Received Date : 14-Oct-2016

Revised Date : 20-Dec-2016

Accepted Date : 23-Jan-2017

Article type : Standard Paper

Handling Editor: Nathan Swenson

Ecological interactions and coexistence are predicted by gene expression similarity in freshwater green algae

Anita Narwani [a]*, Bastian Bentlage [b, c], Markos A. Alexandrou [d], Keith J. Fritschie [e], Charles Delwiche [b,f], Todd H. Oakley [g], and Bradley J. Cardinale [h]

[a] LA-H60.1 Überlandstrasse 133, Department of Aquatic Ecology, Eawag (Swiss Federal Institute of Aquatic Science and Technology), 8600 Dübendorf, Switzerland

*Correspondence author

email: anita.narwani@eawag.ch

phone: +41 (0) 79 466 1257

fax: +41 (0) 58 765 5802

[b] CMNS-Cell Biology and Molecular Genetics, 2107 Bioscience Research Building, University of Maryland, College Park, MD 20742-4407, USA

email: bastian.bentlage@gmail.com

[c] 303 University Drive, Marine Laboratory, University of Guam, Mangilao, GU 96923, USA

[d] Wildlands Conservation Science, LLC P.O. Box 1846, Lompoc, CA 93438
email: markosalexandrou@me.com
[e] 78 College Street, Department of Biological Sciences, Dartmouth College, Hanover, NH 03755 USA
email: keith.j.fritschie.gr@dartmouth.edu
[f] Maryland Agricultural Experiment Station, AGNR, University of Maryland, College Park, MD 20742, USA
email: delwiche@umd.edu
[g] 4101 Life Sciences Building, UCEN Road, Department of Ecology, Evolution and Marine Biology, University of California, Santa Barbara, CA 93106, USA
email: todd.oakley@lifesci.ucsb.edu
[h] 440 Church Street, School of Natural Resources and Environment, University of Michigan, Ann Arbor, MI 48109-1041, USA
email: bradcard@umich.edu

Key words: coexistence, competition, facilitation, gene expression, molecular phenotype, overyielding, phylogenetic distance, species interactions, transcriptomics

Summary

1) Phenotypic variation controls the species interactions which determine whether or not species coexist. Long-standing hypotheses in ecology and evolution posit that phenotypic differentiation enables coexistence by increasing the size of niche differentiation. This hypothesis has only been tested using macroscopic traits to date, but niche differentiation, particularly of microscopic organisms, also occurs at the molecular and metabolic level.
2) We examined how phenotypic variation that arises at the level of gene expression over evolutionary time affects phytoplankton species interactions and coexistence.
3) We predicted that similarity in gene expression among species would decline with phylogenetic distance, and that reduced similarity in gene expression would weaken competition, increase facilitation, and promote coexistence.
4) To test this, we grew eight species of freshwater green algae in monocultures and bicultures for 46 days in a lab microcosm experiment. We quantified the strength of species interactions by: 1) fitting Lotka-Volterra models to time-series densities and estimating interaction

coefficients, and 2) calculating relative densities that compare species' steady-state densities in biculture to those in monoculture. We used Illumina high throughput sequencing to quantify the expression of 1,253 families of homologous genes, including a set of 17 candidate genes that we hypothesized *a priori* to be involved in competition or facilitation.

5) We found that closely related species had greater similarity in gene expression than did distantly related species, but as gene expression became more similar, species experienced weaker competition or greater facilitation, and were more likely to coexist. We identified gene functional categories that were uniquely differentially regulated in association with particular species interaction types.
6) *Synthesis* – Contrary to common thinking in ecology and evolution, similarity in gene expression, and not differentiation, was associated with weaker competition, facilitation and coexistence.

Introduction

Understanding biodiversity and species coexistence continues to be a central goal in community ecology. Theories of biodiversity state that niche differentiation among species enables coexistence by weakening competitive interactions that would otherwise lead to competitive exclusion (Chesson, 2000, MacArthur and Levins, 1967, Volterra, 1928, Lotka, 1920, Lotka, 1925). Niche differentiation and the strength of interactions among species are determined by phenotypic differentiation among species at lower levels of biological organization. Understanding biodiversity and coexistence therefore requires investigations of the evolution of the phenotypic basis of niche differentiation and the strengths of species interactions.

Long held views in biology predict that recently diverged species will be more phenotypically similar to one another ('phylogenetic niche conservatism'), will share similar resource requirements and, in turn, will compete more strongly and be less likely to coexist than species that diverged longer ago (Webb et al., 2002, Ackerly, 2003, Cooper et al., 2010, Wiens et al., 2010, Darwin, 1859). The phenotypic traits considered to date are generally observed at the organismal level and have been chosen because they are thought to mediate species interactions, including competition for resources (e.g. beak shape (Lamichhaney et al., 2015), body shape (Ingram, 2015, Wanek and Sturmbauer, 2015), or body size (Ashton, 2004, Blomberg et al., 2003)). Species also express significant phenotypic variation at the molecular and metabolic

level in response to biotic and abiotic environmental stimuli (Tirosh et al., 2006, Grishkevich and Yanai, 2013). Using high throughput sequencing technology, it is now possible to determine how species' phenotypes are differentiated at the level of gene expression, and to determine how gene expression profiles mediate and respond to the presence of other interacting species (Schulze et al., 2016), and in turn influence coexistence.

To date, investigations of the molecular basis of species interactions at the level of gene expression have largely focused on host-pathogen, -parasite or -symbiont interactions (Schulze et al., 2016), and to a lesser extent, on facilitative interactions (e.g. Amin et al., 2015). They have uncovered significant transcriptomic changes (i.e. gene expression changes measured across numerous genes) occurring for each species during the interactions (Schulze et al., 2016). These transcriptomic changes can result in important functional changes within the organism, including for example, the production of lysosomes used in cell lysis and macromolecular digestion, changes in the cell cycle or the rate of ribosome production (Wohlrab et al., 2016, Schulze et al., 2016), and the detoxification of secondary metabolites (Arfi et al., 2013). Such changes, though only observable at the molecular level, may be of major importance in determining the strength, type and outcome of species interactions (Schulze et al., 2016).

In this study, we systematically investigated how the similarity in gene expression among eight species of freshwater green algae mediates the type, strength and outcome of their interspecific interactions. Interactions among algae are known to be both competitive and facilitative (Venail et al., 2014, Fritschie et al., 2014). Algae experience competition for a limited number of inorganic resources and light (Litchman and Klausmeier, 2008, Hutchinson, 1961), and recent transcriptomic studies of marine phytoplankton have shown complex transcriptomic responses to resource limitation (hundreds of differentially regulated genes) (Dyhrman et al., 2012, Frischkorn et al., 2014). A recent study also showed that two species of co-occurring marine diatoms had functionally unique sets of differentially-regulated genes in response to nitrogen and phosphorus availability, suggesting that they may partition their niches at the metabolic level, enabling coexistence (Alexander et al., 2015). By contrast, relatively little is known about the mechanistic basis of facilitative interactions in algae, although some green algae receive a yield benefit from mixotrophic carbon consumption (Tanoi et al., 2011, Gautam et al., 2013), and some are auxotrophic for particular vitamins (Croft et al., 2006). This suggests that cross-feeding of metabolites or waste products from "leaky" interspecific neighbors may

lead to growth rate and yield benefits in the presence of other species. As a result, we hypothesized that species with lower similarity in gene expression across their transcriptomes would compete less strongly and be more likely to coexist due to greater ecological niche differentiation at the molecular and metabolic level (Levy and Borenstein, 2012, Levy and Borenstein, 2013, Lindemann et al., 2016). We also expected that species with lower similarity in gene expression may be more likely to display facilitative interactions and coexistence due to a greater possibility of metabolic complementarity and cross-feeding (Lindemann et al., 2016).

First, we aimed to test whether patterns of gene expression among species tend to diverge over evolutionary time, as represented by relatedness among species on a phylogenetic tree. While relatively little is known about how transcriptomes evolve as species diverge along a phylogeny, some recent investigations lend support to neutral models of evolution (Khaitovich et al., 2005, Khaitovich et al., 2006, Li et al., 2010, Uebbing et al., 2016), which predict that as species become more distantly related, species similarity in gene expression should decline monotonically (see also Brawand et al., 2011, Yang and Wang, 2013). However, other patterns including gene expression conservatism (Liao and Zhang, 2006) and rapid divergence (Whittle et al., 2014) are also observed, and there is currently no general consensus as to how sequence and expression divergence are related. Second, we aimed to determine whether gene expression similarity influences the type (competition versus facilitation) and strength of species interaction and the likelihood of coexistence among species pairs. Specifically, we tested three predictions: 1) More distantly related species of freshwater green algae have more distinct patterns of gene expression across their transcriptomes than do closely related species, 2) Species with more distinct patterns of gene expression experience weaker competitive interactions and are more likely to display facilitation, and 3) because of weaker competition and more likely facilitation, species with lower gene expression similarity are more likely to coexist with one another. To test these predictions, we used a microcosm experiment in which we grew each of eight species of freshwater algae either in monoculture or biculture, and we measured species interactions and gene expression using high throughput Illumina RNA sequencing. We found that while gene expression similarity did diverge over evolutionary time, competition grew weaker, and facilitation and coexistence more likely when gene expression patterns among species were more similar.

Methods

Species selection and phylogeny

We selected eight species of freshwater green algae: *Chlorella sorokiniana, Closteriopsis acicularis, Cosmarium turpinii, Pandorina charkowiensis, Scenedesmus acuminatus, Selenastrum capricornotum, Staurastrum punctulatum* and *Tetraedron minimum* (Supporting Information Table S1). Cultures were obtained from the University of Texas at Austin or the University of Göttingen (Germany). We chose these algae because they are widespread and abundant in lakes across the United States (Naughton et al., 2015). It was also important that the species were able to grow under laboratory conditions and be morphologically distinguished under the microscope. This subset of species also provided a relatively even cross-section of species from a phylogeny of green algae, and therefore also a good range of phylogenetic distances (Alexandrou et al., 2015). Phylogenetic distance is defined here as the sum of all branch lengths between a group of species on a phylogeny (Faith, 1992, Cavender-Bares et al., 2009, Cadotte et al., 2010), and we estimated the phylogenetic distances for this study based on the molecular phylogeny published by Alexandrou et al. (2015).

Experimental design and sampling

We prepared 108 1L media bottles filled with 1L of modified COMBO growth medium (enriched with 0.1 mM KCl and 30 µM NH_4Cl final concentrations)(Kilham et al., 1998). We inoculated bottles with either one of the 8 monocultures or one of the 28 possible bicultures at a total initial density of 200 cells·mL^{-1}. Inoculations were conducted in a substitutive design such that each species in a given biculture was inoculated at 100 cells·mL^{-1}. All species compositions were replicated in triplicate. Bottles were then placed on Wheaton® (349000-A) roller racks at 20° C under a 16:8 hour light:dark cycle at a light intensity of ca. 81 µEinstein. We exchanged 10% (100 mL) of the culture volume every other day with sterile COMBO starting 4 days after the initial inoculation. We monitored community-level biomass over time by measuring the fluorescence of chlorophyll-a every second day on a well-plate reader (Fluorometer, Synergy H1 Hybrid Reader, Biotek). We used the community-level biomass to gauge when the majority of communities had achieved steady-state biomass, as indicated by no further increase in chlorophyll-a fluorescence. We continued exchanges and sampling for one more week after steady state had been achieved for the majority of communities, before terminating the experiment at 46 days after inoculation (see Appendices A and B). Forty-eight days represents

between 14 and 36 generations of algal growth for the species used here. Samples for the identification and counting of algae were taken from each bottle every other day until day 30, and then every 4 days until day 46. The samples were preserved by pipetting 250 uL of sugared, buffered Formalin into 1 mL of algae (final concentration of 2%) and densities of algal natural units (cells or colonies) were counted on a FlowCam™ (Fluid Imaging Technologies Inc.).

Transcriptomics

On day 46, we took samples from each bottle for mRNA extraction and quantification. We centrifuged between 100 mL and 900 mL of algal culture to obtain a pellet of algal biomass for mRNA extraction. The supernatant was decanted and mRNA was extracted from the algal pellet using the Ambion RNAqueous extraction kit (Life Technologies) following the manufacturer's protocol. RNA was polyA-selected and the libraries were prepared using the Illumina TruSeq RNA Sample Preparation Kit, v2. The RNA was sequenced at the Beijing Genome Institute (BGI; Shenzhen, Guangdong, China) on an Illumina HiSeq2000 sequencer generating 91 basepair (bp) paired-end reads. For detailed descriptions of the methods used for the transcriptome assemblies, read mapping, identification of open reading frames, candidate gene annotations and gene family annotations, see the 'Transcriptomics' section of the Supporting Information.

Estimates of species interactions

We used two approaches to estimate the strength of interactions between species in biculture. First, we used the densities of algae at the final time point of the experiment to compare steady state densities of each species in biculture to those in monoculture:

$$RDi = Di_{biculture} / Di_{monoculture} \quad [1]$$

where RDi is the relative density of species i, $Di_{biculture}$ is the cell density of species i in biculture (cells/mL), and $Di_{monoculture}$ is the density of species i in monoculture. Due to the substitutive design of our experiment, the expected relative density for each species in biculture was 0.5, assuming that each species has the same impact on an individual of another species as it has on itself. The relative density total of the biculture is then the sum of each species' individual relative density:

$$RDT = \sum RDi \quad [2]$$

The expected RDT, given that species have the same impact on others as they do on themselves, is 1. An RDT < 1 indicates competition, and that interspecific interactions are stronger than intraspecific interactions for at least one species. An RDT > 1 indicates that interspecific competition is weaker than intraspecific competition for at least one species, which occurs when species display niche partitioning or facilitation.

Second, we estimated species interactions by fitting Lotka-Volterra competition models to the time-series of each bottle to estimate interaction coefficients. First, we estimated each species' maximum growth rate (r) and carrying capacity (K) by fitting the time-series cell density counts from the 3 replicate monocultures to a logistic growth equation:

$$dI / dt = I \cdot r_i \cdot (K_i - I) / K_i \quad [3]$$

where I is the density of species i in natural units, r_i is the maximum intrinsic growth rate of the population of species i, and K_i is the carrying capacity, or the density of species i at steady state. We then used the estimates of r and K for each species grown in monoculture to populate parameters of the Lotka-Volterra model for bicultures:

$$dI/dt = I \cdot r_i \cdot ((K_i - I - \alpha_{ij} \cdot J) / K_i) \quad [4]$$

$$dJ/dt = J \cdot r_j \cdot ((K_j - J - \alpha_{ji} \cdot I) / K_j) \quad [5]$$

In this model, α_{ij} and α_{ji} are the interaction coefficients, which represent the per capita impact of species j (i) on an individual of species i (j). For further details of the model-fitting procedures and parameter estimates, as well as examples of the model fits (Fig. S2), see the 'Estimates of species interactions' section of the Supporting Information.

Statistical analysis

In order to estimate gene expression similarity between pairs of species (hereafter abbreviated 'GES') we estimated Spearman rank correlations (ρ) between the TPM values (transcripts per kilobase million –i.e. read counts normalized for read length (Wagner et al., 2012)) for the two species in each biculture bottle across all commonly-expressed genes with PANTHER IDs. We also estimated the average of this correlation among all 9 pairwise combinations of the 3 replicate monocultures of both species in a given biculture. In calculating the GES, we considered only genes that were expressed by all species in monoculture (i.e. commonly-expressed genes) so that it would be possible to estimate the level of gene expression similarity among species pairs; it is impossible to compare levels of gene expression for genes

not occurring in all species. Lastly, we estimated the log fold change (logFC) of each PANTHER ID in biculture relative to monoculture for each species, and estimated the correlation coefficients of these fold changes for the two species in each biculture bottle. This is a measure of how similarly two species modified their gene expression in biculture relative to monoculture.

We then tested whether measures of GES were correlated with variation in phylogenetic distance (PD), species interactions strength (interaction coefficients and RD_is), and coexistence. We estimated long-term coexistence by simulating the Lotka-Volterra model forward for 100 days (50 model discrete-time steps) and determining whether both species had positive densities at the end of the model simulations (coexistence =1, competitive exclusion = 0). We tested whether gene expression correlations were a significant predictor of the likelihood of coexistence and positive species interactions (i.e. negative interaction coefficient = 1, positive interaction coefficient = 0) using logistic regression.

Many expressed genes may have little to do with ecological niche differentiation, competitive abilities or facilitation, and may simply be 'house-keeping' genes. As a result, we specifically aimed to identify a number of candidate genes that we hypothesized *a priori* to be involved in competitive or facilitative interactions in algae. These included genes related to the ability of algae to compete for nutrients, light, and trace elements, as well as genes that may be related to facilitative interactions via the cross-feeding of sugars and vitamins. While it is generally thought that green algae are solely autotrophic, some green algae have been observed to benefit from mixotrophic carbon metabolism (Tanoi et al., 2011, Gautam et al., 2013, Li et al., 2014) and/or vitamin supplementation (Croft et al., 2006, Giovannoni, 2012). These genes included: 1) Carbonic anhydrase, 2) Iron permease, 3) Light harvesting complexes A & B (LHCAB), 4) Glutamate semialdehyde transferase (GSA), 5) Nitrate reductase, 6) Nitrate transporter, 7) Nitrite reductase, 8) Nitrite transporter, 9) Nitrogen assimilation regulatory protein (NARP), 10) Nitrogen regulatory protein (NRP), and 11) Phosphate transporter. Gene families related to sugar metabolism included: 12) Glucose, 13) Mannose, and 14) Succinate. Gene families related to vitamin production or metabolism included: 15) Biotin B7, 16) Cobalamin B12, and 17) Thiamine B1. When multiple gene sequences were identified within a given gene family, TPM values within the gene family were summed across all genes (see 'Candidate gene annotation' in the Supporting Information). We then tested whether expression levels of 17 gene families were able to predict relative density across all species and species combinations using

Mantel tests, and across pairwise combinations for a particular focal species using Spearman Rank correlations.

We were also interested in determining whether species that experienced facilitation or overyielding were modulating the expression of genes with different gene functions than those species that experienced competition or underyielding. In order to better understand the putative functions of the gene families that were significantly up- or downregulated under different ecological scenarios, we performed a differential expression analysis and mapped the identified genes back to Gene Ontology (GO) terms (for details, see 'Functional annotations of differentially regulated genes' in the Supporting Information). We split the dataset into two, non-mutually exclusive sets of binary ecological categories: first among species that displayed overyielding (RYi > 0.5) versus underyielding (RYi < 0.5), and separately among species that experienced competition ($\alpha_{ij} > 0$) versus facilitation ($\alpha_{ij} < 0$). In total, we created 8 comparisons for each GO category: 2 types of interaction coefficients + 2 types of relative density outcomes, each crossed by 2 types of gene modulation (upregulated or downregulated). Using these comparisons, we aimed to identify the functional differences among differentially regulated genes in our dataset which were uniquely associated with different population-level responses to growth in biculture (i.e. overyielding versus underyielding, or competition versus facilitation).

Results

Results of our study are consistent with the prediction that more distantly related species are more divergent in their patterns of gene expression. Gene expression was positively correlated for all species pairs, often quite strongly (note that all GES measures were > 0, Fig. 1). However, the magnitude of these positive correlations tended to decline as phylogenetic distance among species pairs increased (P1, Fig. 1a, $\rho = -0.27$, $P = 0.02$). Although we were primarily interested in gene expression among species when they were interacting, species can also differ in gene expression intrinsically (i.e. in monoculture), or may differ in how they up- or downregulate gene expression in biculture relative to monoculture (estimated as the log fold change in TPM, 'logFC'). Therefore, we also tested how phylogenetic distance was related to the GES of species grown in monoculture and to the GES of logFC. We observed the same general trend between phylogenetic distance and GES, regardless of whether we looked at the GES across monocultures of species pairs (Fig. 1b, $\rho = -0.35$, $P = 0.07$), the GES of the logFC of species grown as bicultures ($\rho = -0.28$, $P = 0.01$), or the similarity of expression of just the

candidate genes in biculture (ρ = -0.21, P = 0.07) or monoculture (ρ = -0.18, P = 0.36). This set of analyses indicates that, regardless of the conditions in which gene expression was measured, or the particular estimate of gene expression that was used, more distantly related species had greater differences in their patterns of gene expression than more closely related species.

Contrary to our second prediction, when species were more similar in gene expression (higher values of GES), competition between them was weaker (Fig. 2 top panels), and for other species facilitative interactions became more common (Fig 2 bottom panels). Increasing GES was associated with a decline in interaction coefficients estimated from Lotka-Volterra models fit to population dynamics (Spearman rank correlation of GES and α_{ij} for: *Closteriopsis acicularis* (Fig 2a, ρ=-0.72, P<0.01), *Tetraedron minimum* (Fig 2b, ρ=-0.49, P=0.04), *Selenastrum capricornutum* (Fig 2c, ρ=-0.45, P=0.05), and *Staurastrum punctulatum* (Fig 2d, ρ=-0.79, P<0.01), non-significant correlations not shown). This trend was also supported when we investigated GES across the transcriptome for species grown separately in monoculture, or when we investigated the GES of candidate genes in biculture or in monoculture (Table 1). This again indicates that regardless of whether we consider a large part of the transcriptome or only a set of genes presumed to be important in species interactions, and regardless of whether we investigated gene expression in biculture or in monoculture, species with more similar patterns of gene expression tended to show weaker competition and in some cases, facilitation. Altogether, 13 of the 14 significant correlations between GES and interaction strength were negative, which is significantly greater than expected by chance (χ^2 = 10.29, P<0.01).

Contrary to our third prediction, we found that GES was also associated with a greater likelihood of coexistence among species pairs. This is illustrated in Fig 3, which shows species pairs in which one or both species had a negative interaction coefficient (red dots), indicating that at least one species benefited from the presence of another species (i.e. facilitation). GES was a significant predictor of both the likelihood of positive species interactions (β = 9.60, P = 0.005), as well as the probability of coexistence (β = 9.36, P = 0.006).

Instead of being encoded by similarity in expression levels across multiple genes, it is possible that interaction strengths are determined by the expression of individual, functionally important genes. Indeed, when we investigated whether the expression levels of each particular candidate gene were correlated with the interaction strengths of each individual species across bicultures, we found that almost all candidate gene families were negatively correlated with the

magnitude of the interaction coefficients of at least one of the eight species (Table 2). Of the 32 significant correlations between gene expression and species interaction strength, 30 were negative, which is significantly greater than expected from chance (Table 2, $\chi^2 = 24.5$, $P<0.0001$). This result indicates that the expression of candidate genes tended to be negatively associated with species interaction strengths overall, indicating weaker competition and more frequent facilitation. When we considered correlations between expression of candidate gene families and RD_is for individual species (Supporting Information Table S2), 52 of 56 significant and marginally non-significant correlations were positive ($\chi^2 = 41.14$, $P<0.01$), again indicating that candidate gene expression similarity was generally associated with weaker competition and more frequent facilitation. We found that expression levels of all of the candidate gene families except nitrite reductase and cobalamin were significant predictors of RD_i across species and species combinations (Fig. 4, Supporting Information Table S2). Both the frequency of overyielding ($RD_i > 1$) and the frequency of facilitation ($\alpha_{ij} < 0$) increased as expression levels of the candidate gene families increased (two left most columns in Fig 4).

Finally, to identify other potential genes and gene families that may correlate with species interaction strengths, we searched for genes whose expression patterns were differentially regulated in species experiencing different types of interactions. We referenced these differently regulated gene families against the GO annotation database and found 28 Molecular Process (level 3) GO annotations. The majority of these gene functions were differentially regulated in the same fashion (both up or both down) regardless of whether the species experienced competition or facilitation, or whether they experienced over- or underyielding (Fig. S3, pluses and minuses are black and on the same side of the zero line). This suggests that the majority of differentially expressed genes were not contrastingly regulated in a different fashion in species experiencing different types of species interactions, i.e. competition or facilitation. However, six of the 28 Molecular Function annotations were either upregulated when species experienced competition and underyielding but were downregulated when species experienced facilitation and overyielding, or vice versa (annotations are bolded in the legend of Fig. S3, and are indicated by red plus and minus signs being on opposite sides of the zero line). Because these GO annotations were contrastingly regulated among species experiencing different types of interactions, these gene functions may be involved in determining species interaction strengths. Species experiencing facilitation and overyielding tended to upregulate gene functions generally

associated with transcription (e.g. DNA/RNA binding molecules) and energy metabolism (Fig. S3, GO annotation #17, #26 and #27). More specifically, annotations for these GO terms, using the AmiGO2 portal (amigo.geneontology.org) and restricting search results to only those derived from Viridiplantae and with experimental evidence for gene function, included Ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCo), a key enzyme in the Calvin cycle, as well as pyruvate dehydrogenase kinase and succinate-CoA ligase, enzymes involved in the production of Acetyl-CoA and the Citric Acid Cycle. By contrast, species experiencing competition and underyielding tended to upregulate genes associated with molecular transport, both within the cell and across cell membranes (Fig. S3, GO annotation #7, #21 and #22). For example, GO terms 0022857 and 0022892 (Fig S3, #21 and #22) identified as highly abundant in cultures experiencing negative ecological interactions (competition or density under-yielding) contained annotations to nitrate, ammonium, sugar, silicon, magnesium, and other metal transporters.

Discussion

In this study, we investigated whether patterns of gene expression among freshwater green algae tend to diverge over evolutionary time, and in turn, whether gene expression similarity among species predicts the type, strength and outcome of species interactions in terms of coexistence. Consistent with our first prediction and with previous studies on plants and mammals (Yang and Wang, 2013, Brawand et al., 2011), we found that as species diverge from one another along a molecular phylogeny, their similarity in gene expression tends to decline. This finding holds regardless of whether we investigated similarity of genes expressed among species pairs in monoculture or biculture, and whether we considered all genes or just genes thought to be responsible for species interactions. Contrary to our second prediction, as gene expression similarity increased, interaction strengths tended to decline, indicating weaker competition and more frequent facilitation. Moreover, counter to our third prediction, species with greater similarity in gene expression tended to be more likely to coexist.

A fundamental assumption of our original hypothesis is the idea that as species' gene expression diverges over evolutionary time, greater differences in gene expression cause species to become more ecologically niche differentiated, which weakens the impact of competition and enables coexistence. This idea, sometimes referred to as "phylogenetic limiting similarity", "phylogenetic niche conservatism", or "evolutionary character displacement", has been widely supported and adopted in both ecology and evolution (Violle et al., 2011, MacArthur and Levins, 1967, Pfennig and Pfennig, 2009, Davies et al., 2007, Grant and Grant, 2006, Schluter, 2000). However, evidence for these hypotheses is not universal (Kraft et al., 2015, Kunstler et al., 2012, Best et al., 2013, Venail et al., 2015), and in particular they are unsupported for freshwater green algae (Narwani et al., 2013, Venail et al., 2014, Fritschie et al., 2014, Naughton et al., 2015, Alexandrou et al., 2015). Our findings here support the opposite trend: while more distantly related species have greater differences in gene expression, species with greater similarity in gene expression are more likely to experience weakened competition and coexistence. Clearly these hypotheses are then either incorrect or incomplete for the species and interactions investigated here. While niche differences among species are necessary to mitigate the negative influence of competitive interactions and stabilize long-term, stable coexistence (Chesson, 2000, Narwani et al., 2013, Adler et al., 2007), contemporary coexistence theory tells us that the outcome of competition ultimately depends on a balance between two things – 1) relative fitness

differences among species, which define their competitive inequalities and lead to competitive exclusion, and 2) their niche differences, which overcome competitive inequalities and stabilize coexistence (Chesson, 2000, Narwani et al., 2013, Godoy et al., 2014, Adler et al., 2007). If transcriptomic differences accumulated over evolutionary time contribute more on average to relative fitness differences than niche differences, then they would tend to limit coexistence, not promote it (Hillerislambers et al., 2012, Mayfield and Levine, 2010, Godoy et al., 2014). In this paper however, we did not directly estimate niche and fitness differences, and therefore we currently do not have evidence to directly support this hypothesis, but it is consistent with our results. Furthermore, our data suggest that similarity in expression across the many shared genes in the transcriptomes of these species are important for coexistence, and not just the expression levels of a few candidate genes. This may lend support to the idea that some phenotypes, i.e. those related to ecological fitness, are "degenerate" or generated by multiple molecular or genetic pathways rather than deterministically by single pathways (Greenspan, 2012).

Niche differences are necessary in order for competing species to show long-term stable coexistence, and because our analysis shows that competition weakens with gene expression similarity, and not differences, the transcriptomic basis of niche differentiation must lie either in the expression of genes that were excluded from our analysis (i.e. non-commonly expressed genes), or in the expression of particular genes or gene families with particular functions (e.g. Alexander et al., 2015). It has been previously proposed that many genes in the genome are genes that all species need to express in order to survive and reproduce in a given environment – termed the "core genome" (Cordero and Polz, 2014). These genes are likely to encode essential metabolic and house-keeping functions (Cordero and Polz, 2014). Similarity in the expression of these genes would reflect similarity in the ecological fitness of species in this environment. Genes related to niche differentiation may then be rare (i.e. not observed in all species or populations), part of the "flexible genome" (genes that display turnover in response to local, negative frequency-dependent selection), and would allow species to evolve unique phenotypes and functionalities over time (Cordero and Polz, 2014). This possibility has already been supported for some microbial taxa (Cordero and Polz, 2014), but not yet algae. Unfortunately, logistical constraints in our study limited further investigation of genes whose expression were not detected in all species. We were not able to distinguish low-level expression from the complete absence of expression because we did not implement RNA spikes in the sequencing

443 experiment (they come at the high cost of sequencing depth of actual samples). However, our
444 analysis of genes that were differentially regulated in bicultures did identify particular gene
445 functions that were contrastingly regulated in competition and facilitation (Fig. S3). This analysis
446 lends some support to the idea that niche differences and facilitative interactions may be
447 determined by the expression of a limited number of genes and gene families, rather than by
448 differentiation of expression over the whole transcriptome.

449 One of the more surprising and interesting results of our study is that many algal species
450 experience facilitation in biculture, and that facilitation was associated with a greater likelihood
451 of coexistence. Facilitation does not necessarily lead to coexistence in the Lotka-Volterra model
452 of competition, although it has been shown that positive species interactions can increase the
453 likelihood of coexistence among species in more explicit models of resource competition (Gross
454 2008). Consistent with estimates from prior studies (Venail et al., 2014, Fritschie et al., 2014),
455 almost a third of all possible species interactions resulted in an increase in density at steady state,
456 relative to monoculture. Despite their prevalence, facilitative and co-operative interactions are
457 understudied in phytoplankton, with the vast majority of theory and empirical research in algal
458 ecology being focused on competitive and predatory interactions (Tilman, 1982, Huisman and
459 Weissing, 1995, Schippers et al., 2001, Passarge et al., 2006, Benincà et al., 2009, Litchman and
460 Klausmeier, 2001, Klausmeier et al., 2004, Litchman and Klausmeier, 2008). As a result, little is
461 known about the mechanism by which facilitative interactions might occur. In this study, we
462 identified several molecular process gene functions that were preferentially up- or downregulated
463 in different categories of ecological interactions. These gene functions provide clues as to how
464 competitive and facilitative interactions differ at the molecular level. In particular, transporter
465 gene transcripts for a variety of nutrients including nitrate, sugars, and other micronutrients were
466 highly abundant in cultures experiencing negative ecological interactions, consistent with the
467 notion that these species are competing for inorganic resources. By contrast, cultures
468 experiencing positive interactions expressed a higher abundance of genes associated with the
469 core cellular metabolism (e.g., the Citric Acid Cycle) and carbon fixation through the Calvin
470 cycle, and a relatively low abundance of genes associated with the acquisition of nutrients. This
471 suggests that facilitated and overyielding species generate growth and yield benefits from a boost
472 in core metabolism.
473

Caveats and suggestions for future work

Our study offers new insight into the transcriptomic changes that are associated with phytoplankton species interactions and coexistence; but the study is not without limitations. First, for methodological reasons, our transcriptome-wide analysis of gene expression only allowed comparisons of genes that were expressed in all monocultures. Future work would benefit from the use of RNA spikes to determine detection limits and enable comparisons of absolute expression levels. This would allow an investigation of the relative roles of the expression of shared genes versus uniquely-expressed genes in determining species interactions, and in particular niche differentiation and facilitation among species. Second, we measured gene expression at the final time point of the competition experiment, however, tracking changes in gene expression through time would allow comparisons of gene expression between the exponential growth phase, the onset of density-dependence, and the full effects of competition. Gene expression changes over time would allow comparisons of molecular basis of resource-unlimited growth versus density-dependent growth in the presence of conspecific versus interspecific neighbors. Third, it is important to keep in mind that our inference of gene functions were, by necessity, based on transitive annotations. That is, we inferred the functions of genes of interest by finding a gene with similar sequence (a homologous sequence) that is already annotated. These reference annotations may also have been inferred in a similar way and so on, leading to a daisy chain of annotations (Iliopoulos et al., 2003). At some point along the chain, the gene's function was tested in the lab, but it may have been in a very distantly related organism. Final validation and confirmation of gene functions in any particular organism must still be achieved using gene knock-out or knock-down experiments. Fourth, as always, future studies would benefit from a greater biological replication. We were limited to three replicates in this study, but due to the plasticity and variability in gene expression responses, we recommend up to 10 biological replicates in the future to increase statistical power and the insure against library failure.

Lastly, our study is obviously limited in its scale and complexity. While the use of microcosms allowed us to control the environment and directly investigate associations between species interactions and patterns of gene expression, the homogenous, artificial and simplified nature of our study limits our broader conclusions about the occurrence or prevalence of similar phenomena in nature. For example, does gene expression similarity predict interaction strength

or coexistence of three or more species? Does the relationship between gene expression similarity and coexistence change in patchy, heterogeneous or fluctuating environments? All but one of the species pairs in our study have been known to co-occur in natural lakes across the U.S. according to the 2007 EPA National Lakes Assessment Survey (Naughton et al., 2015), but it is possible that spatial and temporal heterogeneity in climate and resource availability are more important in these natural systems than the strength of species interactions. Species gene expression responses to natural environmental variation may also be important in determining interaction strength and coexistence (e.g. Alexander et al., 2015). The role of temporal and spatial biotic and abiotic environmental complexity on gene expression, species interactions and coexistence is an exciting area for future follow-up research.

In conclusion, in contrast to the widely-held notion that phenotypic similarity leads to competitive exclusion, we found that similarity in gene expression among species across the transcriptome tends to lead to weaker competition, more likely facilitation and greater coexistence. This suggests that the expression of the majority of commonly-expressed genes is required for basic survival and fitness in a particular environment, while niche differences and facilitative interactions may be encoded by just a few, or possibly rare genes. We identified gene functions to be investigated directly for their role in determining different types of species interactions in the future.

Authors' contributions

AN, KJF, CD, THO and BJC designed the study, AN, KJF and BJC performed the research, MA performed the phylogenetic analysis, BB performed the transcriptomic analysis, AN performed the modeling and statistical analysis, AN wrote the manuscript, and all authors contributed substantially to the editing.

Acknowledgements

We would like to thank S. Beaudry, K. Kulacki, M. McCarthy, E. Mitchell, P. Parent, J. Thierer, P. Venail, C. Wilson, C. Zhou, and E. Zimmermann for help in sampling and sample processing. We thank B. Matthews and M. Tamminen for editorial comments on the manuscript. We thank J. Lavine for the production of an R package which compiles the biculture Lotka-Volterra model-fitting algorithm in the C programming language. This work was supported by a US National

535 Science Foundation DIMENSIONS of Biodiversity grant to T.H.O. (DEB-1046307), C.F.D.
536 (DEB-1046075) and B.J.C. (DEB-1046121).

537

538 Data accessibility statement

539 The data and pipelines supporting the transcriptomics results in this paper are available on
540 Github (https://github.com/bastodian). All other data and scripts are available on the Data
541 available from the Dryad Digital Repository (Narwani et al., 2017).

Tables

Table 1. Effects of gene expression similarity (GES) on interaction coefficients. Values in each cell indicate the size of the Spearman rank correlation coefficient (ρ) between GES and the interactions coefficients (α_{ij}) of individual species in biculture. Negative correlation coefficients with α_{ij}s indicate that GES was negatively associated with the strength of competition and, in some cases, positively associated with facilitation (negative α_{ij}s). GES values were measured either in monoculture or in biculture, and either across the shared transcriptome, or only the 17 candidate genes. †indicates the correlation is significant with $0.05<P<0.1$; * indicates the correlation is significant at $P\leq 0.05$.

Species/ Gene expression correlation	**All genes in biculture**	**All genes in monoculture**	**Candidate genes in biculture**	**Candidate genes in monoculture**
Chlorella sorokiniana	0.11	-0.25	**-0.47***	**-0.52***
Closteriopsis acicularis	**-0.72***	**-0.76***	0.21	0.37
Cosmarium turpinii	0.23	0.16	0.42	**0.59***
Pandorina charkowiensis	-0.27	-0.14	**-0.44***	-0.29

Scenedesmus acuminatus	-0.26	**-0.45***	-0.35	**-0.69***
Selenastrum capricornutum	**-0.45***	0.009	0.22	-0.15
Staurastrum punctulatum	**-0.72***	-0.31	-0.009	**-0.38**†
Tetraedron minimum	**-0.49***	**-0.41**†	**-0.55***	0.07

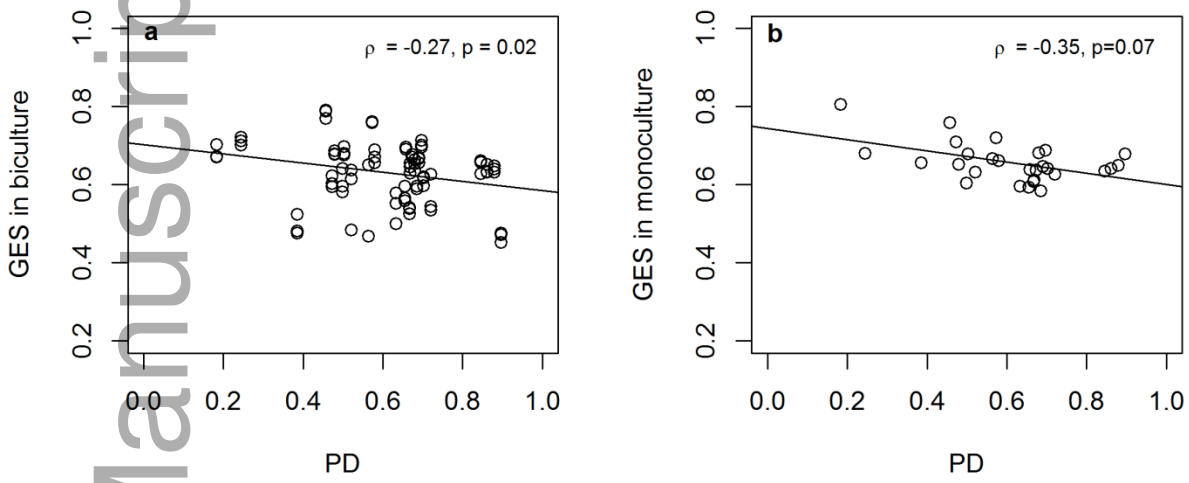
Table 2. Spearman rank correlation coefficients between expression levels of individual gene families (TPM) in individual species when in bicultures with the size of their interaction coefficients. The first 11 genes were proposed due to their ability to impact resource acquisition and metabolism. Genes 12-14 were chosen due to their role in vitamin production or metabolism and their potential ability to mediate facilitative interactions. Genes 15-17 were chosen due to their role in organic sugar production or metabolism and their potential ability to mediate heterotrophic / facilitative interactions. † indicates the correlation is marginally non-significant with 0.05<P<0.1; * significant at P≤0.05.

Gene/ Gene family	**Chlorella**	**Closteriopsis**	**Cosmarium**	**Pandorina**	**Scenedesmus**	**Selenastrum**	**Staurastrum**	**Tetraedron**
1. Carbonic Anhydrase	-0.21	-0.12	-0.08	-0.16	0.25	**-0.71***	NA	0.14
2. Glutamate Semialdehyde Aminetransferase	NA	-0.21	0.11	NA	**-0.72***	**-0.51***	0.003	-0.12
3. Iron Permease	0.28	0.26	NA	-0.23	**-0.46***	-0.58	NA	0.17
4. Light Harvesting Complex AB	**0.59***	0.01	0.02	-0.10	-0.36	-0.28	**-0.66***	0.06
5. Nitrogen Assimilation	0.18	0.25	NA	NA	-0.21	NA	NA	0.13

Regulatory Protein								
6. Nitrate Reductase	0.15	0.19	0.24	-0.30	-0.74	**-0.52***	**-0.46***	-0.06
7. Nitrate Transporter	**-0.63***	0.24	0.23	0.12	0.15	**-0.45***	**-0.50***	0.39
8. Nitrite Reductase	0.19	NA	0.35	**-0.560***	NA	**-0.50***	**-0.53***	-0.37
9. Nitrite Transporter	0.01	0.05	0.26	-0.16	0.04	**-0.51***	NA	0.16
10. Nitrogen Regulatory Protein	-0.01	0.24	0.22	-0.14	-0.19	**-0.42**†	-0.01	0.04
11. Phosphate Transporter	-0.11	0.06	0.17	-0.15	**-0.47***	**-0.53***	-0.05	0.13
12. Biotin, vitamin B7	0.23	0.09	0.20	-0.21	0.01	**-0.65***	0.08	0.21

13. Cobalamin, vitamin B12	0.01	-0.21	0.08	**-0.38**†	**-0.77***	**-0.60***	**-0.46***	-0.00
14. Thiamine, vitamin B1	0.08	0.14	0.12	-0.20	**-0.80***	**-0.57***	**-0.43**†	-0.10
15. Glucose	0.30	0.11	0.19	-0.22	**-0.47***	**-0.45**†	-0.25	0.18
16. Mannose	**-0.63***	0.22	0.12	NA	**-0.48***	**-0.68***	-0.20	0.16
17. Succinate	**0.41**†	0.10	0.14	-0.33	0.23	-0.28	0.02	-0.10

548Figures



549

550**Fig. 1.** The relationship between phylogenetic distance (PD) among species pairs and the gene expression similarity (GES) of all
551commonly expressed genes across two species in (a) biculture (N=84), and b) monoculture (N=28). Values in each panel are the
552Spearman rank correlations (ρ) and the p-value (p).

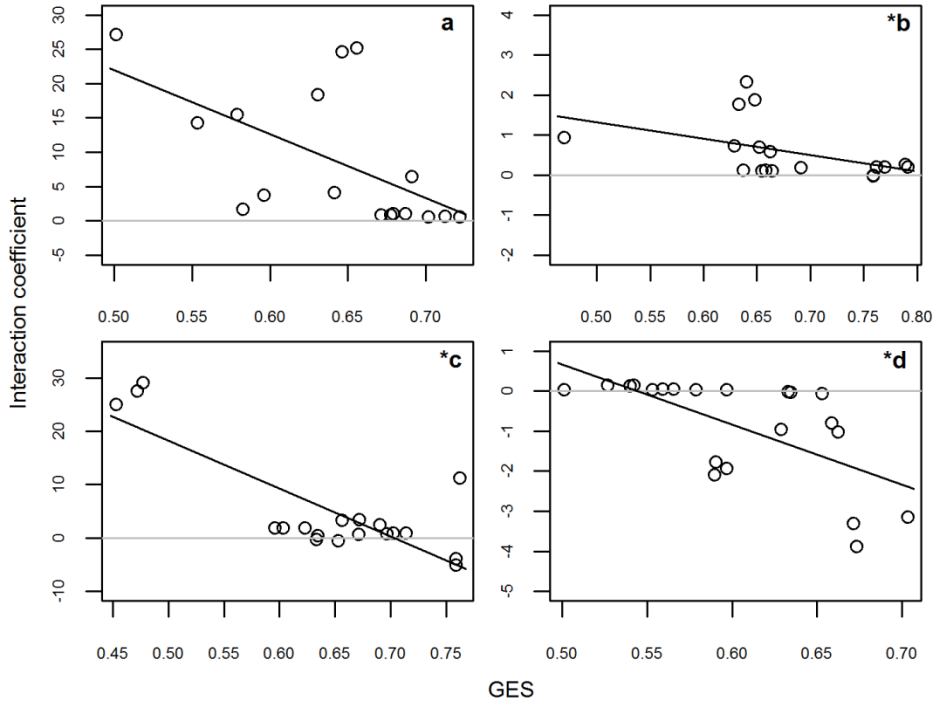


Fig. 2. Species interaction strengths, measured as Lotka-Volterra interaction coefficients (α_{ij}), as a function of gene expression similarity (GES) across all genes expressed in bicultures. Interaction coefficients indicate the per capita size and direction of the impact of individuals of another species in biculture on individuals of the focal species relative to the per capita impact of the focal species on itself. Positive interaction coefficients indicate competitive interactions (a negative effects on steady-state density) and negative coefficients indicate facilitative interactions (a positive effects on steady-state density). The focal species in each panel is: **a)** *Closteriopsis acicularis* (N=17), **b)** *Tetraedron minimum* (N=18), **c)** *Selenastrum capricornutum* (N=19), **d)** *Staurastrum punctulatum* (N=21). The horizontal grey line indicates an interaction coefficient value of zero. Above this line, species interactions are competitive, and below it they are facilitative. In the top two panels, higher values of gene expression similarity are associated with

weaker levels of competition, indicated by the decline in the size of the interaction coefficients (all are positive). In the bottom two panels, the interaction coefficients change from positive to negative, indicating a switch from competition to facilitation with increasing GES. Significant correlations between the gene expression correlation coefficients and the competition coefficients at $P \leq 0.05$ are indicated by an asterisk in the panel label. All interaction coefficients were sign·square-root transformed to aid visual interpretation in the figures.

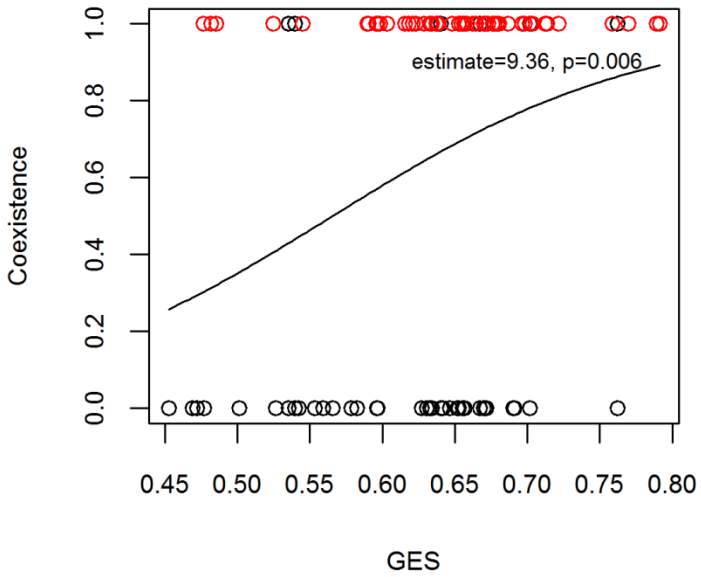


Fig. 3. Coexistence (0=no, 1=yes) as a function of gene expression similarity (GES) across all genes in biculture (N=80). Coexistence was estimated by using fitted interaction coefficients to simulate Lotka-Volterra models forward 100 days or 50 time steps and determining whether both species would have non-zero densities at the end of the simulation. Points in red are species pairs in which

one or both species had a negative interaction coefficient, indicating that the species benefited from the presence of the other species (i.e. facilitation). GES was a significant predictor of both coexistence ($\beta = 9.36$, $P = 0.006$), and the likelihood of positive species interactions ($\beta = 9.60$, $P = 0.005$).

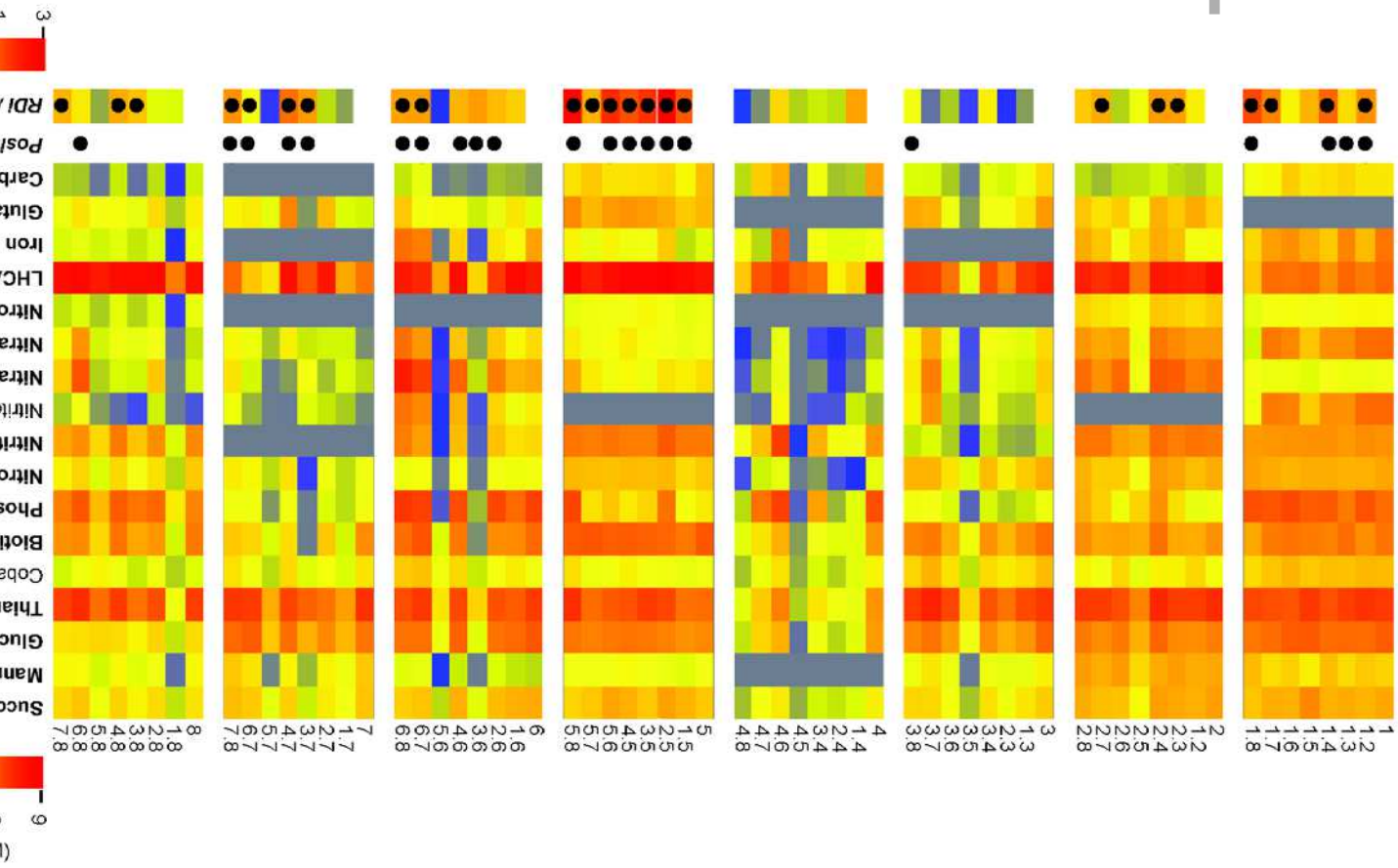


Fig. 4. Heat map representing natural log-transformed TPM values for the 17 candidate gene families in monocultures and all biculture combinations as well as natural log-transformed relative densities of each species (RDi) in biculture, and the presence (dot) or absence (no dot) of overyielding and facilitation (negative competition coefficients). Gene families which were significant predictors of relative density according to the Mantel tests ($P \leq 0.05$) are indicated in bold font, and the strength of the correlation coefficient is indicated in parentheses.

References

Ackerly, D. D. (2003) Community assembly, niche conservatism, and adaptive evolution in changing environments. *International Journal of Plant Sciences,* **164,** S165-S184.

Adler, P. B., Hillerislambers, J. & Levine, J. M. (2007) A niche for neutrality. *Ecology letters,* **10,** 95-104.

Alexander, H., Jenkins, B. D., Rynearson, T. A. & Dyhrman, S. T. (2015) Metatranscriptome analyses indicate resource partitioning between diatoms in the field. *Proceedings of the National Academy of Sciences,* **112,** E2182-E2190.

Alexandrou, M. A., Cardinale, B. J., Hall, J. D., Delwiche, C. F., Fritschie, K., Narwani, A., Venail, P. A., Bentlage, B., Pankey, M. S. & Oakley, T. H. (2015) Evolutionary relatedness does not predict competition and co-occurrence in natural or experimental communities of green algae. *Proceedings of the Royal Society B-Biological Sciences,* **282**.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) Basic local alignment search tool. *Journal of Molecular Biology,* **215,** 403-410.

Amin, S. A., Hmelo, L. R., van Tol, H. M., Durham, B. P., Carlson, L. T., Heal, K. R., Morales, R. L., Berthiaume, C. T., Parker, M. S., Djunaedi, B., Ingalls, A. E., Parsek, M. R., Moran, M. A. & Armbrust, E. V. (2015) Interaction and signalling between a cosmopolitan phytoplankton and associated bacteria. *Nature,* **522,** 98-U253.

Arfi, Y., Levasseur, A. & Record, E. (2013) Differential Gene Expression in Pycnoporus coccineus during Interspecific Mycelial Interactions with Different Competitors. *Applied and Environmental Microbiology,* **79,** 6626-6636.

Ashton, K. G. (2004) Comparing phylogenetic signal in intraspecific and interspecific body size datasets. *Journal of Evolutionary Biology,* **17,** 1157-1161.

Benincà, E., Johnk, K. D., Heerkloss, R. & Huisman, J. (2009) Coupled predator-prey oscillations in a chaotic food web. *Ecology letters,* **12,** 1367-78.

Best, R. J., Caulk, N. C. & Stachowicz, J. J. (2013) Trait vs. phylogenetic diversity as predictors of competition and community composition in herbivorous marine amphipods. *Ecology Letters,* **16,** 72-80.

Blomberg, S. P., Garland, T. & Ives, A. R. (2003) Testing for phylogenetic signal in comparative data: Behavioral traits are more labile. *Evolution; international journal of organic evolution,* **57,** 717-745.

Brawand, D., Soumillon, M., Necsulea, A., Julien, P., Csardi, G., Harrigan, P., Weier, M., Liechti, A., Aximu-Petri, A., Kircher, M., Albert, F. W., Zeller, U., Khaitovich, P., Grutzner, F., Bergmann, S., Nielsen,

R., Paabo, S. & Kaessmann, H. (2011) The evolution of gene expression levels in mammalian organs. *Nature,* **478,** 343-+.

Cadotte, M. W., Davies, T. J., Regetz, J., Kembel, S. W., Cleland, E. & Oakley, T. H. (2010) Phylogenetic diversity metrics for ecological communities: integrating species richness, abundance and evolutionary history. *Ecology Letters,* **13,** 96-105.

Cavender-Bares, J., Kozak, K. H., Fine, P. V. A. & Kembel, S. W. (2009) The merging of community ecology and phylogenetic biology. *Ecology letters,* **12,** 693-715.

Chesson, P. (2000) Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics,* **31,** 343-366.

Coggill, P., Finn, R. D. & Bateman, A. (2008) Identifying protein domains with the Pfam database. *Current Protocols in Bioinformatics,* **suppl. 23**.

Conesa, A. & Gotz, S. (2008) Blast2GO: A comprehensive suite for functional analysis in plant genomics. *International journal of plant genomics,* **2008,** 619832-619832.

Cooper, E. D., Bentlage, B., Gibbons, T. R., Bachvaroff, T. R. & Delwiche, C. F. (2014) Metatranscriptome profiling of a harmful algal bloom. *Harmful Algae,* **37,** 75-83.

Cooper, N., Jetz, W. & Freckleton, R. P. (2010) Phylogenetic comparative approaches for studying niche conservatism. *Journal of Evolutionary Biology,* **23,** 2529-2539.

Cordero, O. X. & Polz, M. F. (2014) Explaining microbial genomic diversity in light of evolutionary ecology. *Nature Reviews Microbiology,* **12,** 263-273.

Croft, M. T., Warren, M. J. & Smith, A. G. (2006) Algae need their vitamins. *Eukaryotic Cell,* **5,** 1175-1183.

Darwin, C. (1859) On the Origin of Species. John Murray, London.

Davies, T. J., Meiri, S., Barraclough, T. G. & Gittleman, J. L. (2007) Species co-existence and character divergence across carnivores. *Ecology Letters,* **10,** 146-152.

Dyhrman, S. T., Jenkins, B. D., Rynearson, T. A., Saito, M. A., Mercier, M. L., Alexander, H., Whitney, L. P., Drzewianowski, A., Bulygin, V. V., Bertrand, E. M., Wu, Z., Benitez-Nelson, C. & Heithoff, A. (2012) The transcriptome and proteome of the diatom *Thalassiosira pseudonana* reveal a diverse phosphorus stress response. *PLoS ONE,* **7,** e33768.

Eddy, S. R. (2011) Accelerated Profile HMM Searches. *Plos Computational Biology,* **7**.

Faith, D. P. (1992) Conservation evaluation and phylogenetic diversity. *Biological Conservation,* **61,** 1-10.

Frischkorn, K. R., Harke, M. J., Gobler, C. J. & Dyhrman, S. T. (2014) De novo assembly of *Aureococcus anophagefferens* transcriptomes reveals diverse responses to the low nutrient and low light conditions present during blooms. *Frontiers in Microbiology,* **5,** 375.

Fritschie, K. J., Cardinale, B. J., Alexandrou, M. A. & Oakley, T. H. (2014) Evolutionary history and the strength of species interactions: testing the phylogenetic limiting similarity hypothesis. *Ecology,* **95,** 1407-1417.

Gautam, K., Pareek, A. & Sharma, D. K. (2013) Biochemical composition of green alga Chlorella minutissima in mixotrophic cultures under the effect of different carbon sources. *Journal of Bioscience and Bioengineering,* **116,** 624-627.

Giovannoni, S. J. (2012) Vitamins in the sea. *Proceedings of the National Academy of Sciences of the United States of America,* **109,** 13888-13889.

Godoy, O., Kraft, N. J. B. & Levine, J. M. (2014) Phylogenetic relatedness and the determinants of competitive outcomes. *Ecology Letters,* **17,** 836-844.

Grant, P. R. & Grant, B. R. (2006) Evolution of character displacement in Darwin's finches. *Science,* **313,** 224-6.

Greenspan, R. J. (2012) Biological Indeterminacy. *Science and Engineering Ethics,* **18,** 447-452.

Grishkevich, V. & Yanai, I. (2013) The genomic determinants of genotype x environment interactions in gene expression. *Trends in Genetics,* **29,** 479-487.

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B., Lieber, M., MacManes, M. D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C. N., Henschel, R., Leduc, R. D., Friedman, N. & Regev, A. (2013) De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols,* **8,** 1494-1512.

Hillerislambers, J., Adler, P. B., Harpole, W. S., Levine, J. M. & Mayfield, M. M. (2012) Rethinking community assembly through the lens of coexistence theory. *Annual Review of Ecology and Systematics,* **43,** 227-248.

Huisman, J. & Weissing, F. J. (1995) Competition for nutrients and light in a mixed water column - A theoretical analysis. *American Naturalist,* **146,** 536-564.

Hutchinson, G. E. (1961) The paradox of the plankton. *American Naturalist,* **95,** 137-145.

Iliopoulos, I., Tsoka, S., Andrade, M. A., Enright, A. J., Carroll, M., Poullet, P., Promponas, V., Liakopoulos, T., Palaios, G., Pasquier, C., Hamodrakas, S., Tamames, J., Yagnik, A. T., Tramontano, A., Devos, D., Blaschke, C., Valencia, A., Brett, D., Martin, D., Leroy, C., Rigoutsos, I., Sander, C. & Ouzounis, C. A. (2003) Evaluation of annotation strategies using an entire genome sequence. *Bioinformatics,* **19,** 717-726.

679Ingram, T. (2015) Diversification of body shape in Sebastes rockfishes ofthe north-east Pacific. *Biological*
680 *Journal of the Linnean Society,* **116,** 805-818.
681Khaitovich, P., Enard, W., Lachmann, M. & Paabo, S. (2006) Evolution of primate gene expression.
682 *Nature Reviews Genetics,* **7,** 693-702.
683Khaitovich, P., Paabo, S. & Weiss, G. (2005) Toward a neutral evolutionary model of gene expression.
684 *Genetics,* **170,** 929-939.
685Kilham, S. S., Kreeger, D. A., Lynn, S. G., Goulden, C. E. & Herrera, L. (1998) COMBO: a defined
686 freshwater culture medium for algae and zooplankton. *Hydrobiologia,* **377,** 147-159.
687Klausmeier, C. A., Litchman, E. & Levin, S. A. (2004) Phytoplankton growth and stoichiometry under
688 multiple nutrient limitation. *Limnology and Oceanography,* **49,** 1463-1470.
689Kraft, N. J. B., Godoy, O. & Levine, J. M. (2015) Plant functional traits and the multidimensional nature of
690 species coexistence. *Proceedings of the National Academy of Sciences of the United States of*
691 *America,* **112,** 797-802.
692Kunstler, G., Lavergne, S., Courbaud, B., Thuiller, W., Vieilledent, G., Zimmermann, N. E., Kattge, J. &
693 Coomes, D. A. (2012) Competitive interactions between forest trees are driven by species' trait
694 hierarchy, not phylogenetic or functional similarity: implications for forest community assembly.
695 *Ecology letters,* **15,** 831-840.
696Lamichhaney, S., Berglund, J., Almen, M. S., Maqbool, K., Grabherr, M., Martinez-Barrio, A., Promerova,
697 M., Rubin, C. J., Wang, C., Zamani, N., Grant, B. R., Grant, P. R., Webster, M. T. & Andersson, L.
698 (2015) Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature,*
699 **518,** 16.
700Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. (2009) Ultrafast and memory-efficient alignment of
701 short DNA sequences to the human genome. *Genome Biology,* **10**.
702Levy, R. & Borenstein, E. (2012) Reverse Ecology: From Systems to Environments and Back. *Evolutionary*
703 *Systems Biology* (ed S. O. Soyer), pp. 329-345. Springer New York, New York, NY.
704Levy, R. & Borenstein, E. (2013) Metabolic modeling of species interaction in the human microbiome
705 elucidates community-level assembly rules. *Proceedings of the National Academy of Sciences,*
706 **110,** 12804-12809.
707Li, B. & Dewey, C. N. (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without
708 a reference genome. *Bmc Bioinformatics,* **12**.
709Li, T., Zheng, Y., Yu, L. & Chen, S. (2014) Mixotrophic cultivation of a Chlorella sorokiniana strain for
710 enhanced biomass and lipid production. *Biomass & Bioenergy,* **66,** 204-213.

711Li, X. G., Wu, H. X. & Southerton, S. G. (2010) Comparative genomics reveals conservative evolution of
712 the xylem transcriptome in vascular plants. *Bmc Evolutionary Biology,* **10,** 14.
713Liao, B.-Y. & Zhang, J. (2006) Low rates of expression profile divergence in highly expressed genes and
714 tissue-specific genes during mammalian evolution. *Molecular Biology and Evolution,* **23,** 1119-
715 1128.
716Lindemann, S. R., Bernstein, H. C., Song, H.-S., Fredrickson, J. K., Fields, M. W., Shou, W., Johnson, D. R. &
717 Beliaev, A. S. (2016) Engineering microbial consortia for controllable outputs. *ISME J*.
718Litchman, E. & Klausmeier, C. A. (2001) Competition of phytoplankton under fluctuating light. *American*
719 *Naturalist,* **157,** 170-187.
720Litchman, E. & Klausmeier, C. A. (2008) Trait-Based Community Ecology of Phytoplankton. *Annual*
721 *Review of Ecology Evolution and Systematics,* **39,** 615-639.
722Lotka, A. J. (1920) Analytical note on certain rhythmic relations in organic systems. *Proceedings of the*
723 *National Academy of Sciences,* **6,** 410-415.
724Lotka, A. J. (1925) *Elements of Physical Biology.* Williams and Wilkins.
725MacArthur, R. & Levins, R. (1967) The limiting similarity, convergence, and divergence of coexisting
726 species. *American Naturalist,* **101,** 377-385.
727Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads.
728 *EMBnet.journal,* **17,** 10-12.
729Mayfield, M. M. & Levine, J. M. (2010) Opposing effects of competitive exclusion on the phylogenetic
730 structure of communities. *Ecology letters,* **13,** 1085-1093.
731Mi, H. Y., Lazareva-Ulitsky, B., Loo, R., Kejariwal, A., Vandergriff, J., Rabkin, S., Guo, N., Muruganujan, A.,
732 Doremieux, O., Campbell, M. J., Kitano, H. & Thomas, P. D. (2005) The PANTHER database of
733 protein families, subfamilies, functions and pathways. *Nucleic Acids Research,* **33,** D284-D288.
734Narwani, A., Alexandrou, M. A., Oakley, T. H., Carroll, I. T. & Cardinale, B. J. (2013) Experimental
735 evidence that evolutionary relatedness does not affect the ecological mechanisms of
736 coexistence in freshwater green algae. *Ecology letters,* **16,** 1373-1381.
737Narwani, A., Bentlage, B., Alexandrou, M. A., Fritschie, K. J., Delwiche, C., Oakley, T. H. & Cardinale, B. J.
738 (2017) Data from: Ecological interactions and coexistence are predicted by gene expression
739 similarity in freshwater green algae. Dryad Digital Repository. doi:10.5061/dryad.2959s.
740Naughton, H. R., Alexandrou, M. A., Oakley, T. H. & Cardinale, B. J. (2015) Phylogenetic distance does not
741 predict competition in green algal communities. *Ecosphere,* **6**.

Passarge, J., Hol, S., Escher, M. & Huisman, J. (2006) Competition for nutrients and light: Stable coexistence, alternative stable states, or competitive exclusion? *Ecological Monographs,* **76,** 57-72.

Pfennig, K. S. & Pfennig, D. W. (2009) Character displacement: ecological and reproductive responses to a common evolutionary problem. *Quarterly Review of Biology,* **84,** 253-276.

Robinson, M. D., McCarthy, D. J. & Smyth, G. K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics,* **26,** 139-140.

Robinson, M. D. & Oshlack, A. (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology,* **11**.

Schippers, P., Verschoor, A. M., Vos, M. & Mooij, W. M. (2001) Does "supersaturated coexistence" resolve the "paradox of the plankton"? *Ecology letters,* **4,** 404-407.

Schluter, D. (2000) Ecological character displacement in adaptive radiation. *American Naturalist,* **156,** S4-S16.

Schulze, S., Schleicher, J., Guthke, R. & Linde, J. (2016) How to Predict Molecular Interactions between Species? *Frontiers in Microbiology,* **7**.

Tanoi, T., Kawachi, M. & Watanabe, M. M. (2011) Effects of carbon source on growth and morphology of Botryococcus braunii. *Journal of Applied Phycology,* **23,** 25-33.

Tilman, D. (1982) *Resource Competition and Community Structure.* Princeton University Press, Princeton, New Jersey.

Tirosh, I., Weinberger, A., Carmi, M. & Barkai, N. (2006) A genetic signature of interspecies variations in gene expression. *Nature Genetics,* **38,** 830-834.

Uebbing, S., Kunstner, A., Makinen, H., Backstrom, N., Bolivar, P., Burri, R., Dutoit, L., Mugal, C. F., Nater, A., Aken, B., Flicek, P., Martin, F. J., Searle, S. M. J. & Ellegren, H. (2016) Divergence in gene expression within and between two closely related flycatcher species. *Molecular ecology,* **25,** 2015-28.

Venail, P., Gross, K., Oakley, T. H., Narwani, A., Allan, E., Flombaum, P., Isbell, F., Joshi, J., Reich, P. B., Tilman, D., van Ruijven, J. & Cardinale, B. J. (2015) Species richness, but not phylogenetic diversity, influences community biomass production and temporal stability in a re-examination of 16 grassland biodiversity studies. *Functional Ecology,* **29,** 615-626.

Venail, P. A., Narwani, A., Fritschie, K., Alexandrou, M. A., Oakley, T. H. & Cardinale, B. J. (2014) The influence of phylogenetic relatedness on species interactions among freshwater green algae in a mesocosm experiment. *Journal of Ecology,* **102,** 1288-1299.

774Violle, C., Nemergut, D. R., Pu, Z. C. & Jiang, L. (2011) Phylogenetic limiting similarity and competitive
775 exclusion. *Ecology letters,* **14,** 782-787.
776Volterra, V. (1928) Variations and fluctuations of the number of individuals in animal species living
777 together. *Journal du Conseil Permanent International pour l'Exploration de la Mer,* **3,** 3–51.
778Wagner, G. P., Kin, K. & Lynch, V. J. (2012) Measurement of mRNA abundance using RNA-seq data:
779 RPKM measure is inconsistent among samples. *Theory in Biosciences,* **131,** 281-285.
780Wanek, K. A. & Sturmbauer, C. (2015) Form, function and phylogeny: comparative morphometrics of
781 Lake Tanganyika's cichlid tribe Tropheini. *Zoologica Scripta,* **44,** 362-373.
782Webb, C. O., Ackerly, D. D., McPeek, M. A. & Donoghue, M. J. (2002) Phylogenies and community
783 ecology. *Annual Review of Ecology and Systematics,* **33,** 475-505.
784Whittle, C. A., Sun, Y. & Johannesson, H. (2014) Dynamics of transcriptome evolution in the model
785 eukaryote Neurospora. *Journal of Evolutionary Biology,* **27,** 1125-1135.
786Wiens, J. J., Ackerly, D. D., Allen, A. P., Anacker, B. L., Buckley, L. B., Cornell, H. V., Damschen, E. I.,
787 Davies, T. J., Grytnes, J. A., Harrison, S. P., Hawkins, B. A., Holt, R. D., McCain, C. M. & Stephens,
788 P. R. (2010) Niche conservatism as an emerging principle in ecology and conservation biology.
789 *Ecology letters,* **13,** 1310-1324.
790Wohlrab, S., Tillmann, U., Cembella, A. & John, U. (2016) Trait changes induced by species interactions in
791 two phenotypically distinct strains of a marine dinoflagellate. *ISME J*.
792Yang, R. L. & Wang, X. F. (2013) Organ Evolution in Angiosperms Driven by Correlated Divergences of
793 Gene Sequences and Expression Patterns. *Plant Cell,* **25,** 71-82.

794