Niche of harmful alga Aureococcus anophagefferens revealed through ecogenomics

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

Citation

As Published
http://dx.doi.org/10.1073/pnas.1016106108

Publisher
National Academy of Sciences (U.S.)

Version
Final published version

Accessed
Sat Apr 06 16:05:00 EDT 2019

Citable Link
http://hdl.handle.net/1721.1/66233

Terms of Use
Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.

Detailed Terms

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.
Harmful algal blooms (HABs) cause significant economic and ecological damage worldwide. Despite considerable efforts, a comprehensive understanding of the factors that promote these blooms has been lacking, because the biochemical pathways that facilitate their dominance relative to other phytoplankton within specific environments have not been identified. Here, biogeochemical measurements showed that the harmful alga *Aureococcus anophagefferens* outcompeted co-occurring phytoplankton in estuaries with elevated levels of dissolved organic matter and turbidity and low levels of dissolved inorganic nitrogen. We subsequently sequenced the genome of *A. anophagefferens* and compared its gene complement with those of six competing phytoplankton species identified through metaproteomics. Using an ecogenomic approach, we specifically focused on gene sets that may facilitate dominance within the environmental conditions present during blooms. *A. anophagefferens* possesses a larger genome (56 Mbp) and has more genes involved in light harvesting, organic carbon and nitrogen use, and encoding selenium- and metal-requiring enzymes than competing phytoplankton. Genes for the synthesis of microbial deterrents likely permit the proliferation of this species, with reduced mortality losses during blooms. Collectively, these findings suggest that anthropogenic activities resulting in elevated levels of turbidity, organic matter, and metals have opened a niche within coastal ecosystems that ideally suits the unique genetic capacity of *A. anophagefferens* and thus, has facilitated the proliferation of this and potentially other HABs.

This article contains supporting information online at www.pnas.org/cgi/doi/10.1073/pnas.1016106108 and D5Supplemental.


This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

**Data deposition:** The sequence report published in this paper has been deposited in the GenBank database (accession no. AC69900000).

1To whom correspondence may be addressed. E-mail: christopher.gobler@stonybrook.edu or IVGrigoriev@lbl.gov.

2D.B., S.D., and S.W. contributed equally to this work.

This article contains supporting information online at www.pnas.org/cgi/doi/10.1073/pnas.1016106108/D5Supplemental.

---

**Niche of harmful alga *Aureococcus anophagefferens* revealed through ecogenomics**

Christopher J. Gobler,a,b,1 Dianna L. Berry,a,b,2 Sonya T. Dyhrman,2 Steven W. Wilhelm,a,b,2 Asaf Salamov,a Alexei V. Lobanov,a Yan Zhang,a Jackie L. Collier,a Louie L. Wurch,a Adam B. Kustka,a Brian D. Dill,a,2 Manesh Shah,a Nathan C. Verberkmoes,a Alan Kuoa,v Astrid Terry,a Jasmin Pangilinan,a Erika A. Lindquist,a Susan Lucas,a Ian T. Paulsen,j Theresa K. Hattenrath-Lehmann,a,b Stephanie C. Talmage,a,b Elyse A. Walker,a,b, Florian Koch,a,b Amanda M. Burson,a,b MariaAlejandra Marcoval,a,b Ying-Zhong Tang,a,b Gary R. LeCleir4, Kathy J. Coyne6, Gry M. Berg1, Erin M. Bertrand1, Mak A. Saito,a,m, Vadim N. Gladyshev,d, and Igor V. Grigoriev,1

---

Harmful algal blooms (HABs) are caused by phytoplankton that have a negative impact on ecosystems and coastal fisheries worldwide (1–4) and cost the US economy alone hundreds of millions of dollars annually (5). The frequency and impacts of HABs have intensified in recent decades, and anthropogenic processes, including eutrophication, have been implicated in this expansion (1–3). Although there is great interest in mitigating the occurrence of HABs, traditional approaches that have characterized biogeochemical conditions present during blooms do not identify the aspects of the environment that are favorable to an individual algal species. Predicting where, when, and under what environmental conditions HABs will occur has further been inhibited by a limited understanding of the cellular attributes that facilitate the proliferation of one phytoplankton species to the exclusion of others.

*Aureococcus anophagefferens* is a pelagophyte that causes harmful brown tide blooms with densities exceeding 10^6 cells mL^-1 for extended periods in estuaries in the eastern United States and South Africa (6). Brown tides do not produce toxins that poison humans but have decimated multiple fisheries and seagrass beds because of toxicity to bivalves and extreme light attenuation, respectively (6). Brown tides are a prime example of the global expansion of HABs, because these blooms had never been documented before 1985 but have recurred in the United States and South Africa annually since that time (6). Like many other HABs, *A. anophagefferens* blooms in shallow, anthropogenically modified estuaries when levels of light and inorganic nutrients are low and organic carbon and nitrogen concentrations are elevated (1–3). For this study, we used an ecogenomic approach to assess the extent to which the gene set of *A. anophagefferens* may permit its dominance under the environmental conditions present in estuaries during brown tides. We characterized the biogeochemical conditions present in estuaries before, during, and after *A. anophagefferens* blooms. Sequencing this HAB genome (*A. anophagefferens*), we compared its gene content to those of six phytoplankton species identified through metaproteomics to co-occur with this alga during blooms events. Using this ecogenomic approach, we investigated how the gene sets of *A. anophagefferens* differ from the six comparative phytoplankton species and how these differences may affect the ability of *A. anophagefferens* to compete in the physical (e.g., light harvesting), chemical (e.g., nutrients, organic matter, and trace metals), and ecological (e.g., defense against predators and allelopathy) environment present during brown tides.
Results and Discussion

During an investigation of a US estuary, Quantuck Bay, NY, from 2007 to 2009, brown tides occurred annually from May to July, achieving abundances exceeding $10^6$ cells mL$^{-1}$ or $5 \times 10^6$ μm$^3$ mL$^{-1}$ (Fig. 1). *A. anophagefferens* was observed to bloom after spring diatom blooms and outcompeted small (<2 μm) eukaryotic and prokaryotic phytoplankton (e.g., *Ostreococcus* and *Synechococcus*) during summer months (Fig. 1D), a pattern consistent with prior observations (7, 8). Concurrently, dissolved inorganic nitrogen levels were reduced to <1 μM during blooms, whereas dissolved organic nitrogen levels and light extinction were elevated, resulting in a system with decreased light availability and concentrations of dissolved organic nitrogen far exceeding those of dissolved inorganic nitrogen (Fig. 1C). Metaproteomic analyses of planktonic communities were performed to identify phytoplankton that *A. anophagefferens* may compete with during blooms by quantifying organism-specific peptides among the microbial community. Performing such analyses on the plankton present in this estuary highlights the dominance of *A. anophagefferens* and coexistence of the six phytoplankton species for which complete genome sequences have been generated (Fig. 1E): two coastal diatom species, *Phaeodactylum tricornutum* (clone CCMP632) (9) and *Thalassiosira pseudonana* [clone CCMP 1335 (10) isolated from an embayment that now hosts brown tides (6)], and coastal zone isolates of *Ostreococcus* (*O. lucimarinus* and *O. tauri*) (11) and *Synechococcus* [clones CC9311 (12) and CC9902] small eukaryotic and prokaryotic phytoplankton, respectively, (Fig. 1 and Table 1). To assess the extent to which the gene set of *A. anophagefferens* may permit its dominance within the geochemical environment found in this estuary (Fig. 1C), the gene complement of *A. anophagefferens* was determined by genome sequencing and was compared with those of the six competing phytoplankton species (Fig. 1E and Table 1).

Although phytoplankton genome size generally scales with cell size (15, 16), *A. anophagefferens* (2 μm) has a larger genome (56 Mbp) and more genes (~11,590) than the six competing phytoplankton species (2.2–32 Mbp and 2,301–11,242 genes) (Table 1 and SI Appendix, Tables S1, S2, S3, and S4). Its small cell size and thus larger surface area to volume ratio allows it to kinetically outcompete larger phytoplankton for low levels of light and nutrients (17), whereas its large gene content and more complex
genetic repertoire may provide a competitive advantage over other small phytoplankton with fewer genes. The *A. anophagefferens* genome contains the largest number of unique genes relative to the six competing phytoplankton examined here (209 vs. 12–79 unique genes) (Table 1). Many of these unique or enriched genes in *A. anophagefferens* are associated with light harvesting, organic matter use, and metalloenzymes as well as the synthesis of microbial predation and competition deterrents (*SI Appendix*, Tables S5-S8). Twenty-six *A. anophagefferens* LHC genes belong to a group that has only six representatives in *P. tricornutum* (branch PHYMKG in Fig. 3 and *SI Appendix*, Fig. S1) but are similar to the multicellular brown macroalgae, *Ectocarpus siliculosus* (20). Similar LHC genes in the microalgae *Emiliania huxleyi* have recently been shown to be up-regulated under low light (21). We hypothesize that these LHC genes encode the major light-harvesting proteins for *A. anophagefferens* and that the enrichment of these proteins imparts a competitive advantage in acquiring light under the low-irradiance conditions that prevail during blooms (Fig. 1C).

**Light Harvesting.** Phytoplankton rely on light to photosynthetically fix carbon dioxide into organic carbon, but the turbid, low-light environment characteristic of estuaries and intense shading during dense algal blooms (Fig. 1 B and C) can strongly limit photosynthesis. *A. anophagefferens* is better adapted to low light than the comparative phytoplankton species, which requires at least threefold higher light levels to achieve maximal growth rates (Fig. 2A). Its genome contains the full suite of genes involved in photosynthesis, including 62 genes encoding light-harvesting complex (LHC) proteins (Fig. 2A). This is 1.5–3 times more than other eukaryotic phytoplankton sequenced thus far (Fig. 2A and *SI Appendix*, Table S7) and a feature that likely enhances adaptation to low and/or dynamic light conditions found in turbid estuaries. LHC proteins bind antenna chlorophyll and carotenoid pigments that augment the light-capturing capacity of the photosynthetic reaction centers (18, 19). Twenty-six *A. anophagefferens* LHC genes belong to a group that has only six representatives in *T. pseudonana* and one representative in *P. tricornutum* (branch PHYMKG in Fig. 3 and *SI Appendix*, Fig. S1) but are similar to the multicellular brown macroalgae, *Ectocarpus siliculosus* (20). Similar LHC genes in the microalgae *Emiliania huxleyi* have recently been shown to be up-regulated under low light (21). We hypothesize that these LHC genes encode the major light-harvesting proteins for *A. anophagefferens* and that the enrichment of these proteins imparts a competitive advantage in acquiring light under the low-irradiance conditions that prevail during blooms (Fig. 1C).

**Organic Matter Use.** In addition to being well-adapted to low light, *A. anophagefferens* also outcompetes other phytoplankton in estuaries with elevated organic matter concentrations (6) (Fig. 1C), and can survive extended periods with no light (22). Consistent with these observations, the genome of *A. anophagefferens* contains a large number of genes that may permit the degradation of organic compounds to support heterotrophic metabolism. For example, its genome encodes proteins involved in the transport of oligosaccharides and sugars that are not found in competing phytoplankton, including genes for glycerol, glucose, and d-xylose uptake (*SI Appendix*, Table S8). The *A. anophagefferens* genome also encodes more nucleoside sugar transporters and major facilitator family sugar transporters than other comparative phytoplankton species (*SI Appendix*, Table S8). It is highly enriched in genes associated with the degradation of polysaccharides (*A. anophagefferens* possesses 47 sulfatase genes, including those targeting sulfonated polysaccharides such as glucosamine-(N-acetyl)-6-sulfatases, whereas the diatoms contain a total of three to four sulfatases and the comparative picoplankton contain none (*SI Appendix*, Table S9). *A. anophagefferens* also possesses many more genes involved in carbohydrate degradation than competing phytoplankton (85 vs. 4–29 genes in comparative phytoplankton), including 29 such genes present only in *A. anophagefferens* (Fig. 4 and *SI Appendix*, Tables S10 and S11). Collectively, these genes ( *SI Appendix*, Tables S9, S10, S11, and S12) provide this alga with unique metabolic capabilities regarding the degradation of an array of organic carbon compounds, many of which may not be accessible to other phytoplankton. In an ecosystem setting, such a supplement of organic carbon would be critical for population proliferation within the low-light environments present in estuaries, particularly during dense algal blooms (Fig. 1C).

Table 1. Major features of the genomes of *A. anophagefferens* and six competing algal species: *P. tricornutum* (9), *T. pseudonana* (10), *O. tauri* (11), *O. lucimarinus* (11), *Synechococcus* (CC9311) (12), and *Synechococcus* (CC9902)

<table>
<thead>
<tr>
<th></th>
<th><em>A. anophagefferens</em></th>
<th><em>P. tricornutum</em></th>
<th><em>T. pseudonana</em></th>
<th><em>O. tauri</em></th>
<th><em>O. lucimarinus</em></th>
<th><em>Synechococcus</em> (CC9311)</th>
<th><em>Synechococcus</em> (CC9902)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell diameter (µm)</td>
<td>2.0</td>
<td>11.0</td>
<td>5.0</td>
<td>1.2</td>
<td>1.3</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Cell volume (µm³)</td>
<td>6</td>
<td>61</td>
<td>88</td>
<td>1.8</td>
<td>2.0</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>Genome size (Mb)</td>
<td>57</td>
<td>27</td>
<td>32</td>
<td>13</td>
<td>13</td>
<td>2.6</td>
<td>2.2</td>
</tr>
<tr>
<td>Predicted gene number</td>
<td>11,501</td>
<td>10,402</td>
<td>11,242</td>
<td>7,892</td>
<td>7,651</td>
<td>2,892</td>
<td>3,201</td>
</tr>
<tr>
<td>Genes with known functions</td>
<td>8,560</td>
<td>6,239</td>
<td>6,797</td>
<td>5,090</td>
<td>5,322</td>
<td>1,607</td>
<td>1,469</td>
</tr>
<tr>
<td>Genes with Pfam domains</td>
<td>6,908</td>
<td>5,398</td>
<td>5,791</td>
<td>4,763</td>
<td>4,214</td>
<td>1,636</td>
<td>1,488</td>
</tr>
<tr>
<td>Genes with unique Pfam domains</td>
<td>209</td>
<td>79</td>
<td>75</td>
<td>23</td>
<td>51</td>
<td>55</td>
<td>12</td>
</tr>
</tbody>
</table>

*Genes with known functions were identified using Swiss-Prot, a curated protein sequence database, with an e-value cutoff of <10⁻5 (13). Pfam domains are sequences identified from a database of protein families represented by multiple sequence alignments and hidden Markov models (14). The compressed nature of *P. tricornutum* cells (11 x 2.5 µm) makes its biovolume smaller than *T. pseudonana*."

4354 | www.pnas.org/cgi/doi/10.1073/pnas.1016106108

Gobler et al.
petitors, a hypothesis supported by its dominance in systems with elevated ratios of dissolved organic nitrogen to dissolved inorganic nitrogen and the reduction in dissolved organic nitrogen concentrations often observed during the initiation of brown tides (6, 25).

**Metalloenzymes.** *A. anophagefferens* blooms in shallow, enclosed estuaries (6) where the concentrations of metals and elements like selenium are elevated (26–28), but it never dominates deep estuaries or continental shelf regions (6) that are characterized by lower metal and trace element inventories (26–28). *A. anophagefferens* has a large and absolute requirement for some trace elements, such as selenium (Fig. 2C). In comparison, phytoplankton, such as *Synecochoccus*, do not require this element, whereas others, such as *T. pseudonana* and *P. tricornutum*, have lower selenium requirements for maximal growth (Fig. 2C). The *A. anophagefferens* genome is consistent with these observations, being enriched in numerous classes of proteins that require metals and elements like selenium as cofactors (Fig. 2C). It possesses at least 56 genes encoding selenocysteine-containing proteins, two times the number present in the *O. lucimarinus* genome, which previously had the largest known eukaryotic selenoproteome (11, 29), and fourfold more than the diatom genomes (Fig. 2C). The *A. anophagefferens* selenoproteome includes nearly all known eukaryotic selenoproteins as well as selenoproteins that were previously described only in bacteria (29) and several selenoproteins that have never been described in any other organism (SI Appendix, Table S14). In addition, several selenoprotein families are represented by multiple isoforms (SI Appendix, Table S14). One-half of the selenoproteins are methionine sulfoxide reductases, thioredoxin reductases, glutathione peroxidases, glutaredoxins, and peroxiredoxins (SI Appendix, Table S14). Together, these enzymes help protect cells against oxidative stress in the dynamic and ephemeral conditions present in estuaries through the removal of hydroperoxides and the repair of oxidatively damaged proteins. Moreover, selenocysteine residues are often superior catalytic groups compared with cysteine (30–32), and thus, they allow *A. anopha-
Fig. 4. Genes encoding for enzymes involved in degrading organic carbon compounds in *A. anophagefferens*. The graph displays the portion and names of the genes encoding for functions that are unique to *A. anophagefferens* (red: 53%), enriched in *A. anophagefferens* relative to the six comparative phytoplankton (34%, green), and present at equal or lower numbers in *A. anophagefferens* relative to the six comparative phytoplankton (13%; blue). The number of genes present in multiple copies in *A. anophagefferens* is shown in parentheses. Further details regarding these genes are presented in SI Appendix, Tables S10 and S11.

gers to more efficiently execute multiple metabolic processes and increase its competitiveness relative to other phytoplankton in the anthropogenically modified estuaries where it blooms. The *A. anophagefferens* genome is also enriched in genes encoding for molybdenum-, copper-, and nickel-containing enzymes (Fig. 2C). For example, the *A. anophagefferens* genome includes two times the number of genes encoding molybdenum-containing oxidases found in competing species (6 vs. 1–3 genes) (Fig. 2C and SI Appendix, Tables S15 and S16) and has the largest number of molybdenum-specific transporters (SI Appendix, Table S8). Similarly, *A. anophagefferens* possesses four times more genes that encode copper-containing proteins than its competitors (27 vs. 1–6 genes) (Fig. 2C), including 5 multicopper oxidases and 20 tyrosinase-like proteins (SI Appendix, Tables S15 and S16). Several of the *A. anophagefferens* tyrosinase and multicopper oxidase family proteins are heavily glycosylated (more than four glycosylation sites) (SI Appendix, Table S16) and thus, are likely secretory proteins, whereas the few present in the other comparative algal species are not. These copper-containing enzymes degrade lignin, catalyze the oxidation of phenolics, and can have antimicrobial properties (33, 34) and thus, may provide nutrition or confer protection to *A. anophagefferens* cells. *A. anophagefferens* is also the only phytoplankton species with a homolog of the CtrC copper homeostasis protein, which permits efficient cellular trafficking of this metal (SI Appendix, Table S8). With three nickel-requiring ureases, *A. anophagefferens* has more nickel-containing enzymes than other comparative phytoplankton (Fig. 2B and C). Consistent with its ecogonomic profile, these ureases allow *A. anophagefferens* to meet its daily N demand from urea, whereas other phytoplankton do not (35). Perhaps to support the synthesis and use of ureases, *A. anophagefferens* is the only comparative phytoplankton species with a high-affinity nickel transporter (HoxN) (36). *A. anophagefferens* is not universally enriched in metalloenzymes, because other phytoplankton contain equal numbers of cobalt-containing enzymes (Fig. 2C). However, the formation of blooms exclusively in shallow estuaries ensures that *A. ano-

**Microbial Defense.** Although genes associated with the adaptation to low light, the use of organic matter, and metals permit *A. anophagefferens* to dominate a specific geochemical niche found within estuaries, genes involved in the production of compounds that inhibit predators and competitors may further promote blooms (2). Although specific toxins have yet to be identified in *A. anophagefferens*, it is grazed at a low rate during blooms (2, 6), and its genome contains two to seven times more genes involved in the synthesis of secondary metabolites than the comparative phytoplankton genomes (SI Appendix, Fig. S2). *A. anophagefferens* also possesses a series of genes involved in the synthesis of putative antimicrobial compounds that are largely absent from the competing phytoplankton species (SI Appendix, Table S17). For example, *A. anophagefferens* has five berberine bridge enzymes involved in the synthesis of toxic isosiquinoline alkaloids (38, 39) (SI Appendix, Table S17). *A. anophagefferens* uniquely possesses a membrane attack complex gene and multiple phenazine biosynthetase genes (SI Appendix, Table S17) that encode enzymes that may provide defense against microbes and/or protistan grazers (40, 41). There are two to fourfold more ATP-binding cassette (ABC) transporters in *A. anophagefferens* compared with competing species (112 vs. 30–54 ABC transporters) (SI Appendix, Table S8), and it is specifically enriched in ABC multidrug efflux pumps that protect cells from toxic xenobiotics and endogenous metabolites (42, 43). Finally, the *A. anophagefferens* genome encodes 16-fold more Sel-1 genes (130 vs. 0–8 genes) (Table S6), 4-fold more ion channels (82 vs. 1–19 ion channels) (SI Appendix, Table S8), 4-fold more protein kinases, and 2-fold more WD40 domain genes than other phytoplankton (SI Appendix, Table S6). These genes may collectively mediate elaborate cell signaling and sensing by dense bloom populations (44–46), processes that would be important for detecting competitors, predators, other *A. anophagefferens* cells, and the environment. Together, genes involved in the synthesis of microbial deterrents, export of toxic compounds, and cell signaling may contribute to the proliferation of this species with reduced population losses and thus, assist in promoting these HABs (2).

**Conclusions.** The global expansion of human populations along coastlines has led to a progressive enrichment in turbidity (47), organic matter, including organic nitrogen (1, 47, 48), and metals (26, 28) in estuaries. Matching the expansion of HAB events around the world in recent decades, *A. anophagefferens* blooms were an unknown phenomenon before 1985 but have since become chronic, annual events in US and South African estuaries (6), with the potential for further expansion. The unique gene complement of *A. anophagefferens* encodes a disproportionately greater number of proteins involved in light harvesting and organic matter use as well as metal and selenium-requiring enzymes relative to competing phytoplankton. Collectively, these genes reveal a niche characterized by conditions (low light, high organic matter, and elevated metal levels) that have become increasingly prevalent in anthropogenically modified estuaries, suggesting that human activities have enabled the proliferation of these HABs. In estuaries that host *A. anophagefferens* blooms, anthropogenic nutrient loading promotes algal growth and as a result, elevated levels of organic matter and turbidity (6), whereas high concentrations of metals have been attributed to maritime paints and some fertilizers (27, 49). Collectively, these findings establish a context within which to prevent and control HABs, specifically by ameliorating anthropogenically altered aspects of marine environments that harmful phytoplankton are genomically predisposed to exploit. Like *A. anophagefferens*, many HAB-forming dinoflagellates are known to exploit organic
forms of carbon and nitrogen for growth (1–4), grow well under low light (50), and have elevated requirements of copper, molybdenum, and selenium (51, 52). Continued ecogenomic analyses of HABs will reveal the extent to which these events can be attributed to human activities that have transformed coastal ecosystems to suit the genetic capacity of these algae.

**Materials and Methods**

The environmental conditions and plankton community composition within a brown tide-prone estuary (Quantuck Bay, NY) were monitored biweekly from spring to fall of 2007, 2008, and 2009. Nutrient levels were assessed by wet chemical and combustion techniques, whereas the composition of the plankton community was assessed by immunofluorescent assays, flow cytometry, and standard microscopy. Metaproteomes were generated using 2D gel electrophoresis, proteomics analysis of fractionated samples, and MALDI-TOF analysis. The performance of the metabolic model was evaluated by least-squares regression analysis. The model was then validated using a holdout group of experimental data. The model was then used to predict the potential impact of climate change on brown tide dynamics.