



**HAL**  
open science

# The transcriptional regulation of the glyoxylate cycle in SAR11 in response to iron fertilization in the Southern Ocean

Sara Beier, Maria J Galvez, Veronica Molina, Géraldine Sarthou, Fabien Quéroùé, Stephane Blain, Ingrid Obernosterer

## ► To cite this version:

Sara Beier, Maria J Galvez, Veronica Molina, Géraldine Sarthou, Fabien Quéroùé, et al.. The transcriptional regulation of the glyoxylate cycle in SAR11 in response to iron fertilization in the Southern Ocean. *Environmental Microbiology Reports*, 2015, 7 (3), pp.427-434. 10.1111/1758-2229.12267 . hal-01130266

**HAL Id: hal-01130266**

**<https://hal.sorbonne-universite.fr/hal-01130266>**

Submitted on 11 Mar 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The transcriptional regulation of the glyoxylate cycle in SAR11 in response to iron fertilization in the Southern Ocean

Sara Beier,<sup>1\*†‡</sup> María J. Gálvez,<sup>1,2‡</sup> Veronica Molina,<sup>3</sup> Géraldine Sarthou,<sup>4</sup> Fabien Quérroué,<sup>4,5,6</sup> Stephane Blain<sup>1</sup> and Ingrid Obernosterer<sup>1</sup> <sup>1</sup>CNRS, Sorbonne Universités, UPMC Univ Paris 06, Laboratoire d'Océanographie Microbienne (LOMIC), Observatoire Océanologique, F-66650, Banyuls/mer France.

<sup>2</sup>Graduate program in Oceanography, Department of Oceanography, University of Concepción, Chile. <sup>3</sup>Departamento de Biología, Facultad de Ciencias Naturales y Exactas, Universidad de Playa Ancha, Avda. Leopoldo Carvallo 270, Playa Ancha, Valparaíso, Chile.

<sup>4</sup>LEMAR-UMR CNRS UBO IRD 6539, Technopole Brest Iroise, Place Nicolas Copernic, Plouzané F29280, France. <sup>5</sup>Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Tas 7001, Australia.

<sup>6</sup>Antarctic Climate & Ecosystems Cooperative Research Centre, University of Tasmania, Hobart, Tas 7001, Australia.

## Summary

The tricarboxylic acid (TCA) cycle is a central metabolic pathway that is present in all aerobic organisms and initiates the respiration of organic material. The glyoxylate cycle is a variation of the TCA cycle, where organic material is recycled for subsequent assimilation into cell material instead of being released as carbon dioxide. Despite the importance for the fate of organic matter, the environmental factors that induce the glyoxylate cycle in microbial communities remain poorly understood. In this study, we assessed the expression of isocitrate lyase, the enzyme that induces the switch to the glyoxylate cycle, of the ubiquitous SAR11 clade in response to natural iron fertilization in the Southern Ocean. The cell-specific transcriptional regulation of the glyoxylate cycle, as determined by the ratio between copy numbers of isocitrate lyase gene transcripts and isocitrate genes, was consistently lower in iron fertilized than in high-nutrient, low chlorophyll waters (by 2.4- to 16.5-fold). SAR11 cell-specific isocitrate lyase gene transcription was negatively correlated to chlorophyll *a*, and bulk bacterial heterotrophic metabolism. We conclude that the glyoxylate cycle is a metabolic strategy for SAR11 that is highly sensitive to the degree of iron and carbon limitation in the marine environment.

## Introduction

Marine systems store an estimated 700 Pg of organic carbon and its processing by aerobic heterotrophic bacterioplankton has important implications for global carbon cycling (Hedges *et al.*, 2002). Moreover, a single group like the SAR11 clade from the Alphaproteobacteria could significantly contribute to rates of carbon processing in the ocean, considering its worldwide distribution and high abundance in marine ecosystems (Morris *et al.*, 2002).

The switch between the tricarboxylic acid (TCA) cycle and the glyoxylate cycle may play a prominent role for the fate of carbon substrates: both cycles start with acetyl coenzyme A derived from carbohydrate, fatty acid or protein degradation and share a number of common reactions (Berg *et al.*, 2012). However, the glyoxylate cycle bypasses two decarboxylation steps and the coupled release of CO<sub>2</sub> and reducing equivalents (NADH<sub>2</sub>) of the TCA cycle. Instead, organic molecules are recycled and can be assimilated into cell biomass (Fig. 1).

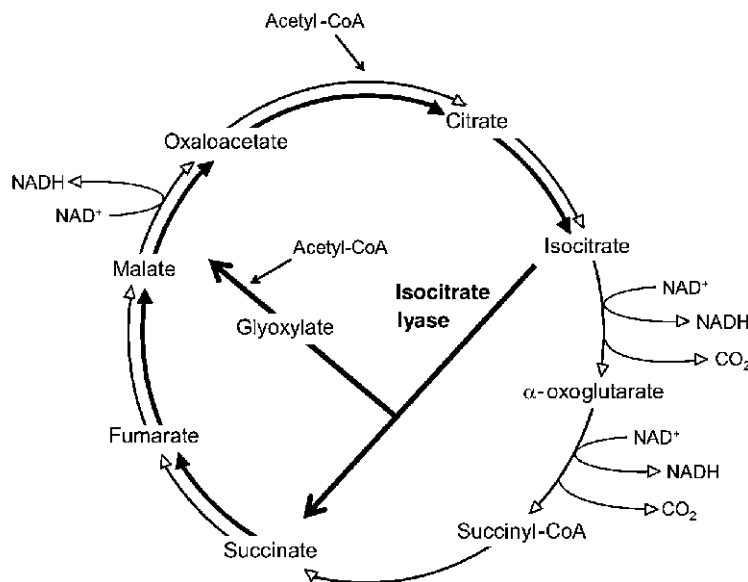
Up-expression of the glyoxylate cycle was usually described for organisms growing on carbon sources as, i.e. acetate, if other compounds that are essential for cell growth are not available. The glyoxylate cycle is hereby used to transform these carbon substrates into organic compounds as sugars or amino acids (Kretzschmar *et al.*, 2008; Tripp *et al.*, 2009; Berg *et al.*, 2012). Using metatranscriptomics, the expression of genes involved in the glyoxylate cycle were enriched in bacterioplankton communities grown under non-bloom experimental conditions as compared with communities grown in the presence of an induced phytoplankton bloom (Rinta-Kanto *et al.*, 2012). The latter treatment was likely accompanied by higher concentrations of labile carbon sources required for bacterial growth (Rinta-Kanto *et al.*, 2012).

Further studies that were based on transcriptional or proteomic responses in cultures suggest an upregulation of the glyoxylate cycle in the diatom *Thalassiosira oceanica* (Lommer *et al.*, 2012), in members of the SAR11 clade (Smith *et al.*, 2010) and in the gammaproteobacterium *Alteromonas macleodii* (Fourquez *et al.*, 2014) also in response to iron limitation. In these studies, the authors found an increase of the enzyme involved in the initiation of the glyoxylate cycle (isocitrate lyase) in iron-depleted cultures by means of gene transcripts (Smith *et al.*, 2010; Lommer *et al.*, 2012) or proteins (Fourquez *et al.*, 2014). Iron has a fundamental role in cell biology acting as cofactor of several proteins, i.e. involved in respiration or photosynthesis (Morel and Price, 2003). A strategy to deal with iron limitation is the reduction in iron cell quota (Tortell *et al.*, 1996). This is in part accomplished by a lower number and expression of iron-containing enzymes in the TCA cycle and the respiratory chain. A reduced number of iron enzymes, like cytochromes, in the respiratory chain has implications for the oxidation of NADH<sub>2</sub> derived from the decarboxylation steps of the TCA cycle back to its oxidized form NAD, while providing energy in form of adenosine triphosphate (ATP) (Fourquez *et al.*, 2014). Because ATP production via the respiration chain is inhibited during iron depletion, it may be beneficial for organisms to process organic material via the glyoxylate cycle in order to prevent the loss of carbon. Such adaptive upregulation of the glyoxylate cycle and the

coupled decrease of respiration in response to iron and carbon limitation could noticeably influence the fate of organic carbon processed by bacteria.

Prominent iron-limited regions are the eastern subarctic Pacific, the eastern equatorial Pacific and the Southern Ocean, and together they occupy one third of the global ocean (Boyd, 2004). Among these, the Southern Ocean represents the largest high-nutrient, low chlorophyll (HNLC) area on earth. Natural iron fertilization was shown to occur in the wake of islands with the region around Kerguelen Island representing the largest naturally fertilized area (Blain *et al.*, 2007). These regions provide contrasting sites with respect to iron supply and bloom development that have important consequences on the production of dissolved organic matter and the related microbial heterotrophic activity (Zubkov *et al.*, 2007; Obernosterer *et al.*, 2008; Christaki *et al.*, 2014).

The aim of our study was to investigate the TCA to glyoxylate cycle switch (Fig. 1) by means of isocitrate lyase gene expression in members of the abundant SAR11 clade inhabiting areas characterized by contrasting trophic conditions and iron availability of the Southern Ocean.

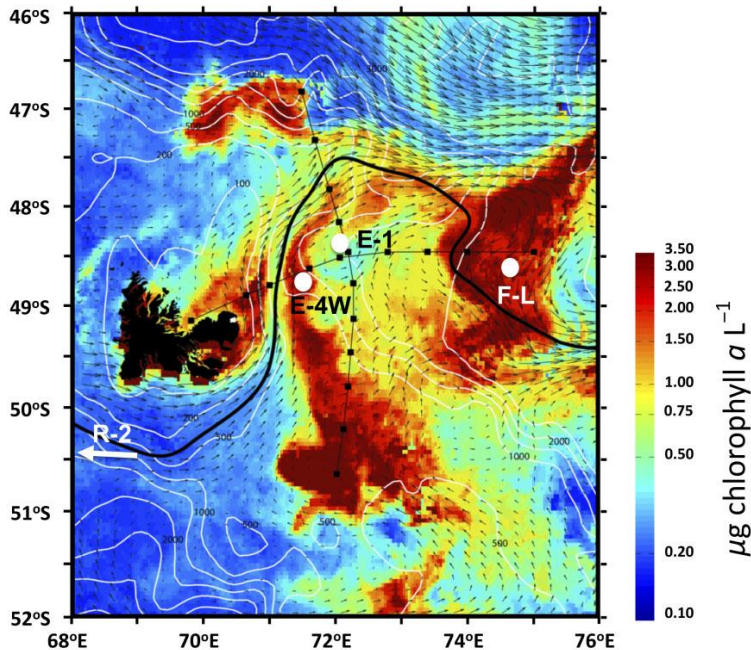


**Fig. 1.** Schematic illustration of the TCA cycle (thin line) and glyoxylate cycle (bold line) where the isocitrate lyase initiates the glyoxylate cycle by cleaving isocitrate into glyoxylate and succinate.

## Results and discussion

### *Environmental setting*

The region east of Kerguelen Island is characterized by intense mesoscale activity (Park *et al.*, 2014) resulting in diverse biogeochemical responses to the input of iron-rich water masses. Concentrations of dissolved iron (DFe) were 0.08 nM at station R-2 (mean mixed layer), in HNLC surface waters, and they varied between 0.17 and 0.22 nM (mean mixed layer) at the fertilized stations F-L and E-4W respectively (Table S1, Qu  rou   *et al.*, 2015). Concentrations of Chla were 0.3  $\mu\text{g Chla L}^{-1}$  in surface waters at station R-2, and stations F-L and E-4W, located North and South, respectively, of the Polar Front revealed intermediate (E-4W, 1.33  $\mu\text{g Chla L}^{-1}$ ) and high (F-L, 4.00  $\mu\text{g Chla L}^{-1}$ ) Chla concentrations in overall shallow mixed layers (61–38 m) (Fig. 2, Table 1). Station E-1 is located within a complex meander South of the Polar Front. No DFe data are available for station E-1. However, this site was revisited twice in a quasi-Lagrangian manner after 1 and 4 days, and the DFe concentrations determined during these visits (range 0.08 to 0.38  $\text{nmol l}^{-1}$  in the mean mixed layer of stations E-2 and E-3), and more importantly the threefold higher Chla concentrations (1.0  $\mu\text{g Chla l}^{-1}$ ) than in HNLC waters clearly place station E-1 in an iron-fertilized water parcel. Despite the large variability in Chla, dissolved organic carbon (DOC) concentrations were, except for station F-L, identical among stations, probably because the labile dissolved organic matter (DOM) released by phytoplankton was rapidly consumed. This was reflected by the large range in cell-specific bacterial heterotrophic production (0.02  $\text{fmol C cell}^{-1} \text{day}^{-1}$  at station R-2 to 0.22  $\text{fmol C cell}^{-1} \text{day}^{-1}$  at station F-L), and in cell-specific bacterial respiration (range 0.88  $\text{fmol O}_2 \text{cell}^{-1} \text{day}^{-1}$  at station R-2 to 2.25  $\text{fmol O}_2 \text{cell}^{-1} \text{day}^{-1}$  at station F-L) (Christaki *et al.*, 2014). Using Catalyzed reporter deposition Fluorescence In Situ Hybridization (CARD-FISH), the abundance of SAR11 varied between  $1.12 \times 10^8$  and  $2.79 \times 10^8 \text{ cells L}^{-1}$  and accounted for 36–49% of total bacterial abundance, indicating that members of the SAR11 clade were abundant at all stations (I. Obernosterer, unpubl. data).



**Fig. 2.** Composite MODIS/MERIS satellite image of the KEOPS2 study area in October–November 2011. Shown are chlorophyll *a* (color scale), surface velocity fields (arrows), the Polar Front (black line), the north-south and east-west transects carried out during the cruise and the position of the three stations sampled in naturally iron-fertilized waters (stations E-1, E-4W and F-L). The reference station R-2 in HNLC waters is only indicated by an arrow as it is out of the area of the map (longitude 66°41.570'E, latitude 50°23.370'S). Map is courtesy F. d'Ovidio and collaborators. The chlorophyll content represented on the map corresponds to the last week of the KEOPS2 cruise.

### *Specificity of SAR11 clade isocitrate lyase genes amplification assays*

The up-expression of the isocitrate lyase gene under iron-depleted conditions could be a common trait in marine plankton (Smith *et al.*, 2010; Fourquez *et al.*, 2014). Our alignment analyses indicate that this gene presented a high sequences variance within the referential microbial groups analysed; therefore, we designed specific-group primers for the amplification of isocitrate lyase genes in the common and ubiquitous SAR11 clade (Table S2). Sequences ( $n = 200$ ) derived from randomly picked clones for each of the two developed primer pairs indicated the specific amplification of isocitrate lyase genes derived mainly from members of the SAR11 clade (Fig. S1). The only sequence that was not affiliated to SAR11 clade isocitrate lyases (iso514\_clone56) still grouped among isocitrate lyases from other organisms [closest relative from RefSeq Database (blastp search, November 2013): isocitrate lyase from *Hydrocarboniphaga effusa*, GI: 494341513, e-value:  $2e-16$ ]. Nonetheless, possible rare amplification events of isocitrate lyase gene sequences that were not derived from members of the SAR11 clade would still allow testing for transcription of isocitrate lyase in response to iron limitation.

The construction of degenerated primers (Table S2) in order to amplify isocitrate lyase genes from possibly all members of the SAR11 clade may lead to unequal amplification of isocitrate lyase genes derived from different SAR11 strains.

However, we assumed transcription rates [ratio of ribonucleic acid : deoxyribonucleic acid (RNA:DNA) copy numbers] in different SAR11 strains to be equal or at least similar in response to the environmental conditions at each sample site, because these different strains are phylogenetically closely related and should function in a similar way. If transcription rates among different strain

are similar, the overall transcription rate is only marginally influenced by different amplification efficiencies of the strain-specific isocitrate lyase genes (Table S3). SAR11 in our sample sites was dominated by a single OTU (at 97% similarity of the 16S rRNA gene) accounting for 60–66% of all SAR11 sequences (Landa, 2013).

#### *Transcription patterns of the isocitrate lyase gene*

We compared the cell specific transcription of the isocitrate lyase gene in members of the SAR11 clade by estimating the ratio of gene copy numbers in RNA and DNA extracts from the four selected stations (Fig. 3). Lowest gene transcription was detected for the iron- fertilized station F-L, while stations E-4W and E-1 that were also affected by iron fertilization revealed intermediate transcription levels. The highest gene transcription was observed at the iron-depleted HNLC station R-2, being 16.5-fold increased compared with station F-L (Fig. 3). The performed analysis of variance revealed a highly significant influence of the chosen stations on the transcription of the isocitrate lyase gene among members of the SAR11 clade ( $P = 0.0003$ ). Post hoc pairwise analyses indicated significant differences between station R-2 and stations F-L and E-4W (Table 2). Even if only four stations were investigated, these cover a large range in responses to iron availability and connected environmental parameters of the isocitrate lyase gene expression. Accordingly, our data demonstrate a pronounced transcriptional upregulation of the glyoxylate cycle in SAR11 at sites that feature iron depletion, but are simultaneously characterized by limited amounts of bioavailable DOM as indicated by low values for chlorophyll *a* and bacterial heterotrophic metabolism (Fig. 4). Our results obtained in the Southern Ocean extend previous observations in experimental studies on the regulation of the isocitrate

**Table 1.** Station locations and environmental and biological data of the mixed layer (ML).

Station	Sample date, time	Longitude (E)	Latitude (S)	RNA/DNA Depth (m)	ML Depth (m)	DFe (nM)	DOC ( $\mu\text{M}$ )	Chlorophyll a ( $\mu\text{g L}^{-1}$ ) <sup>a</sup>	Bacterial abundance ( $\times 10^6$ cells $\text{L}^{-1}$ )	Cell-specific bacterial respiration ( $\text{fmol O}_2 \text{ cell}^{-1} \text{ day}^{-1}$ ) <sup>b</sup>	Cell-specific bacterial production ( $\text{fmol C cell}^{-1} \text{ day}^{-1}$ ) <sup>b</sup>
R-2	26 Oct, 1:18 h	66°41.570'	50°23.370'	60	105 ± 15	0.08 ± 0.01	48 ± 1	0.3 ± 0.0	2.7 ± 0.3	0.88 ± 0.10	0.02 ± 0.00
F-L	7 Nov, 16:00 h	74°48.325'	48°37.174'	20	38 ± 7	0.22 ± 0.06	50 ± 1	4.0 ± 1.6	6.0 <sup>c</sup>	2.25 ± 1.06	0.22 ± 0.01
E-1	30 Oct, 16:20 h	72°10.652'	48°29.867'	20	72 ± 38	NA	48 ± 1	1.0 ± 0.0	4.3 ± 0.0	0.54 ± 0.14	0.07 ± 0.00
E-4W	10 Nov, 12:50 h	71° 25.505'	48°45.928'	30	61 ± 11	0.17 ± 0.03	48 ± 1	1.3 ± 0.1	6.0 ± 0.0	1.48 ± 0.13	0.09 ± 0.02

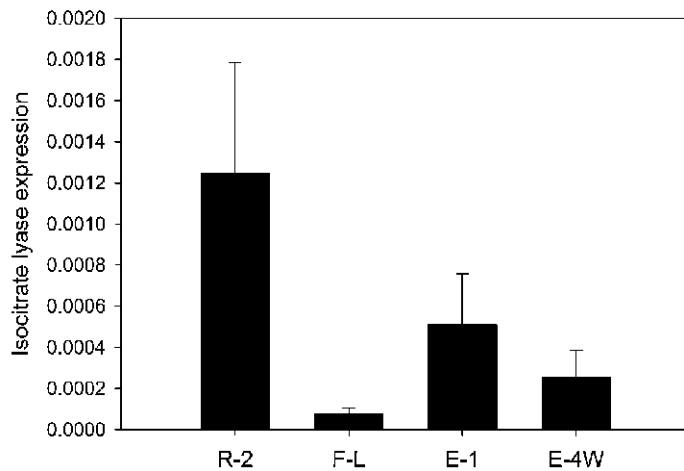
**a.** Data from Lasbleiz and colleagues (2014).

**b.** Data from Christaki and colleagues (2014).

**c.** Only one data point for the ML available. NA, not available.

RNA/DNA depth indicates the depths where water for nucleic acid extractions was sampled. For all environmental parameters, mean values ± standard deviations for the ML are given. Values for individual measurements are given in Table S1. For station E-1, no value for dissolved iron is available because the used cartridges were contaminated. Concentrations of dissolved iron (DFe) and dissolved organic carbon (DOC) were measured using previously described methods (Benner and Strom, 1993; Sarthou *et al.*, 2003). 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; Quéroùé *et al.*, 2015). For all environmental parameters, we computed mean values from several measurements that were taken at different depths within the mixed layer (Table S1). Further details about the sampling procedure are outlined in Appendix S1.





**Fig. 3. Isocitrate lyase gene expression.** Quantification of the isocitrate lyase gene expression estimated by the ratio between isocitrate lyase copy numbers per nanogram RNA and isocitrate lyase copy numbers per nanogram DNA. Both values were obtained by a quantitative PCR approach. Error bars indicate the standard deviation of the RNA/DNA ratios estimated from the three ratios as described in Appendix S1. Further methodological details are outlined in Appendix S1. Information about the developed primers that target the isocitrate lyase gene of Pelagibacter-like organisms is given in Table S2 and Fig. S2.

lyase gene expression in response to either iron (Smith *et al.*, 2010; Lommer *et al.*, 2012; Fourquez *et al.*, 2014) or DOM supply via phytoplankton exudates (Rinta-Kanto *et al.*, 2012). The large range of expression patterns observed in the present study (up to 16.5-fold) compared with the response of the SAR11 member *Candidatus Pelagibacter ubique* (*Cand. P. ubique*) to iron limitation in culture [1.6 fold-change, (Smith *et al.*, 2010)] may in our case indicate a co-regulation mediated simultaneously via iron and DOM supply. Heterotrophic bacterial production was indeed co-limited by iron and organic carbon in the study region (Obernosterer *et al.*, 2014).

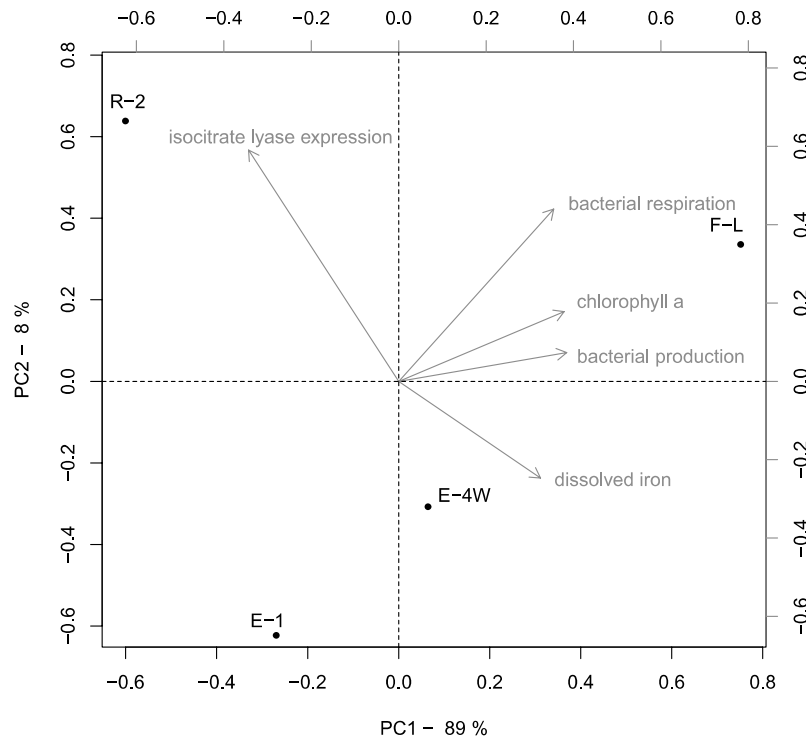
Recent observations from culture and environmental studies suggest a potential role of light in the regulation of the isocitrate lyase gene. SAR11 cells as proteorhodopsin-containing organisms (Giovannoni *et al.*, 2005a) could increase the transcription of the isocitrate lyase gene in the light analogously to *Dokdonia* sp. MED 134 (Palovaara *et al.*, 2014). Diel patterns in the expression of the isocitrate lyase gene in members of the SAR11 clade in natural bacterial communities in the coastal Pacific ocean support this idea (Ottesen *et al.*, 2013, Palovaara *et al.*, 2014). By contrast, the organic carbon-starved *Cand. P. ubique* did not reveal any changes in the expression of the isocitrate lyase gene in response to light/dark regimes (Steindler *et al.*, 2011). Another light-driven process was recently proposed by Carini and colleagues (2014). These authors have shown that *Cand. P. ubique* cannot use exogenous thiamin (vitamin B<sub>1</sub>), an essential coenzyme for the TCA cycle, and that this member of SAR11 is auxotrophic for the thiamin precursor 4-amino-5-hydroxymethyl-2-methylpyrimidine (HMP), a molecule that is degraded by light. The enhanced availability of HMP in the dark could trigger the isocitrate lyase expression. To which extent light influences the expression of the isocitrate lyase gene clearly merits further investigation.

Regulation of the glyoxylate cycle expression was described in detail on the posttranscriptional level (Laporte *et al.*, 1989; White, 2007; Berg *et al.*, 2012). However, specifically *Cand. P. ubique* was shown to feature a posttranscriptional regulation system of the glyoxylate shunt that differs from that described in other bacteria and is based on a riboswitch driven by cellular glycine concentrations (Tripp *et al.*, 2009; Tripp, 2013). The increased transcript abundance of isocitrate lyase genes in response to iron depletion in cultured *Cand. P. ubique* (Smith *et al.*, 2010) as well as culture independent data from our study imply that this gene is also regulated on the transcriptional level. Even though an upregulation of isocitrate lyase was found under iron- and carbon- limited conditions, this gene cannot be considered a general stress response gene since it is not upregulated under nitrogen- and sulfur-limited conditions (Smith, 2011; Smith *et al.*, 2013). To the best of our knowledge, the mechanism underlying transcriptional regulation of the isocitrate lyase gene in *Pelagibacter* sp. has so far not been described.

**Table 2.** Differences in isocitrate lyase gene transcription among the sampled stations were tested for significance by applying analysis of variance (ANOVA) with post-hoc pairwise comparisons (Tukey test).

Sample site	Pairwise <i>P</i> -value
E-4W/E-1	0.1625586
F-L/E-1	0.0023862
R-2/E-1	0.1717241
F-L/E-4W	0.0504277
R-2/E-4W	0.0068310
R-2/F-L	0.0002238

The overall *P*-value for the ANOVA was 0.000288 and pairwise posthoc *P*-values (Tukey test) of the ANOVA comparing isocitrate lyase gene expression at different sample sites are displayed in the table. Further details about the performed statistical analyses are given in Appendix S1.



**Fig. 4.** Principal component analyses (PCA). In order to estimate principal components, we used the non-linear iterative partial least square (NIPALS) algorithm (Wold, 1966) implemented in the *ade4* package of the R program (Chessel *et al.*, 2004). NIPALS analyses allow the extraction of principal components from a dataset that contains missing variables, as it was the case for our data. In order to estimate the contribution of the two first principal components to the total variance, we estimated in total four principal components, as it would be the case in an analog PCA. The variables for isocitrate lyase expression, bacterial respiration and bacterial production represent cell-normalized values.

Multivariate analyses point to a strong regulation of the isocitrate lyase gene by DFe, Chla and bacterial activity, with varying degrees of influence (Fig. 4). Concentrations of DFe and Chla are indicators of the amount of bioavailable iron, while bacterial production is an indicator of bioavailable DOM. In the Southern Ocean, the iron and carbon cycles are tightly coupled; thus, the SAR11 isocitrate lyase gene expression pattern we observed in the present study points to a response to the availability of both iron and organic carbon. The direct effect of iron on bacterial heterotrophic metabolism was shown experimentally. Bacterial cells grown under iron-limited conditions were shown to contain lower iron quota (Tortell *et al.*, 1996; Fourquez *et al.*, 2014). This could in part be a consequence of the reduced number of iron-containing enzymes within the TCA cycle and the respiratory chain (Fourquez *et al.*, 2014). Channeling organic substrate through the glyoxylate cycle could be a strategy to avoid loss of carbon atoms (Fig. 1) if iron concentrations are scarce and NADH-dependent ATP production via the respiration chain is inhibited. The upregulation of the glyoxylate cycle that may not only occur in SAR11 entails that a smaller fraction of carbon molecules per cell are respired to CO<sub>2</sub>. However, also the limited availability of labile carbon substrates likely causes a decrease in respiration at the HNLC site. Probably both mechanisms cause the negative correlation between DFe and single cell respiration along the principal component analysis axis (Fig. 4).

The otherwise streamlined genome of *Pelagibacter* sp. (Giovannoni *et al.*, 2005b) contains all genes for the glyoxylate cycle. This feature could represent an ecological advantage in vast parts of the oceans where iron is a limiting nutrient and SAR11 dominates the bacterioplankton, including our study sites where the cosmopolitan subclade Ia was dominant (Landa, 2013). Moreover, marine strains unrelated to the geographically ubiquitous distributed SAR11 subclade Ia (HIMB59 and HIMB114) isolated from non-iron-limited coastal zones of the tropical North Pacific lack the isocitrate lyase gene (Grote *et al.*, 2012). Also, this gene was absent in members of the freshwater sister clade LD12 (S. Bertilsson, pers. comm.). Even though freshwater systems contain iron in the micromolar range, this nutrient can be limiting for microbial metabolism (Vrede and Tranvik, 2006). The higher concentrations and qualitative characteristics of DOM in freshwater systems might lead to a stronger iron-binding capacity and result in the lower availability of iron. Whether enhanced iron concentrations in freshwater and marine systems and the lack of the full glyoxylate cycle are related remains to be investigated.

Based on our study, recent metatranscriptomic approaches (Rinta-Kanto *et al.*, 2012; Rivers *et al.*, 2013) and culture-dependent studies (Smith *et al.*, 2010; Lommer *et al.*, 2012; Fourquez *et al.*, 2014), we propose isocitrate lyase as a key indicator gene of the glyoxylate cycle whose transcription is sensitively coupled to environmental changes in marine environments associated with iron and DOM availability. Future studies that investigate in more detail regulation mechanisms of this gene in specific response to iron and/or DOM supply under varying light regimes may help to better understand and predict the processing of carbon through the TCA or glyoxylate cycle.

## **Acknowledgements**

We thank the captain and the crew of the *RV Marion Dufresne* and the chief scientist B. Quéguiner for their support during the KEOPS2 cruise. The help of M. Landa in the collection of seawater for nucleic acid analyses is greatly appreciated. J. Caparros excellently performed DOC analyses. The colour products for the Kerguelen area were produced by Ssalto/Duacs and CLS with support from Cnes. Three anonymous reviewers greatly improved a previous version of the manuscript. We thank Y. Park, M. Zhou and F. d'Ovidio for providing Fig. 2. M.J. Galvez was supported by the LIA MORFUN, the ECOS-SUD project PROMO (C12U01) and CONICYT-PCHA/Magíster Nacional/2014– 221320590. The KEOPS2 project received financial support from the CNRS-INSU-LEFE-CYBER, the ANR-10-BLAN- 0614 and the IPEV.

## References

- Benner, R., and Strom, M. (1993) A critical-evaluation of the analytical blank associated with doc measurements by high-temperature catalytic-oxidation. *Mar Chem* **41**: 153–160.
- Berg, J.M., Tymoczko, J.L., and Stryer, L. (2012) *Biochemistry*, 7th edn. New York, NY, USA: W. H. Freeman and Company.
- Blain, S., Queguiner, B., Armand, L., Belviso, S., Bombled, B., Bopp, L., *et al.* (2007) Effect of natural iron fertilization on carbon sequestration in the Southern Ocean. *Nature* **446**: 1070–U1.
- Boyd, P. (2004) Ironing out algal issues in the southern ocean. *Science* **304**: 396–397.
- Carini, P., Campbell, E.O., Morre, J., Sanudo-Wilhelmy, S.A., Thrash, J.C., Bennett, S.E., *et al.* (2014) Discovery of a SAR11 growth requirement for thiamin's pyrimidine precursor and its distribution in the Sargasso Sea. *ISME J* **8**: 1727–1738.
- Chessel, D., Dufour, A.B., and Thioulouse, J. (2004) The ade4 package – I: one-table methods. *R-News* **4**: 5–10.
- 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). *Biogeosci Discuss* **11**: 6985–7028.
- 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.
- Giovannoni, S.J., Bibbs, L., Cho, J.C., Stapels, M.D., Desiderio, R., Vergin, K.L., *et al.* (2005a) Proteorhodopsin in the ubiquitous marine bacterium SAR11. *Nature* **438**: 82–85.
- Giovannoni, S.J., Tripp, H.J., Givan, S., Podar, M., Vergin, K.L., Baptista, D., *et al.* (2005b) Genome streamlining in a cosmopolitan oceanic bacterium. *Science* **309**: 1242–1245.
- Grote, J., Thrash, J.C., Huggett, M.J., Landry, Z.C., Carini, P., Giovannoni, S.J., and Rappe, M.S. (2012) Streamlining and core genome conservation among highly divergent members of the SAR11 clade. *Mbio* **3**: e00252–12.
- Hedges, J.I., Baldock, J.A., Gelin, Y., Lee, C., Peterson, M.L., and Wakeham, S.G. (2002) The biochemical and elemental compositions of marine plankton: a NMR perspective. *Mar Chem* **78**: 47–63.
- Kretzschmar, U., Khodaverdi, V., Jeoung, J.-H., and Goerisch, H. (2008) Function and transcriptional regulation of the isocitrate lyase in *Pseudomonas aeruginosa*. *Arch Microbiol* **190**: 151–158.
- Landa, M. (2013) Lien entre matière organique dissoute et diversité des communautés bactériennes hétérotrophes marines. PhD thesis. University Pierre et Marie Curie. Paris. France.

- Laporte, D., Stueland, C., and Ikeda, T. (1989) Isocitrate dehydrogenase kinase phosphatase. *Biochimie* **71**: 1051– 1057.
- Lasbleiz, M., Leblanc, K., Blain, S., Ras, J., Cornet-Bathaux, V., He´lias Nunige, S., and Que´guiner, B. (2014) Pig- ments, elemental composition (C, N, P, Si) and stoichiometry of particulate matter, in the naturally iron fertilized region of Kerguelen in the Southern Ocean. *Biogeosciences* **11**: 8259–8324.
- Lommer, M., Specht, M., Roy, A.-S., Kraemer, L., Andreson, R., Gutowska, M.A., *et al.* (2012) Genome and low-iron response of an oceanic diatom adapted to chronic iron limitation. *Genome Biol* **13**: R66.
- 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., Rappe, M.S., Connon, S.A., Vergin, K.L., Siebold, W.A., Carlson, C.A., and Giovannoni, S.J. (2002) SAR11 clade dominates ocean surface bacterioplankton communities. *Nature* **420**: 806–810.
- Obernosterer, I., Christaki, U., Lefevre, D., Catala, P., and Van Wambeke, F. (2008) Rapid bacterial mineralization of organic carbon produced during a phytoplankton bloom induced by natural iron fertilization in the Southern Ocean. *Deep-Sea Res Part 2 Top Stud Oceanogr* **55**: 777–789.
- Obernosterer, I., Fourquez, M., and Blain, S. (2014) Fe and C co-limitation of heterotrophic bacteria in the naturally ferti- lized region off Kerguelen Islands. *Biogeosci Discuss* **11**: 15733–15752.
- Ottesen, E.A., Young, C.R., Eppley, J.M., Ryan, J.P., Chavez, F.P., Scholin, C.A., and DeLong, E.F. (2013) Pattern and synchrony of gene expression among sympatric marine microbial populations. *Proc Natl Acad Sci USA* **110**: E488– E497.
- Palovaara, J., Akram, N., Baltar, F., Bunse, C., Forsberg, J., Pedrós-Alió, C., *et al.* (2014) Stimulation of growth by proteorhodopsin phototrophy involves regulation of central metabolic pathways in marine planktonic bacteria. *Proc Natl Acad Sci* **11**: E3650–E3658.
- Park, Y.-H., Durand, I., Kestenare, E., Rougier, G., Zhou, M., d’Ovidio, F., *et al.* (2014) Polar Front around the Kerguelen Islands: an up-to-date determination and associated circulation of surface/subsurface waters. *J Geophys Res Oceans* **119**: 6575–6592.
- Qu´erou´e, F., Sarthou, G., Planquette, H.F., Bucciarelli, E., Chever, F., van der, Merwe, *et al.* (2015) High variability of dissolved Fe at the vicinity of the Kerguelen Islands (South- ern Ocean). *Biogeosciences Discussions, KEOPS 2 special issue* **12**: 231–270.
- Rinta-Kanto, J.M., Sun, S., Sharma, S., Kiene, R.P., and Moran, M.A. (2012) Bacterial community transcription pat- terns during a marine phytoplankton bloom. *Environ Microbiol* **14**: 228–239.
- Rivers, A.R., Sharma, S., Tringe, S.G., Martin, J., Joye, S.B., and Moran, M.A. (2013) Transcriptional response of bathypelagic marine bacterioplankton to the Deepwater

Horizon oil spill. *ISME J* **7**: 2315–2329.

Sarthou, G., Baker, A.R., Blain, S., Achterberg, E.P., Boye, M., Bowie, A.R., *et al.* (2003) Atmospheric iron deposition and sea-surface dissolved iron concentrations in the eastern Atlantic Ocean. *Deep-Sea Res Part -Oceanogr Res Pap* **50**: 1339–1352.

Smith, D.P. (2011) The proteomic and transcriptomic responses to iron, sulfur, and nitrogen limitation in the abundant marine bacterium *Candidatus Pelagibacter ubique*. PhD thesis, Oregon State University, Corvallis, USA.  
<https://ir.library.oregonstate.edu/xmlui/bitstream/handle/1957/26077/SmithDanielP2011.pdf?sequence=1>.

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.

Smith, D.P., Thrash, J.C., Nicora, C.D., Lipton, M.S., Burnum-Johnson, K.E., Carini, P., *et al.* (2013) Proteomic and transcriptomic analyses of ‘*Candidatus Pelagibacter ubique*’ describe the first P-II-independent response to nitrogen limitation in a free-living alphaproteobacterium. *Mbio* **4**: e00133–12.

Steindler, L., Schwalbach, M.S., Smith, D.P., Chan, F., and Giovannoni, S.J. (2011) Energy starved *Candidatus Pelagibacter ubique* substitutes light-mediated ATP production for endogenous carbon respiration. *PLoS ONE* **6**: e19725.

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.

Tripp, H.J. (2013) The unique metabolism of SAR11 aquatic bacteria. *J Microbiol* **51**: 147–153.

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.

Vrede, T., and Tranvik, L.J. (2006) Iron constraints on planktonic primary production in oligotrophic lakes. *Ecosystems* **9**: 1094–1105.

White, D. (2007) *The Physiology and Biochemistry of Prokaryotes*, 3rd edn. New York, NY, USA: Oxford University Press.

Wold, H. (1966) Estimation of principal components and related models by iterative least squares. In *Multivariate Analysis*. Krishnaiah, P. (ed.). New York, NY, USA: Academic Press, pp. 391–420.

Zubkov, M.V., Holland, R.J., Burkill, P.H., Croudace, I.W., and Warwick, P.E. (2007) Microbial abundance, activity and iron uptake in vicinity of the Crozet Isles in November 2004–January 2005. *Deep-Sea Res Part II-Top Stud Oceanogr* **54**: 2126–2137.

## Supporting information

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

**Fig. S1.** Phylogenetic tree containing isocitrate lyase sequences from cultured reference sequences (RefSeq) and PCR amplicons from primer pairs Iso448f/Iso799r (Iso799\_cloneXX, orange letters) and Iso448f/Iso514r (Iso541\_cloneXX, blue letters). The tree was constructed by calculating a backbone tree with the reference protein sequences using SATE versus 2.2.5. (Treeestimator: FASTTREE, Aligner: MAFFT, merger: MUSCLE, model GTR + G20, default settings) (Edgar, 2004; Katoh *et al.*, 2005; Liu *et al.*, 2009; Price *et al.*, 2010). Sequences of PCR amplicons were integrated a posteriori into the tree using the ARB program (Parsimony, partial sequence option) (Ludwig *et al.*, 2004) and sequences derived from methylisocitrate lyase genes were used as outgroup. Alphaproteobacterial sequences are indicated by bold letters and isocitrate lyases from sequenced *Pelagibacter* strains are labelled in pink.

**Fig. S2.** Alignment of the developed primers, isocitrate lyase gene sequences from cultured strains of the SAR11 clade and sequences with high similarity to SAR11 isocitrate lyases extracted from the GOS metagenomes.

**Table S1.** Individual values for environmental data in the mixed layer depth that were used to calculate mean values and standard deviations given in Table 1. **Table S2.** Sequence information and melting temperature of developed primers.

Station	DOC [ $\mu\text{M}$ ]	Cell-specific bacterial respiration [ $\text{fmol O}_2 \text{ cell}^{-1} \text{ d}^{-1}$ ]	Chlorophyll a [ $\mu\text{g L}^{-1}$ ]	Cell-specific bacterial production [ $\text{fmol C cell}^{-1} \text{ d}^{-1}$ ]	Dissolved Iron [ $\text{nM}$ ]
R-2	48 (40 m)	0.77 (20 m)	0.32 (20 m)	0.018 (20 m)	0.09 (40m)
R-2	47 (70 m)	1.41 (40 m)	0.32 (40 m)	0.023 (40 m)	0.08 (70m)
R-2	48 (100m)	0.90 (80 m)	0.27 (80m)	0.019 (80m)	
R-2		0.43 (100 m)	0.23 (100m)	0.016 (100m)	
F-L	49 (20 m)	1.50 (10 m)	2.88 (10 m)	0.220 (10 m)	0.26 (20 m)
F-L	51 (35 m)	3.00 (20 m)	5.12 (35 m)	0.210 (35 m)	0.17 (35 m)
E-1	49 (20 m)	0.62 (15 m)	1.00 (20 m)	0.074 (20 m)	NA
E-1	49 (40 m)	0.60 (20 m)	0.96 (40 m)	0.069 (40 m)	
E-1	47 (70 m)	0.40 (65 m)			
E-4W	49 (20m)	1.57 (30 m)	1.38 (10 m)	0.11 (10 m)	0.20 (20m)
E-4W	49 (30m)	1.39 (50 m)	1.40 (40 m)	0.09 (40 m)	0.16 (40m)
E-4W	47 (70m)		1.22 (70 m)	0.08 (70 m)	0.15 (70 m)

**Table S2.** Sequence information and melting temperature of developed primers.

Primer	sequence (5'-3')	melting temperature ( $^{\circ}\text{C}$ )	sequence position in the isocitrate lyase gene of <i>Candidatus Pelagibacter</i>
--------	------------------	--	--



			<i>ubique</i> HTCC1062
Iso448f	GGTGAAGCHGGDTTGGWGG	53.8-57.9	448
Iso799r	AGAGGTYTWGCWTAYGCRCCTT	51.1-56.7	799
Iso514r	GCKGCWGGNGTTCACTTGA	51.8-55.9	514

**Table S3.** Demonstration of possible biases to QPCR data that can be introduced by unequal amplification efficiency of degenerated primers.

**Table S4.** Sequence information for clones obtained from the primer pair Iso448f/Iso514r.

## Appendix S1. Experimental procedures

### Appendix S1: Experimental procedures

#### *Sample collection*

Seawater samples were collected during the natural iron fertilization experiment KEOPS2 (Kerguelen Ocean and Plateau Compared Study 2, Oct. 8 – Nov. 30 2011) in the Indian sector of the Southern Ocean. The French Polar Institute (Institut Polaire Emile Victor) approved the sampling and work conducted in the study area.

Four stations were chosen for the present study: One station was located in HNLC waters (Station R-2) and three stations were located in naturally iron-fertilized waters (Stations E-1, E-4W and F-L).

Seawater samples for nucleic acid extractions were collected with 12 L Niskin bottles mounted on a rosette equipped with a CTDO Seabird SBE911-plus. For dissolved iron (DFe) and dissolved organic carbon (DOC) analyses, samples were collected from 10 L Teflon-lined Niskin-1010X bottles mounted on a 1018 rosette system adapted for trace metal clean work. DFe concentrations ( $< 0.2 \mu\text{m}$ ) were measured onboard by Flow Injection Analysis using a modified method of the Obata method (Sarhou *et al.*, 2003). At Station E-1, the cartridge used for  $0.2 \mu\text{m}$  filtration of the samples was contaminated, and therefore no results are available for this site. DOC concentrations were measured on acidified aliquots of GF/F filtered samples using a Shimadzu TOC-VCP analyzer with a Pt catalyst at  $680^\circ\text{C}$  (Benner and Strom, 1993). All DOC analyses were run in duplicate and 3-5 injections were done for each sample.

For nucleic acid extractions, seawater was sampled at one depth in the surface mixed layer (Table 1), and the chemical and biological parameters were collected throughout the water column (Christaki *et al.*, 2014; Lasbleiz *et al.*, 2014). For all environmental parameters we computed mean values from several measurements that were taken at different depths within the mixed layer (Table S1).

### *Nucleic acid sampling and extraction*

For both, DNA and RNA extractions volumes varying between 15 L and 30 L of prefiltered water (200 µm nylon screen and 5 µm Polycarbonate Isopore filters) was collected onto 0.2 µm SuporPlus Membranes using a 142 mm filtration system (geotech equipment inc.) and a peristaltic pump. For RNA samples, the filtration procedure did not exceed 10 minutes. DNA filters were stored at -80°C and to the RNA filters 10 ml of RNA-later was added before storage at -80°C. All nucleic acid extractions were performed in triplicates by dividing the filter in 3 parts. DNA was extracted using the DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany). Cellular lysis was performed with Proteinase K (2 h incubation, 56°C) and followed by mechanical disruption using low binding zirconium beads (OPS Diagnostics, Lebanon, NJ, USA). The extracted DNA was quantified using the Quant-it Picogreen dsDNA Assay (Invitrogen Life Technologies, Carlsbad, CA, USA). 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. 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 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).

### *Primer design*

We extracted 2577 bacterial sequences belonging to the isocitrate lyase protein family from the RefSeq data base. Out of these, reference sequences with a protein sequence similarity of 50% (CD-Hit Program (Li *et al.*, 2001)) and three sequences derived from *Pelgibacter* sp. (GI numbers: 91763004, 71083928, 27377566) as cultured representatives for the SAR11 clade were used for a seed alignment. An initial alignment was produced by the Mafft v7.012 software (homolog option, (Kato *et al.*, 2005)). This alignment was then manually refined using the secondary structure of the isocitrate lyase from *M. tuberculosis* (Protein Data Bank [PDB] Entry 1f61, chain A).

Sequences from uncultured marine bacteria were extracted from the GOS metagenome (Rusch *et al.*, 2007) via the CAMERA Portal (Seshadri *et al.*, 2007). For this purpose GOS data (all open reading frames) were queried by a tBLASTN search using three isocitrate lyases from *Pelgibacter* sp. (GI numbers: 91763004,

71083928, 27377566) and 13 further reference sequences from the remaining RefSeq isocitrate lyases (35% protein sequence similarity, CD-Hit program). All RefSeq protein data and all retrieved GOS sequences with an isocitrate lyase from *Pelagibacter* sp. as top hit (translated to protein sequence, min 300 amino acids,  $e\text{-value} \leq 5.72 \times 10^{-67}$ ) were aligned to the previously described seed alignment by HMMer v2.3 software (Eddy, 1998). Corresponding nucleotide sequences were aligned according to the created protein alignment and used for primer design. The primers were designed to amplify short and long regions for different applications such as Q-PCR and cloning using Primer-BLAST from NCBI (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>) online program with C. *Pelagibacter ubique* (HTCC1062) isocitrate lyase sequence as template. The primers were then manually modified in order to include GOS environmental sequences related to SAR11-like sequences and selected to have a TM close to 60°C and low self-complementarity (Table S2).

All primer combinations were tested for PCR using pooled DNA extracts including all stations from this study. The combination of the primer pairs Iso448f/Iso799r as well as Iso448f/Iso514r resulted in the amplification of PCR products with the expected lengths (approx. 400 bp and 100 bp, respectively).

To test the amplification specificity of the primers, PCR products were subjected for cloning and subsequent sequencing. Hereby reaction were prepared in 20 µl volumes containing 1× PCR buffer (Invitrogen Life Technologies, Carlsbad, CA, USA), 2.5 mmol l<sup>-1</sup> MgCl<sub>2</sub>, 100 nmol L<sup>-1</sup> of the forward and reverse primers, 200 mmol l<sup>-1</sup> of each dNTP (Eurogentec, Liège, Belgium) and 0.8 U Taq polymerase (Invitrogen Life Technologies, Carlsbad, CA, USA). Initial denaturation for 5 min at 94°C was followed by 35 cycles (Iso448f/Iso799r) or 30 cycles (Iso448f/Iso514r) of 94°C for 1 min, 60°C for 1 min and 72°C for 1 min. A step for final elongation was at 72°C for 7 min.

4 PCR reactions of each amplification were pooled and directly used for ligation into pCR 2.1™ TOPO® vector (Invitrogen Life Technologies, Carlsbad, CA, USA) following the manufacturer's instructions. In order to remove salts from the ligation, 4 µl of the reaction were dialyzed for 1h at room temperature on VSWP filters (Millipore, Billerica, MA, USA, 0.025 µm poresize) that were located on top of 20 ml 1 × TE buffer (pH 8) in a petridish. The dialyzed and diluted (1:1, MilliQ-water [Millipore, Billerica, MA, USA]) ligation reaction was used for cloning into electro competent One Shot TOP10® Electrocomp<sup>M</sup> *E. coli* cells (Invitrogen), following the manufacturer's instructions. For both primer combinations 100 clones were picked and send for Sanger Sequencing to Macrogen (Seoul, South Korea). Sequence information is available at GeneBank under accession numbers KF991397 - KF991474 (primerpair Iso448f/Iso799r) and Table S4 (primerpair Iso448f/Iso514r).

### *Quantitative PCR*

We applied a quantitative PCR approach (Q-PCR) in order to determine isocitrate lyase copy numbers in DNA as well as RNA extracts. It turned out difficult to construct probes that would hybridize with all clone-sequences derived from

amplification with the primer pair Iso448f/Iso799r for subsequent TaqMan Q-PCR. We instead decided to use the primer pair Iso448f/Iso514r for SYBRGreen based Q-PCR.

Q-PCR for DNA extracts was performed using SybrGreen PCR Core Reagents in a StepOnePlus Real-Time PCR Systems (Applied Biosystems Life Technologies, Carlsbad, CA, USA). 1  $\mu\text{l}$  of diluted DNA ( $\sim 0.1 \text{ ng } \mu\text{l}^{-1}$ ) was used in a final reaction volume of 10  $\mu\text{l}$  containing 500  $\text{nmol L}^{-1}$  of each primer. The q-PCR reaction was performed with one cycle of 50°C (2 min) and a denaturation step of 95°C (10 min) followed by 40 cycles of 95°C for 15 s, 60°C for 20 s and 72°C for 20 s. A melting curve was created subsequently by heating 95°C (15s) cooling down to 60°C (60s) and increasing temperature again gradually by 0.3°C steps to 95°C. Fluorescence was measured at the elongation step and after each 0.3°C temperature increasing for the melting curve.

To estimate the copy numbers of isocitrate lyase transcripts a reverse transcription q-PCR was performed using the *Power SYBR Green RNA-to-Ct 1-step kit* (Applied Biosystems Life Technologies, Carlsbad, CA, USA). 1  $\mu\text{l}$  of diluted RNA (1-4  $\text{ng } \mu\text{l}^{-1}$ ) was used in a final reaction volume of 10  $\mu\text{l}$  containing 500  $\text{nmol L}^{-1}$  of each primer. A reverse transcription was performed at 48°C for 30 min. PCR started immediately after with 10 minutes at 95°C, followed by 40 cycles with 95°C for 15 s and 60°C for 60s. The reaction was finalized with a melting curve as described above.

All DNA and RNA samples respectively were processed in a single qPCR run in order to obtain best comparability between the samples. Copy numbers of isocitrate lyase genes were estimated using standards with known concentration of the target gene ranging from  $10^6$  to 10 copy numbers  $\mu\text{l}^{-1}$  (10-fold dilutions). The standards were obtained from clone iso514\_clone23 after plasmid extraction (QIAprep Spin Miniprep, Qiagen, Hilden, Germany), plasmid stabilization (Plasmid-Safe™, Epicentre Biotechnologies, Madison, WI, USA) and linearization using the Not1 restriction enzyme (New England Biolabs, Ipswich, Massachusetts, MA, USA). Amplification efficiencies for DNA and RNA amplifications respectively were 106% and 85% respectively. Tests preceding final Q-PCR amplifications with different dilutions of template DNA and RNA demonstrated that amplifications at chosen DNA/RNA concentrations were not affected by inhibitors in the nucleic acid extracts. Isocitrate lyase copy numbers per ng nucleic acid were determined by triplicates in each extract and a single mean value per extraction was used for subsequent analyses.

### *Data Analyses*

All sequenced members of the SAR11 clade (*Pelagibacter* sp.) possess a single copy of the isocitrate lyase gene and we estimated the cell specific transcription activity of the isocitrate lyase gene by calculating the ratio of gene copy number per ng RNA and gene copy number per ng DNA. Differences in isocitrate lyase gene transcription among the sampled stations were tested for significance by applying ANOVA with post-hoc pairwise comparisons (Tukey-test). Hereby we built all 9 possible ratio values by combining triplicates for RNA isocitrate copy number and DNA isocitrate

copy number. We, however, used only the two ratios with the lowest and the highest value and the median intermediate value as input data for the ANOVA, to not artificially increase the degree of freedom that was based on only three measurements for each, RNA and DNA copy numbers. These input data for the ANOVA were log transformed because visual inspection indicated a better fit of residuals to normal distribution after the transformation. In order to estimate principal components we used the non-linear iterative partial least square (NIPALS) algorithm (Wold, 1966) implemented in the ade4 package of the R program (Chessel *et al.*, 2004). NIPALS analyses allow the extraction of principal components from a dataset that contains missing variables, as it was the case for our data. In order to estimate the contribution of the two first principal components to the total variance, we estimated in total 4 principal components, as it would be the case in an analog PCA.

## References (Appendix S1)

- Benner, R. and Strom, M. (1993) A critical-evaluation of the analytical blank associated with DOC measurements by high-temperature catalytic-oxidation. *Mar Chem* **41**: 153–160.
- Chessel, D., Dufour, A., B., and Thioulouse, J. (2004) The ade4 package - I : One-table methods. *R-News* **4**: 5–10.
- 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.
- Eddy, S.R. (1998) Profile hidden Markov models. *Bioinformatics* **14**: 755–763.
- Edgar, R.C. (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**: 1–19.
- Katoh, K., Kuma, K., Toh, H., and Miyata, T. (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* **33**: 511–518.
- 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 Discuss* **11**: 8259–8324.
- Li, W.Z., Jaroszewski, L., and Godzik, A. (2001) Clustering of highly homologous sequences to reduce the size of large protein databases. *Bioinformatics* **17**: 282–283.
- Liu, K., Raghavan, S., Nelesen, S., Linder, C.R., and Warnow, T. (2009) Rapid and accurate large-scale coestimation of sequence alignments and phylogenetic trees. *Science* **324**: 1561–1564.
- Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, *et al.* (2004) ARB: a software environment for sequence data. *Nucleic Acids Res* **32**: 1363–1371.
- Price, M.N., Dehal, P.S., and Arkin, A.P. (2010) FastTree 2-approximately maximum-likelihood trees for large alignments. *Plos One* **5**: e9490.  
doi:10.1371/journal.pone.0009490

- Rusch, D.B., Halpern, A.L., Sutton, G., Heidelberg, K.B., Williamson, S., Yooseph, S., *et al.* (2007) The sorcerer II global ocean sampling expedition: Northwest Atlantic through Eastern Tropical Pacific. *PLoS Biol* **5**: e77.
- Sarthou, G., Baker, A.R., Blain, S., Achterberg, E.P., Boye, M., Bowie, A.R., *et al.* (2003) Atmospheric iron deposition and sea-surface dissolved iron concentrations in the eastern Atlantic Ocean. *Deep-Sea Res I* **50**: 1339–1352.
- Seshadri, R., Kravitz, S.A., Smarr, L., Gilna, P., and Frazier, M. (2007) CAMERA: a community resource for metagenomics. *Plos Biol* **5**: 394–397.
- Wold, H. (1966) Estimation of principal components and related models by iterative least squares. In *Multivariate Analysis*. Krishnaiah, P. (ed). New York: Academic Press, pp. 391–420.