# TRANSCRIPTOME ANALYSIS OF *PSEUDO-NITZSCHIA AUSTRALIS* YIELDS INSIGHTS ON NITROGEN, ISOPRENE AND DOMOIC ACID METABOLISM IN THIS TOXIGENIC DIATOM

**Poster # 174**


G. Jason Smith[1], Thomas J. Savage[2], Raphael Kudela[3], Kendra Negrey[3], April Woods[1], Holly A. Bowers[4] and Deborah Robertson[5]

[1]Moss Landing Marine Laboratories, 8272 Moss Landing Rd, Moss Landing, CA 95039, USA, email: jsmith@mlml.calstate.edu ; [2]Department of Chemistry, California State University Sacramento, 6000 J Street, Sacramento, CA 95819; [3]Ocean Sciences Department, University of California Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA; [4]Monterey Bay Aquarium Research Institute, 7700 Sandholdt Road, Moss Landing, CA 95039, USA; [5]Department of Biology, Clark University, Worcester, MA 01610, USA


## Overview

Numerous physiological studies have pointed to a dependence of domoic acid biosynthesis on the growth status of *Pseudo-nitzschia* spp, such that DA accumulation is generally enhanced during growth limitation by a range of stressors. While DA accumulation can be affected through manipulation of growth rate in this diatom genus, such manipulations confound interpretation of DA biosynthetic patterns and pathway reconstruction. Here we report on transcriptomes generated from a Monterey Bay isolate of *P. australis* (10249_10AB) grown under different N-source supplementations (**NO3, NH4, UREa, GLUtamate**) at a common dilution rate.

Batch subcultures of antibiotic treated 10249_10AB were conditioned for 6 passages (ca. 2 months) with specific N supplements at 40 µM-N and f/2 levels of other constituents. Log phase pre-conditioned cultures were used to establish continuous cultures under constant light at 15°C with the same N supplementation supplied at 0.27/d. Cultures were sampled after 10 days of growth. Total DA accumulation (pg/cell) was highest in the GLU culture (0.253) and lowest under URE supplementation (0.100) with a rank order of GLU>NO3>NH4>URE.

DNA-free RNA was submitted to the National Center for Genome Resources (NCGR) for RNA-seq library construction, sequencing and annotation as part of the **M**arine **M**icrobial **E**ukaryote **T**ranscriptome **S**equencing **P**roject (Project IDs: MMETSP_139_2, 140_2, 141_2, 142_2). An average of 23,317 contigs (2178 bp median length) were machine assembled from each library with 80% having defined protein coding sequences. Of these putative genes an average of 3899 (22%) sequences per library had significant homologies ($E<10^{-10}$) to known proteins thereby enabling pathway reconstruction.

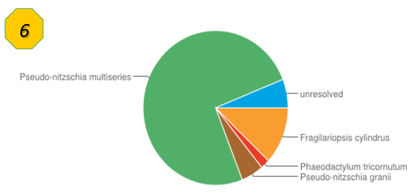
*P. australis* 10249_10AB expressed transcripts encoding enzymes required for complete mevalonate (*hmgr*-dependent) and methyl-erithrytol phosphate (*dxs*-dependent) pathways for isoprene generation. RNA-seq enables use of normalized read frequency (RPKM) as an index of relative transcript abundance. Compared to *hmgr*, *dxs* abundance was strongly correlated with DA accumulation (r=0.96) indicating the MEP pathway as the likely source of the isoprene group on DA. Involvement of the ornithine cycle in supply of substrates for DA biosynthesis is again indicated by correlation of *argD* abundance with DA (r=0.90). Combined with physiological data, this transcriptome database points to the power of RNA-seq analysis for metabolic network reconstruction and analysis of toxin biosynthesis.

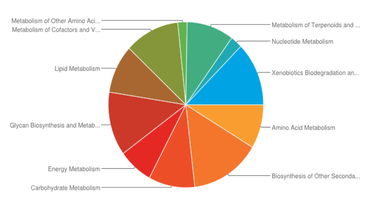
## Characterization of Continuous Culture Physiology and RNA-seq Library Yields

5

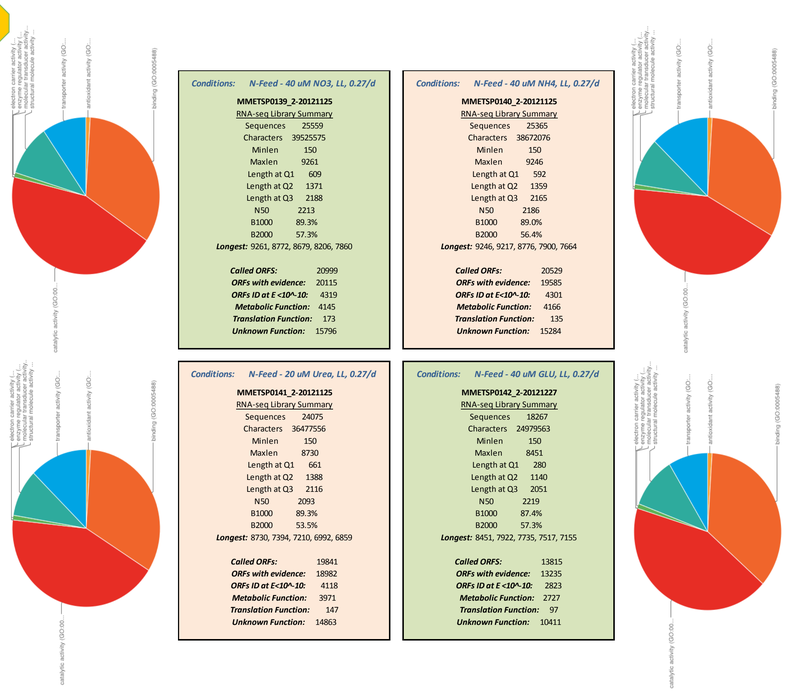
| | Residual Nutrients, µM | | | | | Biomass and Physiology | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Treatment | [ Si ] | [ P ] | [ NO3 ] | [ URE ] | [ NH4 ] | cells / mL | pg Chl / cell | Fv / Fm | pg RNA / cell | RIN |
| FSW base | 109.45 | 31.38 | 24.76 | 0.35 | 1.58 | | | | | |
| NO3 | 83.39 | 26.33 | 33.22 | <0.05 | 2.30 | 22833 | 4.00 | 0.64 | 1.84 | 5.0 |
| NH4 | 60.93 | 27.84 | 12.31 | <0.05 | 0.09 | 51000 | 2.73 | 0.56 | 2.11 | 4.6 |
| URE | 67.58 | 30.15 | 8.70 | <0.05 | 7.95 | 33833 | 2.37 | 0.54 | 2.93 | 3.7 |
| GLU | 69.43 | 25.02 | 8.24 | <0.05 | 9.93 | 28167 | 3.32 | 0.59 | 1.98 | 5.6 |

6




7




DA accumulation is generally stimulated under growth limiting conditions, but is dependent upon nitrogen availability. This study seeks to describe the transcriptomes of N-source acclimated cells. Pseudo-nitzschia australis isolate 10249 10AB were antibiotic treated and acclimated to indicated N-sources (60 µM) for 6 batch culture passages (ca 2 mo). Continuous cultures were utilized to provide cells experiencing equivalent N-supply rates (10.8 µmole-N $d^{-1}$) albeit different N-sources. Cultures were maintained under constant illumination at 1 umol photon m2/s. Cultures were characterized by routine physiological methods and DA content assessed by LC/MS. After 14d of maintenance under these conditions cells were harvested and nucleic acids extracted using TRIZOL reagent. DNA-free RNA (the transcriptome) was sent to NCGR for RNA-seq library preparation, sequencing and automated annotation using their informatics pipeline (see: marinemicroeukaryote.org). RNA-seq libraries were enriched for polyA fraction of total RNA, consequently organelle encoded transcripts are likely under represented.

5. Surprisingly, although sterile filtered natural SW was used (ca 25 µM NO3) there was clear drawdown of N-source supplements in addition to NO3. Cellular Chl**a** yields ranked NO3 > GLU > NH4 > GLU, a similar ranking was observed for DA (see Fig 9). RNA yields followed the inverse ranking.
6. Automated assembly of expressed transcripts was done in absence of a reference genome for P. australis 10249 10AB. Recovered 18S transcript exhibited 100% ID to targeted sequences of this isolate done prior to the experiment. Large transcripts spanning 18S_ITS1_5.8S_ITS2_LSU were also recovered (not shown). Functional annotation of predicted coding sequence (CDS) were based on NCGR gene models and public databases. Subsequent characterization of the predicted proteins expressed by Pnaus using the JCVI PhyloMetaRep tool reveal excellent support (>90%) from recent diatom and chromophyte genome studies. For well supported annotations ($E<10^{-10}$) distribution of metabolic functional classes was similar across the four libraries.
7. Unfortunately, of the called ORFs (average ca. 17,000 / library) nearly 80% of the hypothetical proteins had **no** assigned function. Of the remaining 20% of the annotations, 95% could be assigned metabolic enzyme function while ca 5% contributed to genetic information processing (e.g. translation). Based on current gene ontology (GO) assignments the libraries were again similar in functional distributions, however both NH4 and UREa fed cell transcriptomes (MMETSP_140, _141 respectively) exhibited a higher abundance (ca 5%) of transport associated functionality. The analysis reported here focuses on the subset of known metabolic functions as an initial attempt to identify DA associated metabolic pathways.

## Mining the P. australis Transcriptomes

***Strategy:***

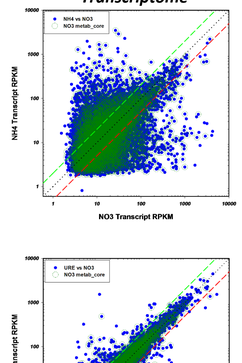
- Set **NO3** culture as **reference growth state**
- Subset library transcript pools on ≥ 98% ID with > 75% coverage of ref transcript
- Independent of annotation, reduced assembly bias

***Outcomes:***

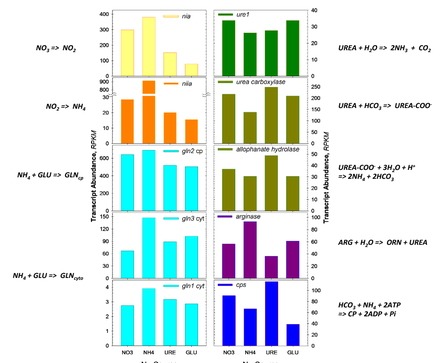
- Enabled characterization of **NO3 Metabolism Core Transcriptome** (6920 contigs) [ **8** ]
- Parallel analysis of non-core transcripts enabled definition of dissolved inorganic N (**DIN**) and organic N (**DON**) metabolism associated transcriptomes (1598 and 2180 contigs respectively) [ **9** ]
- Identifies N-source effects on transcript populations consistent with physiological expectations
- Sampling of cultures acclimated to equivalent N supply rates, transcript abundance (RPKM) can be assumed to represent near steady-state expression levels and likely correlated with corresponding protein abundance, facilitating metabolic pathway reconstruction and comparative flux-balance analysis

8

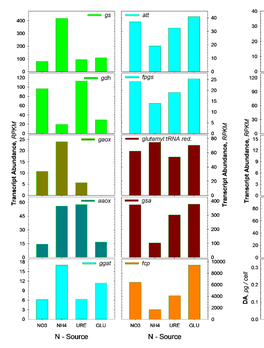
***Nitrate Metabolism Transcriptome***



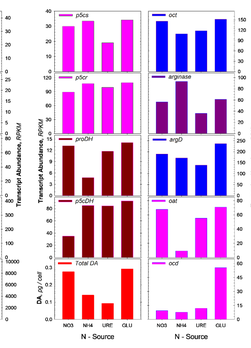

***Nitrate & Urea Metabolism***




***Glutamate Metabolism***



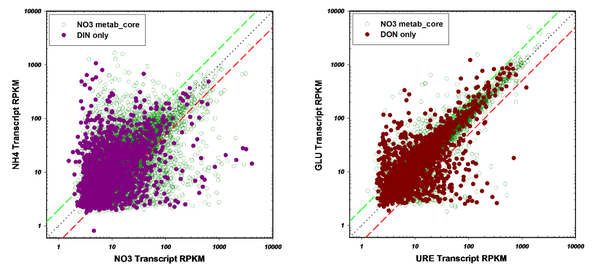

***PRO & ORN Metabolism***




- Range of gene-specific expression patterns associated with aa metabolism
- Multiple strategies for maintaining flux towards GLU
- Novel paths for GLU assimilation and PRO synthesis
- Confirms importance of ORN in PRO metabolism
- NH4 inhibition of ORN dependent PRO synth
- Pathway reconstruction also enables flux – balance modeling, *hypothesis building for metabolite pool sizes*
- *Insights regarding DA metabolism – the story of **argD** functions*

9

**Transcript mapping between specific N-source treatment pairs revealed unique DIN and DON metabolism transcriptomes.**

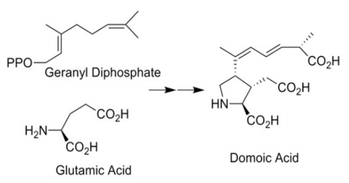



*Unfortunately,* the bulk of these transcripts are derived from the un- or poorly annotated contig pools

**Exploration of the wealth of data in the unknown universe of hypothetical proteins awaits more sleepless nights!**

## Observations and Working Hypotheses

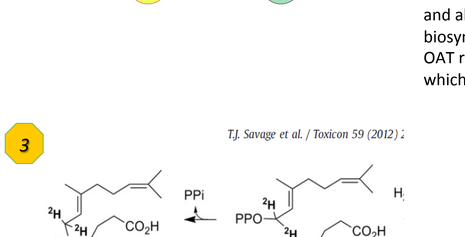
1




In order to fully understand the metabolic role of DA in Pseudo-nitzshia spp., it is necessary to describe its biosynthesis and thereby the regulation of its production. Experimental observations have helped guide the research towards achieving this understanding.

1. $^{13}$C-acetate labeling studies supported the hypothesis that DA is derived from conjugation of an activated isoprenoid and glutamate derived moieties. (Douglas et al. 1992, J Chem Soc Chem Commun 9: 714)
2. DA is a tricarboxylic imino acid. Species producing this derivatized pyrrolidine exhibit altered amino acid metabolism characterized by lower free proline pools and higher ornithine amino transferase (OAT) activities than non-toxic species and strains. (Smith et al. 2001. HABS 2000: 324)
3. $^{2}$H-geraniol feeding studies confirmed the DA biosynthesis hypothesis by demonstrating that DA is synthesized via a nucleophilic attack of geranyl diphosphate onto a glutamate. (Savage et al. 2012. Toxicon 59:25)
4. 'Old Generation' subtractive (OGS) cDNA enrichment libraries of the P. australis transcriptome (Armbrust, Parker, Jenkins in prep) yielded a partial sequence homologous with acetyl-ornithine amino transferase (argD/acOAT). RT-PCR revealed a correlation between argD abundance and cellular DA toxicity. (Kudela et al. 2010. Prog Oceanog 85:108)

2

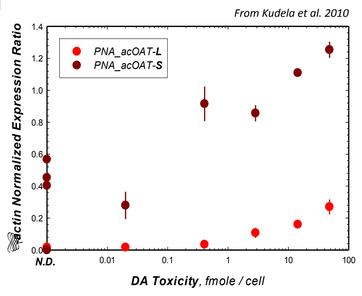



The ornithine-urea cycle (OUC) is critical for arginine biosynthesis and also contributes provides an alternate pathway for proline biosynthesis. Transamination activities of both acOAT (argD) and OAT regulate cellular levels of **activated glutamate derivatives** which could contribute to DA biosynthesis.

3



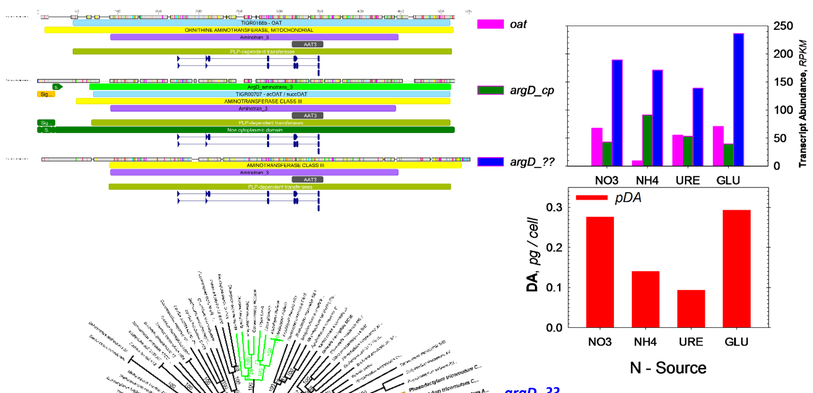

4




## Transcriptome Based Insights on Domoic Acid Metabolism in P. australis

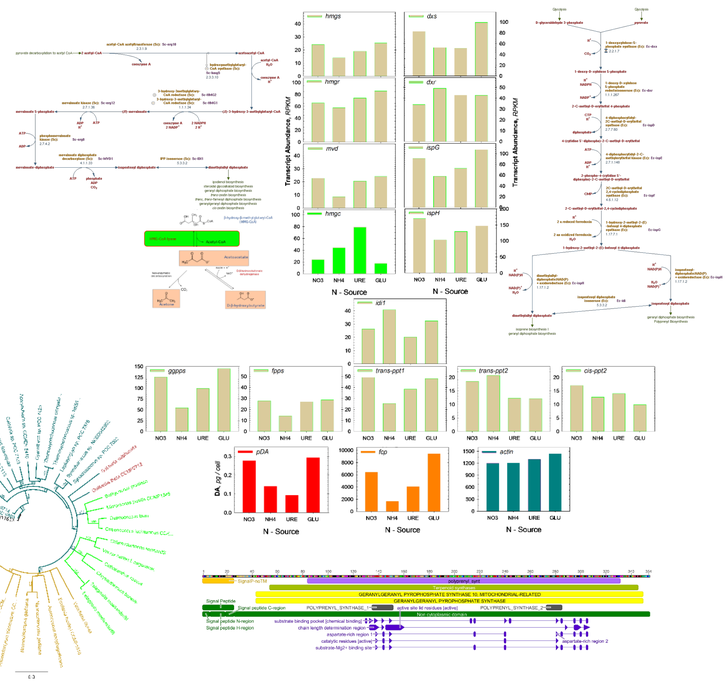
9

**Identification of Three *argD* – like Amino Transferases**




10

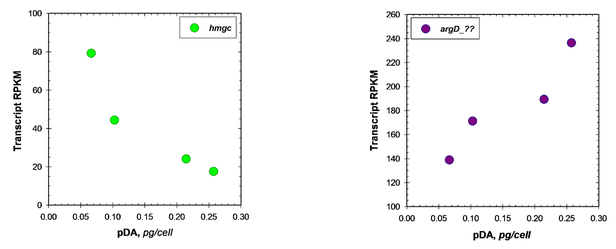
**P. australis Expresses Multiple Pathways for Isoprene Biosynthesis**




11

**Application of Flux-Balance Analysis For Prioritization of Gene Targets for Manipulation of the DA Toxigenicity Phenotype**

- *As we know cellular turnover rate, can assume transcript abundance represents **net rate of expressio**n and by extension enzyme **abundance /activity** (at least for diatoms!)*




***The DA toxigenicity phenotype has been observed to be associated with elevated ornithine amino transferase activities and hypothesized to supply the 'activated glutamate derivative' required for DA biosynthesis***

- Ornithine amino transferases are characterized as ClassIII pyridoxal phosphate-dependent amino transferases
- Machine annotation of the P. australis transcriptomes indentified an ***argD*** homolog (*PNAargD_??*) which exhibited a DA correlated expression similar to EST derived ***argD/acOAT*** from cDNA subtraction (hi-low DA cultures, see above)
- ***However,*** the previously characterized *PNAacOAT* was not called in the MMETSP annotations although full length cds present! – *many layers to machine annotation*
- *PNAacOAT* is plastid targeted, but expression pattern in this experiment was not correlated with content suggesting a likely growth rate dependent expression
- *PNAargD_??* localization is unclear, but homologs present in other *Pseudo-nitzschia* species and diatoms forming a distinct clade with fungal and ciliate forms
- While predicted functions are similar determination of substrate-specificity is critical and subject of future work

***Co-Expression of complete:***

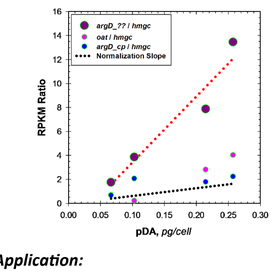
- Hydroxy-methylglutaryl/Mevalonate pathway – *sterol biosynthesis*, likely cytoplasmic localization (*left pathway*)
- Methyl-erythritol phosphate (MEP) pathway – *cartotenoid, terpene biosynthesis*, likely plastid localized (*right pathway)*
- Both lead to IPP required for geranylPP (*shared pool?*)
- Higher expression along MEP, rank order GLU>NO3>NH4>URE in *dxs*
- At least 5 classes of prenylPP transferases identified, expression variable, several map to light harvesting *fcp* expression consistent with carotenoid metabolism
- *ggpps* expression levels similar to MEP pathway activity. Expression pattern also correlated with cellular DA content. PNA*ggpps* forms well supported subclade with *ggpps*' from other *Pseudo-nitzschia* species
- ***However,*** the mevalonate pathway associated *hmgc* encoding an HMG-CoA lyase exhibits an expression pattern *anti-correlated* with DA content supports the hypothesis that mevalonate derived IPP may be rate limiting for DA biosynthesis

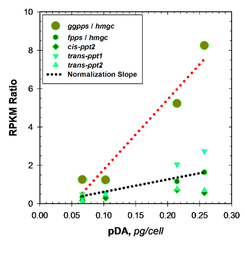
***hmgc:***

- Will limit IPP production
- Mevalonate pathway < MEP activity
- *Regulates supply of GGP to DA?*
- *Localization?*
- **Consistent with $^{13}$C-acetate labeling**

***argD_??:***

- Class III aminotransferase
- Positively correlated with DA
- *Regulates supply of activated dicarboxylic amino acid to DA?*
- *Localization? Substrates?*
- **Consistent with high ORN_AT phenotype of *P. australis***




***Application:***

***- Screening for DA Biosynthetic <u>functions</u>***

- Of the three *P. australis* class III aminotransferases, only ***argD_??*** exhibits a significant RPKM ratio response slope
- For prenyl transferase screening only ***ggpps*** exhibits significant response slope
- Identifies these two as functional candidates
- Approach should be applicable to any expression assay
- Can be made species specific . . . . . .