Fall 12-2020

# The Design and Validation of a qPCR-Based Approach to Cell Quantification of Fungal-Algal Symbioses

Isabelle Katherine Garlotte
*University of Mississippi*

# THE DESIGN AND VALIDATION OF A qPCR BASED APPROACH TO CELL QUANTIFICATION OF FUNGAL-ALGAL SYMBIOSES

By
Isabelle Katherine Garlotte

A thesis submitted to the faculty of The University of Mississippi in partial fulfillment of the requirements of the Sally McDonnell Barksdale Honors College.

Oxford
December 2020

Approved by

DocuSigned by:
Dr. Erik Hom
0E090DC3B5254AE...
Advisor: Dr. Erik Hom

DocuSigned by:
Dr. Wayne Gray
81703FA76AD54B9...
Reader: Dr. Wayne Gray

DocuSigned by:
Dr. Jason Hoeksema
42A8E0C06EA14EA...
Reader: Dr. Jason Hoeksema

# ACKNOWLEDGEMENTS

I would like to thank Dr. Erik Hom for his mentorship throughout this project, commitment to being my thesis advisor, and for allowing me to perform research in his laboratory. I would like to thank the University of Mississippi Biology Department and the Sally McDonnell Barksdale Honors College for their support in both my undergraduate education at Ole Miss, and in the process of this research project. I would also like to thank Michael Clear for all of his tremendous amounts of help and guidance throughout my past two and a half years working in the Hom lab and on this project in particular, and for never getting tired of my many, many questions. I would also like to thank Dr. Wayne Gray and Dr. Jason Hoeksema for agreeing to serve on my thesis committee. Lastly, I would like to thank my family for their undying support of my education throughout my entire life and for fostering my love for learning at a young age. Without them, none of this would be possible.

ABSTRACT

ISABELLE KATHERINE GARLOTTE: The Design and Validation of a qPCR Based Approach to Cell Quantification of Fungal-Algal Symbioses
(Under the direction of Dr. Erik Hom)

Symbiotic relationships are extremely common in virtually all species. These relationships between organisms, particularly mutualistic symbioses, can have an incredibly large impact and can greatly affect the evolution of the species involved. Of particular interest is the suspected mutualistic relationship between green algae and fungi that may hold key evidence of how terrestrial plants evolved. In this experiment we established a method of quantifying the cells of each species involved in these symbioses through qPCR analysis to determine if the relationships are truly mutualistic. We developed a system for using DNA copies as a proxy for cell count in co-cultures of various species of algae and fungi to determine the nature of these symbioses. To perform a successful qPCR analysis, both primers and standards had to be created for each species and a method for quantification had to be determined. In this study, an effective method for producing reliable qPCR copy numbers in both the algal and fungal species was created and validated through preliminary experimental results. The methodology established in this thesis may be used in future studies of the fungal-algal co-culture assays to gain more insight into their possible mutualistic symbioses.

TABLE OF CONTENTS

## LIST OF FIGURES AND TABLES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| PCR | Polymerase Chain Reaction |
| qPCR | Quantitative Polymerase Chain Reaction |
| DNA | Deoxyribose Nucleic Acid |
| NCBI | National Center for Biotechnology Information |
| CDS | Coding DNA Sequences |
| BLAST | Basic Local Alignment Search Tool |
| MUSCLE | Multiple Sequence Comparison by Log-Expectation |
| $T_m$ | Melting Temperatures |
| NEB | New England BioLabs |
| dsDNA | Double Stranded DNA |
| Cq | Quantification Cycle |
| µL | Microliter |
| mL | Milliliter |
| °C | Degrees Celsius |
| ng | Nanograms |
| OPC | One Point Calibration |

## INTRODUCTION

Symbiotic relationships in which organisms are living directly on, within, or closely with another species are of great interest to biologists. In 1878, scientist Heinrich Anton de Bary described symbioses as "the living together of two dissimilar organisms" that are usually "to the benefit of at least one of them" (Dublier et. al 2008, deBary 1887). While many think of symbioses as being mostly beneficial, it was originally used to describe a multitude of symbiotic relationships involving intimate physical association of partners: some benefitting only one species or the other, some harming one species for the other's benefit, and some benefitting both species (Dublier et. al 2008). Symbioses are seen in virtually every known form of life at some level (Paracer and Ahmadjian 2000). In mutualistic symbioses, the interaction benefits the growth and/or fitness of both species involved. Mutualisms are particularly important due to their ability to profoundly shape the evolution of species (Bronstein 2015). Examples of mutualistic relationships are evident in organisms throughout every kingdom, from pollinating insects and flowers, to bacteria living in the human gut; mutualistic relationships are largely responsible for creating the diverse world we live in (Bronstein 2015).

Of special interest are mutualistic and/or symbiotic relationships between various species of endophytic and endolichenic fungi and green algae. Similar to the mutualistic symbioses seen between land plants and arbuscular mycorrhizal fungi, it is thought that a mutualism between green algae and fungi may have led to the early ancestors of land plants that evolved to become terrestrial (Douglas 2010). Endophytic and endolichenic

fungi were chosen for this study due to their known ability to form symbioses with plants and lichens respectively, which may be an indicator of an evolutionary relationship. Fungi were previously isolated from plants or lichen and species were selected that best represented a genetically and geographically diverse subset of fungi. Fungal samples were selected and obtained from Dr. Elizabeth Arnold's lab's culture collection (University of Arizona). Green algae were chosen because they are thought to be the ancestor of land plants (Karol et. al 2001). Various species of charophyte green algae (*Penium margaritaceum*, *Klebsormidium flaccidum*, and *Zygnema circumcarinatum*) and chlorophyte green algae (*Chlamydomonas reinhardtii*, *Chlorella vulgaris*) were selected to include organisms from both lineages of green algae that have been well studied.

These fungi and algae have been shown to grow closely together with tight physical associations and are suspected to form mutualistic relationships (Hom and Murray 2014). Although there appears to be "wall-to-wall contact" between the fungi and algae studied, there doesn't appear to be any extracellular structures connecting the two species (Hom and Murray 2014). Although the two species are seen tightly growing together, it is not clear whether this symbiotic relationship is actually mutualistic (Hom and Murray 2014). We currently have biomass data for various assays of single fungal, single algal, and fungal/algal co-cultures. However, this data alone doesn't give us the whole picture. We do not know what the effect of the symbiotic relationship is on each species individually. We may have a co-culture where both species are growing better than they would in a mono-culture, but we may also have co-cultures where only one species is benefitting from their interactions. This leaves us wondering about the true nature of the relationship between fungal and algal partners in these co-cultures.

In this study, we aimed to better delineate the mutualistic potential of fungal/algal relationships, particularly on a cellular level. We used qPCR as a method to approximate cell numbers of both algae and fungi living in co-culture, and thus better estimate the fitness of each partner. Using this method of cell quantification would allow one to determine if both species grow better together than separately, a sign of a mutualistic relationship. This thesis examines the factors necessary to set up a successful qPCR experiment that can be used to analyze fungal/algal cell counts. Chapter I will focus on the basics of PCR and the qPCR methodology as well as the design and testing of the qPCR primers. Chapter II will describe the creation and use of qPCR standards. Lastly, Chapter III presents details of data analysis and preliminary experimental results on fungal and algal cultures.

# CHAPTER I

## *qPCR Theory and Primer Design*

### *1.1 Understanding Quantitative PCR*

Developed by Kary Mullis in the 1980's, the Polymerase Chain Reaction, known as PCR, won the Nobel Prize for Chemistry in 1993. This reaction takes advantage of the way that DNA replicates itself semi-conservatively to quickly amplify a target gene (Mullis et al. 1986). This discovery has since been revolutionary in molecular biology, being utilized in research, medicine, and criminal forensics. This reaction only requires a simple mixture of DNA polymerase enzyme, free nucleotides, DNA primer sequence, and a sample of template DNA that will be replicated (Green and Sambrook 2019). With these ingredients, and cycles of heating and cooling to help along the steps of the reaction, one copy of a small sequence of DNA can turn into millions of copies fairly quickly (Polymerase Chain Reaction (PCR) Basics). This occurs because during each cycle of heating and cooling every segment of DNA is replicated and made into two identical segments. Therefore, after each cycle the number of copies of that DNA segment should double if working at 100 percent efficiency. This is represented by the equation: $y = 2^n$ where y is equal to the total number of DNA copies and n is equal to the number of cycles.

The Polymerase Chain Reaction occurs in three main steps: Denaturation, Annealing, and Extension (Green and Sambrook 2019). In the Denaturation step the

reaction mixture is placed under high heat, around 95 ℃, to break apart double-stranded DNA (Green and Sambrook 2019). The two strands of DNA are separated from one another as the hydrogen bonds that hold the nucleotide bases together are broken by the addition of heat. In the following step, the Annealing step, the reaction mixture is cooled to about 60 ℃ (Green and Sambrook 2019). In this step the DNA primers, small portions of DNA that match up with the ends of the segment of DNA one wants to replicate, are annealed to the target sequence. Lastly, in the Extension step the free nucleotides that correspond to the sequence being replicated are added onto the strand following the primer by the enzyme DNA polymerase. This is performed by heating up the reaction mixture again to about 75 ℃ (Green and Sambrook 2019). This cycle of Denaturation, Annealing, and Extension is repeated typically for around thirty to forty times to achieve sufficient amplification of the target sequence (Green and Sambrook 2019).

A study describing quantitative PCR (also known as real time PCR), hereby referred to as qPCR, was published by Higuchi et al. (1993), describing how the amplification of a target gene can be observed as it is occurring (in real time). This technique can be used to help determine how many copies of a gene a sample starts out with before PCR occurs. Quantitative PCR includes the addition of fluorescent molecules into the reaction mixture that attach to each replicated DNA segment (Higuchi et. al 1993). The fluorescent molecules (SYBR® Green dye) conformationally change when they bind to double-stranded DNA and emit a higher level of fluorescent light due to this change in conformation (Arya et. al 2005). While the reaction cycles, the fluorescence levels are read by the thermocycler machine and recorded by the computer program accompanying the machine. The amount of fluorescence observed during the reaction

correlates with the number of DNA copies produced. This technology allows for amplification of a target gene to be observed as it is occurring and for an exact number of DNA copies in a sample to be calculated. This copy number calculation is the key to determining an approximate cell number in the samples.

### *1.2 Primer Design and Testing*

A key component in the mixture for a polymerase chain reaction is the oligonucleotide primers. Primers are small segments (15-25 bases) of single stranded DNA that align with the target gene. DNA polymerase needs the primers to have something from which to build the replicated copy. From the primers, the polymerase enzyme used in PCR adds on additional free nucleotides that match the parent strand of DNA at the newly available 3' end of the primer sequence to fill in the remainder of the amplified gene. In this experiment, new primers were designed for five species of algae (*Chlamydomonas reinhardtii*, *Chlorella vulgaris*, *Penium margaritaceum*, *Klebsormidium flaccidum*, and *Zygnema circumcarinatum)* and thirteen universal primers for fungi were designed, analyzed, and tested.

#### **1.2.1** ***Algal Primer Design and Theory***

When using qPCR as a proxy for quantifying cell abundance, it is important to identify a target gene that is only present once in the genome. Being able to know that every cell should have one copy present of this gene will allow us to use the number of copies of that gene present in a sample as a proxy for cell count. The centrin gene, which we chose, is known to be a single copy gene in at least one of our species of algae,

*Chlamydomonas reinhardtii* (Bhattacharya et al. 1993). The centrin gene codes for a protein coming from the EF-hand superfamily and produces a protein critically involved in calcium binding in the cytoskeleton (Bhattacharya et al. 1993).

Primer design began by identifying a copy of the centrin gene in the species and strains of each algae. The genome for *C. reinhardtii* has been well annotated and a coding DNA sequence (CDS) for the centrin gene was obtained by keyword search through the NCBI database. For *C. vulgaris*, we were able to identify the CDS for the centrin gene in strain C-169 using the NCBI genome database. While C-169 was not the strain of *C. vulgaris* we have used in our experiments (UTEX 389), it is thought the two are very closely related and thus should have almost identical copies of the centrin gene. The centrin gene for a closely related *Klebsormidium* species, *K. nitens,* was obtained using OrthoDB's ortholog protein database where we found the centrin protein sequence and CDS which we used as a substitute for the gene in our *Klebsormidium* species, *K. flaccidum* (Kriventseva et al. 2018). Although the sequences weren't for the same species of *Klebsormidium*, we hoped that the two species would be genetically similar enough, due to their shared genetic history, to create functioning primers. For *P. margaritaceum*, the centrin gene was found by BLASTing known centrin genes (from *C. reinhardtii*) against the *P. margaritaceum* genome in a nucleotide to nucleotide blast (blastn) which allowed us to identify the centrin gene through genetic sequence similarity (Jiao et. al 2020). Lastly, we were unable to locate a copy of the centrin gene for *Z. circumcarinatum*. To find a sequence we could use to construct primers we created a consensus sequence from closely related algal species (other charophytes, *P. margaritaceum* and *K. nitens*) to create a centrin gene that would most likely closely

resemble that of *Z. circumcarinatum*. To do this, we then downloaded the sequences as FASTA files and imported them into the UGENE program for sequence alignment. Sequences were aligned via MUSCLE alignment to determine which region of the gene has the fewest nucleotide substitutions and the most similarity across the different species and strains. This alignment is used by the program to create a consensus sequence which we then used as our centrin sequence for *Z. circumcarinatum.*

We used the Primer3 program available through NCBI's online PrimerBLAST tool to design the primers. Parameters were set to control the primer length to between 18-24 base pairs and the melting temperatures ($T_m$) to 57-63 °C. Primers of this length are ideal for PCR reactions because they are just specific enough to ensure you are amplifying the correct target (Singh and Kumar 2001). Of the 10 output pairs from Primer3 we selected two primer pairs for each organism based on the following criteria: 1) The length of the gene target, 2) The location of the gene target, and 3) The number of possible hits on the genome. In a qPCR experiment, ideally, we need to choose primers that amplify a target sequence of as close to 200 base pairs as possible because shorter sequences can be amplified more efficiently and are still long enough to be recognized as being a desired product and not a primer dimer (qPCR Assay Design and Optimization). Secondly, we wanted to choose primers that amplified sequences on different parts of the target sequence in order to ensure that at least one of the chosen primer pairs worked because a nucleotide substitution that happens to fall in the region where the primers anneal would render the primers useless. By choosing primers that are annealing in different locations on the target gene we minimize the chance of this happening. Lastly, we wanted to ensure that the primers we designed would only amplify a segment of the

target gene. If a primer pair could amplify in multiple genes within the genome, we could get multiple PCR products and therefore wouldn't be able to use our qPCR data as a means of cell quantification. We checked this by BLASTing our newly created primers against the whole genomes of our species and ensuring they only had one BLAST result on the centrin gene.

### 1.2.*2 Fungal Primer Design*

The degenerate qPCR primers were designed for our fungal species in order to minimize the number of primers required to analyze all of the fungi in our study. We first used fungal specific primers provided by Francois Lutzoni's lab to amplify the RPB1 and RPB2 regions, which are two genes coding for the two largest RNA polymerase subunits (subunits 1 and 2), as our target genes for our fungal samples (Matheny et al. 2002, Lui et al. 1999). We used these primers to amplify these regions via PCR in all of our fungal species. We then sequenced these regions from our PCR product and then aligned all of the sequences with a MUSCLE alignment in the program Geneious and determined the regions of the genes that were most similar across all species. Regions of the sequences had a better alignment (more consensus) for certain species were used as the sequences to create our new primers for those species. We then designed primers that should work with as many of our fungal species as possible. Many of our primer pairs matched up with only one fungal sample. However, there were a few that matched up with two or three samples. We then recorded which primers matched with which samples as these would be the DNA samples we used to test the primers. We ended up with a total

of 13 primer pairs that needed to be tested with a total of 18 fungal samples (some of which matched with multiple primer pairs).

We then tested all of the primers with their matching fungal samples through qPCR. We calculated the annealing temperature with the NEB online tool with the Q5 product group and a primer concentration of 400nM. In this initial run no DNA amplification was observed. All 13 of the fungal primers failed their initial test. There were several potential causes of this failed qPCR: low quality of template DNA, expired polymerase enzyme, incorrect annealing temperatures used, or the primers are all designed badly. It was unlikely that all 13 primers were bad, so we explored the other potential causes of our failed test.

To ensure the quality of template DNA we were using to test the primers, we quantified the DNA concentration of all the fungal samples we would be using in testing. We did these quantifications using the QuantiFluor dsDNA kit. A few of the samples had low to no detectable concentration of DNA. In these samples that was an obvious reason for why our qPCR had failed. However, several samples had more than enough DNA to have worked. We then extracted a second sample of DNA (using the NucleoSpin soil extraction kit) from those fungi that did not have an adequate concentration; which we deemed as anything with a concentration of less than 2.00 ng/microliter.

Our second potential cause for the failed qPCR was an expired DNA polymerase enzyme. If this was the case it would explain why none of the reactions worked as it is critical to the PCR. To ensure that the enzyme was functioning we tested it with another organism, our algae, *C. reinhardtii*, and its corresponding primer pair. This alga we knew had adequate template DNA and that the primers worked very well. If we saw any

amplification from this sample, we would confidently know that the polymerase enzyme was working. I performed a qPCR with this sample and I did see amplification in that DNA. I was then confident to conclude that our polymerase was not expired and working as it should.

The third potential reason for qPCR failure was that the wrong annealing temperatures were used. If the wrong annealing temperature is used, especially if the temperature is too high, then the primers cannot effectively bind to the DNA template and the reaction will not proceed. To assess this, we tested a few of our primer pairs on a temperature gradient PCR. A range of annealing temperatures was calculated using the NEB online tool with a primer concentration of 400 nM and the Phusion product group. This range was then used as a guide for creating the gradient in the PCR machine. Each column of the machine had a different temperature, ranging from 59 °C to 64 °C. After analyzing the PCR product through gel electrophoresis, I was able to determine that annealing temperatures of either 61 °C or 62 °C were ideal for these primers in a PCR. These temperatures were slightly lower than what we had expected. qPCR uses a slightly higher temperature range (65 °C to 68 °C) than a regular PCR, so we knew we needed to lower the annealing temperature we were using by a couple of degrees since we had been testing them on the higher end at 68 °C.

As the annealing temperature seemed to potentially be the most likely cause of our failed qPCR, we proceeded to test all 13 primer pairs again at a lower temperature of 67 °C. In this second round of qPCR testing, we ran all 13 primer pairs with the fungal samples that matched with the primer and had the highest DNA concentration to ensure the template DNA would not again be an issue. Amplification was seen in four of the

thirteen primer pairs. The other nine samples were tested again with other DNA samples if available. They still did not show amplification. This led us to conclude that those primer pairs that did not show amplification may not work for the fungal samples we have, and new ones may need to be designed. The primer pairs that did work as expected were rpB2_349F & 540R, rpB1_881F & 1059R, rpB1_350F & 564R, and rpB2_69F & 310R (See Table 1).

**Table 1: Fungal Primers and Species -** *This table shows the 13 universal fungal primers, which fungal species they matched with, their annealing temperatures, and whether or not we were able to get them to function properly at the end of our tests.*

| Primer Pair Names | Fungal Species Matches | Working? Y/N? | Annealing Temperatures for qPCR(°C) |
|---|---|---|---|
| rpB1_881F & rpB1_1059R | 6589 | Yes | 68 |
| rpB1_350F & rpB1_564R | ER2484 | Yes | 67 |
| rpB2_69F & rpB2_310R | AK0346 | Yes | 67 |
| rpB2_349F & rpB2_540R | 1608, AK0105, BC0189 | Yes | 67 |
| rpB1_135F & rpB1_255R | FL0176, FL0195 | No | 67 |
| rpB1_278F & rpB1_519R | 6252, MTL0066, JM0138 | No | 68 |
| rpB1_215F & rpB1_302R | 6158 | No | 67 |
| rpB1_109F & rpB1_345R | AZ0081 | No | 67 |
| rpB1_361F & rpB1_609R | MTL0044 | No | 66 |
| rpB2_51F & rpB2_201R | IL1387 | No | 67 |
| rpB2_265F & rpB2_484R | ER2574, MTL0044 | No | 67 |
| rpB2_244F & rpB2_379R | IL0271 | No | 67 |
| rpB2_334F & rpB2_416R | JO181, FL0195 | No | 67 |

# CHAPTER II

## *Preparation of qPCR Standards*

### *2.1 Creation of Standards*

In absolute quantification DNA copy numbers of unknown samples are estimated through the use of a standard curve of absolute, defined values (Plaffl et al. 2013). This method contrasts with relative quantification which compares the expression of a target gene in one organism to the expression of the same gene in another organism (Real Time PCR Handbook 2012). In a qPCR experiment using absolute quantification, producing reliable and accurate standards is an essential part of the experiment. Standards are a dilution series of a purified PCR product of the target gene from the same organism as your unknown DNA samples (Plaffl et al. 2013). Each set of standards is a small amount of target DNA of a known concentration (Plaffl et al. 2013). These standards are critical in absolute quantification methods and assessing the performance of the reaction through its efficiency (Plaffl et al. 2013).

In order to create standards, one must extract DNA from the sample, use PCR to increase the quantity of target DNA in the sample, purify or clean that PCR product, quantify the target DNA remaining, and then use that to perform the dilution series to create known concentrations of the purified PCR product. We would need standards for all of our species that we would perform qPCR with, so for any of our species that had working primer pairs, we created standards.

In order to perform any type of PCR, DNA must be extracted from the samples. Throughout the experiment, two methods were used to extract DNA: a phenol chloroform extraction and extraction using the NucleoSpin Soil DNA extraction kit. DNA was resuspended in TE buffer and frozen in a -80 ℃ freezer. DNA was extracted from algal samples alone, fungal samples alone, and from coculture samples as well. To ensure quality of each extraction, DNA samples were quantified to determine the concentration in ng/uL of DNA dissolved in the TE buffer. To create our standards, we used DNA extracted from the six fungi mentioned in Chapter I for which we had working primers (6589, ER2484, AK0346, 1608, AK0105, and BC0189) and for the algae, *C. reinhardtii*.

In order to increase the quantity of DNA found in our extracted samples we used PCR to amplify the target DNA. PCR allowed us to quickly increase the DNA concentration in our samples without having to extract more DNA and while being able to ensure that our DNA was all the same since it came from the exact same sample. The one drawback of amplifying our DNA through PCR is that the samples had other PCR reactants such as the polymerase enzyme and extra nucleotides or primers in the mixture as well. To remove these extra reactants, we purified/cleaned the PCR product to ensure we just had the DNA in our sample. We then performed another quantification to determine how much DNA was in each sample. The concentration of purified DNA is critical in creating standards of a known concentration.

When quantifying DNA, we obtained a DNA concentration in ng/uL of the sample. However, a calculation must be done to convert that mass concentration into a DNA copy number concentration. To do this we utilized the formula below:

$$Number\ of\ DNA\ copies/uL = \frac{ng\ of\ DNA \cdot 6.023\times10^{23}}{base\ pairs\ in\ the\ sequence \cdot 1\times10^{9} \cdot 650\ daltons} \quad \text{(Eq. 1)}$$

With Eq. 1, we were able to determine the approximate number of copies of our target DNA sequence we had in our samples. Our samples had a much higher concentration of DNA than was needed to create our first standard that would have a concentration of $10^9$ copies of DNA. We chose this as our starting concentration because a serial dilution could easily be performed to create a set of standards that would have concentrations within which our unknown samples would likely fall (Plaffl et. al 2013). It is extremely important that the starting concentration of standards is accurate or all the following standards will be incorrect as well. To create our starting $10^9$ standards, we calculated the volume of sample to add to molecular grade water. Then, once we created the first $10^9$ standard, we were able to create the others ($10^8$-$10^2$) through a simple 10-fold serial dilution.

### *2.2 Standard Testing*

To determine the quality of the standards we had produced, we had to test them using qPCR. A properly functioning set of standard dilutions should amplify at regular intervals and produce a melt curve that shows there is only one product formed from the reaction. A significant fluorescence increase above the baseline level for each standard dilution should occur at regular intervals depending on the dilution factor used to create the standards. Because we did a series of 10-fold dilutions, amplification should be observed approximately every 3.3 cycles of PCR. This interval can be determined using

the equation $y = 2^n$ with y being equal to the dilution factor and n being equal to the number of cycles observed between amplification of the dilutions. If amplification is not observed at these intervals it means that the dilutions may have not been successful and your standards do not contain the concentration of DNA they are thought to. This can also occur due to the degradation of the DNA in the sample after long periods of time or from thawing frozen standards. The known concentrations of your standards are critical to quantification of unknown samples, so if there is a problem with these, then the standards will likely have to be remade.

In some of our standards we had this amplification problem. In our standards for samples 6589 and 1608, we observed amplification every 3.3 cycles for the $10^9$- $10^4$ standards, but after that amplification was occurring at intervals of 4-5 cycles for our last two standards. We hypothesized that this was primarily due to improper dilution technique and not mixing up each previous dilution well enough while performing our serial dilutions. This may have led to our standards not having the exact concentrations we had intended them to. To correct for this error, we re-diluted the last 3 dilutions in our standard sets, ensuring that each dilution in the set was thoroughly mixed before pipetting to the next dilution.

One of the biggest advantages to using PCR or qPCR is that it is highly specific, but it is still possible to accidentally amplify a region other than our target gene or to produce primer dimers where the primers anneal to themselves and thus amplify themselves (Plaffl et al. 2013). We can verify that these errors have not occurred by melt-curve analysis. Because different DNA sequences contain a different nucleotide sequence, they contain a different content of adenine (A), thymine (T), guanine (G), and

cytosine (C) nucleotide bases. This base content will dictate at what temperature the hydrogen bonds between base pairs separate thus melting the DNA. Sequences with a high percentage of guanine and cytosine will melt at a higher temperature because guanine and cytosine base pairs have 3 hydrogen bonds whereas adenine and thymine base pairs only have 2, thus requiring a higher temperature and more energy input to break more bonds.

The melt curve can be produced by heating up and melting all qPCR products after all of the cycles have run. When the melting temperature ($T_m$) of a DNA strand is reached, the strand denatures and the fluorescence being emitted by the fluorescent molecules attached to that strand sharply decreases (Plaffl et al. 2013). This decrease in fluorescence is measured by the qPCR machine. The machine's software generates this melt curve by plotting a measurement of temperature vs fluorescence change. Analyzing the melt curve produced by the qPCR software (ProStudy) showed us at what temperature(s) our PCR product melted. An ideal melt curve will show all of the products melting at the same temperature and will produce a single, clean peak on the graph. This indicates that only one product, likely the desired product, was amplified. All of our standards showed one clear melt curve peak, so we could be confident that we were amplifying just our desired target and no primer dimers were being produced.

### *2.3 Standard Validation*

After using the standards a few times, some slight degradation was observed in the lower concentrations. This was likely due to the standards being frozen and thawed multiple times. If freezing and thawing our standards multiple times was causing a

significant amount of degradation, they would no longer be reliable in the qPCR experiments for quantification. Therefore, we needed to determine the optimal number of freeze-thaws that could be performed before they were not working effectively. To determine the ideal preparation and usage of the standards we conducted a simple experiment.

We applied three different freeze-thaw treatments to our standards to observe the effects of the repetitive freezing and thawing on the degradation of the standards. Our three treatments included: 1) A fresh set of never frozen standards, 2) A set of standards that had been frozen once and then thawed for use, and 3) A set of standards that had been frozen and thawed twice. If freezing and thawing were causing significant damage to our standards then we would see a drop off in the Cq values observed when run in a qPCR, particularly in the standards of lower concentration with multiple freeze-thaws.

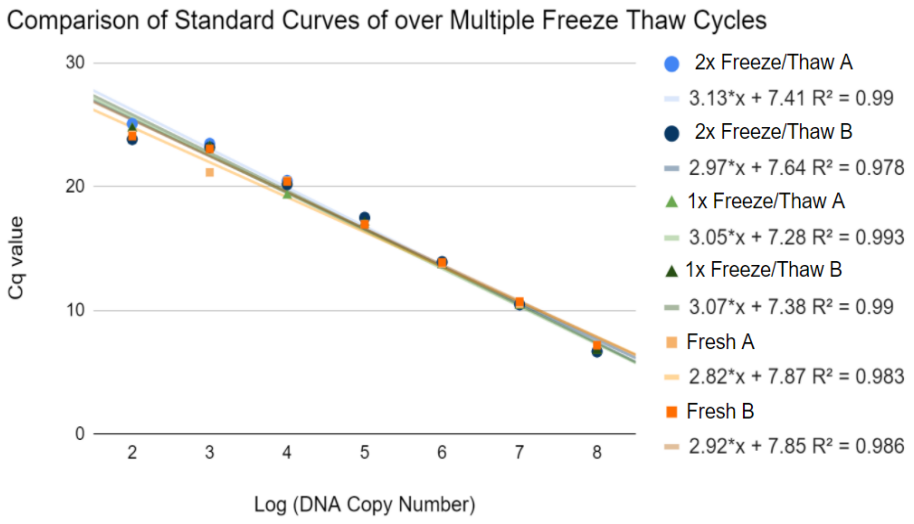


**Figure 1: Comparison of Standard Curves over Multiple Freeze Thaw Cycles** - ***The*** *above figure shows the comparison of what the standards for C. reinhardtii look like after multiple, single, and no freeze/thaw treatments. Each color/shape on the figure indicates a different treatment set. Each treatment was performed in duplicate to ensure quality of the results, these are indicated as A and B replicates on the figure.*

After the qPCR we observed a slight difference in the Cq values for all three treatments. This result was different than expected as there didn't seem to be any significant damage in the samples with multiple freeze/thaws. From this data we could conclude that the lower concentrations, especially the $10^2$ standards, were reducing the efficiency of the standards and were not going to be helpful in calculating quantities in the future. Although there was not a meaningful decrease in the performance of the standards with the 2x freeze/thaw treatment, we still didn't want to risk degradation if we refroze and thawed our samples too many times (more than once or twice). There was also no significant advantage to making fresh standards for each qPCR run, so we concluded that 1x freeze/thaw is optimal for making our standards. Thus, a set of standards should be made once in a large volume, aliquoted into smaller volumes, frozen, and each aliquot will be thawed and used once and then disposed to prevent any degrading over multiple freeze and thaw cycles. Moving forward, this will be the method used for generating and using our qPCR standards.

# CHAPTER III

## *Data Analysis and Quantification*

Results from a qPCR run are plotted on an amplification plot, which projects cycle number against the corrected fluorescence measured. A baseline fluorescence emitted by a secondary fluorophore (ROX dye) is detected by the machine and used to adjust the raw fluorescence of the target dye. This internal correction accounts for differences in raw fluorescence that may occur due to condensation or obstruction between the sample wells and the photodetector. It is assumed that the secondary fluorophore will maintain a constant fluorescence over the course of the qPCR run. In order to determine the point of amplification, a threshold is set at a fluorescence level significantly above the baseline level of fluorescence. This threshold level is automatically set by the ProStudy software accompanying the thermocycler machine. Where the amplification curves cross this threshold is what determines the Cq values to be used in quantification. The Cq value is the cycle number at which the fluorescence produced by the sample has crossed the threshold value (Real Time PCR Handbook 2012). The Cq value has an inverse relationship with the amount of DNA present in a sample (Plaffl et al. 2013), meaning the more DNA present at the start of a reaction, the lower the resulting Cq value will be because it will require fewer cycles of PCR to reach that threshold value. However, the Cq values alone do not provide enough information to determine an actual quantity of DNA in your sample, and absolute or relative

quantification must be performed. This study focuses on absolute quantification, which relies on the use of a standard curve with pre-quantified standards to determine the quantity of DNA present in unknown samples.

### *3.1 Quantification Methods*

There is currently some debate on what is the best method to use for the interpretation of the raw qPCR data (Bustin 2009). Throughout this project I focused on two main methods of absolute quantification: The Standard Curve method of quantification and the One Point Calibration (OPC) method. Both methods are adequate for quantifying DNA after a qPCR, but both methods also have their flaws. We performed an analysis of our data using both of the methods to determine their accuracy and ease of use for this project (See Fig. 3).

#### *3.1.1 The Standard Curve Method*

The Standard Curve method of quantification was the original method used in the first ever qPCR experiments and is what is still most widely used today. It is so common that this quantification method is often included in qPCR machine software. The Standard Curve method relies on the use of standard DNA samples of known concentrations to determine the concentration of the DNA in the unknown samples. This method uses a linear regression of the standards' concentrations vs Cq values to create a standard curve. It then places the unknown samples on that curve based on their Cq values to estimate the concentration of the unknown samples which can then be converted into a DNA copy number.

The standard curve method is easy to understand and can allow for quick and simple analysis of your data. This is the primary reason why it is preferred by many and is the more widely used quantification method. However, this method assumes that the efficiency of the reaction is equal and constant in both the standards and the unknown samples one is testing (Brankatschk et al. 2012). Unfortunately, it has been observed that the efficiency tends to vary from 1.8-2.0, which can leave a large room for error in quantification calculations (Ramakers et al. 2009). This difference of efficiency is more often seen in experiments where the sample DNA one is trying to quantify comes from another source than the standard's DNA, particularly when the sample comes from a co-culture and the standard from a singular, pure culture (Brankatschk et al. 2012).

### *3.1.2 The One Point Calibration Method*

To correct for differences in efficiency observed in the Standard Curve method, Brankatschk et al. (2012) proposed a novel method of quantification, the One Point Calibration (OPC) method. The OPC method takes into account the individual efficiency values of unknown samples to quantify those samples by comparing them to a single point on the standard curve. As previously discussed, an incorrect efficiency estimate can throw off quantification by a large factor, which led us to explore it further in our project. The OPC method uses the following equation in its quantification calculations:

$$N_0 \text{ sample} = N_0 \text{ standard} \times (E \text{ standard}^{Cq \text{ standard}} \div E \text{ sample}^{Cq \text{ sample}}) \qquad \text{(Eq. 2)}$$

A key part of using this method is determining the efficiency (E) values in Eq. 2 from the amplification plot of the qPCR run. To do this we utilized an *R* package called

"qpcR" (Speiss 2018). This program fit all of our samples to a sigmoidal curve to calculate their efficiency by comparing them to this ideal curve. This program produced an Excel file of data, including the newly determined efficiency values, that we could use for our OPC calculations. The largest limitation of the OPC method is the determination of the efficiency values. Because these efficiencies are the key motivation behind using this method, it can be expected that determining them with accuracy is critical.

### *3.2 Results and Conclusions*

qPCR was run and preliminary experimental results were obtained for one of our algae species, *C. reinhardtii* and our fungal species. These results were for qPCR tests of a single sample alone, not a co-culture. The species was tested to determine the concentration of "unknown" samples of a single species and the standards we made were also run for use in data analysis. These preliminary results will also allow us to determine which quantification method will be ideal for future qPCR experiments and if our methodology will work for future analysis of the fungal/algal mutualism assays.

#### *3.2.1 Cell Count Validation of qPCR Quantification in Chlamydomonas reinhardtii*

In order to verify that our qPCR experimental setup and methodology was sufficient and the quantification results were accurate, we performed a simple experiment. *C. reinhardtii* was selected as our first algae to test as it had the most effective primers and consistent and accurate standards. We first obtained a homogenous sample of our algae growing in liquid media and then, adding 10 μL of our sample to a hemocytometer, counted the number of algal cells in 1 square via light microscopy to

determine the approximate concentration of cells/ mL of our sample. We then resuspended our algal sample to a concentration of 7.5 x $10^6$ cells/mL. Serial dilutions were then performed to obtain 6 algal samples (Labeled as Chlamy Dilutions A-F) with decreasing cell concentrations. Calculations were performed to determine the approximate number of cells in each of these dilutions, which would be our "expected quantity" in our data analysis (See Fig.3).

The DNA was then extracted from these 6 samples to use in qPCR analysis. Although we do have an expected number of cells in these samples prior to extraction, and thus an expected number of DNA copies, these samples would serve as our "unknown samples" in our qPCR. qPCR was then run using our "unknown" dilution samples along with our standards we had created for *C. reinhardtii*. All samples were run in triplicate to ensure the reliability of the results of the experiment. Data was gathered from the run, a standard curve was constructed (See Fig. 2), and the "unknown" data were analyzed with both the Standard Curve method and the OPC method to quantify the unknown samples (See Fig. 3).

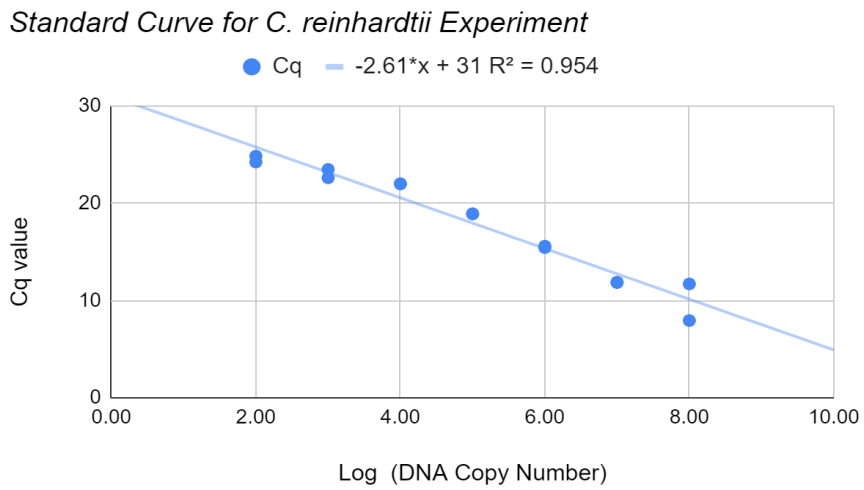


**Figure 2: Standard curve for *C. reinhardtii* Experiment-** *This figure shows the standard curve created for the C. reinhardtii experiment used in the absolute quantification of the unknown samples.*

The *C. reinhardtii* experiment showed some slight variations between the quantification methods (See Fig. 3). The Standard Curve method quantifications seemed to be much closer to the expected quantities as determined by the cell counts, but they were still on average higher. We can also see that our qPCR quantities determined by the Standard Curve method seemed to overestimate the quantity at our higher concentration dilutions (A, B, C and D) and underestimate the quantity at the lower concentration dilutions (E and F) (See Fig.3). This difference in counts may be due to the error described previously due to differences in efficiency of the qPCR of the standards versus the unknown samples. However, this difference may also be due to human error in counting the expected number of cells in the original sample, the small portion of the sample not being representative of the concentration of the whole sample, or due to the counting of dead cells in the sample. The OPC method provided us with copy numbers that widely varied from what was expected by the cell counts, especially in our lower

concentration dilutions. It is also important to note that the OPC method doesn't seem to represent our dilutions very well. The calculations show that the DNA copy number increased across a few of our dilutions (Dilutions A-C), which we know is incorrect. The calculations should indicate a steady decrease in DNA quantity across the dilutions.

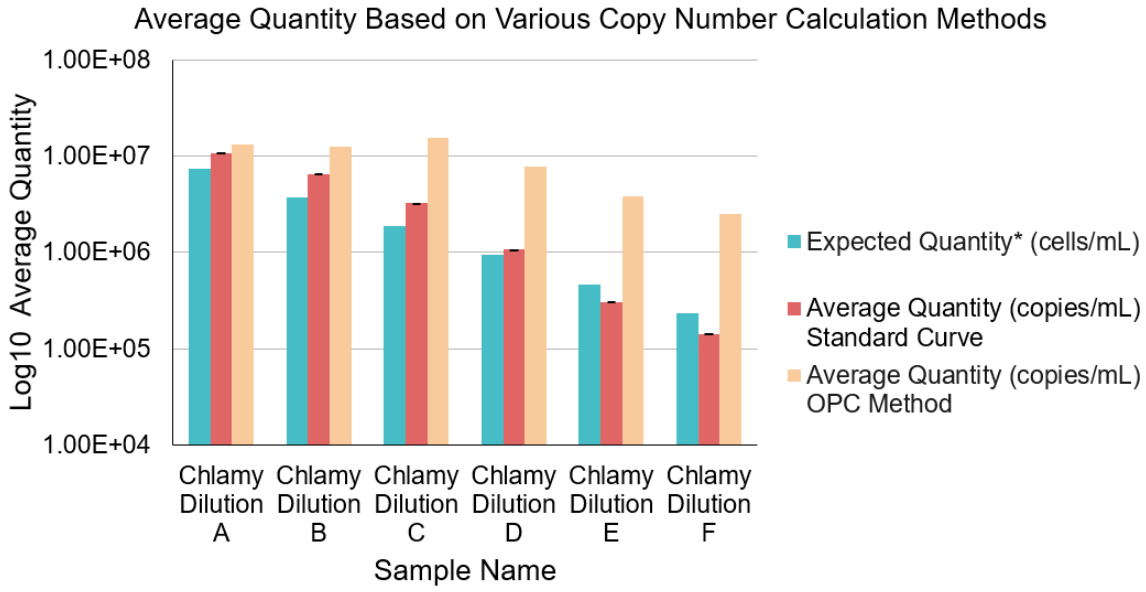


**Figure 3: Average Quantity Based on Various Copy Number Calculation Methods-** *This figure depicts the results we got when analyzing our preliminary results for our C. reinhardtii samples. The asterisk indicates that the expected value for Chlamy Dilution A was determined by cell counts of a small sample through microscopy and a calculation was performed to determine the approximate number of cells in the entire sample. From there, the other expected quantities were obtained by performing a serial dilution of the samples.*

Overall, from this data we can conclude that using the Standard Curve quantification method, included in the qPCR software, provided us with a copy number closest to what we expected, showing close to the expected 1:1 ratio of copies of the amplified sequence to the number of cells. We expected the OPC method to be more

accurate due to it accounting for differences in efficiency; however, other errors in calculating the efficiency value possibly caused by things such as difficulty fitting a sigmoidal curve or dropping off of the plateau, caused the OPC method to actually be a worse representation of our data. If the efficiency is able to be calculated correctly, the OPC method could possibly still be an effective method for quantification of our samples.

These data also show us that the effect of imperfect efficiency may not be as critical as previously thought as the Standard Curve method quantifications are not extremely far from what was expected and are likely to still be useful in making approximations of overall cell counts of a species. From this result we can conclude that the Standard Curve method is satisfactory for quantifying the samples, and is much more user-friendly in that there is less room for human error in calculations.

### *3.2.2 Nuclei Count Validation of qPCR Quantification in Fungi*

Similar to that of *C. reinhardtii*, a preliminary quantification of our fungal samples was also performed. This preliminary quantification served as a way to validate our methodology by comparing qPCR-based DNA copy numbers to fungal nuclei counts. We selected our fungal species from the same list of those we had functioning primers and standards for (ER2484, BC0189, AK0105, AK0346, 1608, and 6589). We obtained a plate of each fungal species that had the most growth over the plate and seemed to not have any contaminants. A sample of two medium-diameter “hole punches” were placed in a microcentrifuge tube containing 1mL of NanoPure water. Samples were then homogenized via sonication to ensure the fungus was well dispersed throughout the sample and there were no large clumps, this 1mL sample was then separated in half to

obtain two separate 500uL samples. One tube of the sample would be used to obtain approximate cell counts via nuclei staining and microscopy and the other to obtain DNA copy numbers via qPCR.

The fungal species used in this experiment tend to grow in clumps together and can create what appears to be one large cell with multiple nuclei. This can cause issues when counting cells underneath the microscope. To correct for this problem and have a cell quantity that could be directly compared to the DNA copy number, we decided to count the nuclei alone. We used a Hoechst 33342 fluorescent dye to stain the nuclei of the cells in our samples. This dye binds to the DNA in the sample to allow us to clearly visualize the nuclei of our cells for counting. A 5uL drop of each stained sample was put on three microscope slides per sample to create a total of 18 slides to image and quantify. Slides were then imaged using a light microscope emitting UV light to visualize our stained cells. Cells were counted from the images captured (See Fig. 4) and calculations were done to obtain an average number of cells per mL of each of our samples.

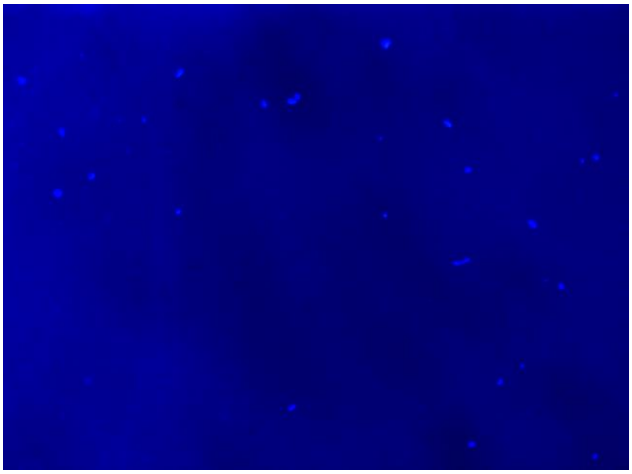

**Figure 4: Fungal Nuclei Staining -** *This image shows a small section (Photo 4/63) of one of the imaged slides of fungi ER2484 after staining of the fungal nuclei for cell quantification.*

DNA was extracted from the samples set aside for qPCR using the NucleoSpin Soil DNA extraction kit. This DNA was then used as our unknown samples in our qPCR assays. The unknown samples and their standards, for each species, were run in triplicate to ensure precision of our qPCR results. We then performed an absolute quantification using the Standard Curve method to determine the copy number of the DNA in our samples. The Standard Curve method was chosen for quantification of our fungal samples due to its showing higher accuracy with *C. reinhardtii* previously.

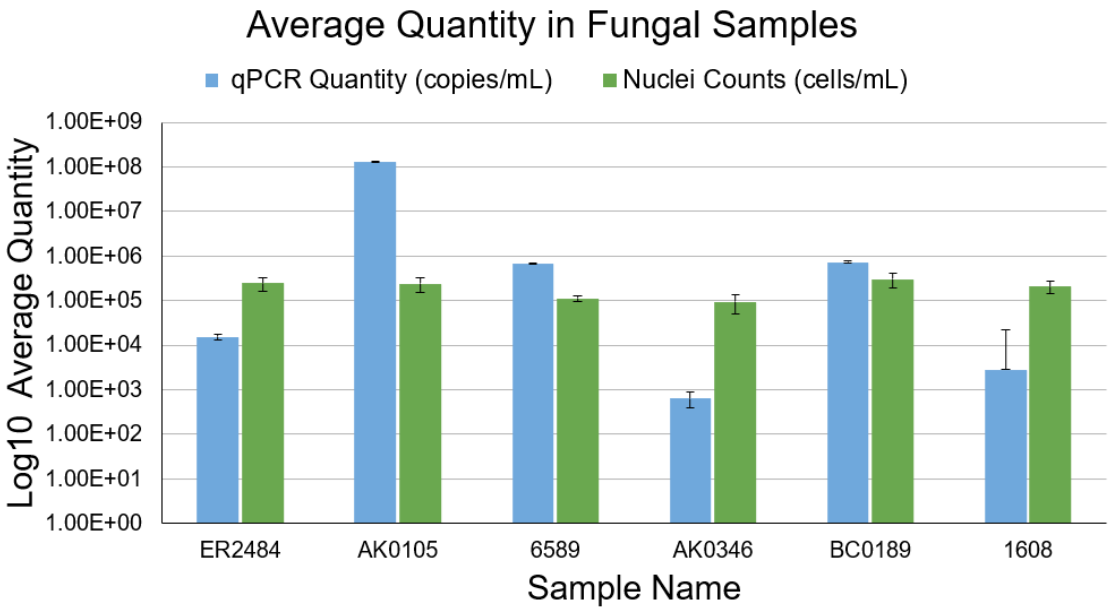


**Figure 5: Average Quantity in Fungal Samples-** *This figure depicts the results of our preliminary qPCR tests with our fungal species. Each point on the graph indicates a fungal species that was tested. Our data for both cell counts and qPCR copies/mL were averaged to achieve these data points seen to account for our technical replicates.*

From this data (See Fig.5) we can determine that while our qPCR counts may be fairly accurate, they do not correlate very strongly with the nuclei counts we obtained. We do not see very much error between the qPCR data within the replicates of each sample, which indicates that there is little variability between our technical replicates in

our qPCR. However, there may still be some error in the quantification of this qPCR data due to any imperfections in creating our standard curve. This error would not be indicated by our error bars in the figure as all of our technical replicates would be off by the same or a similar amount since they used the same standards for calculations. This error is likely from any inaccuracy in creating our standards, not from creating the standard curve itself, as the $R^2$ value of our standard curves were all calculated to be 0.98 and above, indicating that the linear regression equation used in our calculations is accurate.

Unfortunately, we do see a large margin of error particularly when it comes to the nuclei counts. This is likely due to error in our methodology of staining and counting nuclei. We only counted 15 uL and it is possible that this small volume of sample was not representative of the sample as a whole. There is also the possibility that there was human error in effectively counting the nuclei in the slides. Despite the samples having vastly different quantities of DNA, according to the qPCR data, they seem to all have close to the same number of nuclei. This is a good indicator that there is something wrong with the method used to obtain these cell counts as well. To improve this method, we should image a larger sample of our stained nuclei in order to get a more accurate count of the approximate cell count in our total sample.

Throughout this study we have determined and established an effective method for quantifying the number of cells in both our fungal and algal cultures. These experiments have set up a system that should be effective in use in our single cultures and in our co-cultures as well. In the co-cultures, unknown DNA samples should be run with both sets of primers separately and then quantified to determine the DNA copy numbers of each species present in the sample. Moving forward, following this study, these

methods should be able to be used to analyze these co-cultures to get a clearer understanding of the approximate cell counts of each species given the copy numbers are a sufficient proxy for the cell counts. Understanding the approximate numbers of cells of each species in these co-cultures will allow us to better understand the true nature of this suspected mutualism between the algal and fungal species.

## BIBLIOGRAPHY


Akcay, Erol, et al. *Mutualism*. Edited by Judith L. Bronstein, 1st ed., Oxford University Press, 2015.

Arya, Manit, et al. "Basic Principles of Real-Time Quantitative PCR." *Expert Review of Molecular Diagnostics*, vol. 5, no. 2, Mar. 2005, pp. 209–19. *DOI.org (Crossref)*, doi:10.1586/14737159.5.2.209.

Bhattacharya, Debashish, et al. "Molecular Cloning and Evolutionary Analysis of the Calcium-Modulated Contractile Protein, Centrin, in Green Algae and Land Plants." *Plant Molecular Biology*, vol. 23, no. 6, Dec. 1993, pp. 1243–54. *DOI.org (Crossref)*, doi:10.1007/BF00042357.

Brankatschk, Robert, et al. "Simple Absolute Quantification Method Correcting for Quantitative PCR Efficiency Variations for Microbial Community Samples." *Applied and Environmental Microbiology*, vol. 78, no. 12, American Society for Microbiology, June 2012, pp. 4481–89. *aem.asm.org*, doi:10.1128/AEM.07878-11.

Bustin, Stephen A., et al. "The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments." *Clinical Chemistry*, vol. 55, no. 4, Apr. 2009, pp. 611–22. *DOI.org (Crossref)*, doi:10.1373/clinchem.2008.112797.

DeBary, Anton "Ueber Symbiose" [On Symbiosis]. *Daily Journal for the Conference of German Scientists and Physicians.* 14 Sept. 1878, doi:10.1890/0012-9623-96.1.80

Dieffenbach, C. W., et al. "General Concepts for PCR Primer Design." *PCR Methods Appl*, 1993, pp. 30–37.

Douglas, Angela E. The Symbiotic Habit. Princeton University Press, 2010.

Dubilier, N. “"Til Death Do Us Part”: Coming to Terms with Symbiotic Relationships.” *Nature Reviews: Microbiology*, vol. 6, 2008.

Green, Michael R., and Joseph Sambrook. “Polymerase Chain Reaction.” *Cold Spring Harbor Protocols*, vol. 2019, no. 6, Cold Spring Harbor Laboratory Press, June 2019, p. pdb.top095109. *cshprotocols.cshlp.org*, doi:10.1101/pdb.top095109.

Heid, C. A., et al. “Real Time Quantitative PCR.” *Genome Research*, vol. 6, no. 10, Oct. 1996, pp. 986–94. *genome.cshlp.org*, doi:10.1101/gr.6.10.986.

Higuchi, Russell, et al. “Kinetic PCR Analysis: Real-Time Monitoring of DNA Amplification Reactions.” *Bio/Technology*, vol. 11, no. 9, 9, Nature Publishing Group, Sept. 1993, pp. 1026–30. *www.nature.com*, doi:10.1038/nbt0993-1026.

“History of PCR Discoveries - PCR Technologies Guide.” *Sigma-Aldrich*, 15 June 2020.

Hom, Erik F. Y., and Andrew W. Murray. “Niche Engineering Demonstrates a Latent Capacity for Fungal-Algal Mutualism.” *Science (New York, N.Y.)*, vol. 345, no. 6192, July 2014, pp. 94–98. *PubMed Central*, doi:10.1126/science.1253320.

Jiao, Chen, et al. “The Penium Margaritaceum Genome: Hallmarks of the Origins of Land Plants.” *Cell*, vol. 181, no. 5, May 2020, pp. 1097-1111.e12. *PubMed*, doi:10.1016/j.cell.2020.04.019.

Karol, Kenneth G. “The Closest Living Relatives of Land Plants.” Science (New York, N.Y.), vol. 294, Dec. 2001, doi:10.1126/science.1065156.

*Kinetic PCR Analysis: Real-Time Monitoring of DNA Amplification Reactions*. 12 June 2020, http://gmo-qpcr-analysis.info/higuchi-1993.pdf.

Kriventseva EK et al, “OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs” NAR, Nov 2018, doi:10.1093/nar/gky1053

Liu, Y.J., et al. 1999. Phylogenetic relationships among ascomycetes: evidence from an RNA polymerase II subunit. Mol. Biol. Evol. 16(12): 1799-1808

Matheny, P.B., et al. 2002. Using RPB1 sequences to improve phylogenetic inference among mushrooms (Inocybe, Agaricales). Am. J. Bot. 89: 688-698.

Mullis, K., et al. “Specific Enzymatic Amplification of DNA In Vitro: The Polymerase Chain Reaction.” *Cold Spring Harbor Symposia on Quantitative Biology*, vol. 51, Cold Spring Harbor Laboratory Press, Jan. 1986, pp. 263–73. *symposium.cshlp.org*, doi:10.1101/SQB.1986.051.01.032.

Paracer, Surindar, and Vernon Ahmadjian. *Symbioses.* Oxford University Press, 2000.

Pfaffl, Michael W., et al. *Good Practice Guide for the Application of Quantitative PCR (QPCR)*. National Measurement System, 2013.

*Polymerase Chain Reaction (PCR)*. NCBI. 15 June 2020,

*ProStudy Software Operation Manual*. Bibby Scientific. 2014

*QPCR Assay Design and Optimization | LSR | Bio-Rad*. 14 Sept. 2020,

Ramakers, Christian, et al. “Assumption-Free Analysis of Quantitative Real-Time Polymerase Chain Reaction (PCR) Data.” *Neuroscience Letters*, vol. 339, no. 1, Mar. 2003, pp. 62–66. *DOI.org (Crossref)*, doi:10.1016/S0304-3940(02)01423-4.

*Real Time PCR Handbook*. Life Technologies Corporation, 2012.

Ruijter, J. M., et al. “Amplification Efficiency: Linking Baseline and Bias in the Analysis of Quantitative PCR Data.” *Nucleic Acids Research*, vol. 37, no. 6, Apr. 2009, p. e45. *PubMed Central*, doi:10.1093/nar/gkp045.

Singh, Vinay, and Anil Kumar. “PCR Primer Design.” *Molecular Biology Today*, 2001,

Spiess, Andrej-Nikolai. Modelling and Analysis of Real-Time PCR Data. 29 May 2018,