Title:

Pond Crash Forensics: Presumptive identification of pond crash agents by next generation sequencing in replicate raceway mass cultures of *Nannochloropsis salina*.

Laura T. Carney^{ac}, Joshua S. Wilkenfield^{bc}, Pam D. Lane^a, Owen D. Solberg^a, Zachary B. Fuqua^b, Nina G. Cornelius^b, Shaunette Gillespie^b, Kelly P. Williams^a, Tzachi M. Samocha^b, Todd W. Lane^{a*}

^aSystems Biology, Sandia National Laboratories, PO Box 969, MS 9671, Livermore, CA 94551 USA

^bTexas A&M AgriLife Research Mariculture Laboratory at Flour Bluff, 4301 Waldron Road, Corpus Christi, TX, 78418 USA

^cPresent address: Heliae Development, LLC, 578 E. Germann Road, Gilbert, AZ 85297

* Correspondence:

Dr. Todd Lane
Systems Biology
Sandia National Laboratories
PO Box 969, MS 9292
Livermore, CA 94551 USA
+1-925-294-2057
twlane@sandia.gov

ABSTRACT

Productivity of algal mass culture can be severely reduced by contaminating organisms. It is, therefore, important to identify contaminants, determine their effect on productivity and, ultimately, develop countermeasures against such contamination. In the present study we utilized microbiome analysis by second-generation sequencing of small subunit rRNA genes to characterize the predator and pathogen burden of open raceway cultures of *Nannochloropsis salina*. Samples were analyzed from replicate raceways before and after crashes. In one culture cycle, we identified two algivorous species, the rotifer *Brachionus* and gastrotrich *Chaetonotus*, the presence of which may have contributed to the loss of algal biomass. In the second culture cycle, the raceways were treated with hypochlorite in an unsuccessful attempt to interdict the crash. Our analyses were shown to be an effective strategy for the identification of the biological contaminants and the characterization of intervention strategies.

Keywords: algae mass culture, biocontamination, biomass algal crash, second generation sequencing, Microbiome analysis

1. Introduction

One of the major challenges to achieving high rates of long-term production in microalgal mass cultures, is the ability to maintain culture stability. Lost days of production due to pond crashes can significantly lower annualized production rates (McBride et al., 2014; Richardson et al., 2014). Methods exist for the tracking of many of the abiotic parameters, such as temperature, pH, and nutrient depletion, which can contribute to pond instability, however understanding of the contribution of the biological agents has been difficult to achieve. Aside from a few well-known causative agents, the diversity of algal pathogens, parasites, and predators has not been

well characterized. What is understood, however, is that the economic impact of biological agents on the mass culture of algae can be substantial. Algal pathogens, primarily bacteria (Cole, 1982; Fukami et al., 1997) and fungi (Fott, 1967; Hoffman et al., 2008), and zooplankton grazers (e.g., Park et al., 2011) can cause severe algal biomass loss. Such decline can occur rapidly; for example, metazoan rotifers can ingest 200 algal cells min^{-1} rotifer $^{-1}$ (Hirayama and Ogawa, 1972) while doubling their density in 24 hrs (Yúfera and Navarro, 1995).

In addition to being useful in community characterization (e.g., Bik et al., 2012; Li et al., 2011), second generation sequencing (SGS) has brought about a revolution in detection and identification of new etiological agents in diverse systems. SGS of molecular barcode regions, such as ribosomal RNA genes and mitochondrial cytochrome oxidase, can identify agents to various taxonomic levels, providing measures of richness and diversity for bacterial (e.g., Bartram et al., 2011) and eukaryotic (e.g., Dunthorn et al., 2012) communities. Sequencing of specific single variable regions of the small subunit ribosomal RNA (SSU rRNA) gene can, generally, provide genus level identification (Tang et al., 2012). Species level identification can be obtained with sequence data from multiple variable regions or additional molecular barcodes such as the internal transcribed spacer (ITS) region of the ribosomal RNA precursor transcript (Baldwin, 1992). In terms of algal mass culture operations, it may not be necessary to identify an agent beyond the genus level.

There are several sources of bias, including differential nucleic acid extraction and amplification efficiencies that prevent sequencing datasets from being a quantitative measure of the relative abundances of species present in pond samples. Thus sequence data by itself is not sufficient to unambiguously identify the etiological agent of a crash, but rather, provides a presumptive or conjectural, identification of potential etiological agents. Confirmatory

identification requires isolation of the agent, recapitulation of the crash, and re-isolation of the microorganism from the second infection thus fulfilling Koch's postulates (McBride et al., 2014). When isolation of the etiological agent is desired, sequence data can be used to generate genetic probes to track the agent and to facilitate the purification process. The increases in sequencing throughput has enabled multiplexing of libraries and driven down the per-sample costs of analysis and made SGS a practical solution for the identification of unknown deleterious species in algae culture samples.

Although some expertise in the nucleic acid manipulation and analysis is required, SGS does not require expertise in taxonomic identification via microscopy and is more effective at identifying morphological variants or low abundance species within a sample. SGS also has advantages over oligonucleotide-mediated methods such as PCR, in that it is not dependent on *a priori* knowledge of the organisms present. Microbiome analysis can also inform the development of specific probes, primers or biomarkers for both beneficial and deleterious species, which will allow rapid near real time tracking of organisms of interest. However, given the current cost and time required, SGS is unlikely to supplant either of these alternative methods for routine monitoring.

The marine alga, *Nannochloropsis salina*, shows promise as a candidate species for potential conversion of algae biomass to biofuel (Chisti, 2007; Su et al., 2010). Here we report the use of SGS for the presumptive identification of etiological agents likely responsible for the *N. salina* culture instabilities in outdoor open raceways at the Texas A&M AgriLife Research Mariculture Laboratory, Corpus Christi, Texas.

2. Materials and Methods

2.1. Raceway Culture Conditions and Monitoring

Algae culture experiments were conducted in an outdoor algae tanks at the Texas A&M AgriLife Research Mariculture Laboratory at Flour Bluff, Corpus Christi, Texas. Six fiberglass raceways were used for this study (3 m x 1 m; surface area of 2.78 m²; 550 L with 20 cm culture depth). Raceways were equipped with a paddle wheel operating at 9 rpm, resulting in a flow rate of 50-60 cm sec⁻¹. CO₂ was supplied from compressed cylinders at a flow rate of 0.125 L min⁻¹ for 8 hr d⁻¹, during daylight hours. Describe the water source, treatment and other characteristics

Three replicate algae raceways (ARWs) for each cycle were inoculated on July 16, 2011 (Cycle 7B) and on September 17, 2011 (Cycle 8B) to a final *N. salina* cell density of 3.7 x 10⁶ and 12.3 x 10⁶ cells mL⁻¹, respectively. Nutrient concentrations are shown in Table 1 and physical parameters during the two cycles are given in Table 2. Contamination was monitored while making daily cell counts by hemacytometer and light microscopy observations. For cycle 7B, raceway culture samples (250-1,000 mL) were collected on the day of inoculation and again after conditions in two of the raceways had deteriorated significantly. For cycle 8B, samples were taken from the stock culture used as the inoculum because it was observed to be contaminated, and from three replicate raceways after contamination was observed, and then on day 17 after replicates were treated on days 10, 12 and 14, post inoculation with sodium hypochlorite (see figure 1b) in an attempt to control the observed contamination.

Samples were shipped overnight on ice to Sandia National Laboratories for SGS analysis. Biomass was harvested from each sample by centrifugation (10,000 x g for 10 minutes) and filtration of the supernatant (0.2 um). Biomass, harvested by these two sequential methods, was combined and stored at -80°C.

2.2. Library Preparation and Second Generation Sequencing

Genomic DNA was purified from 200 mg of pelleted biomass from each sample (ZR Fungal/Bacterial DNA MiniPrepTM, Zymo Research, Irvine, CA USA). For each sample, three replicate sequencing libraries were prepared for the hyper-variable V4 region of the eukaryotic SSU rRNA gene and the hyper-variable V6 region of the bacterial SSU rRNA using conditions and primers described in Carney et al. (2013). Multiplexed V4 and V6 libraries containing equimolar concentrations of PCR product from each sample were sequenced on Illumina GAI, HiSeq or MiSeq platforms, according to manufacturer's instructions (Illumina, San Diego CA),

2.3. Second Generation Sequencing Data Analysis

SGS data processing and Silva based analysis and classification was carried out according to the methods described in Carney et al. (2013). In an attempt to characterize the read pairs that remained unaligned after our Silva-based classification phase, we adopted a "join, cluster, and BLAST" approach. First, in order to join overlapping reads, read pairs were processed with FLASH (Magoc and Salzberg, 2011) using default settings. Next, reads were processed with CD-HIT (Li and Godzik, 2006) and summarized at the sequence cluster level. Finally, we used BLAST to query these clustered sequences against the full Silva 111 and Genbank nr databases. Significant alignments were manually inspected and reported at whatever taxonomic level was possible given the available annotations associated with the BLAST hits.

Correlation analysis was performed to determine the reproducibility of microbiome analysis. For each sample/timepoint, the observed raw taxonomic count prior to cluster analysis was used to calculate the Pearson correlation coefficient between technical replicates. The calculations were performed at all taxonomic levels.

3. Results and Discussion

We carried out microbiome analysis by SGS on samples taken from two sets of open raceway pond growth cycles referred to as Cycle 7B and 8B. In the analysis of the eukaryotic microbiome, of > 160 million paired reads that passed initial quality filters and were analyzed by the pipeline, 92% were mapped to genus-level in our edited version of the Silva 111 rRNA database (Table 3). In the analysis of the bacterial microbiome of > 140 million paired reads that passed initial quality filters and were analyzed by the pipeline, 73% hit to the Silva 111 rRNA database, but only 17% on average were mapped to genus-level (Table 3). Replicate datasets 2 and 3 generated for each variable region were highly correlated ($r > 99.9\%$), thus only correlation analyses between rep 1 and 2 are reported. Replicates for eukaryotic variable region V4 was highly correlated at the genus level and above (mean correlation genus [$\pm$ 1SD]: $r = 99.1 \pm 1.7\%$). However, the V6 datasets were only moderately correlated at the phylum and genus level (mean correlation for phylum [$\pm$ 1SD]: $r = 78.8 \pm 24.6\%$; for genus: $r = 68.0 \pm 34.6\%$). The weaker correlation for V6 may be attributed to the fact that, for some of the samples, a majority of reads were not aligned to a sequence in the Silva database. In other words, the greater correlation for V4 versus V6 is explainable by the fact that Silva release 111 provides a much better survey of Eukaryotic diversity than it does Prokaryotic diversity. Because, currently, there are no known reports of archaeal pathogens of eukaryotic microalgae (Mayali and Azam, 2004), the archaeal community structure was not surveyed in this study. In addition, we attempted but were unable to detect phycodnaviruses in the samples by PCR (data not shown).

3.1. Forensic analysis of Cycle 7B

For Cycle 7B, microbiome analysis was carried out on samples from three biological

replicate raceways (referred to as ARW 6, ARW 8, and ARW 11) collected on the day of inoculation and on day 17 post inoculation (PI), after cell counts indicated severe algal biomass loss in two of the raceways (Figure 1a; ARW 8 and ARW 11) and moderate loss was occurring in the third (Figure 1a: ARW 6).

The eukaryotic SSU rRNA gene sequencing datasets derived from each of the three raceways in cycle 7B were initially dominated by *Nannochloropsis* (Figure 2a; $86.3 \pm 2.8\%$ of read pairs). Sequencing hits from competing algae genera, chrysophytes, fungi and an unknown stramenopile were also present but in much lower proportions (Figure 2b). By day 17 post inoculation (PI) the makeup of the sequencing dataset from each of the three replicate raceways had diverged from those derived from initial samples as well as from each other (Figure 2a). The relative abundance of *Nannochloropsis* sequencing hits decreased in all three datasets (to 56.4% in ARW 6, to 3.9% in ARW 8, and to 12.4% in ARW 11) while those from potentially deleterious algivores had increased: the suctorian ciliate *Acineta* (6.8%) and the potentially parasitic vampyrellid, *Vampyrella* (8.6%), in ARW 6, the rotifer *Brachionus* in ARW8 (72.2%), and the gastrotrich *Chaetonotus* in ARW11 (51.0%). In the cases of ARW8 and ARW11, the presence of *Brachionus* and *Chaetonotus* was confirmed by microscopy starting on days 11 and 12 PI, respectively (data not shown).

The ciliate, *Acineta*, mainly preys on other ciliates but also ingests microalgae (Fernandez-Leborans et al., 2000). Although Vampyrellids are primarily described as parasites of freshwater microalgae, some have been isolated from marine environments (reviewed by Hess et al., 2012). The grazer, *Brachionus*, can cause severe losses of algae biomass due to its high ingestion rates (Hirayama and Ogawa, 1972) and ability to reach high densities in a matter of days (>100 indiv. mL^{-1} , Yúfera and Navarro, 1995). *Chaetonotus* is a very common unsegmented worm-like

gastrotrich that feeds on bacteria, protists and algae (Strayer et al 2010).

Reads from other genera were found in datasets from day 17 PI but at lower abundances than the three main grazers (Figure 2a). A thraustochytrid, *Aplanochytrium* (synonymous with *Labyrinthuloides*), was found in all three datasets (> 2.8% of read pairs). This marine protist is known to parasitize macroalgae and is often associated with decaying algal tissue (Raghukumar, 2002). Vampyrellids were also found in the datasets from ARW8 by day 17 PI. Vorticella, a peritrich ciliate known to ingest bacteria and small phytoplankton (Kankaala and Eloranta, 1987, Fernandez-Leborans et al., 2009) was also detected in the ARW11 dataset (>7%). Hits from competing algae were found in ARW6 dataset (the raphidophyte *Heterosigma*) and ARW11 (the diatoms *Cymbella* or *Amphora*) and may have been less palatable than *Nannochloropsis* to the ciliate, *Acineta*, and the gastrotrich *Chaetonotus*. The rotifer *Brachionus* is known to be a non-specific grazer (Harzevili et al., 1997; Hotos, 2002) and likely reduced biomass of all algae in ARW8.

Our analysis suggested that the bacterial community in the three raceways from cycle 7B followed similar patterns as the eukaryotic community, showing almost identical microbiome sequencing results from the day of inoculation and then significant and unique divergences by day 17 PI (Figure 3). A large fraction of the read pairs from each sample on the day of inoculation mapped to uncultured organisms, primarily a δ -proteobacterium (Figure 3; Accession #: AB015242; $42.1 \pm 0.8\%$ of read pairs for all three samples). By day 17 PI in ARW6, when sequencing hits to the ciliate, *Acineta*, were beginning to dominate the eukaryotic libraries, sequencing hits to *Flavobacteriia* and the unknown δ -proteobacterium had increased in abundance (Figure 3; 11.7% and 16.4% of read pairs, respectively). There have been few studies of the algal associated bacteria in marine algal mass culture systems. The bacterial community

structure that we observed, with predominance of *δ-proteobacterium* in the sequencing library, was not typical of that previously reported for *Nannochloropsis* open mass culture systems (Nakase and Eguchi, 2007) *Flavobacteriia* may be stimulatory to algal productivity (Suminto Hirayama, 1996) but also plays an algicidal role at the end of algal blooms (Fukami et al., 1992; Kirchman, 2002). In ARW8, where a crash in algae biomass may have been caused by *Brachionus* grazing, sequencing hits for *Balneola*, the unknown *δ-proteobacterium*, *Pseudomonas*, and *α-proteobacterium* *Hyphomonas* or *Anderseniella* were in high relative abundance (Figure 3; (11.8%, 8.7%, 5.3% and 6.9% of read pairs, respectively). In ARW11 where the gastrotrich, *Chaetonotus*, was the dominant eukaryote in the dataset, the *Cytophagia*, some of which are known to lyse algae (Cole, 1982), were largely abundant (16% of read pairs) primarily due to the presence of *Aureibacter* (Figure 3). Hits for multiple *α-proteobacteria*, *Balneola*, and *Thermonema* also increased in abundance. A small fraction of hits to unclassified bacterial strains were also detected in all three raceways on day 17 PI (Figure 3).

Our analysis also suggests that there may be signature bacterial communities associated with different algivores. In cycle 7B, the bacterial sequencing data sets derived from the microbiome of samples that displayed high read counts of either the rotifer, *Brachionus*, or the gastrotrich, *Chaetonotus*, contained a large fraction of *Bacteroidetes*; the *Sphingobacterium Balneola* in the case of *Brachionus* (ARW8) and the *Cytophagia Aureibacter*, in the case of *Chaetonotus* (ARW11). In addition, both ARW8 and ARW 11 contained *Pseudomonas* and representatives of the *α-proteobacteria*. These results are consistent with previously published analyses of the bacterial community associated with *Brachionus* (Ishino et al., 2012) but there are no reports for *Chaetonotus*-specific communities in the literature. Understanding of predator-associated microbiomes could conceivably lead to detection of or countermeasures against those predators.

Predator-specific bacterial are likely to be present in higher numerical abundance than the predator itself and could be used as an indicator species, particularly in the early stages of infestation. It is also conceivable that pairwise, predator/bacterial, interactions are beneficial to the predator and, thus, pond microbiomes might be designed and controlled to disrupt those interactions.

3.2. Forensic analysis of Cycle 8B

For Cycle 8B, we analyzed samples from three biological replicate raceways (referred to as ARW3, ARW7, and ARW12) and the culture used to inoculate the raceways (inoculum), for which some contamination by bacteria and flagellates was initially observed at the time of raceway inoculation. The inoculum was sampled on the day the raceways were inoculated (day 0) and again on day 9 PI, after the inoculum was observed by microscopy to be contaminated with achloroplastic species. Samples from each of the parallel raceways were taken on day 10 PI after an initial decrease in cell density was observed in ARW 12 (Figure 1b) and on day 17 PI after all three raceways had been treated with hypochlorite on days 10, 12 and 14 PI. In ARW3 and ARW7, the eukaryotic community structure was different on day 17 than that observed on day 10 and all three raceways suffered significant algae biomass loss (Figure 1b). Because of staffing limitations at the culture facility, cell counts were not available for ARW 7 and ARW 12 on days 16 and 17.

The relative abundance of *Nannochloropsis* reads remained high 9-10 days PI in both the inoculum and raceways (Figure 4a; $87.9 \pm 6.6\%$ of read pairs across samples) with only minor proportions of potentially deleterious organisms (Figure 2b; Thraustochytrids and ciliates). However, on day 17 PI, higher numbers of sequencing hits on non-photosynthetic eukaryotes were observed (Figure 4). In ARW 3, large fractions of potential algal parasites *Aplanochytrium*

and vampyrellids (primarily *Vampyrella*) were detected (Figure 4a and b; 8.8% and 18% of read pairs, respectively). Reads that matched an unknown freshwater eukaryote (Accession #: AB622340) were also detected (Figure 4a; 20.5%). In ARW 7, the free-living limax amoeba, *Nolandella*, was detected (Figure 4a; 9.8%), as well as competing algae and the *Labrynthulid*, *Aplanochytrium*, (Figure 4b). *Nolandella* are known to prey on a wide variety of organisms, including unicellular algae (Page, 1983). In ARW 12, very small fractions of *Aplanochytrium* (0.3%) and *Vampyrella* (1.3%) were detected (Figure 4b).

The bacterial community changed considerably after hypochlorite treatment (Figure 5). Prior to day 17 PI, the bacterial communities in the inoculum and raceways were dominated by a small number of species (Figure 5; >50% of read pairs with database matches). Similar to cycle 7B, the most abundant bacterium was an uncultured δ -proteobacterium. *Criblamydia* was also detected in ARW 3 and 7 day 10 PI (> 7%), which were populated by various amoebae on day 17 PI (see Figure 4). *Criblamydia* is an emerging human/animal pathogen and often associated with complex aquatic microbial communities where it may survive inside of amoebae (Thomas et al., 2006). On day 17 PI, the raceways were dominated by γ -proteobacteria, primarily *Alteromonas* (Figure 5; 9.9% for ARW 3, 51.6% for ARW 7 and 74.2% for ARW 12). Some species of *Alteromonas*, are known to lyse algae cells during bloom formation (Imai et al., 1995).

4. Conclusion

Because pond crashes are unpredictable, we chose to track a number of raceways that were in cultivation for other purposes. When crashes did occur we were given access to archived samples and existing, but sometimes limited, ancillary data. Decisions to intervene, in an attempt to reverse incipient crashes, were made by the raceway operators.

Our analysis of cycle 7B, demonstrates the partial fulfillment of Koch's postulates for pond

crashes. In the sequencing datasets derived from each crashed pond (ARW8 and ARW11) we have identified one or two agents present at high abundance in sequencing datasets from crashed ponds that were either undetectable, or present in very low abundance, in datasets from previous time points taken from the same pond or from a relatively healthy neighboring pond (ARW6). As with other forms of molecular analysis, SGS alone cannot satisfy Koch's postulates and conclusively identify the etiological agent of a pond crash, however, in many cases the sequence hits from the presumptive agent are in high abundance relative to other species and the known characteristics of the agent are not inconsistent with a potential role in algal biomass loss. With these limitations in mind, our study has identified a number of species, not previously implicated in pond crashes, which may be of interest to pond operators for consideration in monitoring schemes.

Along with killing algae, broad-spectrum chemical agents such as hypochlorite can affect the underlying microbial community structure in hitherto uncharacterized and potentially deleterious manner. In the cycle 8B raceways, attempts were made to reverse an incipient crash by treatment with increasing concentrations of hypochlorite. Given the limited and incomplete data from this cycle it is not possible to separate the effects of the hypochlorite treatments from those of the biological agents present and thus determine the root cause of the algal biomass loss. Although there was a clear impact on both the eukaryotic and prokaryotic community structures, the treatment failed to prevent, and may have been the cause of, continued algal biomass loss in all three ponds. In two of three raceways, this algal biomass loss was not accompanied by a large increase in the relative abundance of sequencing hits on eukaryotic heterotrophs such as observed in cycle 7B.. Thus, it is evident that crash interdiction strategies require additional study.

We have demonstrated microbiome analysis by SGS as a tool for the presumptive identification of eukaryotic and bacterial agents associated with pond crashes and chemical interventions. Such analysis may also be used to inform the modification of algae production management strategies in order to avoid unintended deleterious effects on the target biomass and supporting microbial community. In addition to known deleterious organisms such as *Brachionus*, we have provided initial evidence for additional organisms including *Chaetomonas*, *Vampyrella*, *Nolandella* and *Acineta*. Although Koch's postulates have not been satisfied, the presence of these organisms may be of concern to algal production facility operators.

Acknowledgements

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. Microbiome analysis at Sandia National Laboratories was supported by the BioEnergy Technology Office, U.S. Department of Energy under Award #NL0022897.

References

1. Baldwin, B.G., 1992. Phylogenetic utility of the internal transcribed spacers of nuclear ribosomal DNA in plants: an example from the Compositae. *Molec. Phylo. Evol.* 1, 3-16.
2. Bartram, A.K., Lynch, M.D.J., Stearns, J.C., Moreno-Hagelsieb, G., Neufeld, J.D., 2011. Generation of multimillion-sequence 16S rRNA gene libraries from complex microbial communities by assembling paired-end Illumina reads. *App. Env. Microbiol.* 77, 3846-3852.

3. Bik, H.M., Halanych, K.M., Sharma, J., Thomas, W.K., 2012. Dramatic shifts in benthic microbial eukaryote communities following the deepwater horizon oil spill. PLoS ONE 7:e38550. doi:10.1371/journal.pone.0038550.
4. Carney, L.T., Reinsch, S.S., Lane, P. D., Solberg, O.D., Jansen, L.S., Williams, K.P., Trent, J. D. and Lane, T.W., 2013. Biocontaminant Analysis of A Microalgal Mass Culture Growing in Municipal Wastewater in a Prototype OMEGA Photobioreactor. Algal Res. dx.doi.org/10.1016/j.algal.2013.11.006.
5. Chisti, Y., 2007. Biodiesel from microalgae. Biotechnol. Adv. 25, 294-306.
6. Cole, J.J., 1982. Interactions between bacteria and algae in aquatic ecosystems. Ann. Rev. Ecol. Syst. 13, 291-314.
7. Dunthorn, M., Klier, J., Bunge, J., Stoeck, T., 2012. Comparing the hyper-variable V4 and V9 regions of the small subunit rDNA for assessment of ciliate environmental diversity. J. Eukaryotic Microbiol. 59, 185-187.
8. Fernandez-Leborans, G., Tato-Porto, M.L., 2000. A review of the species of protozoan epibionts on crustaceans. II. Suctorian ciliates. Custaceana 73, 1205-1237.
9. Fott, B., 1967. *Phlyctidium scenedesmi* spec. nova, a new chytrid destroying mass cultures of algae. Z. Allg. Mikrobiol. 7, 97-102.
10. Fukami, K., Nishijima, T., Ishida, Y., 1997. Stimulative and inhibitory effects of bacteria on the growth of microalgae. Hydrobiol. 358, 185-191.
11. Fukami, K., Yuzawa, A., Nishijima, T., Hata, Y., 1992. Isolation and properties of a bacterium inhibiting the growth of *Gymnodinium nagasakiense*. Nihon Suisan Gakkai-shi 58, 1073–1077.

12. Harzevili, A.R.S., Van Duffel, H., Defoort, T., Dhert, P., Sorgeloos, P., Swings, J., 1997. The influence of a selected bacterial strain *Vibrio anguillarum* TR 27 on the growth rate of rotifers in different culture conditions. *Aquaculture Internat.* 5, 183-188.
13. Hess, S., Sausen, N., Melkonian, M., 2012. Shedding light on vampires: the phylogeny of Vampyrellid amoebae revisited. *PLoS ONE* 7, e31165.
14. Hirayama, K., Ogawa, S., 1972. Fundamental studies on the physiology of the rotifer for its mass culture. I. Filter feeding of rotifer. *Bull. Jap. Soc. Sci. Fish* 38, 1207-1214.
15. Hoffman, Y., Aflalo, C., Zarka, A., Gutman, J., James, T.Y., Boussiba, S., 2008. Isolation and characterization of a novel chytrid species (phylum Blastocladiomycota), parasitic on the green alga *Haematococcus*. *Mycol. Res.* 112, 70-81.
16. Hotos, G.N., 2002. Selectivity of the rotifer *Brachionus plicatilis* fed mixtures of algal species with various cell volumes and cell densities. *Aquaculture Res.* 33, 949-957.
17. Imai, I., Ishida, Y., Sakaguchi, K., Hata, Y., 1995. Algicidal bacteria isolated from northern Hiroshima Bay, Japan. *Fish Sci.* 61, 628–636.
18. Ishino, R., Iethata, S., Nakano, M., Tanaka, R., Yoshimatsu, T., Maeda, H., 2012. Bacterial diversity associated with the rotifer *Brachionus plicatilis* sp. complex determined by culture-dependent and –independent methods. *Biocontrol Sci.* 17, 51-56.
19. Kankaala, P., Eloranta, P., 1987. Epizooic ciliates (*Vorticella* sp.) compete for food with their host *Daphnia longispina* in a small polyhumic lake. *Oecologia* 73, 203-206.
20. Kirchman, D.L., 2002. The ecology of *Cytophaga-Flavobacteria* in aquatic environments. *FEMS Microbiol. Ecol.* 39, 91-100.
21. Li, W., Godzik, A., 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22, 1658-9.

22. Li, N., Zhang, L., Li, F., Wang, Y., Zhu, Y., Hui, H., Wang, S., Qin, S., 2011. Metagenomics of microorganisms associated with the toxic Cyanobacteria *Microcystis aeruginosa* analyzed using the 454 sequencing platform. Chinese J. Oceanog. Limnol. 29, 505-513.
23. Magoc, T., Salzberg, S., 2011. FLASH: Fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27, 2957-63.
24. Mayali, A., Azam, F., 2004. Algicidal bacteria in the sea and their impact on algal blooms. J. Euk. Microbiol. 51, 139-44
25. McBride, R.C., Lopez, S., Meenach, C., Burnett, M., Lee, P.A., Nohilly, F., Behnke, C., 2014. Contamination management in low cost open algae ponds for biofuels production. Ind. Biotech. 10, 221-227
26. Nakase, G., Eguchi., 2007. Analysis of bacterial communities in *Nannochloropsis* sp. cultures used for larval fish production. Fisheries Sci. 73, 543-549
27. Page, F.C., 1983. Marine Gymnamoebae. Institute of Terrestrial Ecology, Cambridge, pp. 57.
28. Park, J.B.K., Craggs, R.J., Shilton, A.N., 2011. Wastewater treatment high rate algal ponds for biofuel production. Bioresour. Technol. 102, 35-42.
29. Raghukumar, S., 2002. Ecology of the marine protists, the Labyrinthulomycetes (Thraustochytrids and Labyrinthulids). Europ. J. Protistol. 38, 127-145.
30. Richardson, J.W., Johnson, M.D., Zhang, X.Z., Zemke, P., Chen, W., Hu, Q., 2014. A financial assessment of two alternative cultivation systems and their contributions to algae biofuel economic viability. Algal Research. 4, 96-104.
31. Strayer, D., Hummon, W. Hochberg, R. 2010. Gastrotricha. Pp. 163-172 in J Thorp, A Covich, eds. *Ecology and Classification of North American Freshwater Invertebrates*,

385 3rd Edition. London, England: Academic Press

- 386
- 387 32. Su, C.H., Chien, L.J., Gomes, J., Lin, Y.S., Yu, Y.K., Liou, J.S., Syu, R.J., 2010. Factors
- 388 affecting lipid accumulation by *Nannochloropsis oculata* in a two-stage cultivation
- 389 process. J. Appl. Phycol. DOI 10.1007/s10811-010-9609-4.
- 390 33. Suminto Hirayama K., 1996. Effects of bacterial coexistence on the growth of a marine
- 391 diatom *Chaetoceros gracilis*. Fish. Sci. 62, 40-43.
- 392 34. Tang, C.O., Leasi, F., Obertegger, U., Kieneke, A., Barraclough, T.G., Fontaneto, D.,
- 393 2012. The widely used small subunit 18S rDNA molecule greatly underestimates true
- 394 diversity in biodiversity surveys of the meiofauna. Proc. Natl. Acad. Sci. USA
- 395 doi/10.1073/pnas.1209160109
- 396 35. Thomas, V., Casson, N., Greub, G., 2006. *Criblamydia sequanensis*, a new intracellular
- 397 Chlamydiales isolated from Seine river water using amoebal co-culture. Env. Microbiol.
- 398 8, 2125-2135.
- 399 36. Yúfera, M., Navarro, N., 1995. Poulation growth dynamics of the rotifer *Brachionus*
- 400

plicatilis cultured in non-limiting food conditions. Hydrobiol. 313/314, 399-405.

Figure Captions

Figure 1: Cell density over time. Cell densities for a) cycle 7B and b) cycle 8B. Raceways for cycle 8B were treated with sodium hypochlorite 10 (2.5 ppm), 12 (5 ppm), and 14 (10 ppm) days after inoculation (denoted using *).

Figure 2. Eukaryotic community analysis of cycle 7B. A) Genus-level assignments for the eukaryotic hyper-variable region V4 in order of decreasing abundance for three raceways from cycle 7B on day 0 and 17 PI. B) Close-up of genus level assignments on day 0 PI, excluding *Nannochloropsis*. Genera comprising A) < 0.2% and B) 0.02% of at least one sample were grouped into ‘minor genera’; read pairs that were assigned a taxa in our edited version of Silva with unresolved classification were grouped in ‘classification uncertain’; when large fractions of read pairs were assigned to unknown organisms in full Silva or in Genbank they are listed with accession numbers. Read pairs unhit in any of the databases (not shown) consisted of less than 10% of pairs analyzed.

Figure 3. Bacterial community analysis of cycle 7B. Genus-level assignments for the bacterial hyper-variable region V6 gene for cycle 7B in order of decreasing abundance. Genera comprising < 0.7% of at least one sample were grouped into ‘minor genera’; read pairs not assigned to any taxa (*i.e.*, no hit) are indicated by ‘unhit in databases’; when large fractions of read pairs were manually assigned to unknown organisms in full Silva or in Genbank they are listed with accession numbers.

Figure 4. Eukaryotic community analysis of cycle 8B. A) Genus-level assignments for the eukaryotic hyper-variable region V4 in order of decreasing abundance. B) Close-up of all samples, excluding taxa representing > 10% of reads. Genera comprising A) < 0.01% and B) < 0.03% of at least one sample were grouped into ‘minor genera’ for both plots; when large fractions of read pairs were assigned to unknown organisms in full Silva or in Genebank they are listed with accession numbers. Read pairs unhit in any of the databases (not shown) consisted of less than 10% of pairs analyzed.

Figure 5. Bacterial community analysis of cycle 8B. Genus-level assignments for the bacterial hyper-variable region V6 in order of decreasing abundance. Genera comprising < 0.4% of at least one sample were grouped into ‘minor genera’; read pairs not assigned to any taxa (*i.e.*, no hit) are indicated by ‘unhit in databases’; when large fractions of read pairs were assigned to unknown organisms in full Silva or in Genebank they are listed with accession numbers.

TABLES

Table 1. Nutrient formulations used in relevant algae raceways (ARWs), Cycles 7B and 8B.

Nutrient	Chemical Formula	Units	Nutrient L ⁻¹ of Culture Media	
			7B	8B
			(ARW #s 6, 8 & 11)	(ARW #s 3, 7 & 12)
Ammonium Sulfate	(NH ₄) ₂ SO ₄	g	0.2640	0.1320
Phosphoric Acid (85%)	H ₃ PO ₄	ml	0.0078	0.0240
Ferrous Sulfate	FeSO ₄ ·7H ₂ O	g	0.0200	0.0100

446 Table 2. Physical parameters averaged across raceways and sample days with standard
 447 deviations during Cycles 7B and 8B.

Cycle	Date		Temperature (°C)		Salinity (ppt)		Dissolved Oxygen (mg/L)		pH	
	Start	End	AM	PM	AM	PM	AM	PM	AM	PM
7B	07/16/11	08/13/11	26.44	33.12	36.45	35.67	6.71	10.69	6.61	6.27
			±	±	±	±	±	±	±	±
			0.47	1.24	0.87	0.63	1.21	2.78	0.66	0.66
8B	09/17/11	10/04/11	23.55	29.71	35.07	34.49	7.12	9.91	7.16	6.85
			±	±	±	±	±	±	±	±
			2.24	2.16	0.57	0.52	1.21	2.16	0.58	0.84

448

449

450 Table 3. Read pair counts (in million) for sequences generated for eukaryotic V4 SSU rRNA and
 451 bacterial V6 SSU rRNA.

	All Samples		Cycle 7B		Cycle 8B	
Read pairs:	V4	V6	V4	V6	V4	V6
Generated	178.2	265.8	5.9	14.9	3.7	13.8
Analyzed with pipeline	164.1	147.1	5.1	8.0	3.0	7.6
Hit in full Silva	151.7	107.5	4.2	5.3	2.4	4.7
Hit to a genus	151.1	25.5	4.1	2.6	2.4	1.2
Analyzed per sample						
(median)	1.3	1.4	0.9	1.4	0.6	1.4

452

453

455

456 Figure 1: Cell densities for a) cycle 7B and b) cycle 8B. Raceways for cycle 8B were treated

457 with sodium hypochlorite 10 (2.5 ppm), 12 (5 ppm), and 14 (10 ppm) days after inoculation

458 (denoted using *).

Figure 2. Eukaryotic community analysis of cycle 7B. A) Genus-level assignments for the eukaryotic hyper-variable region V4 in order of decreasing abundance for three raceways from cycle 7B on day 0 and 17 PI. B) Close-up of genus level assignments on day 0 PI, excluding *Nannochloropsis*. Genera comprising A) < 0.2% and B) 0.02% of at least one sample were grouped into ‘minor genera’; read pairs that were assigned a taxa in our edited version of Silva with unresolved classification were grouped in ‘classification uncertain’; when large fractions of read pairs were assigned to unknown organisms in full Silva or in Genebank they are listed with accession numbers. Read pairs unhit in any of the databases (not shown) consisted of less than 10% of pairs analyzed.

Figure 3. Bacterial community analysis of cycle 7B. Genus-level assignments for the bacterial hyper-variable region V6 gene for cycle 7B in order of decreasing abundance. Genera comprising < 0.7% of at least one sample were grouped into ‘minor genera’; read pairs not assigned to any taxa (*i.e.*, no hit) are indicated by ‘unhit in databases’; when large fractions of read pairs were manually assigned to unknown organisms in full Silva or in Genebank they are listed with accession numbers.

Figure 4. Eukaryotic community analysis of cycle 8B. A) Genus-level assignments for the eukaryotic hyper-variable region V4 in order of decreasing abundance. B) Close-up of all samples, excluding taxa representing > 10% of reads. Genera comprising A) < 0.01% and B) < 0.03% of at least one sample were grouped into ‘minor genera’ for both plots; when large fractions of read pairs were assigned to unknown organisms in full Silva or in Genbank they are listed with accession numbers. Read pairs unhit in any of the databases (not shown) consisted of less than 10% of pairs analyzed.

Figure 5. Bacterial community analysis of cycle 8B. Genus-level assignments for the bacterial hyper-variable region V6 in order of decreasing abundance. Genera comprising < 0.4% of at least one sample were grouped into 'minor genera'; read pairs not assigned to any taxa (*i.e.*, no hit) are indicated by 'unhit in databases'; when large fractions of read pairs were assigned to unknown organisms in full Silva or in Genebank they are listed with accession numbers.