Global Biogeochemical Cycles

Supporting Information for

Mesoscale eddies as natural iron fertilization experiments to the deep chlorophyll maximum

Nicholas J. Hawco[1,2], Benedetto Barone[2], Matthew J. Church[3], Lydia Babcock-Adams[4], Daniel J. Repeta[4], Emma K. Wear[3], Rhea K. Foreman[2], Karin M. Björkman[2], Shavonna Bent[4], Benjamin A. S. Van Mooy[4], Uri Sheyn[2], Edward F. Delong[2], Marianne Acker[4], Rachel L. Kelley[1], Alex Nelson[2], John Ranieri[3], Tara Clemente[2], David M. Karl[2], and Seth G. John[1]

[1]Department of Earth Sciences, University of Southern California, Los Angeles, CA USA.

[2]Department of Oceanography, University of Hawaii at Mānoa, Honolulu, HI, USA.

[3]Flathead Lake Biological Station, University of Montana, Polson, MT, USA.

[4]Woods Hole Oceanographic Institution, Woods Hole, MA, USA.

Contents of this file

Supplemental Methods

16S rRNA amplicon sequencing

Genomic DNA was extracted from filters collected from the long-term incubations as described in the main text. Samples were lysed by freeze-thawing and bead-beating in a Biospec Mini-Beadbeater-16, then DNA was extracted using the MasterPure Complete DNA and RNA Purification Kit (Lucigen). Genomic DNA was quantitated using a Qubit v4 with High-Sensitivity DNA Kit.

Bacterioplankton community composition was assessed by sequencing 16S rRNA gene amplicons using primers targeting the V4-V5 hypervariable region: 515F-Y, 5'-GTGYCAGCMGCCGCGGTAA-3', and 926-R, 5'-CCGYCAATTYMTTTRAGTTT-3', as recommended by Parada et al. (2016) and with multiplexing indexes as designed by the Earth Microbiome Project (Caporaso et al. 2012). Triplicate 25 µL PCR reactions consisted of: 10 µL 2x Invitrogen Platinum II Hot-Start PCR Master Mix, 0.5 µL, 10 µM indexed forward primer (515F-Y), 0.5 µL, 10 µM reverse primer (926R), 12 µL PCR water (Invitrogen Ultra Pure Distilled Water), and 2 µL genomic DNA. PCR reactions were cycled in a Bio-Rad Tetrad 2 thermal cycler on the following program: 1) 94° C, 2 hr; 2) 35 cycles of: 94° C, 45 min, 58° C, 1hr, 72° C, 1.5 hr; 3) 72° C, 10 hr. Pooled triplicate amplicons were visually verified on an agarose gel, cleaned with an ENZA Cycle Pure Kit (Omega Bio-tek), quantitated using a Qubit v3 with High-Sensitivity DNA Kit, and pooled at equimolar proportions. The pooled library was sequenced using an Illumina MiSeq with PE250 v2 chemistry at the University of Montana Genomics Core.

Amplicon sequence variants (ASVs) were generated in DADA2 v1.14.1 (Callahan et al. 2016) and classified using the SILVA v138 database (Quast et al. 2013). Sequences identified as plastids, mitochondria, and eukaryotes were removed. Samples were subsampled to 20,000 sequences using the "rrarefy" function in the R package vegan (Oksanen et al. 2019). Sequences were aligned using the R package DECIPHER (Wright 2016); aligned sequences were used to generate a phylogenetic tree using the R package phangorn (Schliep 2011); and a weighted UniFrac distance matrix (Lozupone and Knight 2005) was calculated using the R package phyloseq (McMurdie and Holmes 2013). Multivariate statistics were conducted in Primer-E v6 (Clarke and Gorley 2006). Full bioinformatics code and ASV results are available at: https://github.com/ekwear/AlohaFe16S

Transcriptomics

Approximately 1 L of seawater for metatranscriptomic samples was filtered onto a 0.2 µm Supor® Membrane Disc filters (Pall) housed in SwinnexTM filter manifolds (MilliporeSigma) using a peristaltic pump. The filtration time ranged from 15 to 20 min. Immediately following filtration, filters were placed in RNALater (Invitrogen, Waltham, MA, USA) and stored at -80°C until processing.

RNA extractions were performed by first removing RNALater (via centrifugation and pipetting), adding 300 µL of Ambion denaturing solution directly onto the filter, and spiking in External RNA Controls Consortium (ERCC) ExFold RNA Spike-In Mixes (Mix #1, 4456739 , Invitrogen) followed by vortexing for 1 min. 750 µL of nuclease-free water was added to the sample, which were then purified and DNase-treated using a Chemagen MSM I instrument with the tissue RNA CMG-1212A kit (PerkinElmer). Samples were enriched for mRNA by removal of rRNA using Ribozero (Illumina, San Diego, CA, USA). The quality of purified RNA was assessed using Fragment Analyzer high sensitivity reagents (Agilent, Santa Clara, CA, USA) and quantified using Ribogreen (Invitrogen). cDNA was synthesized and sequencing libraries were produced using the ScriptSeq v2 RNA-Seq kit (Illumina #SSV21124). Unique single-plex barcodes were annealed onto cDNA fragments during the PCR enrichment for Illumina sequencing primers over

12 cycles, following the manufacturer's guidelines. Libraries were normalized to 4 nmol L^{-1} final DNA concentration, pooled in equal volumes, and sequenced using an Illumina® NextSeq 500 system with a V2 high output 300 cycle reagent kit. A phiX quality control (Illumina) reagent was added to an estimated final contribution of 5% of the total estimated sequence density. A total of ~2 million, 150 bp paired-end reads were produced for each sample.

Sequence reads were verified and screened for quality using BBMap (v38.73, www.sourceforge.net/projects/bbmap/) to remove adapters and phiX, BFC r181 (Li, 2015) to correct sequencing errors, and Trimmomatic v0.39 (Bolger et al., 2014)) to remove low quality bases. Raw sequence reads were submitted to the NCBI SRA under project number PRJNA596510. Cleaned reads were assembled using RNA-SPAdes v3.13.2 (Bushmanova et al., 2019), and genes were predicted from assembled transcripts using Prodigal v2.6.3 (Hyatt et al., 2010). A combined gene database was created from predicted genes and the ALOHA 2.0 gene catalog (Luo et al., 2020) using CD-HIT v4.8.1 (Fu et al., 2012) to dereplicate genes at a 97% amino acid identity cutoff. Transcripts were identified and counted by mapping cleaned reads against this combined gene catalog using BWA-MEM (v0.7.17, www.github.com/lh3/bwa). Transcript counts were normalized to the ERCC standards to account for any methodological biases, including library preparation and sequencing, between samples.

Transcript counts were normalized to the volume of water filtered for each sample to calculate transcripts per mL for each gene. To use the ERCC spike-in as an internal standard in each sample, the correction factor was calculated as follows: for each ERCC standard, the known quantities of spiked-in RNA standards were compared to that same standard's read counts recovered by read mapping using BBMap (v38.73, www.sourceforge.net/projects/bbmap/). A standard curve was generated for each sample, and the corresponding correction factor calculated as the slope of the best fit line going through these pairs of values with the intersection forced to the origin. The detection limit corresponded to the lowest molar amount of the ERCC sequence detectable in each sample. Values above detection limits were further normalized to the total sample expression sum of each specific genus level annotation, using GTDB (Parks et al., 2018) to identify prokaryote transcripts. This step was required in order to compare eddy differential gene expression based on cellular regulation rather than the total number of cells.

Supplemental References

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120.

Bushmanova, E., Antipov, D., Lapidus, A., & Prjibelski, A. D. (2019). rnaSPAdes: a de novo transcriptome assembler and its application to RNA-Seq data. *GigaScience*, *8*(9), giz100.

Callahan, B.J., P.J. McMurdie, M.J. Rosen, A.W. Han, A.J.A. Johnson, and S.P. Holmes. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods 13: 581-583. doi: 10.1038/nmeth.3869

Caporaso, J.G, C.L. Lauber, W.A. Walters, D. Berg-Lyons, J. Huntley, N. Fierer, S.M. Owens, J. Betley, L. Fraser, M. Bauer, N. Gormley, J.A. Gilbert, G. Smith, and R. Knight. 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. The ISME J 6: 1621-1624. doi: 10.1038/ismej.2012.8

Clarke, K.R., and R.N. Gorley. 2006. PRIMER v6: user manual/tutorial. New Brunswick, Canada: PRIMER-E.

Fu, L., Niu, B., Zhu, Z., Wu, S., & Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics*, *28*(23), 3150–3152.

Hyatt, D., Chen, G.-L., LoCascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, *11*(1), 1–11.

Li, H. (2015). BFC: correcting Illumina sequencing errors. *Bioinformatics*, *31*(17), 2885–2887.

Lozupone, C., and R. Knight. 2005. UniFrac: a new phylogenetic method for comparing microbial communities. Appl. Environ. Microbiol. 71: 8228-8235. doi: 10.1128/AEM.71.12.8228-8235.2005

Luo, E., Eppley, J. M., Romano, A. E., Mende, D. R., & DeLong, E. F. (2020). Double-stranded DNA virioplankton dynamics and reproductive strategies in the oligotrophic open ocean water column. *The ISME Journal*, *14*(5), 1304–1315.

McMurdie, P.J., and S. Holmes. 2013. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. PloS ONE 8: e61217. doi: 10.1371/journal.pone.0061217

Oksanen, J., F.G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P.R. Minchin, R.B. O'Hara, G.L. Simpson, P. Solymos, M.H.M. Stevens, E. Szoecs, and H. Wagner. 2019. vegan: Community ecology package. R package version 2.5-6. https://CRAN.R-project.org/package=vegan

Parada, A.E., D.M. Needham, and J.A. Fuhrman. 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. doi: 10.1111/1462-2920.13023

Parks, D. H., Chuvochina, M., Waite, D. W., Rinke, C., Skarshewski, A., Chaumeil, P.-A., & Hugenholtz, P. (2018). A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nature Biotechnology*, *36*(10), 996–1004.

Quast, C., E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies, F.O. Glöckner. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 41: D590-D596. doi:10.1093/nar/gks1219

Schliep, K.P. 2011. phangorn: phylogenetic analysis in R. Bioinformatics 27: 592-593. doi: 10.1093/bioinformatics/btq706

Wright, E.S. 2016. Using DECIPHER v2.0 to analyze big biological sequence data in R. The R Journal 8: 352-359.

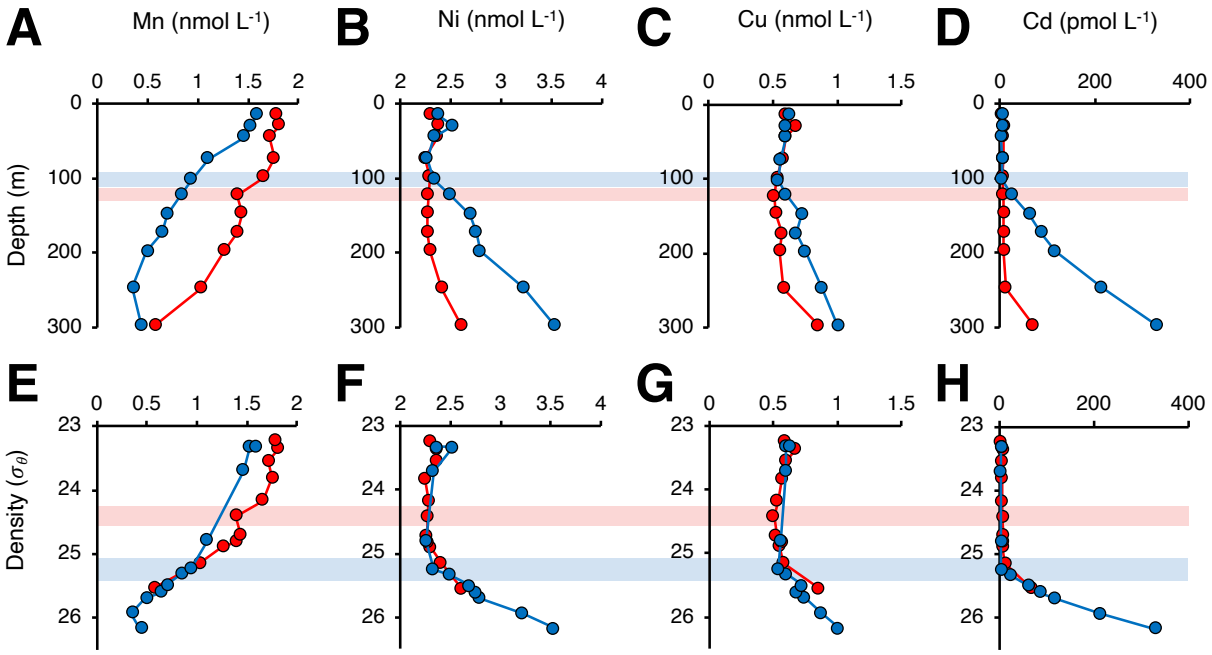

Figure S1. Density driven variation in Mn (a, e), Ni (b, f), Cu (c, g) and Cd (d, h) for the cyclonic eddy (blue) and anticyclonic eddy (red). Top panels (a-d) are plotted against depth and bottom panels (e-h) are plotted against potential density with the DCM layers for each eddy highlighted in blue and red shading. See Fig. 3 in the main text.

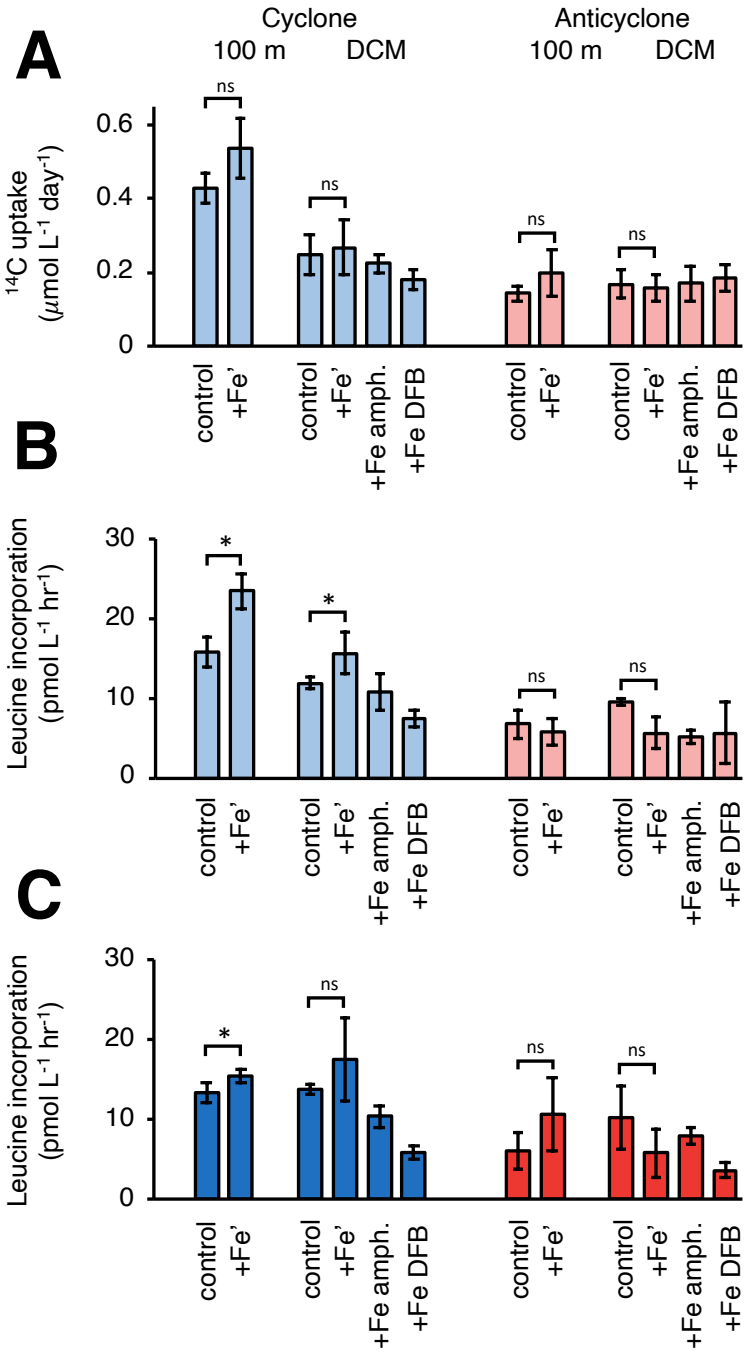

Figure S2. Fe addition experiments at 100 m and the DCM. ^{14}C primary production (a) and leucine incorporation under ambient light (b) and in darkness (c) for the cyclonic eddy (blue) and anticyclonic eddy (red). Statistical comparisons for 'control' and '+Fe' treatments reflect pairwise comparisons (Student's t-test) at 100 m and one way ANOVAs with Fisher least significant difference (LSD) tests for DCM experiments (light and dark experiments considered separately). Asterisks (*) denote $p < 0.05$, 'ns' denotes $p > 0.05$.

Table S1. Isopycnal analysis of biogeochemical parameters with SLA_{corr} at Station ALOHA using the analyses of Barone et al. (2019). Median values for each bin are shown for all fields. NA reflects no measurements available.

24.50

SLA bin cm	SLA cm	Depth m	NO_3 μmol L^{-1}	Si μmol L^{-1}	$Si - NO_3$ μmol L^{-1}	Chl μg L^{-1}	O_2 μmol L^{-1}	^{14}C uptake μmol L^{-1} day^{-1}
< -10	-13.4	99	0.04	1.53	1.368	0.19	217	0.28
-10 to -5	-7.3	102	0.19	1.57	1.323	0.19	214	0.12
-5 to -1'	-3.5	122	0.30	1.53	1.1695	0.21	208	0.14
-1 to +1'	-0.1	125	0.54	1.59	1.0835	0.16	204	0.07
+1 to +5	3.4	132	0.57	1.70	1.1665	0.13	206	0.06
+5 to +10	7.2	149	0.65	1.57	0.962	0.11	203	0.07
> +10	12.8	157	0.60	1.75	1.092	0.08	202	0.04
> +10	12.8	213	2.51	2.88	0.56	0.01	197	NA

24.75

SLA bin cm	SLA cm	Depth m	NO_3 μmol L^{-1}	Si μmol L^{-1}	$Si - NO_3$ μmol L^{-1}	Chl μg L^{-1}	O_2 μmol L^{-1}	^{14}C uptake μmol L^{-1} day^{-1}
< -10	-13.4	125	0.52	1.81	1.38	0.20	209	0.13
-10 to -5	-7.3	131	0.85	1.88	0.99	0.12	204	0.10
-5 to -1'	-3.5	144	1.18	1.91	0.80	0.10	202	0.06
-1 to +1'	-0.1	151	1.26	2.04	0.75	0.06	200	0.03
+1 to +5	3.4	166	1.34	2.09	0.82	0.05	200	0.00
+5 to +10	7.2	175	1.31	2.01	0.66	0.04	197	NA
> +10	12.8	188	1.33	2.24	0.73	0.04	198	NA

25.00

SLA bin cm	SLA cm	Depth m	NO_3 μmol L^{-1}	Si μmol L^{-1}	$Si - NO_3$ μmol L^{-1}	Chl μg L^{-1}	O_2 μmol L^{-1}	^{14}C uptake μmol L^{-1} day^{-1}
< -10	-13.4	150	1.72	2.45	0.75	0.07	201	0.01
-10 to -5	-7.3	164	1.88	2.64	0.51	0.04	200	0.01
-5 to -1'	-3.5	175	2.19	2.50	0.40	0.04	198	NA
-1 to +1'	-0.1	177	2.30	2.73	0.34	0.02	196	0.01
+1 to +5	3.4	200	2.41	2.66	0.32	0.02	195	NA
+5 to +10	7.2	201	2.27	2.70	0.26	0.02	196	NA
> +10	12.8	213	2.51	2.88	0.56	0.01	197	NA