutrient ratios governed by plankton biogeography. *Nature* **467**: 550–554.
West, N.J., Obernosterer, I., Zemb, O., and Lebaron, P. (2008) Major differences of bacterial diversity and activity inside and outside of a natural iron-fertilized phytoplankton bloom in the Southern Ocean. *Environmental Microbiology* **10**: 738–756.
Whitman, W.B., Coleman, D.C., and Wiebe, W.J. (1998) Prokaryotes: The unseen majority. *PNAS* **95**: 6578–6583.
Wienhausen, G., Noriega-Ortega, B.E., Niggemann, J., Dittmar, T., and Simon, M. (2017) The Exometabolome of Two Model Strains of the Roseobacter Group: A Marketplace of Microbial Metabolites. *Front Microbiol* **8**:.
Wilhelm, S.W. and Trick, C.G. (1994) Iron-limited growth of cyanobacteria: Multiple siderophore production is a common response. *Limnology and Oceanography* **39**: 1979–1984.
Wilkins, D., Lauro, F.M., Williams, T.J., Demaere, M.Z., Brown, M.V., Hoffman, J.M., et al. (2013) Biogeographic partitioning of Southern Ocean microorganisms revealed by metagenomics. *Environmental Microbiology* **15**: 1318–1333.
Wilkins, D., Yau, S., Williams, T.J., Allen, M.A., Brown, M.V., DeMaere, M.Z., et al. (2013) Key microbial drivers in Antarctic aquatic environments. *FEMS Microbiol Rev* **37**: 303–335.
Williams, T.J., Long, E., Evans, F., Demaere, M.Z., Lauro, F.M., Raftery, M.J., et al. (2012) ORIGINAL ARTICLE A metaproteomic assessment of winter and summer bacterioplankton from Antarctic Peninsula coastal surface waters. *The ISME Journal* **6**: 1883–1900.
Williams, T.J., Wilkins, D., Long, E., Evans, F., Demaere, M.Z., Raftery, M.J., and Cavicchioli, R. (2013) The role of planktonic Flavobacteria in processing algal organic matter in coastal East Antarctica revealed using metagenomics and metaproteomics. *Environmental Microbiology* **15**: 1302–1317.
Winkelmann, G. (1991) CRC handbook of microbial iron chelates, Boca Raton : CRC Press.
Yooseph, S., Sutton, G., Rusch, D.B., Halpern, A.L., Williamson, S.J., Remington, K., et al. (2007) The Sorcerer II Global Ocean Sampling Expedition: Expanding the Universe of Protein Families. *PLoS Biology* **5**: e16.
Zeebe, R.E. and Archer, D. (2005) Feasibility of ocean fertilization and its impact on future atmospheric CO2 levels. *Geophysical Research Letters* **32**:.

Online references

Figure 2:
http://omix.aori.u-tokyo.ac.jp/en/public-research/h28-h29-member/kondo-yoshiko/

Chapter 1

Microbial iron metabolism as revealed by gene expression profiles in contrasted Southern Ocean regimes

Pavla Debeljak[1,2], Eve Toulza[3], Sara Beier[4], Stéphane Blain[1] and Ingrid Obernosterer[1]

[1] Sorbonne Université, CNRS, Laboratoire d'Océanographie Microbienne, LOMIC, F-66650 Banyuls/mer, France

[2] Department of Limnology and Bio-Oceanography, University of Vienna, A-1090 Vienna, Austria

[3] Université Perpignan Via Domitia, IHPE UMR 5244, CNRS, IFREMER, Univ. Montpellier, F-66860 Perpignan, France.

[4] Leibniz Institute for Baltic Sea Research, Warnemünde, Germany

environmental microbiology

Environmental Microbiology (2019) **00**(00), 00–00 doi:10.1111/1462-2920.14621

Microbial iron metabolism as revealed by gene expression profiles in contrasted Southern Ocean regimes

Pavla Debeljak,[1,2*] **Eve Toulza**,[3] **Sara Beier**,[4] **Stephane Blain**[1] **and Ingrid Obernosterer**[1]

[1]*Sorbonne Université, CNRS, Laboratoire d'Océanographie Microbienne, LOMIC, F-66650 Banyuls/mer, France.*

[2]*Department of Limnology and Bio-Oceanography, University of Vienna, A-1090, Vienna, Austria.*

[3]*Université Perpignan Via Domitia, IHPE UMR 5244, CNRS, IFREMER, Univ. Montpellier, F-66860, Perpignan, France.*

[4]*Leibniz Institute for Baltic Sea Research, Warnemünde, Germany.*

Summary

Iron (Fe) is a limiting nutrient in large regions of the ocean, but the strategies of prokaryotes to cope with this micronutrient are poorly known. Using a gene-specific approach from metatranscriptomics data, we investigated seven Fe-related metabolic pathways in microbial communities from high nutrient low chlorophyll and naturally Fe-fertilized waters in the Southern Ocean. We observed major differences in the contribution of prokaryotic groups at different taxonomic levels to transcripts encoding Fe-uptake mechanisms, intracellular Fe storage and replacement and Fe-related pathways in the tricarboxylic acid (TCA) cycle. The composition of the prokaryotic communities contributing to the transcripts of a given Fe-related pathway was overall independent of the *in situ* Fe supply, indicating that microbial taxa utilize distinct Fe-related metabolic processes. Only a few prokaryotic groups contributed to the transcripts of more than one Fe-uptake mechanism, suggesting limited metabolic versatility. Taxa-specific expression of individual genes varied among prokaryotic groups and was substantially higher for all inspected genes in Fe-limited as compared to naturally fertilized waters, indicating the link between transcriptional state and Fe regime. Different metabolic strategies regarding low Fe concentrations in the Southern Ocean are discussed for two abundant prokaryotic groups, *Pelagibacteraceae* and *Flavobacteriaceae*.

Received 14 December, 2018; revised 18 March, 2019; accepted 1 April, 2019. *For correspondence. E-mail pavla.debeljak@gmail.com. Tel. +33468887354; Fax +33468887301

Introduction

Since John Martin's 'iron hypothesis' was introduced in the late 1980s to solve the high nutrient low chlorophyll (HNLC) paradox in the ocean, the micronutrient iron (Fe) has been recognized as a major factor in the regulation of ocean primary productivity (Martin, 1990; Tagliabue *et al*., 2017). The Southern Ocean as the largest HNLC area has been subject to multiple mesoscale artificial Fe fertilization studies focusing on enhanced phytoplankton blooms through Fe input (reviewed in the study by Boyd *et al*., 2007). Together with investigations in naturally Fe-fertilized regions (Blain *et al*., 2007; Pollard *et al*., 2009), the control by Fe of primary productivity and subsequent carbon dioxide (CO_2) drawdown in this ocean has been confirmed.

Heterotrophic microorganisms rapidly respond to phytoplankton blooms induced by Fe-fertilization (Cochlan, 2001; Hall and Safi, 2001; Oliver *et al*., 2004; Obernosterer *et al*., 2008) and they remineralize a substantial fraction of phytoplankton-derived dissolved organic matter (Christaki *et al*., 2014). Fe is essential for microbial heterotrophic metabolism, the access to this micronutrient by various taxa will therefore affect the processing of organic carbon. The limited number of measurements indicates that heterotrophic prokaryotes have cellular Fe quotas that are similar or higher than those of phytoplankton (Tortell *et al*., 1999; Sarthou *et al*., 2008; Fourquez *et al*., 2012). The majority of Fe (> 90%) in heterotrophic prokaryotic cells is located in the respiratory chain (Andrews *et al*., 2003) and as a consequence Fe limitation results in a prokaryotic reduction in prokaryotic respiration and growth rates (Tortell *et al*., 1999; Smith *et al*., 2010; Fourquez *et al*., 2014; Koedooder *et al*., 2018). Experimental studies testing the effect of Fe on natural prokaryotic communities have revealed both positive and negative bulk metabolic responses

(summarized in the study by Obernosterer *et al*., 2015), likely reflecting temporal and spatial variability of the bioavailability and the cellular requirements of this micronutrient. In the region off Kerguelen Island, heterotrophic prokaryotic growth and production were limited by Fe and organic carbon in early spring (Obernosterer *et al*., 2015), leading to competition between heterotrophic and phototrophic microorganisms for this micronutrient (Fourquez *et al*., 2015).

These observations raise the question of the mechanisms used by microbial taxa to acquire and to metabolize this micronutrient in cellular processes. An increasing number of prokaryotic genomes and of metagenomes originating from global ocean surveys has provided insights to the inventories of Fe-related pathways (Desai *et al*., 2012; Hopkinson and Barbeau, 2012; Toulza *et al*., 2012; Hogle *et al*., 2016). These studies have shown that the genomic potential for Fe-uptake mechanisms varies among prokaryotic taxa (Hopkinson and Barbeau, 2012; Hogle *et al*., 2016) and that the prevalence of Fe-related pathways in prokaryotes reflects Fe concentrations across ocean regions (Toulza *et al*., 2012). The single-cell approach MICRO-CARD-FISH using ^{55}Fe revealed that the community taking up Fe in the Southern Ocean was dominated by *Gammaproteobacteria* and FCB, while SAR11 and *Roseobacter* had overall lower contributions (Fourquez *et al*., 2016). Fe limitation has been shown to induce the glyoxylate shunt in heterotrophic bacterial model organisms (Smith *et al*., 2010; Fourquez *et al*., 2014; Koedooder *et al*., 2018), a pattern that was also observed for SAR11 in the HNLC Southern Ocean (Beier *et al*., 2015). The glyoxylate shunt bypasses two decarboxylation steps and the coupled release of CO_2 and reducing equivalents ($NADH_2$) of the TCA cycle, with important consequences on ATP production and processing of organic carbon (Koedooder *et al*., 2018).

The aim of the present study was to extend these observations, by providing a detailed picture on the expression of genes responsible for Fe-uptake and Fe-related downstream processes in Southern Ocean natural prokaryotic communities. We screened the total mRNA from metatranscriptomes against a database that contained 10,411 protein sequences corresponding to seven Fe-related metabolic pathways connected to Fe transport, Fe storage and central carbon metabolism. We explored the functional expression profiles of different taxa in microbial communities originated from HNLC and naturally Fe-fertilized waters in the region off Kerguelen Island.

Results

Environmental context

The three stations considered in the present study were part of the Kerguelen Ocean and Plateau compared

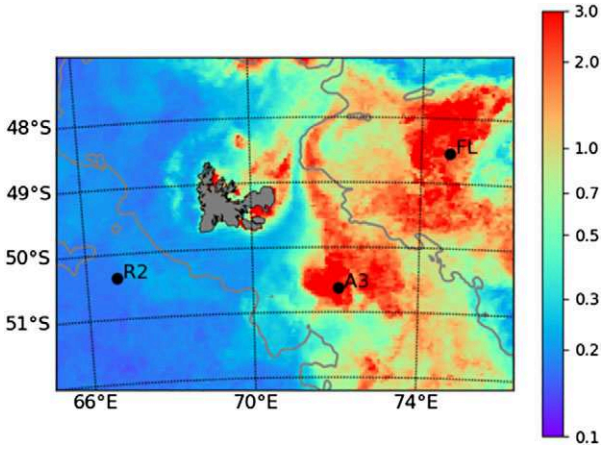


Fig. 1. Location of sampling sites positions superimposed on the monthly composite satellite image of chlorophyll (μg L^{-1}) provided by Copernicus Marine Service for (November 2011, 4 × 4 km). The grey line denotes 1000 m isobaths.

study 2 (KEOPS2)-cruise that took place in November 2011. Station (R-2) was located in HNLC waters west of Kerguelen Island (Fig. 1), and two stations were situated east of the island in naturally Fe-fertilized waters south (Station A3-2) and north (Station F-L) of the polar front. Concentrations of dissolved Fe in the surface mixed layer varied between 0.13 ± 0.05 nM and 0.22 ± 0.06 nM (Quéroué *et al*., 2015) (Table 1). Chlorophyll *a* concentrations were low at station R-2 (Chl *a*, 0.25 ± 0.08 μg L^{-1}) and up to 16-fold higher in naturally fertilized waters (Lasbleiz *et al*., 2014) (Table 1). The abundance, production and respiration of heterotrophic prokaryotes were several fold enhanced at the fertilized sites as compared with HNLC waters. Concentrations of DOC, DON and DOP were not enhanced at the Fe-fertilized sites, probably due to the rapid consumption of organic matter by heterotrophic prokaryotes. Despite these pronounced differences in biomass and production, heterotrophic prokaryotes were shown to be limited by Fe at the HNLC-site and in Fe-fertilized waters (Obernosterer *et al*. 2015).

Contribution of prokaryotic groups to specific gene expression

Filtering and rRNA removal of raw reads resulted in a total of 29–36 Million reads per sample (Supporting Information Table S1). The prokaryotic proportions were calculated for each duplicate based on Blastx results and ranged from 9.7% ± 1.1 at Station R-2, 14.5% ± 1.0 at A3-2%, to 19.9% ± 2.9 at F-L (Supporting Information Fig. S1). The retrieved sequences cover a broad range of families out of 30 bacterial phyla as well as archaea. Highest overall pathway specific contributions were observed for the classes of *Alpha-, Beta-* and *Gammaproteobacteria* as well as the *Fibrobacteres, Chlorobi, Bacteroidetes* (FCB) (Figs. 2 and 3,

Table 1. Location, date, biogeochemical properties and bulk prokaryotic parameters at the three study sites. All parameters are mean ± SD for the surface mixed layer.

	R-2	F-L	A3-2
Latitude S	50.3590	48.5222	51.0333
Longitude E	66.7170	74.6500	72.0833
Date of sampling	26 October 2011	7 November 2011	16 November 2011
Sampling depth (m)	60	20	20
Surface mixed layer (m)	105 ± 15	38 ± 7	153 ± 15
Dissolved and particulate nutrients			
DOC (μM)[a]	47.8 ± 0.4	49.6 ± 1.3	51.3 ± 1.5
DON (μM)[b]	6.1 ± 0.04	5.47 ± 1.33	6.44 ± 2.2
DOP (μM)[b]	0.3 ± 0.02	0.26 ± 0.12	0.36 ± 0.04
POC (μM)[a]	6.5 ± 1.8	11.5 ± 1.2	13.5 ± 1.8
DFe (nmol L^{-1})[c]	0.13 ± 0.05	0.22 ± 0.06	0.16 ± 0.03
Chlorophyll a (μg L^{-1})[d]	0.25 ± 0.08	4.00 ± 1.58	2.03 ± 0.34
Prokaryotic abundance ($\times 10^5$ cells mL^{-1})[e]	2.72 ± 0.3	6.06[c,f]	3.16 ± 0.5
Prokaryotic production (ng C L^{-1} h^{-1})[e]	2.59 ± 0.53	65.7 ± 1.62	19.9 ± 3.4
Prokaryotic respiration (μmol O_2 L^{-1} d^{-1})[e]	0.25 ± 0.12	1.37 ± 0.64	0.63 ± 0.45

a. From Tremblay *et al*. (2015).
b. From Blain *et al*. (2015).
c. From Queroue *et al*. (2016).
d. From Lasbleiz *et al*. (2014).
e. From Christaki *et al*. (2014).
f. Only one measurement for the surface mixed layer available.

Supporting Information Figs. S2 and S3, Supporting Information Table S2). These bacterial phyla were shown to be abundant also in data sets derived from 16S rRNA amplicon sequencing (Fig. 4). We observed major differences in the contribution of prokaryotic groups to the different Fe-uptake mechanisms and this pattern was largely independent of site. Siderophore-uptake was dominated by *Gammaproteobacteria* (43% of total transcripts) and FCB (25%), while Fe^{2+}- and Fe^{3+}-uptake revealed an increased contribution of alphaproteobacterial groups, most pronounced for Fe^{3+}-uptake (29%). Within *Gammaproteobacteria* noticeable differences in the contribution of different phylotypes to the three uptake mechanisms could be observed. While *Alteromonadaceae*, *Cellvibrionaceae* and *Shewanellaceae* were abundant contributors to siderophore-uptake, *Shewanellaceae* and *Enterobacteriaceae*, accounted for most gammaproteobacterial Fe^{2+}-uptake transcripts and *Pseudomonadaceae, Chromatiaceae* and *Piscirickettsiaceae* were major contributors to Fe^{3+}-uptake transcripts (Fig. 3, Supporting Information Table S2). FCB were mostly represented by *Flavobacteriaceae* for siderophore- and Fe^{2+}-uptake, and *Bacillaceae* contributed additionally to Fe^{3+}-uptake. Within *Alphaproteobacteria, Rhodobacteracea* substantially contributed to Fe^{2+}- and Fe^{3+}-uptake, but this group had low siderophore uptake transcripts (<1.6% of alphaproteobacterial transcripts). A contrasting pattern was observed for *Sphingomonadaceae* and *Erythrobacteraceae* that contributed to siderophore uptake transcripts but neither Fe^{2+} nor Fe^{3+}. *Pelagibacteraceae* transcripts were not detectable for siderophore- and Fe^{2+}-uptake, and Fe^{3+}-uptake transcripts belonging to this group accounted for 0.8% of alphaproteobacterial transcripts. *Actinobacteriaceae* and *Archaea* were almost absent from siderophore-uptake but contributed to Fe^{2+}- and Fe^{3+}-uptake.

The flavodoxin switch and bacterioferritin transcripts revealed both high contributions of gammaproteobacterial groups, and smaller proportions of alphaproteobacterial and FCB transcripts, in particular for flavodoxin transcripts. Pronounced differences in the contribution of different gammaproteobacterial phylotypes to the two Fe-related processes were detectable. *Shewanellaceae, Alcanivoracaceae* and *Aeromonadaceae* were the most important contributors to flavodoxin switch transcripts, while *Cellvibrionaceae, Halieaceae* and *Alteromonadaceae* dominated the bacterioferritin transcripts. *Flavobacteriaceae* and *Rhodobacteraceae* contributed each substantially to bacterioferritin transcripts, while flavodoxin switch transcripts belonging to these groups were almost absent. *Pelagibacterales* did not contribute to flavodoxin switch and bacterioferritin transcripts. *Cyanobacteria* accounted for a large portion of flavodoxin switch transcripts, in particular at the HNLC-site. The relative contribution of cyanobacteria to 16S rRNA sequences varied between 0.37% at the HNLC and 0.41% at the Fe-fertilized stations respectively (Landa *et al*. 2016). Taken together, these results construe specific Fe-uptake and Fe-processing mechanisms for several prokaryotic groups.

We identified two categories of prokaryotic groups contributing to transcripts of aconitase and isocitrate lyase, genes associated with the TCA cycle. Alphaproteobacterial phylotypes, in particular *Pelagibacteraceae* and *Actinobacteriaceae*, each had similar contribution to aconitase and isocitrate lyase. In contrast, the proportions of

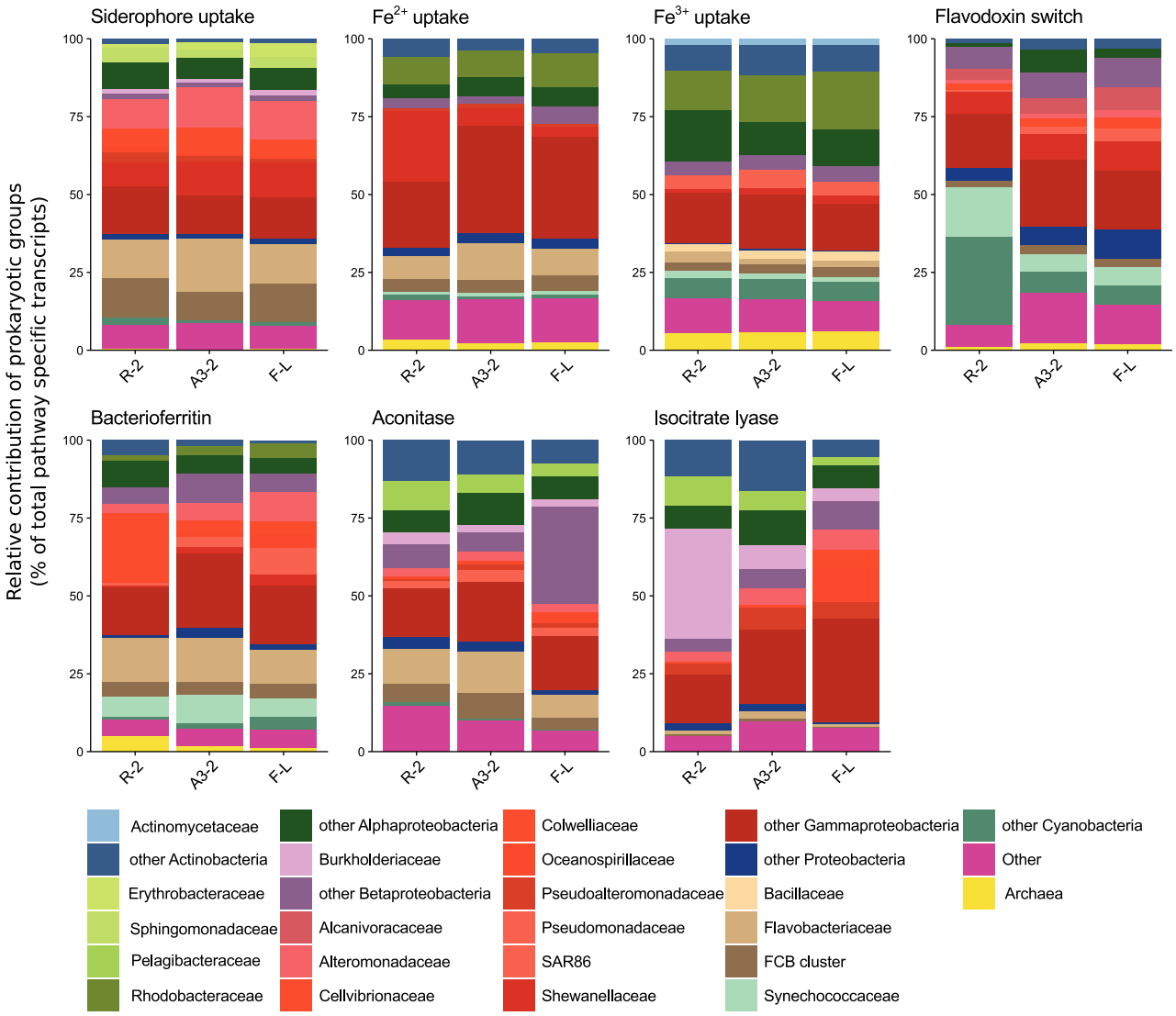


Fig. 2. Relative contribution of prokaryotic groups to pathway specific transcripts. Prokaryotic group is defined until taxonomic family level. For clarity, one replicate per station is shown, duplicates as well as mean values and error estimates are shown in the Supporting Information (Fig. S2 and Table S2).

transcripts accounted for by *Flavobacteriacaea* and gammaproteobacterial phylotypes varied considerably between the two genes. *Flavobacteriacaea* had higher contributions to aconitase transcripts than to those of isocitrate lyase, and gammaproteobacterial phylotypes, such as *Pseudoalteromonadaceae, Moraxellaceae, Oceanosprillaceae*, *Alteromonadaceae* and *Colwelliaceae* showed the opposite pattern. *Burkholderiaceae*, a group that had low contributions to all other transcripts, accounted for up to 30% of the aconitase and isocitrate lyase transcripts.

Transcriptional activity in HNLC and Fe-fertilized waters

Among site comparison of pathway-specific transcripts revealed that Fe^{2+}, Fe^{3+}, the flavodoxin switch and bacterioferritin had higher proportions at R-2 as compared to the Fe-fertilized sites (Fig. 5). In contrast, no such pattern was observed for siderophore-uptake, aconitase and isocitrate lyase. Comparison among genes encoding for the different Fe-uptake mechanisms was possible due to the similar gene length (Supporting Information Table S4). The relative transcript abundance of siderophore-uptake was roughly 6- and 2-fold higher than those of Fe^{2+}- and Fe^{3+}-uptake respectively.

In order to assess a potential per cell transcriptional activity, results from qPCR expression of isocitrate lyase in *SAR11* at two sites, R-2 and F-L, were used as correction factors (see Experimental procedures). This normalization step was chosen in order to answer the following question: Does gene expression vary between HNLC and Fe-fertilized waters? And further, how variable is cell specific expression of a given gene among the different prokaryotic

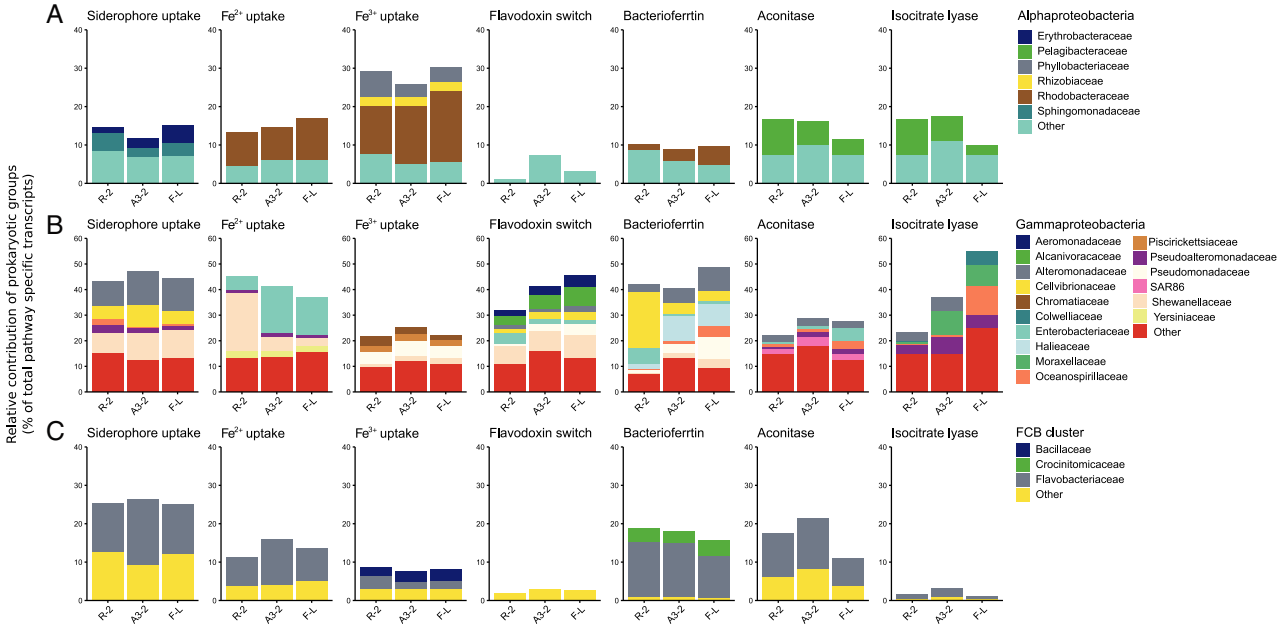


Fig. 3. Detailed view of relative contribution of Alpha (A) and Gammaproteobacteria (B) and FCB cluster (C) to pathway specific transcripts. Note the different y-axis for *Gammaproteobacteria.* For clarity, one replicate per station is shown, duplicates as well as mean values and error estimates are shown in the Supporting Information (Fig. S3 and Table S2).

groups? This step was done for representatives from the prokaryotic phyla (presented in Fig. 2) for which the gene transcripts and 16S relative abundances were available for a given phylogenetic level (Fig. 4 and Supporting Information Fig. S5). To verify the approach, ribosomal proteins, essential in cellular processes of translation, were retrieved from Kyoto Encyclopedia of Genes and Genomes (KEGG) and screened against our data sets (Fig. 6 and Supporting Information Fig. S4).

Cell-specific expression of all genes considered in the present study was considerably higher (10- to 1000-fold) at Station R-2 in comparison to F-L for all prokaryotic groups (Fig. 6 and Supporting Information Table S3). Exceptions were *Pseudomonadales* that consistently revealed an inverse pattern, and *Burkholderiales* for which differences between sites were in most cases small. In comparison, the ribosomal proteins showed the lowest fold-change (1.5) between R-2 and F-L. For clarity, the following comparison of taxa-specific expression levels among prokaryotic groups is focused on the HNLC site R-2. *Alteromonadales* had the highest taxa-specific expression levels for all genes, while taxa-specific expression levels of a given gene was more variable for the other prokaryotic groups. *Flavobacteriales*, *Oceanospirillales* and *Rhodobacterales* had similar taxa-specific siderophore-uptake expression while *Rhodobacterales* and *Actinomycetales* had substantially higher taxa-specific Fe^{3+}-uptake as compared with *Flavobacteriales* (8- to 6-fold respectively) and *Oceanospirillale*s (15- to 12-fold). *Rhodobacterales, Flavobacteriales* and *Oceanospirillale*s had similar bacterioferritin expression levels, and flavodoxin switch expression levels belonging to *Flavobacteriales* were considerably lower than those of the other groups. Taxa-specific gene expression of aconitase was similar for *Rhodobacterales, Flavobacteriales* and *Oceanospirillale*s. In contrast, taxa-specific expression levels of isocitrate lyase were highly variable among groups, with *Actinomycetales* and *Pelagibacterales* at the higher and lower ranges respectively.

Discussion

We present here taxon-specific strategies of Fe-uptake and intracellular processes dependent on Fe, a crucial, yet growth-limiting element for microbial heterotrophs in large areas of the ocean. By investigating the *in situ* expression patterns of candidate genes from metatranscriptomics data, we provide novel insights into these metabolic traits of Southern Ocean microbial communities in contrasting Fe- and C-regimes. We observe major differences in the contribution of prokaryotic groups to the pathways investigated, indicating distinct metabolic capabilities for Fe-related processes and downstream carbon metabolism for microbial taxa. Our finding that taxa-specific expression levels are substantially higher under Fe-limited conditions suggests that the transcriptional states related to Fe-uptake and Fe cell-content reduction or control are associated with the *in situ* Fe-supply.

Our observation that *gamma-* and *flavobacterial* groups accounted for a large fraction of siderophore-

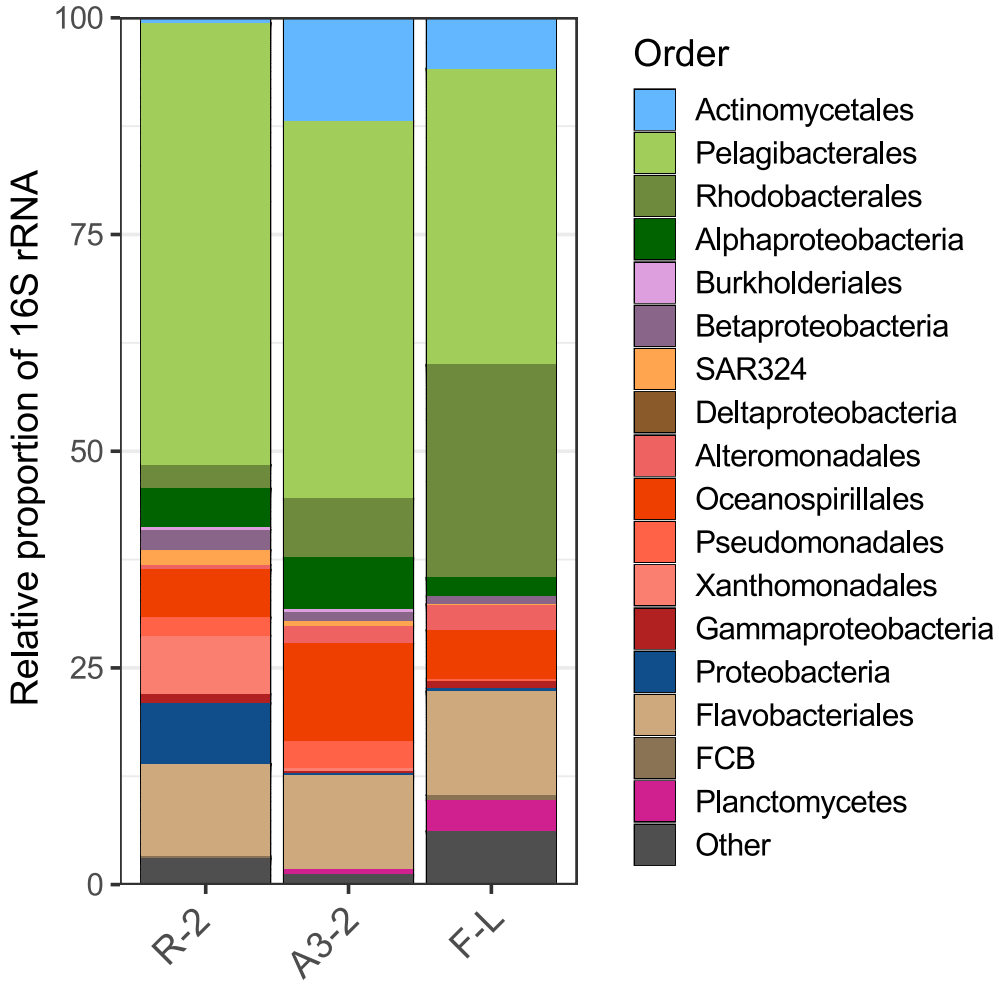


Fig. 4. Relative proportion of 16S rRNA gene sequences on the order level (data are from Landa *et al*. 2016).

uptake transcripts, while *alphaproteobacterial* groups had an increased contribution to Fe^{3+}-uptake transcripts corroborates previous findings on the genomic potential of representatives of these classes (Hopkinson and Barbeau, 2012; Tang *et al*., 2012; Hogle *et al*., 2016). A survey of 206 bacterial genomes revealed that Ton-B-dependent transporters (TBDTs), many of which are known as siderophore type transporters, are common in *Gammaproteobacteria* and *Bacteroidetes* but less abundant in *Alphaproteobacteria*, and absent in *Pelagibacter ubique* (Hopkinson and Barbeau, 2012; Tang *et al*., 2012). Within *Alphaproteobacteria*, *Erythrobacteracea* and *Sphingomonadacea* were major contributors to siderophore-uptake transcripts in the present study, an observation supported by the genomic potential of representative strains (Tang *et al*., 2012). Fe^{3+}-uptake gene expression revealed a contrasting pattern, with high contributions of *Rhodobacteraceae* and *Actinomycetales* and a minor contribution of FCB. While Fe^{3+} transporters were abundant in many of the 206 bacterial genomes investigated, only 1 of 16 *Bacteroidetes* genomes contained this type of transporter (Hopkinson and Barbeau, 2012). Despite its dependency on inorganic Fe-uptake, *Pelagibacteraceae* had a minor contribution to the overall Fe^{3+}-uptake gene expression (0.8% of total Fe^{3+} transcripts) and the lowest cell-specific Fe^{3+}-uptake (Fig. 6). The lower copy numbers of Fe^{3+} transporters in *SAR11* as compared with *Roseobacter* genomes (Hogle *et al*., 2016) and potentially lower Fe requirements could explain this observation.

The dominating microbial contributors to the transcripts of a given Fe-uptake mechanism were overall similar among sites and thus largely independent of the *in situ* Fe supply. This contrasts with the different relative contributions of the observed groups at the sites (Fig. 4, Landa *et al*., 2016) and suggests that our observations on the transcriptome level are mainly driven by the metabolic potential of the respective prokaryotic groups (see Supporting Information Fig. S5). In support of this conclusion, our results illustrate that only a few prokaryotic

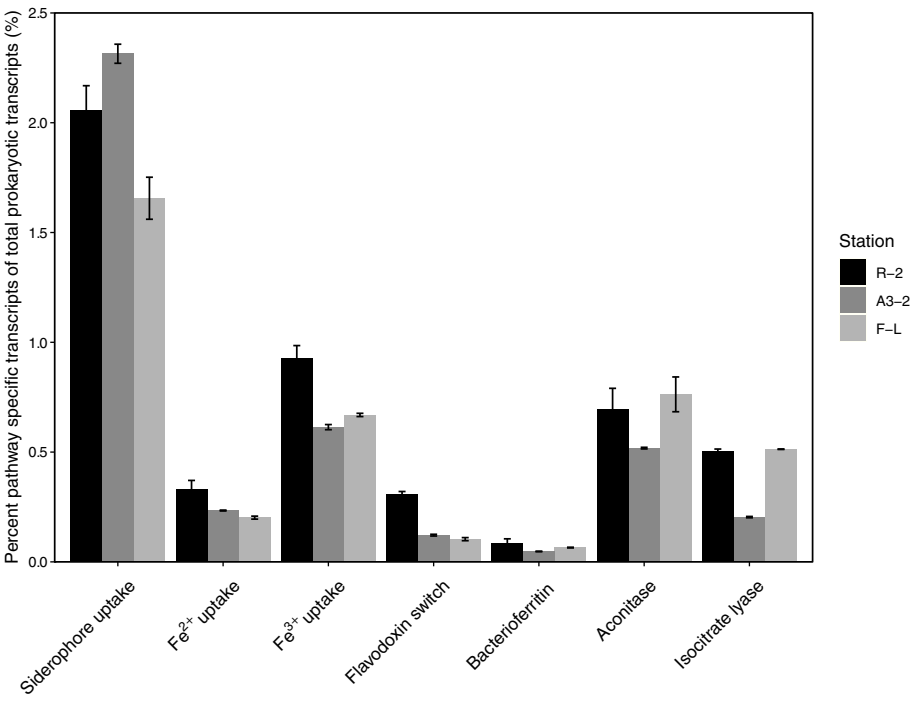


Fig. 5. Percent contribution of pathway-specific transcripts to total prokaryotic transcripts at a given site. Bars represent mean values and error bars represent minimum and maximum values.

groups contribute substantially to the transcripts of more than one Fe-uptake mechanism. This was the case for *Shewanellaceae* transcripts that were present for the three Fe-uptake mechanisms, and for *Rhodobacteraceae* and the FCB cluster that accounted both for substantial proportions of Fe^{2+}- and Fe^{3+}-uptake transcripts.

The 2- to 6-fold higher proportions of siderophore-uptake gene expression as compared with Fe^{3+} and Fe^{2+} could indicate that this mechanism was a more efficient pathway for prokaryotes to acquire Fe. Most dissolved Fe (99%) (Gledhill and van den Berg, 1994; Rue and Bruland, 1995) is complexed by organic ligands leaving extremely low steady-state concentrations of inorganic Fe^{3+} noted as Fe'. The steady-state concentrations of Fe^{2+} are also typically extremely low in oxygenated surface waters although reduction mediated by different processes could locally produce enhanced concentrations. Fe' has been considered a more bioavailable and thus more important form than siderophore-bound Fe for phytoplankton (Morel *et al*., 2008; Lis *et al*., 2015). Our observations suggest that heterotrophic prokaryotes favour the uptake of Fe bound to siderophores, which presents several advantages. First, once complexed by siderophores, Fe is hardly available for most phytoplankton (Lis *et al*., 2015), even though the idea that phytoplankton cannot directly use Fe-siderophore complexes has recently been challenged (Kazamia *et al*., 2018; McQuaid *et al*., 2018). Heterotrophic prokaryotes thereby avoid competition with phytoplankton for this scare resource. Second, siderophores can contribute to access initially not available forms of Fe, such as particulate Fe or Fe-organic complexes (Kraemer, 2004). Third, Fe-siderophore-uptake could be stimulated by siderophore-production of the same microbial cell and thus provide an advantage to certain taxa (Martinez *et al*., 2003; Hopkinson and Barbeau, 2012; Sijerčić and Price, 2015; Boiteau *et al*., 2016). Our taxa-specific transcripts point out that the expression of the genes encoding for siderophore-uptake was increased (20- to 135-fold) under Fe-limited conditions when the competition for the acquisition of this resource was highest. This latter observation agrees well with the previously reported global scale inverse relationship between siderophore-uptake gene occurrence and Fe concentrations (Toulza *et al*., 2012). Under strong Fe limitation, the Fe-uptake rate can only be increased by the number of Fe transporters as has been demonstrated for phytoplankton (Hudson and Morel, 1993). Our observations point to a similar conclusion for heterotrophic prokaryotes. Besides the capacity to optimize Fe acquisition in limited conditions, the second global strategy for prokaryotes is to decrease the Fe cellular content.

To explore differences in intracellular Fe-regulatory mechanisms among prokaryotic groups, we investigated the two genes encoding for flavodoxin and bacterioferritin. The non-Fe containing protein flavodoxin is an iso-functional protein, which can replace ferrodoxin, an electron shuttle harbouring Fe-sulfur clusters. The expression of this protein by marine autotrophic plankton was proposed as a proxy for Fe scarcity in the oceans (Roche *et al*., 1996), and insight into the

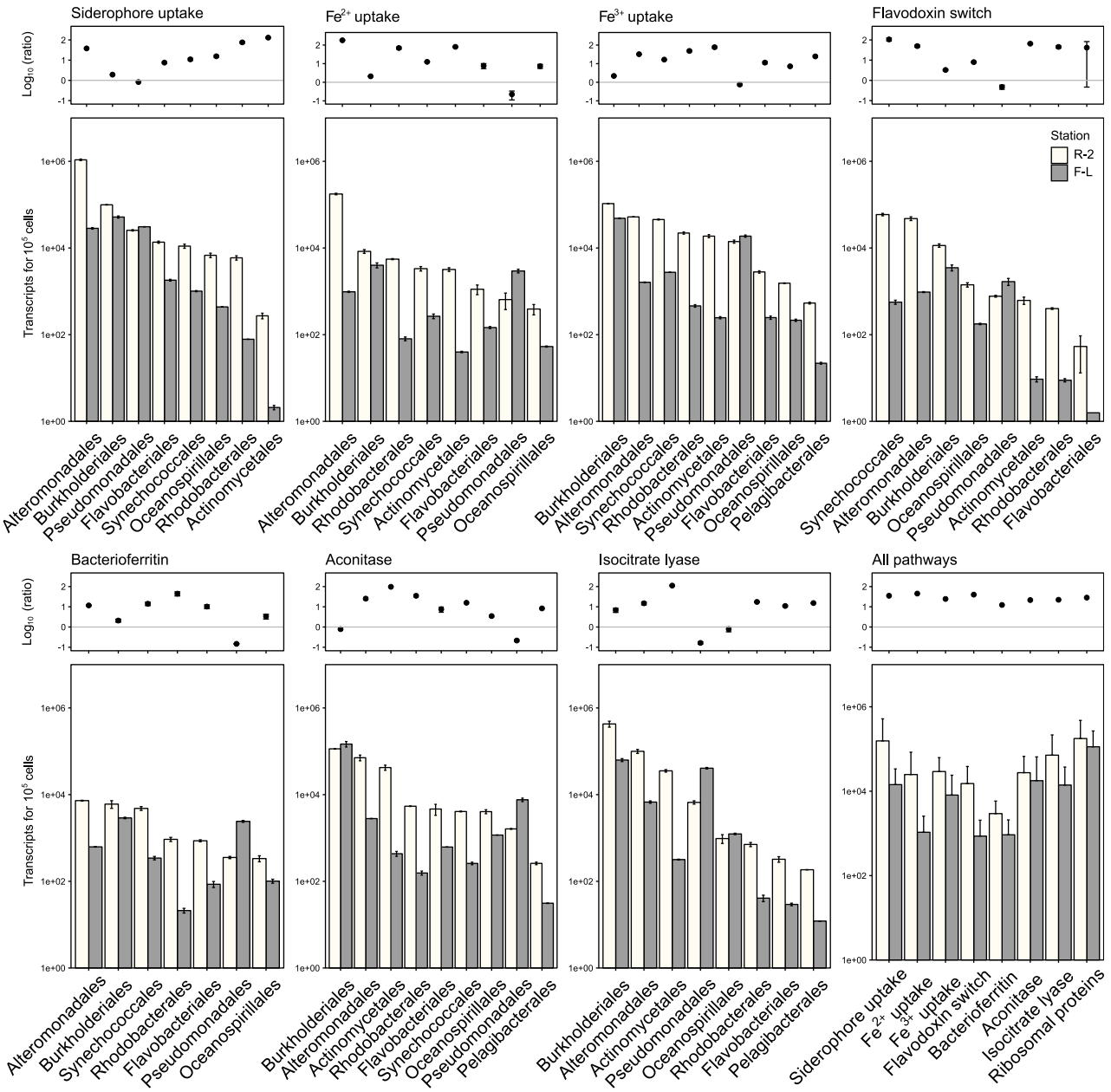


Fig. 6. Taxa-specific transcript abundance (per 10^5 cells) of a given pathway in HNLC (Station R-2, white bars) and Fe-fertilized waters (Station F-L, grey bars), and the Log_{10} of the ratio between the cell-specific transcripts at R-2 to F-L. The order of prokaryotic groups is from high to low cell-specific transcripts at Station R-2. The panel 'All pathways' shows the cell-specific transcripts for all prokaryotic groups combined and the Log_{10} of the R-2/F-L ratios for the respective pathways, including ribosomal proteins. Bars represent mean values and error bars represent minimum and maximum values. Error estimates are given in the Supporting Information (Table S3). Ratios were calculated by dividing each duplicate pair.

mechanisms of how Fe availability regulates this protein was obtained from studies on various temporal and spatial scales (Erdner *et al*., 1999; Saito *et al*., 2011; Tara Oceans Coordinators *et al*., 2018). A metagenomic analysis revealed that prokaryotes lacking this flavoprotein are confined to coastal areas where Fe supply is high, while flavodoxin-containing marine prokaryotes are preferably located in open ocean sites (Toulza *et al*., 2012). Given the roughly 35-fold higher proportion of total flavodoxin transcripts and the higher cell-specific expression in HNLC waters as compared with Fe-fertilized sites, our results extend the understanding of the regulation of this switch by Fe-availability for a large range of prokaryotic groups. In the present study, the dominant contributors were *gammaproteobacterial* taxa and Cyanobacteria, while *Alphaproteobacteria* and the FCB cluster had minor contributions to flavodoxin

transcripts. These latter groups revealed also lowest cell-specific expression patterns, suggesting that many FCB and *alphaproteobacterial* members might make use of other strategies to cope with Fe-limitation.

Ferritins are compounds that were shown to regulate the storage and the release of intracellular Fe in a number of eukaryotic microorganisms (Marchetti *et al*., 2009; Botebol *et al*., 2015). Bacterioferritins are known to be involved in the storage of Fe in *Bacteria*; however, the regulation and the exact physiological mechanisms of these compounds are not clear (Andrews *et al*., 2003; Carrondo, 2003). In the GOS data set, bacterioferritin gene abundance was higher at coastal sites with overall high Fe concentrations (Toulza *et al*., 2012). In the present study, all prokaryotic groups contributed to bacterioferritin transcripts. The differences in bacterioferritin transcripts between HNLC and Fe-fertilized waters were far less pronounced than the other Fe-related metabolisms. This could indicate that bacterioferritin-related processes are occurring at background levels, for instance, as a control of Fe homeostasis, rather than as storage of Fe in response to episodic Fe supply.

The higher cell-specific expression of siderophore-, Fe^{2+}-, Fe^{3+}-uptake, flavodoxin and bacterioferritin transcripts in HNLC as compared with Fe-fertilized waters highlight the increased investment in Fe-related metabolism when prokaryotic growth and production are limited by Fe and organic carbon (Obernosterer *et al*., 2015). But how can this affect cellular carbon metabolism? We addressed this question by investigating the two enzymes aconitase and isocitrate lyase belonging to a central metabolic pathway, the TCA cycle. Aconitase, an Fe-containing enzyme, transforms citrate to isocitrate, which can either serve as a substrate for the enzyme isocitratedehydrogenase (IDH) in the TCA cycle or as a substrate for isocitrate lyase, a non-Fe-containing enzyme that induces the glyoxylate shunt, a bypass of the TCA cycle (Supporting Information Fig. S6). While the regulation of the glyoxylate shunt can be driven by a number of factors, its induction by Fe-limitation has been demonstrated in bacterial model organisms (Fourquez *et al*., 2014; Koedooder *et al*., 2018). Using genetic tools and bioreporters demonstrated that the isocitrate lyase knock-out strains of gammaproteobacterium *Photobacterium angustum* S14 had significantly lower growth and respiration rates as compared with the wild type under Fe-limited conditions (Koedooder *et al*., 2018). Using qPCR, *SAR11* cell-specific isocitrate lyase gene expression was higher at Station R-2 as compared with F-L (Beier *et al*., 2015). The increased cell-specific isocitrate-lyase expression in Fe-limited as compared with Fe-fertilized waters observed in the present study extends this previous observation to several prokaryotic groups.

In the context of these recent findings, we focus in the following discussion on two prokaryotic groups with distinct patterns in the expression of aconitase and isocitrate lyase. *Pelagibacteraceae* contributed similarly to the expression of both genes; in contrast, *Flavobacteriaceae* revealed substantially higher contributions to aconitase (7%–13% of prokaryotic transcripts) as compared with isocitrate lyase transcripts (0.7%–2.2% of prokaryotic transcripts). In addition, *Flavobacteriaceae* had 18.5-fold higher cell-specific aconitase expression than *Pelagibacteraceae*, but both groups had similar cell-specific isocitrate lyase expression. These observations could indicate that the entire TCA cycle is more preferentially used in *Flavobacteriaceae* than in the members of *Pelagibacteraceae* with consequences on the production of NADH and ATP equivalents.

The combined information obtained by the present results and previous knowledge on characteristics of these bacterial groups leads us to propose two distinct ecological strategies with respect to Fe-related processes for *Pelagibacteraceae* and *Flavobacteriaceae.* Members of *Pelagibacteraceae* appear to be the thriftiest group, characterized by the unique use of Fe^{3+}-uptake, performed by ABC-type transporters that do not require the costly outer membrane receptors (Andrews *et al*., 2003) but do require ATP. Additionally, members of this group lack Fe storage and the flavodoxin switch for which no transcripts were detectable in the present study. Despite its dependency on Fe^{3+}-uptake, *Pelagibacteraceae* had a minor contribution to the total transcripts, suggesting low Fe requirements of this group. These characteristics extend those described previously of the most prominent representatives such as the streamlined *SAR11* (reviewed in the study by Giovannoni, 2017). The *SAR11* clade has been shown to possess high-affinity uptake systems for a range of small molecules present at low concentrations, including two-carbon compounds, known to induce the glyoxylate shunt. This strategy allows to maintain cellular metabolism with low Fe requirements and to efficiently metabolize small molecules.

Members of *Flavobacteriaceae* appear to be characterized by different features. Our results point out that members of this group display the most competitive Fe-uptake systems. Also, they can potentially regulate homeostasis with bacterioferritin in particular under Fe limited conditions. This group has a moderate use of the glyoxylate shunt, because the biosynthesis of TBDTs, and their transport to the cytoplasm renders the acquisition of siderophore-bound Fe a process that is costlier in terms of energy and carbon requirements than that of Fe^{3+}-uptake. Siderophore biosynthesis, coupled in many bacterial genomes to TBDTs (Hopkinson and Barbeau, 2012), adds further energy requirements (Sijerčić and Price, 2015). Besides Fe-siderophore-uptake, TBDTs were associated with the uptake of a range of substrates, such as carbohydrates, amino acids, amino sugars or vitamin B_{12} (Schauer *et al*., 2008; Noinaj *et al*., 2010).

The highest number and most diverse types of TBDTs were associated with *Gammaproteobacteria* and *FCB* in the GOS data set (Tang *et al*., 2012). This genomic information, in combination with whole genome sequencing, culture-based studies and single-cell approaches have led to the characterization of members of FCB to be efficient degraders of polymeric organic matter (Kirchman *et al*., 2003; Bauer and Blodau, 2006; Kabisch *et al*., 2014). Even though members of both groups are well equipped to thrive in Fe-limited environments, *Pelagibacteraceae* are likely to have an advantage over *Flavobacteriaceae* when organic carbon is limiting. These contrasting characteristics for members of *Pelagibacteracea* and *Flavobacteriaceae* extend those known for other metabolic traits to clear prevalence for Fe^{3+}- and siderophore-uptake, respectively, and could be considered as ecological strategies in an ocean region where microbial activity is limited by Fe and organic carbon.

Experimental procedures

Sample collection

Seawater samples were collected during the KEOPS2 cruise (Kerguelen Ocean and Plateau Compared Study 2, 8 October to 30 November2011) in the Indian sector of the Southern Ocean.

Seawater samples were collected with 12 l Niskin bottles mounted on a rosette equipped with a CTDO Seabird SBE911-plus. For nucleic acid extractions, seawater was sampled at one depth in the surface mixed layer, and the chemical and biological parameters were collected throughout the water column (Christaki *et al*., 2014; Lasbleiz *et al*., 2014; Blain *et al*., 2015; Quéroué *et al*., 2015; Tremblay *et al*., 2015) (Table 1).

RNA extraction

For RNA extractions, volumes varying between 15 l and 30 l of pre-filtered water (200 μm nylon screen and 5 μm polycarbonate isopore filters) were collected onto 0.2 μm SuporPlus Membranes using a 142 mm filtration system (geotech equipment) and a peristaltic pump. The filtration procedure did not exceed 10 min and 10 ml of RNA-later was added before storage at −80°C. All nucleic acid extractions were performed in triplicates by cutting the filter in three parts.

Total prokaryotic and eukaryotic RNA was extracted using the NucleoSpin® RNA Midi kit (Macherey-Nagel, Düren, Germany). Filters stored in RNA later were defrosted, removed from the RNA later solution, refrozen in liquid nitrogen and shattered using a mortar. The obtained 'powder-like' filter-pieces were added together with low binding zirconium beads (OPS Diagnostics, Lebanon, NJ, USA) to the denaturing lysis buffer supplied by the NucleoSpin® RNA Midi kit and cells were disrupted by vortexing for 2 min. Beads were discarded by centrifugation. The extraction with the NucleoSpin® RNA Midi kit include an on-column DNA digestion step. However, in order to ensure the absence of DNA in the sample, a control PCR reaction was performed without the retrotranscription (RT) step. Samples with DNA contamination, as indicated by amplification products were treated with a second DNA digestion step using the Turbo DNA-*free* kit (Ambion Life Technologies, Carlsbad, CA, USA). This additional DNAse treatment was followed by purification with the RNeasy MinElute Clean Up kit (Qiagen, Hilden, Germany). The extracted RNA was quantified with the Agilent 2100 Bioanalyzer/Agilent RNA 6000 Nano Kit (Agilent, Santa Clara, CA, USA) and duplicates were chosen for sequencing.

Library preparation and sequencing

Prior to sequencing, ribosomal RNA was treated enzymatically with the RiboZero rRNA stranded RNA protocol to ensure sequencing of primarily messenger RNA followed by cDNA library construction using Illumina TruSeq Stranded mRNA Library Prep kit (Fasteris SA). Libraries were sequenced using paired-end 2 × 125 read length on one Illumina HiSeq 2500 lane.

Bioinformatic analysis

The raw Illumina reads were checked with FastQC (Andrews 2010; http://www.bioinformatics.babraham.ac.uk/projects/fastqc) and adapters were eliminated using Cutadapt (Martin, 2011). Remaining ribosomal RNA sequences were removed by the riboPicker (Schmieder *et al*., 2012) tool and sequences were checked by interlacing and de-interlacing paired-end reads ensuring that the same sequences were removed from each R1 and R2 files and finally retaining only R1 (performed in Galaxy, Afgan *et al*., 2016). Randomized subsets of 1% of the data were affiliated using BLASTX (Altschul *et al*., 1990) against the non-redundant (nr) protein database followed by the visualization in MEGAN6 (Huson *et al*., 2007) and proportions of unassigned to not assigned sequences as well as prokaryotic to eukaryotic sequences were retrieved (Supporting Information Fig. S1). These sequence data have been submitted to the EMBL databases under accession number PRJEB30315.

Database construction

A database containing sequences of genes involved into Fe-related metabolic pathway was retrieved from the study by Toulza *et al*. (2012). This database was constructed by screening for bacterial sequences from NCBI

with the gene name as query, as well as the protein sequences from the Moore Microbial Genome database (http://www.moore.org/microgenome/) for genes involved in Fe metabolism. For the purpose of this study, the specific sequences for the following pathways were retrieved from the database and updated by searching for these in NCBI protein clusters: Flavodoxin switch (FL), Fe^{2+}-uptake (F2), Fe^{3+}-uptake (F3), siderophore-uptake (SU), and storage (ST) (Supporting Information Table S2). These five together are further named as 'Fe' database.

In addition, protein sequences for two supplementary enzymes were chosen for pathway specific analysis. Aconitase that catalyses the isomerization of citrate to isocitrate via cis-aconitate in the tricarboxylic acid cycle and isocitrate lyase an enzyme in the glyoxylate shunt which catalyses the cleavage of isocitrate to succinate and glyoxylate. For maximal phylogenetic coverage, all available bacterial protein sequences were retrieved from protein clusters (proteins grouped on taxonomic groups which are non-redundant) using the NCBI search tool with the protein name as query.

In order to include more sequences from environmental marine bacteria the two databases were aligned using BLASTX tool (Altschul *et al*., 1990) against the Global Ocean Sampling (GOS) protein database (retrieved from /iplant/home/shared/imicrobe/projects/26/CAM_PROJ_GOS .read_pep.fa, Yooseph *et al*., 2007) containing peptide sequences predicted from long reads from Sanger sequencing and thus more robust for the purpose of the analysis. The retrieved GOS sequences were then checked against the KEGG (Kanehisa and Goto, 2000; Kanehisa *et al*., 2016; 2017) and for each custom database sequences with an e-value of $<1e^{-5}$ were chosen for further analysis and annotated from their KEGG-Id (Supporting Information Table S3). Finally, each database contained FASTA sequences with the taxonomic affiliation in the header as well as the KEGG-Id. For SU only Fe-related siderophore KEGG-Ids were retained.

Additionally, all ribosomal protein sequences (n = 261 980) stored in the KEGG database (download January 2018) were broadly (order level) annotated by retrieving taxonomic information for sequences from KEGG. These sequences were used to recruit ribosomal protein transcripts from our metatranscriptome data that served as verification of the normalization approach.

Sequence alignment

The final curated databases (n = 4, Fe, aconitase, isocitrate lyase and ribosomal proteins) containing information on the annotated taxonomic levels (phylum, class, order, family and genus) were aligned to the short-read translated DNA query sequences for each station and duplicate using diamond blastx (parameters used -k 1 -e 10 -p 12) (Buchfink *et al*., 2015). Total counts per phylum, class, order, family and genus for each database, and in case of Fe, for each pathway, were summed and relative proportions to all prokaryotic reads for bacterial groups were calculated (Supporting Information Fig. S7). Bacterial groups defined at the taxonomic order level with the highest abundances of pathway specific transcripts were taken for further analysis.

Normalization approach

In an attempt to estimate the absolute number of transcripts per sample, we followed the principals published elsewhere (Satinsky *et al*., 2013): the number of reads per sample is normalized by the number of reads obtained from an internal standard added with a known number of RNA molecules to the RNA extraction. However, in our case, we did not add internal standard RNA molecules, but instead based our calculations on the number of SAR11 isocitrate lyase transcripts for normalization, which were quantified earlier via qPCR (Beier *et al*., 2015).

While the number of SAR11 isocitrate lyase transcripts per L water derived from the qPCR approach might be biased, for instance, due to primer miss matches, such biases are strongly reduced for the ratio of SAR11-isocitrate lyase gene transcripts to SAR11-isocitrate lyase gene copies (Beier *et al*., 2015). For the normalization step, we therefore assumed that the above-mentioned ratio derived from qPCR data equals the ratio of SAR11-isocitrate lyase gene transcripts (metatranscriptome; RNA) to the number of SAR11 cell per L estimated by CARD-FISH (as described in the study by Fourquez *et al*. 2016):

$$qT_{iso}/qC_{iso} = mT_{iso}/nC_{SAR11} \quad (1)$$

where, for the purpose of this study, qT_{iso}/qC_{iso} is the ratio of SAR11-isocitrate lyase gene transcripts L^{-1} (qT_{iso}) to SAR11-isocitrate lyase gene copies L^{-1} (qC_{iso}) estimated by qPCR (Beier *et al*., 2015), mT_{iso} is the number of SAR11-isocitrate lyase gene transcripts L^{-1}, and nC_{SAR11} is the SAR11 cells L^{-1} estimated by CARD-FISH (Fourquez *et al*. 2016)

The ratio qT_{iso}/qC_{iso} as well as nC are known variables and Eq. 1 can accordingly be resolved by mT_{iso}. We subsequently related mT_{iso} to the number of metatranscriptome reads coding for SAR11 isocitrate lyase gene transcripts (mT_{iso}) and used this factor to estimate the absolute transcript numbers per L water or all remaining genes (mT).

qPCR data were only available for Station R-2 and F-L, thus samples from A3-2 were excluded from these calculations. Operational taxonomic units (OTUs) obtained by 16S rRNA gene sequencing were retrieved for Station R-2 and F-L from an already published data set from the same sampling date (Landa *et al*., 2016). The OTUs were

corrected for copy numbers of 16S rRNA gene per cell per specific taxa obtained from the ribosomal RNA database (Stoddard *et al*., 2015). Total cell numbers per bacterial group were then calculated with 16S rRNA gene relative proportions and counts from the Eub228-I, -II and -II catalysed reporter deposition–fluorescence *in situ* hybridization (CARD-FISH) probe from the same samples (data from the study by Fourquez *et al*., 2016, Supporting Information Table S6).

In order to obtain the number of transcripts per cell, we divided mT for the inspected genes derived from a certain bacterial group (based on the number of hits from the specific databases) by the total cell numbers per L for the respective bacterial group shown in Supporting Information Table S6. While we tried to minimize biases introduced by the qPCR approach, it should be considered that our method is not fully free of biases, also because additional biases might be introduced by combing multiple techniques, such as qPCR, CARD-FISH and amplicon sequencing data. Accordingly, the retrieved and presented values (Fig. 6) represent a rough estimate of transcripts per cell.

Analysis and visualization of data were performed in the R language [https://cran.r-project.org/, version 3.4.0 (21 April 2017)] using customized colour palettes. Following packages and versions were used: phyloseq_1.22.3, gdtools_0.1.7, gridExtra_2.3, cowplot_0.9.3, gtable_0.2.0, ggpubr_0.2, magrittr_1.5, ggplot2_3.0.0, plyr_1.8.4. Figures were using the open-source vector graphics editor Inkscape (http://inkscape.org/). Codes and databases are provided through the following link https://github.com/PavlaDe/emi14621/.

Acknowledgements

We thank the chief scientist Bernard Quéguiner for leading the KEOPS2 expedition. We also thank the officers and the crew aboard the *R/V Marion Dufresne* for their help in the successful completion of the cruise. We thank Marine Landa for her help with the seawater collection and Francois-Yves Bouget for financial support for the sequencing of the samples. The authors thank the Roscoff Bioinformatics platform ABiMS (http://abims.sb-roscoff.fr) for providing computational resources. Extremely important comments on this data set have come from Barbara Bayer, Laurie Bousquet, Coco Koedooder and Yan Liu during multiple discussions and have greatly improved this work. Two anonymous reviewers provided insightful comments on a previous version of the manuscript. The KEOPS 2 project was supported by the French Research program of INSU-CNRS LEFE–CYBER ('Les enveloppes fluids et l'environnement' – 'Cycles biogéochimiques, environnement et ressources'), the French ANR ('Agence Nationale de la Recherche', SIMI-6 program, ANR-2010-BLAN-614 KEOPS2 and, ANR-10-JCJC-606 ICOP), the French CNES program ('Centre National d'Etudes Spatiales') and the French Polar Institute IPEV (Institut Polaire Paul-Émile Victor). This work was additionally funded by the ANR project PhytoIron 11BSV7 018 02 and the Austrian FWF grant under the number P28781-B21, and is part of the PhD thesis of PD.

References

Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Čech, M., *et al*. (2016) The galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res* **44**: W3–W10.

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

Andrews, S.C., Robinson, A.K., and Rodríguez-Quiñones, F. (2003) Bacterial iron homeostasis. *FEMS Microbiol Rev* **27**: 215–237.

Andrews, S. (2010) FastQC: a quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Bauer, M., Kube, M., Teeling, H., Richter, M., Lombardot, T., Allers, E., *et al*. (2006) Whole genome analysis of the marine Bacteroidetes 'Gramella forsetii' reveals adaptations to degradation of polymeric organic matter. *Environ Microbiol* **8**: 2201–2213.

Bauer, M., and Blodau, C. (2006) Mobilization of arsenic by dissolved organic matter from iron oxides, soils and sediments. *Sci Total Environ* **354**: 179–190.

Beier, S., Gálvez, M.J., Molina, V., Sarthou, G., Quéroué, F., Blain, S., and Obernosterer, I. (2015) The transcriptional regulation of the glyoxylate cycle in SAR11 in response to iron fertilization in the Southern Ocean. *Environ Microbiol Rep* **7**: 427–434.

Blain, S., Capparos, J., Guéneuguès, A., Obernosterer, I., and Oriol, L. (2015) Distributions and stoichiometry of dissolved nitrogen and phosphorus in the iron-fertilized region near Kerguelen (Southern Ocean). *Biogeosciences* **12**: 623–635.

Blain, S., Queguiner, B., Armand, L.K., Belviso, S., and Bomb, B. (2007) Effect of natural iron fertilization on carbon sequestration in the Southern Ocean. *Nature* **446**: 1070–1074.

Boiteau, R.M., Mende, D.R., Hawco, N.J., McIlvin, M.R., Fitzsimmons, J.N., Saito, M.A., *et al*. (2016) Siderophore-based microbial adaptations to iron scarcity across the eastern Pacific Ocean. *PNAS* **113**: 14237–14242.

Botebol, H., Lesuisse, E., Šuták, R., Six, C., Lozano, J.-C., Schatt, P., *et al*. (2015) Central role for ferritin in the day/night regulation of iron homeostasis in marine phytoplankton. *PNAS* **112**: 14652–14657.

Boyd, P.W., Jickells, T., Law, C.S., Blain, S., Boyle, E.A., Buesseler, K.O., *et al*. (2007) Mesoscale iron enrichment experiments 1993-2005: synthesis and future directions. *Science* **315**: 612–617.

Buchfink, B., Xie, C., and Huson, D.H. (2015) Fast and sensitive protein alignment using DIAMOND. *Nat Methods* **12**: 59–60.

Carrondo, M.A. (2003) Ferritins, iron uptake and storage from the bacterioferritin viewpoint. *EMBO J* **22**: 1959–1968.

Carradec, Q., Pelletier, E., Da Silva, C., Alberti, A., Seeleuthner, Y., Blanc-Mathieu, Y., *et al*. (2018) A global ocean atlas of eukaryotic genes. *Nat Commun.* **9**.

Christaki, U., Lefèvre, D., Georges, C., Colombet, J., Catala, P., Courties, C., *et al*. (2014) Microbial food web dynamics during spring phytoplankton blooms in the naturally iron-fertilized Kerguelen area (Southern Ocean). *Biogeosciences* **11**: 6739–6753.

Cochlan, W.P. (2001) The heterotrophic bacterial response during a mesoscale iron enrichment experiment (IronEx II) in the eastern equatorial Pacific Ocean. *Limnol Oceanogr* **46**: 428–435.

Desai, D.K., Desai, F., and LaRoche, J. (2012) Factors influencing the diversity of iron uptake systems in aquatic microorganisms. *Front Microbiol* **3**: 362.

Erdner, D.L., Price, N.M., Doucette, G.J., Peleato, M.L., and Anderson, D.M. (1999) Characterization of ferredoxin and flavodoxin as markers of iron limitation in marine phytoplankton. *Mar Ecol Prog Ser* **184**: 43–53.

Fourquez, M., Beier, S., Jongmans, E., Hunter, R., and Obernosterer, I. (2016) Uptake of Leucine, chitin, and iron by prokaryotic groups during spring phytoplankton blooms induced by natural iron fertilization off Kerguelen Island (Southern Ocean). *Front Mar Sci* **3**: 256.

Fourquez, M., Devez, A., Schaumann, A., Guéneuguès, A., Jouenne, T., Obernosterer, I., and Blain, S. (2014) Effects of iron limitation on growth and carbon metabolism in oceanic and coastal heterotrophic bacteria. *Limnol Oceanogr* **59**: 349–360.

Fourquez, M., Obernosterer, I., and Blain, S. (2012) A method for the use of the radiotracer 55Fe for microautoradiography and CARD-FISH of natural bacterial communities. *FEMS Microbiol Lett* **337**: 132–139.

Fourquez, M., Obernosterer, I., Davies, D.M., Trull, T.W., and Blain, S. (2015) Microbial iron uptake in the naturally fertilized waters in the vicinity of the Kerguelen Islands: phytoplankton–bacteria interactions. *Biogeosciences* **12**: 1893–1906.

Giovannoni, S.J. (2017) SAR11 bacteria: the Most abundant plankton in the oceans. *Annu Rev Mar Sci* **9**: 231–255.

Gledhill, M., and van den Berg, C.M.G. (1994) Determination of complexation of iron(III) with natural organic complexing ligands in seawater using cathodic stripping voltammetry. *Mar Chem* **47**: 41–54.

Hall, J.A., and Safi, K. (2001) The impact of in situ Fe fertilisation on the microbial food web in the Southern Ocean. *Deep-Sea Res II Top Stud Oceanogr* **48**: 2591–2613.

Hogle, S.L., Cameron Thrash, J., Dupont, C.L., and Barbeau, K.A. (2016) Trace metal acquisition by marine heterotrophic bacterioplankton with contrasting trophic strategies. *Appl Environ Microbiol* **82**: 1613–1624.

Hopkinson, B.M., and Barbeau, K.A. (2012) Iron transporters in marine prokaryotic genomes and metagenomes. *Environ Microbiol* **14**: 114–128.

Hudson, R.J.M., and Morel, F.M.M. (1993) Trace metal transport by marine microorganisms: implications of metal coordination kinetics. *Deep-Sea Res I Oceanogr Res Pap* **40**: 129–150.

Huson, D.H., Auch, A.F., Qi, J., and Schuster, S.C. (2007) MEGAN analysis of metagenomic data. *Genome Res* **17**: 377–386.

Kabisch, A., Otto, A., König, S., Becher, D., Albrecht, D., Schüler, M., *et al*. (2014) Functional characterization of polysaccharide utilization loci in the marine *Bacteroidetes* '*Gramella forsetii*' KT0803. *ISME J* **8**: 1492–1502.

Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017) KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* **45**: D353–D361.

Kanehisa, M., and Goto, S. (2000) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**: 27–30.

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016) KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* **44**: D457–D462.

Kazamia, E., Sutak, R., Paz-Yepes, J., Dorrell, R.G., Vieira, F.R.J., Mach, J., *et al*. (2018) Endocytosis-mediated siderophore uptake as a strategy for Fe acquisition in diatoms. *Sci Adv* **4**: eaar4536.

Kirchman, D.L., Yu, L., and Cottrell, M.T. (2003) Diversity and abundance of uncultured Cytophaga-like bacteria in the Delaware estuary. *Appl Environ Microbiol* **69**: 6587–6596.

Koedooder, C., Gueneugues, A., Van Geersdaele, R., Verge, V., Bouget, F.-Y., Labreuche, Y., *et al*. (2018) The role of the Glyoxylate shunt in the acclimation to iron limitation in marine heterotrophic bacteria. *Front Mar Sci* **5**: 435.

Kraemer, S.M. (2004) Iron oxide dissolution and solubility in the presence of siderophores. *Aquat Sci* **66**: 3–18.

Landa, M., Blain, S., Christaki, U., Monchy, S., and Obernosterer, I. (2016) Shifts in bacterial community composition associated with increased carbon cycling in a mosaic of phytoplankton blooms. *ISME J* **10**: 39–50.

Lasbleiz, M., Leblanc, K., Blain, S., Ras, J., Cornet-Barthaux, V., Hélias Nunige, S., and Queguiner, B. (2014) Pigments, elemental composition (C, N, P, and Si), and stoichiometry of particulate matter in the naturally iron fertilized region of Kerguelen in the Southern Ocean. *Biogeosciences* **11**: 5931–5955.

Lis, H., Shaked, Y., Kranzler, C., Keren, N., and Morel, F.M.M. (2015) Iron bioavailability to phytoplankton: an empirical approach. *ISME J* **9**: 1003–1013.

Marchetti, A., Parker, M.S., Moccia, L.P., Lin, E.O., Arrieta, A.L., Ribalet, F., *et al*. (2009) Ferritin is used for iron storage in bloom-forming marine pennate diatoms. *Nature* **457**: 467–470.

Martin, J.H. (1990) Glacial-interglacial CO2 change: the iron hypothesis. *Paleoceanography* **5**: 1–13.

Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* **17**: 10.

Martinez, J.S., Carter-Franklin, J.N., Mann, E.L., Martin, J. D., Haygood, M.G., and Butler, A. (2003) Structure and membrane affinity of a suite of amphiphilic siderophores produced by a marine bacterium. *PNAS* **100**: 3754–3759.

McMurdie, P. J., & Holmes, S., *et al*. (2013) phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE* **8**(4): e61217. https:/doi.org/:10.1371/journal.pone.0061217

McQuaid, J.B., Kustka, A.B., Oborník, M., Horák, A., McCrow, J.P., Karas, B.J., *et al*. (2018) Carbonate-sensitive phytotransferrin controls high-affinity iron uptake in diatoms. *Nature* **555**: 534–537.

Morel, F.M.M., Kustka, A.B., and Shaked, Y. (2008) The role of unchelated Fe in the iron nutrition of phytoplankton. *Limnol Oceanogr* **53**: 400–404.

Noinaj, N., Guillier, M., Barnard, T.J., and Buchanan, S.K. (2010) TonB-dependent transporters: regulation, structure, and function. *Annu Rev Microbiol* **64**: 43–60.

Obernosterer, I., Christaki, U., Lefèvre, D., Catala, P., Van Wambeke, F., and Lebaron, P. (2008) Rapid bacterial mineralization of organic carbon produced during a phytoplankton bloom induced by natural iron fertilization in the Southern Ocean. *Deep-Sea Res II Top Stud Oceanogr* **55**: 777–789.

Obernosterer, I., Fourquez, M., and Blain, S. (2015) Fe and C co-limitation of heterotrophic bacteria in the naturally fertilized region off the Kerguelen Islands. *Biogeosciences* **12**: 1983–1992.

Oliver, J.L., Barber, R.T., Smith, W.O., and Ducklow, H.W. (2004) The heterotrophic bacterial response during the Southern Ocean iron experiment (SOFeX). *Limnol Oceanogr* **49**: 2129–2140.

Pollard, R.T., Salter, I., Sanders, R.J., Lucas, M.I., Moore, C. M., Mills, R.A., *et al*. (2009) Southern Ocean deep-water carbon export enhanced by natural iron fertilization. *Nature* **457**: 577–580.

Quéroué, F., Sarthou, G., Planquette, H.F., Bucciarelli, E., Chever, F., van der Merwe, P., *et al*. (2015) High variability of dissolved iron concentrations in the vicinity of Kerguelen Island (Southern Ocean). *Biogeosciences Discuss* **12**: 231–270.

Roche, J.L., Boyd, P.W., McKay, R.M.L., and Geider, R.J. (1996) Flavodoxin as an in situ marker for iron stress in phytoplankton. *Nature* **382**: 802–805.

Rue, E.L., and Bruland, K.W. (1995) Complexation of iron(III) by natural organic ligands in the central North Pacific as determined by a new competitive ligand equilibration/adsorptive cathodic stripping voltammetric method. *Mar Chem* **50**: 117–138.

Saito, M.A., Bertrand, E.M., Dutkiewicz, S., Bulygin, V.V., Moran, D.M., Monteiro, F.M., *et al*. (2011) Iron conservation by reduction of metalloenzyme inventories in the marine diazotroph Crocosphaera watsonii. *PNAS* **108**: 2184–2189.

Sarthou, G., Vincent, D., Christaki, U., Obernosterer, I., Timmermans, K.R., and Brussaard, C.P.D. (2008) The fate of biogenic iron during a phytoplankton bloom induced by natural fertilisation: impact of copepod grazing. *Deep-Sea Res II Top Stud Oceanogr* **55**: 734–751.

Satinsky, B.M., Gifford, S.M., Crump, B.C., and Moran, M.A. (2013) Use of internal standards for quantitative Metatranscriptome and Metagenome analysis. *Methods Enzymol* **531**: 237–250.

Schauer, K., Rodionov, D.A., and de Reuse, H. (2008) New substrates for TonB-dependent transport: do we only see the 'tip of the iceberg'? *Trends Biochem Sci* **33**: 330–338.

Schmieder, R., Lim, Y.W., and Edwards, R. (2012) Identification and removal of ribosomal RNA sequences from metatranscriptomes. *Bioinformatics (Oxford, England)* **28**: 433–435.

Sijerčić, A., and Price, N. (2015) Hydroxamate siderophore secretion by Pseudoalteromonas haloplanktis during steady-state and transient growth under iron limitation. *Mar Ecol Prog Ser* **531**: 105–120.

Smith, D.P., Kitner, J.B., Norbeck, A.D., Clauss, T.R., Lipton, M.S., Schwalbach, M.S., *et al*. (2010) Transcriptional and translational regulatory responses to iron limitation in the globally distributed marine bacterium Candidatus Pelagibacter ubique. *PLoS One* **5**: e10487.

Stoddard, S.F., Smith, B.J., Hein, R., Roller, B.R.K., and Schmidt, T.M. (2015) rrnDB: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. *Nucleic Acids Res* **43**: D593–D598.

Tagliabue, A., Bowie, A.R., Philip, W., Buck, K.N., Johnson, K. S., and Saito, M.A. (2017) The integral role of iron in ocean biogeochemistry. *Nature* **543**: In Press: 51–59.

Tang, K., Jiao, N., Liu, K., Zhang, Y., and Li, S. (2012) Distribution and functions of TonB-dependent transporters in marine bacteria and environments: implications for dissolved organic matter utilization. *PLoS ONE* **7**: e41204.

Tortell, P.D., Maldonado, M.T., Granger, J., and Price, N.M. (1999) Marine bacteria and biogeochemical cycling of iron in the oceans. *FEMS Microbiol Ecol* **29**: 1–11.

Toulza, E., Tagliabue, A., Blain, S., and Piganeau, G. (2012) Analysis of the global ocean sampling (GOS) project for trends in iron uptake by surface ocean microbes. *PLoS ONE* **7**: e30931.

Tremblay, L., Caparros, J., LEBLANC, K., and Obernosterer, I. (2015) Origin and fate of particulate and dissolved organic matter in a naturally iron-fertilized region of the Southern Ocean. *Biogeosciences* **12**: 607–621.

Yooseph, S., Sutton, G., Rusch, D.B., Halpern, A.L., Williamson, S.J., Remington, K., *et al*. (2007) The sorcerer II Global Ocean sampling expedition: expanding the universe of protein families. *PLoS Biol* **5**: e16.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix S1. Manuscript_Codes

Supplementary Figure 1. Blast results of subsets for assigned to non-assigned reads.

Supplementary Figure 2. Relative contribution of prokaryotic groups to pathway specific transcripts. Prokaryotic group is defined until taxonomic family level. Number 1 and 2 refer to the sequencing results of duplicates per station.

Supplementary Figure 3 A-C. Detailed view of relative contribution of *Alpha- and Gammaproteobacteria* and FCB cluster to pathway specific transcripts. Note the different y-axis for *Gammaproteobacteria.* Number 1 and 2 refer to the sequencing results of duplicates per station.

Supplementary Figure 4. Percentage of ribosomal protein transcripts to all prokaryotic transcripts.

Supplementary Figure 5 A-G. Relative contribution of prokaryotic groups to pathway specific transcripts and

relative abundance of 16S rRNA gene sequences. Prokaryotic group is defined until taxonomic order level.

Supplementary Figure 6. Simple illustration of the glyoxylate shunt (inside of the circle) in which isocitrate lyase cleaves isocitrate into glyoxylate and succinate.

Supplementary Figure 7. Detailed plots for duplicates on the phylum, class and order level for each pathway.

Supplementary Table 1. General information on sequencing results and reads.

Supplementary Table 2. Relative contribution of prokaryotic groups to pathway specific transcripts. For each group, mean value $\pm$ standard deviation of 2 replicates are shown. Error estimates are provided for all prokaryotic groups illustrated in Figs 2 & 3, and groups are listed by alphabetic order.

Supplementary Table 3. Taxa-specific transcript abundance (per 10^5 cells) of a given pathway at Station F-L and R-2. For each group, mean value $\pm$ standard deviation of 2 replicates are shown. Error estimates are provided for all prokaryotic groups illustrated in Figs 6, and groups are listed by alphabetic order.

Supplementary Table 4. Information of databases constructed or modified from Toulza *et al*. (2012) and retrieved by NCBI

Supplementary Table 5. Pathways and corresponding KEGG-Id numbers that were chosen for further analysis

Supplementary Table 6. Cells per L for prokaryotic groups in Fig. 6 for station R-2 and F-L, calculated as described in experimental procedures.

Chapter 1 Supplementary Tables and Figures

Supplementary Table 1. *General information on sequencing results and reads.*

Sample Names	Raw Reads from paired-end sequencing (fastq replicates)	Interlaced reads after trimming and rRNA removal	Percent of prokaryotic reads from BLAST (%)	Number of prokaryotic reads from total reads
R_1	25 795 611 25 795 611	31 872 360	8.9	2 860 032
R_2	24 750 021 24 750 021	31 519 356	10.5	3 308 037
A3_2_1*	23 871 679 23 871 679	29 824 202	15.3	5 514 089
A3_2_2*	21 853 684 21 853 684	34 620 685	13.8	4 438 123
FL_1*	20 460 960 20 460 960	36 070 585	18.0	5 373 036
FL_2*	23 556 644 23 556 644	32 099 894	21.7	7 504 148

* The higher number of prokaryotic reads and their higher relative contributions to total reads at A3-2 and F-L as compared to R-2 can be explained by the elevated prokaryotic cell abundances at these sites (Table 1) and differences in the phytoplankton community composition. While larger diatom cells, abundant at the 2 bloom sites, were retained by the 5 µm filter used for the pre-filtration step, phytoplankton biomass at station R-2 was dominated by picoeukaryotes (Lasbleiz et al. 2016) that pass this pore size resulting in a higher number of eukaryote-assigned reads.

Supplementary Table 2. *Relative contribution of prokaryotic groups to pathway specific transcripts. For each group, mean value ± standard deviation of 2 replicates are shown. Error estimates are provided for all prokaryotic groups illustrated in Figures 2 & 3, and groups are listed by alphabetic order.*

Station	Prokaryotic group	Aconitase	F^{2+} uptake	F^{3+} uptake	Flavodoxin switch	Bacterioferritin	Isocitrate lyase	Siderophore uptake
A3_2	Actinomycetaceae	0	0	1.9 ± 0.09	0	0	0	0
A3_2	Aeromonadaceae	0	0	0	3.33 ± 0.16	0	0	0
A3_2	Alcanivoracaceae	0	0	0	5.18 ± 0.07	0	0	0
A3_2	Alteromonadaceae	3.11 ± 0.06	0	0	1.48 ± 0.21	5.04 ± 0.82	5.45 ± 0.26	13.17 ± 0.25
A3_2	Aphanizomenonaceae	0	0	0	4.37 ± 0.24	0	0	0
A3_2	Archaea	0	2.42 ± 0.13	6.15 ± 0.19	2.33 ± 0.04	1.74 ± 0.21	0	0.26 ± 0.01
A3_2	Bacillaceae	0	0	2.63 ± 0.11	0	0	0	0
A3_2	Burkholderiaceae	2.49 ± 0.14	0	0	0	0	7.45 ± 0.65	1.12 ± 0.01
A3_2	Cellvibrionaceae	0	0	0	2.55 ± 0.12	4.08 ± 0.05	0	8.81 ± 0.14
A3_2	Chromatiaceae	0	0	2.92 ± 0.34	0	0	0	0
A3_2	Chromobacteriaceae	0	0	0	0	1.72 ± 0.48	0	0
A3_2	Colwelliaceae	0	0	0	0	0	0.31 ± 0.11	0
A3_2	Coriobacteriaceae	0	0.6 ± 0.06	0	0	0	0	0
A3_2	Corynebacterineae	0	0	0	0	0	7.65 ± 0.04	0
A3_2	Crocinitomicaceae	0	0	0	0	2.7 ± 0.49	0	0

A3_2	Enterobacteriaceae	1.17 ± 0.02	18.72 ± 0.51	0	1.94 ± 0.09	0.98 ± 0.36	0	0
A3_2	Erythrobacteraceae	0	0	0	0	0	0	3.52 ± 1.37
A3_2	FCB cluster	8.03 ± 0.38	4.25 ± 0.37	2.88 ± 0.06	2.96 ± 0.17	1.09 ± 0.21	1.1 ± 0.16	8.87 ± 0.54
A3_2	Flavobacteriaceae	12.78 ± 0.69	10.79 ± 1.65	2.04 ± 0.01	0	13.97 ± 0.28	2.33 ± 0.17	15.95 ± 1.49
A3_2	Halieaceae	0	0	0	0	9.83 ± 0.15	0	0
A3_2	Moraxellaceae	0	0	0	0	0	9.53 ± 0.34	0
A3_2	Nitrosomonadaceae	0.18 ± 0.05	0	0	3.07 ± 0.16	0	0	0
A3_2	Oceanospirillaceae	0.98 ± 0.11	0	0	0	1.32 ± 0.11	0.76 ± 0.28	0.47 ± 0.01
A3_2	Other	10.13 ± 0.23	14.12 ± 0.1	10.13 ± 0.42	16.24 ± 0.13	5.83 ± 0.31	9.44 ± 0.4	8.04 ± 0.67
A3_2	other Actinobacteria	11.05 ± 0.26	3.1 ± 0.17	9.5 ± 0.08	3.34 ± 0.02	1.94 ± 0.09	8.55 ± 0.02	0.88 ± 0.06
A3_2	other Alphaproteobacteria	10.4 ± 0.3	6.25 ± 0.15	5.55 ± 0.64	7.46 ± 0.05	5.59 ± 0.52	11.04 ± 0.18	7 ± 0.25
A3_2	other Betaproteobacteria	5.22 ± 0.2	2.21 ± 0.43	4.78 ± 0.25	4.92 ± 0.07	6.88 ± 0.45	6.34 ± 0.34	1.56 ± 0.08
A3_2	other Cyanobacteria	0.74 ± 0.06	0.78 ± 0.13	3.44 ± 0.04	1.15 ± 0.05	2.12 ± 0.64	0.09 ± 0.05	0.94 ± 0.01
A3_2	other Gammaproteobacteria	17.96 ± 0.23	14.19 ± 0.91	11.62 ± 0.49	16.05 ± 0.29	14.49 ± 1.61	14.59 ± 0.4	12.65 ± 0.14
A3_2	other Proteobacteria	3.36 ± 0.22	3.21 ± 0.04	0.3 ± 0.06	6.04 ± 0.04	3.35 ± 0.26	2.24 ± 0.14	1.41 ± 0.01
A3_2	Oxalobacteraceae	1.03 ± 0.18	0	0	0	0	0	0
A3_2	Pelagibacteraceae	5.84 ± 0.44	0	0	0	0	6.37 ± 0.15	0
A3_2	Phyllobacteriaceae	0	0	3.16 ± 0.15	0	0	0	0
A3_2	Piscirickettsiaceae	0	0	2.81 ± 0.03	0	0	0	0
A3_2	Prochloraceae	0	0	3.05 ± 0.1	1.46 ± 0.35	0	0	0
A3_2	Pseudoalteromonadaceae	2.04 ± 0.19	1.18 ± 0.03	0	0	0	6.88 ± 0.09	1.89 ± 0.06
A3_2	Pseudomonadaceae	0	0	5.74 ± 0.03	2.66 ± 0.07	3.71 ± 0.41	0	0
A3_2	Rhizobiaceae	0	0	2.43 ± 0.13	0	0	0	0
A3_2	Rhodobacteraceae	0	8.34 ± 0.23	15.37 ± 0.4	0	2.84 ± 0.11	0	0
A3_2	SAR86	3.58 ± 0.02	0	0	0	0	0	0
A3_2	Shewanellaceae	0	5.92 ± 0.33	2.22 ± 0.03	8.08 ± 0.12	1.71 ± 0.32	0	10.97 ± 0.41
A3_2	Sphingomonadaceae	0	0	0	0	0	0	2.57 ± 0.1
A3_2	Synechococcaceae	0	1.36 ± 0.06	1.5 ± 0.15	5.48 ± 0.18	9.16 ± 0.13	0	0
A3_2	Yersiniaceae	0	2.65 ± 0.08	0	0	0	0	0
FL	Actinomycetaceae	0	0	1.93 ± 0.08	0	0	0	0
FL	Aeromonadaceae	0	0	0	4.67 ± 0.47	0	0	0
FL	Alcanivoracaceae	0	0	0	7.68 ± 0.55	0	0	0
FL	Alteromonadaceae	2.76 ± 0.22	0	0	2.65 ± 0.33	8.74 ± 0.67	6.79 ± 0.09	12.41 ± 0.31

FL	Aphanizomenonaceae	0	0	0	2.57 ± 0.53	0	0	0
FL	Archaea	0	2.86 ± 0.44	6.01 ± 0.28	1.74 ± 0.19	1.26 ± 0.08	0	0.44 ± 0.03
FL	Bacillaceae	0	0	2.93 ± 0.25	0	0	0	0
FL	Burkholderiaceae	2.37 ± 0.25	0	0	0	0	4.31 ± 0.09	1.5 ± 0.05
FL	Cellvibrionaceae	0	0	0	3.17 ± 0.18	4.2 ± 0.12	0	5.18 ± 0.11
FL	Chromatiaceae	0	0	1.64 ± 0.29	0	0	0	0
FL	Chromobacteriaceae	0	0	0	0	1.9 ± 0.35	0	0
FL	Colwelliaceae	0	0	0	0	0	5.92 ± 0.66	0
FL	Coriobacteriaceae	0	0.62 ± 0.21	0	0	0	0	0
FL	Corynebacterineae	0	0	0	0	0	2.3 ± 0.17	0
FL	Crocinitomicaceae	0	0	0	0	4.81 ± 0.97	0	0
FL	Enterobacteriaceae	4.61 ± 0.37	14.67 ± 0.18	0	1.3 ± 0.01	0.83 ± 0.04	0	0
FL	Erythrobacteraceae	0	0	0	0	0	0	3.94 ± 0.97
FL	FCB cluster	4.22 ± 0.5	4.95 ± 0.12	3.09 ± 0.11	2.43 ± 0.3	0.66 ± 0.03	0.36 ± 0.03	12.6 ± 0.63
FL	Flavobacteriaceae	7.38 ± 0.34	8.84 ± 0.38	2.14 ± 0.14	0	11.44 ± 0.65	0.72 ± 0.05	13.55 ± 0.94
FL	Halieaceae	0	0	0	0	9.28 ± 0.86	0	0
FL	Moraxellaceae	0	0	0	0	0	7.57 ± 0.92	0
FL	Nitrosomonadaceae	13.99 ± 1.42	0	0	3.96 ± 0.83	0	0	0
FL	Oceanospirillaceae	3.55 ± 0.34	0	0	0	4.36 ± 0.35	11.47 ± 0.47	0.82 ± 0.04
FL	Other	6.72 ± 0.07	13.71 ± 0.57	9.53 ± 0.17	12.93 ± 0.06	5.15 ± 0.93	7.73 ± 0.11	7.31 ± 0.17
FL	other Actinobacteria	7.46 ± 0.12	3.8 ± 0.04	8.72 ± 0.02	2.74 ± 0.38	0.9 ± 0.11	2.97 ± 0.1	1.22 ± 0.02
FL	other Alphaproteobacteria	8.14 ± 0.97	6.09 ± 0.24	5.81 ± 0.21	3.58 ± 0.5	4.58 ± 0.3	7.34 ± 0.04	7.23 ± 0.01
FL	other Betaproteobacteria	3.26 ± 0.05	5.77 ± 0.08	4.96 ± 0.12	6.24 ± 0.39	4.31 ± 0.25	8.05 ± 1.3	1.83 ± 0.07
FL	other Cyanobacteria	0.5 ± 0.02	1.06 ± 0.18	2.9 ± 0.18	1.41 ± 0.21	4.04 ± 0.38	0.03 ± 0.01	1.38 ± 0.06
FL	other Gammaproteobacteria	12.76 ± 0.38	15.74 ± 0.05	11.07 ± 0.04	13.12 ± 0.22	9.97 ± 0.9	25.6 ± 1.03	13.57 ± 0.21
FL	other Proteobacteria	1.79 ± 0.13	2.84 ± 0.36	0.25 ± 0.02	9.19 ± 0.51	1.36 ± 0.33	0.86 ± 0.03	1.56 ± 0.01
FL	Oxalobacteraceae	12.22 ± 1.55	0	0	0	0	0	0
FL	Pelagibacteraceae	4.09 ± 0.22	0	0	0	0	2.56 ± 0.03	0
FL	Phyllobacteriaceae	0	0	3.78 ± 0.05	0	0	0	0
FL	Piscirickettsiaceae	0	0	2.27 ± 0.08	0	0	0	0
FL	Prochloraceae	0	0	3.06 ± 0.01	1.58 ± 0.11	0	0	0
FL	Pseudoalteromonadaceae	1.84 ± 0.1	1.02 ± 0.06	0	0	0	5.5 ± 0.11	1.4 ± 0.02
FL	Pseudomonadaceae	0	0	4.6 ± 0.16	4.15 ± 0.35	7.88 ± 0.68	0	0

FL	Rhizobiaceae	0	0	2.37 ± 0.02	0	0	0	0
FL	Rhodobacteraceae	0	11.47 ± 1.07	18.96 ± 0.76	0	5.14 ± 0.27	0	0
FL	SAR86	2.43 ± 0.08	0	0	0	0	0	0
FL	Shewanellaceae	0	2.91 ± 0.36	2.34 ± 0.13	9.47 ± 0.4	4.02 ± 0.33	0	10.81 ± 0.32
FL	Sphingomonadaceae	0	0	0	0	0	0	3.34 ± 0.05
FL	Synechococcaceae	0	1.32 ± 0.13	1.75 ± 0.03	5.52 ± 0.68	5.26 ± 0.6	0	0
FL	Yersiniaceae	0	2.42 ± 0.02	0	0	0	0	0
R2	Actinomycetaceae	0	0	1.88 ± 0.1	0	0	0	0
R2	Aeromonadaceae	0	0	0	2.37 ± 0.08	0	0	0
R2	Alcanivoracaceae	0	0	0	3.91 ± 0.57	0	0	0
R2	Alteromonadaceae	2.36 ± 0.07	0	0	1.13 ± 0.34	3.51 ± 0.61	3.09 ± 0.6	10.23 ± 0.87
R2	Aphanizomenonaceae	0	0	0	13.61 ± 0.15	0	0	0
R2	Archaea	0	3.46 ± 0.13	5.64 ± 0.12	1.02 ± 0.21	4.35 ± 0.79	0	0.64 ± 0.04
R2	Bacillaceae	0	0	2.23 ± 0.18	0	0	0	0
R2	Burkholderiaceae	3.74 ± 0.18	0	0	0	0	39.68 ± 6.11	1.41 ± 0.05
R2	Cellvibrionaceae	0	0	0	2.06 ± 0.4	13.02 ± 12.37	0	5.45 ± 0.12
R2	Chromatiaceae	0	0	3.5 ± 0.1	0	0	0	0
R2	Chromobacteriaceae	0	0	0	0	1.74 ± 0.09	0	0
R2	Colwelliaceae	0	0	0	0	0	0.07 ± 0.03	0
R2	Coriobacteriaceae	0	3.15 ± 0.61	0	0	0	0	0
R2	Corynebacterineae	0	0	0	0	0	3.39 ± 0.23	0
R2	Crocinitomicaceae	0	0	0	0	4.52 ± 1.45	0	0
R2	Enterobacteriaceae	1.1 ± 0.09	5.35 ± 0.04	0	4.04 ± 0.19	7.33 ± 1.12	0	0
R2	Erythrobacteraceae	0	0	0	0	0	0	1.68 ± 0.36
R2	FCB cluster	5.96 ± 0.32	3.93 ± 0.12	2.84 ± 0.13	2.21 ± 0.34	1.41 ± 0.49	0.37 ± 0.15	12.57 ± 0.13
R2	Flavobacteriaceae	9.61 ± 2.28	6.56 ± 1.09	3.48 ± 0.17	0	15.91 ± 2.37	1.15 ± 0.2	12.24 ± 0.44
R2	Halieaceae	0	0	0	0	2.16 ± 0.8	0	0
R2	Moraxellaceae	0	0	0	0	0	0.88 ± 0.03	0
R2	Nitrosomonadaceae	0.82 ± 0.11	0	0	1.57 ± 0.07	0	0	0
R2	Oceanospirillaceae	0.88 ± 0.04	0	0	0	0.7 ± 0.19	0.23 ± 0.03	2.02 ± 0.15
R2	Other	14.91 ± 0.03	12.59 ± 0.41	11.06 ± 0.04	6.32 ± 1.04	5.59 ± 0.07	5.31 ± 0.27	7.46 ± 0.36
R2	other Actinobacteria	12.99 ± 0.01	3.08 ± 0.11	8.28 ± 0.07	1.43 ± 0.05	5.26 ± 0.67	6.65 ± 2.05	1.39 ± 0.14
R2	other Alphaproteobacteria	7.71 ± 0.6	4.79 ± 0.25	7.49 ± 0.13	1.11 ± 0.06	7.43 ± 1.64	6.82 ± 0.88	8.52 ± 0.15

R2	other Betaproteobacteria	6.44 ± 0.61	3.04 ± 0.03	4.65 ± 0.22	5.26 ± 0.13	4.72 ± 1.51	4.26 ± 0.15	1.86 ± 0.11
R2	other Cyanobacteria	0.98 ± 0.08	1.42 ± 0.28	2.56 ± 0.14	6.37 ± 0.42	0.92 ± 0.01	0.08 ± 0.01	1.76 ± 0.47
R2	other Gammaproteobacteria	14.69 ± 0.04	13.04 ± 0.44	9.96 ± 0.26	10.35 ± 1.09	7.8 ± 0.99	13.69 ± 1.85	15.29 ± 0.06
R2	other Proteobacteria	3.66 ± 0.15	2.67 ± 0.06	0.48 ± 0.01	4.44 ± 0.48	1.36 ± 0.63	1.85 ± 0.32	1.66 ± 0.07
R2	Oxalobacteraceae	1.2 ± 0.13	0	0	0	0	0	0
R2	Pelagibacteraceae	9.63 ± 0.35	0	0	0	0	9.11 ± 0.07	0
R2	Phyllobacteriaceae	0	0	6.69 ± 0.22	0	0	0	0
R2	Piscirickettsiaceae	0	0	2.52 ± 0.25	0	0	0	0
R2	Prochloraceae	0	0	4.18 ± 0.29	8.05 ± 0.39	0	0	0
R2	Pseudoalteromonadaceae	0.86 ± 0.05	1.11 ± 0.08	0	0	0	3.44 ± 0.23	3.33 ± 0.15
R2	Pseudomonadaceae	0	0	4.28 ± 0.14	0.82 ± 0.01	1.28 ± 0.41	0	0
R2	Rhizobiaceae	0	0	2.44 ± 0.21	0	0	0	0
R2	Rhodobacteraceae	0	9.18 ± 0.58	12.49 ± 0.33	0	1.88 ± 0.45	0	0
R2	SAR86	2.56 ± 0.46	0	0	0	0	0	0
R2	Shewanellaceae	0	23.28 ± 1.11	1.24 ± 0.1	6.65 ± 0.24	0.65 ± 0.12	0	8.14 ± 0.54
R2	Sphingomonadaceae	0	0	0	0	0	0	4.42 ± 0.12
R2	Synechococcaceae	0	1.12 ± 0.07	2.21 ± 0.04	17.36 ± 1.73	8.56 ± 2.87	0	0
R2	Yersiniaceae	0	2.31 ± 0.65	0	0	0	0	0

Supplementary Table 3. *Taxa-specific transcript abundance (per 10^5 cells) of a given pathway at Station F-L and R-2. For each group, mean value ± standard deviation of 2 replicates are shown. Error estimates are provided for all prokaryotic groups illustrated in Figures 6, and groups are listed by alphabetic order.*

Station	Species	Pathway	Mean ± SD
F-L	Actinomycetales	Aconitase	433.5 ± 55.87
F-L	Actinomycetales	F2+ uptake	39.81 ± 1.46
F-L	Actinomycetales	F3+ uptake	246.32 ± 14.14
F-L	Actinomycetales	Flavodoxin switch	9.4 ± 1.28
F-L	Actinomycetales	Isocitrate lyase	317 ± 7.08
F-L	Actinomycetales	Ribosomal proteins	24.1 ± 0.43
F-L	Actinomycetales	Siderophore uptake	2.09 ± 0.25
F-L	Alteromonadales	Aconitase	2804.42 ± 4.66
F-L	Alteromonadales	F2+ uptake	975.05 ± 29.47
F-L	Alteromonadales	F3+ uptake	1617.63 ± 24.8
F-L	Alteromonadales	Flavodoxin switch	961.79 ± 23.12
F-L	Alteromonadales	Isocitrate lyase	6742.55 ± 413.26
F-L	Alteromonadales	Ribosomal proteins	93235.65 ± 1342.4
F-L	Alteromonadales	Bacterioferritin	624.06 ± 8.33

F-L	Alteromonadales	Siderophore uptake	28433.39 ± 1163.07
F-L	Burkholderiales	Aconitase	146820 ± 21094.74
F-L	Burkholderiales	F2+ uptake	4002.1 ± 522.9
F-L	Burkholderiales	F3+ uptake	48495.46 ± 882.19
F-L	Burkholderiales	Flavodoxin switch	3505.73 ± 544.77
F-L	Burkholderiales	Isocitrate lyase	62830.09 ± 4936.64
F-L	Burkholderiales	Ribosomal proteins	314522.75 ± 48995.2
F-L	Burkholderiales	Bacterioferritin	2895.88 ± 107.54
F-L	Burkholderiales	Siderophore uptake	51850.49 ± 2905.22
F-L	Flavobacteriales	Aconitase	618 ± 2.83
F-L	Flavobacteriales	F2+ uptake	145.98 ± 7.65
F-L	Flavobacteriales	F3+ uptake	247.73 ± 21.77
F-L	Flavobacteriales	Flavodoxin switch	1.58 ± 0.59
F-L	Flavobacteriales	Isocitrate lyase	29.32 ± 1.95
F-L	Flavobacteriales	Ribosomal proteins	15925.98 ± 1438.44
F-L	Flavobacteriales	Bacterioferritin	85.51 ± 13.5
F-L	Flavobacteriales	Siderophore uptake	1809.17 ± 84.7
F-L	Oceanospirillales	Aconitase	1167.52 ± 0.07
F-L	Oceanospirillales	F2+ uptake	53.27 ± 1.51
F-L	Oceanospirillales	F3+ uptake	215.21 ± 12.69
F-L	Oceanospirillales	Flavodoxin switch	176.71 ± 6.33
F-L	Oceanospirillales	Isocitrate lyase	1239.78 ± 46.2
F-L	Oceanospirillales	Ribosomal proteins	40928.03 ± 290.1
F-L	Oceanospirillales	Bacterioferritin	101.06 ± 10.67
F-L	Oceanospirillales	Siderophore uptake	437.12 ± 2.45
F-L	Pelagibacterales	Aconitase	31.2 ± 0.43
F-L	Pelagibacterales	F3+ uptake	21.94 ± 1.08
F-L	Pelagibacterales	Isocitrate lyase	12.1 ± 0
F-L	Pelagibacterales	Ribosomal proteins	444.63 ± 4.14
F-L	Pseudomonadales	Aconitase	7598.56 ± 707.9
F-L	Pseudomonadales	F2+ uptake	2950.06 ± 291.64
F-L	Pseudomonadales	F3+ uptake	18795.22 ± 1129.27
F-L	Pseudomonadales	Flavodoxin switch	1675.58 ± 308.57
F-L	Pseudomonadales	Isocitrate lyase	40737.78 ± 1594.53
F-L	Pseudomonadales	Ribosomal proteins	441176.88 ± 3248.82
F-L	Pseudomonadales	Bacterioferritin	2417.7 ± 117.73
F-L	Pseudomonadales	Siderophore uptake	30927.55 ± 129.2
F-L	Rhodobacterales	Aconitase	155.47 ± 16.66
F-L	Rhodobacterales	F2+ uptake	80.35 ± 8
F-L	Rhodobacterales	F3+ uptake	461.17 ± 30.76
F-L	Rhodobacterales	Flavodoxin switch	8.9 ± 0.78
F-L	Rhodobacterales	Isocitrate lyase	40.91 ± 6.87

F-L	Rhodobacterales	Ribosomal proteins	4508.99 ± 712.13
F-L	Rhodobacterales	Bacterioferritin	21.15 ± 2.62
F-L	Rhodobacterales	Siderophore uptake	78.39 ± 0.19
F-L	Synechococcales	Aconitase	258.87 ± 17.67
F-L	Synechococcales	F2+ uptake	266.41 ± 32.84
F-L	Synechococcales	F3+ uptake	2748.01 ± 16.98
F-L	Synechococcales	Flavodoxin switch	563.17 ± 60.54
F-L	Synechococcales	Ribosomal proteins	105576.5 ± 6701.15
F-L	Synechococcales	Bacterioferritin	343.67 ± 30.96
F-L	Synechococcales	Siderophore uptake	1011.01 ± 30
R-2	Actinomycetales	Aconitase	42193.15 ± 5722.64
R-2	Actinomycetales	F2+ uptake	3194.07 ± 308.89
R-2	Actinomycetales	F3+ uptake	18782.59 ± 1570.95
R-2	Actinomycetales	Flavodoxin switch	617.87 ± 117.26
R-2	Actinomycetales	Isocitrate lyase	35417.43 ± 2302.06
R-2	Actinomycetales	Ribosomal proteins	5095.92 ± 161.52
R-2	Actinomycetales	Siderophore uptake	272.33 ± 40.42
R-2	Alteromonadales	Aconitase	71008.88 ± 10653.34
R-2	Alteromonadales	F2+ uptake	176175.89 ± 7129.55
R-2	Alteromonadales	F3+ uptake	52198.77 ± 161.71
R-2	Alteromonadales	Flavodoxin switch	47663.54 ± 4736.17
R-2	Alteromonadales	Isocitrate lyase	99877.15 ± 10599.23
R-2	Alteromonadales	Ribosomal proteins	275256.83 ± 8632.97
R-2	Alteromonadales	Bacterioferritin	7273.67 ± 136.36
R-2	Alteromonadales	Siderophore uptake	1082336.94 ± 46623.43
R-2	Burkholderiales	Aconitase	114020.04 ± 1920.06
R-2	Burkholderiales	F2+ uptake	8350.3 ± 823.95
R-2	Burkholderiales	F3+ uptake	105845.73 ± 1467.22
R-2	Burkholderiales	Flavodoxin switch	11413.09 ± 1024.97
R-2	Burkholderiales	Isocitrate lyase	427429.08 ± 66848.53
R-2	Burkholderiales	Ribosomal proteins	138255.51 ± 9595.33
R-2	Burkholderiales	Bacterioferritin	6040.08 ± 1212.87
R-2	Burkholderiales	Siderophore uptake	99233.46 ± 1133.47
R-2	Flavobacteriales	Aconitase	4657.41 ± 1319.02
R-2	Flavobacteriales	F2+ uptake	1117.31 ± 279.88
R-2	Flavobacteriales	F3+ uptake	2802.47 ± 173.62
R-2	Flavobacteriales	Flavodoxin switch	53.45 ± 40.38
R-2	Flavobacteriales	Isocitrate lyase	322.85 ± 47.01
R-2	Flavobacteriales	Ribosomal proteins	58880.52 ± 1988.95
R-2	Flavobacteriales	Bacterioferritin	862.24 ± 44.98
R-2	Flavobacteriales	Siderophore uptake	13630.66 ± 752.65
R-2	Oceanospirillales	Aconitase	4052.72 ± 426.93

R-2	Oceanospirillales	F2+ uptake	390.74 ± 104.7
R-2	Oceanospirillales	F3+ uptake	1541.95 ± 22.16
R-2	Oceanospirillales	Flavodoxin switch	1414.24 ± 158.46
R-2	Oceanospirillales	Iso	969.54 ± 221.75
R-2	Oceanospirillales	Ribosomal proteins	21389.3 ± 671.28
R-2	Oceanospirillales	Bacterioferritin	335.14 ± 53.06
R-2	Oceanospirillales	Siderophore uptake	6786.09 ± 801.14
R-2	Pelagibacterales	Aconitase	259.79 ± 18.48
R-2	Pelagibacterales	F3+ uptake	537.07 ± 26.72
R-2	Pelagibacterales	Isocitrate lyase	184.85 ± 0
R-2	Pelagibacterales	Ribosomal proteins	4193.84 ± 261.32
R-2	Pseudomonadales	Aconitase	1619.08 ± 34.74
R-2	Pseudomonadales	F2+ uptake	645.59 ± 266.08
R-2	Pseudomonadales	F3+ uptake	14064.86 ± 1143.46
R-2	Pseudomonadales	Flavodoxin switch	771.49 ± 37.5
R-2	Pseudomonadales	Isocitrate lyase	6623.61 ± 587.56
R-2	Pseudomonadales	Ribosomal proteins	47072.01 ± 805.73
R-2	Pseudomonadales	Bacterioferritin	354.52 ± 20.8
R-2	Pseudomonadales	Siderophore uptake	25575.43 ± 921.16
R-2	Rhodobacterales	Aconitase	5404.12 ± 67.84
R-2	Rhodobacterales	F2+ uptake	5531.83 ± 155.66
R-2	Rhodobacterales	F3+ uptake	22151.77 ± 1272.97
R-2	Rhodobacterales	Flavodoxin switch	400.02 ± 17.43
R-2	Rhodobacterales	Isocitrate lyase	708.99 ± 77.04
R-2	Rhodobacterales	Ribosomal proteins	78689.53 ± 2421.74
R-2	Rhodobacterales	Bacterioferritin	930.15 ± 100.77
R-2	Rhodobacterales	Siderophore uptake	5932.9 ± 680.68
R-2	Synechococcales	Aconitase	4082.61 ± 80.74
R-2	Synechococcales	F2+ uptake	3334.5 ± 389.07
R-2	Synechococcales	F3+ uptake	45383.65 ± 1140.23
R-2	Synechococcales	Flavodoxin switch	59113.52 ± 3779.32
R-2	Synechococcales	Ribosomal proteins	966188.35 ± 321.28
R-2	Synechococcales	Bacterioferritin	4810.25 ± 484.25
R-2	Synechococcales	Siderophore uptake	11052.27 ± 1219

Supplementary Table 4. *Information of databases constructed or modified from Toulza et al. (2012) and retrieved by NCBI.*

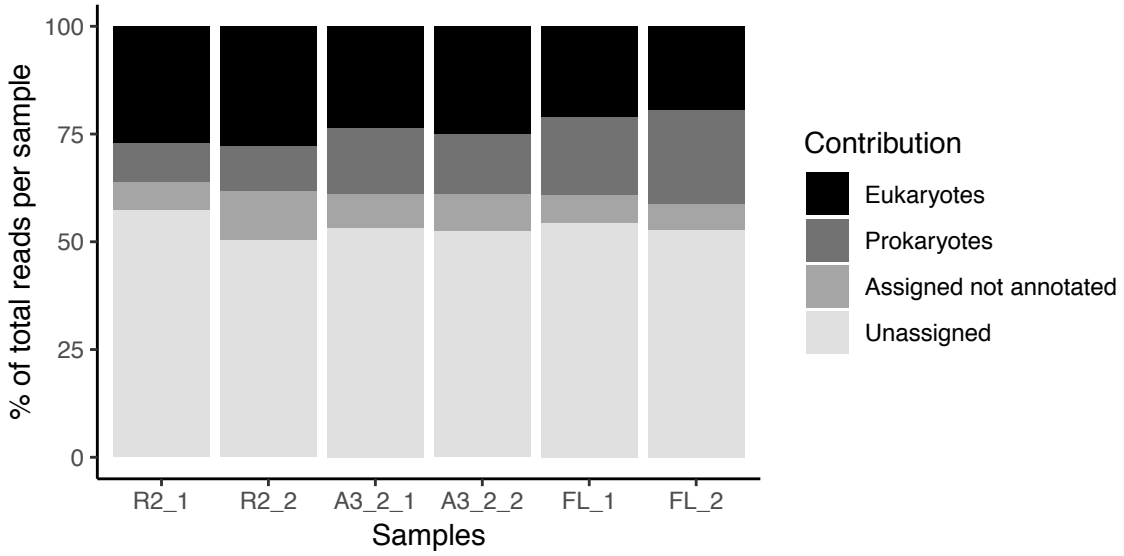
Database Name	Specific processes referred to	Sum retrieved from NCBI/Moore Database	Global Ocean Sampling (GOS) protein database retrieval	Size of created and functionally verified (KEGG) databases in Fasta sequences	Mean length of sequences in customized database (bp)
Aconitase	TCA cycle	527	807	1334	534
Isocitrate lyase	Glyoxylate shunt	1123	591	1714	412
Fe	Fe^{3+}	101	1150	1251	396
	Fe^{2+}	243	757	1000	513
	Siderophore uptake	561	2934	3495	513
	Flavodoxin switch	181	886	1067	209
	Fe Storage	123	427	550	170

Supplementary Table 5. *Pathways and corresponding KEGG-Id numbers that were chosen for further analysis. For Siderophore uptake only Fe related KEGG-Ids were chosen.*

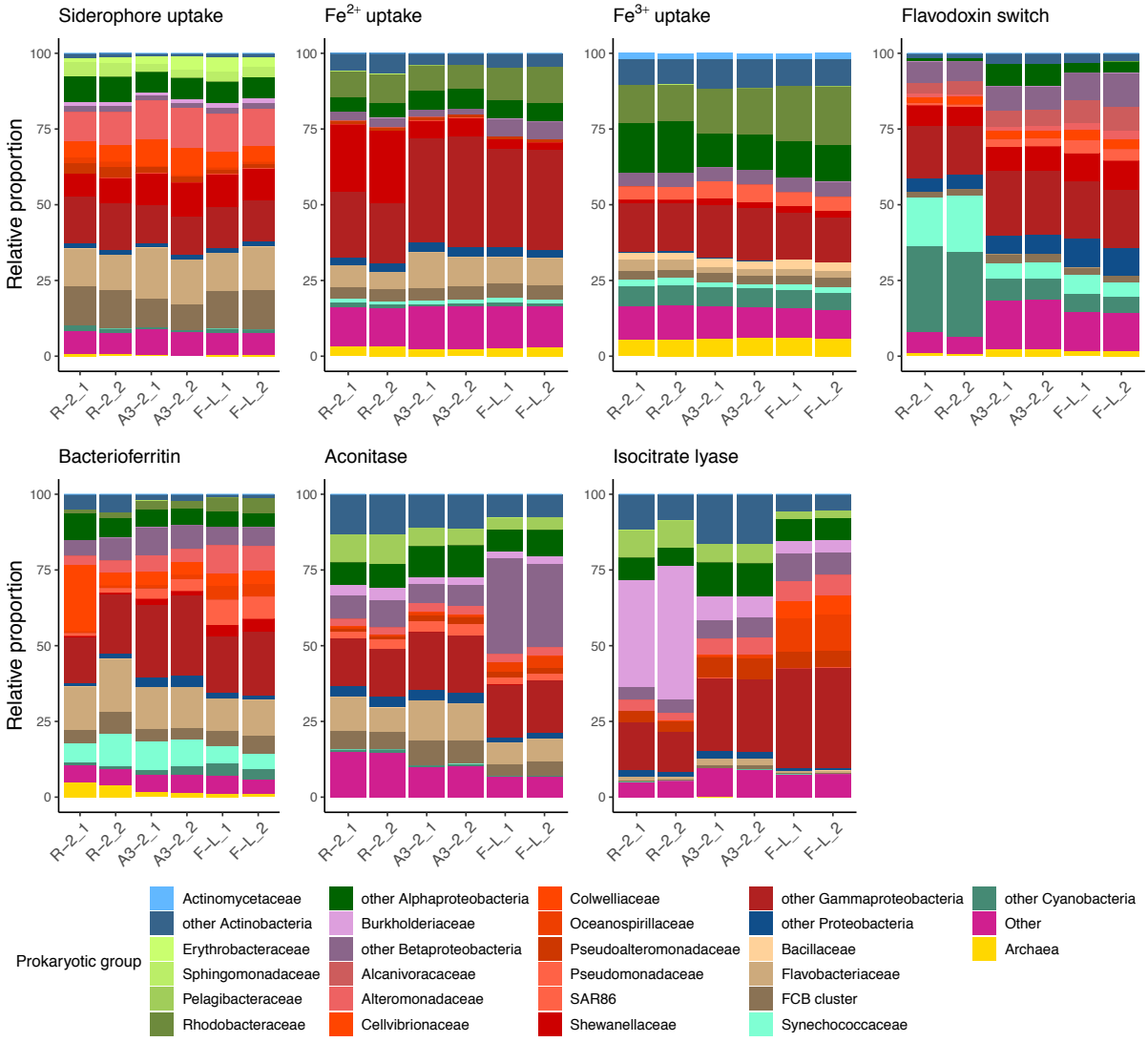
Gene product	**KEGG-ID**
Isocitrate lyase	K01637
Aconitase	K01681, K01682
Fe^{3+}	K02010, K02011, K02012, K02013, K11710
Fe^{2+}	K07224, K07243, K11604, K11605, K11606, K11607, K11707, K16301, K04758, K04759
Siderophore uptake	K02003, K02014, K02015, K02016, K02074, K02075, K02077, K03559, K03560, K03832, K06858, K16087, K16088, K16089
Flavodoxin switch	K00230, K00380, K03809, K03839, K03840, K05524, K06205, K12264
Fe Storage	K00522, K02217, K02255, K03594, K04047
Ribosomal proteins	K003010

Supplementary Table 6. *Cells per L for prokaryotic groups in Figure 6 for station R-2 and F-L, calculated as described in experimental procedures.*

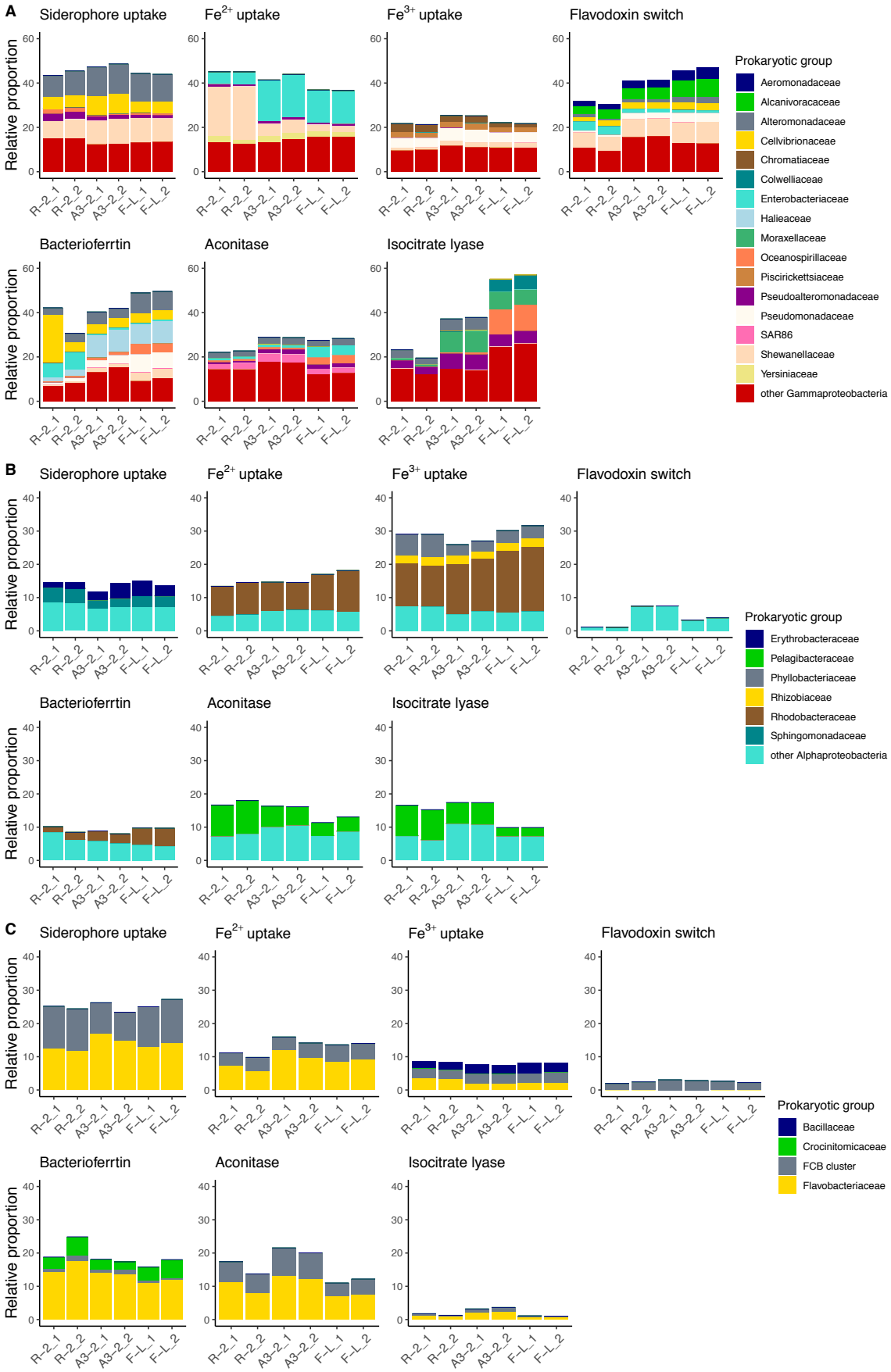
Prokaryotic Order	**R-2 cells L^{-1}**	**F-L cells L^{-1}**
Alteromonadales	1045863.3	13983087.5
Oceanospirillales	11147372.2	27410029.5
Pseudomonadales	4336506.3	1032841.7
Pelagibacterales	102775200.0	165731367.0
Flavobacteriales	21478460.8	58156932.1
Actinomycetales	1096881.0	28125073.7
Rhodobacterales	5433387.3	119253491.0
Burkholderiales	162973.3	110346.3
Synechococcales	440264.9	1715609.3



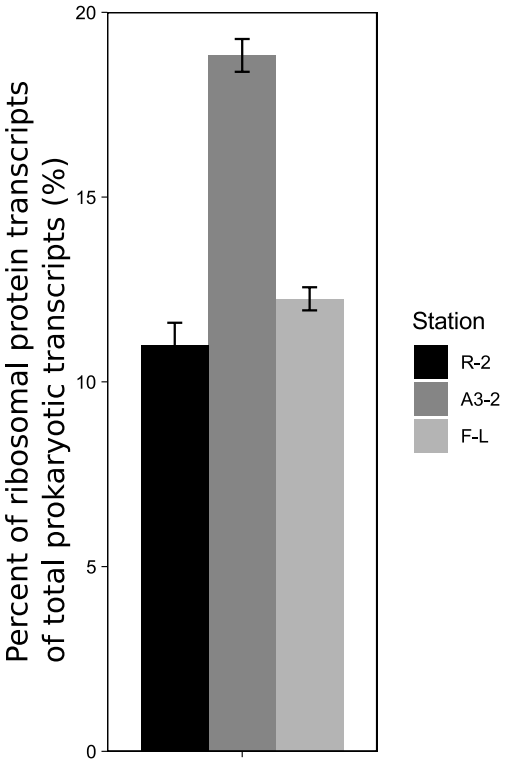
Supplementary Figure 1. *Blast results of subsets for assigned to non-assigned reads.*



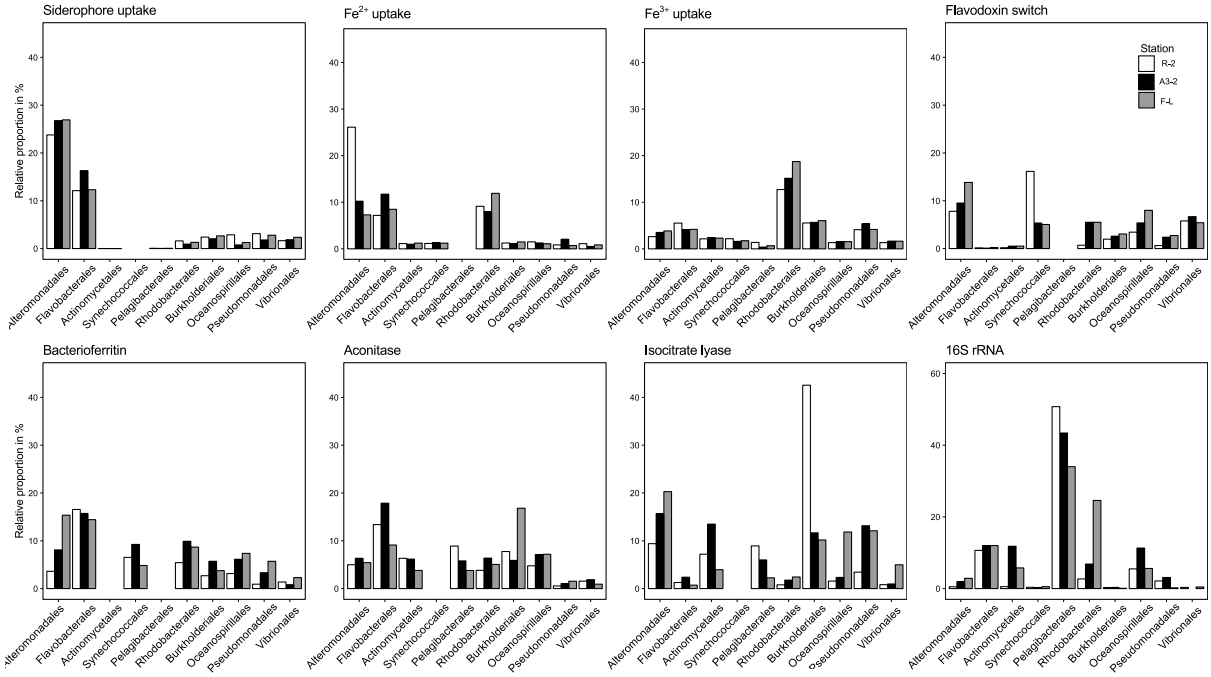
Supplementary Figure 2. *Relative contribution of prokaryotic groups to pathway specific transcripts. Prokaryotic group is defined until taxonomic family level. Number 1 and 2 refer to the sequencing results of duplicates per station.*



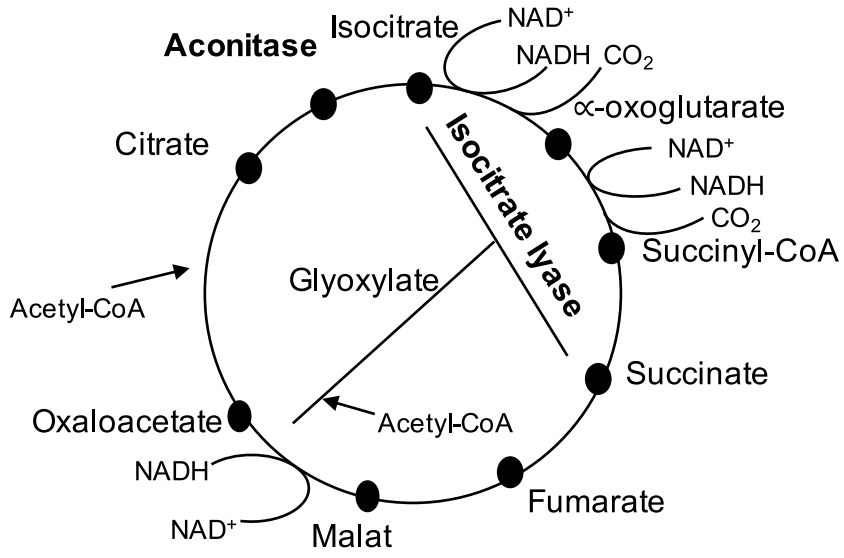
Supplementary Figure 3. A-C *Detailed view of relative contribution of Alpha- and Gammaproteobacteria and FCB cluster to pathway specific transcripts. Note the different y-axis for Gammaproteobacteria. Number 1 and 2 refer to the sequencing results of duplicates per station.*



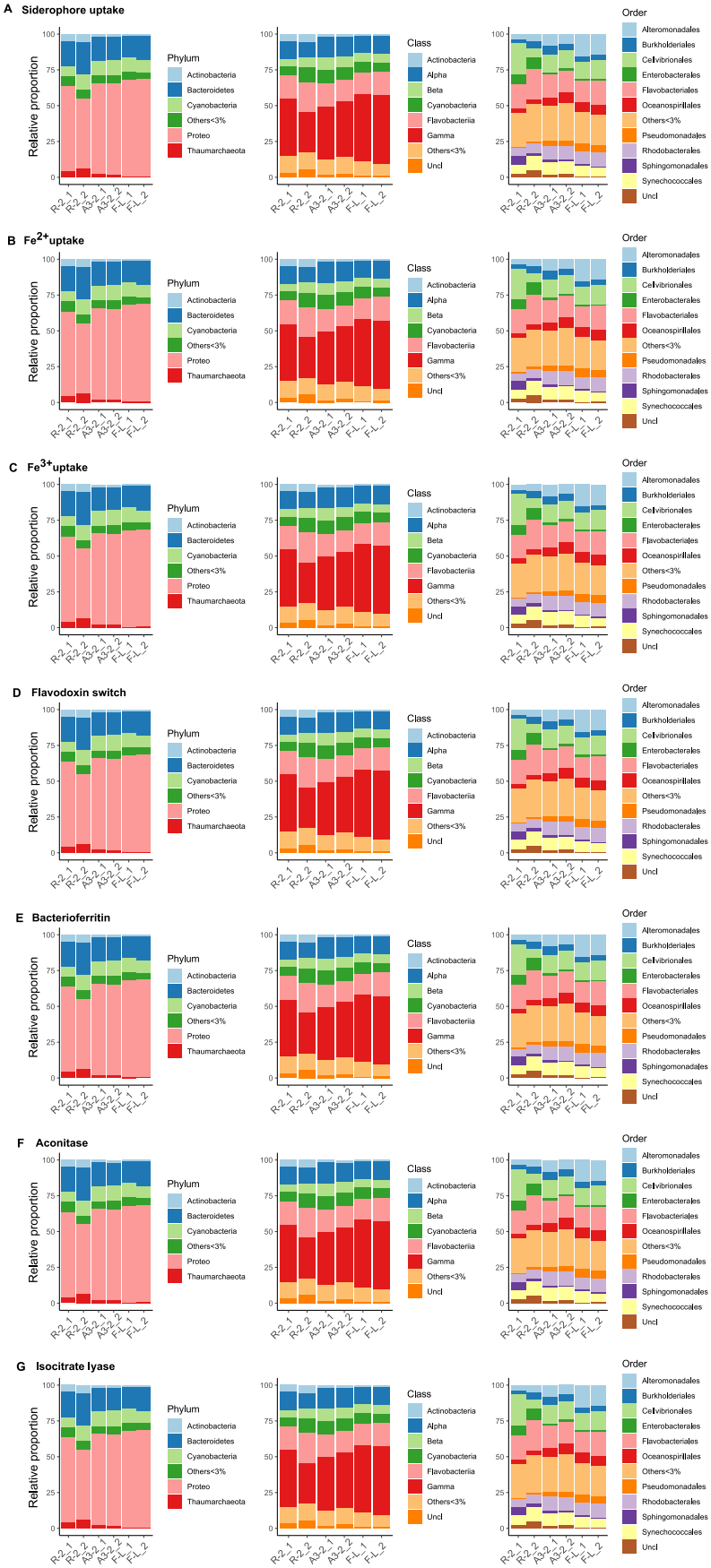
Supplementary Figure 4. *Percentage of ribosomal protein transcripts to all prokaryotic transcripts.*



Supplementary Figure 5 *Relative contribution of prokaryotic groups to pathway specific transcripts and relative abundance of 16S rRNA gene sequences. Prokaryotic group is defined until taxonomic order level.*



Supplementary Figure 6. *Simple illustration of the glyoxylate shunt (inside of the circle) in which isocitrate lyase cleaves isocitrate into glyoxylate and succinate.*



Supplementary Figure 7. *Detailed plots for duplicates on the phylum, class and order level for each pathway.*

Chapter 2

Marine particles as source of trace metals for prokaryotes:

A metatranscriptomic view

Pavla Debeljak[1,2], Stéphane Blain[1], Pier van der Merwe[3], Andrew R. Bowie[3,4] and Ingrid Obernosterer[1]

[1] Sorbonne Université, CNRS, Laboratoire d'Océanographie Microbienne, LOMIC, F-66650 Banyuls/mer, France

[2] Department of Limnology and Bio-Oceanography, University of Vienna, A-1090 Vienna, Austria

[3]Antarctic Climate and Ecosystems Cooperative Research Centre, University of Tasmania, Hobart, TAS, Australia,

[4]Institute for Marine and Antarctic Studies, College of Sciences and Engineering, University of Tasmania, Hobart, TAS, Australia,

Abstract

Trace metal availability in marine ecosystems controls biological activity in vast oceanic regions. The extremely low levels of essential metals have led microorganisms to develop efficient mechanisms of uptake and transformation for specific biochemical functions. While the amendment of the trace metal iron has shown to increase autotrophic and heterotrophic activity in large scale studies as well as in cultures, the expressional response in heterotrophic prokaryotes remains less understood. In this study we used metatranscriptomics on a prokaryotic community from a low iron Southern Ocean site after incubation with marine particles from the vicinity of MacDonalds Island. Bacteria and Archaea attached to particles as well as free-living cells incubated in the presence of particles were enriched in transcripts of transporters for multiple trace metals, siderophore synthesis and uptake, cellular carbon metabolism and several metal co-factors containing pathways. Our results illustrate that marine particles are a source of trace metals to heterotrophic prokaryotes in the iron-deplete Southern Ocean.

Introduction

Several elements from the first row of the transition metals (atomic mass above 50), are known to have an important biological role. They are often co-factors or part of co-factors in enzymes and structural elements in proteins, therefore being essential for living organisms. Of those, the trace metals —Mn, Fe, Co, Ni, Cu, Zn, and Cd— have been best studied by oceanographers (Donat and Bruland, 1995; Morel and Price, 2003). Investigations of their distributions and cycling in the ocean rely on the operational partitioning between dissolved and particulate (operationally defined by membrane filtration, cut-off of 0.2-0.4 μm) fractions. The dissolved pool is considered to be the main reservoir of bioavailable trace metals for microorganism as microbial uptake typically results in "nutrient like" vertical profiles with a clear depletion in surface waters for most trace metals (Morel *et al.*, 2003). In some regions dissolved trace metal concentrations are reduced to such low levels that microorganisms are limited, the most pronounced and studied case being iron (Fe). For example, the Southern Ocean, where our study was conducted is known to be the largest high nutrient low chlorophyll (HNLC) region due to the limitation in dissolved iron (dFe, < 0.2 μm) in surface waters (Martin *et al.*, 1990; Boyd *et al.*, 2000; Blain *et al.*, 2007).

Distributions of the trace metal particulate fraction (> 0.2 - 0.4 μm) in the ocean are highly variable because they are influenced by a large number of processes. When concentrations of particulate forms are compared to dissolved forms contrasted patterns emerge depending on the trace metal considered. For example, in the surface water of the central north Pacific Gyre, Bruland *et al.* (1994) reported ratios of dissolved to particulate metals of 82, 15, 1.9 and 0.9 for Mn, Zn, Cd and Fe, respectively. Thus, for some trace metals such as Fe, the particulate forms can represent a stock that equals or even exceeds the dissolved one (Boyd and Ellwood, 2010; van der Merwe *et al.*, 2019) and therefore might represent an important potential resource for microbes. This raises the question of the bioavailability of particulate trace metals.

This question has thus far been less studied compared to the bioavailability of the dissolved pool. It is not straightforward to address because the nature and the origin of the particles present in the ocean are extremely diverse and variable. Marine particles can be divided by their origin. Particles can derive from lithogenic external sources such as those resulting from atmospheric deposition, from delivery by rivers, glaciers and hydrothermal vents, and sediment re-suspension (Blain and Tagliabue, 2016). The biological production of particles is an additional important source, and authigenic precipitation can also be a significant source for some trace metals. The suspended particles (SP, operationally defined as the fraction 0.2 < SP < 53 μm) can abiotically or biologically aggregate into larger particles (LP, > 53 μm)

that sink to depth (Gardner *et al.*, 1993; Steinberg *et al.*, 2008; Armstrong *et al.*, 2009; Sigman and Hain, 2012; Herndl and Reinthaler, 2013). During their transfer in the water column both exchange with the dissolved pool, and differential remineralization length of the trace metals contribute to modifying their composition (Boyd *et al.*, 2017). All these processes lead to the final complex composition of trace metals in marine particles.

Geochemical approaches were proposed to split the trace metal content of particles in different pools with the ultimate goal to relate them to their reactivity or bioavailability. These approaches rely mainly on leaching protocols that sequentially separate different fractions and were recently compared (Rauschenberg and Twining, 2015). These authors recommend using the acetic-acid hydroxylamine leach to estimate the biogenic fraction and the associated labile particulate phases. They suggest that this fraction is likely cycled through biota, an assumption that implies that this fraction or part of it is bioavailable. However, it is important to keep in mind that the so-called refractory fraction that is left after this leaching protocol might also contain phases that are likely bioavailable.

Besides this pure geochemical view of bioavailability, a more biologically oriented approach was proposed which relies on incubations of ambient microbial communities with particles of different origin. Changes in the chemical composition of the media and of the microbial community activity during the incubation can provide some insight into the bioavailability of the particles. For example, the release of Fe-binding ligands coupled with an increase in dFe has been observed when prokaryotes were incubated with particles (Boyd *et al.*, 2010; Velasquez *et al.*, 2016).

A third perspective to investigate the bioavailability of particles is to look at the metabolic response of the cells in natural environments that are enriched in particles. The expression of microbial Fe transporters was observed in waters close to a hydrothermal plume suggesting a "microbial iron pump", rendering hydrothermal Fe bioavailable through microbial processing (Li *et al.*, 2014). Alternatively, the metabolic profiles of natural microbial communities or strains incubated in different treatments can also be used. While several studies have tackled carbon and nitrogen cycling in marine systems (Baker *et al.*, 2013; Gifford *et al.*, 2016; Li *et al.*, 2016) and the use of different carbon compounds (Beier *et al.*, 2015; Landa *et al.*, 2017), to our knowledge this approach has not been used to focus on metal availability.

Finally, another important feature to consider when investigating the bioavailability of trace metals in particles is that they are inhabited by prokaryotes. Prokaryotic activity on marine particles is shown to be elevated in terms of extracellular enzymes and metabolic functioning compared to the free-living community in multiple studies (Karner and Herndl, 1992; DeLong

et al., 1993; Simon *et al.*, 2002; Azam and Malfatti, 2007; Satinsky *et al.*, 2010; Simon *et al.*, 2014). The complex microstructure of marine particles can also influence prokaryotic community composition as observed through 16S rRNA and metagenomic surveys (Smith *et al.*, 2013; Ganesh *et al.*, 2014). While a clear separation of particle attached and free-living communities has been proposed (DeLong *et al.*, 1993), recent studies suggest that processes on particles are more dynamic. The presence of particulate matter enhances heterotrophic productivity and this can be due to multiple factors. Particles are proposed to have plumes as they sink and can therefore be hotspots for prokaryotes that are able to actively sense these (Kiørboe, 2001; Stocker *et al.*, 2008; Stocker, 2012). These processes are happening on the microscale and thus the 100 μm length scale has been proposed as it characterizes the chemical, physical and biological processes happening in the surrounding of a prokaryotic cell (Stocker, 2015). These plumes or spheres itself produce gradients from which not directly attached microorganism can profit.

In the present study we were particularly interested to investigate how important particles are as a source of trace metals to microorganisms in the Southern Ocean. The specific questions of the present study were i) Are marine particles a source of available trace metals for heterotrophic prokaryotes? and ii) How are the particle-derived trace metals used for prokaryotic metabolism?

To address these questions, an experiment was designed to investigate prokaryotes attached to particles (particle – attached, PA), present in the particle-sphere (free-living particle-sphere, FLP) and those out of reach of marine particles (free-living control, FLC) (Fig. 1). The three different approaches described above were applied to describe the bioavailability of trace metals with a focus on Fe, the proximate limiting trace element of the Southern Ocean microbial community. The particles were characterized by the geochemical approach, the bulk response of the prokaryotic community was followed during incubations with and without particles, and the messenger RNA of microbial communities at the end of the incubations was sequenced to provide a functional fingerprint of gene expression in the different fractions (PA, FLP and FLC).

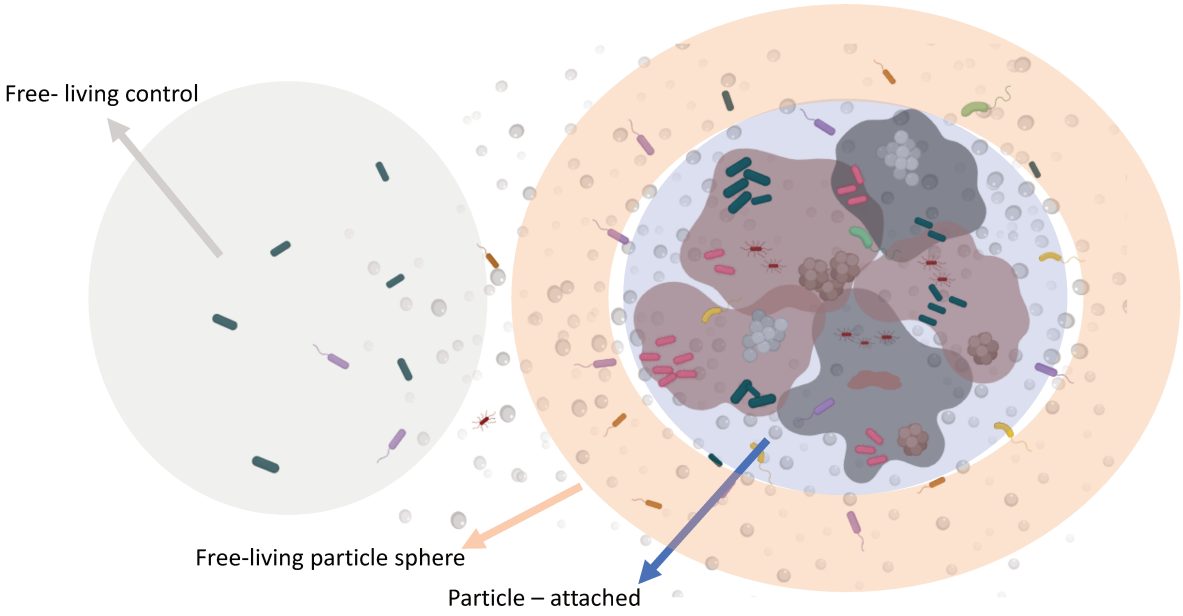


Figure 1. *Conceptual visualization of the experimental design. Particle-attached prokaryotes (>0.8 µm size fraction); Free-living prokaryotes in particle-sphere (<0.8 µm size fraction in the presence of particles); Free-living prokaryotes in the absence of particles*

Material and methods

The experiment was performed during the Heard Earth-Ocean-Biosphere Interactions (HEOBI) cruise (8 January - 27 February 2016) aboard the Australian R/V *Investigator* (INV2016-V01). Particles were collected on January 29th 2016 in shallow waters (overall depth 210 m) in the vicinity of Mac Donald Island (53.0349 S, 72.55 E) at 100 m by McLane high volume in-situ pumps (WTS-LV) (here on referred to as ISPs) (Fig. 2).

Sample handling, collection and processing was performed in accordance with the general GEOTRACES protocols (http://www.geotraces.org/) and specific methodologies outlined in (Bowie *et al.*, 2010). Briefly, suspended particles for the incubations and the trace metal analysis were collected using paired, acid washed, 142 mm, 0.8 µm Supor filters to obtain a 0.4 µm effective pore size (Bishop *et al.*, 2012). The pumped volume was 96 L. The top filter was then split into punches of 47 mm diameter each. For the experimental incubations, particles were rinsed off from 2 filter punches with each 20 mL 0.2 µm-filtered seawater collected following trace-metal clean procedures (Holmes *et al.*, 2019). The remaining punches served for chemical analyses of the particulate matter as described in Van der Merwe et al. (van der Merwe *et al.*, 2019).

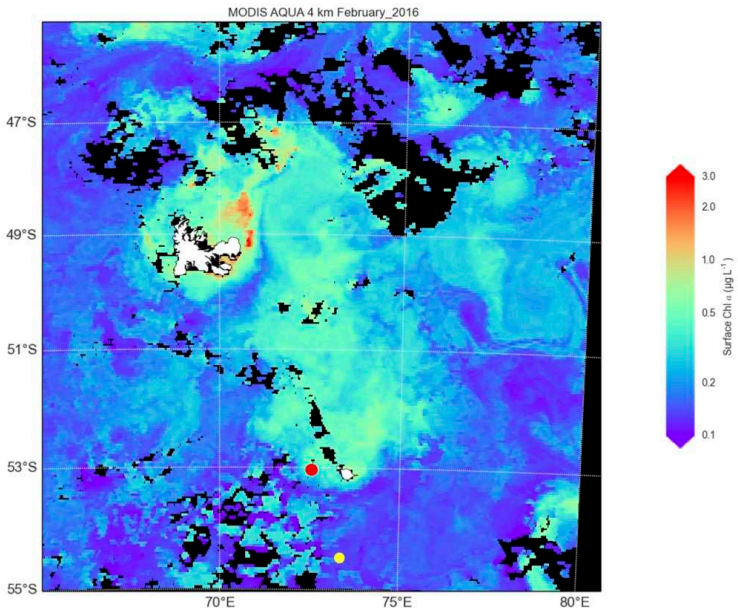


***Figure 2**. Sampling location of particles (red dot) close to Mac Donald Island and of seawater used for the incubations (yellow dot) in HNLC waters. Chlorophyll map of 4 Km resolution for February 2016.*

Experimental design

To investigate the effect of suspended particles on prokaryotes, on-board incubation experiments were performed (Fig. 3).

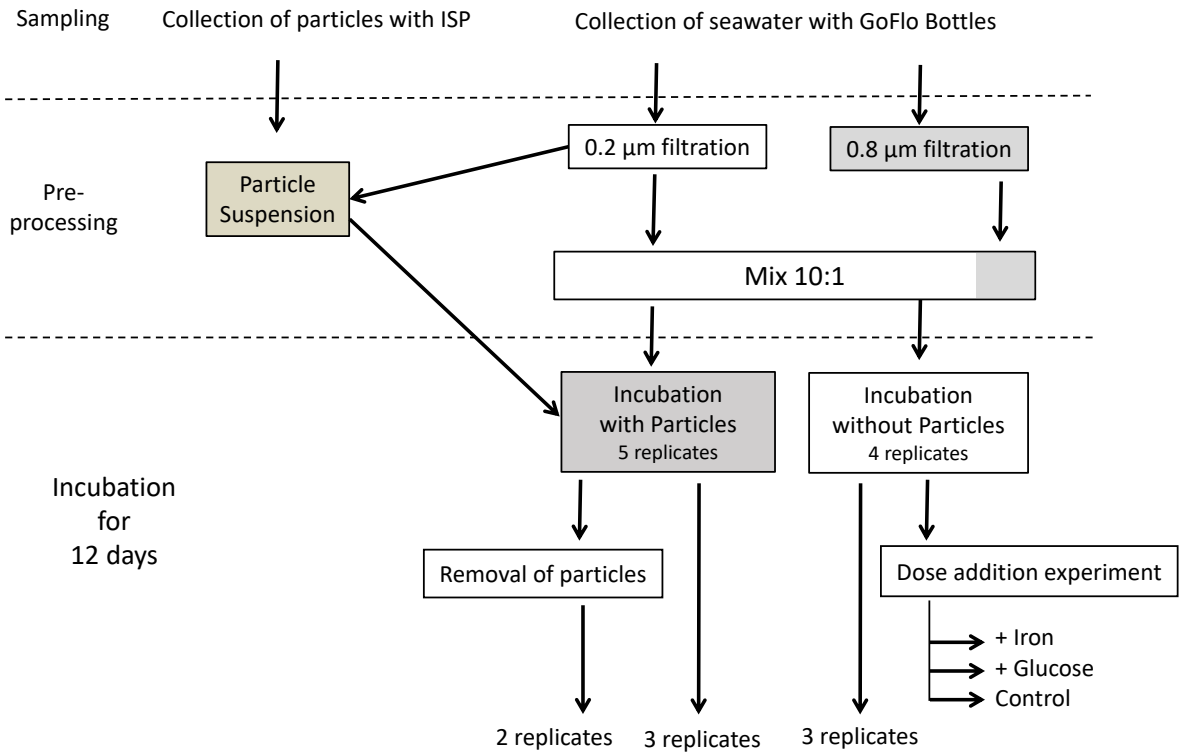


Figure 3. *Experimental design of the on-board incubations. ISP-In situ pump*

Batch cultures, containing a natural prokaryotic consortium (< 0.8 µm size fraction) and 0.2 µm-filtered seawater (ratio of 1:10, final volume 2L), were amended with particles and incubated for 12 days. Seawater for the < 0.2 µm- and < 0.8 µm-size fractions was collected at 45 m depth from HNLC waters south of Heard Island (54 10.061 S/ 73 40.092 E) (Fig. 2), using trace-metal clean procedures. Particles were added to the cultures from the suspension described above (6.5 mL of suspension to 2L of batch culture). Batch cultures without particle addition served as control. Because of the expected variability induced by the addition of the particle suspension, five replicate particle-amended cultures were prepared. The control incubations were done in quadruplicates. To investigate short- and medium-term dynamics of the suspended matter, particles were removed after 7.7 days of incubation in two of the five replicate particle-amended cultures. The cultures were therefore filtered through an 0.8 µm filter, and the incubation containing < 0.8 µm filtered seawater continued until the end of the experimental period (Fig. 3). To keep the particles in suspension, each culture was continuously stirred at slowest speed using a magnetic stirrer. The cultures were kept in polycarbonate

carboys, in the dark in a temperature-controlled incubator at 8°C. All steps of the preparation of the cultures were done under a trace metal clean laminar flow hood (ISO class 5).

During the 12d incubation period, the cultures were subsampled every 2-3 days for prokaryotic heterotrophic abundance and production. Both parameters were determined for the bulk community (unfiltered water from the batch cultures) and in the < 0.8 µm size fraction, the latter considered as free-living community. The difference in prokaryotic heterotrophic production between these fractions was considered as particle-attached. Prior to microbial measurements, subsamples of each culture were filtered through an 0.8 µm polycarbonate filter (25 mm diameter), that was treated with a mild HCl-wash (0.1 N) and several rinses with Milli-Q water. At the end of the experiment, additional samples were taken for the determination of dissolved and particulate organic carbon (DOC and POC), prokaryotic diversity and metatranscriptomics.

Dose addition experiment

To determine whether Fe or organic carbon (C) were the limiting factors for prokaryotic heterotrophic activity in the control incubations, we sacrificed one of the 4 control incubations after 8 days to perform a dose addition experiment. We therefore split the control incubation into the following three treatments: +Fe, +C and an unamended control (Fig. 3; + Iron, + Glucose, Control). Fe was added as $FeCl_3$ (final concentration 1 nM of $FeCl_3$), and C was added as trace-metal clean glucose (final concentration 6 µM of glucose). To eliminate trace metal contamination, the working solution of glucose was passed over a Chelex 100 ion exchange resin (Bio-rad, 200-400 mesh). Triplicate incubations (250 mL each) were done for the Fe- and C-amended treatments. Because of the high reproducibility in prokaryotic heterotrophic abundance and production in the cultures without particles over the 8 days of incubation, a single incubation bottle served as control for the dose addition experiment. The incubations were done in the dark in the same temperature-controlled incubator as the particle-incubation experiments. For subsampling, incubation bottles were opened under a laminar flow hood (ISO class 5). Prokaryotic abundance and production were determined after 1 and 3 days of incubation.

Particulate organic carbon concentrations

In situ suspended particles for elemental CHN analysis were collected using pre-combusted (450° C for 12 h) quartz micro-fibre (QMA, Sartorius) filters with a nominal 1 µm pore size. 16 mm subsamples were taken from the 142mm QMA filters and frozen until analysis at -20° C. Total carbon, nitrogen and hydrogen were determined using a Thermo Finnigan EA 1112 Series Flash Elemental Analyzer (estimated precision ~ 1%). In the cultures, subsamples (500 mL) for particulate organic carbon analyses were collected on pre-combusted GF/F filters (25 mm diameter). Samples were analysed with a Carlo Erba Instruments EA1108 elemental analyser using an acetanilide standard as reference.

Enumeration of heterotrophic prokaryotes and determination of particle size distribution

Heterotrophic prokaryotic abundance of the bulk community and in the < 0.8 µm fraction were enumerated by flow cytometry. Subsamples (1.8 mL) were fixed with glutaraldehyde (1% final concentration), incubated for 30min at 4°C and stored at -80°C until analyses on a BD FACS Canto (Marie *et al.*, 2000) equipped with an air-cooled laser, providing 15mW at 488 nm with the standard filter set-up. Heterotrophic prokaryotes were stained with SYBRGreen I, and determined by flow cytometry as described in detail in (Obernosterer *et al.*, 2008). For microscopic observations of the particles and the enumeration of particle-attached cells, the 0.8 µm polycarbonate filters that served for the size-fractionation (see Experimental design) were placed on pads (Millipore) soaked with 5% PFA overnight at 4°C. The filters were then rinsed with 0.2 µm filtered Milli-Q water, dried and stored in petri dishes at -20°C. Back in the home lab, each filter was split into about 6 pieces, and stained with DAPI for microscopic observations using an Olympus AX70 epifluorescence microscope equipped with an image analyses system as described in Cottrell and Kirchman (Cottrell and Kirchman, 2000). For each observation, two images were acquired using excitation with white and blue light. The images were then processed with the software image J. The blue image was used to count by eye the cells attached to the different particles. The white image was split into the three channels (Green, Blue and Red). The red one presented the best characteristics (brightness, contrast) for the determination of the surface of the particles. The particles were contoured and the surface determined after calibration of the pixel size. For each replicate, at least 100 particles were observed on different randomly selected parts of the filter. For each particle, the surface and the number of attached prokaryotic cells was recorded to estimate the size distribution of the particles and the cell areal abundance.

Prokaryotic heterotrophic production

Prokaryotic production was estimated by [^{3}H]leucine incorporation applying the centrifugation method (Smith and Azam, 1992). Briefly, 1.5-mL samples were incubated with a mixture of [3,4,5-^{3}H(N)] leucine (Perkin Elmer, 123.8 Ci mmol^{-1}; 7 nM final concentration) and nonradioactive leucine (13 nM final concentration). Controls were fixed with trichloroacetic acid (TCA; Sigma) at a final concentration of 5%. Samples were incubated for 2-3h in the dark. Incubations were terminated with TCA (5% final concentration). The radioactivity incorporated into prokaryotic cells was measured aboard in a HIDEX 300 SL liquid scintillation counter.

Sampling for prokaryotic diversity and metatranscriptomics

At the end of the incubation, 500 mL of each replicate culture were sequentially filtered on 0.8 µm and 0.2 µm polycarbonate filters, and both filters were stored at -80°C until further diversity and metatranscriptomic analysis. The cells extracted from the 0.8 µm filters were considered as particle-attached (PA). The cells extracted from the 0.2 µm filters of the same incubation were considered as free-living in the presence of particles (FLP). The cells extracted from the 0.2 µm filters of the incubations without particles were considered as free-living control (FLC). The cells extracted from the 0.2 µm filters obtained from the cultures from which particles were removed after 7.7 days of incubation were also considered as free-living controls. This treatment is referred to as FLC-P in Fig. 8 and in Suppl. Table 1 and 2. In order to calculate the concentration of mRNA standard that needed to be added, two control samples (FLC) were used for trial extractions. Therefore, the control triplicates consisted of one FLC and two FLC-P samples (Fig. 8 and Suppl. Table 2). In total 9 samples (PA, FLP and FLC each in triplicates) were sequenced for metatranscriptomics.

Nucleic acid extraction

In order to enhance efficiency prior to extraction, each filter (Whatman 47 mm) was cut into smaller pieces of several millimetres on a sterile petri dish with a sterile razor blade. The pieces were transferred back to the Eppendorf tube (2 ml) where lysis buffer (ZYMO research) was added immediately. DNA from all experiment samples (n = 14) was extracted using a ZR Fungal/Bacterial DNA MiniPrepTM Kit (ZYMO research) following the manufacturer's guidelines with the addition of a 45 min lysozyme digestion at 37°C and a 1 h Proteinase K digestion (20 mg ml^{-1} final concentration) at 55°C.

Amplicon sequencing

DNA was amplified and pooled as described in Parada et al. (2016) with a modification to the PCR amplification step. Briefly, the V4 - V5 region of the 16S rRNA gene from DNA samples was amplified with the primer sets 515F-Y (5' - GTGYCAGCMGCCGCGGTAA) and 926R (5' - CCGYCAATTYMTTTRAGTTT). Triplicate 10 µl reaction mixtures contained 2 µg DNA, 5 µl KAPA2G Fast HotStart ReadyMix, 0.2 µM forward primer and 0.2 µM reverse primer. Cycling reaction started with a 3 min heating step at 95 °C followed by 22 cycles of 95 °C for 45 s, 50 °C for 45 s, 68 °C for 90 s, and a final extension of 68 °C for 5 min. The quality of amplification products was examined on a 1 % agarose gel electrophoresis and when bands were visible triplicate reactions were pooled. Unique paired barcodes were added to each sample. The Master 25 µl mixtures contained 1 µl PCR product, 12.5 µl KAPA2G Fast HotStart ReadyMix (Kapa Biosystems, USA), 0.5 µl barcode 1 and 0.5 µl barcode 2. The cycling program included a 30 s initial denaturation at 98°C followed by 8 cycles of 98 °C for 10 s, 60 °C for 20 s, 72 °C for 30 s, and a final extension of 72 °C for 2 min. 3 µl PCR product were used to check for amplification on 1% agarose gel electrophoresis. The remaining 22 µl barcoded amplicon product was cleaned to remove unwanted dNTPs and primers by Exonuclease I and Shrimp Alkaline Phosphatase at 37 °C, 30 min for treatment and at 85 °C, 15 min to inactivate. The concentration of double-stranded DNA was quantified by PicoGreen fluorescence assay (Life Technologies). After calculating the PCR product concentration of each samples, they were pooled at equal nanogram manually. The pooled PCR amplicons were concentrated using Wizard SV gel and PCR clean-up system (Promega, USA) according to the manufacturer's protocol. Amplicons were sequenced with MiSeq Illumina 2 × 300 bp chemistry by Fasteris SA sequencing service (Switzerland) in a batch of 300 samples on one lane. Mock community DNA (LGC standards, UK) was used as a standard for subsequent analyses and considered as a DNA control for all treatments. All samples from the platform have been demultiplexed and barcodes have been trimmed off. A total number of 764,658 raw sequences was obtained for 14 samples.

RNA Extraction with customized standards

RNA was extracted according to the protocol of Roey (2012) with the following modifications for the use of filters. Tubes containing the filters (48 mm filters, 2ml Eppendorf tubes) were submerged in liquid nitrogen and ground into small pieces with a metal spatula. Phosphate buffer (750 µL, 120 mM, pH 5.5) and 250 µL TNS buffer (500 mM Tris base, 100 mM NaCl, 10% SDS) and internal standards (described below) were added to the filter pieces, and cells were lysed with 5x freeze and thaw cycles. After centrifuging at 20 000 x g at 4°C for

3 min, the supernatant was added to a fresh tube and one volume of TE saturated phenol (pH 5.5) was added. RNA was purified using standard phenol/chloroform/isoamyl alcohol and chloroform/isoamyl alcohol purification, and the supernatant was transferred to non-stick RNase-free microfuge tubes (Eppendorf). Nucleic acids were precipitated with 30% polyethylene glycol (PEG) in 1.6 M NaCl and 2 μL glycogen (Thermo Scientific), washed once with ice-cold 75% EtOH and resuspended in low TE buffer. DNA was digested with TURBO DNase (Ambion) and purified using the RNeasy MinElute Cleanup Kit (Qiagen). Complete DNA removal was validated via PCR amplification of 30 cycles and quality of RNA was checked with Bioanalyzer RNA chip profiles (Agilent Technologies, Santa Clara, CA, USA). Ribo-Zero Gold rRNA Removal Kit (Epidemiology) was used to remove cytoplasmic and mitochondrial ribosomal RNA and sequencing was performed in pair-end using HiSeq (3000/4000) Illumina 2 × 150 bp chemistry on one full lane by Fasteris SA.

Internal standard addition

Prior to the lysis step of the extraction, known copy numbers of two artificial internal mRNA (MTST5 = 3.70 x 108 molecules; MTST6 = 3.72 x 108 molecules) standards ~ 1000 nt in length were added to each sample individually (Satinsky *et al.*, 2013); and dx.doi.org/10.17504/protocols.io.ffwbjpe). Standards were synthesized using custom templates that were transcribed in vitro to RNA (Satinsky *et al.*, 2013).

Amplicon analysis

The analysis of the raw sequences was done by following the standard pipeline of the DADA2 package (Callahan et al. 2016; https://benjjneb.github.io/dada2/index.html, version 1.6) in 'R' (R Core Team (2019), https://www.R-project.org/) with the following parameters: trimLeft = c(19, 20), maxN = 0, maxEE = c(2, 5), truncQ = 2. Briefly, the package includes the following steps: filtering, dereplication, sample inference, chimera identification, and merging of paired-end reads. DADA2 infers exact amplicon sequence variants (ASVs) from sequencing data, instead of building operational taxonomic units from sequence similarity. In total, from 14 samples we obtained an average of 23,000 reads per sample (Supplementary Table 1). The sequence data were normalized by dividing counts by sample size. The taxonomy assignments were done with the SILVA v.128 database (https://www.arb-silva.de/documentation/release-128/) and the "assignTaxonomy" function in DADA2 that implements the RDP naiveBayesian classifier method described in Wang et al. (2007). Visualization of results was performed in

phyloseq package in R (McMurdie and Holmes, 2013) and plots were enhanced using the open source tool inkscape (https://inkscape.org/).

Metatranscriptome read processing, assembly and annotation

Nine metatranscriptome shotgun libraries were prepared using a strand-specific protocol and sequenced on the Illumina's HiSeq 3000/4000 system (Fasteris SA, Inc.). An initial round of read processing was provided by the company, using trimmomatic (Bolger *et al.*, 2014), an integrated quality-control tool for high-throughput Illumina NGS data (Fig. 4). The standard Illumina adapters and low-quality bases were removed with the following parameters "2:30:10 SLIDINGWINDOW:4:5", and reads for which no insert was found or ambiguities were also dropped, resulting in 91.84 - 97.12% of 2×150 bp paired-end reads (Supplementary Table 2). Further, we performed another round of quality control and refinement with Trim Galore v0.5.0 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) for automatic detection and removal of overrepresented adapters, using a phred score of 20 and a length cutoff of 50 bp. To focus on protein-coding RNAs, we computationally eliminated ribosomal RNA and internal standards-derived reads using SortMeRNA v 2.1b (Kopylova *et al.*, 2012). A BLASTN homology search against a custom database (Altschul *et al.*, 1990), which consists of representative prokaryotic rRNAs and tRNAs from NCBI RefSeq along with the template sequences of internal standards, was also implemented using a bit score cut-off of 50 as suggested in previous studies (Satinsky *et al.*, 2013).

Given that similar starting communities were observed among samples according to amplicon sequencing data (Supplementary Fig. 4), all reads from the 3 samples (9 triplicates) were pooled together and used for co-assembly in MegaHit v1.2.7 (Li *et al.*, 2015) with default parameters. The assembly yielded 819,328 contigs with an N50 of 577 bp. Contigs shorter than 300 bp were discarded from further analysis, and 732,369 contigs were retained. Genes were predicted using Prodigal v2.6.3 (Hyatt *et al.*, 2010). The resulted proteins with more than 100 aa were taxonomically annotated by DIAMOND BLASTP against nr database (January 2019) with options "--more-sensitive -f 100 -k 1 -p 10" (Buchfink *et al.*, 2015), and functionally annotated using eggNOG 5.0 (Huerta-Cepas *et al.*, 2016, 2019) and GhostKOALA (Kanehisa and Goto, 2000; Kanehisa, Sato, and Morishima, 2016) with default parameters. The taxonomic affiliation of each protein was inferred based on its top hit with the aid of MEGAN6 (Huson *et al.*, 2007). A total of 420,218 prokaryotic protein sequences (66% of all, 9% eukaryotic and 15% unassigned) were finally retrieved, among which 65 % could be assigned to KEGG orthologous groups (KO) (Kanehisa, Sato, Kawashima, *et al.*, 2016).

Gene differential expression and functional enrichment analyses

Reads were mapped back to the assembled contigs using Bowtie2 (Langmead and Salzberg, 2012) with default parameters (Fig. 4), resulting an overall 54.3% recruitment rate. Reads that were aligned to the selected prokaryotic gene features were counted using featureCounts (http://bioinf.wehi.edu.au/featureCounts/) with the following parameters for stranded specific reads -F 'GTF' -t 'gene' -Q 1 -s2 -p -C (Liao *et al.*, 2014). The R package DESeq2 (Love, Huber et al. 2014) was implemented to detect differential expression based on the count table generated by featureCounts. The control sample FLC-P_14 (Suppl. Table 2) had to be excluded from the further analysis as it was significantly different from the triplicate set recruiting less reads than all other samples (DESeq function Cooks distance). Differential expression analysis was thus calculated with a duplicate set of control samples.

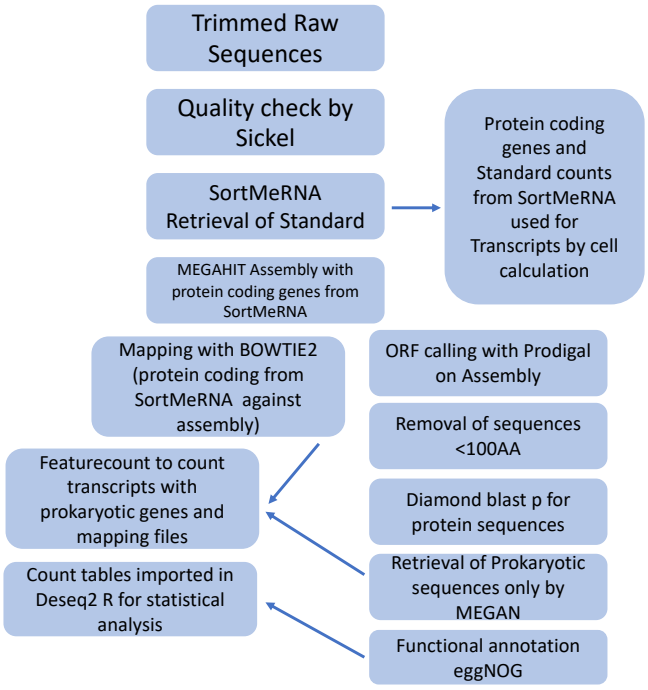


Figure 4. *Bioinformatic pipeline for the analysis of metatranscriptomes.*

Correction for cell abundance

Enumeration of prokaryotes revealed about 10 times higher cell abundance of free-living than particle-attached cells per litre (see Results Section below). To avoid a potential bias in the interpretation of the DESeq analysis results, we accounted for this difference in cell abundance by multiplying the metatranscriptomics raw reads of the PA treatment by a conservative factor of five prior to the analysis.

Cell specific calculation with internal standards

Total transcripts per sample (500 ml) were calculated by multiplying protein encoding reads of the transcriptome library with the number of molecules of the internal standard added to the samples; this value was then divided by the recovered reads of the internal standard in the sequence library (Satinsky *et al.*, 2013). Total transcripts per samples were then divided by the number of total cells per sample to obtain an estimate for transcripts per cell.

DESeq2 normalization with internal standards

We have used the Likelihood Ratio Test (LRT) as implemented in the DESeq2 software (Love *et al.*, 2014) to screen each gene for significant changes in transcription activity of the particle-attached (PA) treatment against free-living particle-sphere (FLP) and free-living control (FLC) as well as the FLP treatment against FLC (fitType: local). After adjusting for multiple testing, the P-value padj < 0.05 was used as threshold to identify genes with significant changes in transcription activity. The raw count data were normalized based on the count data of the internal standards by applying the controlGenes option of the DESeq2 software to statistically test the dynamics of absolute transcriptional activity. A further normalization step was performed in order to account for differences in the average feature length of each gene in each sample using the normMatrix option. Genes with a total count of less than 50 were eliminated. This was done by KEGG Ids and they were normalized by length of genes and retrieved standard reads.

Metal co-factor analysis

KEGG Ids of co-factor related enzymes were retrieved from the enzyme by cofactor function on the bioinformatics resource portal ExPASy (https://enzyme.expasy.org/enzyme-bycofactor.html). Co-factors were retrieved for Co^{2+}, Cob(II)alamin, cobalt cation, Cu cation, Cu^{2+}, Fe cation, Fe^{2+}, Fe^{3+}, Ferriheme b, Iron-sulphur, Mn^{2+}, Ni^{2+}, Zn^{2+}.

Results

Particle characterization

The in situ POC concentration at 100 m depth, as determined from the material collected with the ISP, was 6.17 µM. A similar POC concentration was determined in the control cultures at the end of the experiment (5.5 ± 1.7 µM; mean ± SD of duplicate cultures). By contrast, POC concentrations in the three replicate particle-amended cultures were substantially higher and more variable, amounting to 22.3, 48.0 and 76.4 µM. In the cultures filtered through 0.8 µm after 7.7 days, the POC concentration was 10.4 ± 3.9 µM (mean ± SD of duplicate cultures) at the end of the incubation.

We used microscopic observations to characterize the particle size spectrum in the range 25 - 1000 µm for one of the treatments (Figure 5; Suppl. Fig. 1). Small particles (25 - 150 µm) accounted for roughly 55% of total particle surface, while each of the considered larger size classes accounted for < 10% to the total particle surface. The particle size spectra were similar at the start and at the end of the incubations. The abundance of cells increased with particle size, ranging from less than 10 cells per particle in the size range 25-50 µm to an average of about 30 cells on the largest particles (500-1000 µm). Prokaryotic cell abundance normalized to particle surface revealed no preferential colonization of a given size class. The number of cells per surface unit remained stable over the course of the experiment, indicating that the particles were already colonized prior to collection and that there was no further colonization of particles during the 12-d incubation period. Based on the total particle surface area and the prokaryotic cell abundance per surface unit, we estimated 0.6 x 10^5 cells mL^{-1} to be attached to particles at the end of the experiment.

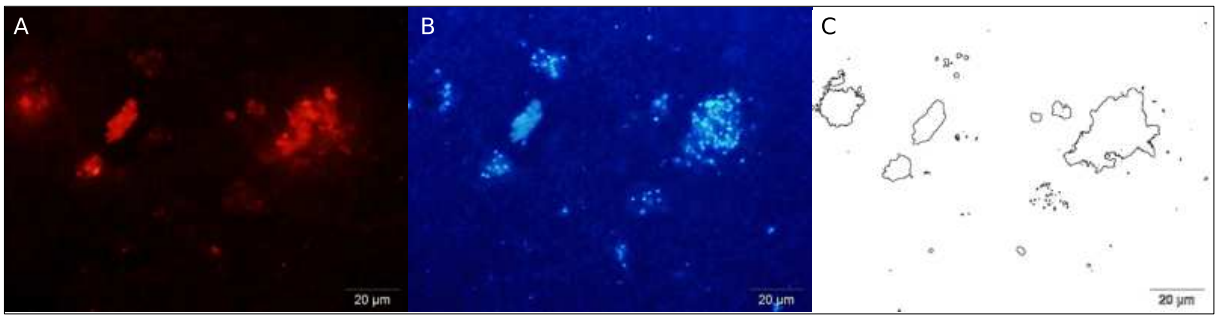


Figure 5. *Microscopic observation of particles collected at the end of the incubation. (left panel) (A) Red channel of the image produced with white light excitation, showing particles of different size and nature as revealed by the intensity of the fluorescence. (B) Image produced with excitation at 365 nm showing the attached cells stained with DAPI. (C) Image shown in left panel treated with the image J software to delineate the edges of particles before measurement of the surface area.*

Prokaryotic abundance and heterotrophic production

Prokaryotic abundance, as determined by flow cytometry, in the < 0.8 µm size fraction, referred to as free-living, accounted for 96 ± 5% (n = 22) and 92 ± 5% (n = 30) of the abundance in the unfiltered seawater in the control- and the particle-amended treatments, respectively. The slightly higher abundances in the unfiltered seawater as compared to the < 0.8 µm fraction in the particle-amended treatments are likely a combined effect of the detachment of cells from particles during the manipulation of the samples prior to flow cytometric analysis and the retention of cells larger than the filter pore size used. Because we added prokaryotic cells concurrently with particles, the abundance of free-living prokaryotes in the particle-amended treatments were 3.6 -fold higher than in the control treatment at the start of the experiment. The abundance of free-living prokaryotes remained higher in the presence of particles throughout the experimental period (Fig. 6A).

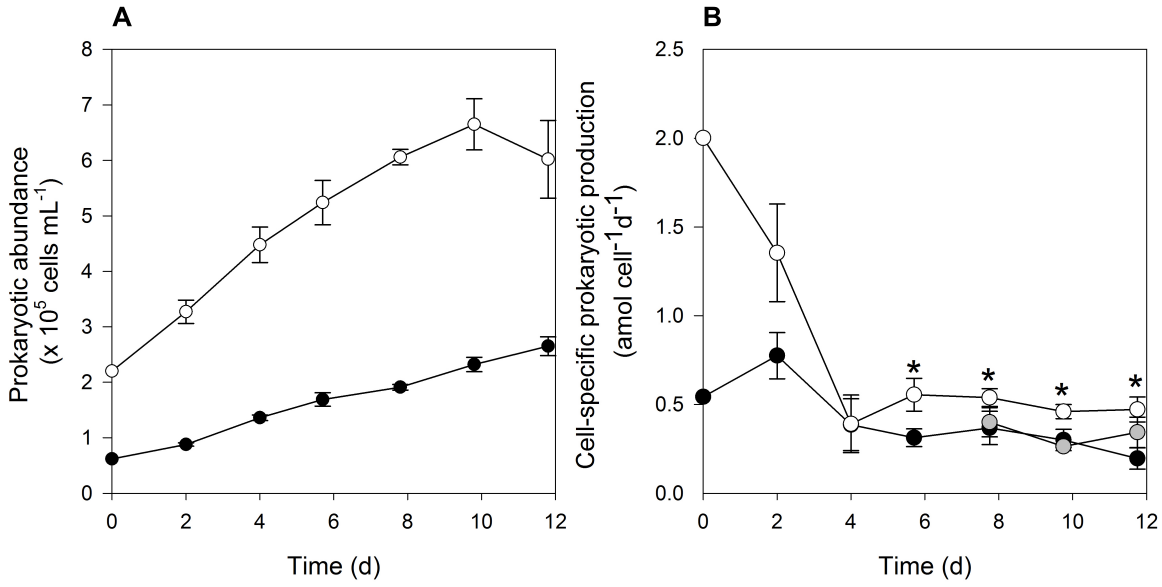


Figure 6. *Abundance (A) and cell-specific production (B) of free-living prokaryotes over the 12-day incubation period. Mean±SD of biological replicates are given. White circles: FLP, Black circles FLC, Grey Circles FLC-P. Asterisk denote significant difference between FLP and FLC (B).*

Despite the variability in the concentration of particulate organic carbon in the particle-amended incubations, the abundance of free-living cells was similar in the three replicate incubations. At the final time point, free-living prokaryotes amounted to 6.02 ± 0.7 x 10^5 cells mL^{-1}. This corresponds to about 10 times the abundance of particle-attached prokaryotes as determined in one of the replicates of the respective treatment.

Particle-attached prokaryotes revealed highly active over the course of the experiment (Suppl. Fig. 2). Particle-associated prokaryotic heterotrophic production ranged between 660 and 981 pmol leucine L^{-1} d^{-1} during the first 6 days of incubation and decreased to values

ranging between 265 and 556 pmol leucine L^{-1} d^{-1} during the last 4 days of incubation. Based on the number of attached cells determined in one of the replicate bottles of the respective treatment, we calculated the cell-specific production for attached cells to be 11.0 and 4.4 amol leucine $cell^{-1}$ d^{-1} at the initial and the final time points, respectively. After 7.7 days of the experiment, particles were filtered out in two of the five replicates. In these incubations, the particle-attached prokaryotic production dropped to values below 150 pmol leucine L^{-1} d^{-1}. In the control treatment, prokaryotic production in the particle-associated size fraction (> 0.8 μm) was undetectable. Bulk prokaryotic production remained below 73 pmol leucine L^{-1} d^{-1} during the experimental period in the control treatments.

Prokaryotic production of free-living cells was normalized to abundance, because of the higher cell abundances in the particle-amended treatments as compared to the control. In the presence of particles, prokaryotic production of free-living cells can be distinguished in two phases. During the first phase, cell-specific production rapidly decreased from 2.0 to 0.39 amol leucine $cell^{-1}$ d^{-1} (Fig. 6B). The higher cell-specific activity at the start of the experiment is likely due to the presence of active prokaryotes detached from particles, and an increased availability of inorganic elements, such as Fe, and organic carbon released from the particles by abiotic leaching processes. At the end of phase 1 (day 4), cell-specific activity in the particle-amended treatments equalled those in the control treatments. During the second phase of the experiment, a consistently higher prokaryotic cell-specific activity could be observed in the presence of particles (0.50 ± 0.06 amol leucine $cell^{-1}$ d^{-1}, n = 4) as compared to the control (0.30 ± 0.08 amol leucine $cell^{-1}$ d^{-1}, n = 3). In the incubations where particles were filtered out after 7.7 days, cell-specific production of free-living bacteria was 0.38 ± 0.14 amol leucine $cell^{-1}$ d^{-1}, n = 3). Cell-specific production was 6- and 14- times higher for particle-attached than for the free-living prokaryotes at the initial and final time points, respectively.

Dose addition experiment

To determine the growth limiting factor of the free-living prokaryotic community in the absence of particles, we followed prokaryotic production and abundance over 3 days in incubations amended either with Fe or C. Both Fe and C individually stimulated prokaryotic heterotrophic production with a more pronounced effect of C than Fe (Fig. 7). The addition of solely Fe lead to a significant increase in production after 3 days ($p < 0.001$, ANOVA and post hoc Tukey), while C stimulated production after 1 day ($p < 0.02$) and 3 days ($p < 0.001$) (Fig. 7). By contrast, significant increases in prokaryotic abundance were only observed in the C-amended treatment after 3 days ($p < 0.05$).

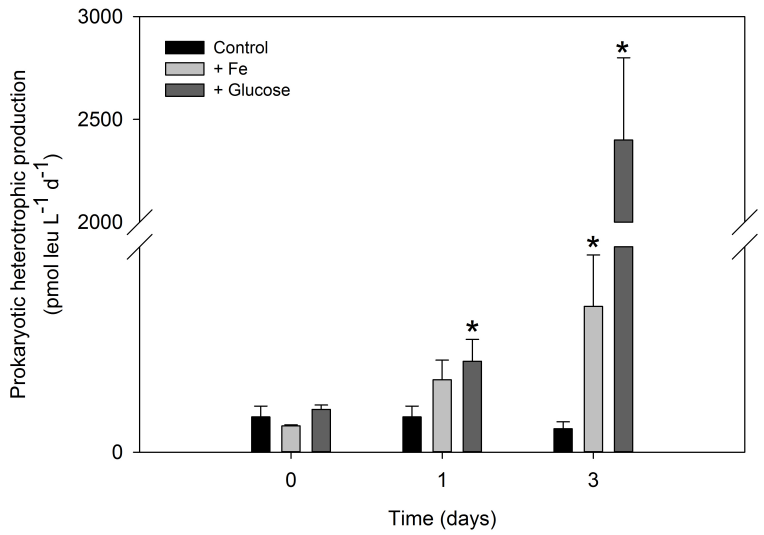


Figure 7. *Prokaryotic heterotrophic production during the dose addition experiment. Mean values ±SD of biological replicates are shown. Asterisk denote significant difference between treatment and control.*

Cell-specific prokaryotic production was significantly enhanced by C after 1 day ($p = 0.025$) and 3 days ($p = 0.006$), and Fe-addition significantly enhanced prokaryotic production after 3 days ($p = 0.013$) (Suppl. Fig 3).

Prokaryotic community composition

Observations of 16S rRNA sequences on the order level revealed treatment specific differences in prokaryotic community composition. *Flavobacteriales* showed highest contribution in all treatments peaking in FLP (35%; Fig. 8).

Other alphaproteobacterial groups with dominant contribution in the treatments were *Rhodobacterales*, the SAR11 clade, *Sphingomonadales*, followed by gammaproteobacterial groups *Oceanospirillales*, *Cellvibrionales, Thiotrichales*. Members of Betaproteobacteria of the order *Methylophilaceae* and deltaproteobacteria of *Bdellovibionales* were also among the highest relative abundances in all treatments. FLP and FLC were especially different in the relative contribution of the SAR11 clade compared to PA which was higher in FLP (11%) and FLC (13%) as compared to PA (2%), and *Alteromonadales* which was lower in FLP (1 %) compared to PA (9%).

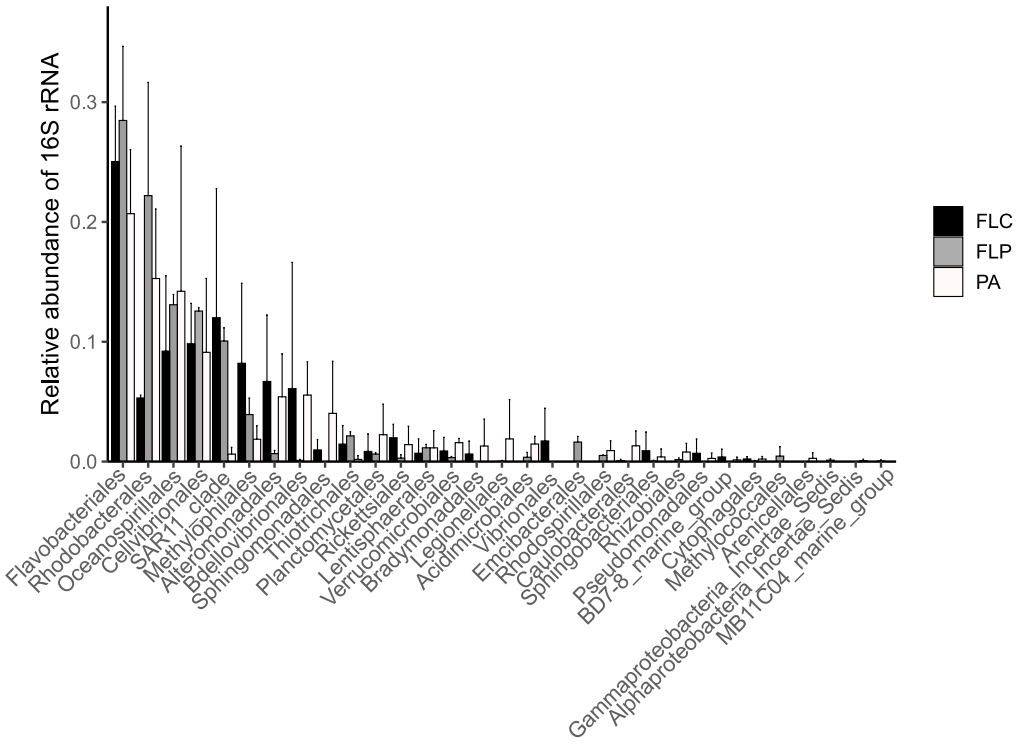


Figure 8. *Relative abundance of 16S rRNA gene sequences on order level representing 89 – 99 % of relative abundance per sample. Mean values ±SD of triplicate amplicons are shown.*

On the phylum level, the taxonomic affiliation of the assembly mirrored amplicon 16S rRNA community composition (Fig. 9., Suppl. Fig. 4).

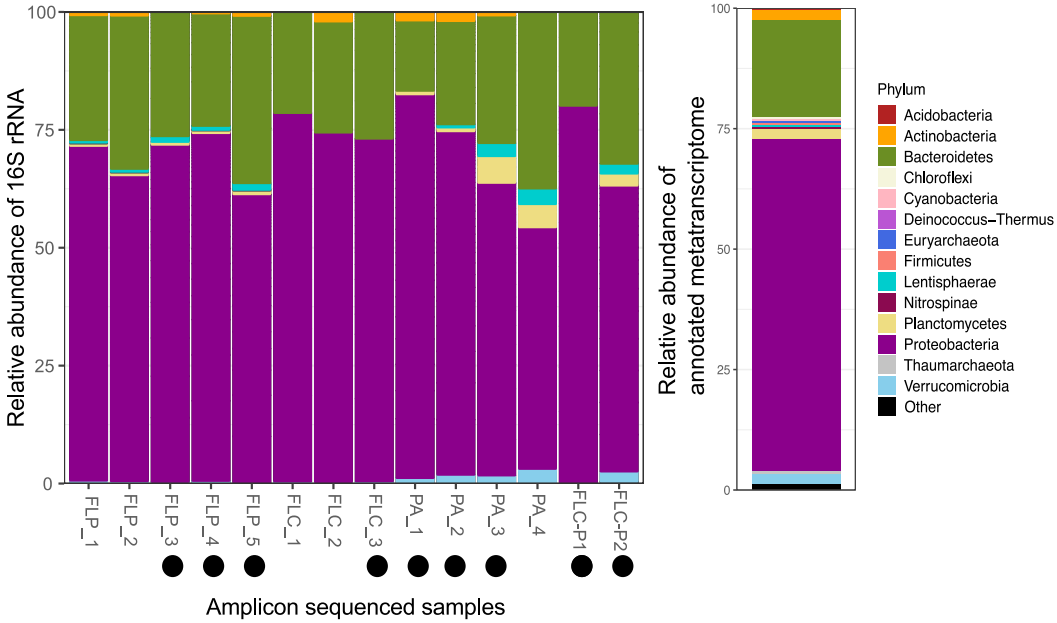


Figure 9. *Broad taxonomic affiliation of samples based on 16S rRNA genes and the assembled metatranscriptome. Samples chosen for RNA sequencing are highlighted by black dots. FLP- free-living prokaryotes in the presence of particles; FLC- free-living prokaryotes in the control (no particles added); PA - particle-attached prokaryotes; FLC-P - free-living prokaryotes in the treatment from which particles were removed after 7.7 d of incubation*

Differentially expressed transcripts

Assembly of metatranscriptomic reads and recovery of prokaryotic open reading frames resulted in 420,218 contigs. Quality assessed raw reads of PA, FLP and FLC, each in triplicates, were mapped back to the assembly and mapped reads were counted. One control sample (FLC-P_14; Suppl. Table 2) was significantly different from the triplicate set recruiting less reads than all other samples (DESeq function Cooks distance) and this sample had to be excluded from further analysis. For differential expression analysis, the control therefore consisted of one FLC and one FLC-P sample (Suppl. Table 2). In order to analyse functional differences, assigned KEGG-Ids were summed by normalized gene length and differential expression analyses were performed for the following pairs: PA vs FLP, FLP vs FLC and PA vs FLC. Significantly differentially expressed transcripts were chosen with an adjusted pvalue < 0.05 (Suppl. Fig. 5). Bray Curtis dissimilarities of transcript counts revealed grouping by treatment and control in non-metric space (Fig. 10).

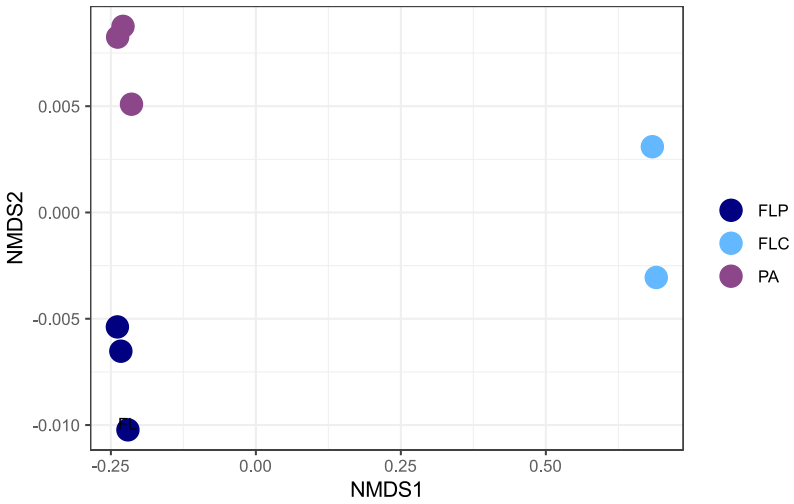


Figure 10. *NMDS plot of Bray Curtis dissimilarities calculated from count matrix for DESeq2. PA – Particle attached, FLP – Free living particle sphere, FLC – Free living control.*

Highest numbers of over-expressed transcripts were found for FLP versus FLC (4860) and PA vs FLC (4789), and this number was substantially lower for PA versus FLP (875). Comparison of FLP versus PA and FLC versus PA revealed 532 and 10, respectively, over-expressed transcripts (Table 1). A large number (4560) of KEGG Enzymes were shared between PA and FLP, and 300 and 229 were unique in PA and FLP, respectively.

Table 1. *Number of transcripts differentially expressed between particle-attached and free-living prokaryotes. PA – Particle attached, FLP – Free living particle sphere, FLC – Free living control.*

Testing conditions	Differentially expressed transcripts	Higher abundance in FLP	Higher abundance in PA	Higher abundance in FLC
FLP vs. FLC	4860	4860		
PA vs. FLC	4799		4789	10
FLP vs. PA	1407	532	875	

The highest transcript counts, corrected for cell abundance, were determined for the iron complex outer membrane receptor protein, for both PA and FLP. Flagellin (K02406) and the outer membrane protein OmpU (K08720) had also among the highest transcript counts in PA and FLP, respectively (Fig. 11, Suppl. Fig. 6, Suppl. Table of DESeq2 results provided as pdf). In the following, trace-metal related transcripts were investigated in more detail, including trace metal transporters and siderophore biosynthesis genes, genes implicated in the Fe-S cluster assembly, Fur transcriptional regulators and iron homeostasis, oxidative phosphorylation and trace metal co-factors. Other transcripts of interest, such as genes involved in the two-component regulatory system, transporters of organic compounds, genes implicated in the nitrogen metabolism and vitamin B biosynthesis were additionally considered. In an attempt to describe the underlying prokaryotic diversity for specific transcripts of interest, taxonomic affiliation was retrieved by aligning all sequences by a KEGG-Id to the non-redundant protein database of NCBI.

Trace metal transporter and siderophore synthesis

A total of 1780 protein families under KEGG Brite ID 02000 (Transporters) were retrieved and matched with the differentially expressed transcripts. From the resulting 610 transporters, all trace metal transporters (Fe, Zn, Mn, Ni, Cu, Co, W, Mo) were extracted (Fig. 11). In addition, KEGG Ids from the pathway 01055 corresponding to the biosynthesis of siderophore group non-ribosomal peptides were retrieved, and 15 out of a total of 30 enzymes matched with the DESeq2 results.

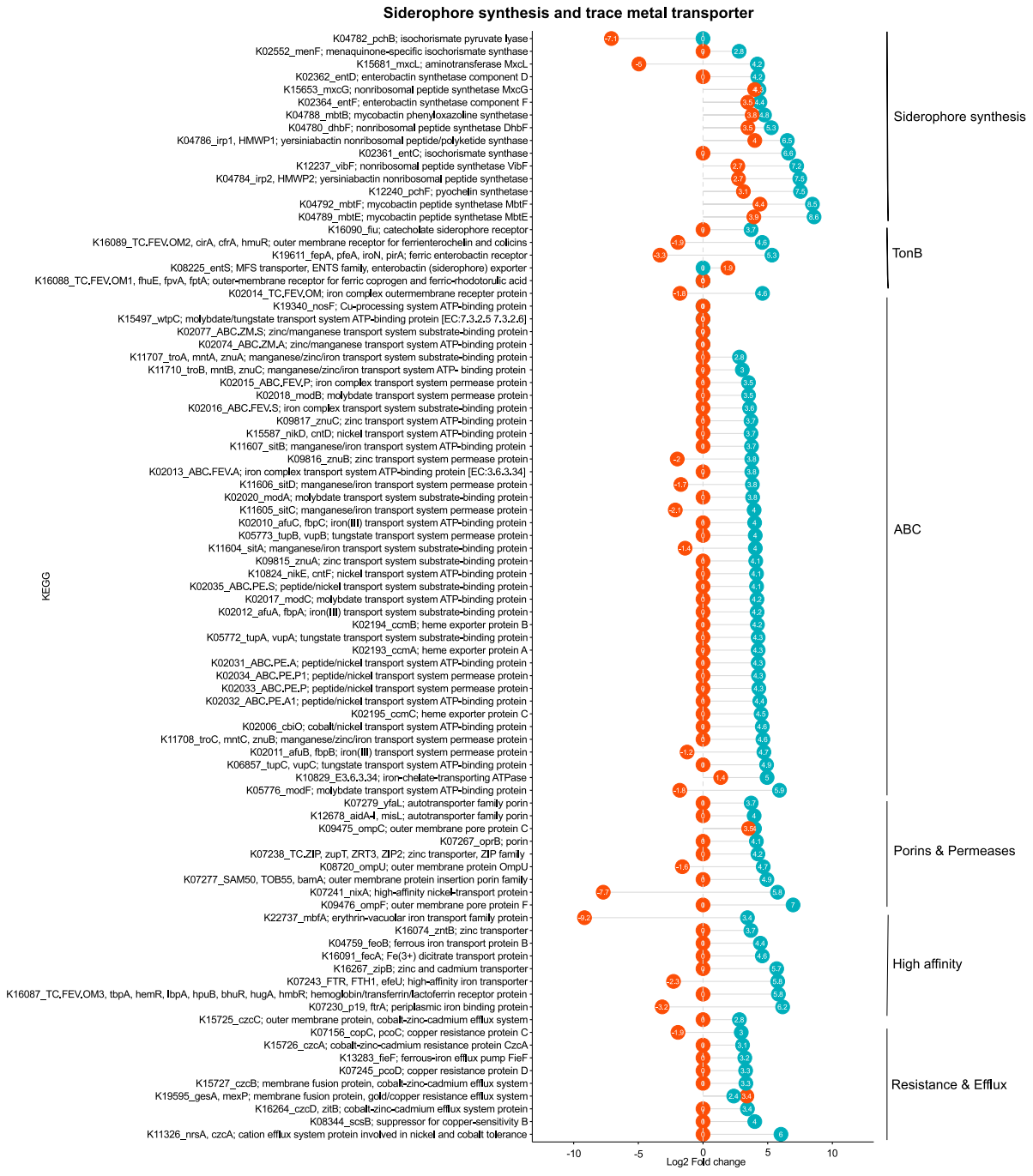


Figure 11. *Log2 Fold changes for differentially expressed transcripts. Positive values indicate higher expression in PA than in FLP (orange circles) and in FLP than in FLC (blue circles). Negative values indicate lower expression in the respective comparisons.*

Most siderophore synthesis transcripts were higher expressed in PA compared to FLP and in FLP to FLC (Log2 fold change 2.7 to 4.4, Suppl. Fig. 6 & 7). By contrast, siderophore transporters such as TonB receptors were mostly higher expressed in FLP vs PA (Log2 fold change -1.8 to -3.3, K02014, K16089, 19611). One siderophore exporter (enterobactin) was found highly expressed in PA compared to FLP (Log2 fold change 1.9, K08225) (Fig. 11).

Most ABC transporter complexes for multiple trace metals were higher expressed in FLP vs FLC (Log2 fold change 2.8 to 5.9), while most of them were not different between PA and FLP (Fig. 11, Supplementary Fig. 8). An exception was the iron-chelate transporting ATPase K11326 with a Log2 fold change of 1.4 for PA to FLP, and this transporter was highly expressed in FLP to FLC (Log2 fold change 5). The structural complex of the iron (III) transport system was expressed with a Log2 fold change of 4 to 4.7 in FLP to FLC. ACB transporters for manganese/iron (K11604, K11606, K11605), molybdate (K05776) and zinc (K09816) showed lower expression in PA compared to FLP (Log2 fold change -1.2 to -2.1). By contrast, several ABC transporters for copper, molybdate/tungstate and zinc/manganese were differentially expressed in PA to FLP and FLC (Log2 fold change 2.8 – 5.9, K19340, K15497, K02077, K02074).

Several high affinity trace metal permeases and porins were higher expressed in FLP compared to PA (Log2 fold change -1.6 to -7.7, K08720, K07241, K07230). One of the most pronounced changes was observed for the vacuolar erythrin iron transport, which was highly expressed in FLP to PA (Log2 fold change -9.2) and FLP to FLC (Log2 fold change 3.4, K22737). Copper and cobalt-zink-cadmium resistance proteins were higher expressed in FLP to FLC (Log2 fold change 3.1 to 4, K15726, K08344, K07245) as well as partially in FLP to PA (Log2 fold change -1.9, K07156). The same could be observed for efflux systems proteins for cobalt-zink-cadmium, ferrous iron, nickel and cobalt which were higher expressed in FLP over FLC (Log2 fold change 2.8 to 6, K16264, K15727, K15725, K13283, K11326) while a cobalt/gold efflux system was higher expressed in PA to FLP (Log2 fold change 3.4, K19595).

Intracellular regulation

All of the observed Ferric uptake regulator (Fur) transcripts showed high differential expression in FLP to FLC, but there was no difference observed in PA to FLP (Fig. 12). This indicates the enhanced regulation of intracellular Fe concentrations in PA and FLP, as compared to FLC. Members of the Fur family are also involved in homeostasis of other ions.
In our dataset we found higher expression of transcripts related to the homeostasis of zinc (Zur), iron or nickel (Nur) or manganese (Mur), and peroxide (PerR, occurring in zinc/iron and zinc/manganese forms) in PA and FLP (Supplementary DESeq2 table). Iron homeostasis transcripts such as for ferritin, bacterioferritin and dps were upregulated in PA and FLP compared to FLC (K02217, K03594, K04047, Log2 fold change: 5.3, 4.1, 3.4; Suppl. Fig. 9) and ferritin and dps were also upregulated in PA compared to FLP with Log2 fold change of 1.5 and 1.3, respectively (Fig. 12).

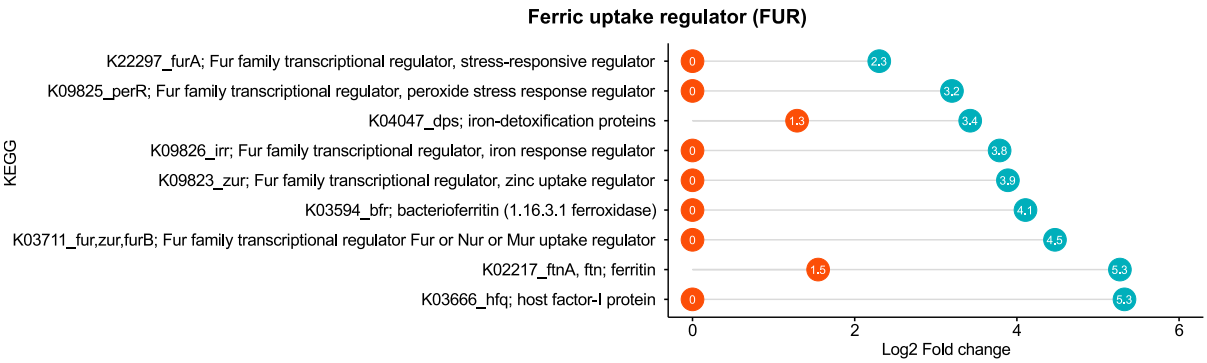


Figure 12. *Log2 Fold changes for differentially expressed transcripts. Positive values indicate higher expression in PA than in FLP (orange circles) and in FLP than in FLC (blue circles).*

Iron being an essential cofactor for respiration, we investigated transcripts necessary for the assembly of Fe-S clusters which require specific biosynthetic pathways (Fig. 13). The cysteine desulfurase enzyme (Suf) which is required to liberate sulphur atoms from free cysteine to be used in cluster assembly and a corresponding transcriptional regulator were highly expressed in FLP compared to FLC (Fig. 13, Log2 fold change 3.7 - 4.3, 3.2 - 4.6, SufE, SufS, iscS). For the Fe-S cluster assembly all other necessary transcripts were higher expressed in PA and FLP compared to FLC (Log2 fold change 2.7 – 4.7 SufA, SufC, SufD, SufB).

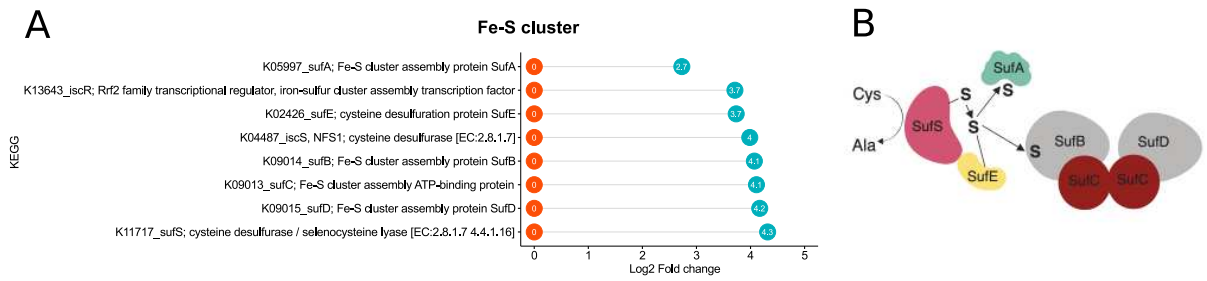


Figure 13 A. *Log2 Fold changes for differentially expressed transcripts. Positive values indicate higher expression in PA than in FLP (orange circles) and in FLP than in FLC (blue circles).* ***B.*** *Simplified scheme of Fe/S cluster assembly of Suf, modified from (Ayala-Castro et al. 2008)*

Oxidative phosphorylation

The respiratory chain consists of 5 complexes that contain one or more Fe atoms as well as Cu cofactors. The majority of the genes involved in the oxidative phosphorylation, a key pathway for ATP production, were higher expressed in FLP than FLC (Fig. 14, Suppl. Figure 10). Especially in complex II all types of succinate dehydrogenase/fumarate reductase have been found over expressed in FLP and PA compared to FLC (SdhC, SdhD, SdhA, SdhB). By

contrast, only a few of these genes were differentially expressed between FLP and PA (Fig. 14).

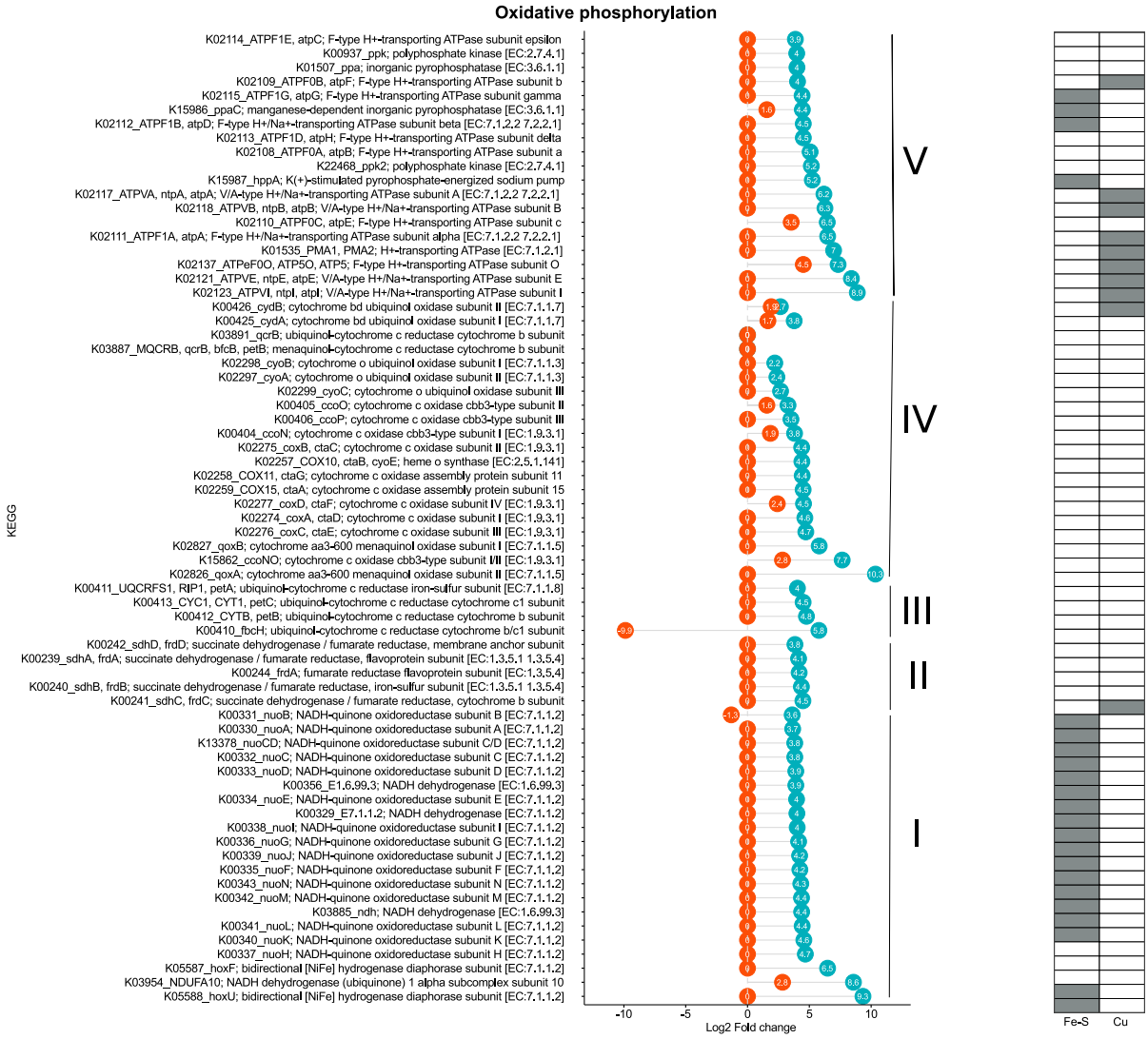


Figure 14*. Log2 Fold changes for differentially expressed transcripts in the 5 complexes of the respiratory chain. Positive values indicate higher expression in PA than in FLP (orange circles) and in FLP than in FLC (blue circles). Negative values indicate lower expression in the respective comparisons.*

Metal co-factors

Trace metal co-factors identified by KEGG-Id were retrieved from ExpaZy and when analysed with differential expression, resulted in 233 Fe co-factor enzymes, 93 Zn, 78 Mn, 25 Co, 24 Cu and 13 Ni. Of the 233 Fe enzymes the majority were Fe-S clusters followed by Fe^{2+} and Fe^{3+} co-factors. Detailed analysis of zinc co-factor enzymes revealed multiple metalloproteinases to be highly expressed in PA compared to FLP as well as FL to FLC (Suppl. Fig. 11).

Two-component systems

Given the pronounced response of prokaryotes attached to or in the vicinity of particles, we investigated transcripts with the potential for sensing changes in growth conditions. We identified a total of 65 sensor histidine kinases in our dataset, and most (62) of them were higher expressed in PA and FLP compared to FLC (Fig. 15; Suppl. Fig. 12). The ferric enterobactin receptor (fepA) is part of the PfeS-PfeR-PfeA enterobactin-dependent Fe acquisition two-component regulatory system and was highly expressed in FLP compared to both FLC and PA (Log2 fold change 5.3 and -3.3 respectively, Fig. 15).

The CRP/FNR family transcriptional regulator, nitrogen fixation regulation protein (K15861, L2FC 4.6) was highly expressed in PA compared to FLP. The complete tetrathionate respiration trS-TtrR two-component regulatory system could be found highly expressed in PA to FLP (K13040, TtrS, L2FC 3.8; K13041, TtrR, L2FC 3.4; K08357, TtrA 6.9; K08358, TtrB, Log2 fold change 4). The RegA (redox response) two-component regulatory system as well as ubiquinol-cytochrome c PetB and PetB were highly expressed in FLP to PA which indicates regulation of the electron transfer system (K15012, regA, Log2 fold change -2; K00410, PetB, PetC, Log2 fold change -9).

In both PA and FLP multiple biofilm formation sensory elements were highly expressed compared to FLC. Multiple quorum sensing proteins were among these such as K18304, (LasR, Log2 fold change -3.4), K11711 (LuxR, Log2 fold change 4.1), K13040 (Log2 fold change 3.8), K13040 (Log2 fold change 3.4), K14978 (Log2 fold change 2.7) and K11712 (Log2 fold change 2.7). The LuxO protein was highly expressed on PA to FLC (Log2 fold change 5.8) and FLP to C (Log2 fold change 4.2) as well as high for PA compared to FLP (Log2 fold change 1.6).

Flagellar proteins MotX and MotY were highly expressed in PA and FLP versus FLC (K21217 Log2 fold change 4.5, K21218, Log2 fold change 4.3) and PA versus FLP (MotY, Log2 fold change 2; Suppl. Fig. 13). Flagellin had highest transcript contribution for PA and was highly expressed for both PA and FLP versus FLC (K0206, Log2 fold change 4.3 and 3.6). Chemotaxis proteins MotA and MotB had a Log2 fold change of 4 for both PA and FLP compared to FLC (Suppl. Fig. 13).

Figure 15. *Log2 Fold changes for differentially expressed transcripts. Positive values indicate higher expression in PA than in FLP (orange circles) and in FLP than in FLC (blue circles). Negative values indicate lower expression in the respective comparisons.*

Porphyrin synthesis pathways

Both PA and FL expressed partially pathways for Vitamin B1, B8, B12 biosynthesis compared to FLC. Vitamin B12 contains cobalt thus the ability of synthesis points to the fact that cobalt was present and that the energetic need for these biosynthetic pathways was covered in presence of particles (Fig. 16; Suppl. Fig. 9).

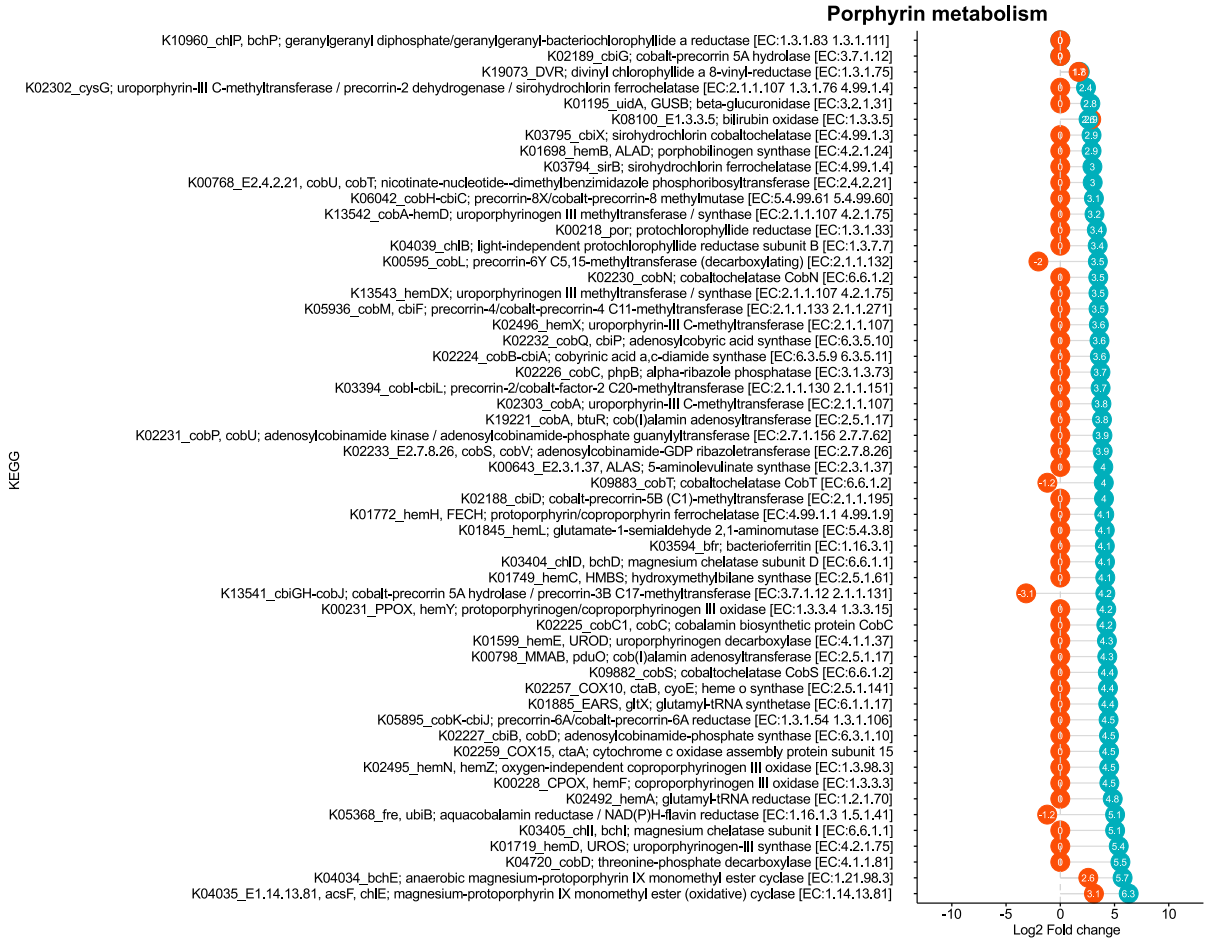


Figure 16. *Log2 Fold changes for differentially expressed transcripts. Positive values indicate higher expression in PA than in FLP (orange circles) and in FLP than in FLC (blue circles). Negative values indicate lower expression in the respective comparisons.*

Other ABC transporters

Numerous amino acids, simple sugars and complex carbohydrates, as well as lipid ABC transporters were identified to be higher expressed in PA and FLP compared to FLC, and several sugars showed higher expression in FLP to PA (Suppl. Fig. 8 & 14). Among the ABC transporters for amino acid derivatives such as taurine, spermidine and putrescine were present. Furthermore, compound specific transporters were identified for Vitamin B1 (thiamine), B8 (biotin) and B12 (cobalamin). The ABC transporter for Vitamin B12, K06073 was highly expressed in PA to FLP (Log2 fold change 5.9) as well as FLP to FLC (Log2 fold change 5.2).

All thiamine transporters were highly expressed in FLP to FLC (K02064 Log2 fold change 3.3, K15598 Log2 fold change 3.8, K02062 Log2 fold change 4.5) and the permease protein was higher expressed in FLP to PA (K02063, Log2 fold change -2.5) (Suppl. Fig. 8 & 14). The biotin ATP transporter (K16784) was differentially expressed in FLP and PA to FLC while the biotin permease was only expressed in FLP.

Nitrogen metabolism

Cu and Fe co-factor analysis revealed proteins present in nitrogen metabolism expressed in different treatments. Several Cu containing nitrite reductases (e.g. K02567, K03385) were found expressed in FLP compared to FLC while nitrogenase transcripts which are Fe-S clusters were found higher expressed in PA (Suppl. Fig 15 & 16). As nitrogen fixation is known to be oxygen sensitive this might indicate anoxic zones in the particles. Nitrate/nitrite transporters were found to be expressed in FLP and PA compared to FLC, however no transporters for ammonium were differentially expressed. Taken together this means that the energy for nitrogen cycling was available to PA and FLP prokaryotes.

Discussion

Iron and other trace metals control the biological activity in large parts of the ocean. While chemical analyses and an unprecedented international effort for sampling (GEOTRACE program) have provided a detailed picture of the concentration of a range of metals across ocean provinces (Schlitzer *et al.*, 2018), understanding the bioavailability of different chemical forms and size fractions of trace metals remains a major challenge. Through gene expression, we observed in the present study that particles collected in the upper water column of the Southern ocean are a highly available source of Fe and other bioactive trace metals (Cu, Ni, Co, Zn, Mo, W) to heterotrophic prokaryotes, stimulating the metabolic activity of cells attached to or in the sphere of particles. For Fe, the limiting factor of primary production, we were able to identify different mechanisms of Fe acquisition and describe how this element affects major metabolic cellular functions. Our study provides a new perspective of particulate Fe availability with implications for other trace metals or nutrients in the Southern Ocean.

Bioavailability of Fe

For a better understanding of the availability of trace metals to microbes, geochemical and biological approaches are used. For the characterization of the particles of the present study, the leaching protocol revealed that the trace metal content was dominated by the refractory fraction with a labile fraction of particulate Fe of roughly 10% (van der Merwe *et al.*, 2019). However, this particulate labile Fe fraction was substantially higher than the dFe inventory at our study site (van der Merwe *et al.*, 2019), suggesting particles are an important potential source of Fe.

A biological approach to determine Fe bioavailability was proposed based on uptake rates by phototrophic microbes (Lis *et al.*, 2015). This has led to the definition of a 'bioavailability envelope', ranging from highly bioavailable Fe, corresponding to the pool Fe', to strongly chelated Fe (i.e. by the terrestrial siderophore DFOB) at the lower limit. This concept suggests that the expression of genes coding for different transport pathways of Fe into the cell are candidates to study Fe bioavailability and/or how Fe is made bioavailable for heterotrophic prokaryotes. The same approach of bioavailability could in principle also be applied to other trace metals. Therefore in order to evaluate whether and how prokaryotes take advantage from the particulate source of metals, we first focused on membrane uptake systems.

Ion transport

High affinity transporters are used by prokaryotes for the uptake of unchelated cationic or anionic forms of metals. In the present study, we observed a higher expression of ABC transporters in cells attached to or in the vicinity of particles as compared to the control for the seven essential metals with well-known biological function in prokaryotes. Cationic transporters were found for Co, Cu, Fe, Mn, Ni and Zn whereas one anionic transporter was present for Mo in the form of molybdate. Additionally, an anionic transporter was also found for W, a metal with a similar biological function as Mo. All these metals, with the exception of W, which was not analysed in our study, were abundant in the particles (data not shown). Our observations suggest that significant amounts of these metal ions were produced in the particles or more likely at the interface between the particles and the ambient seawater. Although our data clearly show that an additional ionic pool of these metals exist and is actively used by the prokaryotic community when particles are present, we are not able to infer its exact origin. The inorganic ions could result from either thermodynamically or biologically mediated dissolution of the particles. However, the similar expression pattern of the ABC transporters suggest that this pool is accessible in a same way for prokaryotes attached to particles and present in the

particle sphere. This is easily conceivable, as in our experiment the rapid dispersion of small ionic species was favoured by mild but continuous mixing during the incubation. In the environment, the sphere affected by the ionic pool will be controlled by the diffusion of the different ions. Interestingly, ABC transporters for Fe(II) and Fe (III) were both found differentially expressed for particle-attached and cells in the particle-sphere. Our results suggest that both redox forms are significant sources of bioavailable Fe. Fe(II) is usually present at vanishing concentrations in oxygenated open ocean seawater and does not constitute a significant source of Fe for microbes (Croot and Heller, 2012). Metagenomic analysis (Hopkinson and Barbeau, 2012; Toulza *et al.*, 2012) provided evidence that in the open ocean Fe(III) prevails over Fe(II) uptake, the opposite pattern being observed in coastal environments (Toulza *et al.*, 2012). Our results show that the presence of particles modify this view and that significant amounts of Fe(II) are accessible. Fe(II) is likely produced in anoxic micro-environments that are present in particles (Ploug *et al.*, 1997; Ploug, 2001; Grossart *et al.*, 2003; Ploug and Bergkvist, 2015) and is rapidly utilised by prokaryotes before oxidation. The fraction of Fe(II) that did not escape oxidation in our study likely contributed to build up of the Fe(III) pool that is also available for prokaryotes as shown by the expression of Fe(III) transporters.

Additional interesting information can also be derived from the taxonomic assignation of the prokaryotes that express genes for metal ion uptake pathways. Based on KEGG annotation we identified the taxa with a high contribution (> 10% to total gene transcripts) to a number of ABC transporters. *Rhodobacterales* had the highest contribution to the ABC transporters for Fe(III), Fe(II), Zn(II), Mn(II)/Fe(II), both in particle-attached and free-living cells (Fig. 17 & Suppl. Fig. 17). This observation is coherent with previous studies that have identified *Rhodobacteracaea* (Tang *et al.*, 2012; Debeljak *et al.*, 2019) and an alphaproteobacterial strain (Weaver *et al.*, 2003) among the highest contributors to Fe(III) ABC mediated transport.

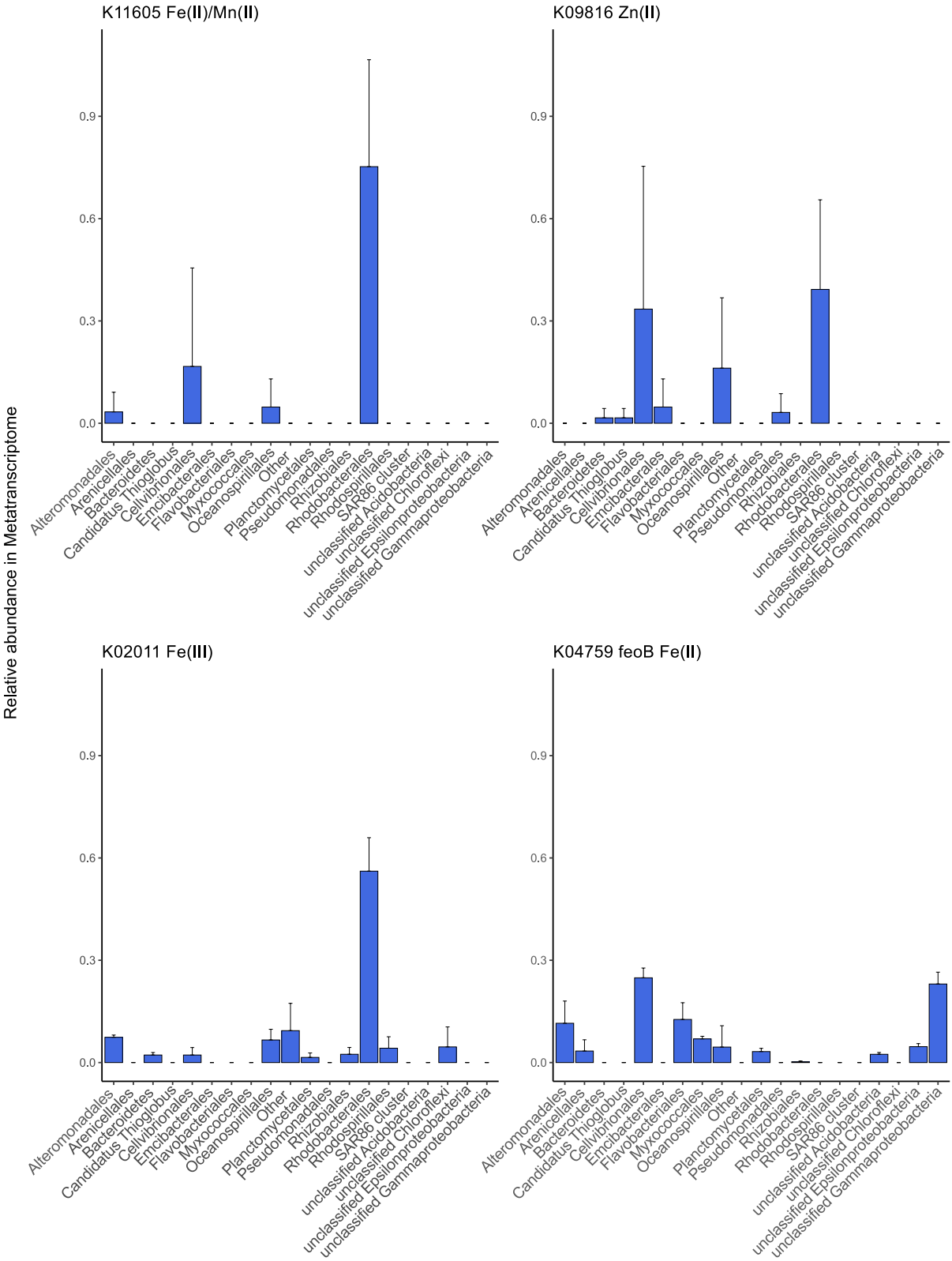


Figure 17. *Examples of ion transporter genes and the relative contribution of taxa for particle attached cells. Mean values ± SD of triplicates are shown (FLP is presented in Supplementary Information).*

Siderophore synthesis and transport of Fe

Siderophores are chelating substances and high-affinity Fe carriers, and siderophore uptake is a well-known pathway for Fe acquisition in heterotrophs (Neilands, 1995; Crosa *et al.*, 2004; Sandy and Butler, 2009; Cornelis and Andrews, 2010; Hider and Kong, 2010). In the marine environment a few siderophores were characterized on a molecular level as compared to terrestrial siderophores (Butler and Theisen, 2010). The siderophore synthesis genes that could be identified in our study were highly expressed in the presence of particles (PA and FLP) as compared to the control, indicating this mechanism is utilised to render a fraction of the particulate Fe bioavailable. Siderophore synthesis was observed to be enhanced both, under Fe limited conditions in the ocean (Völker and Wolf-Gladrow, 1999; Cornelis and Andrews, 2010; Boiteau *et al.*, 2016), and upon pulsed Fe additions in an experimental study (Sijerčić and Price, 2015). Because siderophore synthesis is a costly cellular undertaking as energy needs to be derived from the proton motive force of the inner membrane (Völker and Wolf-Gladrow, 1999; Ferguson and Deisenhofer, 2004; Wiener, 2005; Noinaj *et al.*, 2010), the availability of organic carbon could be an additional regulating factor. In the present study, siderophore synthesis genes were among the few Fe-related pathways that were higher expressed by cells attached to particles as compared to cells in the particle-sphere. The particles we used provided access not only to Fe, but also to labile organic carbon. The high gene expression together with the enhanced cell-specific prokaryotic production rates suggests that particles present a hotspot for the biosynthesis of siderophores.

How accessible are these siderophores to cells in the particle-sphere? The link between the biosynthesis of the large variety of siderophores (Sandy and Butler, 2009) and their respective utilization by diverse prokaryotes is not well understood. The production of siderophores by a cell does not guarantee scavenging of Fe bound to the same compound (Hopkinson and Morel, 2009). As a consequence, cells producing siderophores could potentially be wasting energy, while others that do not possess pathways of production can take advantage (Boiteau *et al.*, 2016). A different scenario could be that siderophore production is a strategy to capture Fe while preventing uptake by competitors that lack these acquisition pathways (Boiteau *et al.*, 2016). We observed that most of the identified transcripts for TonB receptors required for siderophore uptake were higher expressed by free-living cells in the particle-sphere than by attached cells. These results could indicate the transfer of Fe bound to siderophores, produced by attached cells, to the particle sphere.

To address the question of who the siderophore producers are and which taxa utilize siderophore-bound Fe we identified the taxa contributing to siderophore synthesis and TonB receptors required for their uptake (Fig. 18 & 19). Taking as an example the four siderophores Enterobactin, Yersiniabactin, Myobactin and Pyochelin, our analysis shows that these are each produced by a distinct assemblage of taxa, independent of whether they are particle-attached or free-living. Synthesis of Myobactin and Pyochelin were dominated by 3-4 taxa, while a larger range of taxa contributed to the synthesis of Enterobactin and Yersiniabactin. This illustrates variability in the specificity of siderophore producing prokaryotes (Fig. 18 & Suppl. Fig. 18).

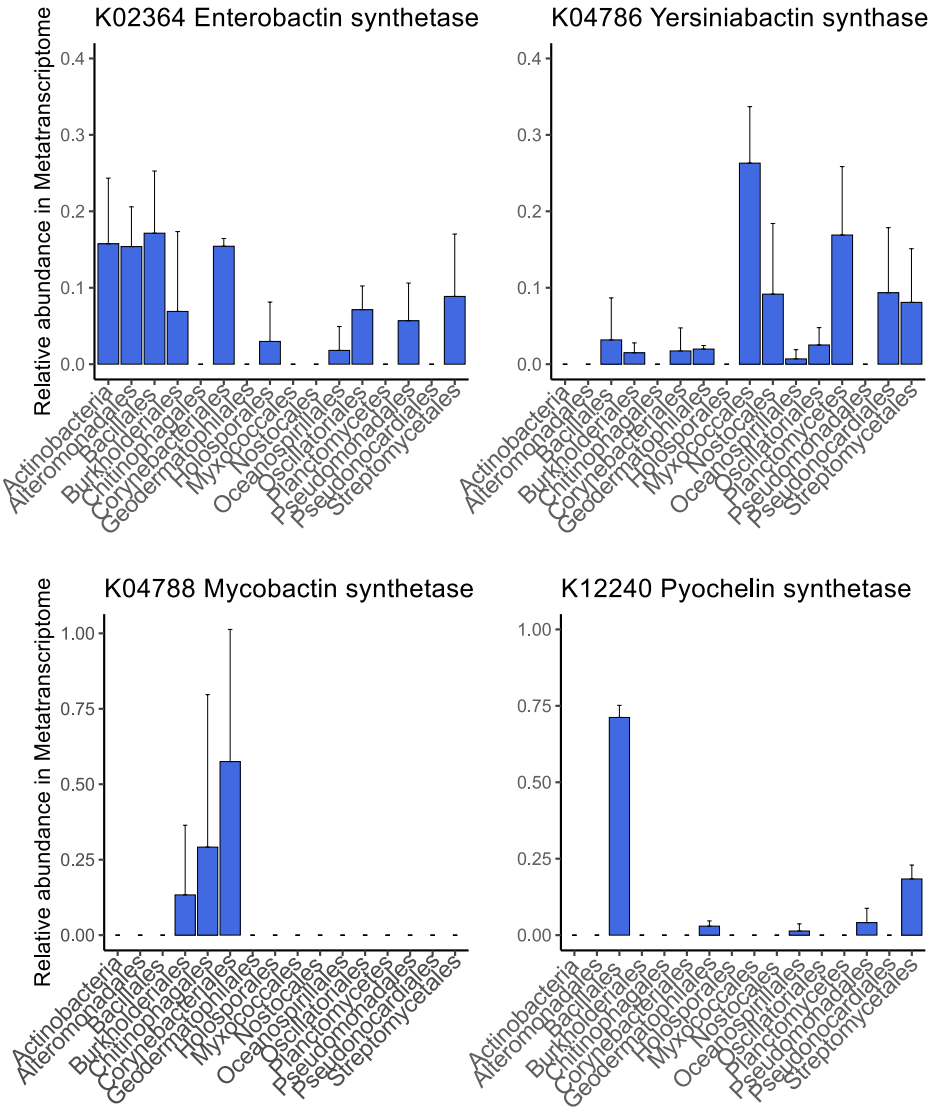


Figure 18. *Examples of siderophore synthesis pathway genes and the relative contribution of taxa for particle attached cells. Mean values ± SD of triplicates are shown (FLP in supplementary).*

Our taxonomic analysis of the TonB transcripts allowed us to identify, among the four siderophores described above, only the taxa contributing to enterobactin uptake (Fig. 19 & Suppl. Fig. 19). SAR86 and SAR92 dominated the TonB transcripts for enterobactin, two

groups that were not identified as producers (Fig. 18). In addition, the siderophore exporter for enterobactin was found highly expressed by representatives of the order *Flavobacterales* in particle-attached compared to free-living cells. These results suggest that in our experimental setting enterobactin is produced by a range of taxa and that SAR92 and SAR86 are the most competitive in the utilization of Fe bound to this siderophore.

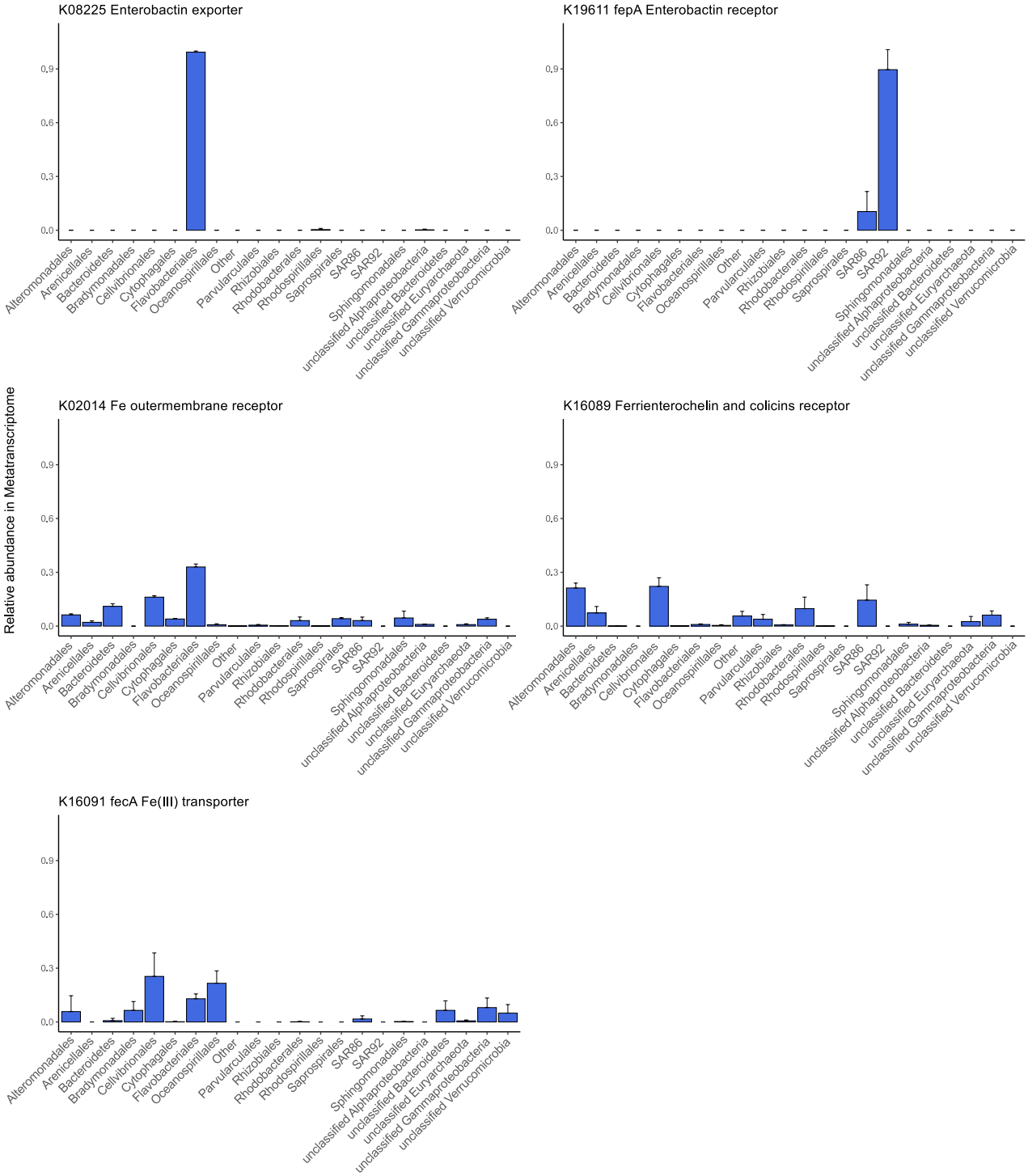


***Figure 19**. Examples of TonB-dependent receptor genes and relative contribution of taxa for particle attached cells. Mean values ± SD of triplicates are shown (FLP in supplementary).*

Sensing and cell-cell communication

Two-component system kinases are responsible for the sensing of nutrients, including trace metal or other cells and have been described as important indicators for biogeochemical cycling in genomes and metaproteomes of marine prokaryotes (Panzeca *et al.*, 2006; Carini *et al.*, 2014; Held *et al.*, 2019). In the present study, cellular signalling processing could potentially explain the energetic investment into siderophore synthesis. The high expression of PFeA a TonB dependent outer-membrane protein, related to FepA, that binds and translocates Fe(III)-enterobactin into the periplasm (Poole *et al.*, 1990), could indicate that free-living prokaryotes were actively sensing for Fe in the vicinity of particles. A recent study confirmed that PfeA recognises enterobactin in *Pseudomonas aeruginosa* (Moynié *et al.*, 2019). Furthermore, chemotaxis and biofilm regulation genes were highly expressed by particle-attached cells (Supplementary Figure 13). Potential cell-to-cell communication could in part explain the higher siderophore production on particles. In a recent study, the model organism *Vibrio harveyi* was observed to use quorum sensing as well as a Fe repressed gene cluster to produce cell-associated siderophores and other related soluble siderophores (McRose *et al.*, 2018). The production of acylated homoserine lactones was observed in *Roseobacter* strains isolated from marine snow and indicating quorum sensing on these particles (Gram *et al.*, 2002).

Intracellular processing of Fe

Once a metal has been taken up, its intracellular concentration must be carefully managed as high concentrations can become toxic. For example, an excess of Fe can lead to the formation of reactive oxygen species inside the cell (Andrews *et al.*, 2003). The ferric iron uptake regulator (Fur) family is responsible for the regulation of Fe uptake and homeostasis in prokaryotes (Andrews *et al.*, 2003; Cornelis and Andrews, 2010). We did not observe any significant differences in the expression of the observed Fur transcripts between particle-attached and free-living cells (Fig. 12). The expression of a Fur transcriptional regulator in excess of the ion inhibits downstream expression of the gene and results in noncoding RNA called *Ryh*B (Fig. 20A&B; Massé and Gottesman, 2002; Andrews *et al.*, 2003).

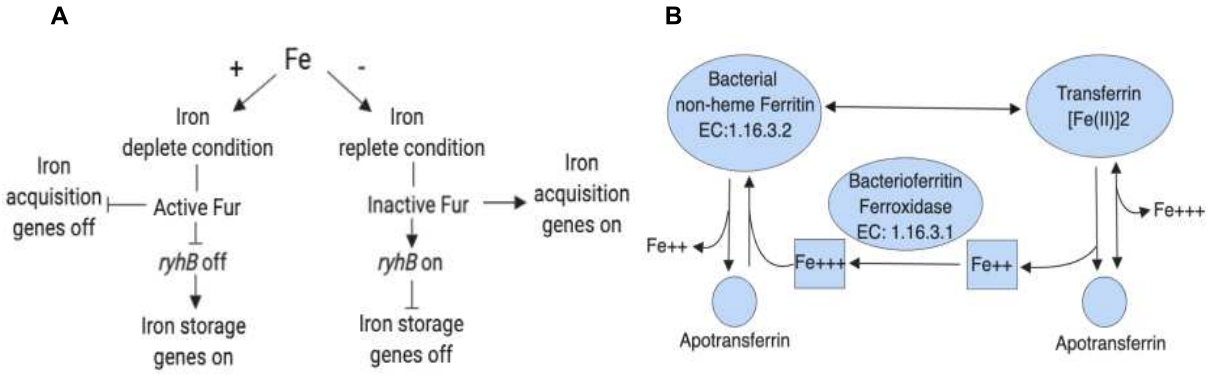


Figure 20 A. *Simplified Fur mechanisms modified from Massé and Gottesman 2002.* ***B.*** *Simplified scheme of ferritin pathway from KEGG.*

*Ryh*B was shown to be inversely correlated with proteins known to be positively regulated by Fur (Massé and Gottesman, 2002). These include the two ferritin genes ftnA and bfr, which, in the present dataset were both higher expressed in the presence of particles (Fig 12 & Suppl. Fig. 9). Three types of Fe storage proteins are described for prokaryotes: the haem-containing bacterioferritins, the haem-free ferritins and the smaller dps proteins (Andrews *et al.*, 2003; Cornelis and Andrews, 2010). In our study, dps and ferritin were higher expressed by cells attached to particles as compared to free-living ones in the particle sphere. This could indicate high Fe concentrations in particle-attached cells that needed to be regulated by homeostasis genes. *Alteromonadales, Cytophagales* and *Flavobacteriale*s were the main contributors to bacterioferritin transcripts, and heam-free ferritin transcripts belonged mainly to *Alteromonadales, Bdellovibrionales, Cellvibrionales* and *Oceanospirillales* (Fig. 21 & Fig. 20).

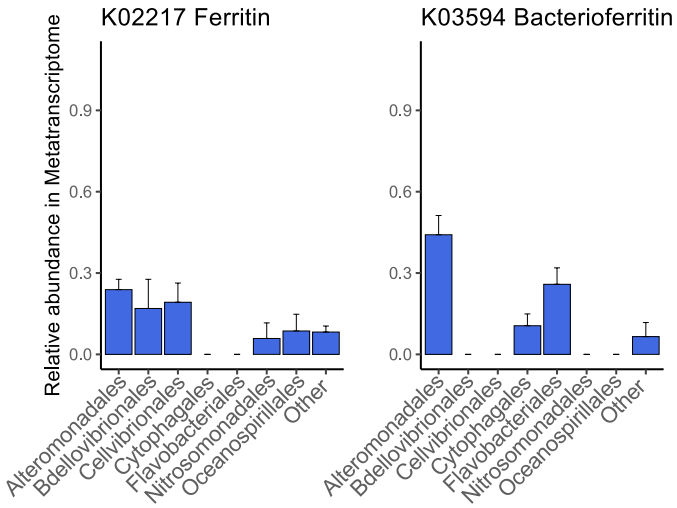


Figure 21. *Examples of bacterioferritin and ferritin genes genes and relative contribution for particle attached cells. Mean values ± SD of triplicates are shown (FLP in supplementary).*

These results complement those on siderophore uptake (previous section) and are in line with previous observations of some prokaryotic groups in the Kerguelen region. In situ gene expression profiles revealed that *Cellvibrionaceae* contributed about 15% to the total siderophore uptake transcripts in early spring (Debeljak *et al.*, 2019). During summer, SAR92 (belonging to *Cellvibrionaceae*) was identified as a main contributor to bulk leucine incorporation (30% of total leucine-active cells) throughout the upper mixed layer (125m) (Obernosterer *et al.*, 2011), a possible consequence of the competitive utilization of siderophore bound Fe. Using MICRO-CARD-FISH and ^{55}Fe, CFB and *Gammaproteobacteria*, in particular SAR86, were shown to be the dominant contributors to total Fe-active prokaryotic cells in early spring (Fourquez *et al.*, 2016). The possible storage of Fe in the form of bacterioferritin, as suggested by the gene expression patterns of the present study, could in part explain the detection of a high number of Fe-active cells belonging to these prokaryotic groups. These previous observations and the present study let us conclude that members of these groups have an important contribution to the Fe-cycle in Southern Ocean waters.

Intracellular utilization of Fe

The function of many proteins depends on the presence of metal cofactors. One of the most ancient and versatile inorganic cofactors in biology are Fe/S clusters (Wächtershäuser, 1992; Ayala-Castro et al., 2008). They range from multiple forms (e.g. 2Fe-2S, 4Fe-4S, 3Fe-4S) and are needed in proteins from varying functions from the electron transfer to regulatory processes. Described assembly structures are broadly divided into 3 categories, the housekeeping cluster assembly (Isc), the cluster assembly under stress conditions (Suf) and those for nitrogen fixation (Nif) (Ayala-Castro *et al.*, 2008). The in vivo process of Fe donation to the cluster complex remains poorly understood. The diverse evolutionary spread of these systems has been proposed as a possible biochemical advantage through the specific pathway. High rates of cluster assembly could support rapid growth of r-strategists and thus be of advantage while for K-strategist a protected and resistant function might be ecologically favoured (Ayala-Castro et al., 2008). In the present study, the Fe/S assembly was not differentially expressed between particle-attached and free-living cells, but highly different to the cells in the control treatment. These results could indicate that the Fe quota was reached when particles were present, similar to the Fur threshold. As a consequence, protein assembly was taking place in cells attached to particles or present in the particle-sphere.

Our results illustrate the important role of marine particles as a source of trace metals. In the case of Fe, we have shown that the production of siderophores contributed to transfer a

fraction of the particulate form to heterotrophic prokaryotes, active players of the microbial loop. This observation reinforces the idea of the "microbial ferrous wheel" (Kirchman, 1996), highlighting the role of heterotrophic prokaryotes in the ocean Fe cycle. Our gene expression patterns suggest increased prokaryotic siderophore production, Fe uptake and storage on marine particles, indicating that prokaryotes inhabiting this micro-environment are key players in Fe chemical speciation and the associated bioavailability of particulate Fe.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology* **215**: 403–410.

Andrews, S.C., Robinson, A.K., and Rodríguez-Quiñones, F. (2003) Bacterial iron homeostasis. *FEMS Microbiology Reviews* **27**: 215–237.

Armstrong, R.A., Peterson, M.L., Lee, C., and Wakeham, S.G. (2009) Settling velocity spectra and the ballast ratio hypothesis. *Deep Sea Research Part II: Topical Studies in Oceanography* **56**: 1470–1478.

Ayala-Castro, C., Saini, A., and Outten, F.W. (2008) Fe-S Cluster Assembly Pathways in Bacteria. *Microbiol Mol Biol Rev* **72**: 110–125.

Azam, F. and Malfatti, F. (2007) Microbial structuring of marine ecosystems. *Nature Reviews Microbiology* **5**: 782–791.

Baker, B.J., Sheik, C.S., Taylor, C.A., Jain, S., Bhasi, A., Cavalcoli, J.D., and Dick, G.J. (2013) Community transcriptomic assembly reveals microbes that contribute to deep-sea carbon and nitrogen cycling. *ISME J* **7**: 1962–1973.

Beier, S., Rivers, A.R., nn Moran, M.A., and Obernosterer, I. (2015) The transcriptional response of prokaryotes to phytoplankton-derived dissolved organic matter in seawater. *Environmental microbiology* **17**: 3466–3480.

Bishop, J.K.B., Lam, P.J., and Wood, T.J. (2012) Getting good particles: Accurate sampling of particles by large volume in-situ filtration. *Limnology and Oceanography: Methods* **10**: 681–710.

Bižić-Ionescu, M., Ionescu, D., and Grossart, H.-P. (2018) Organic Particles: Heterogeneous Hubs for Microbial Interactions in Aquatic Ecosystems. *Front Microbiol* **9**:.

Blain, S., Queguiner, B., Armand, L.K., Belviso, S., and Bomb, B. (2007) Effect of natural iron fertilization on carbon sequestration in the Southern Ocean. *Nature* **446**: 1070–1074.

Blain, S. and Tagliabue, A. (2016) Iron Cycle in Oceans, 1st ed. Wiley-ISTE.

Boiteau, R.M., Mende, D.R., Hawco, N.J., McIlvin, M.R., Fitzsimmons, J.N., Saito, M.A., et al. (2016) Siderophore-based microbial adaptations to iron scarcity across the eastern Pacific Ocean. *PNAS* **113**: 14237–14242.

Bolger, A.M., Lohse, M., and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Bowie, A.R., Townsend, A.T., Lannuzel, D., Remenyi, T.A., and van der Merwe, P. (2010) Modern sampling and analytical methods for the determination of trace elements in marine particulate material using magnetic sector inductively coupled plasma-mass spectrometry. *Analytica Chimica Acta* **676**: 15–27.

Boyd, P.W. and Ellwood, M.J. (2010) The biogeochemical cycle of iron in the ocean. *Nature Geoscience* **3**: 675–682.

Boyd, P.W., Ellwood, M.J., Tagliabue, A., and Twining, B.S. (2017) Biotic and abiotic retention, recycling and remineralization of metals in the ocean. *Nature Publishing Group* **10**:.

Boyd, P.W., Ibisanmi, E., Sander, S.G., Hunter, K.A., and Jackson, G.A. (2010) Remineralization of upper ocean particles: Implications for iron biogeochemistry. *Limnology and Oceanography* **55**: 1271–1288.

Boyd, P.W., Watson, A.J., Law, C.S., Abraham, E.R., Trull, T., Murdoch, R., et al. (2000) A mesoscale phytoplankton bloom in the polar Southern Ocean stimulated by iron fertilization. *Nature* **407**: 695–702.

Bruland, K.W., Orians, K.J., and Cowen, J.P. (1994) Reactive trace metals in the stratified central North Pacific. *Geochimica et Cosmochimica Acta* **58**: 3171–3182.

Buchfink, B., Xie, C., and Huson, D.H. (2015) Fast and sensitive protein alignment using DIAMOND. *Nature Methods* **12**: 59–60.

Butler, A. and Theisen, R.M. (2010) Iron(III)–siderophore coordination chemistry: Reactivity of marine siderophores. *Coordination Chemistry Reviews* **254**: 288–296.

Carini, P., Campbell, E.O., Morré, J., Sañudo-Wilhelmy, S.A., Cameron Thrash, J., Bennett, S.E., et al. (2014) Discovery of a SAR11 growth requirement for thiamin's pyrimidine precursor and its distribution in the Sargasso Sea. *The ISME Journal* **8**: 1727–1738.

Cornelis, P. and Andrews, S.C. eds. (2010) Iron uptake and homeostasis in microorganisms, Norfolk, UK: Caister Academic Press.

Cottrell, M.T. and Kirchman, D.L. (2000) Community composition of marine bacterioplankton determined by 16S rRNA gene clone libraries and fluorescence in situ hybridization. *Appl Environ Microbiol* **66**: 5116–5122.

Croot, P.L. and Heller, M.I. (2012) The Importance of Kinetics and Redox in the Biogeochemical Cycling of Iron in the Surface Ocean. *Front Microbiol* **3**:.

Crosa, J.H., Mey, A.R., and Payne, S.M. (2004) Iron Transport in Bacteria, American Society of Microbiology.

Debeljak, P., Toulza, E., Beier, S., Blain, S., and Obernosterer, I. (2019) Microbial iron metabolism as revealed by gene expression profiles in contrasted Southern Ocean regimes. *Environmental Microbiology*.

DeLong, E.F., Franks, D.G., and Alldredge, A.L. (1993) Phylogenetic diversity of aggregate-attached vs. free-living marine bacterial assemblages. *Limnology and Oceanography* **38**: 924–934.

Donat, J.R. and Bruland, K.W. (1995) Trace elements in natural waters. *ed B Salbu and E Steinness*.

Ferguson, A.D. and Deisenhofer, J. (2004) Metal Import through Microbial Membranes. *Cell* **116**: 15–24.

Fourquez, M., Beier, S., Jongmans, E., Hunter, R., and Obernosterer, I. (2016) Uptake of Leucine, Chitin, and Iron by Prokaryotic Groups during Spring Phytoplankton Blooms Induced by Natural Iron Fertilization off Kerguelen Island (Southern Ocean). *Frontiers in Marine Science* **3**:.

Ganesh, S., Parris, D.J., DeLong, E.F., and Stewart, F.J. (2014) Metagenomic analysis of size-fractionated picoplankton in a marine oxygen minimum zone. *ISME J* **8**: 187–211.

Gardner, W.D., Walsh, I.D., and Richardson, M.J. (1993) Biophysical forcing of particle production and distribution during a spring bloom in the North Atlantic. *Deep Sea Research Part II: Topical Studies in Oceanography* **40**: 171–195.

Gifford, S.M., Becker, J.W., Sosa, O.A., Repeta, D.J., and DeLong, E.F. (2016) Quantitative Transcriptomics Reveals the Growth- and Nutrient-Dependent Response of a Streamlined Marine Methylotroph to Methanol and Naturally Occurring Dissolved Organic Matter. *mBio* **7**: e01279-16.

Gram, L., Grossart, H.-P., Schlingloff, A., and Kiørboe, T. (2002) Possible Quorum Sensing in Marine Snow Bacteria: Production of Acylated Homoserine Lactones by Roseobacter Strains Isolated from Marine Snow. *Appl Environ Microbiol* **68**: 4111–4116.

Grossart, H.-P., Kiørboe, T., Tang, K., and Ploug, H. (2003) Bacterial Colonization of Particles: Growth and Interactions. *Appl Environ Microbiol* **69**: 3500–3509.

Held, N.A., McIlvin, M.R., Moran, D.M., Laub, M.T., and Saito, M.A. (2019) Unique Patterns and Biogeochemical Relevance of Two-Component Sensing in Marine Bacteria. *mSystems* **4**:.

Herndl, G.J. and Reinthaler, T. (2013) Microbial control of the dark end of the biological pump. *Nature Geoscience* **6**: 718–724.

Hider, R.C. and Kong, X. (2010) Chemistry and biology of siderophores. *Nat Prod Rep* **27**: 637–657.

Holmes, T.M., Wuttig, K., Chase, Z., van der Merwe, P., Townsend, A.T., Schallenberg, C., et al. (2019) Iron availability influences nutrient drawdown in the Heard and McDonald Islands region, Southern Ocean. *Marine Chemistry* **211**: 1–14.

Hopkinson, B.M. and Barbeau, K.A. (2012) Iron transporters in marine prokaryotic genomes and metagenomes. *Environmental Microbiology* **14**: 114–128.
Hopkinson, B.M. and Morel, F.M.M. (2009) The role of siderophores in iron acquisition by photosynthetic marine microorganisms. *BioMetals* **22**: 659–669.
Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M.C., et al. (2016) eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res* **44**: D286–D293.
Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S.K., Cook, H., et al. (2019) eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* **47**: D309–D314.
Huson, D.H., Auch, A.F., Qi, J., and Schuster, S.C. (2007) MEGAN analysis of metagenomic data. *Genome Research* **17**: 377–386.
Hyatt, D., Chen, G.-L., LoCascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**: 119.
Kanehisa, M. and Goto, S. (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research* **28**: 27–30.
Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016) KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Research* **44**: D457–D462.
Kanehisa, M., Sato, Y., and Morishima, K. (2016) BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. *J Mol Biol* **428**: 726–731.
Karner, M. and Herndl, G.J. (1992) Extracellular enzymatic activity and secondary production in free-living and marine-snow-associated bacteria. *Marine Biology* **113**: 341–347.
Kiørboe, T. (2001) Marine snow, organic solute plumes, and optimal chemosensory behavior of bacteria. *Limnology and Oceanography* **46**: 1309–1318.
Kirchman, D.L. (1996) Microbial ferrous wheel. *Nature* **383**: 303–304.
Kopylova, E., Noé, L., and Touzet, H. (2012) SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**: 3211–3217.
Landa, M., Burns, A.S., Roth, S.J., and Moran, M.A. (2017) Bacterial transcriptome remodeling during sequential co-culture with a marine dinoflagellate and diatom. *ISME J* **11**: 2677–2690.
Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**: 357–359.
Li, D., Liu, C.-M., Luo, R., Sadakane, K., and Lam, T.-W. (2015) MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**: 1674–1676.
Li, M., Jain, S., and Dick, G.J. (2016) Genomic and Transcriptomic Resolution of Organic Matter Utilization Among Deep-Sea Bacteria in Guaymas Basin Hydrothermal Plumes. *Front Microbiol* **7**:.
Li, M., Toner, B.M., Baker, B.J., Breier, J.A., Sheik, C.S., and Dick, G.J. (2014) Microbial iron uptake as a mechanism for dispersing iron from deep-sea hydrothermal vents. *Nature Communications* **5**: 3192.
Liao, Y., Smyth, G.K., and Shi, W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923–930.
Lis, H., Shaked, Y., Kranzler, C., Keren, N., and Morel, F.M.M. (2015) Iron bioavailability to phytoplankton: an empirical approach. *Isme Journal* **9**: 1003–1013.
Love, M.I., Huber, W., and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**:.
Marie, D., Simon, N., Guillou, L., Partensky, F., and Vaulot, D. (2000) Flow Cytometry Analysis of Marine Picoplankton. In, Diamond,R.A. and Demaggio,S. (eds), *In Living Color: Protocols in Flow Cytometry and Cell Sorting*, Springer Lab Manuals. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 421–454.
Martin, J.H., Gordon, R.M., and Fitzwater, S.E. (1990) Iron in Antarctic waters. *Nature* **345**: 156.
Massé, E. and Gottesman, S. (2002) A small RNA regulates the expression of genes involved in iron metabolism in Escherichia coli. *Proc Natl Acad Sci USA* **99**: 4620–4625.

McMurdie, P.J. and Holmes, S. (2013) phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE* **8**: e61217.
McRose, D.L., Baars, O., Seyedsayamdost, M.R., and Morel, F.M.M. (2018) Quorum sensing and iron regulate a two-for-one siderophore gene cluster in Vibrio harveyi. *PNAS* **115**: 7581–7586.
van der Merwe, P., Wuttig, K., Holmes, T., Trull, T.W., Chase, Z., Townsend, A.T., et al. (2019) High Lability Fe Particles Sourced From Glacial Erosion Can Meet Previously Unaccounted Biological Demand: Heard Island, Southern Ocean. *Front Mar Sci* **6**:.
Morel, F.M.M., Milligan, A.J., and Saito, M.A. (2003) 6.05 Marine Bioinorganic Chemistry: The Role of Trace Metals in the Oceanic Cycles of Major Nutrients. 31.
Morel, F.M.M. and Price, N.M. (2003) The biogeochemical cycles of trace metals in the oceans. *Science* **300**: 944–947.
Moynié, L., Milenkovic, S., Mislin, G.L.A., Gasser, V., Malloci, G., Baco, E., et al. (2019) The complex of ferric-enterobactin with its transporter from Pseudomonas aeruginosa suggests a two-site model. *Nat Commun* **10**: 1–14.
Neilands, J.B. (1995) Siderophores: Structure and Function of Microbial Iron Transport Compounds. *J Biol Chem* **270**: 26723–26726.
Noinaj, N., Guillier, M., Barnard, T.J., and Buchanan, S.K. (2010) TonB-Dependent Transporters: Regulation, Structure, and Function. *Annual Review of Microbiology* **64**: 43–60.
Obernosterer, I., Catala, P., Lebaron, P., and West, N.J. (2011) Distinct bacterial groups contribute to carbon cycling during a naturally iron fertilized phytoplankton bloom in the Southern Ocean. *Limnology and Oceanography* **56**: 2391–2401.
Obernosterer, I., Christaki, U., Lefèvre, D., Catala, P., Wambeke, F. Van, and Lebaron, P. (2008) Rapid bacterial mineralization of organic carbon produced during a phytoplankton bloom induced by natural iron fertilization in the Southern Ocean. *Deep Sea Res II* **55**: 777–789.
Panzeca, C., Tovar-Sanchez, A., Agustí, S., Reche, I., Duarte, C.M., Taylor, G.T., and Sañudo-Wilhelmy, S.A. (2006) B vitamins as regulators of phytoplankton dynamics. *Eos, Transactions American Geophysical Union* **87**: 593.
Parada, A.E., Needham, D.M., and Fuhrman, J.A. (2016) Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental Microbiology* **18**: 1403–1414.
Ploug, H. (2001) Small-scale oxygen fluxes and remineralization in sinking aggregates. *Limnology and Oceanography* **46**: 1624–1631.
Ploug, H. and Bergkvist, J. (2015) Oxygen diffusion limitation and ammonium production within sinking diatom aggregates under hypoxic and anoxic conditions. *Marine Chemistry* **176**: 142–149.
Ploug, M, K., B, B.-C., and Bb, J. (1997) Anoxic aggregates - an ephemeral phenomenon in the pelagic environment? *Aquatic Microbial Ecology* **13**: 285–294.
Poole, K., Young, L., and Neshat, S. (1990) Enterobactin-mediated iron transport in Pseudomonas aeruginosa. *J Bacteriol* **172**: 6991–6996.
Rauschenberg, S. and Twining, B.S. (2015) Evaluation of approaches to estimate biogenic particulate trace metals in the ocean. *Marine Chemistry* **171**: 67–77.
Roey (2012) Total Nucleic Acid Extraction from Soil.
Sandy, M. and Butler, A. (2009) Microbial Iron Acquisition: Marine and Terrestrial Siderophores. *Chem Rev* **109**: 4580–4595.
Satinsky, B.M., Gifford, S.M., Crump, B.C., and Moran, M.A. (2013) Use of Internal Standards for Quantitative Metatranscriptome and Metagenome Analysis. In, *Methods in enzymology*., pp. 237–250.
Satinsky, B.M., Zielinski, B.L., Doherty, M., Smith, C.B., Sharma, S., Paul, J.H., et al. (2010) The Amazon continuum dataset: quantitative metagenomic and metatranscriptomic inventories of the Amazon River plume, June 2010.
Schlitzer, R., Anderson, R.F., Dodas, E.M., Lohan, M., Geibert, W., Tagliabue, A., et al. (2018) The GEOTRACES Intermediate Data Product 2017. *Chemical Geology* **493**: 210–223.
Sigman, D.M. and Hain, M.P. (2012) The Biological Productivity of the Ocean. **3**: 16.

Sijerčić, A. and Price, N. (2015) Hydroxamate siderophore secretion by Pseudoalteromonas haloplanktis during steady-state and transient growth under iron limitation. *Marine Ecology Progress Series* **531**: 105–120.
Simon, H.M., Smith, M.W., and Herfort, L. (2014) Metagenomic insights into particles and their associated microbiota in a coastal margin ecosystem. *Front Microbiol* **5**:.
Simon, M., Grossart, H.-P., Schweitzer, B., and Ploug, H. (2002) Microbial ecology of organic aggregates in aquatic ecosystems. *Aquatic Microbial Ecology* **28**: 175–211.
Smith, D.C. and Azam, F. (1992) A simple, economical method for measuring bacterial protein synthesis rates in sea water using 3H -Leucine. *Mar Microb Food Webs* **6**: 107–114.
Smith, M.W., Zeigler Allen, L., Allen, A.E., Herfort, L., and Simon, H.M. (2013) Contrasting genomic properties of free-living and particle-attached microbial assemblages within a coastal ecosystem. *Front Microbiol* **4**: 120.
Steinberg, D.K., Mooy, B.A.S.V., Buesseler, K.O., Boyd, P.W., Kobari, T., and Karl, D.M. (2008) Bacterial vs. zooplankton control of sinking particle flux in the ocean's twilight zone. *Limnology and Oceanography* **53**: 1327–1338.
Stocker, R. (2012) Marine Microbes See a Sea of Gradients. *Science* **338**: 628–633.
Stocker, R. (2015) The 100 µm length scale in the microbial ocean. *Aquatic Microbial Ecology* **76**: 189–194.
Stocker, R., Seymour, J.R., Samadani, A., Hunt, D.E., and Polz, M.F. (2008) Rapid chemotactic response enables marine bacteria to exploit ephemeral microscale nutrient patches. *PNAS* **105**: 4209–4214.
Tang, K., Jiao, N., Liu, K., Zhang, Y., and Li, S. (2012) Distribution and functions of tonb-dependent transporters in marine bacteria and environments: Implications for dissolved organic matter utilization. *PLoS ONE* **7**:.
Toulza, E., Tagliabue, A., Blain, S., and Piganeau, G. (2012) Analysis of the global ocean sampling (GOS) project for trends in iron uptake by surface ocean microbes. *PLoS ONE* **7**:.
Velasquez, I.B., Ibisanmi, E., Maas, E.W., Boyd, P.W., Nodder, S., and Sander, S.G. (2016) Ferrioxamine Siderophores Detected amongst Iron Binding Ligands Produced during the Remineralization of Marine Particles. *Front Mar Sci* **3**:.
Völker, C. and Wolf-Gladrow, D.A. (1999) Physical limits on iron uptake mediated by siderophores or surface reductases. *Marine Chemistry* **65**: 227–244.
Wächtershäuser, G. (1992) Groundworks for an evolutionary biochemistry: the iron-sulphur world. *Prog Biophys Mol Biol* **58**: 85–201.
Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007) Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Appl Environ Microbiol* **73**: 5261–5267.
Weaver, R.S., Kirchman, D.L., and Hutchins, D.A. (2003) Utilization of iron/organic ligand complexes by marine bacterioplankton. *Aquatic Microbial Ecology* **31**: 227–239.
Wiener, M.C. (2005) TonB-dependent outer membrane transport: going for Baroque? *Current Opinion in Structural Biology* **15**: 394–400.

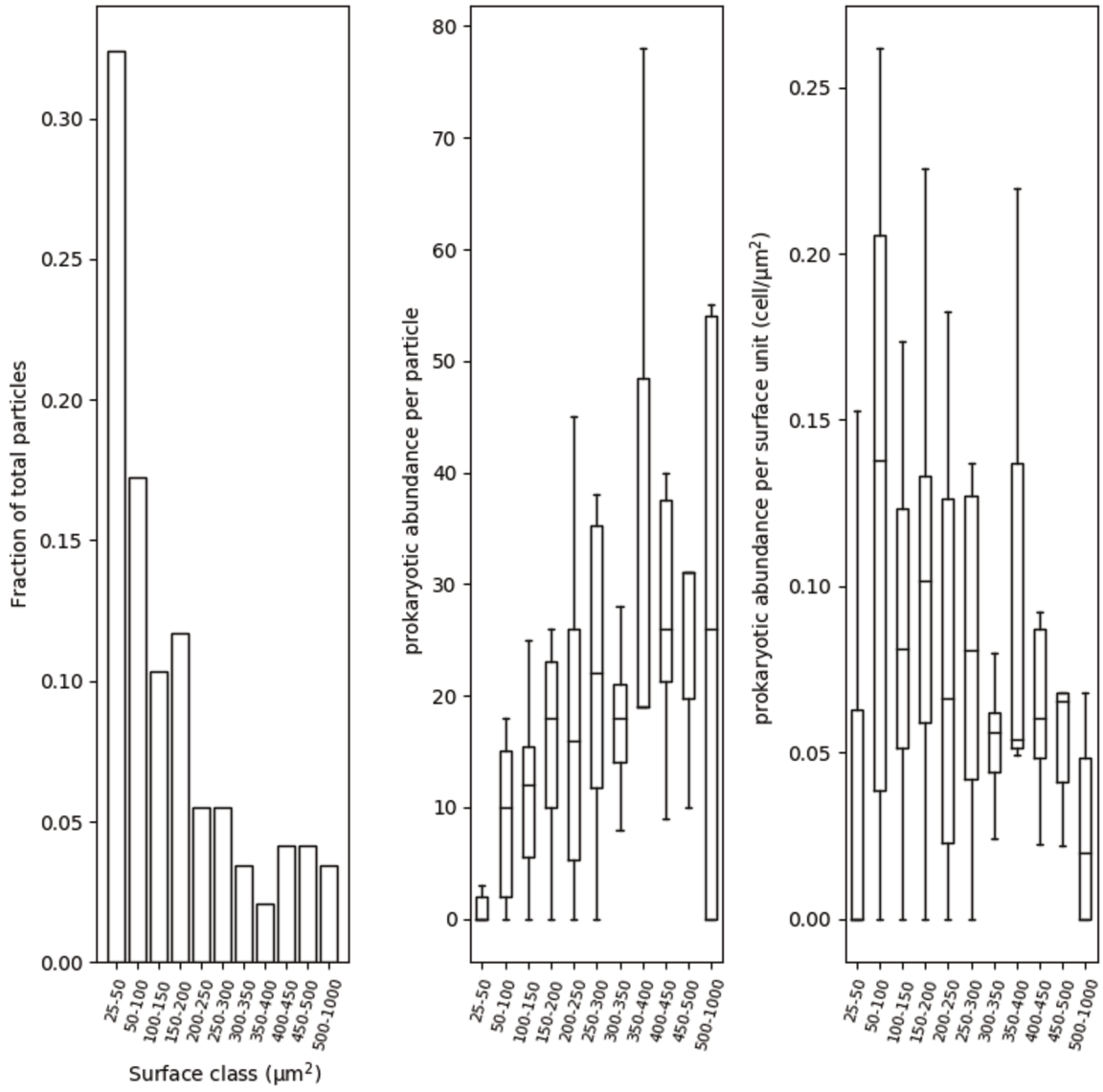
Chapter 2 Supplementary Tables & Figures

Supplementary Table 1 *Detailed information on amplicon sequences through the steps of the DADA2 workflow.*

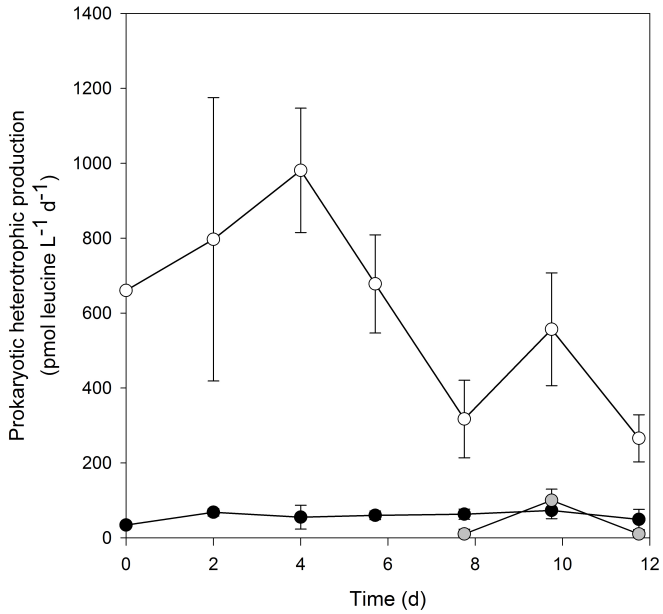
Sample code	Raw reads	Filtered	Denoised	Merged	Tabled	Nonchimeric
PA_1	43516	27080	27080	16459	16459	16459
PA_2	64753	40805	40805	25553	25553	25553
PA_3	49667	30994	30994	14620	14620	14593
PA_4	49258	30725	30725	13492	13492	13296
FLC_1	52876	34004	34004	24627	24627	24362
FLC_2	54477	34985	34985	25861	25861	25861
FLC_3	54926	35706	35706	27431	27431	27426
FLC-P1	36679	22689	22689	9482	9482	9482
FLC-P2	18136	11458	11458	6236	6236	5799
FLP_1	53138	33559	33559	23581	23581	23533
FLP_2	59776	37481	37481	25038	25038	25028
FLP_3	65775	41693	41693	29917	29917	29917
FLP_4	90847	58623	58623	43993	43993	43954
FLP_5	70834	45176	45176	32155	32155	32104

Supplementary Table 2 *Samples chosen for RNA sequencing with raw read counts before and after trimming.*

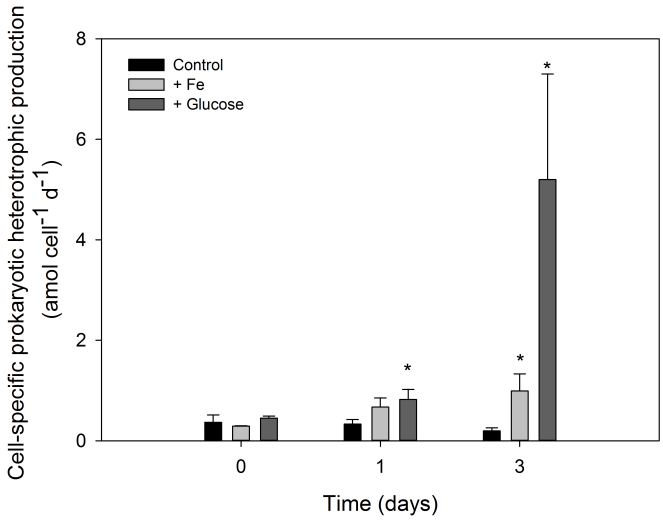
Sample code	Read	Raw	Adapter Trimmed	Sickel Trimmed	% of surviving reads
PA 7	R1	27853639	27244786	25024024	91.85
PA 7	R2	27853639	27244786	25024024	91.85
PA 8	R1	29149587	29039594	27678745	95.31
PA 8	R2	29149587	29039594	27678745	95.31
PA 9	R1	38714036	38535793	37292499	96.77
PA 9	R2	38714036	38535793	37292499	96.77
FLP 10	R1	32557386	32410197	30851772	95.19
FLP 10	R2	32557386	32410197	30851772	95.19
FLP 11	R1	21940087	21894483	21263978	97.12
FLP 11	R2	21940087	21894483	21263978	97.12
FLP 12	R1	26928106	26898976	25453528	94.63
FLP 12	R2	26928106	26898976	25453528	94.63
FLC-P 13	R1	33884697	33190763	32040019	96.53
FLC-P 13	R2	33884697	33190763	32040019	96.53
FLC-P 14	R1	32289912	31454280	30094793	95.68
FLC-P 14	R2	32289912	31454280	30094793	95.68
FLC 15	R1	27590625	27439358	26599999	96.94
FLC 15	R2	27590625	27439358	26599999	96.94



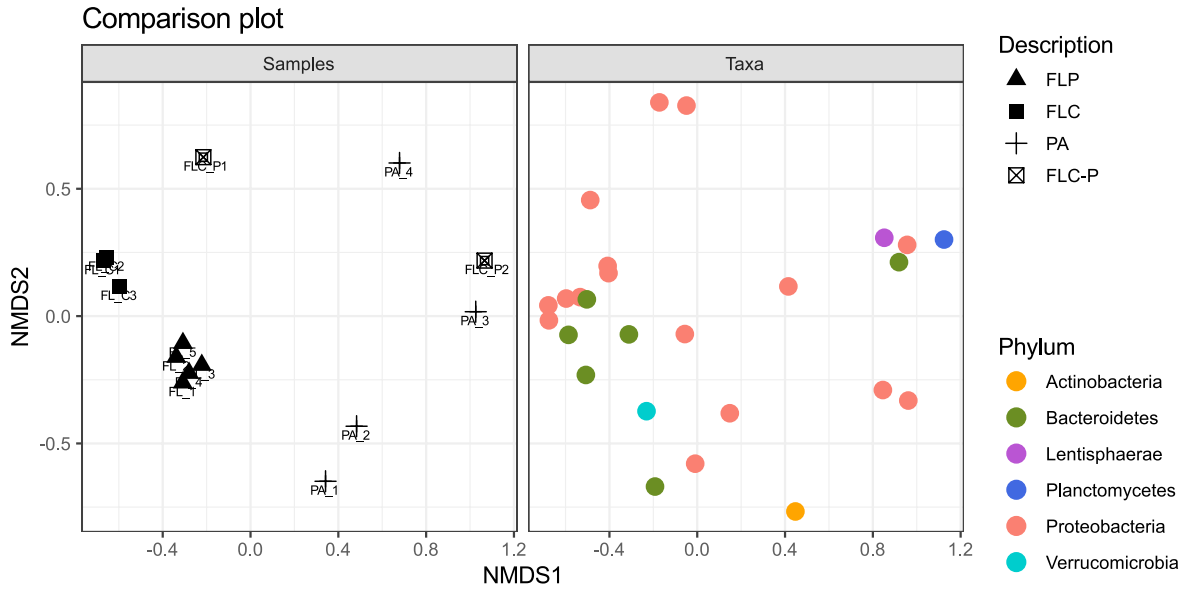
Supplementary Figure 1. Size class distribution of particles and prokaryotic abundance at the end of the experiment in one replicate treatment



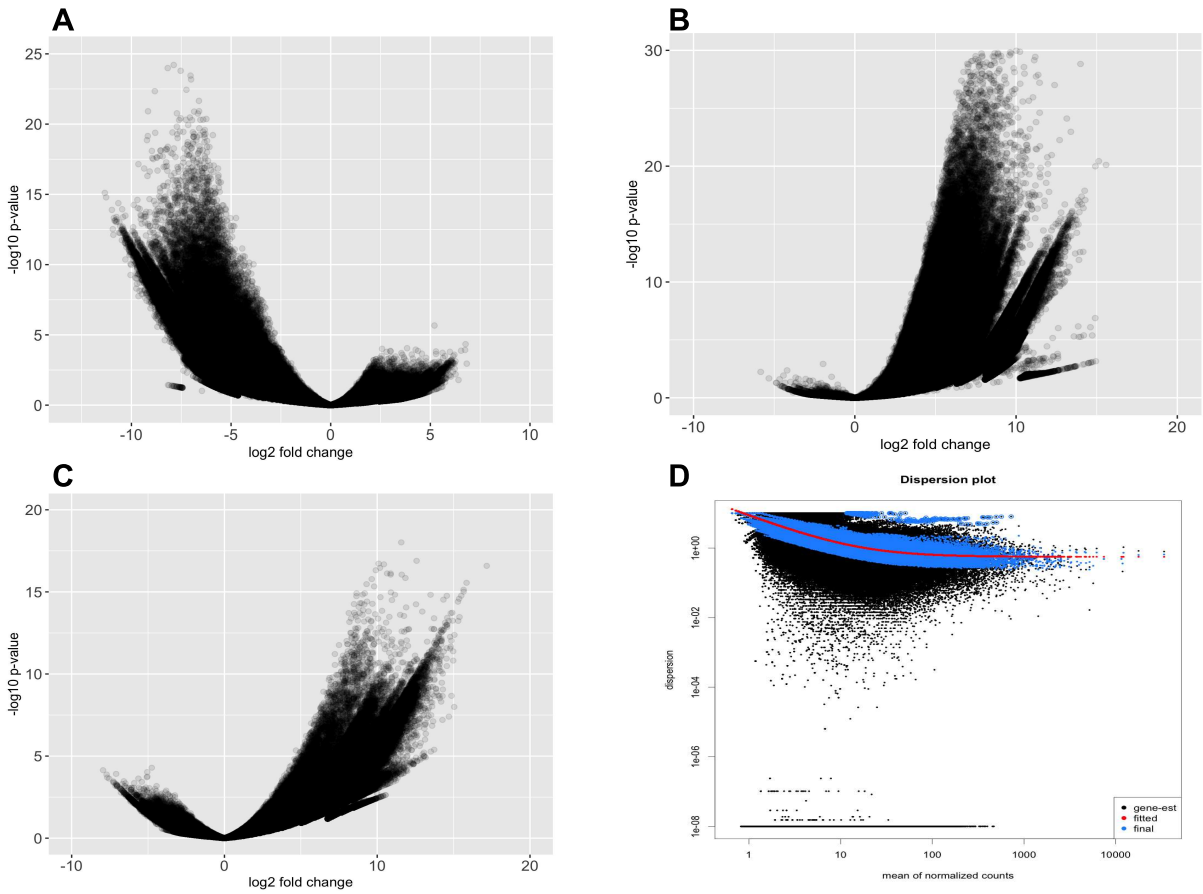
Supplementary Figure 2. *Prokaryotic heterotrophic production of particle-attached cells over the 12-day period. Bulk heterotrophic production is given for the control treatment. White circles – particle-amended treatment, grey circles- particle-amended treatment after 0.8 µm filtration, black circles – control treatment*



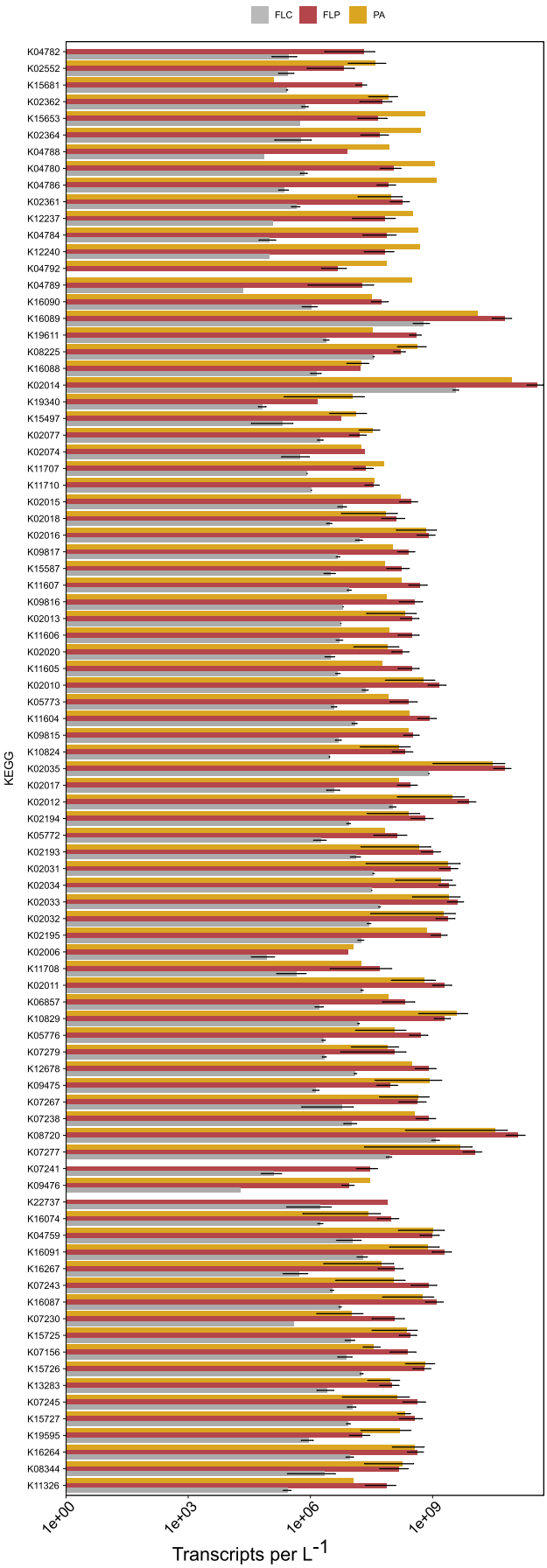
Supplementary Figure 3. *Cell-specific bacterial heterotrophic production for the dose addition experiment.*



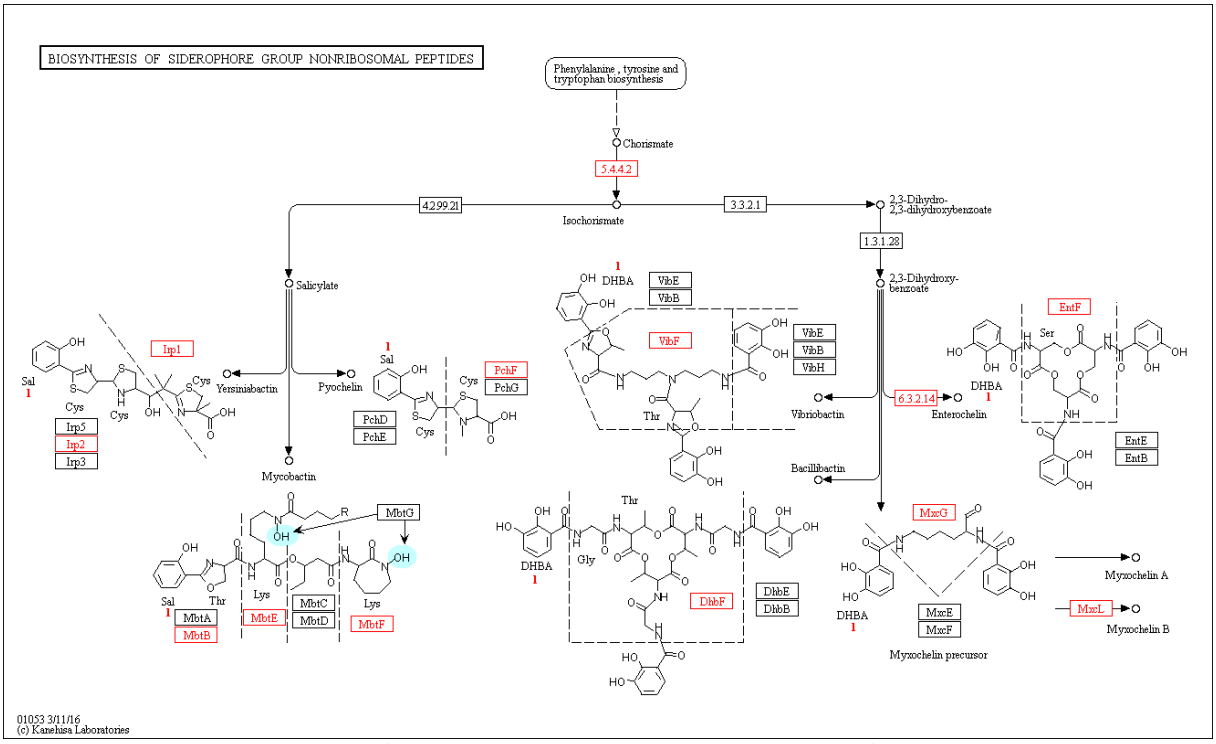
***Supplementary Figure 4**. Amplicon sequencing results. NMDS plot of bray Curtis similarities of 16S Amplicon sequencing variant in a comparison plot with associated phyla distribution.*



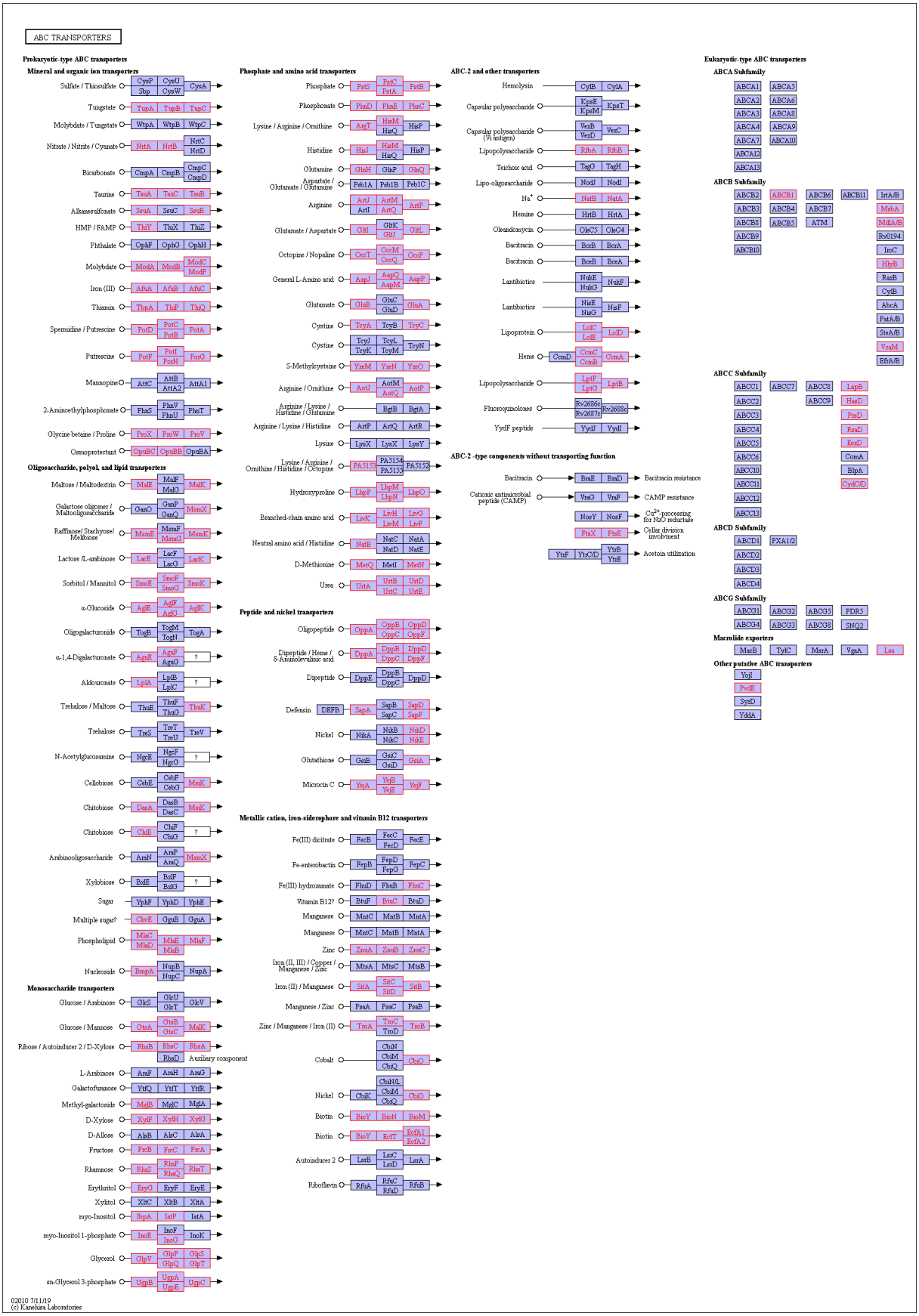
***Supplementary Figure 5.** DESeq2 result plots. Volcano plots of differentially expressed genes for specific comparisons at pvalue < 0.05 (Please note different scales) **A.** FLP vs PA. **B.** FLP vs FLC **C.** PA vs FLC **D.** Dispersion plot of fitted, estimated and final distribution.*



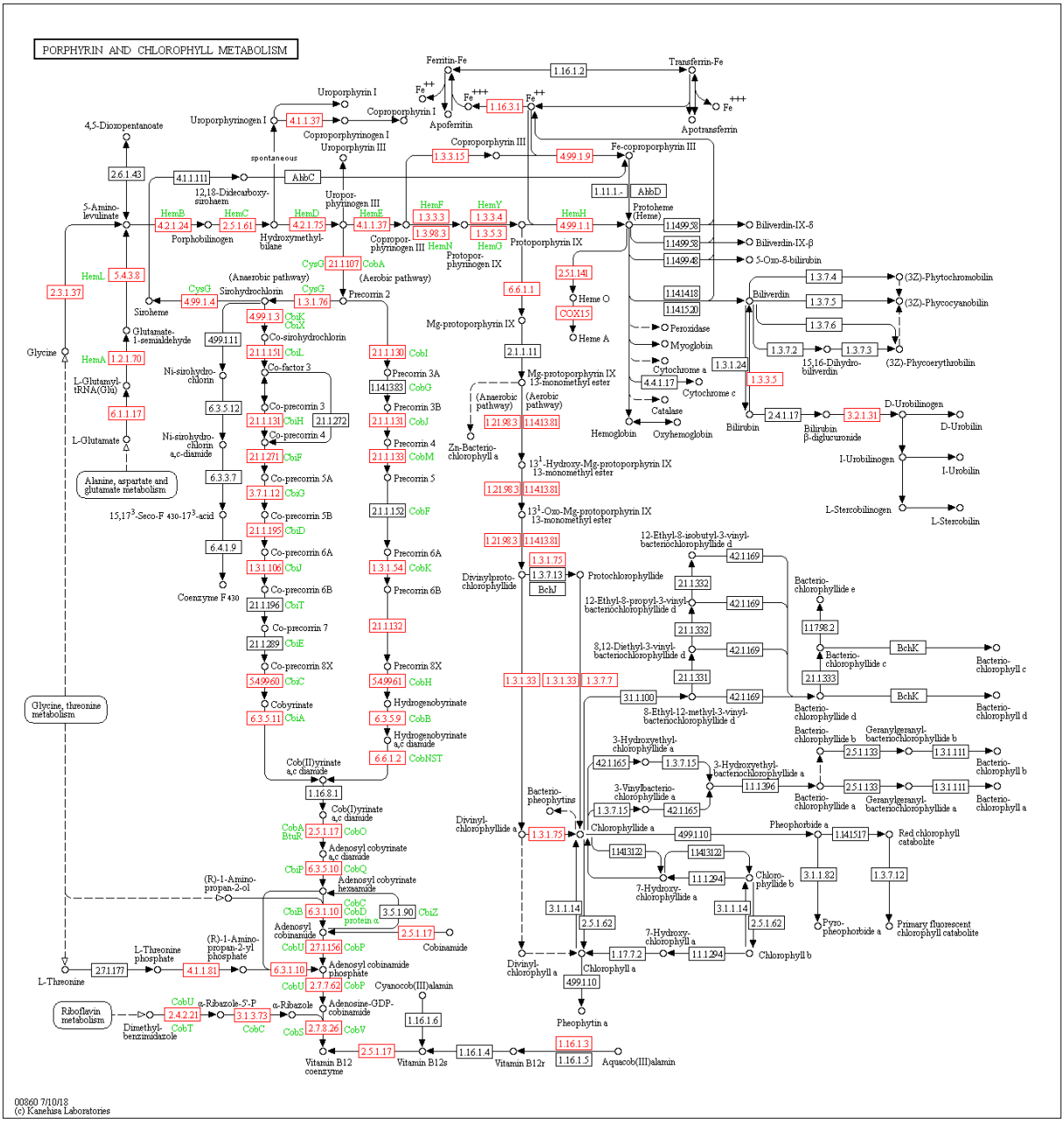
Supplementary Figure 6. *Transcripts per litre for Metal transporters and siderophore synthesis genes calculated using the formula in Satinsky et al. 2013. Mean values ± SD of triplicate are shown.*



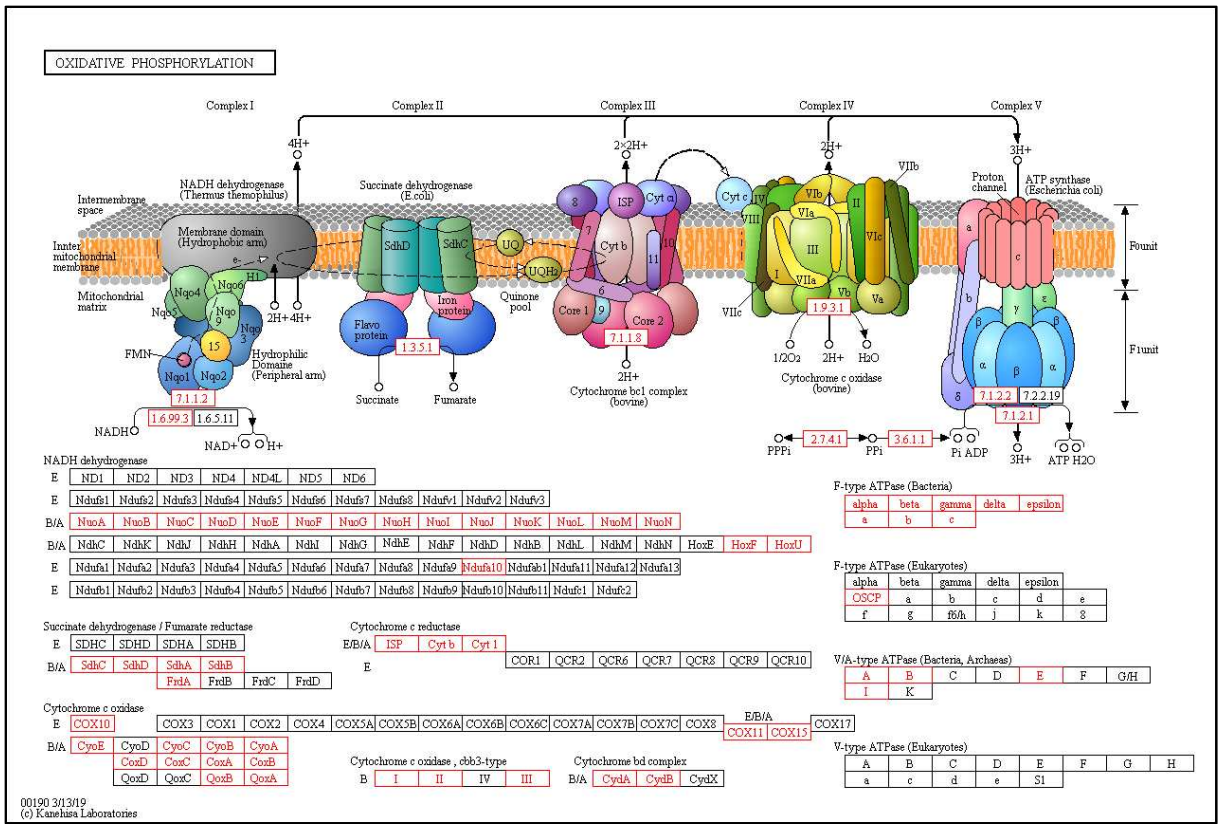
***Supplementary Figure** 7. KEGG Map for siderophore synthesis. Highlighted in red significantly higher expressed genes in FLP compared to FLC.*



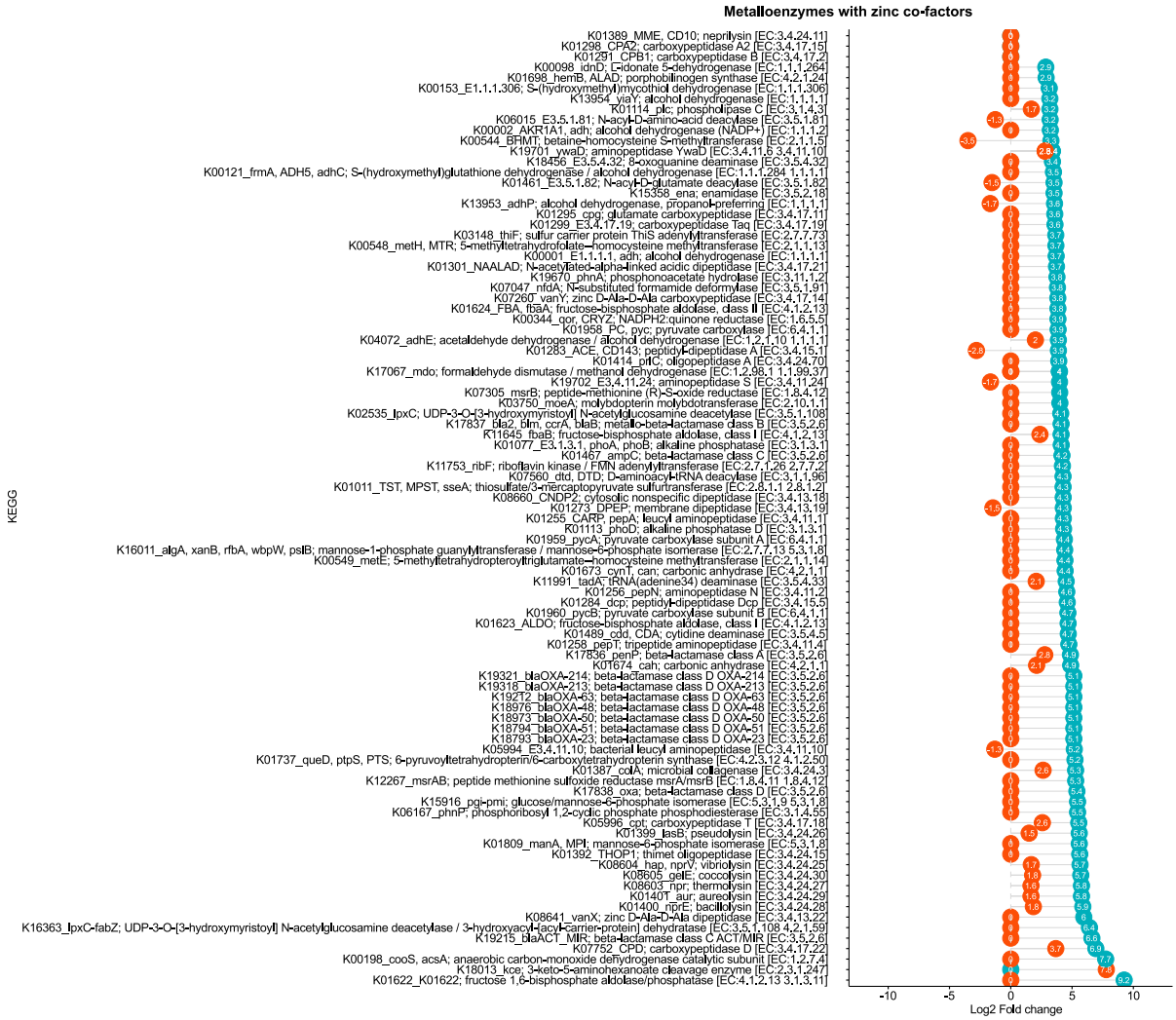
Supplementary Figure 8. *KEGG Map for ABC transporter. Highlighted in red significantly higher expressed genes in FLP compared to FLC.*



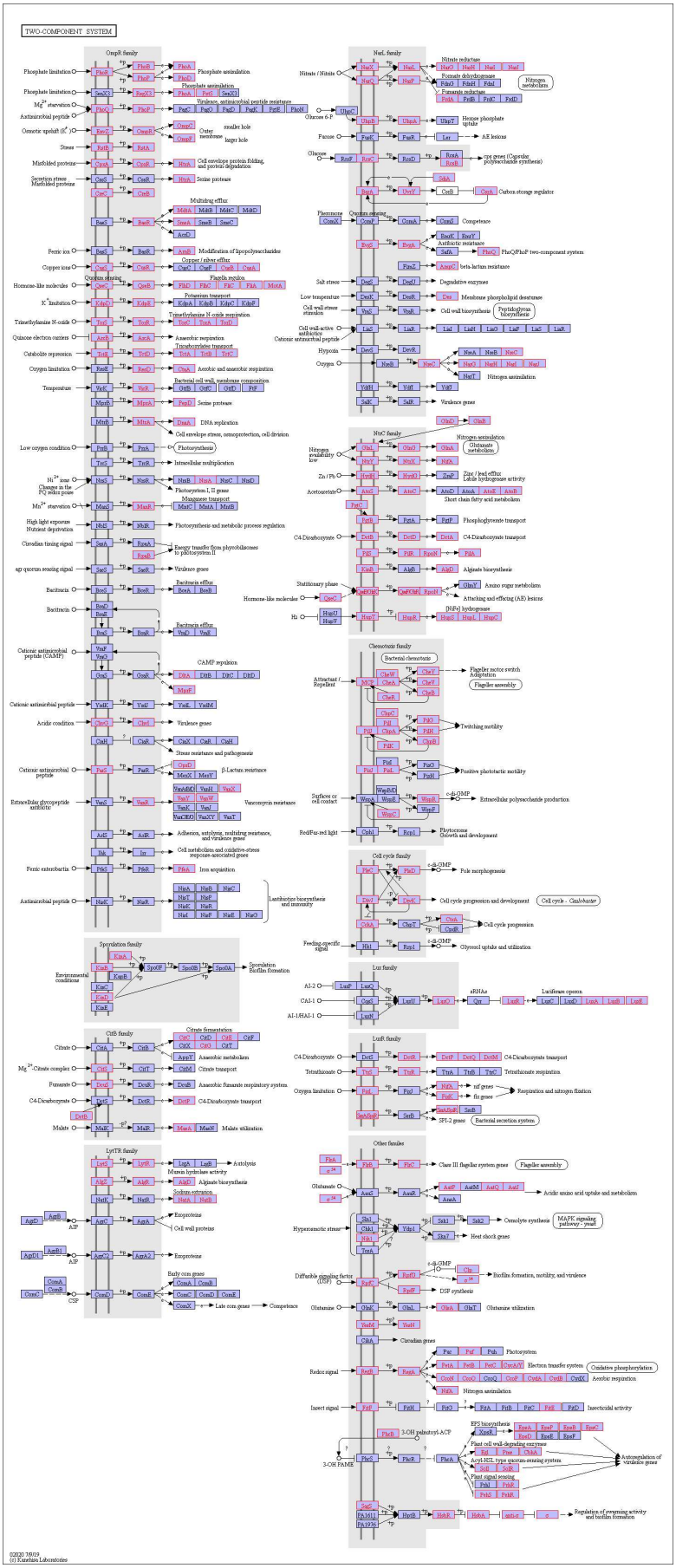
Supplementary Figure 9. *KEGG Map for Porphyrin and chlorophyll metabolism. Highlighted in red significantly higher expressed genes in FLP compared to FLC.*



Supplementary Figure 10. *KEGG Map for Oxidative phosphorylation. Highlighted in red significantly higher expressed genes in FLP compared to FLC.*

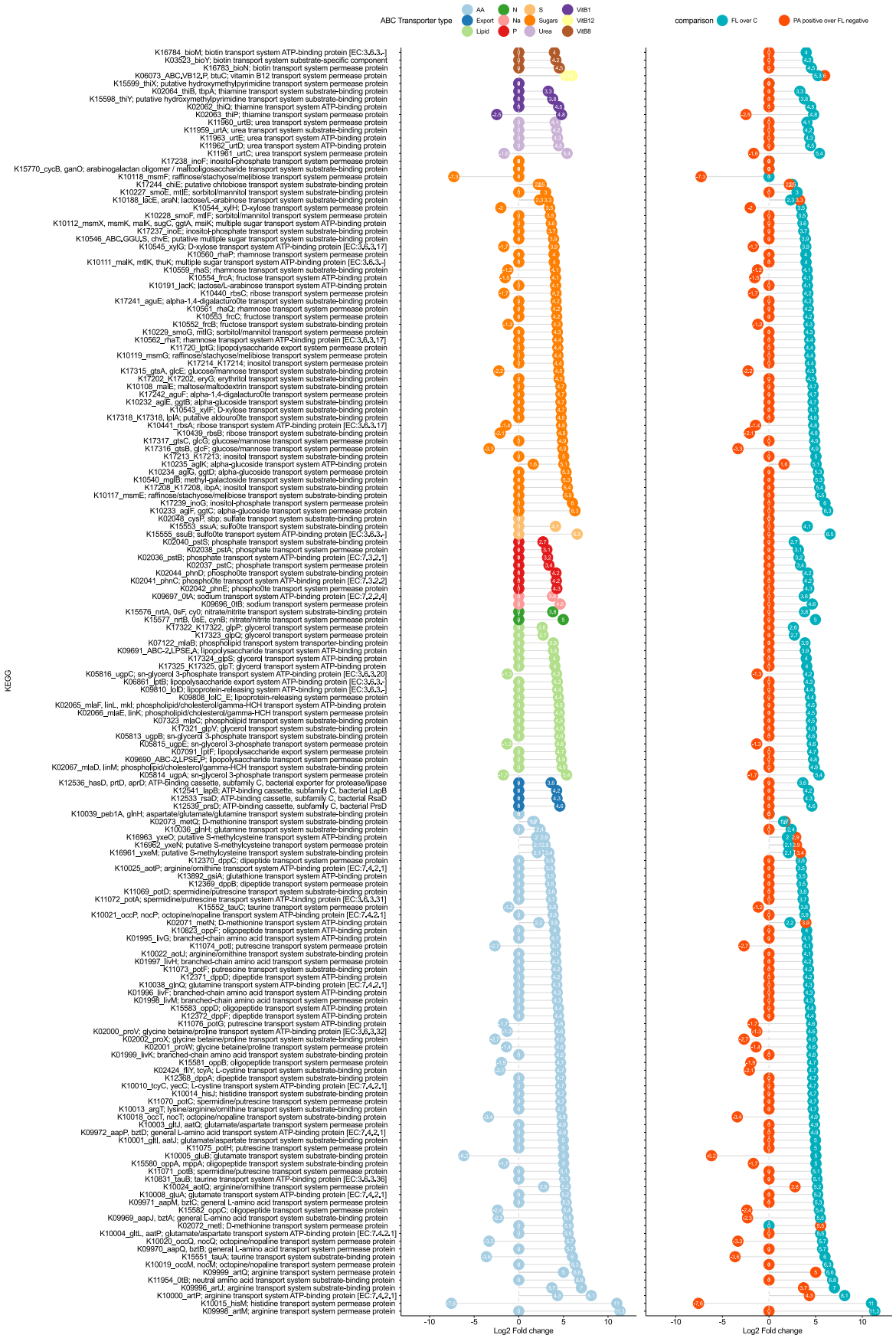


Supplementary Figure 11. *Log2 fold changes for differentially expressed genes for Zink as co-factor including several Metalloproteases. Positive values indicate higher expression in PA than in FLP (orange circles) and in FLP than in FLC (blue circles). Negative values indicate lower expression in the respective comparisons.*

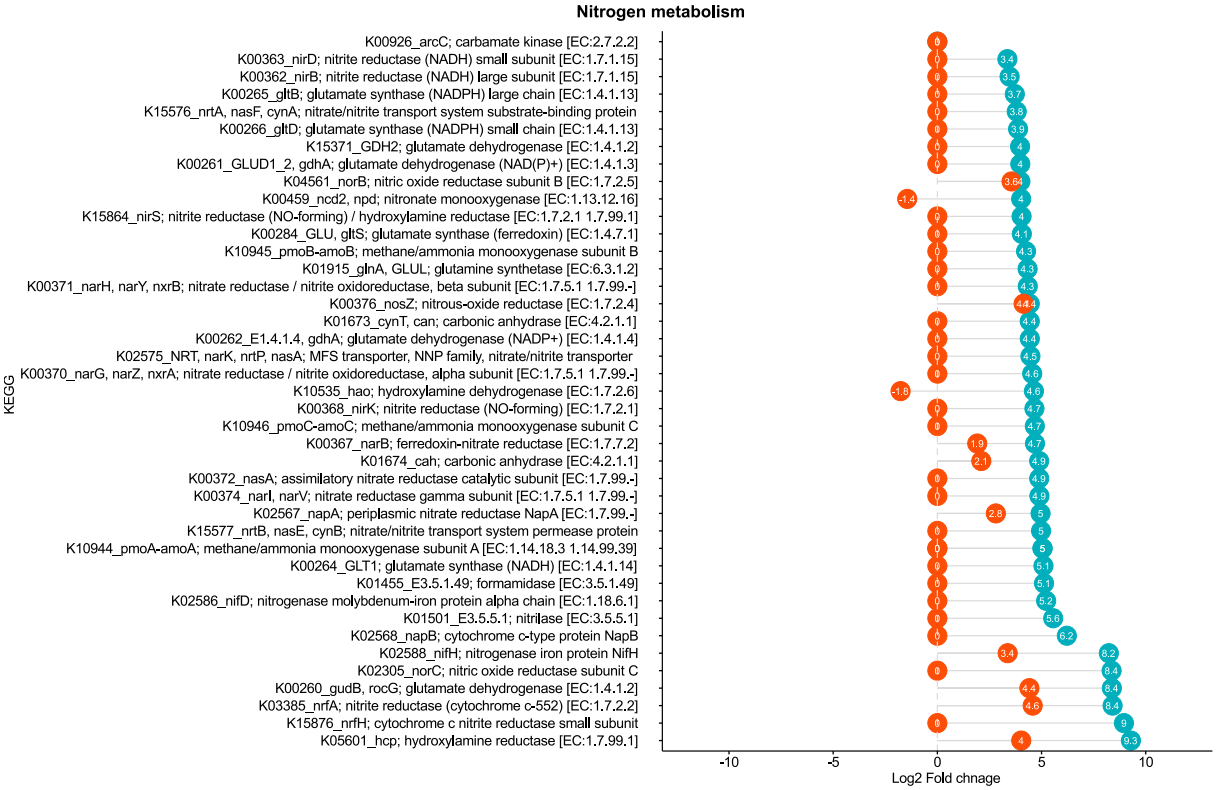


Supplementary Figure 12. *KEGG Map for Two-component regulatory systems. Highlighted in red significantly higher expressed genes in FLP compared to FLC.*

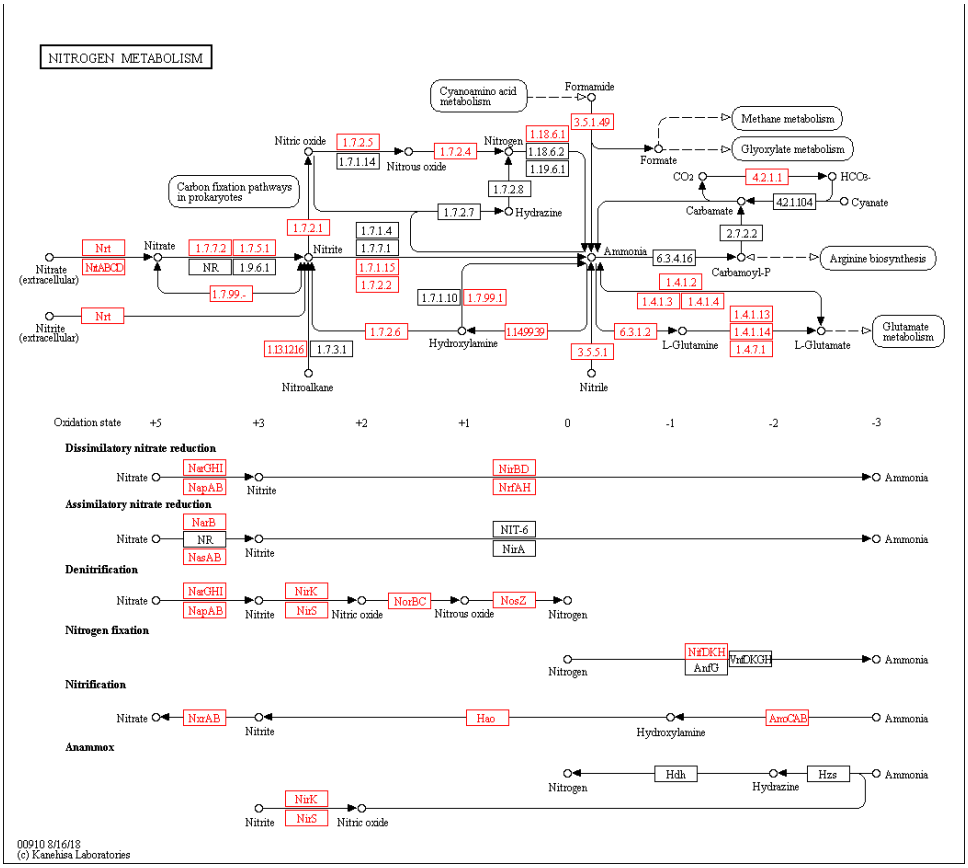
Supplementary Figure 13. *Chemotaxis and Biofilm proteins and their corresponding Log2 Fold change. Positive values indicate higher expression in PA than in FLP (orange circles) and in FLP than in FLC (blue circles). Negative values indicate lower expression in the respective comparisons.*



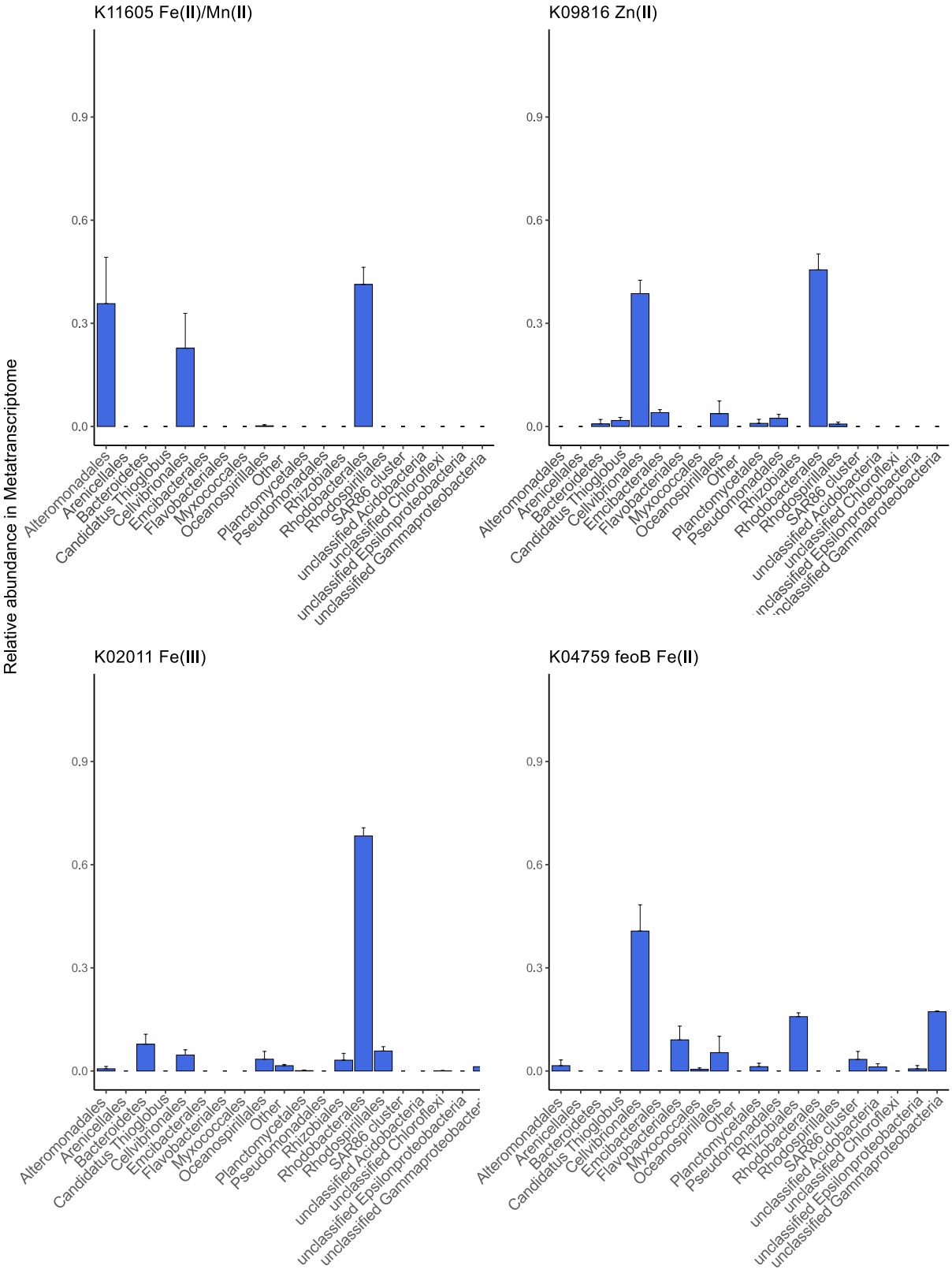
Supplementary Figure 14. *Detailed plot of ABC transporters and corresponding Log2 Fold changes in comparisons. Positive values indicate higher expression in PA than in FLP (orange circles) and in FLP than in FLC (blue circles). Negative values indicate lower expression in the respective comparisons. Compound specific transporters are ordered by colour.*



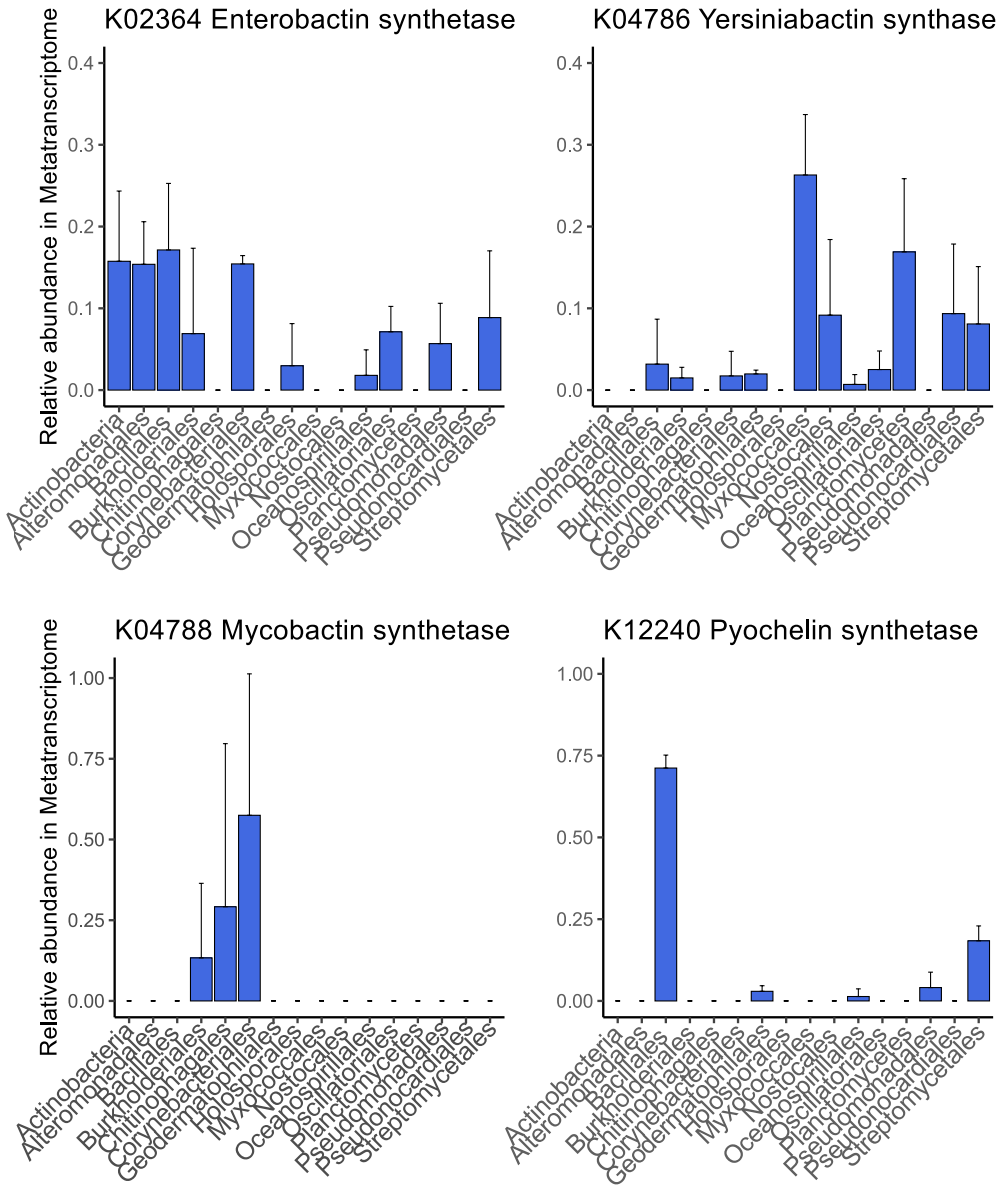
Supplementary Figure 15. *Detailed plot of genes in Nitrogen metabolism and corresponding Log2 Fold changes in comparisons. Positive values indicate higher expression in PA than in FLP (orange circles) and in FLP than in FLC (blue circles). Negative values indicate lower expression in the respective comparisons.*



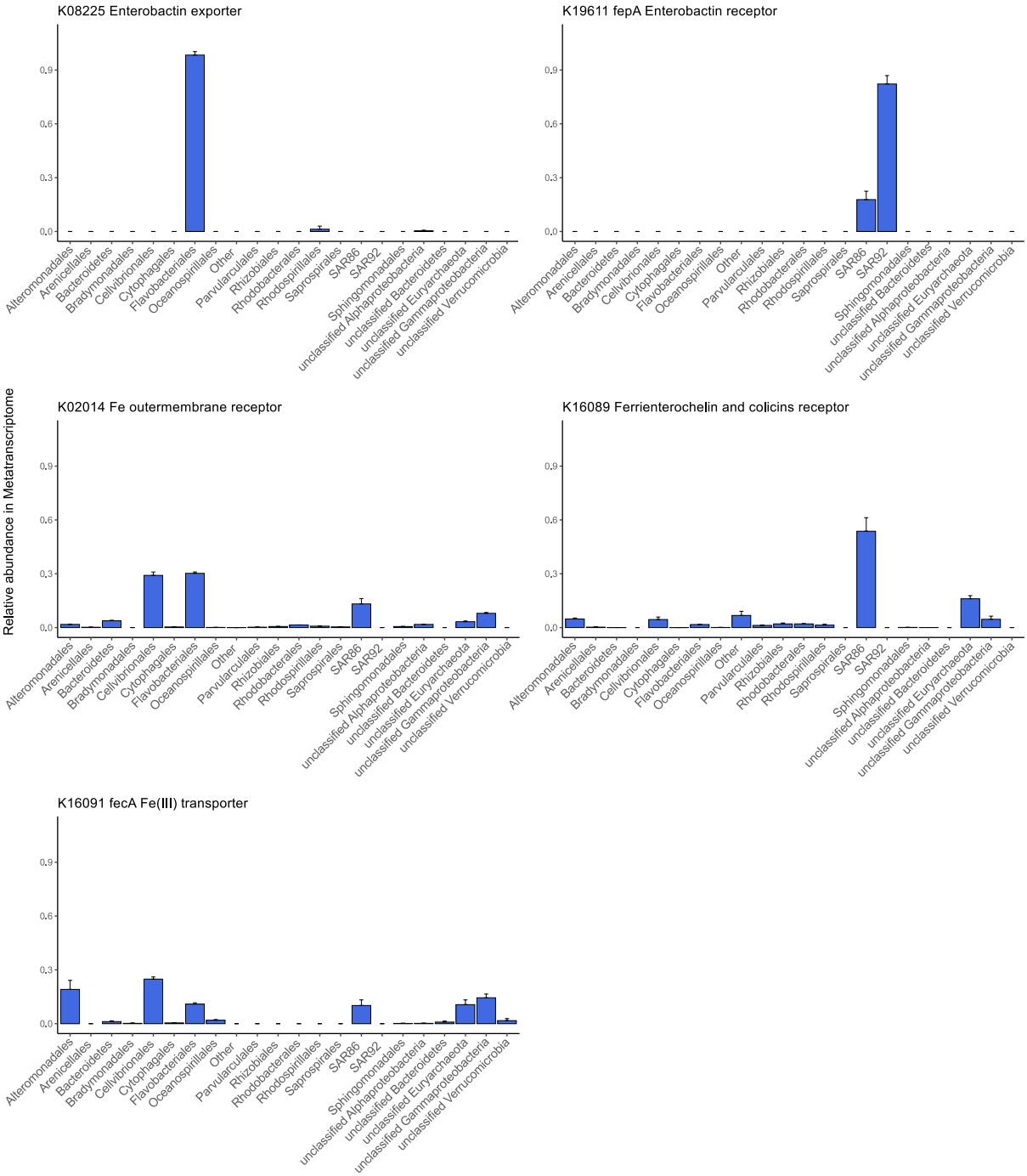
Supplementary Figure 16. *KEGG Map for Nitrogen metabolism. Highlighted in red significantly higher expressed genes in FLP compared to FLC*



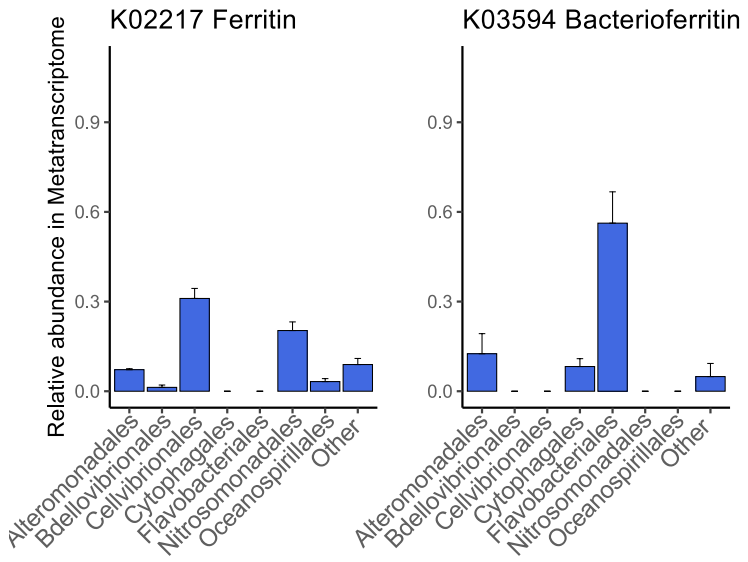
Supplementary Figure 17. *Examples of ion transporter genes and the relative contribution of taxa FLP (Free-living particle sphere) cells. Mean values ± SD of triplicates are shown.*



Supplementary Figure 18. *Examples of siderophore synthesis pathway genes and relative contribution of taxa FLP (Free-living particle sphere) cells. Mean values ± SD of triplicates are shown.*



Supplementary Figure 19. *Examples of TonB-dependent receptor genes and relative contribution of FLP (Free-living particle sphere) cells. Mean values ± SD of triplicates are shown.*



Supplementary Figure 20. *Bacterioferritin and Ferritin genes and relative contribution of FLP (Free-living particle sphere) cells. Mean values ± SD of triplicates are shown.*

Chapter 3

Deciphering organic matter utilization through prokaryotic membrane transporter expression at two Southern Ocean sites in early spring and late summer

Pavla Debeljak[1,2], Barbara Bayer[2], Ying Sun[1], Ingrid Obernosterer[1] & Gerhard J. Herndl[2,3]

[1] Sorbonne Université, CNRS, Laboratoire d'Océanographie Microbienne, LOMIC, F-66650 Banyuls/mer, France

[2] Department of Limnology and Bio-Oceanography, University of Vienna, A-1090 Vienna, Austria

[3]Department of Biological Oceanography, Royal Netherlands Institute for Sea Research (NIOZ), 1790 AB Den Burg, The Netherlands

Abstract

The availability of dissolved organic matter constrains heterotrophic microbial activity in surface waters of the Southern Ocean, characterized by generally low concentrations of organic carbon. In the region southeast of Kerguelen Island, the island mass effect leads to elevated iron concentrations in otherwise high-nutrient low-chlorophyll (HNLC) waters. Thus, it represents an attractive site to study locally occurring spring phytoplankton blooms and the response of heterotrophic microorganisms to phytoplankton-derived organic matter. To obtain insight into the types of organic substrates used by the microbial communities, we collected samples at sites with contrasting organic matter supply during two seasons, early austral spring and late austral summer. In order to analyse organic matter utilization, we focused on the expressed membrane transporter proteins. Comparative proteomics of transporter proteins suggest that ABC transporters are present throughout both seasons but increase in relative abundance in spring, while outer membrane receptors were detected at higher relative abundance in late summer, in both HNLC and iron-fertilized waters. The highest contribution of compound-specific transporters showed amino acid transporters, followed by sugar, polyamine, taurine and urea transporters, which were detected at high relative abundance in early spring metaproteomes. In contrast, iron transporters exhibited the highest relative abundances during late summer. Despite a high functional redundancy, more compound-specific transporters were expressed during early spring than in late summer by diverse taxa as revealed by high quality metagenome assembled genomes.

Introduction

The Southern Ocean has been extensively studied as a high-nutrient low-chlorophyll area, primarily limited by the micronutrient iron (Martin *et al.*, 1990a; 1990b; Tagliabue *et al.*, 2017). This limitation constrains microbial growth and results in an inefficient biological pump (Tortell *et al.*, 1996; Sarmiento and Gruber, 2006). In addition, the Southern Ocean exhibits low concentrations of dissolved organic carbon (DOC) (Hansell *et al.*, 2009), primarily due to low productivity which hinders the provision of fresh, labile carbon (Hansell, 2013; Bercovici and Hansell, 2016). This dual nutritional constraint has been tested in co-limitation experiments (Church *et al.*, 2000; Obernosterer *et al.*, 2015; Baltar *et al.*, 2018).

Although the largest fraction of DOC remains uncharacterized, phytoplankton exudates consist to a considerable extent of labile compounds such as carbohydrates (mono-, and polysaccharides), proteins and amino acids (Obernosterer and Herndl, 1995; Myklestad, 2000; Meon and Kirchman, 2001; Granum *et al.*, 2002). Hence, extracellular release of DOC by phytoplankton supports a major fraction of the labile carbon flux in the surface ocean thereby supporting heterotrophic microorganisms (Fogg, 1983; Baines and Pace, 1991; Morán *et al.*, 2013). The area around the Kerguelen island plateau is characterized by annually re-occurring phytoplankton blooms due to the supply of Fe by the island mass effect in otherwise HNLC waters. Diatom DOC has been observed as a key factor shaping bacterial communities in Southern Ocean during spring and summer (Straza *et al.*, 2010; Luria *et al.*, 2014, 2016; Landa *et al.*, 2018; Liu *et al.*, 2019).

While the chemical characterization of DOC remains challenging despite recent analytical advances, high throughput sequencing methods offer the possibility to describe DOC uptake indirectly via the expression of microbial transporter proteins (Green *et al.*, 2014; Lechtenfeld *et al.*, 2014; Medeiros *et al.*, 2015; Osterholz *et al.*, 2015). In particular, metaproteomic studies have contributed to insights into microbial organic matter utilization in the open ocean (Morris *et al.*, 2010; Bergauer *et al.*, 2018).

Seasonal studies of coastal Antarctic communities based on metagenomic and metaproteomic data indicated the occurrence of different functional processes in diverse taxa (Grzymski *et al.*, 2012; Williams *et al.*, 2012, 2013). Furthermore, meta-omics approaches are able to identify structure-function relationships between microbial communities within complex environments (Kieft *et al.*, 2018; Li *et al.*, 2018). Until recently, linking the taxonomic affiliation to functional data could only be achieved by sequence alignment. Draft genomes from metagenomic assemblies are an essential step towards the taxonomic classification of microbial communities and their corresponding functions (Sharon and Banfield, 2013).

Here, we used a meta-omics approach to study organic matter and Fe uptake patterns, and mapped gene expression to metagenomic assembled genomes to link phylogenetic with functional information. The aim of this study was to answer the following questions: i) Are compound-specific transporter protein differentially expressed in HNLC *vs*. iron-fertilized waters and during different seasons? ii) Does transporter expression indicate major changes in the uptake of different organic compounds by the inherent microbial communities at these two locations during different seasons?

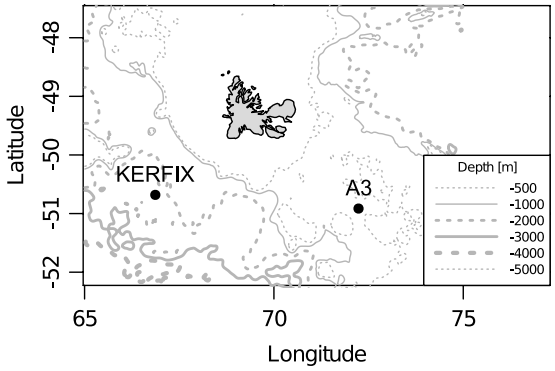


Figure 1. *Bathymetry of the Kerguelen plateau. Position along depth gradients of station KERFIX (1707m) and A3 (527m).*

Material & Methods

Sample collection

Surface seawater samples (10 – 15 m) were collected during the Southern Ocean Climate (SOCLIM, Oct. 6th – Nov. 1st 2016, austral early spring (ES)) and Marine Ecosystem Biodiversity and Dynamics of Carbon around Kerguelen (MOBYDICK, Feb. 18th – Mar. 29th 2018, austral late summer (LS)) cruise in the Indian sector of the Southern Ocean. Sampling and work procedures conducted in the study area were approved by the French Polar Institute (Institut Polaire Emile Victor) and IFREMER aboard the *RV* Marion Dufresne. Seawater was transferred from Niskin bottles mounted on a CTD frame into polycarbonate carboys of 10 – 20 L using sterile tubing with a 60 µm nylon screen.

Two stations were chosen for the present study (Fig. 1): One station was located in deeper off-plateau HNLC waters (KERFIX; 50°40’ S – 68°25’ E) and one station was located in naturally iron-fertilized waters on the plateau of Kerguelen (A3; 50°38’ S – 72°02’ E). Both stations were sampled in early spring and late summer. Except for station KERFIX in early spring, all stations were revisited twice and, A3 in late summer a third time. The number of samples taken per station is shown in Table 1. Extractions were performed as described below

with the same kits and methods for both cruises and revisits (Table 1). Sequencing depth was kept the same for all replicates.

Table 1. *Sample and replicate description*

Station	**Site**	**Season**	**Visits per cruise track**	**Prefilter**	**Metagenomes** Total DNA	**Metatranscriptomes** Ribosomal-depleted RNA	**Metaproteomes** Total protein, trypsin digested
A3	Fe-fert.	ES	2	0.8 µm	1 (pooled extracts from all revisits)	3 (biological triplicates from second visit)	2 (biological duplicates from each visit)
KERFIX	HNLC	ES	1	0.8 µm	NA	3 (biological triplicates)	2 (biological duplicates)
A3	Fe-fert.	LS	3	0.8 µm	1 (pooled extracts from all revisits)	3 (biological triplicates from first visit)	6 (biological replicates from each visit)
KERFIX	HNLC	LS	2	0.8 µm	1 (pooled extracts all revisits)	3 (biological triplicates from first visit)	4 (biological replicates from each visit)

Sampling & extraction of DNA for metagenomics

For DNA analysis, 6 L seawater from each station was pre-filtered (0.8 µm polycarbonate filter) through a 47 mm filtration system and collected on a 0.2 µm Sterivex cartridge (Millipore). Sampling was performed in triplicate resulting in 3 cartridges per station. DNA was extracted from each Sterivex filter unit using the AllPrep DNA/RNA kit (Qiagen, Hiden, Germany) with the following modifications. Filter units were thawed and closed with a sterile pipette tip at the outflow. Lysis buffer was added (40 mM EDTA, 50 mM Tris, 0.75 M sucrose) and three freeze-and-thaw cycles were performed using liquid nitrogen and a water bath at 65°C. Lysozyme solution (0.2 mg mL^{-1} final concentration) was added and filter units were placed on a rotary mixer at 37°C for 45 min. Proteinase K (0.2 mg mL^{-1} final concentration) and SDS (1% final concentration) were added and filter units were incubated at 55°C with gentle agitation every 10 min for 1 h. To each filter unit, 1550 µl RLT plus buffer was added and inverted to mix. The lysate was recovered by using a sterile 5 ml syringe and loaded in three additions onto the DNA columns by centrifuging at 10,000x *g* for 30 sec. DNA purification was performed following manufacturer's guidelines. The concentration of double-stranded DNA was quantified by PicoGreen fluorescence assay (Life Technologies). Triplicate extracts were pooled in equimolar amounts to achieve a final concentration of 1 µg in 30 µL Tris. Quality was checked on an Agilent 2100 Bioanalyzer/Agilent Nano DNA chip (Agilent, Santa Clara, CA, USA). Shotgun library preparation was performed by Fasteris SA using the illumina Nano library preparation kit with 550 bp size selection. Each metagenome was

sequenced on one full lane of HiSeq 4000 with 150 bp paired-end reads yielding between 285 – 339 million reads per metagenome.

Sampling & extraction of RNA for metatranscriptomes

For RNA extractions, volumes of 10 L of pre-filtered seawater (0.8 μm Isopore polycarbonate filters) were collected onto 0.2 μm SuporPlus membranes using a 142 mm filtration system (Geotech equipment inc.) and a peristaltic pump. The filtration procedure did not exceed 10 min and 10 ml of RNA-later was added to the filter in a Greiner tube prior to storage at -80°C. Sampling was performed in triplicate resulting in 3 filters per station.
RNA was extracted using the NucleoSpin® RNA Midi kit (Macherey-Nagel, Düren, Germany). Filters stored in RNA-later were thawed, removed from the RNA-later solution, re-frozen in liquid nitrogen and shattered using a mortar. The obtained 'powder-like' filter-pieces were added to the denaturing lysis buffer supplied by the NucleoSpin® RNA Midi kit and cells were disrupted by vortexing for 2 min. At this step two artificial internal mRNA standards in known copy numbers (MTST5 = 3.70×10^8 molecules; MTST6 = 3.72×10^8 molecules) ~1, 000 nt in length were added to each sample individually (Satinsky et al., 2013; and dx.doi.org/10.17504/protocols.io.ffwbjpe). Standards were synthesized using custom templates that were transcribed in vitro to RNA (Satinsky et al., 2012). The extraction with the NucleoSpin® RNA Midi kit includes an on-column DNA digestion step. However, to ensure the absence of any DNA in the sample, a PCR reaction was performed without the re-transcription step. Samples with DNA contamination, as indicated by amplification products were treated with a second DNA digestion step using the Turbo DNA-*free* kit (Ambion Life Technologies, Carlsbad, CA, USA). This additional DNAse treatment was followed by purification with the RNA Clean & Concentrator™-5 (Zymo Research, OZyme, France). The extracted RNA (N=12) was quantified with the Agilent 2100 Bioanalyzer/Agilent RNA 6000 Nano Kit (Agilent, Santa Clara, CA, USA). Prior to sequencing, ribosomal RNA was treated enzymatically with the RiboZero rRNA stranded RNA protocol to ensure sequencing of primarily messenger RNA followed by cDNA library construction using Illumina TruSeq Stranded mRNA Library Prep kit (Fasteris SA). All metatranscriptomes (n=12) were sequenced on one lane of HiSeq 4000 with 150 bp paired-end reads yielding between 26 – 36 million reads per metatranscriptome.

Sampling & extraction of proteins for metaproteomes

For protein extractions 20 L seawater were pre-filtered seawater (0.8 µm Isopore polycarbonate filters) and collected on 0.2 µm SuporPlus Membranes using a 142 mm filtration system (Geotech equipment inc.) and a peristaltic pump. Filtrations were performed in duplicate for early spring samples and for late summer samples and frozen immediately at -80°C. Whole protein extractions from filters were performed using a modified protocol from Bayer et al. (2019) and subjected to denaturing polyacrylamide gel electrophoresis (SDS-PAGE) followed by overnight trypsin in-gel digestion. For higher yield of extraction, filters in Greiner tubes were cooled in liquid nitrogen and smashed with a sterile metal spoon until a powder-like sample of small filter pieces was achieved. Proteins were extracted from filter pieces by adding 10 mL lysis buffer (100 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM DTT, 1 % SDS, 10 mM EDTA) to the tubes followed by three freeze-and-thaw cycles and 5 cycles of sonication with a sonication probe (Sonopuls HD 2070, Bandelin) for 30 sec intervals at 20% intensity. Tubes were centrifuged at 20,000x *g* at 4°C for 10 min and the supernatant containing the cell lysate was collected in a fresh tube. Proteins were co-precipitated with 0.015% deoxycholate and 6% trichloroacetic acid (TCA) on ice for 1h and washed once with ice-cold acetone according to the protocol of Bensadoun and Weinstein (Bensadoun and Weinstein, 1976). Dried protein pellets were resuspended in 50 µL of 8 M urea containing 4% SDS and protein concentrations were measured with the bicinchoninic acid (BCA) assay using BSA as a standard. Ten-20 µg of protein were subjected to denaturing polyacrylamide gel-electrophoresis (SDS-PAGE), stained with Commassie staining solution and de-stained in 40% (v/v) methanol containing 2% (v/v) acetic acid as described in Valledor and Weckwerth (Valledor and Weckwerth, 2014). Gel bands were excised without fractionating the bands, cut into 1 x 1 mm pieces on a sterile glass plate and transferred into protein-low binding tubes (Eppendorf). Gel pieces were de-stained with 200 mM ammonium bicarbonate containing 50% acetonitrile at 37°C for 30 min, and cysteines were reduced and alkylated with 10 mM DTT and 55 mM iodoacetoamide (IAA), respectively, as described by Shevchenko et al. (Shevchenko *et al.*, 2007) prior to overnight trypsin digestion according to Valledor and Weckwerth (Valledor and Weckwerth, 2014). Peptides were extracted and desalted using 96-well plates (Spec 96-Well C18, Agilent) (Valledor and Weckwerth, 2014) and then resuspended in 2 % acetonitrile containing 0.1 % formic acid to a concentration of 0.2 µg μL^{-1} prior to injection into a one-dimensional nanoflow LC–MS/MS.

LC-MS/MS analysis and peptide identification

Five µL of desalinated peptides were eluted using an Easy-spray PepMap RSLC column (ThermoFisher Scientific, C18, 500 mm x 75 µm, pore size 2.0 µm), during a 270-min gradient from 5 to 40% (v/v) acetonitrile and 0.1% (v/v) formic acid with a controlled flow rate of 300 nl min^{-1}. MS analysis was performed on a Orbitrap Elite and Q Exactive mass spectrometer (Thermo Fisher Scientific). Specific tune settings for the MS were as follows: Mass resolution for precursor ion analysis in FTMS (Fourier transform mass spectrometry): 60,000; full-scan mode, mass window for precursor ion: 1 m/z; spray voltage: 1.9 kV; temperature of the heated transfer capillary: 275°C, covering the range 350–1,800 m/z, and cyclomethicone used as lock mass (m/z 371.101230). Each full MS scan was followed by 20 dependent MS/MS scans in the Ion Trap using rapid mode with centroid data in which the 20 most abundant peptide molecular ions were dynamically selected. The dynamic exclusion window was set to 30 sec and an exclusion list of 500 entries. Dependent fragmentations were performed in CID (collision-induced dissociation) mode, with a normalized collision energy of 35, iso width of 2.0, activation Q of 0.250, and activation time of 30 msec. Ions with a +1 or unidentified charge state in the full MS were excluded from MS/MS analysis.

In Proteome Discoverer 2.2 (Thermo Fisher Scientific), the mass tolerance was set to 5 ppm for precursor masses and 0.6 Da for the fragment masses. Cysteine carbamidomethylation was set as static-, methionine oxidation and protein *N*-terminal acetylation as dynamic modifications. A maximum false discovery of 1% was allowed for both peptide and protein levels. Trypsin was specified as the proteolytic enzyme and 2 missed cleavages were allowed with a maximum of 3 equal post-translational modifications (PTMs) per peptide. The obtained peptide spectrum matches (PSMs) were filtered with the Percolator tool based on a scoring of maximum delta Cn 0.05, a strict false discovery rate (FDR) of 0.01 with validation based on q-values.

Acquired MS/MS spectra were analyzed using the SEQUEST-HT algorithm implemented in Proteome Discoverer 2.2 software (Thermo Fisher Scientific), and spectra were searched against the entire set of metagenomic database (described below in bioinformatic analysis, Figure 2, B&D). Protein matches were accepted if they were identified by at least one unique peptide and with high confidence. Proteins were quantified using the normalized spectral abundance factor (NSAF) approach (Zhang *et al.*, 2015) as follows:

$$NSAF_k = \left(\frac{PSM}{L}\right)_k / \sum_{i=l}^{N} \left(\frac{PSM}{L}\right)_k$$

where the total number of spectral counts for the matching peptides from protein k (PSM) was divided by the protein length (L) and then divided by the sum of PSM/L for all N proteins.

Metagenome analysis

An initial round of read processing was provided by the company, using trimmomatic (Bolger *et al.*, 2014), an integrated quality-control tool for high-throughput Illumina NGS data. The standard Illumina adapters and low-quality bases were removed with the following parameters "2:30:10 SLIDINGWINDOW:4:5", and reads for which no insert was found or ambiguities were also dropped, resulting in million 2×150 bp paired-end reads.

The raw Illumina reads were checked with FastQC (Andrews 2010; http://www.bioinformatics.babraham.ac.uk/projects/fastqc). Decontamination was assured by running bbduk with the following parameters k=31 hdist=1 ftm=5 and checking again with FastQC. BBNorm was used with the following parameters target=100 min=5 prefilter=t bits=16 ecc=t to normalizes coverage by down-sampling reads to a depth of 100 for each library and facilitate computational effort give the large size of the files (96 – 225 GB) (Bushnell 2014).

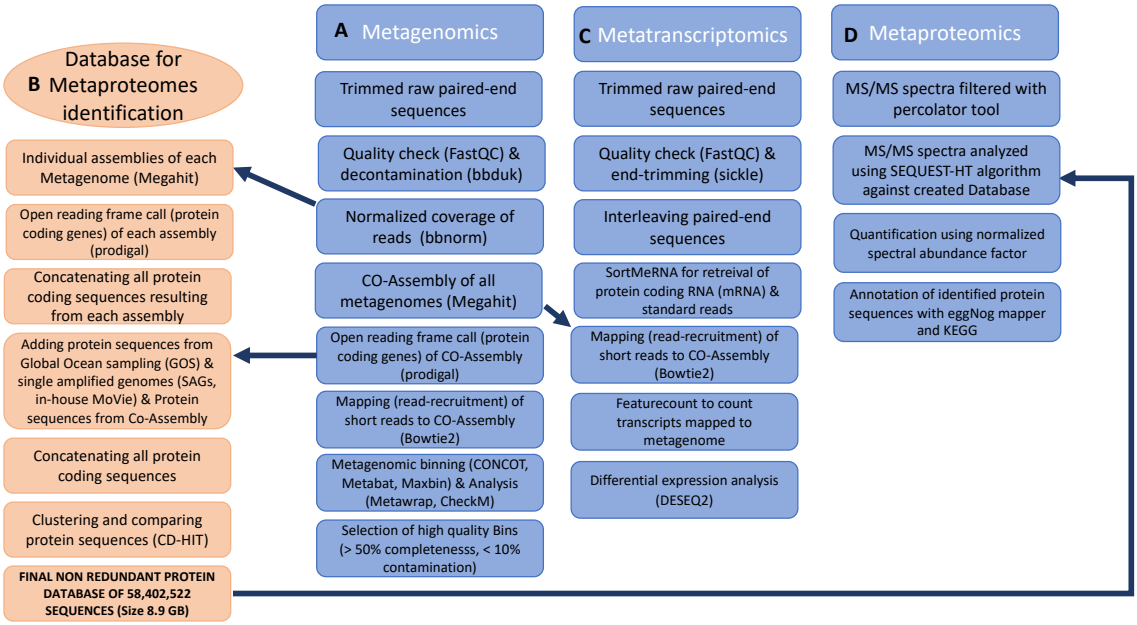


Figure 2. *Bioinformatic flowchart for the analysis of all three 'omics' levels. A. Metagenomic assembly (individual and Co-Assembly) and binning B. Construction of Ocean protein database out of metagenomic assemblies (MoVie – Microbial Oceanography Group, University of Vienna) C. Metatranscriptomic mapping to metagenomes after extraction of mRNA with SortMeRNA D. Metaproteomic analysis with metagenomic database and annotation in eggNOG and KEGG.*

Decontaminated, trimmed and bbnormed metagenomic sequences from station A3 in early spring as well as A3, KERFIX, and an additional available sample M4 form late summer were co-assembled using megahit v1.2.7 (Li *et al.*, 2015) with the default parameters and the --presets 'meta-large' option resulting in 949,228 contigs of at least 1,000 bp (Supplementary Table 2&3). Prodigal, a protein-coding gene prediction software tool for bacterial and archaeal genomes, was used to produce a gene coordinate file (gtf) and a translated protein as well as nucleotide file with the option -p meta (Hyatt *et al.*, 2010).

Binning for metagenomic assembled genomes (MAGs)

We applied a combination of three binning tools, CONCOCT (Alneberg *et al.*, 2014), MaxBin v2.0 (Wu *et al.*, 2016) and MetaBAT v2.0 (Kang *et al.*, 2019), to recover genomes from fragmented metagenomic assemblies with a mimimum contig size of 2,500 bp. CONCOCT bins the contigs by employing sequence composition and contig coverage across multiple samples (Alneberg *et al.*, 2014). Besides the two features used by CONCOCT, MaxBin also takes advantage of phylogenetic marker genes to perform binning through an Expectation-Maximization (EM) algorithm (Wu *et al.*, 2016). MaxBin was performed multiple times with different values for "-markerset" (40 or 107) and "-prob_threshold" (0.5 or 0.8). The resulted MAGs by distinct parameter combinations were compared based on the completeness and contamination estimations in CheckM (Parks *et al.*, 2015), by using an R script provided in the MetaBAT workflow (https://portal.nersc.gov/dna/RD/Metagenome_RD/MetaBAT/Files/benchmark.R). The parameter set (-markerset 40 –prob_threshold 0.5), which yielded slightly more bins fulfilling given recall (>=0.5) and precision (>=0.9) cutoffs, was selected. MetaBAT2 is a fully automated metagenome binner equipped with an adaptive binning algorithm based on iterative graph partitioning for contig clustering, which saves users from the brain-racking parameter tuning procedure (Kang *et al.*, 2019). All three tools are widely used on current metagenomics studies, but their extents of success vary with characteristics of the underlying datasets (Sczyrba *et al.*, 2017). Therefore, we integrated and refined MAGs generated by the three tools by using MetaWRAP, which proposes a scoring function based on the completion and contamination metrics estimated with CheckM (Uritskiy *et al.*, 2018). Finally, 133 high-quality MAGs with >= 50% completeness and < 10% redundancy were obtained.

Protein-coding genes of each MAG were retrieved from the aforementioned Prodigal annotations based on sequence headers. For phylogenomic inference, single-copy orthologous gene families shared by at least 20 (out of 133) MAGs were identified by OrthoFinder v2.2.3

(Emms and Kelly, 2015), and aligned with MAFFT v7.313 (Katoh and Standley, 2013). Then, the protein sequence alignments of each gene family were concatenated into a super alignment. Maximum likelihood (ML) phylogenetic inference of the 133 MAGs was performed using IQ-Tree v1.6.8 (Nguyen *et al.*, 2015) with 1,000 bootstrap replicates, and the amino acid substitution model for each gene family partition was automatically selected by the ModelFinder feature (-m MFP) (Kalyaanamoorthy *et al.*, 2017). Furthermore, taxonomic classification of the 133 MAGs was performed according to the Genome Taxonomy Database (Parks *et al.*, 2018) using GTDB-Tk v0.3.3 (https://github.com/Ecogenomics/GTDBTk) with the classify_wf workflow.

Metatranscriptomic read-recruitment (mapping)

An initial round of read processing was provided by the sequencing company using trimmomatic (Bolger, Lohse et al. 2014), an integrated quality-control tool for high-throughput Illumina NGS data. The standard Illumina adapters and low-quality bases were removed with the following parameters "2:30:10 SLIDINGWINDOW:4:5" and reads for which no insert was found or ambiguities were dropped resulting in million 2×150 bp paired-end reads. The raw Illumina reads were checked with FastQC. A further quality trimming was performed with Sickle resulting in between 99.47 - 100% of remaining reads which is a tool that uses sliding windows along with quality and length thresholds to determine when quality is sufficiently low to trim the 3'-end of reads and also determines when the quality is sufficiently high to trim the 5'-end of reads (https://github.com/najoshi/sickle) (Suppl. Table 4). To focus on protein-coding RNAs, we computationally eliminated ribosomal RNA and internal standard-derived reads using SortMeRNA v 2.1b (Kopylova, Noe et al. 2012). A BLASTN homology search against a custom database, which consists of representative prokaryotic rRNAs and tRNAs from NCBI RefSeq along with the template sequences of internal standards, was also implemented using a bit score cut-off of 50 as suggested in previous studies (Satinsky, Gifford et al. 2013). The files with protein coding genes from SortMeRNA were unmerged again to be mapped to the reference metagenomic co-assembly using Bowtie2 for read recruitment with default parameters. Overall 59.8 – 62.6 % of reads could be mapped on the metagenomic contigs (Supplementary Table 5). For read counting featureCounts (http://bioinf.wehi.edu.au/featureCounts/) was used with the following parameters for stranded specific reads -F 'GTF' -t 'gene' -Q 1 -s2 -p -C. Count tables were further analysed with DESeq2 (Love et al. 2014) in R. Differential expression analysis was thus calculated with triplicates of each station as the spring samples for reference.

Protein database construction and annotation of transporter protein sequences

Each metagenome was individually assembled by megahit and protein sequences were predicted by prodigal using the same parameters as mentioned above. The script anvi-script-reformat-fasta was run on each resulting protein fasta file changing header names to include the sample ID for further analysis. Sequences were compared with cd-hit-2d and the parameters -c 0.9 -n 5 -d 0 -S2. Additionally, the GOS peptide database was added as well as the SAG database available in-house. Once one file was created and cd-hit was run with the following parameters -c 1 -n 5 -d 0 to create a non-redundant protein sequence database. This resulted in a database of 58,403.522 sequences (Fig. 1B).

In order to identify the transporter sequences resulting from the individual metagenomic assemblies, predicted protein sequences were aligned eggNOG 5.0 (Huerta-Cepas, Szklarczyk et al. 2019) and GhostKOALA (Kanehisa, Sato et al. 2016) with default parameters. All KO number defined as "Transporter" in KEGG were retrieved and this list was used for merging with annotated sequences for metatranscriptomes and metaproteomes. Transporter families were verified by manually checking the KO number with assigned classification from the Transporter Classification Database. For definite assignment of a protein sequence with a MAG, the sequences corresponding to transporters were aligned against the high quality curated metagenomic bins and those with 99% identity and 90% coverage were kept for MAG specific further analysis.

Data visualization

All figures were produced using ggplot2 package in R version 3.6.0 (2019-04-26) and enhanced in the open-source software inkscape. Figure 6 was produced using Anvi'o visualization tool version (5.2.0) in the anvio-interactive manual mode.

Results

Environmental context

DOC concentrations ranged from 50 – 52 µM for both seasons and at both sites (Table 2). Chlorophyll *a* concentrations at station A3 ranged from 1.40 – 1.64 µg L^{-1} during both visits in early spring and were lower in late summer with 0.27 – 0.57 µg L^{-1}. At station KERFIX, chlorophyll *a* concentrations were higher in early spring (0.32 µg L^{-1}) as compared to those in late summer (0.14 – 0.19 µg L^{-1}). Total nitrogen (nitrite + nitrate) and phosphate concentrations in early spring at both stations ranged between 28.11 – 28.23 µM and 1.84 – 1.97 µM,

respectively, and were slightly higher compared to those in late summer (21.05 – 22.27 µM ∑N, 1.45 – 1.62 µM PO_4^{3-}) (Table 1).

Prokaryotic cell abundance was about two-fold higher at both stations in late summer ranging from 6.68 – 11.8 x10^8 cells L^{-1} at station A3 and 4.46 – 6.96 x10^8 cells L^{-1} at station KERFIX, as compared to 3.66 – 4.88 x10^8 cells L^{-1} at A3 and 2.89 x10^8 cells L^{-1} at KERFIX in early spring. Dissolved Fe (DFe) was not measured at either cruise, however, reference data from previous cruises at both sites collected during the same seasons are available, indicating that DFe was higher at station A3 during both seasons (Blain *et al.*, 2008; Quéroué *et al.*, 2015).

Table 2. *Environmental parameters for the two stations and revisits. Parameters at sampling depth of 10-15 m for each site. ES – early spring, LS – late summer. T – temperature mean value of mixed layer depth, DOC – dissolved organic carbon, PA – prokaryotic abundance, PP – prokaryotic production, *for early spring HNLC reference site R-2 (November 2011) and A3 from Christaki et al. (2014); PR – prokaryotic respiration, DFe – Dissolved Iron** for early spring from Quéroué et al. (2016) for KEOPS-2 project with HNLC reference site R-2 (November 2011), for late summer from Blain et al. (2007) & Lefèvre et al. (2008) for KEOPS-1 with HNLC reference site C11 (January-February 2005). Mean values and standard deviation are shown for parameters that were measured in triplicates.*

	SEASON	**T** (°C)	**DOC** (µM)	**Chl *a*** (µg L^{-1})	**PO_4^{3-}** (µM)	**∑N** (µM)	**PA** (x$10^8 L^{-1}$)	**PP*** (nmol C L^{-1} d^{-1})	**PR** (µmol O_2 L^{-1} d^{-1})	**DFe**** (nmol L^{-1})
A3	ES	2.19	52	1.40	1.85	28.21	3.66	9.9 ± 0.3	NA	NA (0.16 ± 0.03, KEOPS-2)
A3_2	ES	2.06	51	1.64	1.84	28.11	4.88			
KERFIX	ES	2.38	51	0.32	1.97	28.23	2.89	5.2 ± 0.2	NA	NA (Station R-2, 0.13 ± 0.05)
A3	LS	5.01	52.5	0.27	1.45	23.27	11.80	22.16	0.68 ± 0.36	NA (0.13, KEOPS-1)
A3_2	LS	5.24	52.5	0.32	1.47	21.55	8.37	46.23		
A3_3	LS	5.26	52.5	0.57	1.51	22.05	6.65	56.91		
KERFIX	LS	5.62	50.4	0.19	1.62	23.24	6.96	15.59	0.29 ± 0.13	NA (Station C11, 0.09)
KERFIX_2	LS	5.63	50.4	0.14	1.61	23.25	4.46	13.18		

Metatranscriptomic and metaproteomic assessment of transporter proteins

In order to identify changes in organic matter utilization comparing the two contrasting stations and seasons, membrane transporter proteins from metatranscriptomes and metaproteomes were analyzed using all available transporter proteins identified by KEGG (Supplementary Table 1 & 6). Significant differences in relative transporter protein abundances between the two seasons were identified as shown by dissimilarity analysis for metatranscriptomes and metaproteomes (Fig 3. A&B).

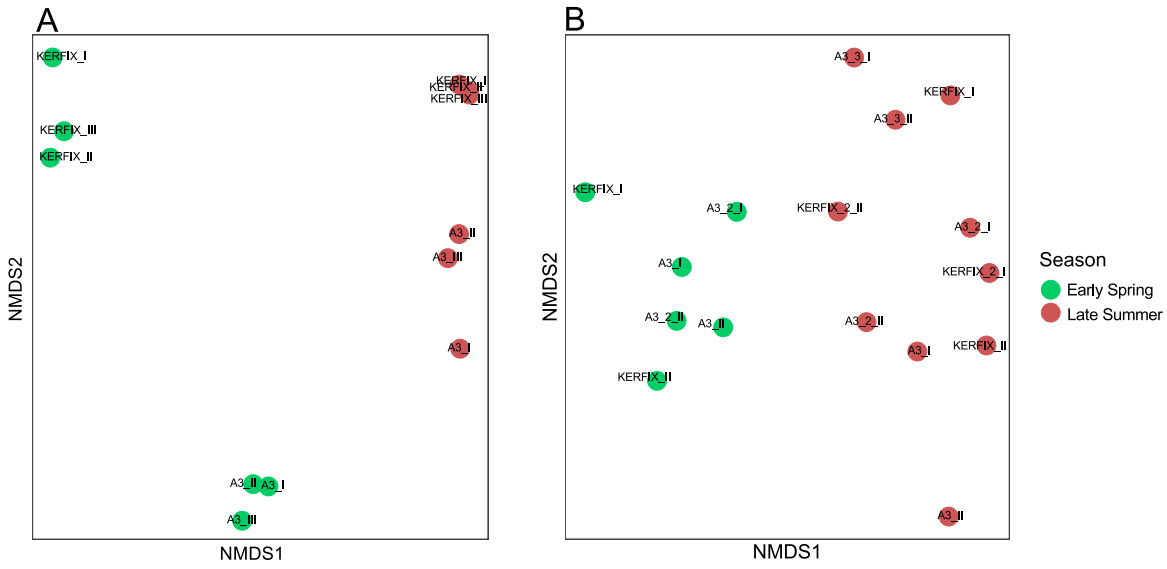


Figure 3. *Nonmetric multidimensional scaling (NMDS) of transporter proteins from **A**. Metatranscriptomes and **B**. Metaproteomes based on Hellinger transformed Bray–Curtis dissimilarity. Biological replicates are shown for each station and revisit. Samples from the two seasons were significantly different (Metatranscriptomes: 1D stress = 0.00001, ANOSIM, R = 0.88, P =0.001; Metaproteomes: 2D stress = 0.1; ANOSIM, R = 0.84, P = 0.001).*

During early spring, transporter proteins made up 64.9 ± 16.8% of all detected proteins compared to 48.47 ± 4.5% in late summer metaproteomes (Fig. 4A) while the relative proportions in metatranscriptomes made up 4.6 ± 1.1% in early spring and 11.0 ± 1.2% in late summer of total normalized counts (Fig. 4B). In both metaproteomes and metatranscriptomes, membrane transporters belonging to the ATP-binding cassette transporter family (ABC) had highest contributions of 3.9 ± 1.1% to 9.9 ± 1.0% respectively, followed by outer-membrane receptors (OMR) ranging from 0.3 ± 0.0% in early spring to 0.5 ± 0.1% in late summer.

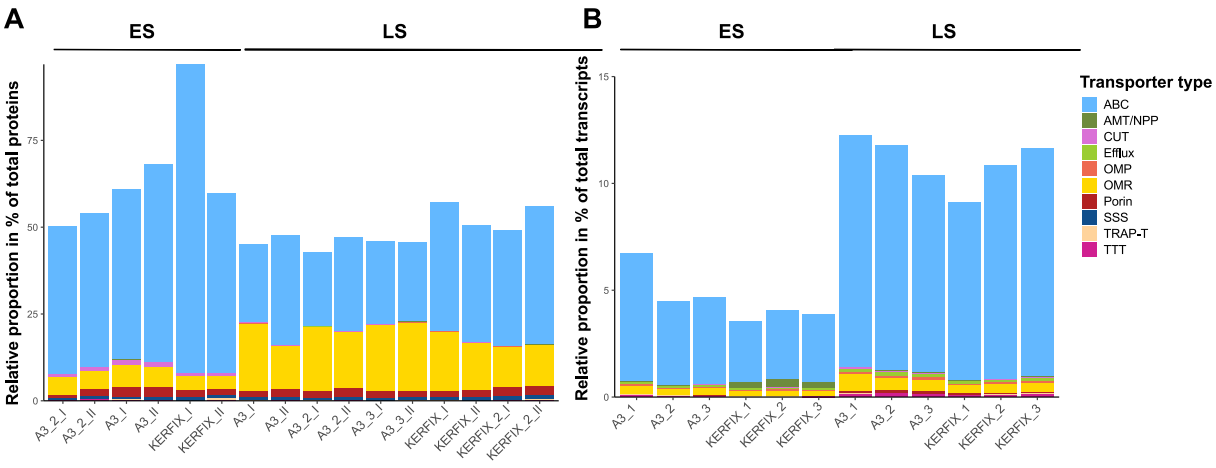


Figure 4. *Relative proportions of transporter families. Bar-plot showing the relative proportions of different transporter families according to the transporter database (TBDB) for* ***A.*** *Metaproteomes from NSAF and* ***B.*** *Metatranscriptomes based on the total normalized transcripts. All duplicates and triplicates are shown, note difference in scale. ES -early spring, LS –late summer, Transporter types are indicated by different colors: ABC – ATP-binding cassette transporter complex, AMT/NPP – Ammonium Channel Transporters/Nitrate, Nitrate Porters, CUT – Carbohydrate Uptake Transporters, OMP – Outer Membrane Proteins, OMR – Outer Membrane Receptors, SSS – Solute Sodium Symporters, TRAP-T - Tripartite ATP-independent periplasmic transporters, TTT – Tripartite Tricarboxylate Transporters*

We found 202 transporter proteins of which 89 were shared in the metatranscriptomic and metaproteomic datasets which are visualized in Figure 5 (Supplementary Figure 1). In both seasons, the highest normalized transcript counts of transporter proteins belonged to those encoding for amino acid transporters (ABC, K09969), branched amino acid transporters (ABC, K01999), spermidine/putrescine transporters (ABC, K02055), iron outer-membrane complex receptors (OMR, K02014), taurine transporters (ABC, K15551), glycine-betaine transporters (ABC, K02002), peptide/nickel transporters (ABC, K02035) and iron (III) transporters (ABC, K02012). Furthermore, transcripts encoding glycerol transporters (ABC, K17321) were among the ten highest normalized counts in late summer whereas higher counts of transcripts encoding ammonium transporters (AMT, K03320) were detected in early spring, particularly at station KERFIX. Additionally, three sugar transporters (K10559, K10111, K10117) and transporters for thiamine (vitamin B1), phosphonate (K02041) were higher expressed in early spring and were detected in metaproteomes from both seasons.

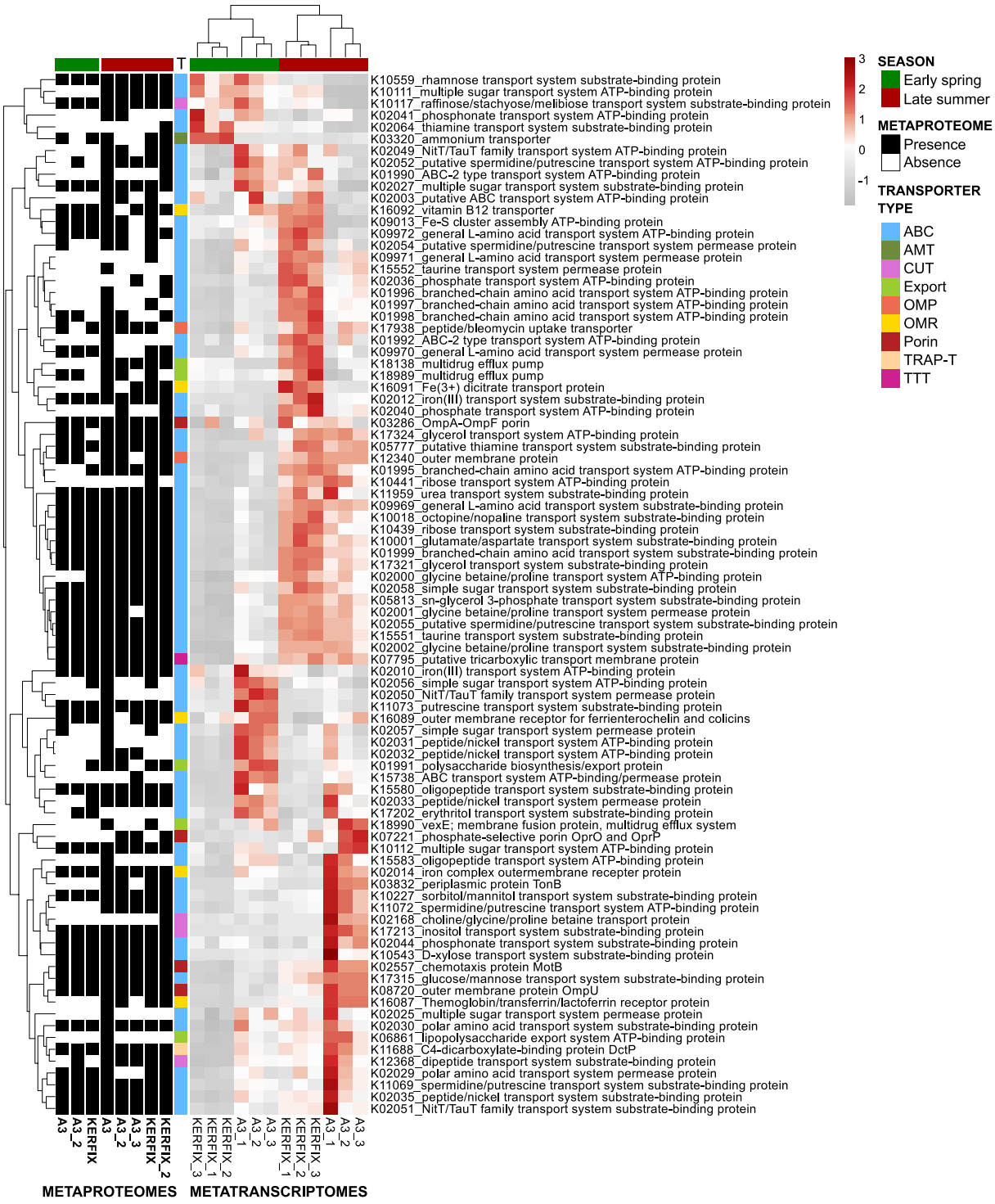


Figure 5. *Heatmap of shared KEGG transporter proteins. All metatranscriptomes are shown in triplicates. Z-scaling of normalized transcript counts by rows and Euclidian clustering by row and column. The presence of proteins in metaproteomes is defined by their presence in one duplicate. Transporter types are indicated by different colors: ABC – ATP-binding cassette transporter complex, AMT/NPP – Ammonium Channel Transporter/Nitrate, Nitrate Porter, CUT – Carbohydrate Uptake Transporter, OMP – Outer Membrane Protein, OMR – Outer Membrane Receptor, SSS – Solute Sodium Symporter, TRAP-T - Tripartite ATP-independent periplasmic transporters, TTT – Tripartite Tricarboxylate Transporter*

Carbon uptake transporters (CUT) were higher expressed in late summer, specifically transporters for peptides and sugars (K02168, K017213, K012368). TRAP-T (Tripartite ATP-independent periplasmic transporters associated extra-cytoplasmic Immunity transporter) and TTT (Tripartite Tricarboxylate Transporter) accounted for the smallest fraction of prokaryotic transporters and had a higher number of transcripts in late summer than in early spring, while detected in metaproteomes of both seasons (K07795, K11688). Porins were expressed at both stations during both seasons, however, membrane export proteins were detected at higher relative abundance during early spring than in late summer (Fig. 5).

The substrate active community

Metagenome assembled metagenomes (MAGs) were constructed and transporter proteins from metatranscriptomes and metaproteomes were mapped to high quality metagenomic bins (completeness >=50%, redundancy <=10%) to identify the substrate active members of the prokaryotic community. MAGs were constructed from the Co-assembly of all metagenomic libraries (Supplementary Table 2 & 3). A total of 133 high quality MAGs spanning 11 phyla (according to https://gtdb.ecogenomic.org/, Genome Taxonomy Database) were used for mapping of metaproteomic and metatranscriptomic sequences (Fig. 6, Supplementary Table 6). Out of 133 MAGs, 39 belonged to the phylum of *Bacteroidota* with the dominant order of *Flavobacteriales* (32), and 69 MAGs belonged to *Proteobacteria* of which 34 belonged to *Alphaproteobacteria* and 35 to *Gammaproteobacteria.* While transcripts of transporter proteins could be mapped to the majority of MAGs (122 out of 133), peptide sequences of transporter proteins could be mapped to 67 MAGs (Fig. 6).

ABC, OMR, porins and CUT could be assigned to specific MAG transcriptomes and proteomes while for AMT/NNPs, TTTs and TRAP-Ts only transcriptomes were assigned.

Several phyla-specific patterns could be observed for transport protein expression. Most transporter proteins present in metatranscriptomes and metaproteomes were mapped to gammaproteobacterial MAGs, including MAG31 of the order *Enterobacterales* and MAGs 130, 56, 44, 13, 115, 117, 62 and 103 of the order *Pseudomonadales*. Furthermore, the alphaproteobacterial MAG137 of the order *Sphingomonadales* expressed many transporter proteins, and MAGs of the order *Rhodobacterales* nearly all expressed TRAP-Ts in the metatranscriptomes. For all MAGs of the order *Flavobacterales* sequences for ABC, OMR and AMT/NNPs transporters were detected in the metatranscriptomes, whereas the euryarchaeal MAG137 expressed an ammonium transporter. Several MAGs did not appear to express

transporter proteins, including MAG16 (*Flavobacterales*), all MAGs of the order *Caulobacterales* (MAG120, 86, 68, 37) and MAG25 (*Burkholderiales*).

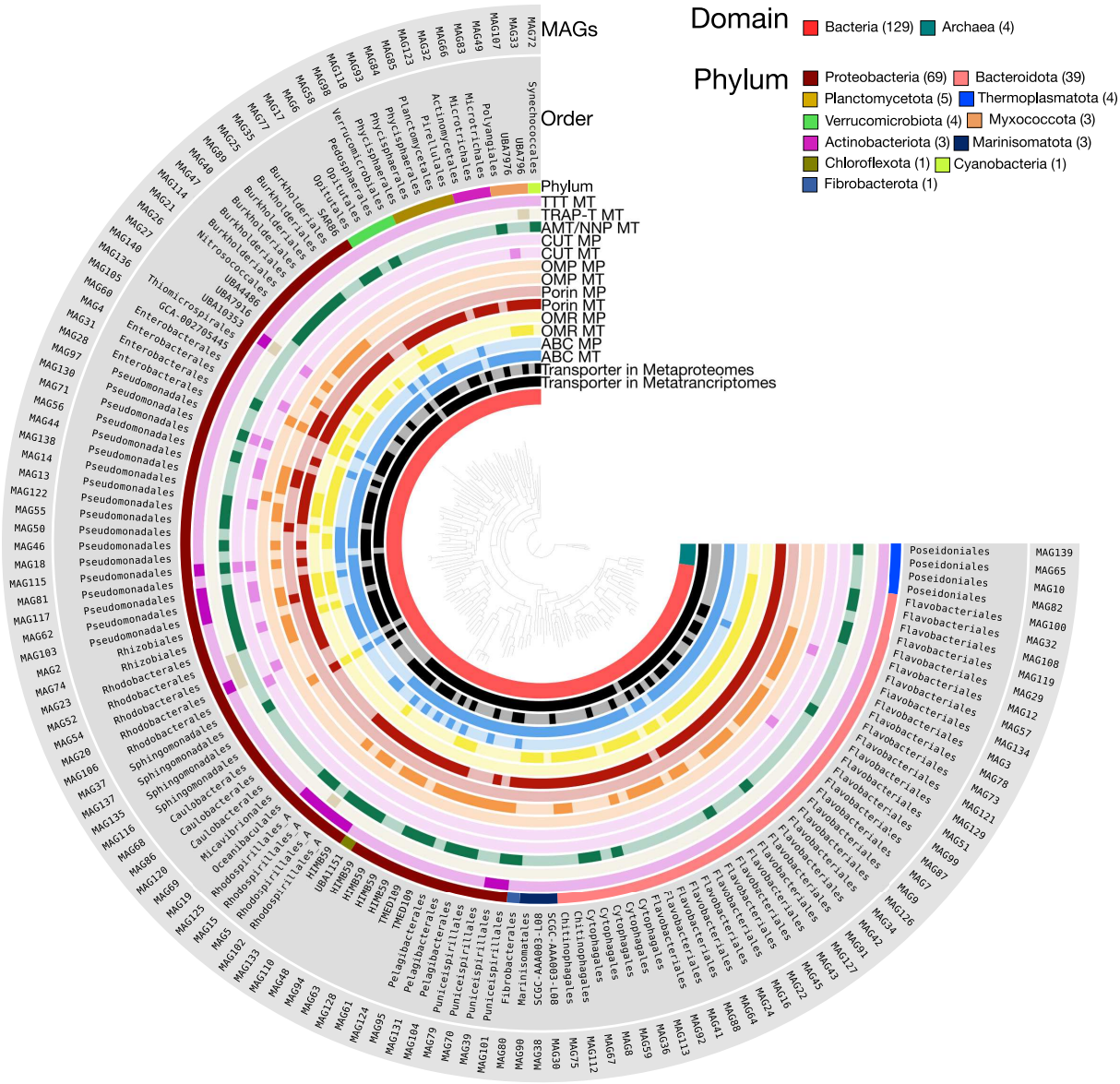


Figure 6. *Phylogenetic tree of 133 metagenome assembled genomes from the Co-Assembly with additional information on the expression of transporter proteins in metatranscriptomes and metaproteomes. Tree calculated from 163 single-copy genes shared by at least 20 out of the 133 MAGs. Taxonomy from GTDB. Additional layer represents presence of specific transporter type in MT – Metatranscriptomes and MP - Metaproteomes*

Organic matter utilization as revealed by transporter proteins

Compound specificity for a transporter protein was derived from sequence similarity to transporter families. Most of the detected compound-specific transporters were identified at higher relative abundance in early spring than in late summer, with the exception of Fe and other trace metal transporters detected in metaproteomes (Fig. 7A). Contrary to metaproteomes, Fe transporters showed no difference in the relative contribution throughout the season for all

stations ranging from 7.9 – 9.1% (Fig. 7B). In relative contribution to all assigned transporter transcripts, amino acids (Fig. 7B; 23.6 – 28.3%) as well as branched amino acids (12.2 – 14.7%) were the most abundant compound-specific transporters at all stations followed by polyamines spermidine and putrescine transporters (8.7 – 10.2%). Amino acid transporters represented 21.8 – 56.6% of the identified transporter proteins in early spring compared to 12.1 – 24.7% in late summer metaproteomes (Fig. 7A).

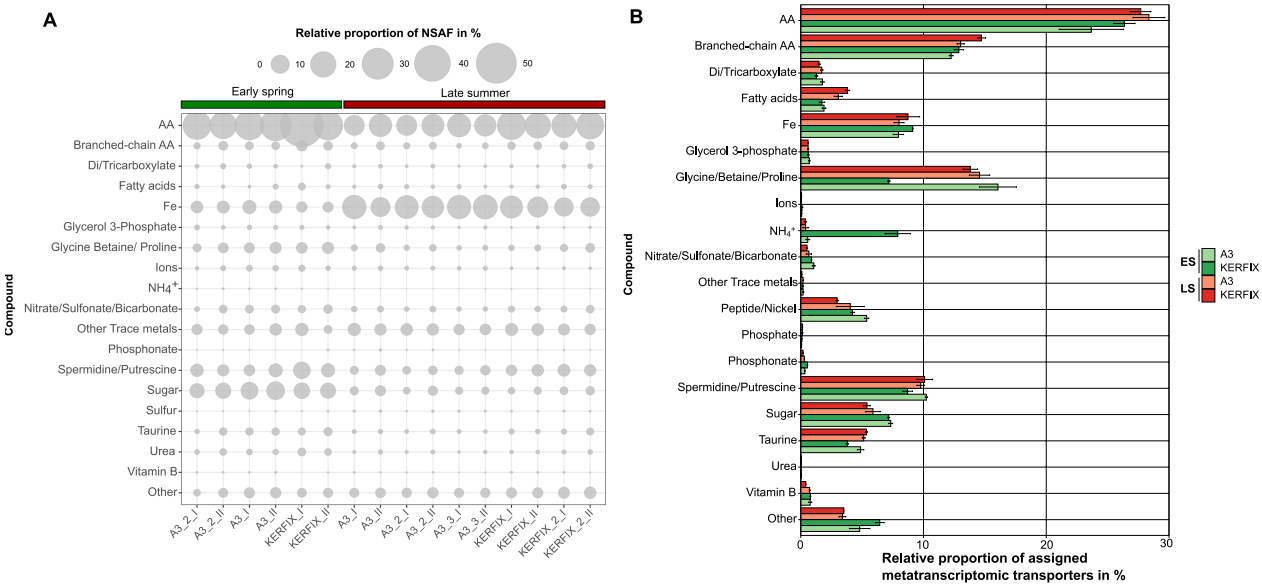


Figure 7. *Compound-specific transporter proteins. A. Relative proportion by NSAF by compound and metaproteomes. All replicates and revisits are shown. B. Relative proportion of compound-specific transporters from assigned transporter transcripts shown as mean values with SD for metatranscriptomes. Green bar – early spring samples, Red bar – late summer samples, AA – amino acids.*

Hence, amino acids potentially comprised the largest pool of transported substrates. Diverse sugar transporters were detected in early spring samples comprising 6.1 – 7.4% of the identified transporter proteins, while comprising 1.2 – 3.1% of transporter proteins in late summer. Glycine/betaine/proline transporter transcripts (Fig. 7B) were in the same range at all stations ranging from 13.8 – 16.1% of the total transporter proteins except for KERFIX in early spring (7.1%). In contrast ammonium transporter transcripts were highest in abundance at KERFIX in early spring (7.9%). Glycine betaine transport proteins were more abundant in early spring samples contributing 1.8 – 3.8% of all transporter proteins compared to 0.5 – 2.2% in late summer (Fig. 7A). Furthermore, polyamine transporters comprised 3.9 – 8.5% of the detected transporter proteins in early spring and 1.8 – 4.3% in late summer, indicating the presence of a labile source of carbon and nitrogen.

The number of taurine transporter transcripts was lowest at KERFIX in early summer at 3.8% and reached up to 5.4% at this station in late summer. Urea transporter transcripts were

among the transporters with the lowest relative abundances ranging from 0.02 – 0.03% (Fig. 7B) while urea transporters ranged from 0.6 – 1.8 % in early spring to 0.2 – 0.6 % in late summer metaproteomes (Fig. 7A). Sugar transporters were roughly similar in abundance in the metatranscriptomes as in the metaproteomes with higher contribution of transporters in early spring samples with 7.5 to 5.5%, respectively.

Differential expression between early spring and late summer

In order to identify seasonal differences between the active members of the prokaryotic community, we retrieved all transporter proteins identified in the metaproteomes and used the respective transcripts for differential expression analyses (see Material & Methods).

Metatranscriptomic coding sequences were mapped against the co-assembly and the resulting count tables obtained for spring and summer samples were compared (Supplementary Table 4 & 5). Of all identified transporter proteins, 12% exhibited a higher relative abundance in late summer, whereas 21% were detected at higher relative abundance in early spring (Supplementary Figures 4 – 6; Supplementary Table 7). The observed differential expression patterns of transport proteins were represented by distinct phylogenetic groups (Fig. 7).

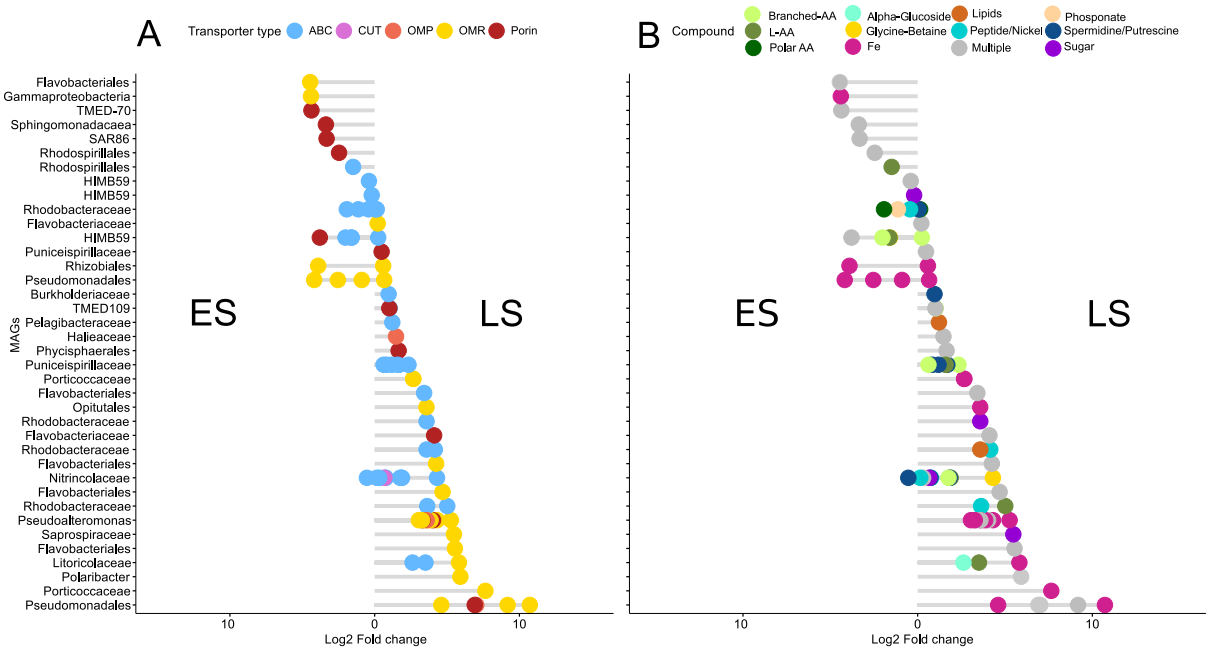


Figure 8. *Differentially expressed transporter protein during early spring and late summer. Left side values represent expression in early spring and on the right side, values represent expression in late summer metatranscriptomes. MAG Ids to lowest phylogenetic level possible.* ***A.*** *By transporter type,* ***B.*** *By compound.*

All early spring ABC transporters could be matched to alphaproteobacterial MAGs (Fig. 7A; *Rhodospirillales, HIMB59, Rhodobacteracea*). Of these, *Rhodobacteracea* differentially expressed transporters for multiple compounds in both seasons indicating metabolic plasticity

within members of this family (Fig. 7B; Polar AA, L-AA, Phosphonate, Spermidine/Putrescine, Peptide/Nickel, lipids). OMRs were expressed at higher relative abundances in late summer than in early spring by *Flavobacteriales* and *Gammaproteobacteria* (*Pseudoalteromonas, Pseudomondales, Litoricolacea*).

Discussion

Although DOC concentrations were similar in early spring and late summer (Table 2), prokaryotic heterotrophic production rates were two- to three-fold higher in late summer than early spring as indicated by ^{3}H-leucine incorporation rates. The higher prokaryotic production observed at station A3 in late summer indicates a post-bloom scenario as previously observed in this season and station (Christaki *et al.*, 2008; Obernosterer *et al.*, 2008). Chlorophyll *a* concentrations were higher in early spring than in summer at both sites in accordance with chlorophyll a data from a remote autonomous sampler at station A3 (Supplementary Figure 5). The low surface DOC values are a characteristic feature of Southern Ocean waters exhibiting some of the lowest world's ocean DOC values in surface waters (Hansell *et al.*, 2009). The availability of Fe in surface waters above the Kerguelen plateau appears to be enhanced by the greater winter stock supplied by a higher continuous vertical mixing which can explain the slightly higher DFe concentrations during spring (Blain *et al.*, 2007). Vertical profiles have shown that deep winter mixing sustains surface water Fe supplies in the Southern Ocean(Tagliabue *et al.*, 2014). The strong seasonal difference in Fe and carbon supply, regardless of specific site, is of greater importance for transporter proteins as indicated by our data.

Membrane proteins are more difficult to solubilize from biomass than periplasmic binding proteins without the application of specific methods of enrichment (Burg *et al.*, 2010; Morris *et al.*, 2010; Williams *et al.*, 2010). However, we detected high relative abundances of different membrane transporter types in all metaproteomes. In contrast, the detected membrane transporters showed a substantially lower relative contribution in the metatranscriptomes. The major advantage of applying metatranscriptomics and -proteomics compared to community diversity analysis, is to infer the active part of the community, and to decipher which genes are transcribed and translated into proteins. RNA expression studies provide a more comprehensive insight into the cell metabolism while proteomic studies suffer from inherent technical shortcomings associated with, for example, protein insolubility or fractionation losses. Accordingly, conclusions from proteomic studies are usually based on smaller datasets. They

also give insight only into protein accumulation and/or protein synthesis. On the other hand, mRNA levels do not necessarily correlate with protein expression status. Therefore, combined mRNA and proteomic investigations are expected to provide maximal insight into cell physiology. An overall comparison of metatranscriptome and metaproteome data showed only a qualitative agreement regarding expression differences. Absolute values differed greatly, as has been found in an earlier study (Scherl *et al.*, 2004; Resch *et al.*, 2006). Therefore, the absolute difference factor in protein expression is very difficult to measure. Given these limitations, the overall similarities in the trends of differential expression data determined by metatranscriptomic and metaproteomic approaches are quite high.

Additionally, prokaryotic mRNA half-life times are short and highly variable (ranging from 1 – 46 min (Steiner et al. 2019)). Messenger RNA is also about four orders of magnitude less abundant than proteins (Ingraham et al., 1983; Bernstein et al., 2002; Hambraeus et al., 2003; Selinger et al., 2003; Taniguchi et al., 2010; Moran et al., 2013) and prokaryotic proteins reside longer in the ocean (up to 20 h) than mRNA. Thus, there is a temporal decoupling that needs to be considered as proteins persist in a bacterial cell long after the mRNA that encoded them (Koch and Levy, 1955; Borek et al., 1958; Mandelstam, 1958; Moran et al., 2013). However, we detected several transporter proteins in the metatranscriptomes which were absent in the metaproteomes (Fig. 5). This can be explained by the sensibility of the methods used. The RNA-sequencing depth used in our study provides a higher throughput compared to SDS-page separated proteins. The contradictory results might be artifacts caused by false-positive results either on the mRNA or the protein level or might reflect post-transcriptional regulation involving, for example, protein half-life and thereby accumulation of the protein or protein modification.

In line with results from metaproteomic studies in the Antarctic peninsula from 6 – 70 m depth (Williams et al. 2012) and in the North Atlantic (Bergauer et al. 2017), substrate-binding proteins were the most dominant transport proteins at all stations and revisits (Fig. 4A, B). The second most abundant transporter system was the outer-membrane receptors (OMR) comprising all TonB-dependent receptors which play an important role in the uptake of high-molecular-weight organic matter (McCarren *et al.*, 2010), and have been reported in meta-omics datasets from surface waters (Morris *et al.*, 2010; Tang *et al.*, 2012; Li *et al.*, 2014). OMRs are involved in proton motive force dependent outer membrane transport in many Gram-negative bacteria. These transporters are involved in the binding and uptake of iron–siderophore complexes, cobalamin (Koebnik, 2005) or carbohydrates (Blanvillain *et al.*, 2007). TonB-dependent receptors from metaproteomes of a coastal Southern Ocean site increased in

summer compared to winter which was interpreted as substrate scavenging, potentially representing an important strategy for some prokaryotic clades at this site (Williams *et al.*, 2012). TRAP-T and TTTs are secondary active transporters of Bacteria and Archaea that retain solute binding proteins to bind ligands (Kelly *et al.*, 1999; Kelly and Thomas, 2001; Winnen *et al.*, 2003; Mulligan *et al.*, 2011). They were only detected at low relative abundance in both datasets similar to previous reports, potentially due to the lack of reference sequences in public databases and the limited knowledge of the substrate range of these transporters (Antoine *et al.*, 2005; Bergauer *et al.*, 2018). Porins were found in all samples and seasons and can create diffusion channels of different size (Nikaido, 2003). Members of *Vibrionaceae* employ a chitoporin channel for the uptake of chitin, the most abundant biopolymer in marine ecosystems (Suginta *et al.*, 2013). Carbohydrates with a size of ~ 600 Da can either diffuse freely into cells of Gram-negative bacteria (Decad and Nikaido, 1976), or are channeled into the cells via porins (Weiss et al., 1991). *Flavobacterial* SusD-like TonB-dependent receptors have been observed coupled with SusC-like TonB-dependent transporter porins, increasing significantly in abundances during phytoplankton blooms (Teeling *et al.*, 2012, 2016).

The exact annotation of compounds specific for transporters remains a challenge due to the complexity and multitude of protein domains and classification databases (Kloppmann *et al.*, 2012; Chiang *et al.*, 2015; Ijaq *et al.*, 2015). ABC transporters can transport several compounds depending on their size and charge, while TonB transporters were initially assumed to exclusively transport Fe-related compounds. However, studies suggest that TonB transporters can transport a wide range of substrates (Schauer *et al.*, 2008).

Amino acids represent essential components of the organic matter pool in surface waters of the Southern Ocean, including the study region (Tremblay *et al.*, 2015). In a previous study, ABC transporters for general amino acids, branched-chain amino acids, glycine betaine and sugars were significantly more abundant in winter than in summer in a coastal Antarctic metaproteome (Williams *et al.*, 2012). Glycine betaine is a compatible solute and a potentially important cryoprotectant in the Antarctic environment. Glycine betaine can also be catabolized to glycine which was proposed to be of particular importance for members of the globally abundant SAR11 clade, which are auxotrophic for glycine (Tripp *et al.*, 2009; Williams *et al.*, 2012).

Several of the proposed transported compounds indicate the potential availability of labile organic matter during early spring, including polyamines, taurine and urea. Transporters for polyamines such as spermidine and putrescine have previously been identified in genomes and metaproteomes from surface waters (Sowell *et al.*, 2009; Bergauer *et al.*, 2018). The amino

acid-like compound taurine can be a source of nitrogen, carbon and sulfur, and has recently been suggested as a major carbon and energy source for heterotrophic prokaryotes (Clifford *et al.*, 2019). Furthermore, urea can be used as a carbon and nitrogen source for diverse prokaryotes in the Arctic summer and winter, including autotrophic ammonia-oxidizing archaea (Alonso-Saez *et al.*, 2012; Connelly *et al.*, 2014).

Flavobacteriales and *Gammaproteobacteria* (*Pseudoalteromonas, Pseudomondales, Litoricolacea*) potentially use OMRs to scavenge iron and carbohydrates dissolved in seawater or associated with particulate organic matter (Blanvillain *et al.*, 2007; Tang *et al.*, 2012; Williams *et al.*, 2012). OMR utilization is adapted for the acquisition of biopolymers and other nutrients derived from phytoplankton generally of a size larger than 600 Da (Tang *et al.*, 2012). Our results indicate a greater reliance on OMR mediated uptake of *Bacteroidota* and some alpha- and gammaproteobacterial families, such as *Pseudomondales* and *Flavobacterales,* particularly in the late summer, as compared to members of the *Pelagibacteracea* (HIMB59) and *Rhodobacteracae,* which might rely more on labile solute high-affinity uptake systems such as ABC transporters or porins. In line with these results members of the FCB cluster were found highly represented in their consumption of low-molecular-weight dissolved organic matter while representatives of *Alphaproteobacteria* were observed to largely consume amino acids when incubated from estuarine and coastal environments (Cottrell and Kirchman, 2000).

Transporter proteins for Fe (OMR & ABC) were higher in relative abundance in late summer samples compared to early spring (Fig. 5 & 7A). In relative transcript abundance there was no difference between seasons, while differential expression results showed Fe transporter expression with highest log2 fold changes in late summer (Fig. 8). Heterotrophic prokaryotic and phytoplankton actively compete for Fe (Fourquez *et al.*, 2015). Studies in the Kerguelen area have observed that the phytoplankton community is dominated by small, fast-growing cells that outcompete heterotrophic bacteria for Fe uptake in spring, while the presence of larger diatoms in summer could result in higher organic carbon and Fe availability for heterotrophic prokaryotic (Quéguiner, 2013; Fourquez *et al.*, 2015). However, Fe limitation could also be reduced in summer as a result of enhanced Fe regeneration mediated by biological activity (Bowie *et al.*, 2015). The idea of seasonal changes in resource limitation is further supported by the higher prokaryotic Fe quota and cell-specific Fe uptake rates in spring than in summer, which point to enhanced prokaryotic Fe requirements early in the season (Fourquez *et al.*, 2015; Obernosterer *et al.*, 2015). Resource supply and biological interactions determine both the extent of Fe and carbon limitation of heterotrophic prokaryotes, with possible important feedbacks on the Fe and C cycles in the Southern Ocean.

References

Alneberg, J., Bjarnason, B.S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U.Z., et al. (2014) Binning metagenomic contigs by coverage and composition. *Nat Methods* **11**: 1144–1146.

Alonso-Saez, L., Waller, A.S., Mende, D.R., Bakker, K., Farnelid, H., Yager, P.L., et al. (2012) Role for urea in nitrification by polar marine Archaea. *Proceedings of the National Academy of Sciences* **109**: 17989–17994.

Antoine, R., Huvent, I., Chemlal, K., Deray, I., Raze, D., Locht, C., and Jacob-Dubuisson, F. (2005) The periplasmic binding protein of a tripartite tricarboxylate transporter is involved in signal transduction. *J Mol Biol* **351**: 799–809.

Baines, S.B. and Pace, M.L. (1991) The production of dissolved organic matter by phytoplankton and its importance to bacteria: Patterns across marine and freshwater systems. *Limnology and Oceanography* **36**: 1078–1090.

Baltar, F., Gutiérrez-Rodríguez, A., Meyer, M., Skudelny, I., Sander, S., Thomson, B., et al. (2018) Specific Effect of Trace Metals on Marine Heterotrophic Microbial Activity and Diversity: Key Role of Iron and Zinc and Hydrocarbon-Degrading Bacteria. *Front Microbiol* **9**:.

Bensadoun, A. and Weinstein, D. (1976) Assay of proteins in the presence of interfering materials. *Analytical Biochemistry* **70**: 241–250.

Bercovici, S.K. and Hansell, D.A. (2016) Dissolved organic carbon in the deep Southern Ocean: Local versus distant controls. *Global Biogeochemical Cycles* **30**: 350–360.

Bergauer, K., Fernandez-Guerra, A., Garcia, J.A.L., Sprenger, R.R., Stepanauskas, R., Pachiadaki, M.G., et al. (2018) Organic matter processing by microbial communities throughout the Atlantic water column as revealed by metaproteomics. *PNAS* **115**: E400–E408.

Blain, S., Queguiner, B., Armand, L.K., Belviso, S., and Bomb, B. (2007) Effect of natural iron fertilization on carbon sequestration in the Southern Ocean. *Nature* **446**: 1070–1074.

Blain, S., Sarthou, G., and Laan, P. (2008) Distribution of dissolved iron during the natural iron-fertilization experiment KEOPS (Kerguelen Plateau, Southern Ocean). *Deep Sea Research Part II: Topical Studies in Oceanography* **55**: 594–605.

Blanvillain, S., Meyer, D., Boulanger, A., Lautier, M., Guynet, C., Denancé, N., et al. (2007) Plant Carbohydrate Scavenging through TonB-Dependent Receptors: A Feature Shared by Phytopathogenic and Aquatic Bacteria. *PLOS ONE* **2**: e224.

Bolger, A.M., Lohse, M., and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Bowie, A.R., van der Merwe, P., Quéroué, F., Trull, T., Fourquez, M., Planchon, F., et al. (2015) Iron budgets for three distinct biogeochemical sites around the Kerguelen Archipelago (Southern Ocean) during the natural fertilisation study, KEOPS-2. *Biogeosciences* **12**: 4421–4445.

Burg, D.W., Lauro, F.M., Williams, T.J., Raftery, M.J., Guilhaus, M., and Cavicchioli, R. (2010) Analyzing the hydrophobic proteome of the antarctic archaeon Methanococcoides burtonii using differential solubility fractionation. *J Proteome Res* **9**: 664–676.

Chiang, Z., Vastermark, A., Punta, M., Coggill, P.C., Mistry, J., Finn, R.D., and Saier, M.H. (2015) The complexity, challenges and benefits of comparing two transporter classification systems in TCDB and Pfam. *Brief Bioinform* **16**: 865–872.

Christaki, U., Obernosterer, I., Van Wambeke, F., Veldhuis, M., Garcia, N., and Catala, P. (2008) Microbial food web structure in a naturally iron-fertilized area in the Southern Ocean (Kerguelen Plateau). *Deep Sea Research Part II: Topical Studies in Oceanography* **55**: 706–719.

Church, M.J., Hutchins, D.A., and Ducklow, H.W. (2000) Limitation of Bacterial Growth by Dissolved Organic Matter and Iron in the Southern Ocean. *Appl Environ Microbiol* **66**: 455–466.

Clifford, E.L., Varela, M.M., De Corte, D., Bode, A., Ortiz, V., Herndl, G.J., and Sintes, E. (2019) Taurine Is a Major Carbon and Energy Source for Marine Prokaryotes in the North Atlantic Ocean off the Iberian Peninsula. *Microb Ecol*.

Connelly, T.L., Baer, S.E., Cooper, J.T., Bronk, D.A., and Wawrik, B. (2014) Urea Uptake and Carbon Fixation by Marine Pelagic Bacteria and Archaea during the Arctic Summer and Winter Seasons. *Applied and Environmental Microbiology* **80**: 6013–6022.

Cottrell, M.T. and Kirchman, D.L. (2000) Natural assemblages of marine proteobacteria and members of the Cytophaga-Flavobacter cluster consuming low- and high-molecular-weight dissolved organic matter. *Appl Environ Microbiol* **66**: 1692–1697.

Decad, G.M. and Nikaido, H. (1976) Outer membrane of gram-negative bacteria. XII. Molecular-sieving function of cell wall. *J Bacteriol* **128**: 325–336.
Emms, D.M. and Kelly, S. (2015) OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol* **16**: 157.
Fogg, G.E. (1983) ecological significance of extracellular products of phytoplankton photosynthesis. *Botanica marina*.
Fourquez, M., Obernosterer, I., Davies, D.M., Trull, T.W., and Blain, S. (2015) Microbial iron uptake in the naturally fertilized waters in the vicinity of the Kerguelen Islands: phytoplankton–bacteria interactions. *Biogeosciences* **12**: 1893–1906.
Green, N.W., Perdue, E.M., Aiken, G.R., Butler, K.D., Chen, H., Dittmar, T., et al. (2014) An intercomparison of three methods for the large-scale isolation of oceanic dissolved organic matter. *Marine Chemistry* **161**: 14–19.
Grzymski, J.J., Riesenfeld, C.S., Williams, T.J., Dussaq, A.M., Ducklow, H., Erickson, M., et al. (2012) A metagenomic assessment of winter and summer bacterioplankton from Antarctica Peninsula coastal surface waters. *The ISME Journal* **6**:.
Hansell, D., Carlson, C., Repeta, D., and Schlitzer, R. (2009) Dissolved Organic Matter in the Ocean: A Controversy Stimulates New Insights. *Oceanography* **22**: 202–211.
Hansell, D.A. (2013) Recalcitrant Dissolved Organic Carbon Fractions. *Annual Review of Marine Science* **5**: 421–445.
Hyatt, D., Chen, G.-L., LoCascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**: 119.
Ijaq, J., Chandrasekharan, M., Poddar, R., Bethi, N., and Sundararajan, V.S. (2015) Annotation and curation of uncharacterized proteins- challenges. *Front Genet* **6**:.
Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K., von Haeseler, A., and Jermiin, L.S. (2017) ModelFinder: Fast Model Selection for Accurate Phylogenetic Estimates. *Nat Methods* **14**: 587–589.
Kang, D., Li, F., Kirton, E.S., Thomas, A., Egan, R.S., An, H., and Wang, Z. (2019) MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies, PeerJ Inc.
Katoh, K. and Standley, D.M. (2013) MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol Biol Evol* **30**: 772–780.
Kelly, D.J., Rabus, R., Saier, Jr, M.H., and Jack, D.L. (1999) TRAP transporters: an ancient family of extracytoplasmic solute- receptor-dependent secondary active transporters. *Microbiology* **145**: 3431–3445.
Kelly, D.J. and Thomas, G.H. (2001) The tripartite ATP-independent periplasmic (TRAP) transporters of bacteria and archaea. *FEMS Microbiol Rev* **25**: 405–424.
Kieft, B., Li, Z., Bryson, S., Crump, B.C., Hettich, R., Pan, C., et al. (2018) Microbial Community Structure–Function Relationships in Yaquina Bay Estuary Reveal Spatially Distinct Carbon and Nitrogen Cycling Capacities. *Front Microbiol* **9**:.
Kloppmann, E., Punta, M., and Rost, B. (2012) Structural genomics plucks high-hanging membrane proteins. *Current Opinion in Structural Biology* **22**: 326–332.
Koebnik, R. (2005) TonB-dependent trans-envelope signalling: the exception or the rule? *Trends Microbiol* **13**: 343–347.
Landa, M., Blain, S., Harmand, J., Monchy, S., Rapaport, A., and Obernosterer, I. (2018) Major changes in the composition of a Southern Ocean bacterial community in response to diatom-derived dissolved organic matter. *FEMS Microbiol Ecol* **94**:.
Lechtenfeld, O.J., Kattner, G., Flerus, R., McCallister, S.L., Schmitt-Kopplin, P., and Koch, B.P. (2014) Molecular transformation and degradation of refractory dissolved organic matter in the Atlantic and Southern Ocean. *Geochimica et Cosmochimica Acta* **126**: 321–337.
Li, D., Liu, C.-M., Luo, R., Sadakane, K., and Lam, T.-W. (2015) MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**: 1674–1676.
Li, D.-X., Zhang, H., Chen, X.-H., Xie, Z.-X., Zhang, Y., Zhang, S.-F., et al. (2018) Metaproteomics reveals major microbial players and their metabolic activities during the blooming period of a marine dinoflagellate Prorocentrum donghaiense. *Environmental Microbiology* **20**: 632–644.
Li, M., Toner, B.M., Baker, B.J., Breier, J.A., Sheik, C.S., and Dick, G.J. (2014) Microbial iron uptake as a mechanism for dispersing iron from deep-sea hydrothermal vents. *Nature Communications* **5**: 3192.

Liu, Y., Debeljak, P., Rembauville, M., Blain, S., and Obernosterer, I. (2019) Diatoms shape the biogeography of heterotrophic prokaryotes in early spring in the Southern Ocean. *Environmental Microbiology* **21**: 1452–1465.
Luria, C., Amaral-Zettler, L., Ducklow, H., and Rich, J. (2016) Seasonal succession of bacterial communities in coastal waters of the western Antarctic Peninsula. *Frontiers in Microbiology* 1–13.
Luria, C.M., Ducklow, H.W., and Amaral-Zettler, L.A. (2014) Marine bacterial, archaeal and eukaryotic diversity and community structure on the continental shelf of the western Antarctic Peninsula. *Aquatic Microbial Ecology* **73**: 107–121.
Martin, J.H., Fitzwater, S.E., and Gordon, R.M. (1990) Iron deficiency limits phytoplankton growth in Antarctic waters. *Global Biogeochemical Cycles* **4**: 5–12.
Martin, J.H., Gordon, R.M., and Fitzwater, S.E. (1990) Iron in Antarctic waters. *Nature* **345**: 156.
McCarren, J., Becker, J.W., Repeta, D.J., Shi, Y., Young, C.R., Malmstrom, R.R., et al. (2010) Microbial community transcriptomes reveal microbes and metabolic pathways associated with dissolved organic matter turnover in the sea. *PNAS* **107**: 16420–16427.
Medeiros, P.M., Seidel, M., Powers, L.C., Dittmar, T., Hansell, D.A., and Miller, W.L. (2015) Dissolved organic matter composition and photochemical transformations in the northern North Pacific Ocean. *Geophysical Research Letters* **42**: 863–870.
Morán, X.A.G., Ducklow, H.W., and Erickson, M. (2013) Carbon fluxes through estuarine bacteria reflect coupling with phytoplankton. **489**: 75–85.
Morris, R.M., Nunn, B.L., Frazar, C., Goodlett, D.R., Ting, Y.S., and Rocap, G. (2010) Comparative metaproteomics reveals ocean-scale shifts in microbial nutrient utilization and energy transduction. *ISME J* **4**: 673–685.
Mulligan, C., Fischer, M., and Thomas, G.H. (2011) Tripartite ATP-independent periplasmic (TRAP) transporters in bacteria and archaea. *FEMS Microbiol Rev* **35**: 68–86.
Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015) IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol Biol Evol* **32**: 268–274.
Nikaido, H. (2003) Molecular Basis of Bacterial Outer Membrane Permeability Revisited. *Microbiol Mol Biol Rev* **67**: 593–656.
Obernosterer, I., Christaki, U., Lefèvre, D., Catala, P., Van Wambeke, F., and Lebaron, P. (2008) Rapid bacterial mineralization of organic carbon produced during a phytoplankton bloom induced by natural iron fertilization in the Southern Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography* **55**: 777–789.
Obernosterer, I., Fourquez, M., and Blain, S. (2015) Fe and C co-limitation of heterotrophic bacteria in the naturally fertilized region off the Kerguelen Islands. *Biogeosciences* **12**: 1983–1992.
Osterholz, H., Niggemann, J., Giebel, H.-A., Simon, M., and Dittmar, T. (2015) Inefficient microbial production of refractory dissolved organic matter in the ocean. *Nature Communications* **6**: 7422.
Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarshewski, A., Chaumeil, P.-A., and Hugenholtz, P. (2018) A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nature Biotechnology* **36**: 996–1004.
Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **25**: 1043–1055.
Quéguiner, B. (2013) Iron fertilization and the structure of planktonic communities in high nutrient regions of the Southern Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography* **90**: 43–54.
Quéroué, F., Sarthou, G., Planquette, H.F., Bucciarelli, E., Chever, F., van der Merwe, P., et al. (2015) High variability in dissolved iron concentrations in the vicinity of the Kerguelen Islands (Southern Ocean). *Biogeosciences* **12**: 3869–3883.
Resch, A., Leicht, S., Saric, M., Pásztor, L., Jakob, A., Götz, F., and Nordheim, A. (2006) Comparative proteome analysis of Staphylococcus aureus biofilm and planktonic cells and correlation with transcriptome profiling. *PROTEOMICS* **6**: 1867–1877.
Sarmiento, J.L. and Gruber, N. (2006) Ocean Biogeochemical Dynamics, Princeton University Press.
Schauer, K., Rodionov, D.A., and de Reuse, H. (2008) New substrates for TonB-dependent transport: do we only see the 'tip of the iceberg'? *Trends in Biochemical Sciences* **33**: 330–338.
Scherl, A., Sanchez, J.-C., and Hochstrasser, D.F. (2004) Challenges in the Overall Analysis of Microbial Proteomes. *CPU* **2**: 79–86.

Sczyrba, A., Hofmann, P., Belmann, P., Koslicki, D., Janssen, S., Dröge, J., et al. (2017) Critical Assessment of Metagenome Interpretation—a benchmark of metagenomics software. *Nature Methods* **14**: 1063–1071.
Sharon, I. and Banfield, J.F. (2013) Genomes from Metagenomics. *Science* **342**: 1057–1058.
Shevchenko, A., Tomas, H., Havliš, J., Olsen, J. V., and Mann, M. (2007) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols* **1**: 2856–2860.
Sowell, S.M., Wilhelm, L.J., Norbeck, A.D., Lipton, M.S., Nicora, C.D., Barofsky, D.F., et al. (2009) Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *ISME J* **3**: 93–105.
Steiner, P.A., Corte, D.D., Geijo, J., Mena, C., Yokokawa, T., Rattei, T., et al. Highly variable mRNA half-life time within marine bacterial taxa and functional genes. *Environmental Microbiology* **0**:
Straza, T.R.A., Ducklow, H.W., Murray, A.E., and Kirchman, D.L. (2010) Abundance and single-cell activity of bacterial groups in Antarctic coastal waters. *Limnology and Oceanography* **55**: 2526–2536.
Suginta, W., Chumjan, W., Mahendran, K.R., Janning, P., Schulte, A., and Winterhalter, M. (2013) Molecular Uptake of Chitooligosaccharides through Chitoporin from the Marine Bacterium Vibrio harveyi. *PLOS ONE* **8**: e55126.
Tagliabue, A., Bowie, A.R., Philip, W., Buck, K.N., Johnson, K.S., and Saito, M.A. (2017) The integral role of iron in ocean biogeochemistry. *Nature* **543**: In Press.
Tagliabue, A., Sallée, J.-B., Bowie, A.R., Lévy, M., Swart, S., and Boyd, P.W. (2014) Surface-water iron supplies in the Southern Ocean sustained by deep winter mixing. *Nature Geoscience* **7**: 314–320.
Tang, K., Jiao, N., Liu, K., Zhang, Y., and Li, S. (2012) Distribution and functions of tonb-dependent transporters in marine bacteria and environments: Implications for dissolved organic matter utilization. *PLoS ONE* **7**:.
Teeling, H., Fuchs, B.M., Becher, D., Klockow, C., Gardebrecht, A., Bennke, C.M., et al. (2012) Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* **336**: 608–611.
Teeling, H., Fuchs, B.M., Bennke, C.M., Krüger, K., Chafee, M., Kappelmann, L., et al. (2016) Recurring patterns in bacterioplankton dynamics during coastal spring algae blooms. *eLife* **5**: e11888.
Tortell, P.D., Maldonado, M.T., and Price, N.M. (1996) The role of heterotrophic bacteria in iron-limited ocean ecosystems. *Nature* **383**: 330–332.
Tremblay, L., Caparros, J., LEBLANC, K., and Obernosterer, I. (2015) Origin and fate of particulate and dissolved organic matter in a naturally iron-fertilized region of the Southern Ocean. *Biogeosciences* **12**: 607–621.
Tripp, H.J., Schwalbach, M.S., Meyer, M.M., Kitner, J.B., Breaker, R.R., and Giovannoni, S.J. (2009) Unique glycine-activated riboswitch linked to glycine-serine auxotrophy in SAR11. *Environ Microbiol* **11**: 230–238.
Uritskiy, G.V., DiRuggiero, J., and Taylor, J. (2018) MetaWRAP-a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome* **6**: 158.
Valledor, L. and Weckwerth, W. (2014) An Improved Detergent-Compatible Gel-Fractionation LC-LTQ-Orbitrap-MS Workflow for Plant and Microbial Proteomics. In, Jorrin-Novo,J. V, Komatsu,S., Weckwerth,W., and Wienkoop,S. (eds), *Plant Proteomics: Methods and Protocols*. Totowa, NJ: Humana Press, pp. 347–358.
Weiss, M.S., Abele, U., Weckesser, J., Welte, W., Schiltz, E., and Schulz, G.E. (1991) Molecular architecture and electrostatic properties of a bacterial porin. *Science* **254**: 1627–1630.
Williams, T.J., Burg, D.W., Raftery, M.J., Poljak, A., Guilhaus, M., Pilak, O., and Cavicchioli, R. (2010) Global proteomic analysis of the insoluble, soluble, and supernatant fractions of the psychrophilic archaeon Methanococcoides burtonii. Part I: the effect of growth temperature. *J Proteome Res* **9**: 640–652.
Williams, T.J., Long, E., Evans, F., Demaere, M.Z., Lauro, F.M., Raftery, M.J., et al. (2012) ORIGINAL ARTICLE A metaproteomic assessment of winter and summer bacterioplankton from Antarctic Peninsula coastal surface waters. *The ISME Journal* **6**: 1883–1900.
Williams, T.J., Wilkins, D., Long, E., Evans, F., Demaere, M.Z., Raftery, M.J., and Cavicchioli, R. (2013) The role of planktonic Flavobacteria in processing algal organic matter in coastal East Antarctica revealed using metagenomics and metaproteomics. *Environmental Microbiology* **15**: 1302–1317.
Winnen, B., Hvorup, R.N., and Saier, M.H. (2003) The tripartite tricarboxylate transporter (TTT) family. *Res Microbiol* **154**: 457–465.

Wu, Y.-W., Simmons, B.A., and Singer, S.W. (2016) MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* **32**: 605–607.
Zhang, Y., Wen, Z., Washburn, M.P., and Florens, L. (2015) Improving Label-Free Quantitative Proteomics Strategies by Distributing Shared Peptides and Stabilizing Variance. *Anal Chem* **87**: 4749–4756.

Chapter 3 Supplementary Tables and Figures

Supplementary table 1. *Results of MS Analysis for Metaproteomes*

Sample_ID	Season	Protein groups	>2 unique peptides	Depth	Protein conc. (ug)	MS
A3_2_I	ES	1983	487	10	19.85	Elite
A3_2_II	ES	4966	1487	10	27.08	Elite + QExactive
A3_I	ES	3809	1028	10	25.62	Elite + QExactive
A3_II	ES	5849	1758	10	17.40	Elite + QExactive
KERFIX_I	ES	1147	266	10	6.30	Elite
KERFIX_II	ES	4369	1293	10	25.02	Elite + QExactive
A3_I	LS	5363	1900	10	36.39	Elite + QExactive
A3_II	LS	6458	2105	10	41.89	Elite + QExactive
A3_2_I	LS	6114	2272	10	30.99	Elite + QExactive
A3_2_II	LS	4254	1457	10	22.39	Elite + QExactive
A3_3_I	LS	3408	1172	10	24.67	Elite + QExactive
A3_3_II	LS	4844	1722	10	29.32	Elite + QExactive
KERFIX_I	LS	4218	1436	10	33.06	Elite + QExactive
KERFIX_II	LS	6178	2104	10	45.32	Elite + QExactive
KERFIX_2_I	LS	4902	1647	10	35.44	Elite + QExactive
KERFIX_2_II	LS	2636	765	10	10.58	Elite + QExactive

Supplementary Table 2. *Detailed information on libraries used for metagenomic assembly*

Fasteris Code	Cruise	Sample Name	Paired-end	Sequence length bp	Total sequences	%GC	After bbduk	Kept from total %	Size GB fq	bbnormed size GB .fq	contigs	min (bp)	max (bp)	avg (bp)	N50 (bp)
AFQF-1	Mobydick	A3	R1	150	328265070	36	327826974	99.87%	219	124	3419066	200	413998	519	647
			R2	150	328265070	36	327826974								
AFQF-2	Mobydick	KERFIX	R1	150	336948192	35	336309436	99.81%	224	140	4373463	200	500480	509	608
			R2	150	336948192	35	336309436								
HDC-14	Soclim	A3_2	R1	150	339113726	35	337330662	99.47%	225	139	4410036	200	271226	533	647
			R2	150	339113726	35	337330662								
AFQF-3	Mobydick	M4_10m	R1	150	285051512	35	285050382	100%	190	120	3776178	200	321983	523	660
			R2	150	285051512	35	285050382								

Supplementary Table 3. *Outcome of Co-Assembly*

Fasteris Code	contigs	min (bp)	max (bp)	avg (bp)	N50 (bp)	Scaffolds >2.5Kbp	Genes in scaffolds >2.5kbp
COASSEMBLY	949228	1000	636072	2536	2973	235115	1484910

Supplementary Table 4. *Detailed information on metatranscriptomic libraries.*

Fasteris Code	Cruise	Sample Name	Paired-end	Sequence length bp	Total sequences	%GC	After Quality trim	%rRNA	%Standard M5	%Standard M6	Total coding RNA
JFG-1	SOCLIM	A3_2_1	R1	4-150	27193587	40	25722177				
			R2	4-150	27193587	39	25722177	2.58	0.0016	0.0013	50117198
JFG-2	SOCLIM	A3_2_2	R1	4-150	36020042	39	34788342				
			R2	4-150	36020042	39	34788342	1.92	0.0020	0.0011	68242486
JFG-3	SOCLIM	A3_2_3	R1	4-150	30664893	40	29104670				
			R2	4-150	30664893	39	29104670	2.03	0.0021	0.0017	57030672
JFG-4	SOCLIM	Kerfix_1	R1	4-150	28679440	42	27473352				
			R2	4-150	28679440	41	27473352	3.24	0.0016	0.0012	53165137
JFG-5	SOCLIM	Kerfix_2	R1	4-150	26218171	42	25027199				
			R2	4-150	26218171	42	25027199	2.63	0.0029	0.0018	48738697
JFG-6	SOCLIM	Kerfix_3	R1	4-150	28038253	41	27247120				
			R2	4-150	28038253	41	27247120	2.64	0.0013	0.0001	53056004
AFQF-4	MOBYDICK	A3_2_1	R1	4-150	30639599	41	29491007	0.84	0.0019	0.0034	58486815
			R2	4-150	30639599	41	29491007				
AFQF-5	MOBYDICK	A3_2_2	R1	4-150	34059459	41	32599016	1.33	0.0010	0.0032	64015950
			R2	4-150	34059459	41	32599016				
AFQF-6	MOBYDICK	A3_2_3	R1	4-150	26714546	41	25735518	1.36	0.0010	0.0028	50770154
			R2	4-150	26714546	43	29256664				
AFQF-7	MOBYDICK	Kerfix_1	R1	4-150	30337390	42	29256664	1.39	0.0019	0.0027	57697979
			R2	4-150	30337390	42	29256664				
AFQF-8	MOBYDICK	Kerfix_2	R1	4-150	27749121	42	26789683	0.66	0.0011	0.0041	53227208
			R2	4-150	27749121	42	26789683				
AFQF-9	MOBYDICK	Kerfix_3	R1	4-150	31034721	42	29862712	0.79	0.0007	0.0069	59252758
			R2	4-150	31034721	42	29862712				

Supplementary Table 5. *Detailed information on mapping results of each metatranscriptomic library to Co-Assembly*

Status	CoAss_SOMD_1kbp.MT_A3_S.Rep1	CoAss_SOMD_1kbp.MT_A3_S.Rep2	CoAss_SOMD_1kbp.MT_A3_S.Rep3	CoAss_SOMD_1kbp.MT_KERFIX_S.Rep1.	CoAss_SOMD_1kbp.MT_KERFIX_S.Rep2	CoAss_SOMD_1kbp.MT_KERFIX_S.Rep3	CoAss_SOMD_1kbp.MT_A3_M.Rep1	CoAss_SOMD_1kbp.MT_A3_M.Rep2	CoAss_SOMD_1kbp.MT_A3_M.Rep3	CoAss_SOMD_1kbp.MT_KERFIX_M.Rep1	CoAss_SOMD_1kbp.MT_KERFIX_M Rep2	CoAss_SOMD_1kbp.MT_KERFIX_M.Rep3
Total alignments	8155250	10112926	8660155	8157526	7059359	8153968	6754266	8357229	6517310	7879145	6972736	6789727
Assigned	4868246	6007459	5156719	4907442	4301115	4977428	4229410	5084837	3990214	4755545	4229814	4132173
Successfully aligned	59.6946262	59.403767	59.545343	60.15845981	60.92784062	61.0430161	62.6183511	60.84357626	61.2248612	60.3561046	60.6621848	60.8591921
Unassigned_MappingQuality	2093676	2538416	2078717	1532763	1297943	1626740	1746037	2174159	1738617	2257518	2058932	1997573
Unassigned_Chimera	130599	188716	141977	99364	85529	110109	101759	121534	91689	159378	121959	130097
Unassigned_No Features	831604	1054978	996986	1372639	1160937	1187845	500676	761738	524691	520595	397846	366958
Unassigned_Ambiguity	231125	323357	285756	245318	213835	251846	176384	214961	172099	186109	164185	162926

Supplementary Table 6 Description of 133 Metagenome assembled genomes by GTDB taxonomy

Contig	Completion %	Redundancy	Length bp	GC %	domain	phylum	class	order	family	genus
BIN10	23.74	0.72	1448918	43.08	Archaea	Thermoplasmatota	Poseidoniia	Poseidoniales	Poseidoniaceae	MGIIa-K1
BIN100	56.83	0	1943337	38.00	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	GCA-002722245	
BIN101	91.37	8.63	2110086	43.17	Bacteria	Proteobacteria	Alphaproteobacteria	Puniceispirillales	Puniceispirillaceae	Puniceispirillum
BIN102	78.42	7.91	1729333	40.90	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodospirillales_A	Casp-alpha2	Casp-alpha2
BIN103	52.52	0.72	3003507	48.60	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	HTCC2089	UBA4421
BIN104	68.35	5.04	708557	30.00	Bacteria	Proteobacteria	Alphaproteobacteria	Pelagibacterales	Pelagibacteraceae	
BIN105	58.27	2.16	855007	30.89	Bacteria	Proteobacteria	Gammaproteobacteria	GCA-002705445	GCA-002716945	GCA-2698665
BIN106	94.96	2.16	3667991	66.39	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Limimaricola

BIN107	41.01	15.83	3863468	51.07	Bacteria	Myxococcota	Bradimonadia	UBA7976	UBA1532	
BIN108	40.29	1.44	1723488	32.19	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	GCA-002722245	
BIN110	79.14	1.44	1045386	32.21	Bacteria	Chloroflexota	Dehalococcoidia	UBA1151	TMED-70	GCA-002700125
BIN112	90.65	1.44	4942583	38.64	Bacteria	Bacteroidota	Bacteroidia	Chitinophagales	Saprospiraceae	Aureispira
BIN113	82.01	0.72	1484225	27.21	Bacteria	Bacteroidota	Bacteroidia	Cytophagales	Marinoscillaceae	
BIN114	53.24	0.72	1248114	34.48	Bacteria	Proteobacteria	Gammaproteobacteria	Burkholderiales	Methylophilaceae	BACL14
BIN115	49.64	0.72	2109252	44.89	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Nitrincolaceae	ASP10-02a
BIN116	78.42	0.72	3606987	64.30	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingobium
BIN117	76.98	19.42	2294111	44.49	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Litoricolaceae	
BIN118	82.01	4.32	2784470	51.04	Bacteria	Planctomycetota	Phycisphaerae	Phycisphaerales	SM1A02	
BIN119	71.22	3.6	2041425	32.94	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	GCA-002722245	GCA-002722245
BIN12	41.01	0.72	1762410	32.58	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	koll-22	GCA-002793235
BIN120	34.53	0.72	2048021	68.69	Bacteria	Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	Phenylobacterium
BIN121	53.24	1.44	1168658	42.44	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	UA16	UBA974
BIN122	87.77	0	1157673	36.62	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Porticoccaceae	Porticoccus
BIN123	71.94	0.72	1740423	48.15	Bacteria	Planctomycetota	Planctomycetes	Pirellulales	UBA1268	UBA1268
BIN124	82.73	0.72	1193751	29.17	Bacteria	Proteobacteria	Alphaproteobacteria	TMED109	TMED131	MED722
BIN125	44.6	7.19	1772490	40.53	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodospirillales_A	UBA3470	
BIN126	35.25	1.44	1652466	35.06	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	UBA7428
BIN127	41.73	1.44	2423633	33.20	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	
BIN128	48.2	0	661390	28.20	Bacteria	Proteobacteria	Alphaproteobacteria	HIMB59	HIMB59	HIMB59
BIN129	57.55	0.72	2133249	49.05	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	UA16	UA16
BIN13	54.68	0.72	1900140	48.83	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Porticoccaceae	HTCC2207
BIN130	58.27	0	1901493	43.22	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Psychrobacter
BIN131	28.06	5.04	561945	28.97	Bacteria	Proteobacteria	Alphaproteobacteria	Pelagibacterales	Pelagibacteraceae	Pelagibacter_A
BIN132	76.98	0.72	2073806	69.34	Bacteria	Actinobacteriota	Actinobacteria	Actinomycetales	Microbacteriaceae	UBA1487
BIN133	53.96	0	1200392	28.19	Bacteria	Proteobacteria	Alphaproteobacteria	HIMB59	GCA-002718135	MarineAlpha5-Bin3

BIN134	71.94	0.72	2626422	34.82	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	1G12	
BIN135	41.01	10.07	2888914	64.33	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Sphingopyxis
BIN136	44.6	0	735855	37.76	Bacteria	Proteobacteria	Gammaproteobacteria	Thiomicrospirales	Thioglobaceae	Thioglobus_A
BIN137	46.04	4.32	2856443	60.96	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Erythrobacter_A
BIN138	43.17	11.51	1216247	44.91	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Porticoccaceae	HTCC2207
BIN139	30.94	0.72	1452397	41.83	Archaea	Thermoplasmatota	Poseidoniia	Poseidoniales	Poseidoniaceae	MGIIa-L1
BIN14	51.8	0.72	1948599	49.65	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Porticoccaceae	TMED48
BIN140	50.36	5.04	3560248	51.44	Bacteria	Proteobacteria	Gammaproteobacteria	UBA10353	LS-SOB	
BIN15	54.68	0	2017812	44.21	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodospirillales_A	Magnetovibrionaceae	
BIN16	61.87	0.72	2558606	38.46	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Flavobacterium
BIN17	82.73	0.72	2729546	42.38	Bacteria	Verrucomicrobiota	Verrucomicrobiae	Opitutales	MB11C04	MB11C04
BIN18	76.26	0.72	2507088	45.76	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Nitrincolaceae	Neptunomonas
BIN19	82.73	0	3503641	65.39	Bacteria	Proteobacteria	Alphaproteobacteria	Oceanibaculales	Oceanibaculaceae	Oceanibaculum
BIN2	59.71	0	1102882	32.15	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	TMED25	
BIN20	57.55	2.16	2611026	55.97	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Octadecabacter
BIN21	89.93	1.44	1275699	40.18	Bacteria	Proteobacteria	Gammaproteobacteria	Nitrosococcales	Methylophagaceae	
BIN22	35.97	0.72	1348582	29.91	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	GCA-2733415
BIN23	44.6	0	1474564	36.92	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	GCA-002705045
BIN24	44.6	0	1693122	34.19	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	UBA7433
BIN25	50.36	1.44	2472219	65.75	Bacteria	Proteobacteria	Gammaproteobacteria	Burkholderiales	Rhodocyclaceae	Methyloversatilis
BIN26	95.68	0	1965454	37.35	Bacteria	Proteobacteria	Gammaproteobacteria	UBA4486	UBA4486	UBA7359
BIN27	87.77	0.72	937673	35.94	Bacteria	Proteobacteria	Gammaproteobacteria	UBA7916		
BIN28	47.48	1.44	4192622	38.96	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacterales	Alteromonadaceae	Pseudoalteromonas
BIN29	77.7	2.88	2251434	32.38	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	GCA-002722245	GCA-002722245
BIN3	92.81	1.44	2779914	36.33	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	1G12	UBA6770
BIN30	34.53	0	925484	28.94	Bacteria	Marinisomatota	Marinisomatia	SCGC-AAA003-L08	GCA-002707645	
BIN31	48.92	0	4077292	40.86	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacterales	Alteromonadaceae	Pseudoalteromonas

BIN32	89.21	2.16	2340707	36.24	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	GCA-002722245	
BIN33	62.59	1.44	3119750	41.85	Bacteria	Myxococcota	UBA796	UBA796	UBA796	
BIN34	52.52	0	1555877	32.57	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	UBA7428
BIN35	79.86	2.16	4908483	67.19	Bacteria	Proteobacteria	Gammaproteobacteria	Burkholderiales	Burkholderiaceae	Bordetella_B
BIN36	82.01	2.88	3674298	34.87	Bacteria	Bacteroidota	Bacteroidia	Cytophagales	Marinoscillaceae	
BIN37	92.81	0	2919799	64.82	Bacteria	Proteobacteria	Alphaproteobacteria	Sphingomonadales	Sphingomonadaceae	Citromicrobium
BIN38	44.6	6.47	1438927	27.19	Bacteria	Marinisomatota	Marinisomatia	SCGC-AAA003-L08	TMED6	TMED6
BIN39	38.13	7.91	1248959	44.84	Bacteria	Proteobacteria	Alphaproteobacteria	Puniceispirillales	Puniceispirillaceae	Puniceispirillum
BIN4	49.64	1.44	2956929	45.25	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacterales	Alteromonadaceae	Alteromonas
BIN40	41.73	0.72	892033	32.25	Bacteria	Proteobacteria	Gammaproteobacteria	Burkholderiales	Methylophilaceae	BACL14
BIN41	84.17	0.72	1567552	35.08	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	BACL11	
BIN42	82.01	0	1657135	38.31	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	UBA8316
BIN43	70.5	1.44	2833910	30.87	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Mesoflavibacter
BIN44	71.22	0	2200746	54.77	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Halieaceae	Luminiphilus
BIN45	44.6	0.72	1661154	35.99	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Hel1-33-131
BIN46	72.66	1.44	2337837	57.28	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Halomonadaceae	Halomonas
BIN47	46.76	1.44	652187	34.14	Bacteria	Proteobacteria	Gammaproteobacteria	Burkholderiales	Methylophilaceae	BACL14
BIN48	61.87	3.6	863451	29.35	Bacteria	Proteobacteria	Alphaproteobacteria	HIMB59	HIMB59	HIMB59
BIN49	82.73	2.16	3367393	50.52	Bacteria	Myxococcota	Polyangia	Polyangiales	Polyangiaceae	NIC37A-2
BIN5	59.71	7.91	2159413	49.26	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodospirillales_A	2-02-FULL-58-16_A	
BIN50	51.8	0	3256594	57.97	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Oleiphilaceae	Marinobacter
BIN51	53.96	0	1666934	49.30	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	UA16	UA16
BIN52	91.37	0	3164695	52.12	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	UBA12010
BIN54	50.36	2.16	1990912	52.39	Bacteria	Proteobacteria	Alphaproteobacteria	Rhodobacterales	Rhodobacteraceae	Ascidiaceihabitans
BIN55	78.42	3.6	3579732	57.68	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Oleiphilaceae	Marinobacter
BIN56	89.93	0.72	2717097	53.73	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Halieaceae	Luminiphilus
BIN57	92.09	0.72	2503512	36.16	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	koll-22	UBA1494_A

BIN58	89.21	0.72	3844626	57.25	Bacteria	Verrucomicrobiota	Verrucomicrobiae	Pedosphaerales	UBA1096	UBA1096
BIN59	41.01	0.72	1505708	33.72	Bacteria	Bacteroidota	Bacteroidia	Cytophagales	UBA9547	
BIN6	48.92	3.6	1976222	50.51	Bacteria	Verrucomicrobiota	Verrucomicrobiae	Opitutales	Opitutaceae	
BIN60	79.14	0	5195352	43.97	Bacteria	Proteobacteria	Gammaproteobacteria	Enterobacterales	Alteromonadaceae	Paraglaciecola
BIN61	93.53	1.44	1698277	29.33	Bacteria	Proteobacteria	Alphaproteobacteria	TMED109		
BIN62	41.73	3.6	2775477	53.60	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	HTCC2089	UBA9926
BIN63	49.64	0	472672	30.21	Bacteria	Proteobacteria	Alphaproteobacteria	HIMB59	HIMB59	HIMB59
BIN64	38.85	2.88	1542391	31.68	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Polaribacter
BIN65	38.13	2.88	1740758	44.49	Archaea	Thermoplasmatota	Poseidoniia	Poseidoniales	Poseidoniaceae	MGIIa-K2
BIN66	57.55	2.16	3313573	58.69	Bacteria	Actinobacteriota	Acidimicrobiia	Microtrichales	UBA10347	UBA10347
BIN67	38.85	1.44	1847207	34.88	Bacteria	Bacteroidota	Bacteroidia	Cytophagales		
BIN68	81.29	1.44	3501394	60.92	Bacteria	Proteobacteria	Alphaproteobacteria	Caulobacterales	Hyphomonadaceae	Hyphomonas
BIN69	97.12	0	2891023	58.47	Bacteria	Proteobacteria	Alphaproteobacteria	Micavibrionales	TMED2	TMED2
BIN7	88.49	2.16	2646838	30.83	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	UBA10066	GCA-2723085
BIN70	71.22	4.32	1151054	29.71	Bacteria	Proteobacteria	Alphaproteobacteria	Puniceispirillales	AAA536-G10	AAA536-G10
BIN71	72.66	0.72	2425483	41.33	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	Acinetobacter
BIN72	66.91	1.44	1648523	53.69	Bacteria	Cyanobacteria	Cyanobacteriia	Synechococcales	Cyanobiaceae	Synechococcus_C
BIN73	88.49	0	2638169	39.72	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	1G12	SHAN690
BIN74	44.6	4.32	902842	35.19	Bacteria	Proteobacteria	Alphaproteobacteria	Rhizobiales	TMED25	MED-G09
BIN75	53.96	0	1853715	36.29	Bacteria	Bacteroidota	Bacteroidia	Chitinophagales	Saprospiraceae	UBA1994
BIN77	61.87	10.79	770069	37.39	Bacteria	Proteobacteria	Gammaproteobacteria	SAR86	SAR86	
BIN78	91.37	1.44	2726173	34.51	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	1G12	UBA2108
BIN79	33.81	0.72	689624	29.12	Bacteria	Proteobacteria	Alphaproteobacteria	Puniceispirillales	AAA536-G10	
BIN8	53.96	0.72	2991545	39.89	Bacteria	Bacteroidota	Bacteroidia	Cytophagales	UBA9547	
BIN80	69.06	0.72	2336923	38.69	Bacteria	Fibrobacterota	Fibrobacteria	Fibrobacterales		
BIN81	51.8	0.72	716964	43.72	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Litoricolaceae	
BIN82	30.22	0	1344407	43.93	Archaea	Thermoplasmatota	Poseidoniia	Poseidoniales	Poseidoniaceae	MGIIa-K1

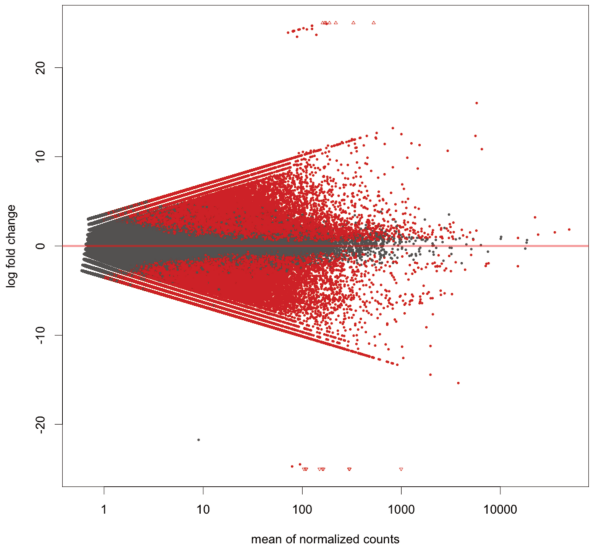
BIN83	74.82	0	2289014	61.98	Bacteria	Actinobacteriota	Acidimicrobiia	Microtrichales	Ilumatobacteraceae	Ilumatobacter
BIN84	48.92	0.72	1668240	56.50	Bacteria	Planctomycetota	Phycisphaerae	Phycisphaerales	SM1A02	UBA1668
BIN85	88.49	5.76	6017329	52.62	Bacteria	Planctomycetota	Planctomycetes	Planctomycetales	Planctomycetaceae	
BIN86	84.17	1.44	2967588	68.35	Bacteria	Proteobacteria	Alphaproteobacteria	Caulobacterales	Caulobacteraceae	Brevundimonas
BIN87	87.05	0	2144075	37.45	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	UA16	ASP10-05a
BIN88	48.92	3.6	3174035	29.15	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	Polaribacter
BIN89	54.68	6.47	893120	34.61	Bacteria	Proteobacteria	Gammaproteobacteria	Burkholderiales	Burkholderiaceae	UBA7377
BIN9	56.12	0.72	2476348	34.36	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	UBA10066	
BIN90	46.76	5.76	2433698	32.05	Bacteria	Marinisomatota	Marinisomatia	Marinisomatales	UBA1611	GCA-2691785
BIN91	59.71	0	1234218	37.05	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	Flavobacteriaceae	UBA8316
BIN92	33.81	0.72	913557	31.31	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	BACL11	BACL11
BIN93	58.99	0.72	2543783	60.32	Bacteria	Planctomycetota	Phycisphaerae	Phycisphaerales	SM1A02	UBA12014
BIN94	57.55	2.88	581506	28.45	Bacteria	Proteobacteria	Alphaproteobacteria	HIMB59	HIMB59	HIMB59
BIN95	50.36	2.16	736857	27.88	Bacteria	Proteobacteria	Alphaproteobacteria	Pelagibacterales	Pelagibacteraceae	AG-414-E02
BIN97	58.27	0	3464936	58.64	Bacteria	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Alcanivoracaceae	Alcanivorax
BIN98	78.42	2.16	3795348	53.97	Bacteria	Verrucomicrobiota	Verrucomicrobiae	Verrucomicrobiales	Akkermansiaceae	SW10
BIN99	69.06	0	2118379	47.61	Bacteria	Bacteroidota	Bacteroidia	Flavobacteriales	UA16	UA16

Supplementary Table 7. *DESeq2 results*

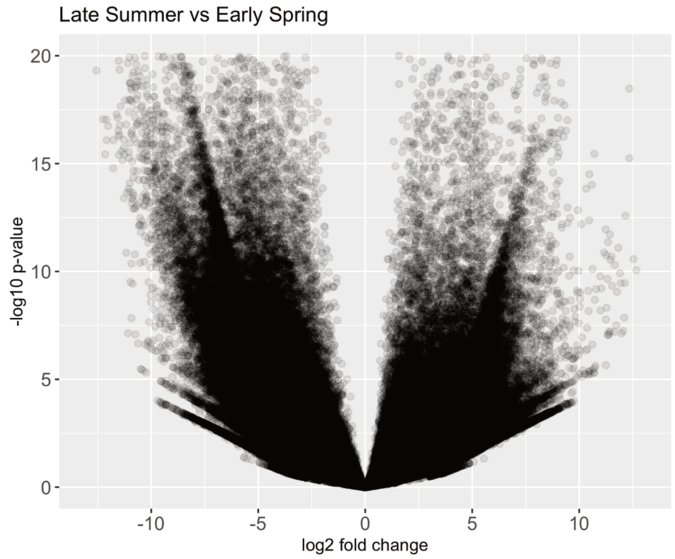
out of 539249 with nonzero total read count adjusted p-value < 0.05	
LFC > 0 (up):	64272, 12%
LFC < 0 (down):	111206, 21%
outliers:	142, 0.026%
low counts:	142, 0.026%


SEASON
Late summer
Early spring
METAPROTEOME
Presence
Absence
3
2
1
0
-1
K02064_thiamine transport system substrate-binding protein
K10559_rhamnose transport system substrate-binding protein
K10111_multiple sugar transport system ATP-binding protein
K10117_raffinose/stachyose/melibiose transport system substrate-binding protein
K02042_phosphonate transport system permease protein
K06515_solute carrier family 44 (choline transporter-like protein)
K16302_metal transporter
K07238_zinc transporter, ZIP family
K03320_ammonium transporter, Amt family
K16075_magnesium transporter
K06580_ammonium transporter Rh
K06902_MFS transporter, UMF1 family
K10823_oligopeptide transport system ATP-binding protein
K02041_phosphonate transport system ATP-binding protein
K19611_ferric enterobactin receptor
K18814_putative inorganic carbon (hco3(-)) transporter
K17329_N,N'-diacetylchitobiose transport system substrate-binding protein
K17325_glycerol transport system ATP-binding protein
K17244_putative chitobiose transport system substrate-binding protein
K17241_alpha-1,4-digalacturonate transport system substrate-binding protein
K17237_inositol-phosphate transport system substrate-binding protein
K17208_inositol transport system substrate-binding protein
K17205_putative xylitol transport system substrate-binding protein
K15727_membrane fusion protein, cobalt-zinc-cadmium efflux system
K15721_pesticin/yersiniabactin receptor
K15598_putative hydroxymethylpyrimidine transport system substrate-binding protein
K15587_nickel transport system ATP-binding protein
K15584_nickel transport system substrate-binding protein
K15577_nitrate/nitrite transport system permease protein
K15576_nitrate/nitrite transport system substrate-binding protein
K15554_sulfonate transport system permease protein
K15553_sulfonate transport system substrate-binding protein
K15497_molybdate/tungstate transport system ATP-binding protein
K13892_glutathione transport system ATP-binding protein
K13889_glutathione transport system substrate-binding protein
K12371_dipeptide transport system ATP-binding protein
K11962_urea transport system ATP-binding protein
K11957_neutral amino acid transport system ATP-binding protein
K11690_C4-dicarboxylate transporter, DctM subunit
K11635_putative ABC transport system ATP-binding protein
K11084_2-aminoethylphosphonate transport system ATP-binding protein
K11081_2-aminoethylphosphonate transport system substrate-binding protein
K11076_putrescine transport system ATP-binding protein
K11075_putrescine transport system permease protein
K11071_spermidine/putrescine transport system permease protein
K10561_rhamnose transport system permease protein
K10562_fructose transport system substrate-binding protein
K10540_methyl-galactoside transport system substrate-binding protein
K10235_alpha-glucoside transport system ATP-binding protein
K10232_alpha-glucoside transport system substrate-binding protein
K10195_oligogalacturonide transport system ATP-binding protein
K10191_lactose/L-arabinose transport system ATP-binding protein
K10188_lactose/L-arabinose transport system substrate-binding protein
K10039_aspartate/glutamate/glutamine transport system substrate-binding protein
K10037_glutamine transport system permease protein
K10036_glutamine transport system substrate-binding protein
K10022_arginine/ornithine transport system substrate-binding protein
K10014_histidine transport system substrate-binding protein
K10013_lysine/arginine/ornithine transport system substrate-binding protein
K10005_glutamate transport system substrate-binding protein
K09996_arginine transport system substrate-binding protein
K06076_long-chain fatty acid transport protein
K05816_sn-glycerol 3-phosphate transport system ATP-binding protein
K05776_molybdate transport system ATP-binding protein
K05685_macrolide transport system ATP-binding/permease protein
K03447_MFS transporter, OPA family, solute carrier family 37 (glycerol-3-phosphate transporter)
K02575_NNP family, nitrate/nitrite transporter
K02424_L-cystine transport system substrate-binding protein
K02062_thiamine transport system ATP-binding protein
K02028_polar amino acid transport system ATP-binding protein
K02017_molybdate transport system ATP-binding protein
K02021_putative ABC transport system ATP-binding protein
K02010_iron(III) transport system ATP-binding protein
K02026_multiple sugar transport system permease protein
K02056_simple sugar transport system ATP-binding protein
K02050_NitT/TauT family transport system permease protein
K11073_putrescine transport system substrate-binding protein
K16089_outer membrane receptor for ferrienterochelin and colicins
K02057_simple sugar transport system permease protein
K02031_peptide/nickel transport system ATP-binding protein
K02032_peptide/nickel transport system ATP-binding protein
K01991_polysaccharide biosynthesis/export protein
K15738_ABC transport system ATP-binding/permease protein
K18889_ATP-binding cassette, subfamily B, multidrug efflux pump
K15580_oligopeptide transport system substrate-binding protein
K02033_peptide/nickel transport system permease protein
K07027_glycosyltransferase 2 family protein
K11720_lipopolysaccharide export system permease protein
K17202_erythritol transport system substrate-binding protein
K02034_peptide/nickel transport system permease protein
K04759_ferrous iron transport protein B
K02004_putative ABC transport system permease protein
K03560_biopolymer transport protein TolR
K03585_membrane fusion protein, multidrug efflux system
K16091_Fe(3+) dicitrate transport protein
K18138_multidrug efflux pump
K18890_ATP-binding cassette, subfamily B, multidrug efflux pump
K02049_NitT/TauT family transport system ATP-binding protein
K02052_putative spermidine/putrescine transport system ATP-binding protein
K02027_multiple sugar transport system substrate-binding protein
K01990_ABC-2 type transport system ATP-binding protein
K07323_phospholipid transport system substrate-binding protein
K02003_putative ABC transport system ATP-binding protein
K16092_vitamin B12 transporter
K09013_Fe-S cluster assembly ATP-binding protein
K09972_general L-amino acid transport system ATP-binding protein
K03641_TolB protein
K02036_phosphate transport system ATP-binding protein
K02195_heme exporter protein C
K03562_biopolymer transport protein TolQ
K02012_iron(III) transport system substrate-binding protein
K02040_phosphate transport system substrate-binding protein
K17324_glycerol transport system ATP-binding protein
K01995_branched-chain amino acid transport system ATP-binding protein
K10441_ribose transport system ATP-binding protein
K05777_putative thiamine transport system substrate-binding protein
K06213_magnesium transporter
K02440_glycerol uptake facilitator protein
K02066_phospholipid/cholesterol/gamma-HCH transport system permease protein
K03498_trk system potassium uptake protein
K02556_chemotaxis protein MotA
K09969_general L-amino acid transport system substrate-binding protein
K02557_chemotaxis protein MotB
K09690_lipopolysaccharide transport system permease protein
K17315_glucose/mannose transport system substrate-binding protein
K02806_nitrogen regulatory IIA component
K03499_trk system potassium uptake protein
K02013_iron complex transport system ATP-binding protein
K08351_biotin/methionine sulfoxide reductase
K17322_glycerol transport system permease protein
K10440_ribose transport system permease protein
K11959_urea transport system substrate-binding protein
K07793_putative tricarboxylic transport membrane protein
K10018_octopine/nopaline transport system substrate-binding protein
K03523_biotin transport system substrate-specific component
K06901_putative MFS transporter, AGZA family, xanthine/uracil permease
K10439_ribose transport system substrate-binding protein
K10001_glutamate/aspartate transport system substrate-binding protein
K01999_branched-chain amino acid transport system substrate-binding protein
K17321_glycerol transport system substrate-binding protein
K02000_glycine betaine/proline transport system ATP-binding protein
K06189_magnesium and cobalt transporter
K05813_sn-glycerol 3-phosphate transport system substrate-binding protein
K02001_glycine betaine/proline transport system permease protein
K02055_putative spermidine/putrescine transport system substrate-binding protein
K15551_taurine transport system substrate-binding protein
K02002_glycine betaine/proline transport system substrate-binding protein
K07795_putative tricarboxylic transport membrane protein
K02058_simple sugar transport system substrate-binding protein
K02067_phospholipid/cholesterol/gamma-HCH transport system substrate-binding protein
K03286_OmpA-OmpF porin, OOP family
K02193_heme exporter protein A
K02054_putative spermidine/putrescine transport system permease protein
K10651_taurine transport system ATP-binding protein
K01992_ABC-2 type transport system permease protein
K09970_general L-amino acid transport system permease protein
K17938_peptide/bleomycin uptake transporter
K09971_general L-amino acid transport system permease protein
K15552_taurine transport system permease protein
K01996_branched-chain amino acid transport system ATP-binding protein
K01997_branched-chain amino acid transport system permease protein
K01998_branched-chain amino acid transport system permease protein
K02053_putative spermidine/putrescine transport system permease protein
K10020_octopine/nopaline transport system permease protein
K03321_sulfate permease, SulP family
K10112_multiple sugar transport system ATP-binding protein
K03543_membrane fusion protein, multidrug efflux system
K07221_phosphate-selective porin OprO and OprP
K11604_manganese/iron transport system substrate-binding protein
K15583_oligopeptide transport system ATP-binding protein
K03561_biopolymer transport protein ExbB
K02014_iron complex outermembrane recepter protein
K07091_lipopolysaccharide export system permease protein
K10543_D-xylose transport system substrate-binding protein
K21572_starch-binding outer membrane protein, SusD/RagB family
K02168_choline/glycine/proline betaine transport protein
K10544_D-xylose transport system permease protein
K02044_phosphonate transport system substrate-binding protein
K03832_periplasmic protein TonB
K17213_inositol transport system substrate-binding protein
K10554_fructose transport system ATP-binding protein
K17214_inositol transport system permease protein
K02011_iron(III) transport system permease protein
K09774_lipopolysaccharide export system protein LptA
K03559_biopolymer transport protein ExbD
K09691_lipopolysaccharide transport system ATP-binding protein
K16087_hemoglobin/transferrin/lactoferrin receptor protein
K08720_outer membrane protein OmpU
K17713_outer membrane protein assembly factor BamB
K10227_sorbitol/mannitol transport system substrate-binding protein
K11072_spermidine/putrescine transport system ATP-binding protein
K02065_phospholipid/cholesterol/gamma-HCH transport system ATP-binding protein
K07277_outer membrane protein insertion porin family
K02030_polar amino acid transport system substrate-binding protein
K06861_lipopolysaccharide export system ATP-binding protein
K11688_C4-dicarboxylate-binding protein DctP
K12368_dipeptide transport system substrate-binding protein
K02025_multiple sugar transport system permease protein
K02029_polar amino acid transport system permease protein
K11069_spermidine/putrescine transport system substrate-binding protein
K02035_peptide/nickel transport system substrate-binding protein
K02051_NitT/TauT family transport system substrate-binding protein
KERFIX_2_II
KERFIX_2_I
KERFIX_I
KERFIX_II
A2_3_II
A3_3_II
A2_2_I
A2_I
A3_I
KERFIX_I
KERFIX_II
A3_2_I
A3_2_II
A3_I
KERFIX_3
KERFIX_1
KERFIX_2
A3_1
A3_2
A3_3
KERFIX_1
KERFIX_2
A2_1
A2_2
A2_3
METAPROTEOMES
METATRANSCRIPTOMES

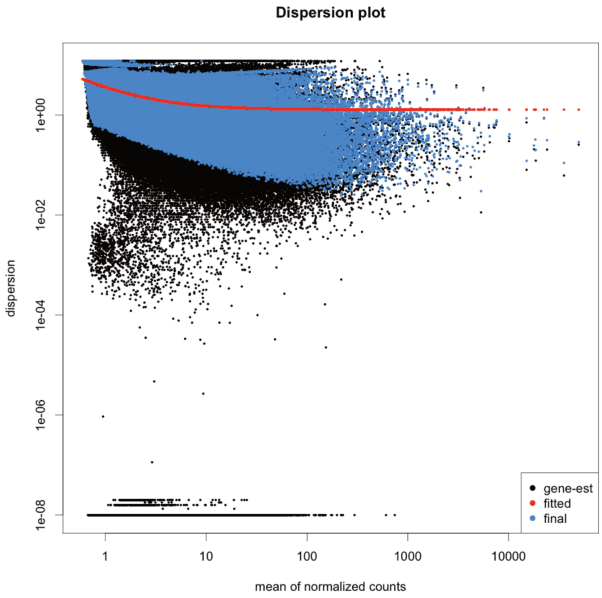
***Supplementary Figure 1**. Heatmap of all KEGG transporter proteins. All metatranscriptomes are shown in triplicates. Z-scaling of normalized transcript counts by rows and Euclidian clustering by row and column, blue denotes absence in metatranscritome. All metaproteomes are shown by their presence in one duplicate.*



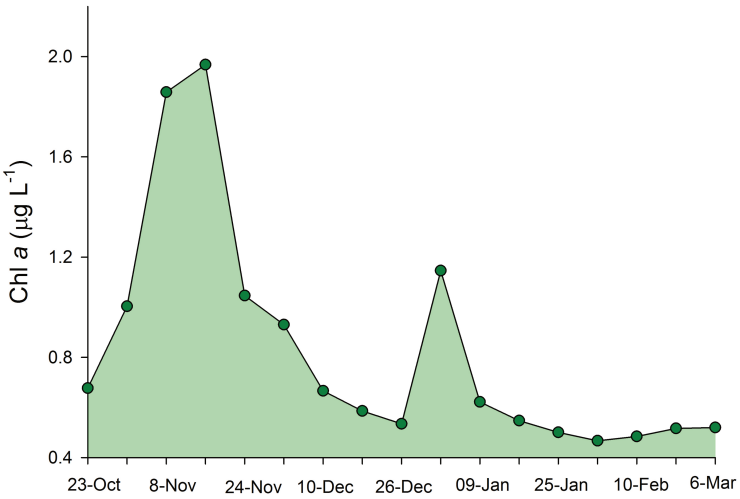
***Supplementary Figure 2.** MA plot for DESeq statistic*



***Supplementary Figure 3.** Volcano plot for comparison of late summer versus early spring*



Supplementary Figure 4. *Dispersion plot for DESeq2 statistic.*



Supplementary Figure 5. *Chlorophyll a data from remote automated sampler at station A3 at 30m depth for a period of 5 months in 2016/17 installed during the SOCLIM cruise.*

Conclusion

This thesis contributes to the understanding of gene expression of microbial communities in contrasting Southern Ocean environments and the potential implications for Fe and C cycling. The following conclusion is intended as a synthesis of the scientific outcome of all three chapters ending with perspectives for the research field.

Taxa specific patterns in heterotrophic Fe metabolism

The analysis on 16S rRNA level for microbial communities in the Kerguelen area from previous studies depict differences in relative abundance of taxa at naturally Fe fertilized and Fe limited sites. During austral summer months (KEOPS 1 cruise, January-February 2005), clone library analysis of the Fe fertilized site in the phytoplankton bloom period revealed dominant operational taxonomic units (OTUs) as the *Roseobacter*, *SAR92* and FCB (*Flavobacteria- Cytophaga-Bacteroides*), while in the HNLC region, *SAR11*, *Roseobacter* and *Polaribacter* dominated (West *et al.*, 2008). These difference are less pronounced in spring (KEOPS 2, November 2011), the relative abundance derived from 454 pyrosequencing of *SAR11* reached up to 25 - 50% at all visited sites followed by *Gammaproteobacteria* at 5 – 15%, FCB at 5 – 40% and at one station up to 30% *Actinomycetales* (Landa *et al.*, 2016). When comparing to metatranscriptomic data for the KEOPS 2 cruise, the taxa contributing to specific Fe metabolism are different (Chapter 1). Siderophore-uptake was dominated by *Gammaproteobacteria* (43% of total pathway specific transcripts) and FCB (25%), while Fe^{2+} - and Fe^{3+} -uptake revealed an increased contribution of alphaproteobacterial groups, most pronounced for Fe^{3+} -uptake (29%). FCB were mostly represented by *Flavobacteriaceae* for siderophore- and Fe^{2+}-uptake, and *Bacillaceae* contributed additionally to Fe^{3+}-uptake. Within *Alphaproteobacteria*, *Rhodobacteracea* substantially contributed to Fe^{2+}- and Fe^{3+}-uptake, but this group had low siderophore uptake transcripts (<1.6% of alphaproteobacterial transcripts).

Regarding prokaryotic communities incubated with particles and sea-water from the Southern Ocean, both particle-attached and free-living were dominated by the three same groups in 16S contribution (Chapter 2): *Flavobacteriales* (25% PA, 35% FL), *Rhodobacterales* (20% PA, 27% FL), *Oceanospirillales* (26% PA, 13% FL). While the non-particle control was dominated by *Flavobacteriales* up to 28%, *SAR*11 up to 20% and *Cellvibrionales* up to 14%. Regarding Fe uptake transcripts for siderophore synthesis were associated to a diverse prokaryotic community of low contribution in 16S rRNA such as for instance *Burkholderiales*

and *Corynebacteriales* while TonB receptors for the respective siderophores were again expressed by other representatives such as SAR86 and SAR92. *Rhodobacteracea* transcripts were again expressed in highest contribution for Fe^{3+} uptake mediated through ABC transporters (Chapter 1 & 2).

When observing the metagenomic co-assembly of sites around the Kerguelen islands and the derived assembled genomes the taxa proportions are close to the previously observed 16S rRNA patterns (Chapter 3). Of 133 MAGs, 4 belonged to Archaea (3.0% of MAGs), 39 to the *Bacteroidota* (29.3%, new Name of FCB cluster by GTDB), 34 to *Alphaproteobacteria* (25.6%) and 27 to *Gammaproteobacteria* (27.8%). Metatranscriptomic data revealed differential expression of Fe transport related genes in MAGs of *Gammaproteobacteria* and *Bacteroidota* and several of these were recovered in metaproteomes as well (Chapter 3, Fig 6&8).

Linking sources of Fe and C with heterotrophic metabolism

Samples from Chapter 1 & 3 include measurements or derived data for dissolved Fe concentrations at open ocean stations and sequencing data is put in context to those. Chapter 2 focuses on different views on dissolved and particulate Fe and the impact on microbial communities from an open ocean HNLC site. The different sources show different processes and taxa at work in terms of gene expression and protein translation.

In Chapter 1 the composition of the prokaryotic communities contributing to the transcripts of a given Fe-related pathway was overall independent of the in-*situ* Fe supply, indicating that microbial taxa utilize distinct Fe-related metabolic processes. Only a few prokaryotic groups contributed to the transcripts of more than one Fe uptake mechanism, suggesting limited metabolic versatility.

The potential of particulate Fe for heterotrophic metabolism remains underexplored. In the particle experiment prokaryotic production increases when only particles are added (Chapter 2). While it cannot be confirmed that this is due to the Fe present in the particles, transcripts for siderophore synthesis and Fe transporters are differentially expressed compared to the control. However, marine particles are a source of complex polymeric substances which can additionally to trace metals induce prokaryotic production increases.

In a multi-omics approach (Chapter 3) transporters for multiple C sources as well as Fe and other trace metals are shown to have different seasonal expression for site A3 and KERFIX. The seasonal difference can be linked to different supplies of Fe and C, while for C it is difficult to assess as DOC is similar for stations and seasons. With this approach it is possible to link the expression of a protein to its expressed gene from a specific species and thus it is possible to

identify who is expressing what, in which season, and thus what sort of organic matter was present.

From gene expression to proteins

The advancement of 'omics' techniques has brought forth an incredible amount of sequence information. Most of the recent oceanic studies have however focused on the genetic potential through metagenomes from the surface to the deep (Sunagawa *et al.*, 2015; Carradec *et al.*, 2018; Acinas *et al.*, 2019). While quantifying and describing metabolic processes is still very challenging, omics techniques can be combined to further describe actual activity of microorganism at a given timepoint (Chapter 3).

Attempts to link sequence data to biogeochemical results have yet to advance the field of microbial oceanography as linking biogeochemistry and omics remains challenging (Caputi *et al.*, 2019). The gap between the scales of microbial and biogeochemical sciences has to be crossed in order to provide quantitative insights into the functioning of biogeochemical cycles based on these data (Schimel, 2016). In order to do so, the flow of biological information from DNA through proteins needs to be converted into actual process rates but at each step, the linkage can be wildly non-linear. One gene can code for thousands of copies of an enzyme or none at all and enzymes may also be present but inactive. It therefore remains unclear how to scale microbial data to biogeochemical process rates or whether it is even possible to do so (Bier *et al.*, 2015). Additionally, there is a mathematical disconnect between the scales: whereas mathematical models of biogeochemical processes rarely have more than twenty separate flows between different chemical forms (for example, DOC in the ocean to microbial biomass), high throughput sequencing identifies millions of individual DNA sequences. How is it possible to link a million independent drivers to a few dependent response variables? These are the challenges facing the attempt to link populations to functions and processes and developing a microbially based picture of system level biogeochemistry. Recent statements and calls of scientist are urging for an integration of microorganisms in to global scale models especially in times of climate change (Baltar *et al.*, 2019; Cavicchioli *et al.*, 2019). Attempts have been taken using gene centric approaches basing the dependence of microbial growth on substrate and nutrient availability through Michaelis–Menten kinetics (Reed *et al.*, 2014) or using simulation models and observing resulting biogeochemical outcomes (Coles *et al.*, 2017). Despite their revolutionary attempts the predictive strength of these remains limited.

Additionally, the increasing amount of sequences in public databases points to fundamental questions that remain unanswered which can be observed through the recent

debates on diversity of prokaryotes and whether conclusions on it can actually be drawn from 16S rRNA data (Martiny, 2019; Steen *et al.*, 2019).

While metagenome assembled genomes (MAGs) provide a new possibility to recreate environmental species with high genomic completeness the rapidly increasing number of MAGs in public databases already competes with the total number of microbial isolate genomes (Bowers *et al.*, 2017). Despite their growing availability, metagenomes are inherently complex and demand researchers to arrange a complex combination of rapidly evolving computational tools and approaches with many alternatives to reconstruct, characterize, and finalize MAGs (Shaiber and Eren, 2019). However, the amount of resolved MAGs from environmental samples does not necessarily increase our knowledge of prokaryotic diversity (commentary in Appendices). Furthermore, the quality of MAGs and the accuracy of their genomic content needs to be of highest precision to truly benefit public databases (Shaiber and Eren, 2019).

Perspectives

Following aspects are yet to be implemented and discussed in the context of methods and concepts of this thesis:

- The reason why Fe concentrations in the ocean only exist since the 1980 (Boyd and Ellwood, 2010) is due to the challenge of obtaining non-contaminated values using trace-metal clean sampling method. While the GEOTRACES framework (international study of the global marine biogeochemical cycles of trace elements and their isotopes, 2007) has tremendously increased our knowledge on trace metals in the ocean, recent developments might allow for trace-metal clean seawater sampling in open-ocean time-series applications (Merwe *et al.*, 2019).
- The results stem from sampling of several liters of seawater to study microorganisms at a given depth and timepoint, such as in other sequencing projects (Yooseph *et al.*, 2007; Duarte, 2015; Sunagawa *et al.*, 2015). The analyzed metabolic processes are occurring on the smallest, so-called Kolomogorov scale or microscale. The observed results however shed light on biogeochemical cycling of nutrients in marine systems happening on larger, even global scales as microorganism are responsible for recycling of crucial elements and reintroducing them into the marine food-web. The discrepancy in discussing this when focusing on ‘global ocean’ projects is problematic. Inspiring works of ecology from the 1960s and 1970s were holistic (Wyatt, 2014 and references therein).

Amongst these inspirations, Margalef's mandala holds a unique place in ecology and in terms of biogeochemical cycles one could raise the question: At what point does a cycle begin and at what point does it end? On a cellular level, cycles are initiated through the uptake of the element in question, and conclude when that said element leaves the cell. However, across larger spatial and temporal scales, this process is only the first of several steps. Once the element of interest leaves the cell, it merely starts the next phase of the cycle. Each phase within a biogeochemical cycle - be it at the cellular level or global scale - is subjected to a plethora of biological, chemical and physical processes. The manner in which these processes interact with each other is key to understanding the finer details of biogeochemical cycling.

Sequencing efforts are however, a snapshot of a very specific water mass at a specific timepoint. Time-series installations and remote automated samplers are providing a more holistic view of changes in marine microbial communities. Coupled with the above-mentioned automated sampler for trace metals, these can significantly advance our understanding on spatial and temporal scales of ocean element cycling.

- In order to derive the importance from sequencing data experiments comprising the whole microbial community need to be thought of. One possibility are chemostats incubating multiple pure culture strains or natural community assemblages under controlled conditions. However, in the case of Fe, trace metal clean sampling and incubation would need to be possible resulting in a great methodological challenge.
- An interesting result from in Chapter 2 is the expression of siderophore synthesis genes by several members of the community and the expression the receptors by other groups. One possibility is the temporal decoupling of these mechanisms in the cell which is not accounted for as the sample is filtered immediately and RNA sequencing only provides a metabolic fingerprint. Another point that is increasingly of importance with the accumulating information on membrane transporters is the question of membrane space. How much space can a prokaryotic membrane provide to fulfill multiple uptake and regulation processes and can all of these be upregulated at the same time? In order to attempt at answering these questions we will need to go back to culture work and design experiments with marine strains that are known siderophore producers. Approaches such as geneFISH might be the most promising at identifying intracellular upregulation coupled with cell activity.
- The current analysis of sequence data relies on alignment with curated or newly created databases. While functional information can be gained from databases targeting a

specific process (Chapter 1) a large proportion remains unclassified. In metagenomic studies up to 70% of raw reads cannot be annotated (Parks *et al.*, 2017).

This approach could be applied on the metagenomes (Chapter 3) using the CAZy database that describes the families of structurally-related catalytic and carbohydrate-binding modules of enzymes that degrade, modify, or create glycoside bonds (Lombard *et al.*, 2014). While the unaccounted part of sequences cannot be integrated with these approaches, the identifiable proportion can be discussed.

- Compared to classical biogeochemical analysis at oceanographic cruises (such as the collection of seawater in vials and the addition of fixatives for subsequent analysis), the analysis of sequence data derived files is methodologically more challenging.

  Measuring a concentration of a solute in seawater is nowadays performed using similar if not identical instruments and protocols (i.e. NH_4^+, DOC). The obtained result from the instrument is recorded and ready to use for interpretation with other parameters of a cruise, a sampling point or a time-series. The result is less subject of debate for the methodology used. With sequence data, the result is only one path taken in a series of complex decisions influencing the outcome. Such as the old Chinese saying claims: 'If you have 4 hours to chop a tree spend 3 hours sharpening your axe', an increasing time is spent on the processing of sequences into results. Even though recent years have resulted in an increase in ready-to-use pipelines, caution needs to be taken as the workflow heavily depends on the research question. More and more scientist applying bioinformatics urge for a consensus in analysis and transparency in the methodology used which will enhance the practice of data sharing.

References

Acinas, S.G., Sánchez, P., Salazar, G., Cornejo-Castillo, F.M., Sebastián, M., Logares, R., et al. (2019) Metabolic Architecture of the Deep Ocean Microbiome. *bioRxiv*.

Baltar, F., Bayer, B., Bednarsek, N., Deppeler, S., Escribano, R., Gonzalez, C.E., et al. (2019) Towards Integrating Evolution, Metabolism, and Climate Change Studies of Marine Ecosystems. *Trends in Ecology & Evolution*.

Bier, R.L., Bernhardt, E.S., Boot, C.M., Graham, E.B., Hall, E.K., Lennon, J.T., et al. (2015) Linking microbial community structure and microbial processes: an empirical and conceptual overview. *FEMS Microbiol Ecol* **91**:.

Bowers, R.M., Kyrpides, N.C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T.B.K., et al. (2017) Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nature Biotechnology* **35**: 725–731.

Boyd, P.W. and Ellwood, M.J. (2010) The biogeochemical cycle of iron in the ocean. *Nature Geoscience* **3**: 675–682.

Caputi, L., Carradec, Q., Eveillard, D., Kirilovsky, A., Pelletier, E., Karlusich, J.J.P., et al. (2019) Community-Level Responses to Iron Availability in Open Ocean Plankton Ecosystems. *Global Biogeochemical Cycles*.

Carradec, Q., Pelletier, E., Da Silva, C., Alberti, A., Seeleuthner, Y., Blanc-Mathieu, R., et al. (2018) A global ocean atlas of eukaryotic genes. *Nature Communications* **9**: 373.

Cavicchioli, R., Ripple, W.J., Timmis, K.N., Azam, F., Bakken, L.R., Baylis, M., et al. (2019) Scientists' warning to humanity: microorganisms and climate change. *Nature Reviews Microbiology* **17**: 569–586.

Coles, V.J., Stukel, M.R., Brooks, M.T., Burd, A., Crump, B.C., Moran, M.A., et al. (2017) Ocean biogeochemistry modeled with emergent trait-based genomics. *Science* **358**: 1149–1154.

Duarte, C.M. (2015) Seafaring in the 21St Century: The Malaspina 2010 Circumnavigation Expedition. *Limnology and Oceanography Bulletin*.

GEOTRACES – An international study of the global marine biogeochemical cycles of trace elements and their isotopes (2007) *Geochemistry* **67**: 85–131.

Landa, M., Blain, S., Christaki, U., Monchy, S., and Obernosterer, I. (2016) Shifts in bacterial community composition associated with increased carbon cycling in a mosaic of phytoplankton blooms. *The ISME journal* **10**: 39–50.

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* **42**: D490-495.

Martiny, A.C. (2019) High proportions of bacteria are culturable across major biomes. *ISME J* **13**: 2125–2128.

Merwe, P. van der, Trull, T.W., Goodwin, T., Jansen, P., and Bowie, A. (2019) The autonomous clean environmental (ACE) sampler: A trace-metal clean seawater sampler suitable for open-ocean time-series applications. *Limnology and Oceanography: Methods*.

Parks, D.H., Rinke, C., Chuvochina, M., Chaumeil, P.-A., Woodcroft, B.J., Evans, P.N., et al. (2017) Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nature Microbiology* **2**: 1533–1542.

Reed, D.C., Algar, C.K., Huber, J.A., and Dick, G.J. (2014) Gene-centric approach to integrating environmental genomics and biogeochemical models. *Proc Natl Acad Sci U S A* **111**: 1879–1884.

Schimel, J. (2016) Microbial ecology: Linking omics to biogeochemistry. *Nature Microbiology* **1**: 15028.

Shaiber, A. and Eren, A.M. (2019) Composite Metagenome-Assembled Genomes Reduce the Quality of Public Genome Repositories. *mBio* **10**: e00725-19.

Steen, A.D., Crits-Christoph, A., Carini, P., DeAngelis, K.M., Fierer, N., Lloyd, K.G., and Thrash, J.C. (2019) High proportions of bacteria and archaea across most biomes remain uncultured. *ISME J* 1–5.

Sunagawa, S., Coelho, L.P., Chaffron, S., Kultima, J.R., Labadie, K., Salazar, G., et al. (2015) Structure and function of the global ocean microbiome. *Science* **348**: 1261359.

West, N.J., Obernosterer, I., Zemb, O., and Lebaron, P. (2008) Major differences of bacterial diversity and activity inside and outside of a natural iron-fertilized phytoplankton bloom in the Southern Ocean. *Environmental Microbiology* **10**: 738–756.

Wyatt, T. (2014) Margalef's mandala and phytoplankton bloom strategies. *Deep Sea Research Part II: Topical Studies in Oceanography* **101**: 32–49.

Yooseph, S., Sutton, G., Rusch, D.B., Halpern, A.L., Williamson, S.J., Remington, K., et al. (2007) The Sorcerer II Global Ocean Sampling Expedition: Expanding the Universe of Protein Families. *PLoS Biology* **5**: e16.

Appendices

Prokaryotic transcription of transport genes elevated in low iron Southern Ocean as revealed through global metatranscriptomic approach

Pavla Debeljak[1,2], Eve Toulza[3], Sara Beier[4], Stephane Blain[1] and Ingrid Obernosterer[1]

[1] Sorbonne Université, CNRS, Laboratoire d'Océanographie Microbienne, LOMIC, F-66650 Banyuls/mer, France

[2] Department of Limnology and Bio-Oceanography, University of Vienna, A-1090 Vienna, Austria

[3] Université Perpignan Via Domitia, IHPE UMR 5244, CNRS, IFREMER, Univ. Montpellier, F-66860 Perpignan, France.

[4] Leibniz Institute for Baltic Sea Research, Warnemünde, Germany

Initially a second analysis was conducted with the data from Chapter 1 as a global approach consisting of an assembled metatranscriptome of the three sites. Functional annotation as well as differential expression analyses of prokaryotic gene transcripts were performed focusing on differences from iron-fertilized to the HNLC site in terms of gene ontology. The following is a short summary of the methodological approach and preliminary results.

While metagenomics is able to identify which microorganisms are present in an environment and what genomic potential they have, metatranscriptomics studies the function and activity of the complete set of transcripts (obtained from RNA-sequencing). In this assembly based metatranscriptomic approach we tried to highlight functional processes that were of highest importance in contrasting environmental sites of the Southern Ocean.

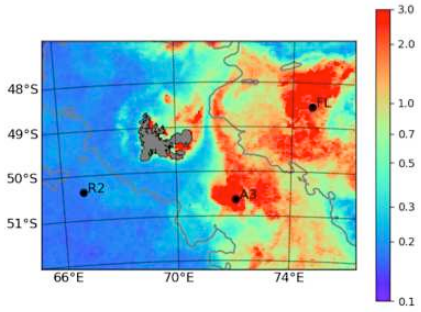


Figure 1 *Location of sampling sites superimposed on the monthly composite satellite image of chlorophyll (µg L^{-1}) provided by Copernicus Marine Service for November 2011, 4 x 4 km. The grey line denotes 1000km isobaths.*

Materials and Methods

Sample collection

Seawater samples were collected during the KEOPS2 cruise (Kerguelen Ocean and Plateau Compared Study 2, Oct. 8 – Nov. 30 2011) in the Indian sector of the Southern Ocean. Sampling and work procedures conducted in the study area were approved by the French Polar Institute (Institut polaire français Paul-Emile Victor) aboard the *RV* Marion Dufresne. Three stations were chosen for the present study: Station R-2 was located in HNLC (high-nutrient low chlorophyll) waters and stations A3-2 and F-L were located in naturally iron-fertilized waters (Figure 1). Seawater samples were collected with 12 L^{-1} Niskin bottles mounted on a rosette equipped with a CTDO Seabird SBE911-plus. For nucleic acid extractions, seawater was sampled at one depth in the surface mixed layer, and the chemical and biological parameters were collected throughout the water column (Christaki *et al.*, 2014; Lasbleiz *et al.*, 2014). For all environmental parameters, mean values from several measurements were calculated from surface deployments (Table 1).

Table 1. *Environmental parameters for the three stations*

	R-2	**F-L**	**A3-2**
Latitude S	50.3590	48.5222	51.0333
Longitude E	66.7170	74.6500	72.0833
Date of sampling	26/10/11	7/11/11	16/11/11
Sampling depth (m)	60	20	20
Surface mixed layer (m)	105±15	38±7	153±15
Dissolved and particulate nutrients			
DOC (µM)[1]	47.8 ± 0.4	49.6± 1.3	51.3 ± 1.5
DON (µM)[2]	6.1 ± 0.04	5.47 ± 1.33	6.44 ± 2.2
DOP (µM)[2]	0.3 ± 0.02	0.26 ± 0.12	0.36 ± 0.04
POC (µM)[1]	6.5 ± 1.8	11.5±1.2	13.5 ± 1.8

DFe (nmol L^{-1}) [3]	0.13 ± 0.05	0.22 ± 0.06	0.16 ± 0.03
Chlorophyll a (µg L^{-1}) [4]	0.25 ± 0.08	4.00 ± 1.58	2.03 ± 0.34
Prokaryotic abundance (x10^5 cells mL^{-1}) [5]	2.72±0.3	6.06*	3.16±0.5
Prokaryotic production (ng C L^{-1} h^{-1}) [5]	2.59±0.53	65.7±1.62	19.9±3.4
Prokaryotic respiration (µmol O_2 L^{-1} d^{-1}) [5]	0.25±0.12	1.37±0.64	0.63±0.45

* Only 1 measurement for the surface mixed layer available

[1]From Tremblay et al. (2015)

[2] From Blain et al. (2015)

[3] From Queroue et al. (2016)

[4] From Lasbleiz et al. (2014)

[5] From Christaki et al. (2014)

RNA extraction

For RNA extractions, volumes varying between 15 l and 30 l of pre-filtered water (200 µm nylon screen and 5 µm polycarbonate isopore filters) were collected onto 0.2 µm SuporPlus Membranes using a 142 mm filtration system (geotech equipment) and a peristaltic pump. The filtration procedure did not exceed 10 min and 10 ml of RNA-later was added before storage at −80°C. All nucleic acid extractions were performed in triplicates by cutting the filter in three parts. Total prokaryotic and eukaryotic RNA was extracted using the NucleoSpin® RNA Midi kit (Macherey-Nagel, Düren, Germany). Filters stored in RNA later were defrosted, removed from the RNA later solution, refrozen in liquid nitrogen and shattered using a mortar. The obtained 'powder-like' filter-pieces were added together with low binding zirconium beads (OPS Diagnostics, Lebanon, NJ,

USA) to the denaturing lysis buffer supplied by the NucleoSpin® RNA Midi kit and cells were disrupted by vortexing for 2 min. Beads were discarded by centrifugation. The extraction with the NucleoSpin® RNA Midi kit include an on-column DNA digestion step. However, in order to

ensure the absence of DNA in the sample, a control PCR reaction was performed without the retrotranscription (RT) step. Samples with DNA contamination, as indicated by amplification products were treated with a second DNA digestion step using the Turbo DNA-free kit (Ambion Life Technologies, Carlsbad, CA, USA). This additional DNAse treatment was followed by purification with the RNeasy MinElute Clean Up kit (Qiagen, Hilden, Germany). The extracted

RNA was quantified with the Agilent 2100 Bioanalyzer/Agilent RNA 6000 Nano Kit (Agilent, Santa Clara, CA, USA) and duplicates were chosen for sequencing.

Library preparation and sequencing

Prior to sequencing, ribosomal RNA was treated enzymatically with the RiboZero rRNA stranded RNA protocol to ensure sequencing of primarily messenger RNA followed by cDNA library construction using Illumina TruSeq Stranded mRNA Library Prep kit (Fasteris SA). Libraries were sequenced using paired-end 2 × 125 read length on one Illumina HiSeq 2500 lane.

Bioinformatic analysis

The raw Illumina reads were checked with FastQC (Andrews 2010) and adapters were eliminated using Cutadapt (Martin, 2011). Remaining ribosomal RNA sequences were removed by the riboPicker (Schmieder *et al.*, 2012) tool and sequences were checked by interlacing and de-interlacing paired-end reads ensuring that the same sequences were removed from each R1 and R2 files (performed in Galaxy, Afgan *et al.*, 2016). These sequence data have been submitted to the EMBL databases under accession number PRJEB30315 (Table 2).

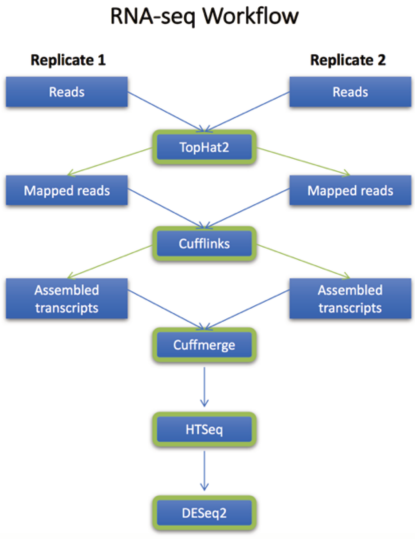


Figure 2. *Simplified bioinformatic pipeline*

Global Metatranscriptome approach

The global Metatranscriptome of pair-end reads from 3 stations in duplicates (n=6) was assembled with Metavelvet (Namiki *et al.*, 2012; Afiahayati *et al.*, 2015). Optimal kmer length was tested using 43, 47, 51, 53, 57, 59 and 61. Based on satisfactory N50 of 321bp with a minimum contig length of 500bp, kmer length of 61 was chosen resulting in 1,294,198 contigs. Each file (n=6) of quality checked short reads was mapped to the Metatranscriptome (read-recruitment) using the short reads aligner Bowtie2 (Langmead and Salzberg, 2012). The assembled contigs were then run through cufflinks and cuffmerge using default parameters (Trapnell *et al.*, 2010, 2013) to integrate reference transcripts into the assembly.

Table 2. *General information on sequencing results and reads.*

Sample Names	Raw Reads from paired-end sequencing (fastq replicates)	Interlaced reads after trimming and rRNA removal	Percent of prokaryotic reads from BLAST (%)	Number of prokaryotic reads from total reads
R_1	25 795 611 25 795 611	31 872 360	8.9	2 860 032
R_2	24 750 021 24 750 021	31 519 356	10.5	3 308 037
A3_2_1*	23 871 679 23 871 679	29 824 202	15.3	5 514 089
A3_2_2*	21 853 684 21 853 684	34 620 685	13.8	4 438 123
FL_1*	20 460 960 20 460 960	36 070 585	18.0	5 373 036
FL_2*	23 556 644 23 556 644	32 099 894	21.7	7 504 148

* The higher number of prokaryotic reads and their higher relative contributions to total reads at A3-2 and F-L as compared to R-2 can be explained by the elevated prokaryotic cell abundances at these sites (Table 1) and differences in the phytoplankton community composition. While larger diatom cells, abundant at the 2 bloom sites, were retained by the 5 µm filter used for the pre-filtration step, phytoplankton biomass at station R-2 was dominated by picoeukaryotes (Lasbleiz et al. 2016) that pass this pore size resulting in a higher number of eukaryote-assigned reads.

Cufflinks was used to assemble the transcripts and cuffmerge was used to produce a GTF file that contains an assembly that merges together the input assemblies creating one metatranscriptome of recruited reads. Obtained mapping regions (sam files converted into bam files) were preprocessed for differential expression analysis by counting the overlap of reads with merged transcripts from cuffmerge using htseq-counts (Anders *et al.*, 2015). Differential expression analysis (DESeq2; Love *et al.*, 2014) was calculated from the count files using the HNLC site (R-2) as reference to the Fe fertilized sites (A3-2 and F-L). The resulting log fold changes (LFC) corresponding to over (positive LFC) or under-presence (negative LFC) of genes at the fertilized sites compared to the reference (R-2) (Table 3).

Table 3. *DESeq2 results for station comparisons.*

Station	adjusted p-value	LFC	Number of Sequences
A3-2 vs R-2	0.05	Positive	84,746
	0.05	Negative	47,174
F-L vs R-2	0.05	Positive	82,856
	0.05	Negative	52,097
A3-2 vs F-L	0.05	Positive	71,529
	0.05	Negative	38,134

In a first attempt only the first 10,000 sequences of each station were chosen for Blastx, however due to prefiltering at 3 µm the large eukaryotic proportion (Fig. 3) led us to retrieve all 376,536 significant sequences from DESeq2 (adjusted p value <0.05) analysis and align them against the non-redundant protein databases using blastx (adjusted parameters: -evalue 0.001 -max_target_seqs 25 -outfmt 5 (xml output); Altschul *et al.*, 1990).

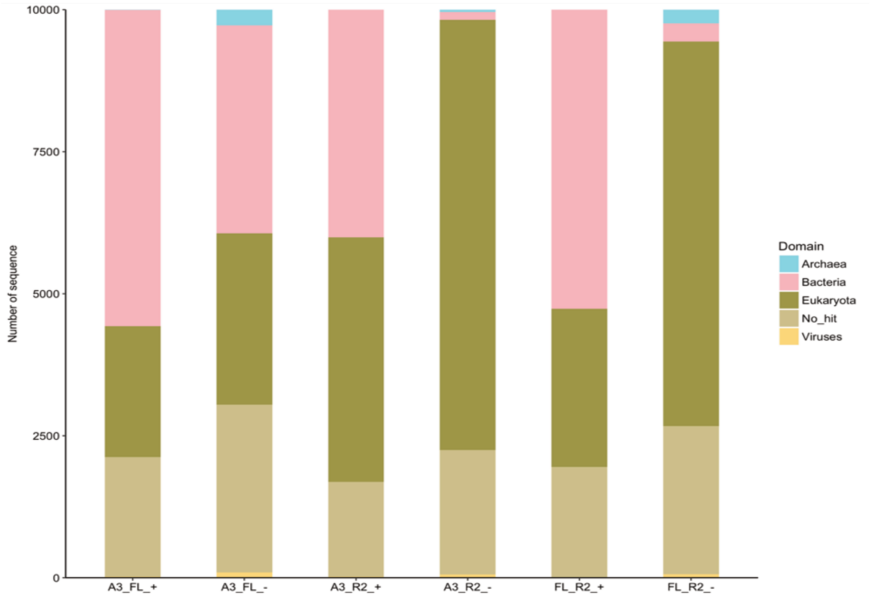


Figure 3. *BlastX results for 10,000 sequences. +/- denotes positive or negative log2 fold change for the comparison.*

Eukaryotic sequences were extracted with MEGAN (Huson *et al.*, 2007) by importing xml output from blastx and marking Bacteria and Archaea sequences. The resulting 85,152 prokaryotic significantly differentially expressed prokaryotic sequences were imported into Blast2GO (Götz *et al.*, 2008) for gene ontology (GO) annotation. Enrichment analysis using Fishers exact test was used to identify enriched functions in cellular and metabolic processes. Tables of Deseq2 results for each comparison were loaded into Blast2GO. Data was visualized by cellular component (by GO definition 'the parts of a cell or its extracellular environment'),

molecular function (by GO definition 'the elemental activities of a gene product at the molecular level such as binding and catalysis') and biological process (by GO definition 'operations or sets of molecular events with a defined beginning and end, pertinent to the functioning of integrated living cells'). Results shown in Figure 4 have been visualized by an R code from REVIGO (Supek *et al.*, 2011). Remaining results were visualized in R from Blast2GO output and enhanced using the open-source program inkscape (www.inkscape.org). Specific GO terms were chosen based on interest in functional meaning (Table 4) and these protein sequences were retrieved from the metatranscriptome and annotated with eggNOG (Huerta-Cepas *et al.*, 2016, 2019).

Table 4. *Specific GO terms chosen for eggNOG annotation.*

GO Number	GO Name	GO Category
GO:0006810	transport	Biological process
GO:0051234	establishment of localization	Biological process
GO:0051179	localization	Biological process
GO:0004872	receptor activity	Molecular function
GO:0060089	molecular transducer activity	Molecular function
GO:0031975	envelope	Cellular component
GO:0030313	cell envelope	Cellular component
GO:0016020	membrane	Cellular component
GO:0015979	photosynthesis	Biological process
GO:0044462	external encapsulating structure part	Cellular component
GO:0009279	cell outer membrane	Cellular component
GO:0019867	outer membrane	Cellular component
GO:0030312	external encapsulating structure	Cellular component
GO:0005509	calcium ion binding	Molecular function
GO:0007034	vacuolar transport	Biological process
GO:0005216	ion channel activity	Molecular function

Results

We used enrichment analysis out of assigned gene ontology terms from gene products. While prokaryotic community changes and specific functional changes have been observed this is to our knowledge the first study using GO assignment of metatranscriptomic data to identify key processes involving sudden environmental changes for the limiting nutrient iron. Our findings observe an increase of genes involving transport and localization at the iron limited site compared to the iron fertilized sites (Fig 4 & 5). These GO terms are part of the biological process ontology and cover "processes involved

in positioning a substance or cellular entity and maintaining it in that location" (GO). A transport protein (variously referred to as a transmembrane pump, transporter, escort protein, acid transport protein, cation transport protein, or anion transport protein) is a protein that serves the function of moving other materials within an organism.

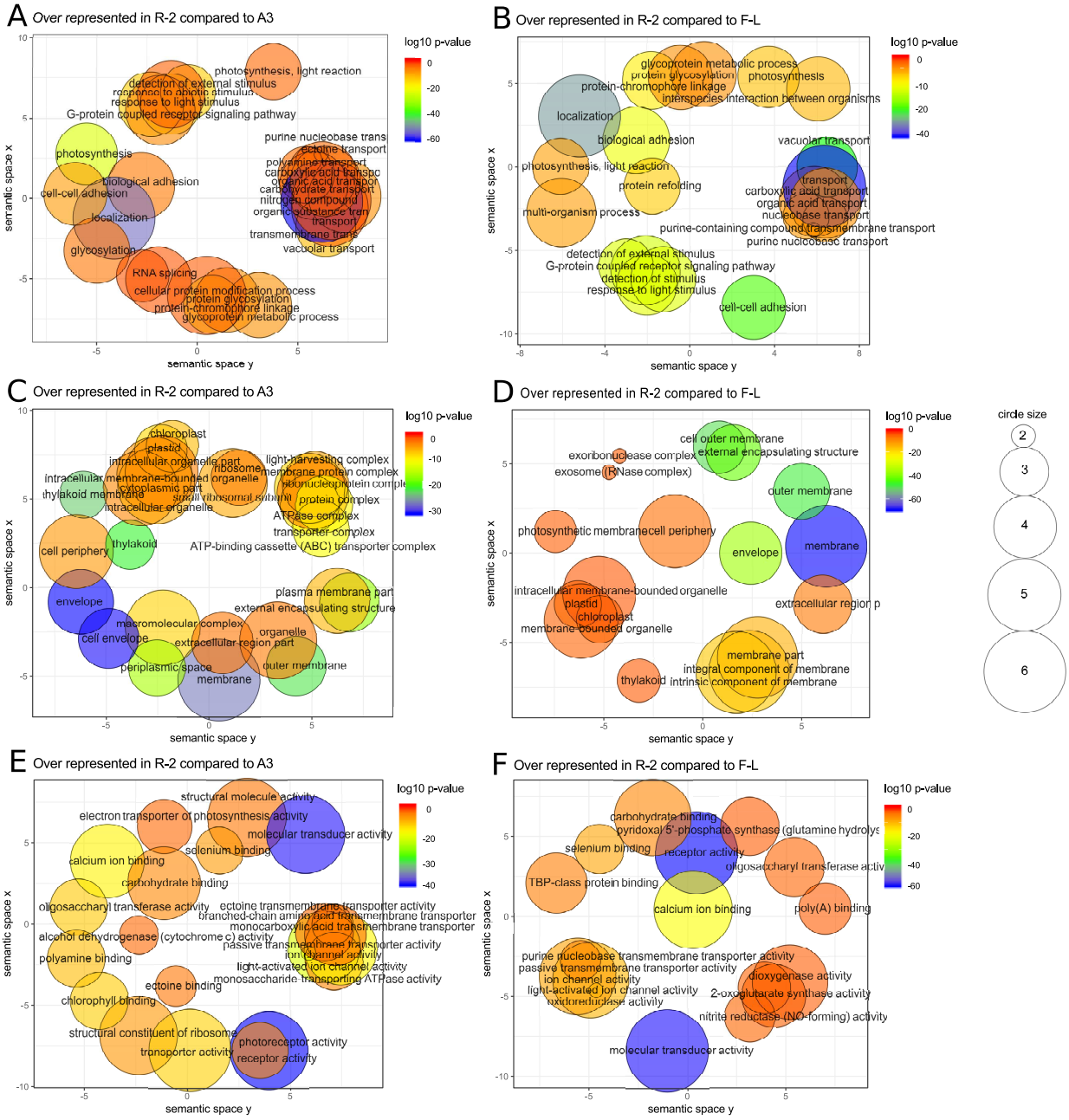


Figure 4 *Gene ontology scatterplots constructed with REVIGO for all significant GO terms from Fisher exact tests using similarity of 0.7, with SimRel as semantic similarity measure. Colours indicate the p-value of enrichment according to the legend. The size of each bubble reflects the count of each term among the enriched term list.* ***A&B.*** *Biological Process,* ***C&D*** *Cellular Component,* ***E&F*** *Molecular Function.*

Transport proteins are vital to the growth and life of all living things. There are several different kinds of transport proteins. Carrier proteins are proteins involved in the movement of ions, small molecules, or macromolecules, such as another protein, across a biological membrane (Ashburner *et al.*, 2000; The Gene Ontology Consortium, 2019). Carrier proteins are integral membrane proteins; that is, they exist within and span the membrane across which they transport substances. The proteins may assist in the movement of substances by facilitated diffusion (i.e., passive transport) or active transport. These mechanisms of movement are known as carrier-mediated transport. Each carrier protein is designed to recognize only one substance or one group of very similar substances (Mi *et al.*, 2019).

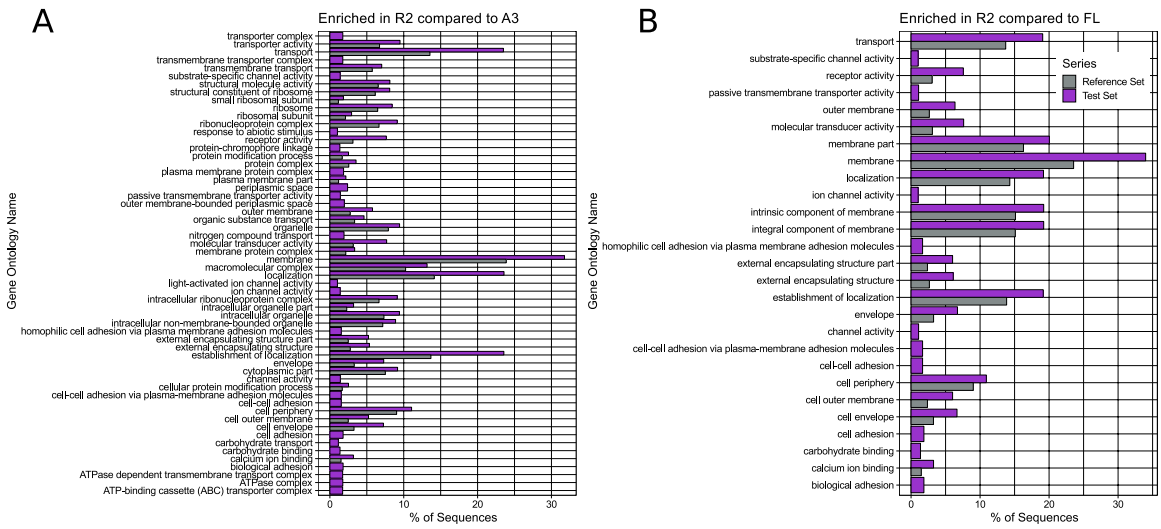


Figure 5. *Percentages of sequences by Gene Ontology terms for* ***A.*** *R2 vs A3 comparison and* ***B.*** *R2 vs FL comparison*

References

Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Čech, M., et al. (2016) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic acids research* **44**: W3–W10.

Afiahayati, Sato, K., and Sakakibara, Y. (2015) MetaVelvet-SL: an extension of the Velvet assembler to a de novo metagenomic assembler utilizing supervised learning. *DNA Res* **22**: 69–77.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *Journal of Molecular Biology* **215**: 403–410.

Anders, S., Pyl, P.T., and Huber, W. (2015) HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**: 166–169.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**: 25–29.

Christaki, U., Lefèvre, D., Georges, C., Colombet, J., Catala, P., Courties, C., et al. (2014) Microbial food web dynamics during spring phytoplankton blooms in the naturally iron-fertilized Kerguelen area (Southern Ocean). *Biogeosciences Discuss* **11**: 6985–7028.

Götz, S., García-Gómez, J.M., Terol, J., Williams, T.D., Nagaraj, S.H., Nueda, M.J., et al. (2008) High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res* **36**: 3420–3435.

Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M.C., et al. (2016) eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res* **44**: D286–D293.

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S.K., Cook, H., et al. (2019) eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res* **47**: D309–D314.

Huson, D.H., Auch, A.F., Qi, J., and Schuster, S.C. (2007) MEGAN analysis of metagenomic data. *Genome Research* **17**: 377–386.

Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**: 357–359.

Lasbleiz, M., Leblanc, K., Blain, S., Ras, J., Cornet-Bathaux, V., Hélias Nunige, S., and Quéguiner, B. (2014) Pigments, elemental composition (C, N, P, Si) and stoichiometry of particulate matter, in the naturally iron fertilized region of Kerguelen in the Southern Ocean. *Biogeosciences (submitted)*.

Love, M.I., Huber, W., and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**:.

Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**: 10.

Mi, H., Muruganujan, A., Ebert, D., Huang, X., and Thomas, P.D. (2019) PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res* **47**: D419–D426.

Namiki, T., Hachiya, T., Tanaka, H., and Sakakibara, Y. (2012) MetaVelvet: an extension of Velvet assembler to de novo metagenome assembly from short sequence reads. *Nucleic Acids Res* **40**: e155–e155.

Schmieder, R., Lim, Y.W., and Edwards, R. (2012) Identification and removal of ribosomal RNA sequences from metatranscriptomes. *Bioinformatics (Oxford, England)* **28**: 433–5.

Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011) REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. *PLOS ONE* **6**: e21800.

The Gene Ontology Consortium (2019) The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res* **47**: D330–D338.

Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., and Pachter, L. (2013) Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nature Biotechnology* **31**: 46–53.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., et al. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* **28**: 511–515.

Diatoms shape the biogeography of heterotrophic prokaryotes in early spring in the Southern Ocean

Yan Liu[1], Pavla Debeljak[1,2], Mathieu Rembauville[1], Stéphane Blain[1], Ingrid Obernosterer[1]

[1] Department of Limnology and Bio-Oceanography, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria

[2] Sorbonne Université, CNRS, Laboratoire d'Océanographie Microbienne, LOMIC, F-66650 Banyuls/mer, France

Summary

The interplay among microorganisms profoundly impacts biogeochemical cycles in the ocean. Culture-based work has illustrated the diversity of diatom–prokaryote interactions, but the question of whether these associations can affect the spatial distribution of microbial communities is open. Here, we investigated the relationship between assemblages of diatoms and of heterotrophic prokaryotes in surface waters of the Indian sector of the Southern Ocean in early spring. The community composition of diatoms and that of total and active prokaryotes were different among the major ocean zones investigated. We found significant relationships between compositional changes of diatoms and of prokaryotes. In contrast, spatial changes in the prokaryotic community composition were not related to geographic distance and to environmental parameters when the effect of diatoms was accounted for. Diatoms explained 30%of the variance in both the total and the active prokaryotic community composition in early spring in the Southern Ocean. Using co-occurrence analyses, we identified a large number of highly significant correlations between abundant diatom species and prokaryotic taxa. Our results show that key diatom species of the Southern Ocean are each associated with a distinct prokaryotic community, suggesting that diatom assemblages contribute to shaping the habitat type for heterotrophic prokaryotes.

Full Manuscript available online:
https://onlinelibrary.wiley.com/doi/epdf/10.1111/1462-2920.14579

Author contribution

YL and PD performed all molecular analyses, including the extraction of DNA and RNA as well as 16S rRNA amplification. YL carried out bioinformatic procedures and statistical analysis to establish links between prokaryotic communities and diatoms. PD collected seawater samples during the Southern Ocean and Climate Project (SOCLIM, PI Stéphane Blain). YL wrote the paper and all authors revised the final manuscript.

Caution in inferring biodiversity from number of metagenomic assembled genomes*

Zihao Zhao[1], Pavla Debeljak[1,2], Gerhard J. Herndl[1,3]

[1] Department of Limnology and Bio-Oceanography, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria

[2] Sorbonne Université, CNRS, Laboratoire d'Océanographie Microbienne, LOMIC, F-66650 Banyuls/mer, France

[3] NIOZ, Netherlands Institute for Sea Research, Department of Marine Microbiology and Biogeochemistry, and Utrecht University, PO Box 59, 1790 AB Den Burg, The Netherlands

*Submitted as *Matters Arising* in *Nature Communications*

Abstract

Metagenomic assembled genomes (MAGs) expanded our knowledge on unculturable microorganisms by shedding light onto the genomic flexibility and functional capacities. Dombrowski et al. reported high biodiversity together with novel functional potentials of the microbial community in the Guaymas Basin hydrothermal sediments using metagenome assembled genomes (MAG) and inferred that the hydrothermal ecosystem was more diverse than background sediments. By re-analyzing the 16S rDNA fragments and phylogenetic marker gene-based operational taxonomic units of the original data, here we show the diversity of microbial community in the hydrothermal vent sediment is lower than that of the background sediment samples, which contrasts their conclusions. These contradicting results indicate a flaw in the method used to estimate microbial community diversity in the study of Dombrowski et al. Given the widespread use of sequencing data, we suggest that more caution is needed when interpreting community diversity based on MAG.

Main

Dombrowski et al.[1] reported a metagenomic analysis of prokaryotic communities in the Guayamas Basin hydrothermal vent sediments. By reconstructing 551 metagenomic assembled genomes (MAGs), this work highlights the metabolic versatility and reports on novel microbes in hydrothermal vent sediments. The authors conclude that the reconstructed MAGs indicate a greater biodiversity in the more extreme environment, since the normalized abundance of the ribosomal protein S3 (RPS3) shows that these MAGs are representing the community as a whole (see their Suppl. Fig. 4). Based on these findings the authors demonstrate that these hydrothermal vent sediments harbor a higher biodiversity than the background sediments. This conclusion, however, is in striking contrast to the findings reported in previous studies[2,3].

While this work provides novel and valuable insight into the genomic flexibility as well as functional capacities of microorganisms in these ecosystems, here we show that the reported higher diversity in the vent sediment might actually be the result of a misleading analysis or somewhat loose interpretation of the term 'diversity'. Re-analyzing their data, we find evidence that the vent sediment is actually less diverse than the background sediment. We argue below that there are fundamental differences between the number of reconstructed MAGs as a proxy of diversity, as used in the study of Dombrowski et al.[1], and the actual community biodiversity as highlighted below.

We focus on the diversity statements from Dombrowski et al. and ask: 1) Does the abundance of a single marker gene like RPS3 reflect the whole microbial community and 2) Can the number of reconstructed MAGs be used as a proxy for community diversity?

We first used the metagenomic 16S rDNA gene[4] (16s miTAG) to explore the diversity. The rarefaction curve of 16S miTAG-based operational taxonomic units (OTUs) revealed that the number of reads of 16S rDNA gene fragments is too limited to represent all the samples, especially for samples from background sediments which are poorly characterized (Fig. 1A). Recovered 16S rDNA from the samples Vent 1b (0 - 3cm), Vent 2 (0 - 3cm) and Vent 3 (4 - 6cm) was on average 0.1 % of the total reads

per samples, which was shown to be sufficient in terms of read recovery to capture the structure of microbial communities[4,5]. For the other samples, recovery of 16S rDNA ranged from 0.06 to 0.08 % (Fig. 1B), which points to a limited sequencing effort when targeting diversity analysis. Only 5 % of the 16S rDNA reads could be used for miTAG analysis (Fig. 1B). Despite of the arguable application for community diversity, 16S rDNA-based analysis is based on the largest existing database for classification of microorganisms, although single marker gene-based approaches are prone to bias and are not able to cover the community as a whole.

To address this limitation, multiple single-copy marker gene-based metagenomic operational taxonomic units[6] (mOTU) were recruited to generate a high-resolution profiling of community diversity (Fig. 1C). Nearly 10 % of the total reads were employed (Fig. 1D) and in addition, as shown in the mOTU-based alpha-diversity analysis, prokaryotic community evenness in the deeper sediment is always lower than in the corresponding shallower sediments in both, the background and hydrothermal vent sediments. Furthermore, at the same depth horizon, the evenness of microbial communities is lower in the vent system, while richness is not drastically affected by both, the depth of the sediment layer and the vent except for sample Vent 3 (4 – 6 cm). These differences are represented in a decreasing Shannon index with depth as well as proximity of sampling site to the vent (Fig. 2). For example, the background samples and samples from the outside vent (Vent 1a) exhibited a higher diversity than the hydrothermal vent system if compared across the same depth layers. Moreover, the decrease in diversity from sample Vent 1a to sample Vent 1c (0 – 3 cm depth) indicates limited influence of the vent system where the sampling location is closer to the core of the vent sediment. Based on these analyses, we conclude that the hydrothermal vent sediments harbor lower microbial diversity towards deeper layers despite their distinct phylogeny due to strong selection. This conclusion differs fundamentally from that reported in Dombrowski et al.[1].

Our analysis points to several pertinent problems in assessing diversity of microbes. While there has been an immense increase in available sequence information over the last two decades leading to novel phyla and functional traits, we are only

starting to grasp the full extent of the diversity of the biosphere and the functional potential of the world's oceans[7]. Binning-based metagenomic studies hold the possibility to recover nearly complete genomes without isolated representatives. Such efforts reveal a high diversification of the microbial community and increase our knowledge on the microbial biodiversity[8]. Special caution should be taken, however, when comparison is made between samples or sites. For example, finding new phyla in the vent system but not in background sediments, as observed in the study of Dombrowski et al.[1] should not be interpreted as higher diversity. Actually, the richness was higher in the background sediments than in the vent sediments (see Fig. 2). The number of recovered MAGs is influenced by sequencing depth as well as the in-situ community complexity, and, within the same sequencing depth, high quality MAGs result from less diverse but phylogenetically distinct samples[9]. Therefore, if completeness and contamination of the bin are used as selection criteria, the number of MAGs is negatively related with community diversity. Instead, the diversity should be estimated prior to the metagenomic binning to provide realistic insights into a given environment[10].

Taken together, the conclusion reached by Dombrowski et al.[1] that the microbial diversity of Guaymas Basin hydrothermal vent sediment is higher than in the adjacent sediments is based on their specific way of analyzing the data. However, while we are aware of the caveats using phylogenetic marker genes and especially 16S rRNA gene analyses, we can conclude from our analysis that the vent system does not maintain a higher diversity than the non-vent sediments which is in contrast to the conclusion reached by Dombrowski et al.[1]. Thus, we provide evidence that more caution needs to be put into the interpretation of diversity when focusing solely on MAGs.

Material & Methods

Metagenomic reads retrieved from Dombrowski et al.[1] were trimmed with AdaptorRemoval[11] for downstream analysis. For 16s miTAG analysis, the quality-filtered reads were merged or interleaved (for unmerged reads) with BBmerge.sh as the input for Meta_RNA[12] to extract 16s rRNA genes. The predicted 16s rRNA gene read

pairs were used to construct miTAG OTU table with the UPARSE[13] function implemented in LotuS[14] at 97% similarity. MATAM[15] (v.1.5.3) was used to assemble 16s rRNA gene related reads from each metagenome and subsequently the relative abundance of 16S was calculated per sample. For mOTU analysis, quality-filtered reads were aligned to mOTU.v2b.nr.padded database[16] with BWA algorism, mOTU table was generated according to mOTU.v2b.map file[15]. Rarefaction curves were generated with rarecurve function from Vegan package and the mOTU-based alpha-diversity was calculated with rtk package on the basis of the average of 1000 rarefied datasets.

Acknowledgements

We would like to thank all members of the Movie (Microbial Oceanography Vienna) team journal club that inspired these comments.

Contributions

Z.Z., P.D. and G.J.H. designed the study and wrote the manuscript. Z.Z. and P.D. analyzed data.

Competing interests

The authors declare no competing interests.

Corresponding author

Correspondence to G.J.H. (email: Gerhard.herndl@univie.ac.at)

References

1 Dombrowski, N., Teske, A. P. & Baker, B. J. Expansive microbial metabolic versatility and biodiversity in dynamic Guaymas Basin hydrothermal sediments. *Nature Communications* **9**, 4999, doi:10.1038/s41467-018-07418-0 (2018).

2 Ruff, S. E. *et al.* Global dispersion and local diversification of the methane seep microbiome. *Proceedings of the National Academy of Sciences* **112**, 4015-4020, doi:10.1073/pnas.1421865112 (2015).

3 Lloyd, K. G. *et al.* Spatial Structure and Activity of Sedimentary Microbial Communities Underlying a Beggiatoa spp. Mat in a Gulf of Mexico Hydrocarbon Seep. *PLOS ONE* **5**, e8738, doi:10.1371/journal.pone.0008738 (2010).

4 Logares, R. *et al.* Metagenomic 16S rDNA Illumina tags are a powerful alternative to amplicon sequencing to explore diversity and structure of microbial communities. *Environmental Microbiology* **16**, 2659-2671, doi:doi:10.1111/1462-2920.12250 (2014).

5 Caporaso, J. G. *et al.* Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences* **108**, 4516-4522, doi:10.1073/pnas.1000080107 (2011).

6 Sunagawa, S. *et al.* Metagenomic species profiling using universal phylogenetic marker genes. *Nature Methods* **10**, 1196, doi:10.1038/nmeth.2693 (2013).

7 Sunagawa, S. *et al.* Structure and function of the global ocean microbiome. *Science* **348**, doi:10.1126/science.1261359 (2015).

8 Hug, L. A. *et al.* A new view of the tree of life. *Nature Microbiology* **1**, 16048, doi:10.1038/nmicrobiol.2016.48 (2016).

9 Papudeshi, B. *et al.* Optimizing and evaluating the reconstruction of Metagenome-assembled microbial genomes. *BMC Genomics* **18**, 915, doi:10.1186/s12864-017-4294-1 (2017).

10 Mohammed, M. H., Mande, S. S. & Ghosh, T. S. Classification of metagenomic sequences: methods and challenges. *Briefings in Bioinformatics* **13**, 669-681, doi:10.1093/bib/bbs054 (2012).

11 Schubert, M., Lindgreen, S. & Orlando, L. AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Research Notes* **9**, 88, doi:10.1186/s13104-016-1900-2 (2016).

12 Huang, Y., Gilna, P. & Li, W. Identification of ribosomal RNA genes in metagenomic fragments. *Bioinformatics* **25**, 1338-1340, doi:10.1093/bioinformatics/btp161 (2009).

13 Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods* **10**, 996, doi:10.1038/nmeth.2604 (2013).

14 Hildebrand, F., Tadeo, R., Voigt, A. Y., Bork, P. & Raes, J. LotuS: an efficient and user-friendly OTU processing pipeline. *Microbiome* **2**, 30, doi:10.1186/2049-2618-2-30 (2014).

15 Pericard, P., Dufresne, Y., Couderc, L., Blanquart, S. & Touzet, H. MATAM: reconstruction of phylogenetic marker genes from short sequencing reads in metagenomes. *Bioinformatics* **34**, 585-591, doi:10.1093/bioinformatics/btx644 (2017).

16 Milanese, A. *et al.* Microbial abundance, activity and population genomic profiling with mOTUs2. *Nature Communications* **10**, 1014, doi:10.1038/s41467-019-08844-4 (2019).

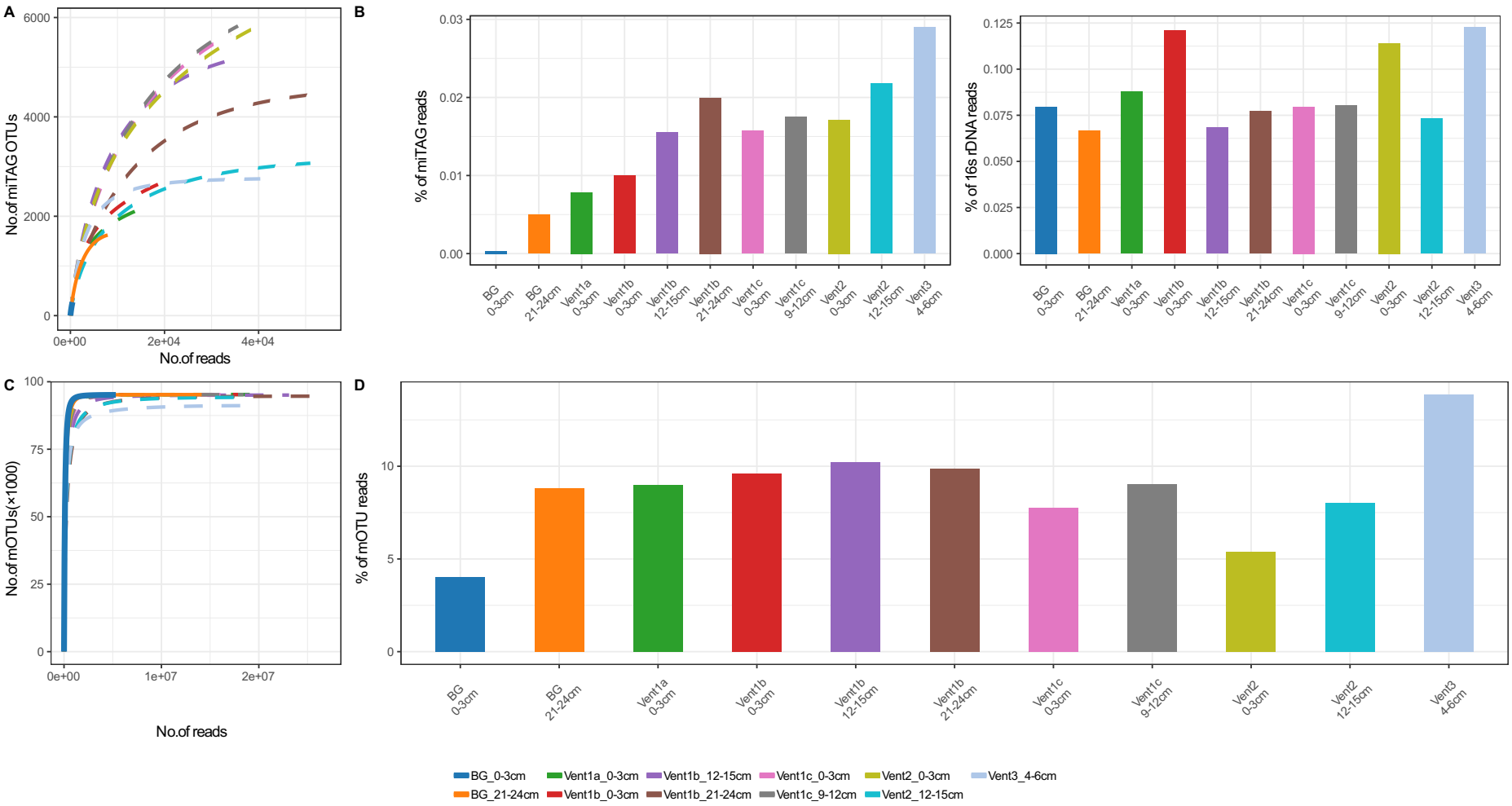


Fig. 1 Rarefaction analysis on 16S miTAG (A) and mOTU (C). Percentage of miTAG reads and assembled 16S rRNA reads per sample (B) as well as mOTU reads (D). Solid lines, samples from background sediment; dashed lines, samples from vent-related samples.

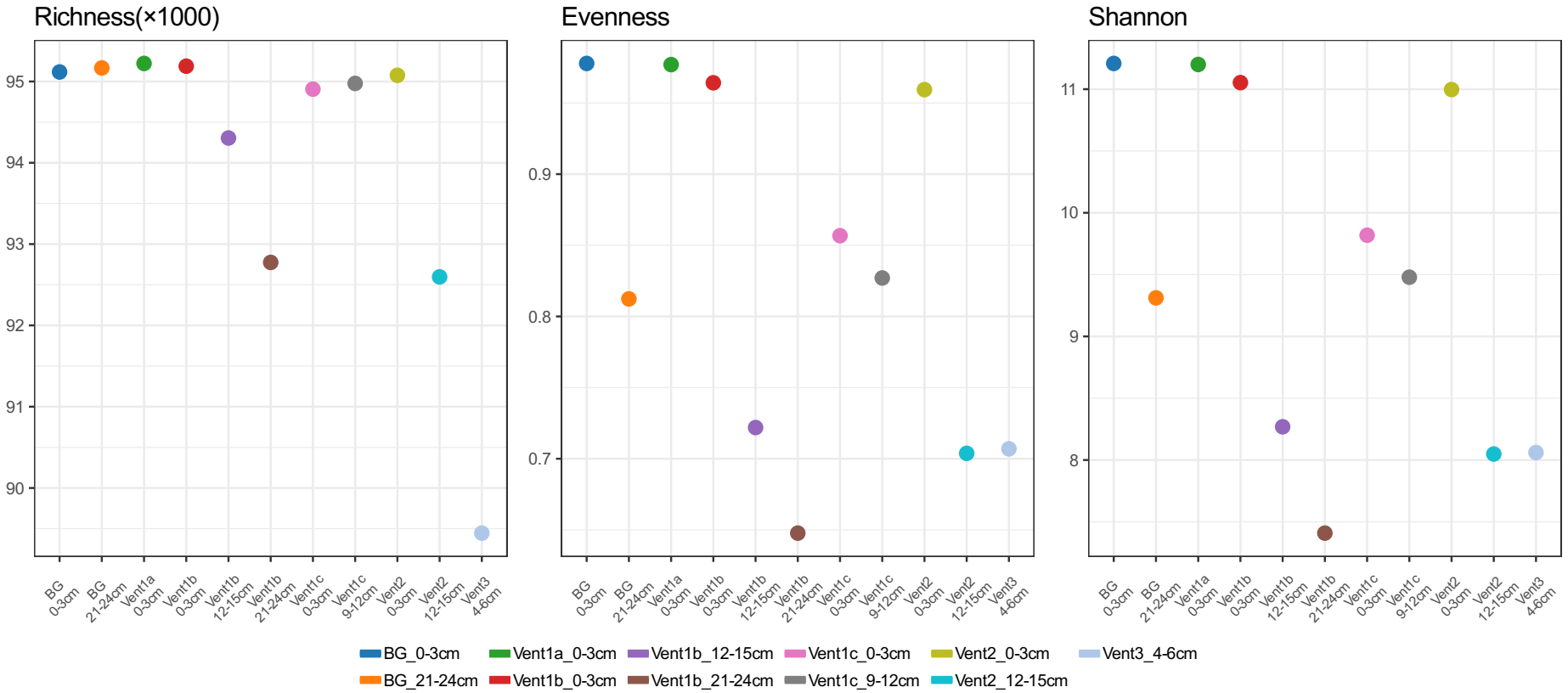


Fig. 2 mOTU based alpha diversity metric

SAR11 Iron and Carbon metabolism in contrasted Southern Ocean regimes

Ying Sun[1], Pavla Debeljak[1,2], Ingrid Obernosterer[1]

[1] Sorbonne Université, CNRS, Laboratoire d'Océanographie Microbienne, LOMIC, F-66650 Banyuls/mer, France

[2] Department of Limnology and Bio-Oceanography, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria

The following poster was presented at the SAME16 in Potsdam in September 2019 by Ying Sun. Ying is working on the metagenomes, metatranscriptomes and metaproteomes from the MOBYDICK cruise that took place February/March 2018. I performed all sampling on board, extracted DNA, RNA and proteins. I then organised sequencing of high-quality metagenomes and metatranscriptomes and prepared all extracted protein samples for MS analysis which was performed at the Vienna Metabolomics platform. Once the sequences were ready I quality filtered the raw reads and co-assembled the metagenomes. From then on Ying worked on the curation of high-quality draft genomes and focused her analysis on the clades of SAR11 and the functional fingerprint of these important marine Bacteria by integrating the information obtained from the metatranscriptomes and metaproteomes.

SAR11 Iron and Carbon Metabolism in Contrasted Southern Ocean Regimes

Ying Sun[1^*], Pavla Debeljak[1,2^], Ingrid Obernosterer[1]

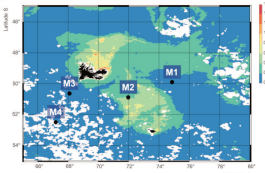
[1] CNRS, Sorbonne Université, Laboratoire d'Océanographie Microbienne, LOMIC, F-66650, Banyuls/mer, France
[2] Department of Limnology and Bio-Oceanography, University of Vienna, A-1090, Vienna, Austria

^These authors contributed equally to this study
*Correspondence to ying.sun@obs-banyuls.fr

Abstract

SAR11 is the most abundant bacterial group (*Pelagibacterales*) in the ocean and it has been intensively studied in temperate and tropical regions. The ecology of SAR11 in the Southern Ocean where iron and organic carbon both constrain microbial heterotrophic activity remains, however, largely unexplored. We present here data from contrasted regions in late summer (March 2018, **MOBYDICK**-project), High-Nutrient-Low-Chlorophyll (HNLC) waters and the naturally iron-fertilized waters off Kerguelen Island, where SAR11 was previously shown to be abundant and active. Using metagenome assembly, binning and refinement techniques, we reconstructed 133 high-quality metagenome-assembled genomes (MAGs) with an estimated >50% completeness and <10% redundancy, of which 8 were affiliated to SAR11. An additional 174 SAR11 draft genomes, with either lower completion or higher redundancy, were also analyzed in our study with careful manual assessment. Phylogenomic analysis based on single-copy orthologous genes shared by the SAR11 MAGs, as well as the average amino acid identity (AAI) at the whole-genome level, confirmed the diversity among SAR11 population sampled from the Southern Ocean. Combining the SAR11 pangenome with metatranscriptome data from each site allowed us to identify the enrichment and differential expression of genes relevant for cellular iron and carbon metabolism under contrasting environmental conditions.

Material and Methods



MOBYDICK sampling sites

Table 1 Sample and replicate descriptions.

Station	Site	Visits	Prefilter	Metagenomes (Total DNA)	Metatranscriptomes (Ribosomal-depleted RNA)
M2	Fe-fert.	3	0.8 µm	1 (pooled all revisits)	3 (biological triplicates from first visit)
M3	HNLC	2	0.8 µm	1 (pooled all revisits)	3 (biological triplicates from first visit)
M4	HNLC	2	0.8 µm	1 (pooled all revisits)	3 (biological triplicates from first visit)

Table 2 Meta-genome/-transcriptome sequencing statistics.

Station	Metagenome # Total Reads	Metatranscriptome # Ribo-depleted Reads	# Transcripts per Sample	# Transcripts per Liter	# Transcripts per Cell
M2	327,826,974	24,854,705	1,715,531,634,861	171,553,163,486	145
		28,285,715	1,199,692,136,678	239,938,427,336	203
		22,320,263	1,158,716,273,142	231,743,254,628	196
		24,522,807	1,763,330,490,035	176,333,049,003	253
M3	336,309,436	21,767,224	847,428,584,719	169,485,716,944	244
		20,801,826	1,571,056,723,086	157,105,672,309	226
		20,371,769	4,211,237,194,521	421,123,719,452	763
M4	285,050,382	22,468,309	1,783,209,594,482	178,320,959,448	323
		16,858,602	1,146,291,928,012	114,629,192,801	208

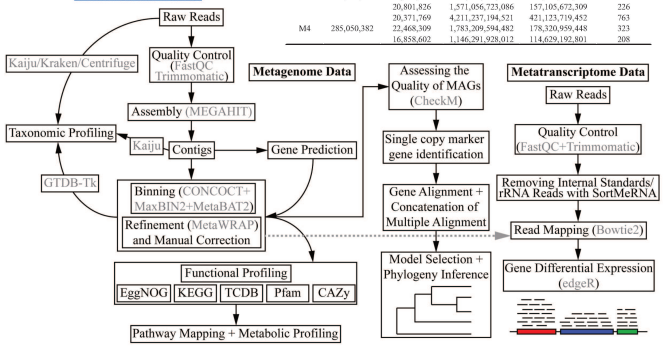


Fig. 1 The pipeline for quality control, assembly, binning, taxonomy profiling, functional annotation of metagenome data, and differential expression analysis of metatranscriptome data.

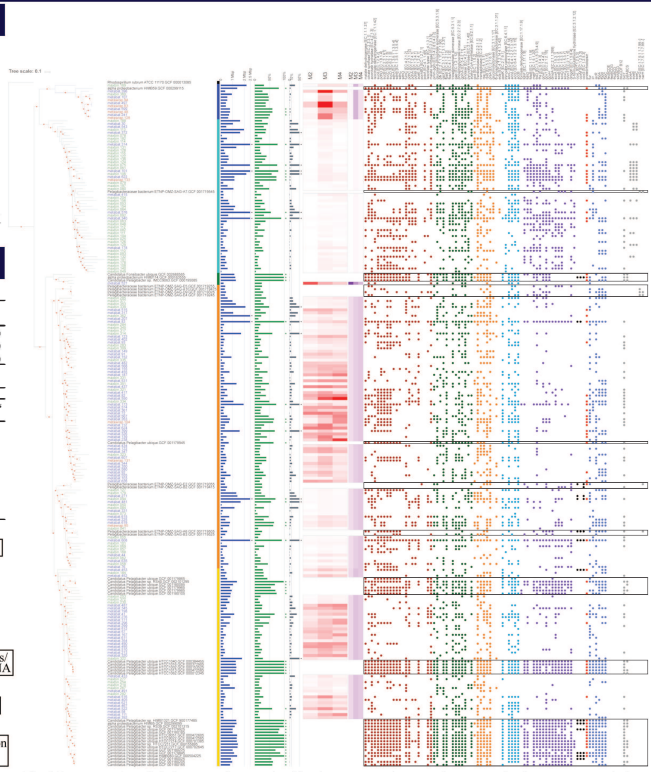


Fig. 2 Representation of phylogeny, subgroup classification, genome size, completeness, redundancy, genome relative abundance (RPKM after TMM normalization), and a survey of key genes involved in the central carbon metabolism pathways (**Fig. 3A**) and iron uptake related transporters/transcription regulators (**Fig. 3B**). Another 41 SAR11 genomes from NCBI RefSeq database were also included as benchmark reference. The phylogenetic tree was constructed based on a concatenation of 695 single-copy CheckM marker gene sets shared among *Pelagibacteraceae* family using IQ-Tree with 1,000 bootstrap replicates. Red dots on the branches represent support values no less than 70. Color bars along the tree indicate 5 major subgroups: "yellow" for subgroup I, "orange" for II, "dark green" for III, "turquoise" for IV, and "blue" for V.

Results and Discussion

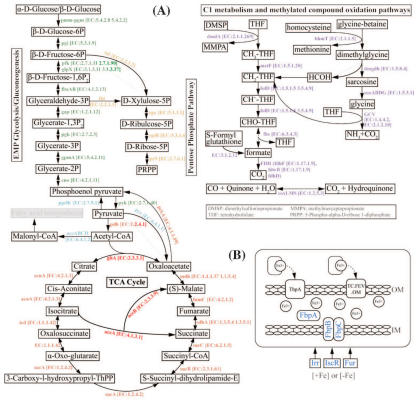


Fig. 3 Colour code for key genes involved in central carbon metabolism pathways, including TCA, EMP, PPP, C1 oxidation (A), and iron uptake related transporter/transcription regulators (B).

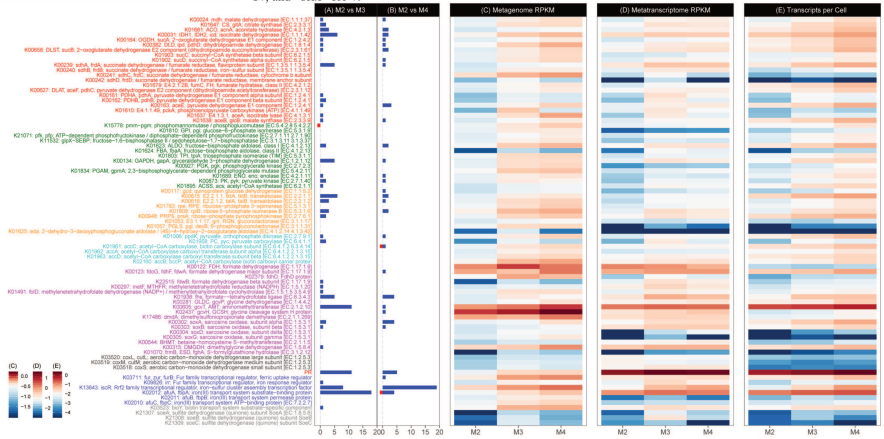


Fig. 4 Metagenome and metatranscriptome profiles of key genes marked in (**Fig. 2** and **3**). (**A** and **B**) Number of genes significantly differentially expressed in each functional category identified by the pairwise comparison between contrasting regions (M2 vs M3 [**A**], or M2 vs M4 [**B**]) using edgeR. Red bars represent the number of genes significantly higher expressed in M2 (Fe-fertilized), while blue bars show the counts of genes significantly higher expressed in M3 or M4 (both HNLC). The majority of these families were characterized by genes significantly higher expressed in HNLC than Fe-fertilized waters. (**C** and **D**) The abundance of each gene family (in RPKM unit after TMM normalization) was evaluated in both metagenome (**C**) and metatranscriptome (**D**) datasets. (**E**) Transcript abundance per cell of each gene family. Internal standard RNA molecules were used to obtain absolute quantification of transcripts. Total transcript pool size and individual transcript abundances were calculated as suggested in (Satinsky et al. 2013). All heatmaps were drawn on log10-scale with corresponding color schemes shown on the left. Pairwise Wilcoxon rank sum test with Bonferroni correction was performed to compare gene family richness between samples based on both the RPKM values and absolute quantification of transcripts per cell. In general, these gene families showed an overall trend of higher richness (metagenome) or expression level (metatranscriptome) in HNLC sites when compared to the Fe-fertilized M2 site (p-value < 0.001 for all Wilcoxon tests). Particularly, two genes involved in iron uptake, including the *iscR* and *fbpA*, are highly enriched in HNLC samples. *iscR* had 4.9 and 16.8 times as many transcripts per cell in M3 and M4 as compared to M2, respectively. *fbpA* showed 1.5 (or 1.9) times as many transcripts per cell in M3 (or M4) as compared to M2.

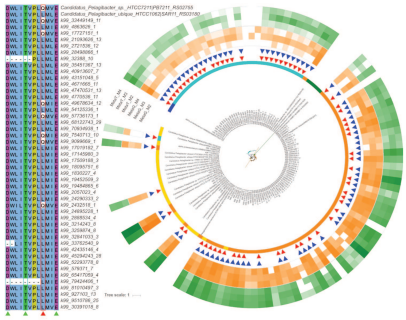


Fig. 5 Phylogenetic tree of proteorhodopsin genes identified in MAGs. PR is another interesting family that had genes significantly highly expressed in M3 (or M4) compared to M2 (**Fig. 4AB**) and, overall, at least 1.5 times as many transcripts per cell in M3 (or M4) compared to M4 (**Fig. 4E**). Circles from inside to outside: (1) subgroups, (2) MAG PR genes verified by Pfam, (3) MAG PR genes verified by EggNOG, (4-5) abundance of each gene (in RPKM unit) in metagenome (orange) or metatranscriptome (green) data. The alignment on the left highlights variations in PR residues involved in spectral tuning (red triangle; amino acid 105) and ion pumping (green).

Reference

Dimasquet, J., M. Landa and I. Obernosterer (2019). "High contribution of Pelagibacterales to bacterial community composition and activity in spring blooms off Kerguelen Island (Southern Ocean)." bioRxiv.
Giovannoni, S. J. (2017). "SAR11 Bacteria: The Most Abundant Plankton in the Oceans." Annual Review of Marine Science.
Satinsky BM, Gifford SM, Crump BC, Moran MA (2013). "Use of internal standards for quantitative metatranscriptome and metagenome analysis." Methods Enzymol.

Acknowledgement

This work was supported by the French Research program of the INSU-CNRS LEFE-CYBER (Les enveloppes fluides et l'environnement - Cycles biogéochimiques, environnement et ressources), the French ANR (Agence Nationale de la Recherche, ANR-17-CE01-0013-02), and the French Oceanographic Fleet (Flotte Océanographique Française, FOF). The MOBYDICK cruise has the following DOI: https://doi.org/10.17600/18000403.

Curriculum vitae

Personal information

Name	Pavla Debeljak
Nationality	Austria

Education

01/10/2016 – 18/10/2019	Doctoral studies in Microbial Oceanography Thesis title: *The coupling of carbon and iron cycles in the Southern Ocean through microbial metabolism* – supervised by Ingrid Obernosterer & Gerhard J. Herndl Sorbonne Université, Banyuls-sur-Mer, France
01/03/2012 – 21/07/2015	Master studies in Gender Studies (MA) Thesis title: *The power of epigenetics: Analysis on perceptions of motherhood and sexuality in scientific literature and popular media* – supervised by Sigrid Schmitz University of Vienna, Austria
01/03/2011 – 14/04/2014	Master studies in Ecology (MSc) Thesis title: *Temperature dependence of prokaryotic autotrophy in the meso- and bathypelagic Atlantic Ocean* – supervised by Roberta L. Hansman & Gerhard J. Herndl University of Vienna, Austria
01/10/2007 – 28/-2/2011	Bachelor studies in Biology (BSc) University of Vienna, Austria

Research experience

30/11/2015 – 31/08/2016	Research assistant The Ocean Cleanup, Delft, The Netherlands
01/09/2014 – 05/03/2015	Research assistant – Leonardo da Vinci Fellowship Université des Antilles et de la Guyane, Guadeloupe, French Caribbean
01/03/2013 – 01/08/2014	Tutor for Ecology University of Vienna, Faculty of Life Science
01/10/2010 – 30/06/2011	Laboratory assistant University of Vienna, Department of Ecosystem Research
01/06/2010 – 31/07/2010	Intern Institute of Oceanography and Fisheries, Split, Croatia

Other peer-reviewed publications

L. Lebreton, B. Slat, F. Ferrari, B. Sainte-Rose, J. Aitken, R. Marthouse, S. Hajbane, S. Cunsolo, A. Schwarz, A. Levivier, K. Noble, **P. Debeljak**, H. Maral, R. Schoeneich-Argent, R. Brambini & J. Reisser. Evidence that the Great Pacifc Garbage Patch is rapidly accumulating plastic. *Scientific Reports*, (2018) 8:4666, 1-15, doi:10.1038/s41598-018-22939-w

P. Debeljak, M. Pinto, M. Proietti, J. Reisser, F. F. Ferrari, B. Abbas, M. C. M. van Loosdrecht, B. Slat and G. J. Herndl. Extracting DNA from ocean microplastics: a method comparison study. *Analytical Methods,* (2017). doi:10.1039/C6AY03119F.

International conferences

June 2019	IMBER Future Ocean Meeting, Brest, France Oral presentation
September 2017	SAME 15 (Symposium of aquatic ecology), Zagreb, Croatia Poster presentation
March 2016	MICRO – 1. Microplastic Conference, Areciffe, Lanzarote, Spain Oral presentation

Field work

February – March 2018 (56 days)	Oceanographic cruise MOBYDICK - Marine Ecosystem Biodiversity and Dynamics of Carbon around Kerguelen, *RV* Marion Dufresne.
October 2016 (28 days)	Oceanographic cruise SOCLIM – Southern Ocean Climate, *RV* Marion Dufresne.
July 2013	Summer school DIMAPLAN on diversity of marine plankton with sampling aboard RV Tethys III, Villefranche-sur-Mer, France
September 2012	Marine ecology project training, The Royal Netherlands Institute of Sea Research (NIOZ), Texel, The Netherlands
July 2011	Introductory course to flora and fauna of the Mediterranean Sea, Institut Ruder Bošković, Rovinj, Croatia
August 2009	Ecological field trip from the steppe zone to the tundra, exploration of carbon and nitrogen cycles in arctic environments. Western Siberia, Russia

Bioinformatic training

October 2017	EBAME Workshop on Computational Microbial Ecogenomics, IUEM, Brest, France
March 2017	Workshop on transcriptomic and proteomic data analysis, EMBL, Heidelberg, Germany

Skills

Software

R programing language
Anvio
Incscape
Unix
Usage of clusters such as slurm & SGE
Microsoft office suite

Languages

English
French
German
Croatian

Acknowledgements

First of all, I would like to thank my supervisor Ingrid Obernosterer. Thank you for your trust in me for my first cruise without even knowing how I will perform and sending me again to this unfathomable place, the Southern Ocean. Thank you for all the effort in funding and making my PhD thesis in France possible. It has been a pleasure to work with you on a basis of interchanging easily in three languages and I know already I will miss this daily fun. Sorry for knocking so often on your door with my quite characteristic way of knocking. Thank you also for lending me your car and helping me buy my own. It greatly improved my life in Banyuls. And thank you for taking me on a my first after-work sailing trip, cannot think of a better way to finish a week!

I have immense appreciation for the wisdom, the kindness and the open-door policy when it comes to science and have hopefully been able to learn from Gerhard J. Herndl. Thank you for inspiring us youngsters, for motivating us even when times as frustrating and for the optimism you share.

Merci à Stéphane Blain, pour le support, les idées, la motivation et la possibilité de faire ma thèse dans le cadre du projet SOCLIM. C'était un très beau voyage et un jour, puisque tu m'as aidé à retrouver mon jean, je te rembourserai ce bidon de 20L (j'espère avec mes propres crédits) !

Merci à Eve Toulza qui m'a appris le principe du RNA-Seq en une matinée et pour me donner la possibilité et le privilège de pouvoir poser toutes les questions possibles.
Danke an Sara die mir anfangs wirklich sehr viel geholfen hat und immer erreichbar war.

Ein großer Dank geht an meine Kollegin und Freundin Barbara Bayer. Die letzten Jahre will ich mir ohne dich gar nicht vorstellen. Danke für deine unglaubliche Motivation und Energie, für die guten Gespräche über unsere Daten, die Aufmunterungen in schweren Zeiten.

Merci à mon unité de recherche le LOMIC pour une atmosphère de travail confortable.
Thanks to my office, Angel, Claire, Justine, we did have a lot of fun.

A big thank you goes to my 'home' department. Thomas, Christian B, Christian, my office that always found a place for me, thank you Paul and Maria! Thanks to Daniele, Eva, Chie, Tinkara, Miguel, Fede, Clara, Lisi and Zihao. It's always good to come home.

Thanks to Roberta for encouraging me and helping me. You are a true mentor to me.

J'ai eu la chance de monter sur le Marion Dufresne pour la première fois en Octobre 2016 lors de la campagne SOCLIM. C'est là-bas que j'ai rencontré deux personnes dont j'ai le privilège aujourd'hui de dire qu'elles sont devenues mes copines. D'abord maman science Julia, merci pour me dire que je suis grande maintenant. Tu m'as appris énormément de choses sur ce que ça veut dire d'être une femme en recherche, sur comment gérer sa vie et profiter du moment. J'espère que l'on aura encore plein de concerts, expos et moments ensembles. Et en suite Yseult, ma journaliste aux choix de musiques formidables. Merci pour tous les week-end parisiens, les expos, les discussions intellectuelles avec Ghislain et les meilleurs brunchs. Un jour je chanterai pour vous accompagner de mon ukulélé !

Merci à Olivier Crispi pour le meilleur labo du monde où j'ai passé pas mal de pauses café. Et merci à ma collègue Audrey, meilleure voisine et copine sur bateau ou sur terre, j'espère te revoir un jour au Finistère ! On ne dit jamais deux sans trois, à bientôt alors sur le Marion.
Coco, the first day I saw you standing in my office and I thought, this is going to be my friend. I told myself if Coco will be there I can do it. Thank you for being an amazing colleague, friend, cook, happy person, motivator and bacteria.
Yan. What to say. Best opposite person I could wish for in these 3 years. It wasn't always easy but having you there with your calmness, your stoic attitude and your wise words made a bad day turn bright again. We will see each other soon on the other side of the world.

Et même si ça nous a pris jusqu'en Avril de vraiment parler une avec l'autre j'ai l'impression qu'on n'a jamais arrêté depuis. Laurie, ma co-vivante ! Tu vois c'est tellement spécial ce que l'on a que j'ai dû inventer un mot et il me plait beaucoup. Tu es une raison pour laquelle je rêverai que ma thèse ne s'arrête pas. Merci de m'avoir donné un des plus beaux cadeaux possibles, une famille française ! Nos aventures viennent juste de commencer.
Banyuls gave me gift. It holds the name of Yan, Coco and Laurie. Thank you for being my family far away from home.

Thanks to Eva for being such a great scientist and all the lunch breaks spent together, I will miss those.

Merci à Stefan. For answering all my (sometimes annoying) questions. Merci à Julie pour ton soutien de loin. Merci à Mathieu, une belle première connaissance à Banyuls. Merci Marc pour ton aide tout au début avec la bio-informatique et pour les discussions sur les gènes, l'évolution et la vie de chinchilla !

Thanks to the 9 caves for being a home for Laurie and me and for adjusting your opening hours to our daily moods.

Merci à Monsieur Emmanuel Lepage qui m'a autorisé d'utiliser les images de son livre pour ma thèse. J'ai eu le privilège de monter deux fois sur le Marion Dufresne au cours des deux dernières années. Ces expériences ont marqué ma vie et en lisant votre livre j'ai eu l'impression de me retrouver à bord. Les couleurs, l'ambiance et même le fromage la nuit, tout y était.
Vous avez réussi à capturer ce qui fait que ce bateau est exceptionnel. Lire le « Voyage aux iles de désolation » a prolongé mes voyages.

The decision to do a PhD in Banyuls-sur-Mer was not an easy one for me. As usually one never knows what is going to happen and in the end it all turns out quite differently than initially thought which makes the journey even more exciting but sometimes also frustrating. I enjoyed these past years very much and I think I will look back to a very creative, stressful, explorative and adventurous period of my life. One thing I did learn for myself is that I am extremely grateful for my family and my friends who have unconditionally supported me throughout all times. How good it is to know that there is something to miss.

Danke an meine besten Freundinnen Babara, Carina, Anna-Lena, Lisa, Anna, Anita und Judith. Das nächste Mal werd ich bestimmt nicht mehr so viel Arbeiten wenn wir gemeinsam auf Urlaub sind. Danke für eure Unterstützung schon über mehr als 15 Jahre. Danke an Andi, mit dem ich zu studieren begonnen habe. Das war jetzt aber auch die letzte Studentenkarte im Kino. Danke für die beste Musik und deine positive Energie. Danke an die liebe Mia für das Yoga.
Und dann gibt es da diese besonderen Menschen in einer grauen Stadt ohne Meer. Manche kenne ich aus einer WG die etwas mit Luftballonen zu tun hatte und seitdem sind sie meine besten Freunde. Danke an Anna, Jeli und Hannes die mir immer das Gefühl gaben da zu sein über die Entfernung. Ihr seid mein Wien das immer mit mir dabei ist.

Danke an Christian für seine „Witze" und die Erinnerung an ‚sharing is caring'. Dann natürlich auch Jörg mein bester Ökologe mit den besten Kaffeepausen an der Althanstraße. Irgendwann sollten wir mal ein Projekt schreiben, über Böden und das Meer.

Liebe Anna, langjährige Begleiterin. Nun ist es soweit, meine also unsere Studententage sind fast gezählt. Jetzt werden wir bald die Damen die sich erzählen wie sie jung waren und zusammen in Wien studiert haben. Aber wir werden bis dahin noch andere Abenteuer bewältigt haben.

Merci à mes deux copine Jennifer et Pauline pour les histoires partagées et le soutien. Incroyable, nous avons toutes fait une thèse maintenant ! Il faut vraiment se revoir pour fêter avec un planteur et les bananes pesées !

Thanks to my former roommates Bruno and Roberto who have since become my friends. Thank you so much for inviting me back to the the house and for some great weekends througout the past years. Merci Bruno pour la porte toujours ouverte si j'avais des soucis ou problèmes ou j'avais besoin d'un logement à Paris. Thanks to my former co-worker and great friend Serena. You are always so amazing to talk to and it has been a pleasure to share these years with you. Thank you also for providing me with sicilian pistaccio pesto which warms my heart.

Hvala Matei i Mariu sto su me udomili u Heidelbergu na tjedan dana i što smo se proveli kako treba sa koktelima. Hvala Matea da si uvijek dostupna i da smo još toliko blizu unatoč tome da vise ne živimo u Beču. Nadam se da se to neće promijeniti.

Thank you Daniele for letting me stay at your place and enjoying a few days together in Japan. Always great this Roman coolness and you relaxed attitude towards science.
Großartige Ulli! Jetzt sind es schon viele Jahre seitdem wir uns gemeinsam in Guadeloupe zum Tauchkurs einschreiben wollten. Toll zu wissen, dass wir uns immer noch so gut verstehen und füreinander da sind. Danke dass du mir deine Seite Japans gezeigt hast.

Danke an Saul, ein toller Freund. Immer zur Stelle, wenn man mal einen schlechten Tag hatte. Für manche Menschen muss man erst auf eine karibische Insel fliegen um sie dann in Wien zu Freunden zu machen.

A tko bi ja bila bez vas. Hvala Mama, Tata i Franka da ste uvijek dostupni sada kada smo svi toliko daleko. Zrno po zrno pogača, kamen po kamen palača.

Jeudi le 21 Juin 2018, très tôt le matin à Banyuls sur Mer

« Jeudi 22 février 2018

Lettre à Hélène

Depuis le début de la semaine, l'aréopage de scientifiques de la mission MOBYDICK envahit les entrailles du Marion Dufresne, et se love dans ses cabines.
Le Marion est un amant généreux, il accueille sans reproche ceux qui l'ont sèchement quitté la dernière fois.
Et puis l'équipage a été averti : Pavla Debeljak se trouve à bord.

La Science a de nouveau de grandes ambitions. Point ne suffit de piéger le phytoplancton au soir de sa vie, dans des petits pièges manufacturés par l'observatoire de Banuyls S/mer. Il faut voir plus grand, plus loin. Que dit le poisson, l'oiseau, le mammifère ? Par où diable le carbone voyage-t-il, avant d'être englouti dans l'océan, ou recraché dans l'atmosphère ?!
Et ces bactéries marines, plus nombreuses que les étoiles de la voie lactée ? On voudrait bien savoir ce qu'elles bricolent de toutes ces chaînes carbonées.

Mais aujourd'hui, les chercheurs feront une pause dans leur quête effrénée de la connaissance. Aujourd'hui c'est l'anniversaire de Pavla, océanographe qui n'a pas froid aux doigts. »

~ Yseult Berger