Figures and figure supplements

Emergence of trait variability through the lens of nitrogen assimilation in Prochlorococcus

Paul M Berube et al
Figure 1. Distributions of the nitrate reductase (narB) and nitrite reductase (nirA) genes across a core marker gene phylogeny of 329 Prochlorococcus and Synechococcus genomes. Selected Prochlorococcus clades are highlighted as pie slices. Red bars indicate genomes with the potential to use nitrate as a nitrogen source.

Reference Culture

Reference HLIII and HLIV single cells were not screened by PCR for narB

Berube et al. eLife 2019;8:e41043. DOI: https://doi.org/10.7554/eLife.41043
nitrate based on the presence/absence of a narB gene. Blue bars indicate genomes with an annotated nirA gene in the genome assembly. The outer ring indicates the estimated percentage of the genomes recovered as a gray bar chart. Reference culture genomes are indicated by gray circles and the results of the PCR screen for narB are indicated by filled (present) and open (absent) squares. Single cells belonging to the HLIII and HLIV clades of Prochlorococcus were not screened by PCR and are only included as additional reference genomes. The nucleotide phylogeny is based on a concatenated alignment of 37 marker genes in the PhyloSift software package and inferred using maximum likelihood in RAxML (GTRCAT model) with automatic bootstopping criteria (250 replicate trees). Filled purple circles on branches indicate that the associated genomes clustered together in at least 75% of trees.

DOI: https://doi.org/10.7554/eLife.41043.002
Figure 2. Hierarchical clustering of presence and absence distributions for flexible CyCOGs found in 83 Prochlorococcus HLII single cell genomes with genome recoveries of at least 75% (median 90%). Genomes are sorted by Atlantic and Pacific Oceans and by the presence/absence of the narB gene – a marker for the capacity to assimilate nitrate. Other than genes in the nitrate assimilation gene cluster (green box), no CyCOGs were over- or under-represented among the flexible genes in genomes containing narB. Genomes from the Atlantic, regardless of whether or not they contained the narB marker gene, were enriched in phosphorus assimilation genes (blue box).

DOI: https://doi.org/10.7554/eLife.41043.004
Hierarchical clustering of presence and absence distributions for flexible CyCOGs found in 22 Prochlorococcus single cell genomes belonging to the LLI clade with genome recoveries of at least 75% (median 87%). Genomes are sorted by Atlantic and Pacific Oceans and by the presence/absence of the narB gene – a marker for the capacity to assimilate nitrate. Other than genes in the nitrate assimilation gene cluster (green box), no CyCOGs were over- or under-represented among the flexible genes in genomes containing narB.

DOI: https://doi.org/10.7554/eLife.41043.005
Figure 3. The core marker protein phylogeny of Prochlorococcus and Synechococcus. (A) in comparison to phylogenies for the nitrate/nitrite transporter, NapA (B), the nitrate reductase, NarB (C), the molybdopterin biosynthesis protein, MoaA (D), the nitrite transporter, FocA (E), and the nitrite reductase, NirA (F). Figure 3 continued on next page.
nitrite reductase, NirA (F). Interclade horizontal gene transfer is minimal for genes encoding proteins in the upstream half of the nitrate assimilation pathway (B–D) since clades defined by the core phylogeny (A) remain separate. Horizontal gene transfer is observed in a few instances for genes encoding proteins in the downstream half of the nitrate assimilation pathway (E–F). One single cell (AG-363-P06, brown circle) from the high-light adapted HLVI clade possesses FocA and NirA proteins most similar to those from low-light adapted Prochlorococcus. Prochlorococcus belonging to the LLI clade possess two types of NirA as indicated by well supported phylogenetic divergence of NirA among LLI cells. Filled purple circles on branches indicate that the associated taxa clustered together in at least 75% of trees. Scale bars are 0.1 changes per site.

DOI: https://doi.org/10.7554/eLife.41043.008
Figure 3—figure supplement 1. The core marker gene phylogeny of Prochlorococcus and Synechococcus (A) in comparison to gene phylogenies for the nitrate/nitrite transporter, napA (B), the nitrate reductase, narB (C), the molybdopterin biosynthesis protein, moaA (D), the nitrite transporter, focA (E), and the nitrite reductase, nirA (F). The phylogenetic groups defined by core marker phylogeny are indicated for each gene.
Figure 3—figure supplement 1 continued

(E), and the nitrite reductase, nirA (F). Filled purple circles on branches indicate that the associated taxa clustered together in at least 75% of trees. Scale bars are 0.1 changes per site.

DOI: https://doi.org/10.7554/eLife.41043.009
Figure 4. The gene order and genomic location of the nitrate assimilation gene cluster in high-light and low-light adapted Prochlorococcus genomes. The location of the nitrate and nitrite assimilation genes is shown relative to conserved core marker genes (black bars). The proportion of genomes in each clade with the indicated location is shown next to the clade names. The percentage of genomes in each group with a specific gene content and order is shown to the left of each gene order plot.

DOI: https://doi.org/10.7554/eLife.41043.011
Figure 4—figure supplement 1. Mauve alignments visualized for representative contigs from HLII Prochlorococcus single cell genome assemblies in comparison to the reference genome Prochlorococcus AS9601 (HLII, non-nitrate assimilating). The nitrate assimilation gene cluster is found in a local collinear block shared with AS9601. CyCOG_60001297 and DNA polymerase I genes are marked as reference points.
DOI: https://doi.org/10.7554/eLife.41043.012
Figure 4—figure supplement 2. Mauve alignments visualized for representative contigs from HLJ and HLVI Prochlorococcus single cell genome assemblies in comparison to the reference genomes *Prochlorococcus* MED4 (HLI) and *Prochlorococcus* AS9601 (HLII). The reference genomes do not contain the nitrate assimilation gene cluster. In single cells, this cluster is found in the genomic island ISL1 (sensu Kettler et al., 2007).

DOI: https://doi.org/10.7554/eLife.41043.013
Figure 4—figure supplement 3. Mauve alignments visualized for representative contigs from LLI Prochlorococcus single cell genome assemblies in comparison to the reference genome Prochlorococcus NATL2A (LLI; nitrite assimilation only). The nitrate assimilation gene cluster is found in a local collinear block shared with NATL2A. The pyrG and ppk genes are marked as reference points.

DOI: https://doi.org/10.7554/eLife.41043.014
Figure 5. Representative phylogenetic patterns for DNA gyrase subunit B (gyrB) in comparison to nitrate reductase (narB) and the phosphate assimilation gene pstB (encoding the ABC transporter ATP binding subunit) from single cells in surface populations at HOT (Hawai‘i Ocean Time-series) and BATS (Bermuda Atlantic Timeseries Study). The gyrB and narB genes do not exhibit significant phylogenetic divergence between the two sites. The pstB gene, in contrast, has significantly diverged into Atlantic-like and Pacific-like clusters of sequences due to frequent recombination, gene transfer, and/or selection (sensu Coleman and Chisholm, 2010).

DOI: https://doi.org/10.7554/eLife.41043.017
Figure 6. Percent nucleotide difference of selected genes for the HLII (A) and LLI (B) clades of Prochlorococcus. The ‘core’ genes are a concatenated alignment of up to 37 PhyloSift marker genes. Genes in the nitrate assimilation cluster are in gray. Center lines show the medians, box limits indicate the 25th and 75th percentiles, whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles, and outliers are represented by dots. For HLII genes (A), n = 5151, 4186, 3321, 3570, 4371, 3240, 378, 351, 351, 351 sample points. For LLI genes (B), n = 496, 406, 435, 435, 406, 406, 153, 136, 136, 171, 210, 28 sample points.
DOI: https://doi.org/10.7554/eLife.41043.020
Figure 7. Proposed model mapping the vertical inheritance of the nitrate assimilation pathway onto a general model of speciation in Prochlorococcus (Braakman et al., 2017). We assume different cost:benefit ratios at top and bottom of the water column based on the energetic requirements of the nitrate assimilation pathway and the selective advantage of carrying this trait in environments with low nitrogen availability (see main text for further discussion). We further assume that ancestral Prochlorococcus, similar to Synechococcus, were capable of nitrate assimilation. As new Prochlorococcus clades/ecotypes emerged to more efficiently harness light energy and facilitate the draw-down of nitrogen at the surface, basal lineages were partitioned to the deeper regions of the euphotic zone (Braakman et al., 2017). In these basal lineages, higher relative costs of the pathway combined with access to other nitrogen sources (e.g. amino acids) hastened the loss of this trait through stochastic gene loss. In more recently emerging lineages (LLI, HL I-III), the trait has been retained with intra-clade frequencies influenced by the specific chemical and physical characteristics of the environment in which they are found (Berube et al., 2016). The founder effect has driven punctuated changes (e.g. genome-wide rearrangements) during speciation, while homologous recombination has acted to constrain the divergence of gene sequence and order within clades.

DOI: https://doi.org/10.7554/eLife.41043.028
Ancestral Prochlorococcus Node

Ancestral Prochlorococcus Node

Ancestral Prochlorococcus Node

Proportional Likelihood at Ancestral Node:
0.65 absent
0.35 present

Prochlorococcus clade LLIV

Prochlorococcus subcluster 5.1

Synechococcus RCC307

Figure 7—figure supplement 1. Reconstruction of the state of the ancestral Prochlorococcus node with regards to the presence or absence of narB. When parsimony is used to infer whether or not ancestral Prochlorococcus possessed the nitrate assimilation trait (A, B), the results are highly dependent on the relative costs associated with the gain or loss of the trait. If these costs are equally weighted (A), all changes are estimated to occur at the leaves of the tree and consist primarily of acquisition events – this explanation seems unlikely given the conservation of gene order and location.
within clades (Figure 4). But, if a bias towards loss of the nitrate assimilation trait is imposed, the more parsimonious explanation is that ancestral Prochlorococcus possessed this trait (B). We further note that maximum likelihood estimates a reasonable probability that ancestral Prochlorococcus were capable of nitrate assimilation (C). Further, maximum likelihood generally supports a higher rate of loss relative to gain of the nitrate assimilation trait in Prochlorococcus given that the estimated forward (trait gain) rate was 14.5 and the estimated backward (trait loss) rate was 27.4 in this asymmetrical model.

DOI: https://doi.org/10.7554/eLife.41043.029