Differentiating toxic and nontoxic congeneric harmful algae using the non-polar metabolome

1. Phytoplankton culturing

Cultures of the non-toxic dinoflagellate *Alexandrium tamarense* strain CCMP 2023 (synonymous with CCMP 115 and CCAP 1119/1), and the toxic dinoflagellates *Alexandrium catenella* strain CCMP 1719 (synonymous with GTCA 28, previously named *Alexandrium fundyense* (John et al. 2014, Prud'homme van Reine 2017)) and *Alexandrium pacificum* strain CCMP 1493 (synonymous with ATCI01, previously named *A. tamarense* (John et al. 2014)), were acquired from NCMA Bigelow Laboratory. Cultures were arbitrarily arranged in a Percival incubator at 20 °C set to a 12:12 h light/dark cycle with irradiance of 100-123 μmol m^{-2} s^{-1} and grown in filtered seawater from the Gulf of Maine (NCMA Bigelow Laboratory, a salinity of 35) amended with full strength K media (Keller et al. 1987). Cell density was quantified after preservation of cells with 1% acidified Lugol's solution by visual enumeration in a Palmer-Maloney chamber on an Olympus IX-50 inverted microscope.

2. Metabolomics experimental design and execution

To explore non-polar molecules associated with toxicity in *Alexandrium*, the non-polar metabolome of a single non-toxic species (*A. tamarense*) was compared with each of two toxic congeners (*A. catenella* and *A. pacificum*). Metabolomes of toxic versus non-toxic species were compared using the following experimental pairings: *A. tamarense* (n=15) with *A. catenella* (n=15) (Experiment 1) and *A. tamarense* (n=15) with *A. pacificum* (n=15) (Experiment 2). The same non-toxic strain of *A. tamarense* was used in both experiments but the two experiments were conducted separately, in different months, to make the experiment manageable based on availability of batches grown from stock cultures. Due to the experiments being performed independently, quantitative comparisons were made only within and not between experiments. For both experiments, *Alexandrium* spp. cultures at a cell density of 12,000 to 13,000 cells mL^{-1} were split into fifteen 300 mL subcultures of each species which grew for two days. At the end of each experiment, during harvesting, a 1.0 mL aliquot from each culture was preserved with Lugol's solution to measure cell concentrations.

3. Harvesting and extraction of *Alexandrium* cultures for metabolomics experiment

Cultures were harvested with handling of replicates and treatments intermixed over approximately 2 h within 6 h of the start of the light cycle. Harvesting and extraction of cells was carried out as reported by Brown et al.(in review), involving a slight modification of methods used by Poulson-Ellestad et al. (2014) in order to broadly extract non-polar metabolites,

including sterols, fatty acids, glycosylated lipids, and various other classes of lipids. An aliquot of 50 mL of each culture was harvest by centrifugation at 3,260 x g for 10 min. The resulting cell pellets were freeze-dried and stored in a -80 °C freezer for later toxin analysis. The remaining cells were harvested under vacuum onto muffled GF/F filters (Whatman #1825-110). Cell metabolism was then quenched with liquid nitrogen and cells were stored at -80 °C in muffled foil pouches until extraction. Cells were extracted with 30 mL of ice-cold 3:2:1 methanol/acetone/acetonitrile (OmniSolv, >99.5%, MilliporeSigma) by grinding the frozen filters with a liquid nitrogen-chilled mortar and pestle. To remove filter particulates, the extracts were centrifuged four times at 800 x g for 15 min. Solid materials were rinsed each time with fresh solvent and rinses combined with the supernatant to create one 53 mL extract. To remove remaining small particulates, extracts were filtered with a 0.2 μm nylon syringe filter (Acrodisc, Pall Laboratory) before being dried in vacuo using a Thermo Savant Speedvac concentrator. Non-polar and polar metabolites were separated by partitioning extracts in a biphasic mixture of 9:10:15 water/methanol/chloroform (OmniSolv, >99.5%, MilliporeSigma). Extracts were then dried again in vacuo with non-polar extracts used for further analysis and polar extracts set aside.

4. NMR Metabolomic sample preparation, spectral data acquisition, and data processing and analysis

(a) Sample preparation and spectral data acquisition

Non-polar extracts were dissolved with volumes calculated to account for different cell densities at the time of harvesting, to a concentration equivalent to 1.62 x 107 *Alexandrium* cells mL-1 in dimethyl sulfoxide-d6 (99.9% atom d6-DMSO; Cambridge Isotope Labs) containing 0.1% trimethylsilane (TMS) as internal standard for 1H nuclear magnetic resonance (NMR) spectroscopy analysis. The extracts were transferred to 3 mm NMR tubes. Extracts from Experiment 1 (*A. tamarense-A. catenella*) were analyzed using a Bruker Avance IIIHD 800 MHz NMR spectrometer equipped with a 3 mm triple resonance broadband cryoprobe. Extracts from Experiment 2 (*A. tamarense-A. pacificum*) were analyzed using a Bruker Avance IIIHD 700 MHz NMR spectrometer equipped with a 5 mm indirect broadband cryoprobe, because the 800 MHz instrument was unavailable at the time of data acquisition. Spectra of non-polar extracts from both experiments were acquired using a Bruker zg30 pulse sequence 1H NMR experiment compiled from 320 scans.

(b) Data processing and analysis

Spectra were pre-processed using Metabolab version 2019.12.08.1237 (Ludwig and Günther 2011) in Matlab R2013a version 8.1.0.604. The spectra from both experiments were processed separately but using the same parameters. Both sets of spectra were aligned using the TMS internal standard (0.00 ppm), manually phased, spline baseline corrected to improve peak integration, and segmentally aligned. The following spectral regions were removed: -0.50 to 0.30 ppm (excess upfield region), 2.49-2.53 ppm (residual DMSO), 3.30-3.45 ppm (water), and 10.70-11.50 ppm (excess downfield region). The spectra were filtered to reduce the impact of noise, binned (0.005 ppm), and probabilistic quotient normalized to minimize minute differences in dilution (Dieterle et al. 2006). The data were generalized logarithmic transformed to minimize

technical variance and highlight biological variance (Parsons et al. 2007) using $\lambda = 7.4690 \times 10\text{-}9$ for Experiment 1 (*A. tamarense-A. catenella*) and $\lambda = 8.2594 \times 10\text{-}9$ for Experiment 2 (*A. tamarense-A. pacificum*). Finally, the data for both experiments were mean centered.

PLS-Toolbox version 8.0.2 (Eigenvector Research) in Matlab was used to generate orthogonalized partial least-squares discriminant analysis (oPLS-DA) models which were cross-validated using Venetian blinds methods with five data splits for each comparison. Due to poor NMR spectral quality one *A. tamarense* replicate spectrum was removed from analysis of Experiment 1 (*A. tamarense-A. catenella*) and two *A. pacificum* and four *A. tamarense* replicate spectra were removed for Experiment 2. The loadings of the first principal component and first latent variable of the PCA and oPLS-DA models, respectively, were used to identify spectroscopic features that were significantly different between the non-polar metabolomes in each experiment.

5. UPLC/MS Metabolomic sample preparation, spectral data acquisition, and data processing and analysis

(a) Sample preparation and spectral data acquisition

Unlike with ^{1}H NMR spectroscopy, extracts for both experiments were prepared and analyzed by ultra-high performance liquid chromatography tandem mass spectrometer (UPLC/MS) together. Using LCMS grade 2:1 isopropyl alcohol/acetonitrile (OmniSolv, >99.5%, MilliporeSigma) extracts were reconstituted at a concentration equivalent to 4.54×10^{6} *Alexandrium* cells mL^{-1} in and centrifuged at 21,100 x *g* for 5 min to pellet minute particulates. Unfortunately, two *A. pacificum* samples from Experiment 2 were lost due to vials breaking during reconstitution. A small portion of all remaining samples were combined to create a pooled quality control sample that was then split into nine samples. The pooled quality samples were analyzed after every 10 injections of the randomly interspersed *Alexandrium* extracts to monitor instrumental drift. Additionally, a solvent blank was analyzed near the beginning and the end of the batch.

Extracts were analyzed using a Orbitrap ID-X Tribrid mass spectrometer (ThermoFisher Scientific) coupled to a Vanquish UPLC system (ThermoFisher Scientific). Chromatographic separation of the extracts was achieved by injecting of 2.0 µL onto an Accucore C30 silica column (2.1 x 150 mm, 2.6-µm particle size; ThermoFisher Scientific), operated at 50 °C, fitted to the UPLC system. Mobile phase A was 40:60 water/acetonitrile and mobile phase B was 10% acetonitrile in isopropyl alcohol with both containing 10 mM ammonium formate and 0.1% formic acid buffer (Optima, LCMS, Fischer Scientific). A flow rate of 0.4 mL min^{-1} was used with the following mobile phase gradient: equilibration at 80% A; 0-1 min ramp to 40% A; 1-5 min ramp to 30% A; 5-5.5 min ramp to 15% A; 5.5-8 min ramp to 10% A; 8-8.2 min ramp to 0% A; 8.2-10.5 min hold at 0% A; 10.5-10.7 min return to 80% A and re-equilibration at 80% A until 12 min.

Mass spectrometric analysis was accomplished using the Orbitrap ID-X Tribrid spectrometer which possesses a resolution power of 500,000 FWHM at 200 *m/z* and mass accuracy of <1 ppm. Extracts were analyzed twice, first in negative ionization mode with a spray voltage of 2.5 kV and then in positive ionization mode with a spray voltage of 3.5 kV. Other than spray voltage, all other instrument settings were the same. The heated electrospray ionization source was run at a vaporizer temperature of 275 ºC with nitrogen gas flows at 40, 8, and 1 L h^{-1} for sheath, auxiliary, and sweep, respectively. Data was acquired using a scan range of 70-1050 *m/z* and an automatic gain control set at 1×10^{5} ions. Tandem mass spectrometry (MS/MS)

data were acquired in a data-dependent manner with a resolution of 30,000 and an isolation window of 0.8 *m/z* with a cycle time of 1.5 s. Dynamic exclusion was set to 10 s and data-dependent MS/MS was activated by higher energy collisional dissociation with stepped normalized collision energies of 15, 30, and 45 % or collision-induced dissociation energy of 30%. AcquireX Deep Scan was performed with four iterations to maximize number of collected MS/MS scans. Data acquisition was controlled using Xcalibur (ThermoFisher Scientific).

(b) Data processing and analysis

After UPLC/MS acquisition, spectral features (retention time, *m/z*) were extracted from chromatograms using Compound Discoverer V3.1 (ThermoFisher Scientific) and XCMS online (Gowda et al. 2014). This was accomplished by aligning chromatograms, picking and integrating peak pickings, followed by extraction and normalization of peak areas. Spectral features were filtered for noise using the following parameters: peak areas had to be five times greater than baseline, as determined by the blank samples, and features had to be present in greater than 50% of the pooled quality control samples and have a relative standard deviation of less than 30%.

Principal component analysis (PCA) was used to examine differences in non-polar metabolomes of *A. tamarense* vs. *A. catenella* (Experiment 1) and *A. tamarense* vs. *A. pacificum* (Experiment 2), using Compound Discoverer. Each dataset was log 10-transformed and PCA models generated based on spectral features with a fold change greater than or equal to 1.1 or less than or equal to 0.9 and $p \leq 0.05$ via t-test when comparing the metabolomes within each experiment. Additionally, spectral features extracted by XCMS online were entered into LIPID Metabolites and Pathways Strategy (LIPID MAPS) Online Tools (Fahy et al. 2007) to generate volcano plots using a fold-change threshold of 1.1 and $p \leq 0.05$ in order to identify significantly different features in each experiment.

6. Metabolite annotation

Lipids whose concentrations were positively associated with both toxic *Alexandrium* species, *A. pacificum* and *A. catenella*, were sought by comparing the non-polar metabolomes of each toxic species with non-toxic *A. tamarense* from their respective experiment. Specifically, metabolites were prioritized for identification if they were more abundant, by similar proportions, in both toxic species relative to *A. tamarense* from their respective experiment and had minimal differences in abundance between *A. tamarense* in both experiments. Metabolite characterization was predominately dependent on UPLC/MS and MS/MS data. (^{1}H NMR spectroscopy data were mainly used for identification of functional groups and metabolite classes.) Using Compound Discoverer, elemental formulae were generated based on exact mass and isotopic patterns from UPLC/MS data. The elemental formulae along with $[M\text{-}H]^-$ and $[M+H]^+$ *m/z* ions, MS/MS fragmentation patterns, retention time, and isotopic patterns were compared to a local spectral database, built from curated experimental data, as well as these publicly available databases: ChemSpider, LIPID MAPS classification and database (Sud et al. 2006, Fahy et al. 2009), MassBank of North America (MoNA)(2020), the human metabolome database (HMDB) (Wishart et al. 2007), mzVault (in house database), and mzCloud. Additionally, the LOBSTAHS database (Collins et al. 2016) and LIPID MAPS Online Tools (Fahy et al. 2007) were used to putatively identify some MS features, by matching exact mass

and adduct patterns. Retention time and characteristic MS/MS fragmentation patterns were also used to place metabolites into broader lipid classes even when full identification was not achieved (Brügger et al. 1997, Piretti et al. 1997, Hsu et al. 2007, Taguchi 2009, Murphy and Axelsen 2011, Zianni et al. 2013, Hsu 2016). KEGG (Ogata et al. 1999) and LIPID MAPS were utilized to map the significant metabolites onto known metabolic pathways.

REFERENCES

2020. MassBank of North America (MoNA) Fiehn lab.

Brügger, B., G. Erben, R. Sandhoff, F. T. Wieland, and W. D. Lehmann. 1997. Quantitative analysis of biological membrane lipids at the low picomole level by nano-electrospray ionization tandem mass spectrometry. Proceedings of the National Academy of Sciences **94**:2339-2344.

Collins, J. R., B. R. Edwards, H. F. Fredricks, and B. A. S. Van Mooy. 2016. LOBSTAHS: An adduct-based lipidomics strategy for discovery and identification of oxidative stress biomarkers. Analytical Chemistry **88**:7154-7162.

Dieterle, F., A. Ross, G. Schlotterbeck, and H. Senn. 2006. Probabilistic quotient normalization as robust method to account for dilution of complex biological mixtures. Application in ^{1}H NMR metabonomics. Analytical Chemistry **78**:4281-4290.

Fahy, E., S. Subramaniam, R. C. Murphy, M. Nishijima, C. R. H. Raetz, T. Shimizu, F. Spener, G. van Meer, M. J. O. Wakelam, and E. A. Dennis. 2009. Update of the LIPID MAPS comprehensive classification system for lipids. Journal of Lipid Research **50**:S9-S14.

Fahy, E., M. Sud, D. Cotter, and S. Subramaniam. 2007. LIPID MAPS online tools for lipid research. Nucleic Acids Research **35**:W606-W612.

Gowda, H., J. Ivanisevic, C. H. Johnson, M. E. Kurczy, H. P. Benton, D. Rinehart, T. Nguyen, J. Ray, J. Kuehl, B. Arevalo, P. D. Westenskow, J. Wang, A. P. Arkin, A. M. Deutschbauer, G. J. Patti, and G. Siuzdak. 2014. Interactive XCMS online: Simplifying advanced metabolomic data processing and subsequent statistical analyses. Analytical Chemistry **86**:6931-6939.

Hsu, F.-F. 2016. Complete structural characterization of ceramides as $[M-H]^-$ ions by multiple-stage linear ion trap mass spectrometry. Biochimie **130**:63-75.

Hsu, F.-F., J. Turk, K. Zhang, and S. M. Beverley. 2007. Characterization of inositol phosphorylceramides from *Leishmania major* by tandem mass spectrometry with electrospray ionization. Journal of the American Society for Mass Spectrometry **18**:1591-1604.

John, U., R. W. Litaker, M. Montresor, S. Murray, M. L. Brosnahan, and D. M. Anderson. 2014. Formal revision of the *Alexandrium tamarense* species complex (Dinophyceae) taxonomy: The introduction of five species with emphasis on molecular-based (rDNA) classification. Protist **165**:779-804.

Keller, M. D., R. C. Selvin, W. Claus, and R. R. L. Guillard. 1987. Media for the culture of oceanic ultraphytoplankton. Journal of Phycology **23**:633-638.

Ludwig, C., and U. L. Günther. 2011. MetaboLab - Advanced NMR data processing and analysis for metabolomics. BMC Bioinformatics **12**:366.

Murphy, R. C., and P. H. Axelsen. 2011. Mass spectrometric analysis of long-chain lipids. Mass spectrometry reviews **30**:579-599.

Ogata, H., S. Goto, K. Sato, W. Fujibuchi, H. Bono, and M. Kanehisa. 1999. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Research **27**:29-34.

Parsons, H. M., C. Ludwig, U. L. Günther, and M. R. Viant. 2007. Improved classification accuracy in 1- and 2-dimensional NMR metabolomics data using the variance stabilising generalised logarithm transformation. BMC Bioinformatics **8**:234.

Piretti, M. V., G. Pagliuca, L. Boni, R. Pistocchi, M. Diamante, and T. Gazzotti. 1997. Investigation of 4-methyl sterols from cultured dinoflagellate algal strains. Journal of Phycology **33**:61-67.

Poulson-Ellestad, K. L., C. M. Jones, J. Roy, M. R. Viant, F. M. Fernández, J. Kubanek, and B. L. Nunn. 2014. Metabolomics and proteomics reveal impacts of chemically mediated competition on marine plankton. Proceedings of the National Academy of Sciences **111**:9009-9014.

Prud'homme van Reine, W. F. 2017. Report of the nomenclature committee for algae: 15. Taxon **66**:191-192.

Sud, M., E. Fahy, D. Cotter, A. Brown, E. A. Dennis, C. K. Glass, A. H. Merrill, Jr, R. C. Murphy, C. R. H. Raetz, D. W. Russell, and S. Subramaniam. 2006. LMSD: LIPID MAPS structure database. Nucleic Acids Research **35**:D527-D532.

Taguchi, R. 2009. Advances in lipid analysis/lipidomics – analyses of phospholipids by recent application of mass spectrometry. Pages 1-20 *in* A. Lajtha, G. Tettamanti, and G. Goracci, editors. Handbook of neurochemistry and molecular neurobiology: Neural lipids. Springer US.

Wishart, D. S., D. Tzur, C. Knox, R. Eisner, A. C. Guo, N. Young, D. Cheng, K. Jewell, D. Arndt, S. Sawhney, C. Fung, L. Nikolai, M. Lewis, M.-A. Coutouly, I. Forsythe, P. Tang, S. Shrivastava, K. Jeroncic, P. Stothard, G. Amegbey, D. Block, D. D. Hau, J. Wagner, J. Miniaci, M. Clements, M. Gebremedhin, N. Guo, Y. Zhang, G. E. Duggan, G. D. MacInnis, A. M. Weljie, R. Dowlatabadi, F. Bamforth, D. Clive, R. Greiner, L. Li, T. Marrie, B. D. Sykes, H. J. Vogel, and L. Querengesser. 2007. HMDB: The human metabolome database. Nucleic Acids Research **35**:D521-D526.

Zianni, R., G. Bianco, F. Lelario, I. Losito, F. Palmisano, and T. R. I. Cataldi. 2013. Fatty acid neutral losses observed in tandem mass spectrometry with collision-induced dissociation allows regiochemical assignment of sulfoquinovosyl-diacylglycerols. Journal of Mass Spectrometry **48**:205-215.