7. Molecular diagnostic solutions in algal cultivation systems

Laura T. Carney[1], Robert C. McBride[2], Val H. Smith[3] and Todd W. Lane[4]

[1]Heliae Development, LLC, 578 E. Germann Road Gilbert, AZ 85297, USA

[2]Sapphire Energy, 3115 Merryfield Row, San Diego, CA, 92121, USA

[3]Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, KS, 66045, USA

[4]Systems Biology Department, Sandia National Laboratories, PO Box 969, MS 9292, Livermore, CA 94551-0969, USA

Abstract:
Contamination of algal mass culture by predators, pathogens and competitors represent one of the major challenges to commercial production of algae. Early detection of, and response to, these deleterious organisms dictates how much of an effect they will have on the algae biomass. Thus, the incorporation of rapid detection tools into an algae production routine is crucial. Molecular methods of identification and detection represent alternative, high throughput tools that can supplement, and in some cases replace, optical methods for the identification and detection of deleterious species in algal mass culture systems. Many of these molecular methods are quite new and their use has been increasingly demonstrated across disciplines in recent years. Here, we review these techniques, including Next Generation Sequencing, qPCR, microarrays, and Fluorescent *in situ* hybridization (FISH), and comment on their promise and limitations. Finally, we discuss how some of these modern methods have been used to detect biocontaminants in algae production systems and the various responses these types of data can direct.

7.1. Introduction

One of the major challenges to achieving high rates of long-term production in microalgal mass cultures is the elimination or reduction of the impact of contamination and crashes in production systems. Although there are both biotic and abiotic root causes of mass culture crashes, infection by deleterious species may be one of the least understood. Aside from a few well-known culprits, the diversity of pathogens, parasites, predators and competing algal species (or 'weed' species) has not been well characterized. Lost production days due to pond crashes can significantly lower annualized production rates. Depending on the scale and type of system, days to weeks of production can be lost while the system is disinfected and new inoculum and growth medium is prepared. Depending on the design and operation of the production facility, there is a risk of spread or persistence of contamination and successive crashes. Despite a paucity of publically available data on the economic impact of biocontaminants on the nascent algae biomass industry, the consensus is that they constitute an economic barrier to commercialization (ANL, NREL, PNNL 2012, Gao et al. 2012). Some insight into the potential magnitude of the financial impact may be gained from the aquaculture-for-food industry which loses several billion US$ annually

(Subasinghe et al. 2001, FAO 2010) due to bacterial and fungal infections (Defoirdt et al. 2004, Ding and Ma 2005, Ramaiah 2006).

There are a wide variety of deleterious species including bacteria (Cole 1982, Fukami et al. 1997), viruses (Dunigan et al. 2006), parasites (e.g., Letcher et al. 2013), fungi (Fott 1967, Hoffman et al. 2008, Li et al. 2010) and herbivorous zooplankton (e.g., Park et al. 2011). In addition there are algal "weed species" (Pienkos and Darzins 2009, Bull and Collins 2012) that can reduce the value of the biomass by supplanting the desired species. Early detection of deleterious species is the key to informed pond management strategies and thus successful large-scale algae cultivation will require routine, detailed, fast and cost-effective identification of potentially deleterious species that have become established in the culture and how their populations may change with time and conditions.

7.2. Diagnostic methods

Identification of algal pests is still largely dependent on microscopy (e.g., Rasconi et al. 2009). Although optical methods can produce near real time data on the presence of contaminants and cannot be avoided when describing novel biocontaminants, microscopy is a method that is labor intensive, low throughput and requires a certain level of expert knowledge to recognize contaminant species. Thus a routine monitoring program relying solely on traditional microscopy is not desirable. As a consequence of these limitations a variety of advanced optical methods have recently been developed such as flow cytometry coupled with digital imaging and image recognition (reviewed by Álvarez et al. 2011, Day et al. 2012) and hyperspectral confocal florescence microscopy (Collins et al. 2014). While such methods represent potential improvement over standard optical assessment they, as is the case with many molecular techniques, are still dependent on *a priori* knowledge of deleterious species.

Molecular diagnosis, in general, can be divided into two separate types of processes. The first is the initial identification of novel etiological agents: viruses, grazers, pathogens and parasites. Once identified and characterized, this information is used to develop reliable detection assays. Initial identification of novel agents is still, fundamentally, based on the postulates first formulated by Robert Koch in the late nineteenth century. Classical methods of fulfilling Koch's postulates have depended upon the isolation and culturing of the etiological agent and the demonstration that the disease-state is unambiguously correlated with the presence of the pathogen via re-isolation. These methods remain the "gold standard" for pathogen identification; however, in the last two decades culture-free molecular methods of identification and detection have become more commonplace in clinical diagnostics. There have been a number of instances in which the etiological agent was not culturable and was only recognized by its molecular signature (for review see Frederick and Relman 1996, Miller et al. 2013). Although, in most cases, identifications that are based solely on molecular signatures are considered to be presumptive unless confirmed by other evidence, the rise of molecular diagnostics has led to a rethinking of Koch's postulates to reflect the use of these methods (Fredericks and Relman 1996). It is now reasonable to reformulate Koch's postulates for molecular diagnosis of algal production system infections:

1. A nucleic acid sequence belonging to the putative pathogen, predator, or parasite should

be present in an infected algal mass culture.

2. In healthy mass cultures, copies of pathogen-associated nucleic acid sequences should occur below a management threshold value specified for that pathogen/pest.

3. When sequence detection predates mass culture infection, increase in sequence copy number should correlate with loss of algal biomass, decrease in productivity or aberrant pond performance.

4. The organism, identified by sequence-based analysis, should have properties that are consistent with, and capable of, generating the phenomena observed in the infected pond.

Several factors can confound the identification of pest species by molecular analysis of algal mass culture infections. 1) Other sequences may become prevalent in crashed or infected ponds including those related to species that feed upon detritus or dead algal cells. In these cases, temporal relationships are critical to establish and thus the frequency of pond sampling is an important consideration. A reasonable strategy is to collect and archive samples on frequent and routine basis to enable the analysis of a time-course of infection should a deleterious event occur. 2) Environmental conditions and abiotic stressors may play a significant role in mediating pond crashes and can be in themselves the root cause of pond crashes. 3) Poly microbial infections are possible although many crashes appear to be caused by a single agent (Hu pers. comm.). 4) The correlation of an increase of a biocontaminant and a decrease of the target strain does not confirm the biocontaminant as the causative agent.

Broadly speaking, there are three forms of molecular techniques that are employed for diagnostics; probe hybridization, target amplification (PCR), and DNA sequencing. Each of these techniques, which are described in greater detail below, take advantage of information contained in specific, well-characterized genetic regions for identification and/or detection of the target organisms. The information contained in these regions can be used to identify novel agents and to develop taxa specific hybridization probes or PCR primers for detection of known agents.

7.2.1 *Molecular targets for diagnostic methods*

There are several molecular markers that have been employed for molecular analysis of clinical and environmental samples (Hoef-Emden 2012) including the large (LSU rRNA) and small subunit ribosomal RNA (SSU rRNA) genes, the internal transcribed spacer (ITS) of the ribosomal RNA genes and the mitochondrial cytochrome oxidase (cox) gene. The best known of these targets is the SSU rRNA gene (see Figure 1). The prokaryotic SSU rRNA gene has nine hypervariable regions dispersed along its length (Van De Peer et al. 1996). Of these, the sequence information contained in the individual hypervariable regions three or six (Huse et al. 2008) or a fragment covering hypervariable regions 1 through 3 or 4 (Kim et al. 2011) are the most useful for phylogenetic determination. Sequence data from individual hypervariable regions is often sufficient for genus level distinction whereas data from multiple hypervariable regions or the entire SSU rRNA gene can result in species level distinction.

The eukarytotic SSU rRNA gene also contains nine variable regions of which V4 and V9 are the most informative and are often used in combination (Amaral-Zettler et al. 2009, Pawlowski et al. 2011, Stoeck et al. 2010, Orsi et al. 2013). The V4 region is the longest with the highest degree of length variation and sequence heterogeneity (Nickrent and Sargent 1991). The information contained in either region is generally sufficient for the genus level identification of an organism. The shorter V9 region is less useful partially because it is sometimes not present in truncated versions of eukaryotic SSU rRNA genes found in sequence databases. Although phylogenetic analysis based on SSU rRNA gene sequence is more common, similar analysis based on the LSU rRNA gene is also carried out (Ludwig and Schleifer 1994, Hunt et al. 2006, Steven et al. 2012). The LSU rRNA gene is longer and contains two variable regions, D1 and D2 (see Figure 2), that can be used for phylogenetic analysis (Sonnenberg et al. 2007, Putignani 2008)

The ITS region, located between the rRNA genes encoding the SSU and the LSU (see Figure 3), is commonly used for genus or species level discrimination in fungal (Lindahl, et al. 2013) and microalgal (Leliaert et al. 2014) phylogenetics. There is a very large database of fungal ITS sequences (see Section 7.2.7). In eukaryotes the ITS region consists of two hypervariable spacers, ITS1and ITS2, that flank the gene encoding the 5.8s ribosomal subunit. A full length amplicon including both ITS1 and ITS2 regions and the 5.8s subunit is approximately 650 bp in length which is beyond the current read length limits of many of the next generation sequencers (see Table 1). Because of this inability to cover the entire ITS region, individual ITS regions have been analyzed by next generation sequencing (Lindner et al. 2013). Reports indicate that community analyses based on ITS1 versus ITS2 yield different taxonomic compositions from each other as well as from those based on the full length ITS region (Blaalid et al. 2013, Bazzicalupo et al. 2013). In the complex communities that may be present in algal mass culture systems, the potential for incorrect identification or detection can be mitigated by the use of multiple assays targeting different regions.

The 5' terminus of the mitochondrial cytochrome oxidase gene *cox1* has been used as a molecular barcode region primarily for metazoans (Bucklin et al. 2011). There are caveats to the use of *cox1*: the amplification of nuclear encoded pseudogenes can lead to overestimation of species diversity (Song et al. 2008) and it has proven difficult to develop universal primers (Saunders and McDevit 2012).

7.2.2. *Sample preparation for molecular diagnostics.*

Most molecular diagnostic methods such as sequencing, hybridization or PCR, require cell lysis and extraction of nucleic acids from the biomass sample. Differential or failure of extraction of nucleic acids from one group of organisms can skew apparent relative abundances. A variety of protocols and commercial kits for the extraction of nucleic acids from recalcitrant organisms have been developed and there are some comparisons of these methods reported in the literature (Purdy 2005, Koid et al. 2012). However, new methods and commercial kits are routinely developed and it is advisable to test the effectiveness of a variety of protocols on the sample types of interest using the intended molecular diagnostic method (qPCR, sequencing, etc.). Many lysis and extraction protocols feature a combination of both chemical and mechanical lysis methods. There are several forms of mechanical lysis including sonication, French pressure cell,

nitrogen bomb, cyropulverization and bead beater. For field applications bead beating is arguably the most convenient form of mechanical lysis. This is largely due to the potential low cost of instrumentation and ability to handle small sample volumes. For particularly recalcitrant samples, an alternative method is cryopulverization in which the sample is first flash frozen in liquid nitrogen then pulverized with a cold mortar and pestle. There are a variety of commercial devices that automate this process. Because of the potential for a high degree of complexity in samples and, therefore, the likelihood that could contain difficult to lyse species, cryopulverization may be a particularly attractive option. However more work is needed to mitigate the bias involved in nucleic acid extraction during microbiome analysis in order to transform these efforts from a relative to a quantitative analyses.

7.2.3. *Identification of contaminants by DNA sequencing*

Next Generation Sequencing (NGS) has been employed extensively to identify unknown pathogens in a diversity of situations ranging from humans, to vineyards, to beehives. NGS based approaches have recently been applied to the characterization of the predator, pathogen and parasite loads of underperforming and crashed pilot scale ponds and photobioreactors (Carney et al. unpublished data, Carney et al. 2014). The sequencing strategies employed in such analyses are dependent on the read length and number which are in turn dependent on the technical specifications of the sequencing system employed. There are a number of sequencing systems in use today and the most common are summarized in Table 1.

Table 1. Current next generation sequencer capabilities

Sequencer	Read lengths	Read number	Run time
Illumina Hi Seq	300 nt	6 $x10^9$	2-11 d
Illumina MiSeq	600 nt	1.2-1.5 $x10^7$	39 h
454 GS	400 nt	7 $x10^4$	10 h
GS FLX	700 nt	1 $x10^6$	23 h
Ion Torrent PGM	Up to 400 nt	Up to 6 $x10^6$	2-7 h
Ion Torrent PI	Up to 200 nt	Up to 3.3 $x10^8$	2-4 h
Pacific Biosciences	4300-5000 nt	2.5-5.0 $x10^3$	~2 h
Applied biosystems Solid 5500	50-100 nt	1 $x10^8$	1-7 d

Generally sequencers fall into two broad categories, high volume machines that are found in core facilities and commercial providers and less expensive, lower volume "personal" sequencers. Personal sequencers including the Illumina MiSeq, 454 GS, and the Ion Torrent PGM have the advantage of relatively rapid run times and lower cost per run than the larger machines. Pacific Biosciences sequencers are targeted for specific applications requiring a relatively small number of long reads. The costs associated with the actual sequencing can be reduced by the creation of multiplexed sequencing libraries allowing for the sequencing of more than one sample per lane. Various systems that take advantage of each type of sequencer have been developed for the creation of multiplexed libraries (Mckenna et al. 2008, Caporaso et al. 2011, Whiteley et al. 2012).

Next generation sequencing strategies have evolved as the read lengths have progressively improved. At the time of writing it is possible to obtain reads that span regions of up to 600 nt (Illumina MiSeq, using paired end kits). Taking prokaryotic SSU rRNA gene analysis as an example, the MiSeq is now capable of completely sequencing amplicons covering hypervariable regions 1-3. An alternative strategy for obtaining full-length rRNA sequence coverage, with next generation sequencers, entails amplifying the full length gene then shotgun sequencing and assembling the amplicon. This strategy is more complex because it requires an assembly step that is not necessary for the analysis of shorter amplicons. The advantage is that more information is available upon which to base phylogenetic assignments.

The handling of the raw sequence data from the sequencer is device-dependent and instruments generally come with data process and analysis software packages. Raw sequencing reads must pass through a series of quality control steps to remove low quality and primer sequences and to mask low complexity regions. Once this initial processing is complete there are a number of software packages for microbiome analysis including QIIME [http://qiime.org/index.html]. (Caporaso et al. 2010), Mothur (Schloss et al. 2009), and the RDP pyrosequencing pipeline designed for the analysis of 454 sequencing datasets.

In samples that contain high algae biomass densities, prokaryotic rRNA gene sequencing libraries may be dominated by amplicons derived from the algal chloroplast. In these cases, primers devised to exclude amplification of plastid SSU rRNA sequences can be used in single step or in nested PCR amplifications (Rastogi et al. 2010, Chelius and Triplett 2001). It should be noted that these primers also exclude the amplification of cyanobacterial sequences. Other strategies for reducing the burden of non-informative sequences include the use of so-called blocking primers that bind to and prevent amplification of unwanted sequences. Such primers contain modified nucleotides that prevent primer extension (Vestheim and Jarman 2008). In theory blocking primer strategies may be more generally applicable and could be applied to additional amplification targets aside from rRNA genes.

The advantages of next generation sequencing based methods of identification are that they are culture independent and do not rely on isolation of the organism. There are three major disadvantages to sequencing based systems for routine pond surveillance. The first is the cost of the equipment and reagents. The second is the technical sophistication required to carry out the library preparation and the data analysis. Finally, the time required to go from sample to answer is too long for routine surveillance. There are a number of commercial entities that provide contract library preparation and sequencing services mitigating some drawbacks. Because sequencing library preparation protocols require amplification, next generation sequencing suffers from the same potential sources of bias as PCR based detection methods. In addition, each next generation sequencing system displays a different level of bias against templates with at high and low GC ratios (Quail et al. 2012).

In cases where a deleterious organism or weed species is of sufficient abundance in the contaminated culture or has been isolated or enriched in culture, it is possible to clone and carry out dideoxynucleotide terminator sequencing (Sanger et al. 1977) of the desired region. With Sanger sequencing it is generally possible to achieve 700 nt of sequence data per primer

extension reaction. Thus, paired end reactions should be sufficient to sequence both DNA strands of the ITS region and single strands of most, if not all, of the full-length SSU rRNA genes. By providing full sequencing coverage of the region of interest, Sanger sequencing can result in a high degree of taxonomic distinction of the target organism and maximal information for the design of PCR primer or oligonucleotide probes for the future detection and quantification of the deleterious species in algal mass culture.

7.2.4. *Detection of contaminants by PCR based methods*

Quantitative PCR (qPCR) is one of the more common molecular diagnostic techniques for the detection of known deleterious species (for review see Botes et al. 2013). qPCR can be an essential tool for conducting reactive management strategies (see Chapter 6.3.2, McBride et al. 2014) where a biocontaminant is isolated, identified and desired to be routinely monitored (Figure 4). Two different reporter systems for qPCR are in general use; fluorogenic dyes or fluorescent oligonucleotide probes. The first system (Ponchel et al. 2003) utilizes dyes, such as Sybr Green, that fluoresce when bound to the double-stranded product of PCR reactions allowing quantification of the product. The limitation of this system is that the dye binds non-specifically to any double-stranded DNA including primer dimers. The major advantage of the fluorgenic dye based system is that it is less expensive than the probe based system, both to design and optimize the probes and to run the reactions. Probe based systems (often referred to as Taqman) utilize a fluorescently labeled oligonucleotide that binds to the desired target product (Holland et al. 1991). This allows for greater specificity and enables quantification of the target even in the presence of non-target amplicons. In addition, probe based systems can be multiplexed for detection of multiple species. The disadvantage of such systems is that the reactions are more expensive because of the requirement for the labeled oligonucleotide probe, and are technically more challenging, as optimizing multiplexed probes is substantively more complex and time consuming than single primers, and more sensitive to interference from unpredictable environmental samples.

qPCR has several advantages, the most important of which is the potential for high sensitivity. Under the appropriate conditions, qPCR can approach single cell detection levels. In practice this level of sensitivity can be difficult to obtain and care must be taken in the preparation of samples to avoid cross-contamination. qPCR reactions can be multiplexed for the detection of multiple species within a single reaction and given the appropriate equipment, thousands of such reactions could be run on a daily basis. In addition to bias introduced in the sample preparation, there are two major sources of PCR bias, the choice of primers for amplification and the reaction conditions used for amplification. Specific primers for detection are designed for either universal or taxa specific amplification. It has, however, proven difficult to create primer sets that amplify all targeted sequences with equal efficiency. This bias can be characterized and potentially limited by choice and testing of a variety of primers against a diversity of targets species or near neighbors. It may be desirable to compare libraries created with alternative primer sets targeting either the same region or different molecular barcodes. PCR amplification bias at high and low GC ratios is also a potentially significant source of error introduced during the generation of sequencing libraries (Aird et al. 2011). This can be controlled or eliminated by adjustment of the PCR cycle parameters and choice of polymerase. An additional limitation to qPCR as a detection technology includes the initial costs of the equipment.

Fulbright et al. (2014) have demonstrated the application of Cleaved Amplified Polymorphic Sequences (CAPS), to the detection of weed species in algal mass culture systems. CAPS is based on restriction fragment length polymorphisms in PCR amplified products. In this procedure a target region, such as the SSU rRNA gene, is amplified from DNA extracted from the mass culture (test sample). The product of the PCR reaction is then cleaved with one or more restriction enzymes and the lengths of the resulting fragments are determined by electrophoretic analysis and compared to those derived from the digestion of the same target amplified from the desired algal strain (control sample). The presence of restriction fragments in the test sample that are not present in the control is indicative of biocontamination. The major advantages of this method are that it can detect contamination without the need for specific probes and that it does not require the more expensive equipment and reagents needed for qPCR. The disadvantages are that the method does not actually identify the contaminant species and it is not as sensitive as qPCR.

7.2.5. *Detection of contaminants by microarray based methods*

Microarrays generally consist of glass slides with oligonucleotide probes arrayed in spots on the surface. Each spot contains a specific probe that is designed to hybridize to a particular target sequence. Samples are prepared by coupling a fluorescent probe to extracted nucleic acids that are then applied to the array. DNA fragments, that are complementary to probes on the array will bind to that specific spot and then are detected by fluorescence. Microarray based systems have been developed using arrayed oligonucleotide probes targeted to the SSU rRNA. These so-called phylochips were originally developed for prokaryotic community structure analysis and there is an extensive literature on their use (Loy et al. 2002). Custom phylopchips can be developed for the detection of organisms of interest and have been adapted to eukaryotic SSU rRNA analysis and applied to the detection of marine eukaryotes, including microalgae and pathogenic protozoa (Metfies et al. 2007). Probe design is a particular challenge in the development of microarrays. All probes on a particular microarray must behave in the same manner under a given set of hybridization conditions. There are a number of databases and tools to assist in the design of probes (see section 7.2.7). The microarray technique, by itself, is not as sensitive as qPCR methods but this can be improved by amplification of the target by PCR or by multiple displacement amplification (Binga et al. 2008). A potential drawback to Phylochip based pond diagnostics is the expense of microarray development and production, particularly in situations when there are several species to be monitored. Macroarray systems, based on dot blot technology can be utilized for rapid optimization and validation of probes or for applications where only a modest number of species are to be detected.

7.2.6. *Detection of contaminants by fluorescent in situ hybridization methods*

Fluorescent *in situ* hybridization (FISH) has been applied to the identification and enumeration of morphologically indistinguishable species such as eukaryotic picoplankton (Not et al. 2002), Bacteria (Amann et a. 1991) and Archaea (DeLong et al 1999). The method entails hybridization of permeabilized samples with fluorescently labeled olignucleotide probes followed by fluorescent imaging or flow cytometry. The analysis can be multiplexed through the utilization of mixtures of probes with differently colored fluorescent tags. Signal amplification

methods such as catalyzed reporter deposition (CARD), also referred to as tyramide signal amplification (TSA), have been developed (LeBaron et al 1997). In CARD-FISH, horseradish peroxidase (HRP) is conjugated to the probe. In the presence of hydrogen peroxide, HRP converts tyramide (conjugated to a fluorophore) to a reactive intermediate that interacts with aromatic compounds in the cell (chiefly tyrosine and tryptophan) forming an insoluble complex that is deposited around the probe, greatly enhancing the signal. This method has been used in conjunction with rRNA targeted probes to eukaryotic algae (Simon et al. 1995), zoopsporic fungi (Jobard et al. 2010) and marine bacteria (Pernthaler et al 2002). The major strengths of FISH techniques are the single cell resolution of the method and the flexibility to utilize either fluorescence microscopy or flow cytometry as the output device. FISH is technically sophisticated method that may limit its application to small scale or research operations, or for the production of high value products. It does not have the same potential for high throughput, multiplexed analysis as other methods such as PCR and, thus, seems unlikely to be applied to routine mass culture diagnostics or surveillance.

7.2.7. *Bioinformatics for molecular diagnosis*

There are a number of sequence data repositories for rRNA and mitochondrial cytochrome oxidase genes. The SILVA database, [http://www.arb-silva.de/], is a curated repository of SSU and LSU rRNA gene sequences from Archaea, Bacteria, and Eukarya (Quast et al 2013). The greengenes database contains LSU rRNA sequences from Bacteria and Archaea [http://greengenes.secondgenome.com/downloads] (DeSantis et al. 2006). Both databases can be downloaded from their websites in a flat file format for use with the various microbiome analysis software packages. The RDP Ribosomal Database project, [http://rdp.cme.msu.edu/], (Cole et al. 2009) maintains a database of bacterial and archaeal SSU sequences. UNITE (Abarenkov et al. 2010) is a searchable database of primarily fungal ITS regions which can be downloaded in flat file format from its home website, [http://unite.ut.ee/], or from the QIIME website. As the name suggests, ITS2 database (Koetschan et al. 2012) is a searchable database of specifically ITS2 sequences, [http://its2.bioapps.biozentrum.uni-wuerzburg.de/] and ITSoneDB, [http://itsonedb.ba.itb.cnr.it/] focuses on the ITS1 region. The BOLD database (Ratnasingham and Hebert 2007), [http://www.barcodinglife.org], contains both sequence data and a registry of primers for *cox1*. Genbank is the NIH depository of publically available annotated sequences, (Benson et al. 2012) [http://www.ncbi.nlm.nih.gov/genbank/] and can be queried to identify sequences that are not found in other databases.

Once the deleterious agent or weed species of concern has been identified and sufficient sequence information is available, bioinformatics analysis is employed in the development of oligonucleotide primers and probes for the detection and quantification of the agent in algal mass culture. ProbeBase [http://www.microbial-ecology.net/probebase/] is an online database of probes and primers targeting rRNA (Loy et al. 2007). The rRNA gene databases SILVA, Greengenes, and RDP host web based tools to assist in the development of probes by checking them for homology to sequences in their respective databases. The web tool probeCheck, [http://131.130.66.200/cgi-bin/probecheck/content.pl?id=home] (Loy et al. 2008) provides a single interface for the comparison of rRNA subunit targeted probe sequences against multiple databases.

7.2.8. *Detection using remote sensing*

There is much work to be done in the detection and recognition of pest challenges, and the most exciting advances will be those technologies which can be implemented in a low cost manner and which provide simultaneous information on multiple problems. Coupled with recent advances in computing and geopositioning technologies, remote sensing data obtained from ground-, air-, and space-based platforms are now capable of providing detailed spatial and temporal information on plant responses to their local environment that is needed for site-specific precision agriculture (Pinter et al. 2003). We suggest here that reflectance monitoring of the open pond surface holds similar promise for the development of rapid strategies for the early detection of grazing- and disease-induced stress in open algal cultures (Reichardt et al. 2014). Multispectral image analysis can be used to measure algal biomass concentrations, detect invasive species, and monitor culture health in real time (Murphy et al. 2014). As algal biomass production moves to commercial scales, we suggest that the use of remote sensing tools will be essential components of real-time crop management systems. In the future, these optical tools could potentially be deployed using small, mobile devices that could quickly sample the entire cultivation system landscape. Following the initial detection of a potential crop growth issue, these devices could immediately be returned to perceived trouble spots for problem confirmation. Mobile devices also could be used for monitoring the behavior of algae in localized subsections of the cultivation facility in which targeted crop protection measures have been implemented.

7.3. Case studies of problematic contaminants

Molecular techniques are very useful in characterizing the community within an algae cultivation system. These techniques are used to not only verify the intended strain but also to detect biocontaminants before they become an irreversible problem. In cases where a natural community of algal strains is allowed to colonize a cultivation system, molecular tools can also be quite valuable. Below we describe some potentially deleterious biocontaminants and give examples where molecular techniques were used to detect them.

7.3.1 *Mesograzers*

Growing algae to scale will undoubtedly lead to some infestations by algivorous predators. Metazoan rotifers and herbivorous ciliates are perhaps the most commonly reported biocontaminants in large scale cultivation systems. For example, once established, rotifers cause extremely rapid biomass loss as they are able to double their density in 24 hrs (Yúfera and Navarro 1995, Sarma et al. 2001) while ingesting 200 algal cells min^{-1} $rotifer^{-1}$ (Hirayama and Ogawa 1972). Rotifers and some ciliates also form thick-walled resting stages (i.e., cysts) that are resistant to desiccation and chemical treatment such as chlorine. However, other lesser known grazers are also problematic, including amoebae, heliozoans, and vampyrellids, and further characterization of large scale cultures will be needed to further investigate the specific threat they pose to commercial algae productivity.

A grazing ciliate has been frequently observed in open raceways in Arizona. A combined strategy of microscopy, isolation, culture and sequencing identified this organism as *Gastrostyla*

steinii. (Figure 5a), informing appropriate measures to reduce its impact (Carney, unpublished data). Replicate raceway ponds in Texas growing *Nannochloropsis salina* were reported to simultaneously crash, presumably from different organisms. Microbiome analysis using NGS confirmed that one crash was due to the outbreak of the rotifer, *Brachionus plicatilis*, while the other crashed raceway had become dominated by a mix of the grazing gastrotrich, *Cheatonotus*, and the Cytophagia bacterium, *Aureibacter* (Carney et al. In prep.). It is interesting that simultaneous crashes would occur in adjacent ponds but due to very different types of biocontaminants.

7.3.2. *Fungal parasites*

Fungal parasites are particularly insidious in algal culture (reviewed by Carney and Lane 2014). While acute fungal infections cause obvious rapid pond crashes, chronic infections cause reduced productivity over longer time periods (Carney et al. 2014). In natural freshwater environments chytrids are well known to parasitize microalgae at infection rates that exceed 90%, causing severe biomass losses (Kagami et al. 2007). Chytrid infections of green algae have been described within cultivation systems as well (both open and closed) and can severely and rapidly reduce algae productivity (Fott 1967, Hoffman et al. 2008). For example, only 3 days may be required for a chytrid parasite to infect 100% of algae cells within a given *Haematococcus* culture (Hoffman et al. 2008).

In a prototype enclosed photobioreactor at a wastewater treatment plant in California, the community was analyzed during a potential crash of a freshwater green alga growing in wastewater effluent that had been chlorinated and dechlorinated (Carney et al. 2014). Microbiome analysis revealed a small spike in the proportion of *Rhizophydium*, a known algal parasite that correlated with a decrease in algae biomass. After a few days, this chytrid was replaced by another chytrid that was saprobic (i.e., feeding on dead algae biomass). This change in chytrid community was consistent with the end of an infection period when dead algal biomass would accumulate.

Recently, a new member of Cryptomycota, *Amoeboaphelidium protococcarum*, was discovered infecting the green alga, *Scenedesmus*, in commercial ponds in New Mexico (Letcher et al. 2013). The abundance of this aphelid was inversely correlated with that of the algae, a potential disaster considering the parasite was previously unknown. It was isolated and cultured and finally a multi-gene analysis was used to place the parasite within the Cryptomycota. Sister taxa included other parasites that had infected algae ponds previously.

Two aquatic fungi parasitized *Haematococcus* growing in open raceways in Arizona, isolate JEL821 identified as *Paraphysoderma* (Blastocladiomycota; Fig. 5b, c & h, sequence identification pending) and isolate JEL812 (Figure 5f & i) identified as belonging in the family Aquamycetaceae in the Rhizophydiales (Chytridiomycota). *Paraphysoderma sedebokerensis* was described by Hoffman et al. (2008) and Gutman et al. (2009) as infecting *Haematococcus pluvialis*. The isolate JEL812 may be cosmopolitan and has been observed infecting mass algal cultures in the past (JEL317; Joyce Longcore, pers. comm.). Both fungi were isolated and examined via microscopy for identifying characteristics. qPCR assays were designed based on

16s SSU rRNA sequence data, which are used routinely to monitor raceways to inform crop protection strategies.

7.3.3. *Pathogenic bacteria*

The role of bacteria in algae dominated communities is not straightforward. Bacteria are reported as both beneficial (Cole 1982, Kazamia et al. 2012) and pathogenic (Cole 1982, Fukami et al. 1997). Deciphering which strains, or which strains under which conditions, are pathogenic will be important for algae cultivation management. *Bdellovibrio*-like bacteria have been reported to lyse the green algae, *Chlorella* (Cole 1982, Coder and Starr 1978). We detected this genus in an enclosed photobioreactor growing freshwater green algae in California (Carney et al. 2014). Genera from the group *Cytophagia* may also be able to infect algae and are known to attach to cyanobacteria, secrete lytic substances to then dissolve the cell wall, causing lysis (reviewed by Rashidan and Bird 2001). These strains are quite common during natural cyanobacterial bloom formation (Rashidan and Bird 2001). We have repeatedly detected cytophage sequences in declining green algae cultures (Carney et al. 2014, Carney et al. In prep.) and it is possible that they have lytic effects on more than just cyanobacteria. The pathogenic bacterium, *Vampirovibrio chlorellavorus*, was detected by metagenomic analysis in open raceways in Arizona and correlated with high levels of bacterial attachment to algal cells and lysis observed using microscopy (Figure 5d, e & g). Because *V. chlorellavorus* is an obligate parasite (Coder and Goff 1986), isolation and culture techniques proved difficult. Once sequence data was obtained, a qPCR assay was designed and used to monitor the presence of *V. chlorellavorus* and validate crop protection strategies to avoid future crashes (Carney, unpublished data).

7.3.4. *Algal viruses*

Algal viruses may be the greatest source of uncharacterized genetic diversity in the world (Dunigan et al. 2006). Those that have been able to be cultured include 40 representatives, infecting 11 microalgae species, and have all been characterized as large dsDNA viruses (family *Phycodnaviridae*; Nagasaki and Bratbak 2010). To date, pond crashes due to viral infection have not been frequently reported, however virus particles have been detected inside algae cells growing in mass culture systems using TEM (Figure 5j; Roberson & Carney, unpublished data). Algal viruses play a significant role in the collapse of algal blooms in natural ecosystems and it is inevitable that viral pathogens will be discovered as impacting algal mass culture once viral presence becomes more routinely monitored for. Although PCR amplification of specific sequences related to the *Phycodnaviridae* has been demonstrated in marine samples (e.g., Chen and Suttle 1995), there have been no universal molecular barcode regions developed for viruses in general. Thus, the identification of novel viruses is dependent on shotgun sequencing of purified viral nucleic acid fractions followed by genomic assembly. Methods have been developed in both freshwater and marine systems to concentrate viral particles and prepare the nucleic acids for sequencing (Lawrence and Steward 2010).

7.5. Conclusions

As algal mass culture systems invariably grow in size, the financial impact of contamination will, concomitantly, become more severe. The development of rapid, sensitive, high-throughput methods of contaminant identification and detection is of paramount importance for the further economic development of the industry. Many of the molecular methods, such as next generation sequencing, PCR and microarray analysis described here have the characteristics that are necessary to fill this unmet need. The choice of methods will largely depend on the value of the product being produced. Operations that are producing low value products, such as fuel, face hard economic constraints and must limit analysis costs, whereas operations producing more valuable products such as nutraceuticals may have greater choice in surveillance methods. Clearly there is significant need for research and development into lower cost alternatives that leverage or employ molecular detection methods.

Two key developments will facilitate the development of molecular approaches in the management of algae biomass production. 1) The uniform extraction of DNA from populations in environmental samples will transform molecular diagnosis of pond communities from qualitative to quantitative. Qualitative analyses, or relative analyses are useful, but developing more advanced extraction technology is sure to speed this field. 2) The development of molecular targets that monitor genetic responses of the algae target to distress will be key moving forward. Instead of monitoring unique pest signatures, tracking algae stress response signatures may circumvent the need to have *a priori* knowledge of challenging organisms.

Once biocontaminants are detected by classical or more advanced monitoring methods, a rapid response is usually required in order to reduce or eliminate them. More detailed information on algae crop protection, including common responses in use today and novel approaches that are currently in development, is presented in Chapter 6.

Acknowledgements:

The manuscript was improved greatly by comments received from M. Strittmatter, S. Slocombe, J. Longcore, and J. Wilkenfeld. Figure 5, specifically, was much improved by edits received from J. Longcore. 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. Research at Sandia National Laboratories was supported by the BioEnergy Tehchnology Office of the Office of Energy Efficiency and Renewable Energy, U.S. Department of Energy under Award #NL0022897.

Literature cited:

Abarenkov, K., R. H. Nilsson, K. H. Larsson et al. 2010. The UNITE database for molecular identification of fungi – recent updates and future perspectives. *New Phytol.* 186:281-285.

Aird, D., M. G. Ross, W. S. Chen et al. 2011. Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries. *Genome Biol.* 12:R18.

Álvarez E., Lopez-Urrutia A., Nogueira E., and S. Fraga. 2011. How to effectively sample the plankton size spectrum? A case study using FlowCAM. *J. Plankton Res*. 33:1119-1133.

Amann R, N. Springer, W. Ludwig, H. D. Görtz, and K. H. Schleifer. 1991. Identification *in situ* and phylogeny of uncultured bacterial endosymbionts. *Nature* 351:161–164.

Amaral-Zettler, L. A., E. A. McCliment, H. W. Ducklow, and S. M. Huse. 2009. A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. *PLoS ONE* 4:e6372. doi:10.1371/journal.pone.0006372.

ANL, NREL, PNNL. 2012. Renewable diesel from algal lipids: an integrated Baseline for cost, emissions and resource potential from a harmonized model. ANL/ESD/12-4; NREL/TP-5100-55431; PNNL-21437. Argonne Il: Argonne National Laboratory; Golden CO: National Renewable Energy Laboratory; Richland WA: Pacific Northwest National Laboratory.

Bazzicalupo, A. L., M. S. Balint, and I. Schmitt. 2013. Comparison of ITS1 and ITS2 rDNA in 454 sequencing of hyperdiverse fungal communities. *Fungal Ecol.* 6:102-109.

Benson, D. A., M., Cavanaugh, K. Clark et al. 2012. GenBank. *Nucleic Acids Res.* 41(Database issue):D36-42. doi: 10.1093/nar/gks1195.

Binga, E. K., R. S. Lasken and J. D. Neufeld. 2008 Something from (almost) nothing: the impact of multiple displacement amplification on microbial ecology. *ISME J.* 2:233E Ji

Blaalid, R., S. Kumar, R. H. Nilsson, K. Abarenkov, P. M. Kirk, and H. Kauserud. 2013. ITS1 versus ITS2 as DNA metabarcodes for fungi. *Mol. Ecol. Resour.* 13:218–224.

Botes, M., M. de Kwaadsteniet, and T. E. Cloete. 2013. Application of quantitative PCR for the detection of microorganisms in water. *Anal. Bioanal. Chem.* 405:91-108.

Bucklin, A., D. Steinke, and L. Blanco-Bercial. 2011. DNA barcoding of marine metazoa. *Annu. Rev. Mar. Sci.* 3:471–508.

Bull, J. J. and S. Collins. 2012. Algae for biofuel: will the evolution of weeds limit the enterprise. *Evolution* 66:2983-2987.

Caporaso, J. G., J. Kuczynski, J. Stombaugh et al. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat. Meth.* 7:335-336. doi:10.1038/nmeth.f.3.

Caporaso, J.G., C. L. Lauber, W. A.Walters et al. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Nat. Acad. Sci. U.S.A* 108:4516-4522

Carney, L. T., and T. W. Lane. 2014. Parasites in algae mass culture. Special issue of *Front. Microbiol.* 5:278. doi: 10.3389/fmicb.2014.00278.

Carney, L. T., Reinsch, S. S., Lane, P. D. et al. 2014. Microbiome analysis of a microalgal mass

culture growing in municipal wastewater in a prototype OMEGA photobioreactor. *Algal Res.*4:52-61. http://dx.doi.org/10.1016/j.algal.2013.11.006

Carney, L. T., J. Wilkenfield, P. Lane et al. 2014. Pond Crash Forensics: Metagenomic analysis of predation on *Nannochloropsis salina* in raceways. *In prep. Bioresource Technol.*

Chelius, M. K. and E. W. Triplett. 2001. The diversity of Archaea and Bacteria in association with the roots of *Zea mays* L., *Microbial Ecol.* 41:252–263.

Chen F. and C. A. Suttle. 1995. Amplification of DNA polymerase gene fragments from viruses infecting microalgae. *Appl. Environ. Microbiol.* 61:1274–1278.

Coder D. M. and M. P. Starr. 1978. Antagonistic association of the chlorellavorous bacterium (*"Bdellovibrio" chlorellavorus*) with *Chlorella vulgaris*. *Current Microbiol.* 1:59-64.

Coder, D. M. and L. J. Goff. 1986. The host range of the chlorellavorous bacterium ("*Vampirovibrio chlorellavorus*"). *J. Phycol.* 22:543-546.

Cole J. J. 1982. Interactions between bacteria and algae in aquatic ecosystems. *Ann. Rev. Ecol. Syst.* 13:291-314.

Cole, J. R., Q. E. Wang, E. Cardenas et al. 2009. The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res.* 37:D141-D145.doi: 10.1093/nar/gkn879

Collins, A. M., H. D. T. Jones, R. C. McBride, C. Behnke, and J. A. Timlin. 2014. Host cell pigmentation in *Scenedesmus dimorphus* as a beacon for nascent parasite infection. *Biotechnol. Bioeng.* DOI: 10.1002/bit.25235.

Day, J. G., Thomas, N. J., Achilles-Day U. E. M., Raymond J. G., and R. J. G. Leakey. 2012. Early detection of protozoan grazers in algal biofuel cultures. *Bioresource Technol.* 114:715–719.

DeLong, E. F., L. T. Taylor, T. L. Marsh, and C. M. Preston. 1999. Visualization and enumeration of marine planktonic archaea and bacteria by using polyribonucleotide probes and Fluorescent In Situ Hybridization. *Appl. Environ. Microbiol*. 65:5554-5563.

Defoirdt, T., Boon, N., Bossier, P., and W. Verstraete. 2004. Disruption of bacterial quorum sensing: an unexplored strategy to fight infections in aquaculture. *Aquaculture.* 240:69-88.

DeSantis, T. Z., P. Hugenholtz, N. Larsen et al. 2006. Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Appl. Environ. Microbiol.* 72:5069-72.

Ding, H. Y. and J. H. Ma. 2005. Simultaneous infection by red rot and chytrid diseases in *Porphyra yezoensis* Ueda. *J. Appl. Phycol*. 17:51-56.

Dunigan, D. D., L. A. Fitzgerald, and J. L. Van Etten. 2006. Phycodnaviruses: a peek at genetic diversity. *Virus Res.* 117:119-132.

FAO, The state of world fisheries and aquaculture 2010, FAO Fisheries and Aquaculture Department, Food and Agriculture Organization of United Nations, Rome, 2010.
Fenchel T., G. F. Esteban, and B. J. Finlay. 1997. Local versus global diversity of microorganisms: cryptic diversity of ciliated protozoa. *Oikos* 80:220-225.

Fott, B. 1967. *Phlyctidium scenedesmi* spec. nova, a new chytrid destroying mass cultures of algae. *Z. Allg. Mikrobiol.* 7:97-102.

Fredricks, D. N. and D. A. Relman. 1996. Sequence-Based Identification of Microbial Pathogens: a Reconsideration of Koch's Postulates. *Clin. Microbiol. Rev*. 9:18–33.

Fukami, K., T. Nishijima, and Y. Ishida. 1997. Stimulative and inhibitory effects of bacteria on the growth of microalgae. *Hydrobiol.* 358:185-191.

Fulbright, S. P., M. K., Dean, G. Wadle, P. J. Lammers and S. Chisolm. 2014. Molecular diagnostics for monitoring contaminants in algal cultivation. *Algal Res*. 4: 41-51

Gao, Y., C. Gregor, Y. Liang, D. Tang, and C. Tweed. 2012. Algae biodiesel – a feasibility report. *Chem. Central J.* 6:S1.

Gutman, J. A. Zarka, and S. Boussiba. 2009. The host-range of *Paraphysoderma sedebokerensis*, a chytrid that infects *Haematococcus pluvialis*. *Eur. J. Phycol.* 44:509-514.

Hirayama, K. and S. Ogawa. 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.

Hoef-Emden, K. 2012. Pitfalls of establishing DNA barcoding systems in the protists: the Cryptophyceae as a text case. *PLoS ONE* 7:e43652. doi:10.1371/journal.pone.0043652

Hoffman, Y., C. Aflalo, A. Zarka, J. Gutman, T. Y. James, and S. Boussiba. 2008. Isolation and characterization of a novel chytrid species (phylum *Blastocladiomycota*), parasitic on the green alga *Haematococus*. *Mycol. Res.* 112:70-81.

Holland, P. M., R. D. Abramson, R. Watson, and D. H. Gelfand. 1991. Detection of specific polymerase chain reaction product by utilizing the 5'----3' exonuclease activity of Thermus aquaticus DNA polymerase. *Proc. Nat. Acad. Sci. U.S.A* 88:7276–7280

Hunt D. E., V. Klepac-Ceraj, S. G. Acinas, C. Gautier, S. Bertilsson, and M. F. Polz. 2006. Evaluation of 23S rRNA PCR Primers for Use in Phylogenetic Studies of Bacterial Diversity. *Appl. Environ. Microbiol.* 72:2221–2225.

Huse, S. M., L. Dethlefsen, J. A. Huber et al. 2008. Exploring Microbial Diversity and Taxonomy Using SSU rRNA Hypervariable Tag Sequencing. *PLoS Genet* 4:e1000255. doi:10.1371/journal.pgen.1000255.

Jobard, M., S. Rasconi and T. Sime-Ngando. 2010. Fluorescence in situ hybridization of uncultured zoosporic fungi: Testing with clone-FISH and application to freshwater samples using CARD-FISH. *J. Microbiol. Meth.* 83:236-243

Kagami, M., A. de Bruin, M. Rijkeboer, B. W. Ibelings, and E. Van Donk. 2007. Parasitic chytrids: their effect on phytoplankton communities and food-web dynamics. *Hydrobiol.* 578: 113-129.

Kazamia, E., H. Czesnick, T. T. Van Nguyen et al. 2012. Mutualistic interactions between vitamin B12-dependent algae and heterotrophic bacteria exhibit regulation. *Env. Microbiol.* 14:1466-1476.

Kim, M., M. Morrison, and Z. Yu. 2011. Evaluation of different partial 16S rRNA gene sequence regions for phylogenetic analysis of microbiomes. *J. Microbiol. Met.* 84:81–87.

Koetschan C., T. Hackl, T. Müller, M. Wolf, F. Förster, and J. Schultz. 2012.ITS2 database IV: Interactive taxon sampling for internal transcribed spacer 2 based phylogenies. *Mol. Phylogenet. Evol.* 63:585-588. doi: 10.1016/j.ympev.2012.01.026.

Koid, A., C. William, W. C. Nelson, A. Mraz, and K. B. Heidelberg. 2012. Comparative analysis of eukaryotic marine microbial assemblages from 18S rRNA gene and gene transcript clone libraries by using different methods of extraction. *Appl. Environ. Microbiol.* 78:3958.

Lawrence, J. E. and G. F. Steward. 2010. Purification of viruses by centrifugation. In *Manual of Aquatic Viral Ecology*, ed. S. W. Wilhelm, M. G. Weinbauer, and C. A. Suttle, 166-181. ASLO.

Lebaron P, P. Catala, C, Fajon, F. Joux, J. Baudart and L. Bernard. 1997. A new sensitive, whole-cell hybridization technique for detection of bacteria involving a biotinylated oligonucleotide probe targeting rRNA and tyramide signal amplification. *Appl. Environ. Microbiol*. 63:3274–3278.

Leliaert, F., H. Verbruggen, P. Vanormelingen et al. 2014. DNA-based species delimitation in algae. *Eur. J. Phycol*. 49:179-196.

Letcher, P. M., S. Lopez, R. Schmieder et al. 2013. Characterization of *Amoeboaphelidium protococcarum*, an algal parasite new to the Cryptomycota isolated from an outdoor algal pond used for the production of biofuel. *PLoS ONE* e56232.

Li, W., Zhang, T., Tang, X., and B. Wang. 2010. Oomycetes and fungi: important parasites on marine algae. *Acta. Oceanol. Sin.* 5:74-81.

Lindahl, B. D., R. H, Nilsson, L. Tedersoo et al. 2013. Fungal community analysis by high-throughput sequencing of amplified markers – a user's guide. *New Phytologist* 199: 288–299 doi: 10.1111/nph.12243.

Lindner D. L., T. Carlsen, R. H. Nilsson, M. Davey, T. Schumacher, and H. Kauserud. 2013. Employing 454 amplicon pyrosequencing to reveal intragenomic divergence in the internal transcribed spacer rDNA region in fungi. *Ecol. Evol.* 3:1751–1764.

Loy A., R. Arnold, P. Tischler, T. Rattei, M. Wagner, and M. Horn. 2008. probeCheck - a central resource for evaluating oligonucleotide probe coverage and specificity. *Environ. Microbiol.* 10:2894-2898.

Loy, A., A. Lehner, N. Lee et al. 2002. Oligonucleotide microarray for 16S rRNA gene-based detection of all recognized lineages of sulfate-reducing prokaryotes in the environment. *Appl. Environ. Microbiol.* 68:5064-5081.

Loy, A., F. Maixner, M. Wagner, M. Horn. 2007. probeBase - an online resource for rRNA-targeted oligonucleotide probes: new features 2007. *Nucleic Acids Res*. 35: D800-D804

Ludwig, W. and K. H. Schleifer. 1994. Bacterial phylogeny based on 16S and 23S rRNA sequence analysis. *FEMS Microbiol. Rev.* 15:155-173.

McBride, R. C., S. Lopez, C. Meenach et al. 2014. Contamination management in low cost open algae ponds for biofuels production. *Industr. Biotechnol.* 10:221-227.

McKenna, P., C. Hoffmann, N. Minkah et al. 2008. The macaque gut microbiome in health, lentiviral infection, and chronic enterocolitis. *PLoS Pathog.* 4:e20. doi:10.1371/journal.ppat.0040020

Metfies, K., M. Berzano, C. Mayer et al. 2007. An optimized protocol for the identification of diatoms, flagellated algae and pathogenic protozoa with phylochips. *Mol. Ecol. Notes* 7:925–936.

Miller R.R., V. Montoya, J.L. Gardy, D.M. Patrick and P. Tang. 2013. Metagenomics for pathogen detection in public health. *Genome Med*. 5:81. doi:10.1186/gm485

Murphy, T. E., K. Macon, and H. Berberoglu. 2014. Rapid algal culture diagnostics for open ponds using multispectral image analysis. *Biotechnol. Prog.* 30: 233-240.

Nagasaki, K. and G. Bratbak. 2010. Isolation of viruses infecting photosynthetic and nonphotosynthetic pro- tests. In *Manual of Aquatic Viral Ecology*, ed. S. W. Wilhelm, M. G. Weinbauer, and C. A. Suttle, 82-91. Waco: American Society of Limnology and Oceanography.

Nickrent, D. L. and M. L. Sargent. 1991. An overview of the secondary structure of the V4 region of eukaryotic small-subunit ribosomal RNA. *Nucleic Acid Res.* 19:227–235.

Not, F., N. Simon, I. C. Biegala, and D. Vaulot. 2002. Application of fluorescent *in situ* hybridization coupled with tyramide signal amplification (FISH-TSA) to assess eukaryotic picoplankton composition. *Aquat. Microb. Ecol.* 28:157-166.

Orsi, W., J. F. Biddle, and V. Edgcomb. 2013. Deep Sequencing of Subseafloor eukaryotic

rRNA reveals active fungi across marine subsurface provinces. *PLoS ONE* 8:e56335. doi:10.1371/journal.pone.0056335.

Park, J. B. K., Craggs R. J., and A. N. Shilton. 2011. Wastewater treatment high rate algal ponds for biofuel production. *Bioresour. Technol.* 102:35-42.

Pawlowski, J., R. Christen, B. Lecroq, D. Bachar, H. R. Shahbazkia, et al. 2011. Eukaryotic richness in the abyss: insights from pyrotag sequencing. *PLoS ONE* 6:e18169. doi:10.1371/journal.pone.0018169.

Pernthaler A, J. Pernthaler and R. Amann. 2002. Fluorescence in situ hybridization and catalyzed reporter deposition for the identification of marine bacteria. *Appl. Environ. Microbiol.* 68:3094–3101.

Pienkos, P.T. and A. Darzins, 2009. The promise and challenges of microalgal-derived biofuels. *Biofuels Bioprod. Biorefin*. 3:431-440.

Pinter, P. J., J. L. Hatfield, J. S. Schepers et al. 2003. Remote sensing for crop management. *Prog. Engin. Rem. Sensing* 69:647-664.

Ponchel, F., C. Toomes, K. Bransfield et al. 2003. Real-time PCR based on SYBR-Green I fluorescence: An alternative to the TaqMan assay for a relative quantification of gene rearrangements, gene amplifications and micro gene deletions *BMC Biotechnol.* 3:18.

Purdy, K. J. 2005 Nucleic acid recovery from complex environmental samples. *Method. Enzymol.* 397:271–292.

Putignani, L., M. G. Paglia, E. Bordi, E. Nebuloso, L. P. Pucillo, and P. Visca. 2008. Identification of clinically relevant yeast species by DNA sequence analysis of the D2 variable region of the 25–28S rRNA gene. *Mycoses* 51:209–227.

Quail, M. A., M. Smith, P. Coupland, T. D. et al. 2012. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* 13:341 doi:10.1186/1471-2164-13-341.

Quast, C., E. Pruesse, P. Yilmaz et al. 2013. The SILVA Ribosomal RNA Gene Database Project: improved data processing and web-based tools. *Nucl. Acids Res.* 41(D1): D590-D596.

Ramaiah, N. 2006. A review on fungal diseases of algae, marine fishes, shrimps and corals. *Indian J. Mar. Sci*. 35:380-387.

Rasconi, S., M. Jobard, L. Jouve, and T. Sime-Ngando. 2009. Use of Calcofluor white for detection, identification and quantification of phytoplanktonic fungal parasites. *Appl. Env. Microbiol.* 75:2545-2553.

Rashidan, K. K. and D. F. Bird. 2001. Role of predatory bacteria in the termination of a cyanobacterial bloom. *Microbiol. Ecol.* 41:97-105.

Rastogi, G., J. J. Tech, G. L. Coaker, and J. H. J. Leveau. 2010. A PCR-based toolbox for the culture independent quantification of total bacterial abundances in plant environments. *J. Microbiol. Method.* 83:127-132.

Ratnasingham, S. and P. D. N. Hebert. 2007. BOLD: The barcode of life data system (www.barcodinglife.org). *Mol. Ecol. Notes* 7:355–364. doi: 10.1111/j.1471-8286.2006.01678.x.

Reichardt, T. A., A. M. Collins, R. C. McBride, C. A. Behnke, and J. A. Timlin. 2014. Spectroradiometric monitoring for open outdoor culturing of algae and cyanobacteria. *Appl. Optics* 53: F31-F45.

Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA Sequencing with chain-terminating inhibitors. *Proc. Nat. Acad. Sci. U.S.A* 74:5463-5467.

Sarma, S. S. S., P. S. L. Jurado, and S. Nandin. 2001. Effect of three food types on the population growth of *Brachionus calyciflorus* and *Brachionus patulus* (Rotifera: Brachionidae). *Rev. Biol. Trop.* 49:77-84.

Saunders, G. W. and D. C. McDevit. 2012. Methods for DNA barcoding photosynthetic protists emphasizing the macroalgae and diatoms. In *DNA Barcodes: Methods and Protocols*, ed. W. J. Kress and D. L. Erickson, Methods in Molecular Biology, vol. 858:207-222.

Schloss. P. D, S. L. Westcott, T. Ryabin et al. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* 75:7537-41

Simon, N., N LeBot, D Marie, F Partensky, and D Vaulot. 1995. Fluorescent in situ hybridization with rRNA-targeted oligonucleotide probes to identify small phytoplankton by flow cytometry. *Appl. Environ. Microbiol.* 61:2506-2513.

Song. H., J.E. Buhay, M.F. Whiting, and K.A. Crandall. 2008. Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. *Proc. Nat. Acad. Sci. U.S.A* 105:13486-13491.

Sonnenberg, R., A. W. Nolte, and D. Tautz. 2007. An evaluation of LSU rDNA D1-D2 sequences for their use in species identification. *Frontiers Zool.* 4:6. doi:10.1186/1742-9994-4-6

Steven, B., S. McCann, and N. L. Ward. 2012. Pyrosequencing of plastid 23S rRNA genes reveals diverse and dynamic cyanobacterial and algal populations in two eutrophic Lakes. *FEMS Microbiol. Ecol.* 82:607–615.

Stoeck T., D. Bass, M. Nebel et al. 2010. Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Mol. Ecol.*

19:21–31.

Subasinghe, R. P., Bondad-Reantaso, M. G., and S. E. McGladdery 2001. Aquaculture development, health and wealth. In *Aquaculture in the Third Millennium. Technical Proceedings of the Conference on Aquaculture in the Third Millennium,* ed. R. P. Subasinghe, P. Bueno, M. J. Phillips, C. Hough, S. E. McGladdery and J. R. Arthur, 167-191. Bangkok: NACA and FAO.

Van de Peer, Y., S. Chapelle, and R. De Wachter. 1996. A quantitative map of nucleotide substitution rates in bacterial rRNA. *Nucleic Acids Res.* 24:3381–3391.

Vestheim H. and S. N. Jarman. 2008. Blocking primers to enhance PCR amplification of rare sequences in mixed samples – a case study on prey DNA in Antarctic krill Stomachs. *Frontiers Zool.* 5:12. doi:10.1186/1742-9994-5-12.

Whiteley A. S, S. Jenkins, I. Waite et al. 2012. Microbial 16S rRNA Ion Tag and community metagenome sequencing using the Ion Torrent (PGM) Platform *J. Microbiol. Meth.* 91:80–88.

Yúfera M. and N. Navarro. 1995. Population growth dynamics of the rotifer *Brachionus plicatilis* cultured in non-limiting food conditions. *Hydrobiol.* 313/314:399-405.

Figure 1. Map of prokaryotic SSU rRNA gene showing variable regions (nucleotide numbering corresponds to the *Escherichia coli* SSU rRNA gene)

Figure 2. Map of the variable regions in the eukaryotic LSU gene.

Figure 3. Map of the entire eukaryotic rRNA transcript structure showing ITS regions.

Figure 4. *Reproduced from McBride et al. 2014.* An example of a reactive pest response strategy aided by the use of quantitative PCR (qPCR). This strategy is triggered by qualitative microscope observations of pests in field samples and/or through bringing a pond sample into the lab and exposing it to conditions that would accelerate pest growth and precipitate a culture crash. If a new or unknown pest is observed, various microbiological techniques (e.g. plaque plating, baiting, selective media) first are used to isolate this pest organism. ITS or SSU rRNA sequencing is then used to identify the pest organism and to develop qPCR primers to track and quantify the genomic DNA of the targeted pest in established ponds. If isolation and culturing is not possible, metagenomic sequencing of the pond community may reveal enough of a pattern to identify a specific pest organism and provide sequence data for qPCR assay design. Specific thresholds can be set for the qPCR-determined abundance of each pest that is being monitored. Once this critical abundance threshold is reached in any given pond, a carefully chosen crop protection strategy (e.g. a selected pesticide application, or the addition of a targeted biocontrol agent) is then implemented and the pest's abundance is consistently monitored to determine whether the control strategy is successful (see chapter 6).

Figure 5. Biocontaminants observed in mass algal culture: a) grazing ciliate, *Gastrostyla steinii*, with ingested whole *Haematococcus* cysts (photo K. Sorensen). b) Sporangia of a parasitic blastocladian fungus, probably *Paraphysoderma* (sequence identification pending), attached to *Haematococcus* growing in open ponds (photo L. T. Carney). c) Higher magnification of *Paraphysoderma* (isolate JEL821), note fungal rhizoids extending into *Haematococcus* cyst, showing signs of cell content depletion (photo J. E Longcore). d) Bacteria attached to lysing algal cells in crashing culture correlated with increasing detection of the pathogenic bacterium, *Vampirovibrio chlorellavorus*, on multiple occasions using SSU rRNA sequencing and qPCR (photo S. Qin). e) Motile *V. chlorellavorus*-like bacteria adjacent to algae cells in culture that was confirmed by qPCR to contain *V. chlorellavorus* (photo J. Wilkenfeld). f) Sporangium of isolate JEL812; arrow points to rhizoidal connection to *Haematococcus* cell (photo J. E. Longcore). g) Algae cells prior to appearance of attachment and detection of *V. chlorellavorus* (photo S. Qin). h) Amoeboid swarmer stages of probable *Paraphysoderma* observed in open ponds growing *Haematococcus* (photo J. Wilkenfeld). i) Motile zoospore stage of the parasitic chytrid isolate JEL812, (Rhizophydiales; Aquamycetaceae; photo K. Sorensen). j) TEM photograph of virus particles inside an algal cell (photo R. Roberson). Note: scale bars = 10 μm where not otherwise indicated.

Fig. 1

SSU rRNA

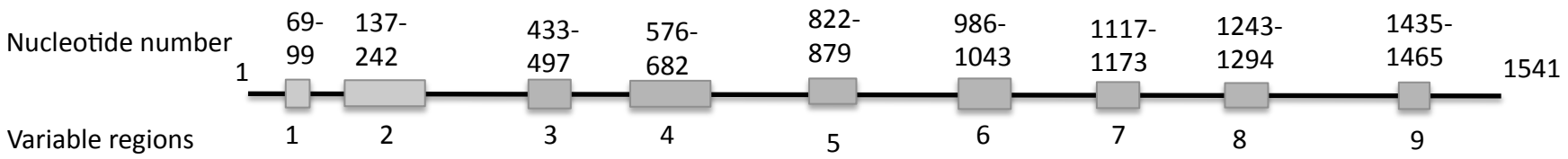


Fig. 2

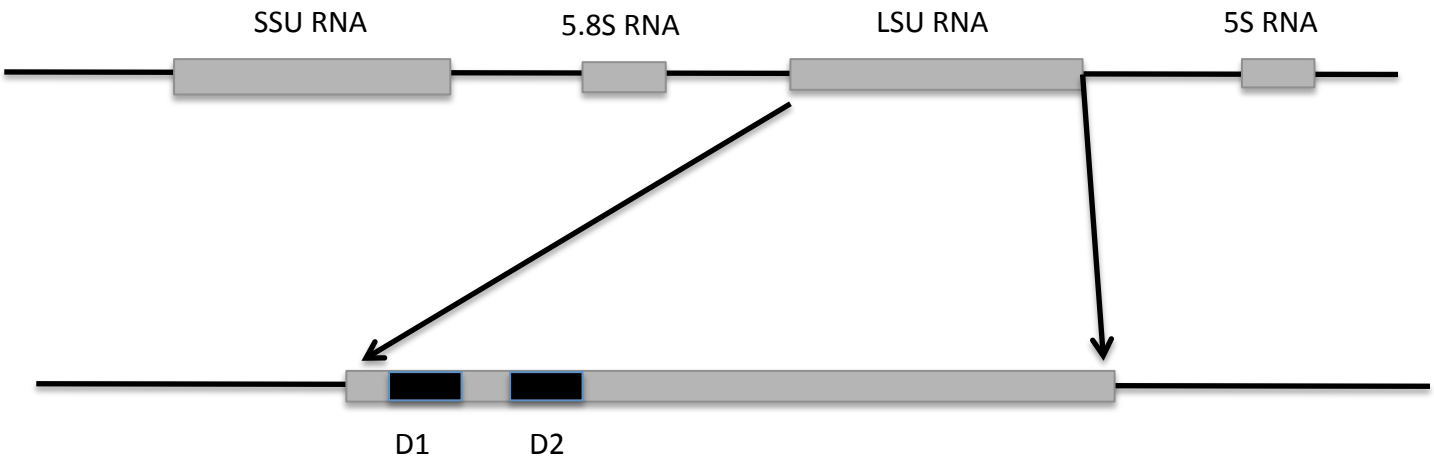
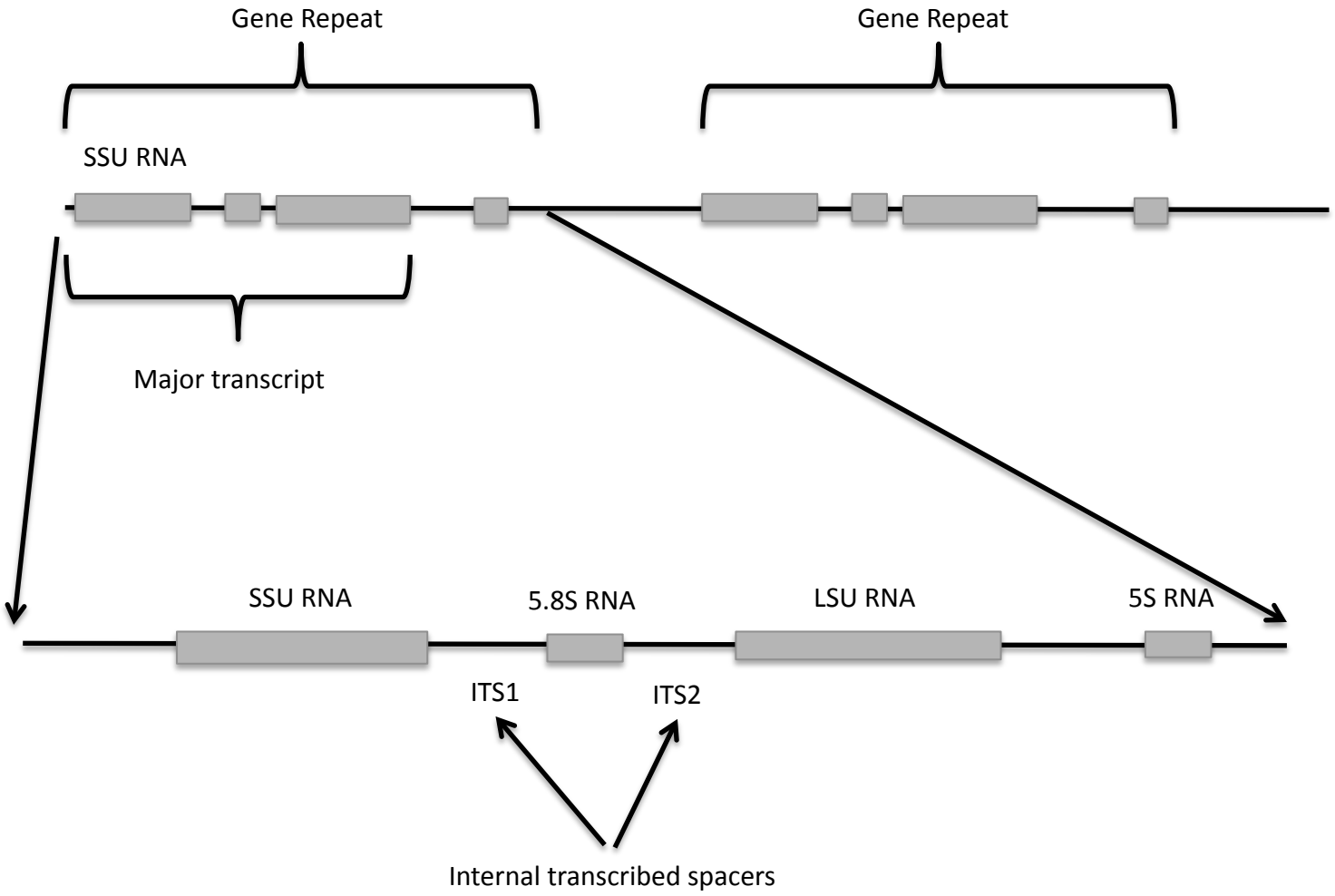
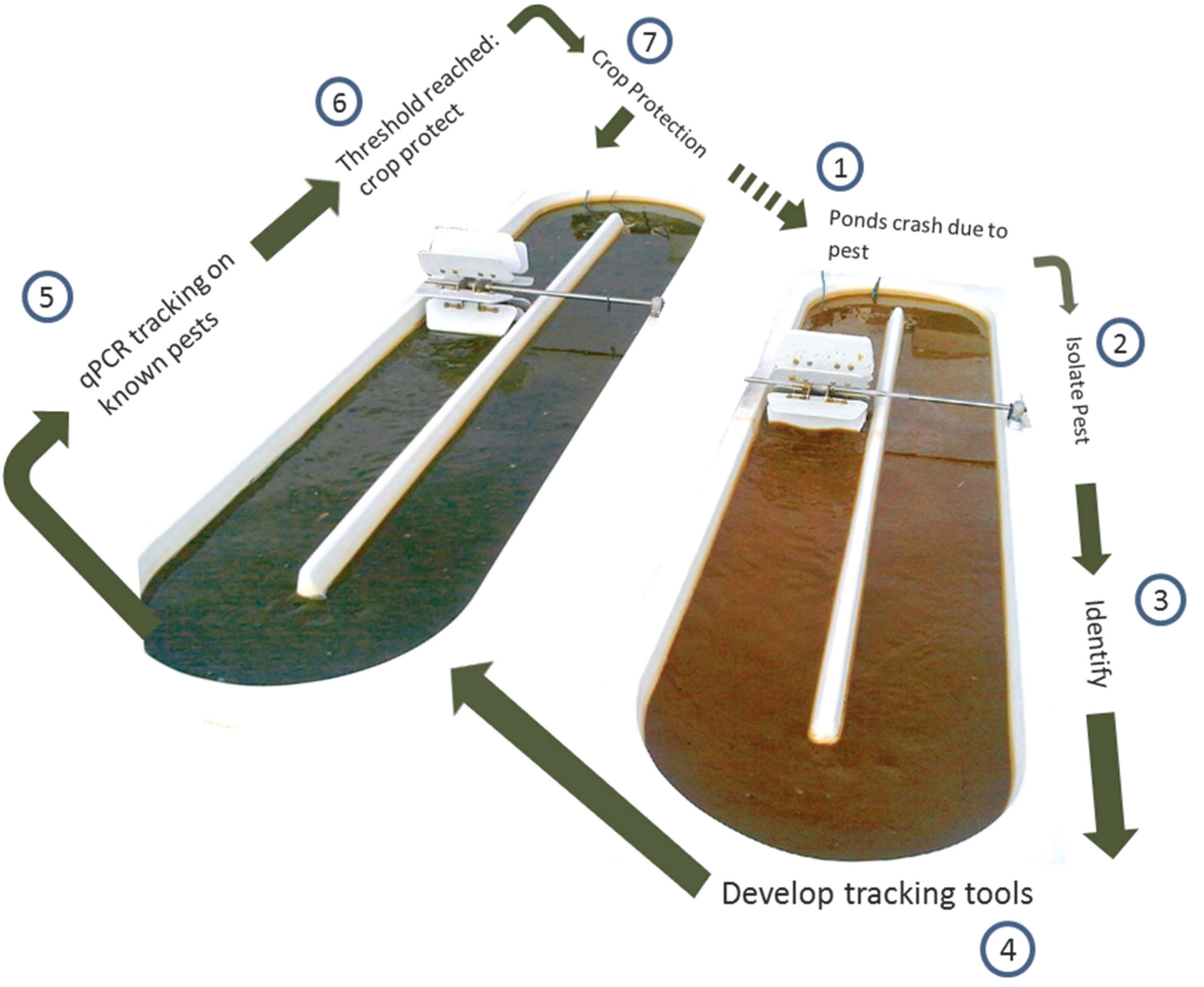
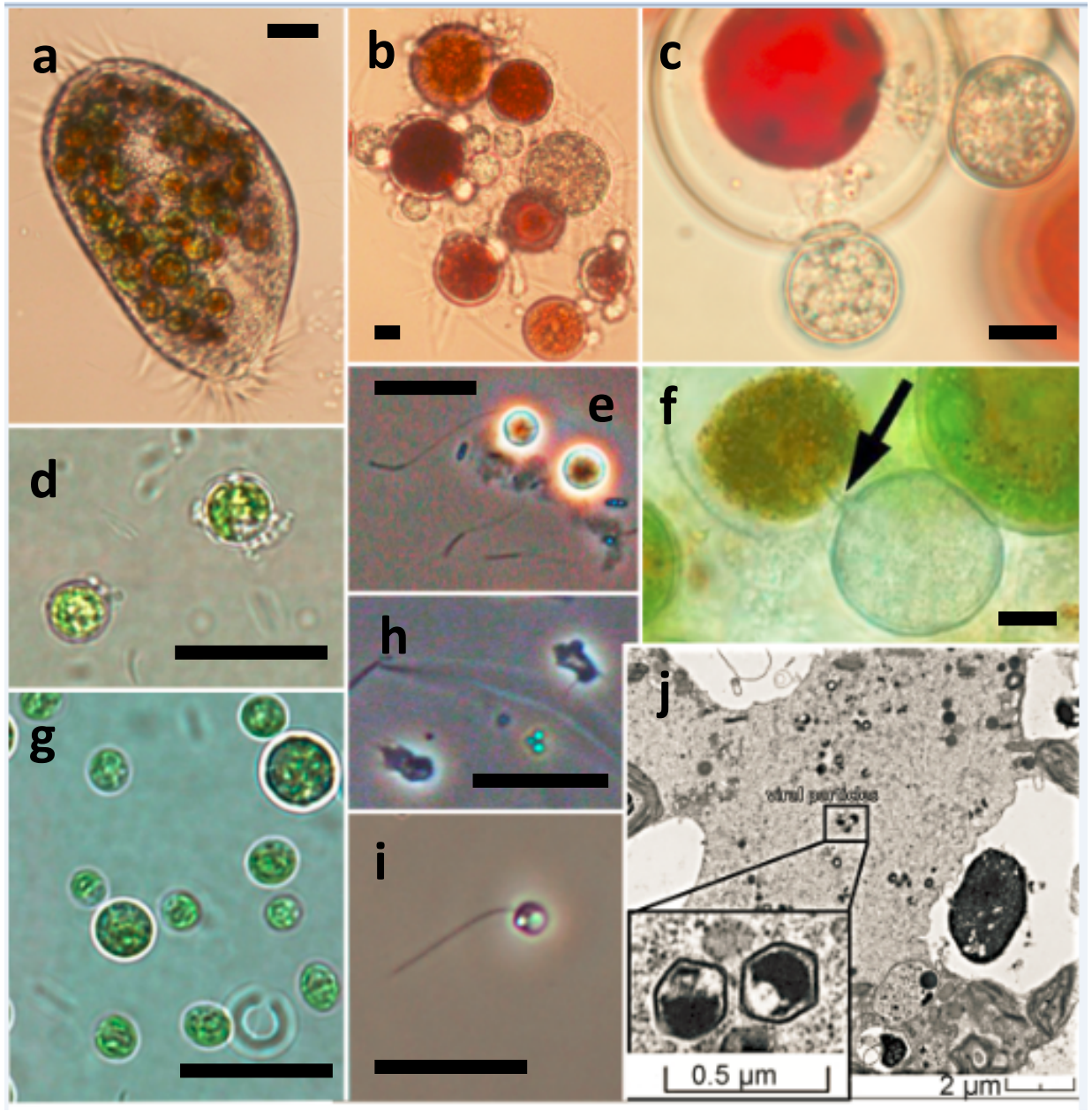


Fig. 3


Gene Repeat
Gene Repeat
SSU RNA
Major transcript
SSU RNA
5.8S RNA
LSU RNA
5S RNA
ITS1
ITS2
Internal transcribed spacers


1
Ponds crash due to pest
2
Isolate Pest
3
Identify
4
Develop tracking tools
5
qPCR tracking on known pests
6
Threshold reached: crop protect
7
Crop Protection


a
b
c
d
e
f
h
g
i
j
viral particles
0.5 µm
2 µm