

## REPORT

## BIOSYNTHESIS

# Biosynthesis of the neurotoxin domoic acid in a bloom-forming diatom

John K. Brunson<sup>1,2\*</sup>, Shaun M. K. McKinnie<sup>1\*</sup>, Jonathan R. Chekan<sup>1</sup>, John P. McCrow<sup>2</sup>, Zachary D. Miles<sup>1</sup>, Erin M. Bertrand<sup>2,3</sup>, Vincent A. Bielinski<sup>4</sup>, Hanna Luhavaya<sup>1</sup>, Miroslav Oborník<sup>5</sup>, G. Jason Smith<sup>6</sup>, David A. Hutchins<sup>7</sup>, Andrew E. Allen<sup>2,8,†</sup>, Bradley S. Moore<sup>1,9,†</sup>

Oceanic harmful algal blooms of *Pseudo-nitzschia* diatoms produce the potent mammalian neurotoxin domoic acid (DA). Despite decades of research, the molecular basis for its biosynthesis is not known. By using growth conditions known to induce DA production in *Pseudo-nitzschia multiseries*, we implemented transcriptome sequencing in order to identify DA biosynthesis genes that colocalize in a genomic four-gene cluster. We biochemically investigated the recombinant DA biosynthetic enzymes and linked their mechanisms to the construction of DA's diagnostic pyrrolidine skeleton, establishing a model for DA biosynthesis. Knowledge of the genetic basis for toxin production provides an orthogonal approach to bloom monitoring and enables study of environmental factors that drive oceanic DA production.

Oceanic harmful algal blooms (HABs) are intensifying in frequency and severity in association with climate change (1–3). This phenomenon was exemplified in the summer of 2015 when the North American Pacific coast experienced the largest HAB ever recorded, spanning the Aleutian Islands of Alaska to the Baja peninsula of Mexico (fig. S1) (4). This bloom caused widespread ecological and economic devastation, resulting in the deaths of marine mammals and the closure of beaches and fisheries. The dominant algal species in the 2015 HAB were pennate diatoms of the globally distributed genus *Pseudo-nitzschia*, which are often associated with high production of the excitatory glutamate receptor agonist domoic acid (DA; 1). Mammalian consumption of DA-contaminated shellfish exerts its toxicity at the AMPA and

kainate ionotropic glutamate receptors of the central nervous system (5, 6). In humans, a high, single-exposure dose of DA can cause amnesic shellfish poisoning (ASP), which involves symptoms of amnesia, seizures, coma, and in extreme cases, death. Even chronic, low-level consumption of DA may lead to kidney damage, cognitive

deficit, and impairment of fetal development, making DA outbreaks an important human health problem (7–10). Similar neurotoxic symptoms have been observed in birds and marine mammals such as sea lions, which suffer spatial memory impairment linked to DA consumption, likely leading to increases in sea lion strandings (11).

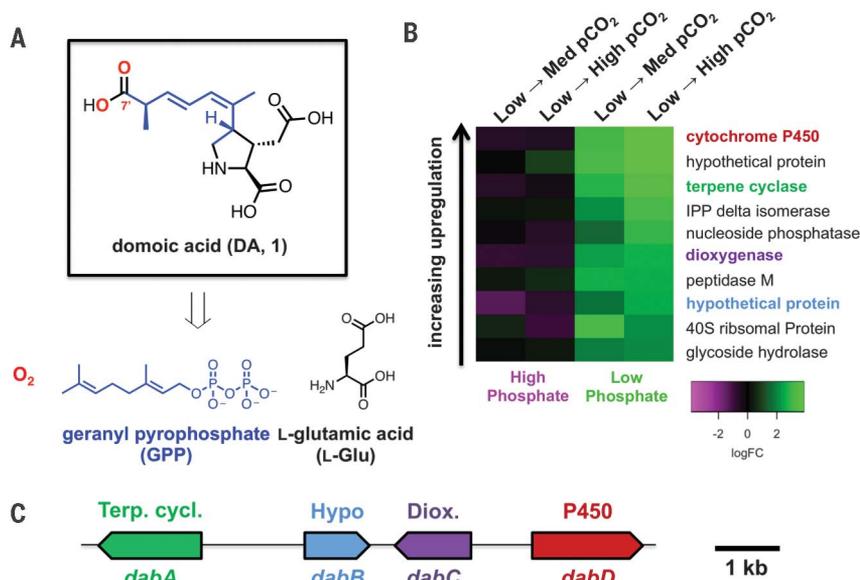
Although abiotic (12) and biotic (13) factors have been shown to affect toxicity in culture, the biological and physicochemical mechanisms underlying DA production in *Pseudo-nitzschia* are unclear. Moreover, not all *Pseudo-nitzschia* blooms produce DA (12). An understanding of the genetic basis of DA biosynthesis in diatoms would facilitate determination of the cellular pathway that controls oceanic DA production and thereby diatom toxicity.

Stable isotope experiments (14, 15) suggest that DA is composed of glutamic acid (Glu) and geranyl pyrophosphate (GPP) building blocks. No pathway that involves these starting materials to construct the characteristic pyrrolidine ring of DA has been previously characterized, and thus, our ability to predict candidate biosynthetic genes using traditional bioinformatic methods was limited. However, we postulated that a redox enzyme, such as a cytochrome P450 (CYP450), generates the 7'-carboxylic acid of DA through three successive oxidations, a reaction common in all branches of life (Fig. 1A) (16).

To identify putative DA biosynthetic genes, we examined patterns of transcriptional activity under previously established conditions—phosphate limitation and elevated CO<sub>2</sub> (17)—that stimulate DA production. We created a differential expression dataset for nearly 20,000 *Pseudo-nitzschia multiseries* RNA transcripts (table S1). A small

<sup>1</sup>Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92093, USA. <sup>2</sup>Microbial and Environmental Genomics Group, J. Craig Venter Institute, La Jolla, CA 92037, USA. <sup>3</sup>Department of Biology, Dalhousie University, Halifax, Nova Scotia B3H 4R2, Canada. <sup>4</sup>Synthetic Biology and Bioenergy Group, J. Craig Venter Institute, La Jolla, CA 92037, USA. <sup>5</sup>Institute of Parasitology, University of South Bohemia and Biology Center CAS, Branišovská 31, 370 05 České Budějovice, Czech Republic. <sup>6</sup>Moss Landing Marine Laboratories, 8272 Moss Landing Road, Moss Landing, CA 95039, USA. <sup>7</sup>Marine and Environmental Biology, Department of Biological Sciences, University of Southern California, Los Angeles, CA 90089, USA. <sup>8</sup>Integrative Oceanography Division, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92037, USA. <sup>9</sup>Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, CA 92093, USA. \*These authors contributed equally to this work.

†Corresponding author. Email: bsmoore@ucsd.edu (B.S.M.); aallen@ucv.org (A.E.A.)



**Fig. 1. Identification of domoic acid biosynthetic genes from transcriptomic and genomic data.**

(A) Structure of DA (1) and its proposed biosynthetic building blocks (14, 15). (B) Ten most up-regulated *P. multiseries* transcripts under previously reported DA-induction conditions (17). Differential expression changes are each shown with reference to the P<sub>CO2</sub> conditions listed across the top of the heatmap, with phosphate concentrations held constant. (C) DA biosynthesis (*dab*) gene cluster in the *P. multiseries* genome.

fraction (~500; 2.5%) showed consistent up-regulation under phosphate limitation (fig. S2 and table S2), and further refinement of this subset to include genes that were also up-regulated with increasing partial pressure of CO<sub>2</sub> (*P*<sub>CO<sub>2</sub></sub>) highlighted only 43 (0.22%) transcripts (table S3). A CYP450 gene showed the highest fold change of all analyzed transcripts under the high *P*<sub>CO<sub>2</sub></sub> and low-phosphate DA-inducing condition (Fig. 1B). This transcript was the only annotated CYP450 gene out of 20 total within the *P. multiseri* genome (220 Mbp; Joint Genome Institute accession no. PRJNA32659) that showed increased transcription (fig. S3 and table S4). When mapped to the public *P. multiseri* genome, this CYP450 was localized to a compact genomic island spanning ~8 kb that possesses three other similarly up-regulated genes that are indicative of canonical gene clustering more typically observed in bacterial- and fungal-specialized metabolism (Fig. 1C) (18). Genomic organization in diatoms is not generally typified by clustering of metabolic genes (19, 20). However, clusters of transcriptionally coregulated genes, sensitive to Fe and Si limitation, have been reported (21, 22). The annotated gene functions suggest involvement in terpenoid and redox biochemistry, which are two predicted hallmarks for DA biosynthesis. These gene candidates for DA biosynthetic (Dab) enzymes were annotated as *dabA* (terpene cyclase), *dabB* (hypothetical protein), *dabC* [ $\alpha$ -ketoglutarate ( $\alpha$ KG)-dependent dioxygenase], and *dabD* (CYP450) (Fig. 1C). We independently sequenced this gene cluster from an environmental isolate of *P. multiseri* in order to validate its conservation (GenBank accession no. MH202990).

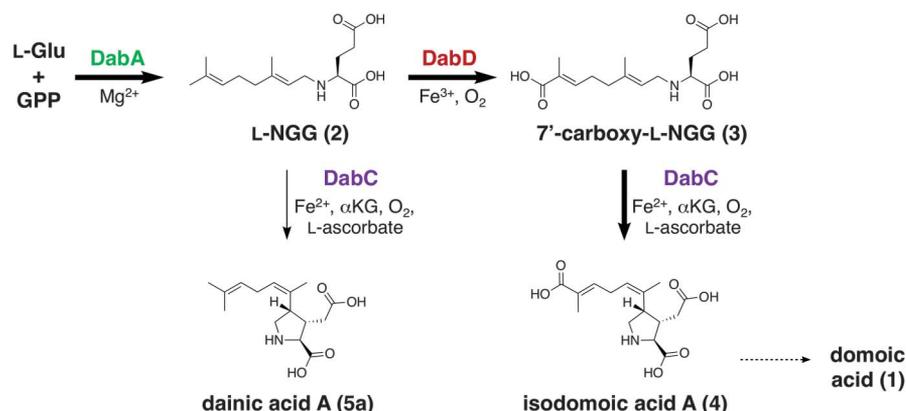
Despite the rarity of the pyrrolidine kainoid skeleton in nature, the putative *dab* gene functions enabled us to initially map our genes onto the suggested biosynthetic pathway (14, 15). We hypothesized that DabA catalyzes the first committed step of DA biosynthesis: *N*-prenylation of L-Glu with GPP to form *N*-geranyl-L-glutamic acid (L-NGG; **2**). DabC and DabD would perform subsequent oxidative reactions (Fig. 2).

We began *in vitro* validation of the *dab* genes with *dabA*, which contains a chloroplast transit peptide sequence and an intron but low similarity to any characterized protein in the National Center for Biotechnology Information (NCBI) database. Structural prediction by using Phyre2 (23) suggested that DabA may possess a terpene cyclase-like fold, and phylogenetic analysis intimated an evolutionary history related to red algal and bacterial genes (fig. S4). Among diatoms, *dabA* appears restricted to *Pseudo-nitzschia* spp. We expressed recombinant His<sub>6</sub>-DabA without the N-terminal transit peptide in *Escherichia coli* and purified the enzyme using Ni<sup>2+</sup> affinity and size-exclusion chromatography (fig. S5). *In vitro*, DabA catalyzes the *N*-geranylation of L-Glu to form L-NGG (**2**) in a Mg<sup>2+</sup>-dependent manner (Fig. 3A). DabA interrogation with structurally similar substrates shows modest promiscuity toward prenyl pyrophosphates but high specificity for Glu (fig. S6). Recently, L-NGG was isolated in low abundance from the DA-containing

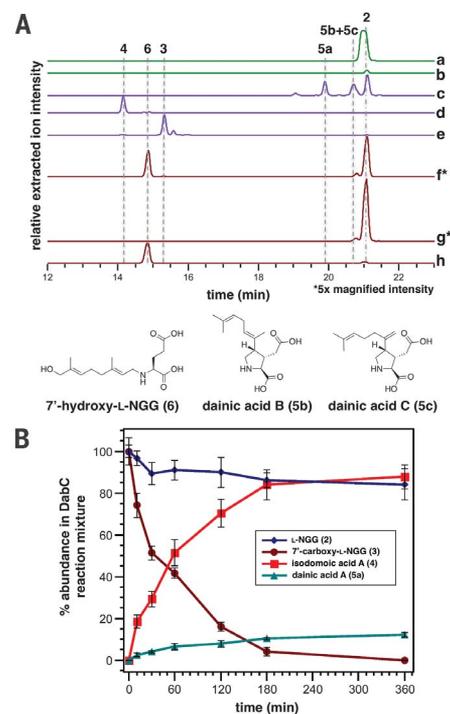
red alga *Chondria armata* and shown through labeling experiments in *P. multiseri* to be a precursor to DA (24). These observations further implicate *dabA* as a gene that encodes a major step in DA biosynthesis.

To investigate the subsequent transformations of L-NGG, we individually interrogated the activity of the DabC and DabD oxygenases toward this substrate. In the presence of Fe<sup>2+</sup>, L-ascorbic acid, and cosubstrate  $\alpha$ KG, recombinant DabC purified from *E. coli* cyclized L-NGG to form three pyrrolidine ring-containing molecules: 7-methylisodomoic acids A, B, and C, termed dainic acids A, B, and C (**5a** to **5c**), respectively (Fig. 3A and figs. S7 and S8), of which **5a** and **5b** had been recently isolated from *C. armata* (24). DabC, however, showed minimal cyclization of other similarly *N*-prenylated glutamic acids (fig. S7).

Enzymatic synthesis of the dainic acids was made more amenable through a one-pot coupled assay using L-Glu, GPP, DabA, DabC, and their requisite cofactors and cosubstrates (fig. S9). Despite DabC generating much of the structural diversity observed within the kainoid metabolites (25), the slow rate of L-NGG consumption and failure to go to completion led us to suspect that this enzyme may instead act on an oxidized substrate, placing DabD oxidation ahead of DabC cyclization in the DA biosynthetic pathway. Incubation of L-NGG with *Saccharomyces cerevisiae* microsomes that possess coexpressed transmembrane proteins DabD (fig. S10) and *P. multiseri* CYP450 reductase (*PmCPR1*) generated small but reproducible quantities of 7'-hydroxy-L-NGG (**6**) and 7'-carboxy-L-NGG (**3**) (Fig. 3A and fig. S11), which was validated through comparison with synthetic



**Fig. 2.** Domoic acid (**1**) biosynthetic pathway based on *dab* gene annotations and *in vitro* enzyme activities.



**Fig. 3.** *In vitro* characterization of Dab enzymes. (A) Relative intensities of negative ionization-extracted ion chromatogram liquid chromatography-mass spectrometry traces [(282.1711, 298.1660, 312.1453, 280.1554, 310.1296) ± 0.01 mass/charge ratio] for *in vitro* Dab reactions. a, DabA, L-Glu (500  $\mu$ M), GPP (500  $\mu$ M), MgCl<sub>2</sub> (10 mM); b, DabA, L-Glu (500  $\mu$ M), GPP (500  $\mu$ M), without MgCl<sub>2</sub>; c, DabC, L-NGG (**2**; 500  $\mu$ M), FeSO<sub>4</sub>,  $\alpha$ KG, L-ascorbic acid; d, DabC, 7'-carboxy-L-NGG (**3**; 500  $\mu$ M), FeSO<sub>4</sub>,  $\alpha$ KG, L-ascorbic acid; e, DabC, 7'-carboxy-L-NGG (**3**; 500  $\mu$ M), FeSO<sub>4</sub>,  $\alpha$ KG, L-ascorbic acid, EDTA (1.0 mM); f, DabD and *PmCPR1* containing *S. cerevisiae* microsomes, L-NGG (**2**; 500  $\mu$ M), NADPH; g, empty vector *S. cerevisiae* microsomes, L-NGG (**2**; 500  $\mu$ M), NADPH; and h, synthetic 7'-hydroxy-L-NGG (**6**) standard. Additional enzyme and cosubstrate concentrations can be found in the supplementary materials. Asterisks indicate that relative extracted ion intensities for traces f and g have been 5 $\times$  magnified. (B) Comparison of *in vitro* DabC substrate consumption of L-NGG (**2**; 1.0 mM) and 7'-carboxy-L-NGG (**3**; 1.0 mM) and relative formation of respective dainic acid A (**5a**) and isodomoic acid A (**4**) products ( $n = 3$  replicate DabC experiments).

standards. Conversely, the dainic acid isomers were not substrates of the DabD P450 (fig. S11). 7'-carboxy-L-NGG incubation with DabC showed rapid cyclization to the *P. multiseriis* natural product isodomoic acid A (4) (Fig. 3A), which did not occur in the presence of metal chelator ethylenediaminetetraacetic acid (EDTA). When comparing both DabC in vitro substrates for physiological relevance, 7'-carboxy-L-NGG was consumed at a much faster rate than L-NGG, further suggesting that 7'-carboxy-L-NGG is an on-pathway intermediate (Fig. 3B and fig. S12). Our cumulative in vitro biochemical results imply a DA biosynthetic pathway that begins with the DabA-catalyzed geranylation of L-Glu to yield L-NGG, likely in the chloroplast. DabD then performs three successive oxidation reactions at the 7'-methyl of L-NGG to produce 7'-carboxy-L-NGG, which is then cyclized by DabC to generate the naturally occurring isodomoic acid A (Fig. 2). A putative isomerase likely converts isodomoic acid A to DA. We biochemically interrogated the coclustered *dabB* gene product but did not observe isomerase activity (fig. S13). Further examination of additionally up-regulated transcripts did not suggest an obvious candidate gene (fig. S14). We are actively investigating this final isomerization reaction to complete the pathway to DA.

In addition to *Pseudo-nitzschia* spp., the kainoid structure has been observed in other diatom, red macroalgal, and fungal compounds (fig. S15) (12, 26, 27). In silico application of the *dab* genes as a kainoid ring biosynthetic query identified *dabA* and *dabC* homologs in several red algal transcriptomes (28, 29), including the known kainic acid producer *Palmaria palmata* (fig. S16). Like DA, kainic acid (26) is a structurally related glutamate agonist constructed from a shorter terpene substrate without the need of the DabD P450 oxidation.

With the establishment of the DabACD-dependent biosynthetic pathway to isodomoic acid A in *P. multiseriis*, we next linked this discovery to the marine environment. We identified *dab* genes with the same genetic organization and high sequence identity in the genome of the known DA-producing *Pseudo-nitzschia multistriata* (fig. S17) (30). Moreover, of the eight publicly available *Pseudo-nitzschia* transcriptomes, only the

highly toxic DA-producing species *Pseudo-nitzschia australis* expressed the *dab* genes (figs. S18 and S19) (31). No *dab* homologs were found in any other sequenced microalgal genera. By virtue of its limited distribution, *dabA* thus presents an opportunity for genetic monitoring of the DA-producing capabilities of *Pseudo-nitzschia* blooms orthogonal to currently established mass spectrometry-based and enzyme-linked immunosorbent assay-based identification approaches. We anticipate that knowledge of the *dab* genes will allow for greater understanding of the basis and diversity of *Pseudo-nitzschia* toxicity, the physiological function of DA, and the environmental conditions that promote HAB formation so that we may better anticipate risk of exposure to this toxic marine natural product.

## REFERENCES AND NOTES

- M. L. Wells et al., *Harmful Algae* **49**, 68–93 (2015).
- S. M. McKibben et al., *Proc. Natl. Acad. Sci. U.S.A.* **114**, 239–244 (2017).
- C. J. Gobler et al., *Proc. Natl. Acad. Sci. U.S.A.* **114**, 4975–4980 (2017).
- R. M. McCabe et al., *Geophys. Res. Lett.* **43**, 10366–10376 (2016).
- G. R. Stewart, C. F. Zorumski, M. T. Price, J. W. Olney, *Exp. Neurol.* **110**, 127–138 (1990).
- J. A. Larm, P. M. Beart, N. S. Cheung, *Neurochem. Int.* **31**, 677–682 (1997).
- J. A. Funk et al., *J. Am. Soc. Nephrol.* **25**, 1187–1197 (2014).
- K. A. Lefebvre et al., *Harmful Algae* **64**, 20–29 (2017).
- J. S. Ramsdell, T. S. Zabka, *Mar. Drugs* **6**, 262–290 (2008).
- L. M. Grattan et al., *Toxins (Basel)* **10**, 103 (2018).
- P. F. Cook et al., *Science* **350**, 1545–1547 (2015).
- A. Lelong, H. Hégaret, P. Soudant, S. S. Bates, *Phycologia* **51**, 168–216 (2012).
- M. P. Sison-Mangus, S. Jiang, K. N. Tran, R. M. Kudela, *ISME J.* **8**, 63–76 (2014).
- U. P. Ramsey, D. J. Douglas, J. A. Walter, J. L. Wright, *Nat. Toxins* **6**, 137–146 (1998).
- T. J. Savage, G. J. Smith, A. T. Clark, P. N. Saucedo, *Toxicol.* **59**, 25–33 (2012).
- X. Zhang, S. Li, *Nat. Prod. Rep.* **34**, 1061–1089 (2017).
- J. Sun et al., *Limnol. Oceanogr.* **56**, 829–840 (2011).
- M. H. Medema et al., *Nat. Chem. Biol.* **11**, 625–631 (2015).
- E. V. Armbrust et al., *Science* **306**, 79–86 (2004).
- C. Bowler et al., *Nature* **456**, 239–244 (2008).
- A. E. Allen et al., *Proc. Natl. Acad. Sci. U.S.A.* **105**, 10438–10443 (2008).
- G. Sapriel et al., *PLOS ONE* **4**, e7458 (2009).
- L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, M. J. E. Sternberg, *Nat. Protoc.* **10**, 845–858 (2015).
- Y. Maeno et al., *Sci. Rep.* **8**, 356 (2018).
- J. Clayden, B. Read, K. R. Hebditch, *Tetrahedron* **61**, 5713–5724 (2005).
- I. Nitta, H. Watase, Y. Tomiie, *Nature* **181**, 761–762 (1958).
- K. Konno, H. Shirahama, T. Matsumoto, *Tetrahedron Lett.* **24**, 939–942 (1983).
- G. W. Saunders, C. Jackson, E. D. Salomaki, *Mol. Phylogenet. Evol.* **119**, 151–159 (2018).
- N. Matasci et al., *Gigascience* **3**, 17 (2014).
- S. Basu et al., *New Phytol.* **215**, 140–156 (2017).
- P. J. Keeling et al., *PLOS Biol.* **12**, e1001889 (2014).

## ACKNOWLEDGMENTS

We acknowledge B. Duggan, A. Mrse, and Y. Su (all University of California, San Diego) for assistance and maintenance of nuclear magnetic resonance and high-resolution mass spectrometry machinery; W. Fenical (Scripps Institution of Oceanography) for helpful discussion and feedback; and M. Kahr (Scripps Institution of Oceanography) for providing processed satellite images. **Funding:** This work was supported by grants from the National Science Foundation (NSF OCE-1313747 to B.S.M., NSF-ANT-1043671 to A.E.A., and OCE 1538525 and OCE 1638804 to D.A.H.), the National Institute of Environmental Health Sciences (NIEHS P01-ES021921 to B.S.M.), the U.S. Department of Energy Genomics Science program (DE-SC0008593 and DE-SC0018344 to A.E.A.), the Gordon and Betty Moore Foundation (GBMF3828 to A.E.A. and GBMF4960 to G.J.S.), and the Czech Science Foundation (18-13458S to M.O.) and funding from the National Institutes of Health (Chemical Biology Interfaces—University of California, San Diego, training grant 5T32GM112584 to J.K.B.), the Dickinson-McCrink Fellowship (J.K.B.), the Natural Sciences and Engineering Research Council of Canada (NSERC-PDF to S.M.K.M.), and the Simons Foundation Fellowship of the Life Sciences Research Foundation (J.R.C.). **Author contributions:** J.K.B., S.M.K.M., J.R.C., A.E.A., and B.S.M. conceived the project, designed the experiments, analyzed the data, and wrote the paper, with input from all authors. J.K.B., S.M.K.M., J.R.C., and Z.D.M. performed protein expression and did the enzymology experiments. S.M.K.M. and J.K.B. performed chemical synthesis and structural characterization of enzymatic products. J.P.M., E.M.B., J.K.B., and A.E.A. performed the transcriptome analyses. V.A.B. designed vectors and expression constructs and assisted with sequencing. H.L. designed the yeast expression experiments. M.O. constructed the phylogeny. G.J.S. provided the *P. multiseriis* isolate 15091C3. D.A.H. provided *P. multiseriis* biomass for RNA sequencing from the DA induction experiments (17) and helped to design RNA sequencing experiments. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Transcriptome sequences are deposited in NCBI's Sequence Read Archive, BioProjectIDs SAMN08773590-SAMN08773622. The sequence of the *dab* cluster as amplified from *P. multiseriis* 15091C3 is deposited in GenBank, accession no. MH202990. Data used to construct fig. S1 are courtesy of the NASA Ocean Biology Processing Group (OBPG) and are available from <https://oceancolor.gsfc.nasa.gov>. All other data needed to evaluate the paper are present in the supplementary materials.

## SUPPLEMENTARY MATERIALS

[www.sciencemag.org/content/361/6409/1356/suppl/DC1](http://www.sciencemag.org/content/361/6409/1356/suppl/DC1)  
Materials and Methods  
Supplementary Text  
Figs. S1 to S19  
References (32–64)  
Tables S1 to S4

3 May 2018; accepted 15 August 2018  
10.1126/science.aau0382

## Biosynthesis of the neurotoxin domoic acid in a bloom-forming diatom

John K. Brunson, Shaun M. K. McKinnie, Jonathan R. Chekan, John P. McCrow, Zachary D. Miles, Erin M. Bertrand, Vincent A. Bielinski, Hanna Luhavaya, Miroslav Obornik, G. Jason Smith, David A. Hutchins, Andrew E. Allen and Bradley S. Moore

*Science* **361** (6409), 1356-1358.  
DOI: 10.1126/science.aau0382

### How algae turn tides toxic

Algal blooms can devastate marine mammal communities through the production of neurotoxins that accumulate within the food web. Brunson *et al.* identified a cluster of genes associated with biosynthesis of the neurotoxin domoic acid in a marine diatom (see the Perspective by Pohnert *et al.*). In vitro experiments established a series of enzymes that create the core structure of the toxin. Knowledge of the genes involved in domoic acid production will allow for genetic monitoring of algal blooms and aid in identifying conditions that trigger toxin production.

*Science*, this issue p. 1356; see also p. 1308

#### ARTICLE TOOLS

<http://science.sciencemag.org/content/361/6409/1356>

#### SUPPLEMENTARY MATERIALS

<http://science.sciencemag.org/content/suppl/2018/09/26/361.6409.1356.DC1>

#### RELATED CONTENT

<http://science.sciencemag.org/content/sci/361/6409/1308.full>

#### REFERENCES

This article cites 61 articles, 7 of which you can access for free  
<http://science.sciencemag.org/content/361/6409/1356#BIBL>

#### PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)