Single cell RT-qPCR based ocean environmental sensing device development

by

Xu Shi

A Dissertation Presented in Partial Fulfillment
of the Requirements for the degree
Doctor of Philosophy

Approved November 2013 by the
Graduate Supervisory Committee:

Deirdre R. Meldrum, Chair
Weiwen Zhang
Shih-hui Chao
Paul Westerhoff

ARIZONA STATE UNIVERSITY

December 2013

ABSTRACT

This thesis research focuses on developing a single-cell gene expression analysis method for marine diatom *Thalassiosira pseudonana* and constructing a chip level tool to realize the single cell RT-qPCR analysis. This chip will serve as a conceptual foundation for future deployable ocean monitoring systems. *T. pseudonana*, which is a common surface water microorganism, was detected in the deep ocean as confirmed by phylogenetic and microbial community functional studies. Six-fold copy number differences between 23S rRNA and 23S rDNA were observed by RT-qPCR, demonstrating the moderate functional activity of detected photosynthetic microbes in the deep ocean including *T. pseudonana*. Because of the ubiquity of *T. pseudonana*, it is a good candidate for an early warning system for ocean environmental perturbation monitoring. This early warning system will depend on identifying outlier gene expression at the single-cell level. An early warning system based on single-cell analysis is expected to detect environmental perturbations earlier than population level analysis which can only be observed after a whole community has reacted. Preliminary work using tube-based, two-step RT-qPCR revealed for the first time, gene expression heterogeneity of *T. pseudonana* under different nutrient conditions. Heterogeneity was revealed by different gene expression activity for individual cells under the same conditions. This single cell analysis showed a skewed, lognormal distribution and helped to find outlier cells. The results indicate that the geometric average becomes more important and representative of the whole population than the arithmetic average. This is in contrast with population level analysis which is limited to arithmetic averages only and highlights the value of single cell analysis. In order to develop a deployable sensor in the ocean, a chip level device

was constructed. The chip contains surface-adhering droplets, defined by hydrophilic patterning, that serve as real-time PCR reaction chambers when they are immersed in oil. The chip had demonstrated sensitivities at the single cell level for both DNA and RNA. The successful rate of these chip-based reactions was around 85%. The sensitivity of the chip was equivalent to published microfluidic devices with complicated designs and protocols, but the production process of the chip was simple and the materials were all easily accessible in conventional environmental and/or biology laboratories. On-chip tests provided heterogeneity information about the whole population and were validated by comparing with conventional tube based methods and by *p*-values analysis. The power of chip-based single-cell analyses were mainly between 65-90% which were acceptable and can be further increased by higher throughput devices. With this chip and single-cell analysis approaches, a new paradigm for robust early warning systems of ocean environmental perturbation is possible.

ACKNOWLEDGEMENT

Completing my PhD degree is probably the most challenging activity of my first 30 years of life. During this doctoral journey, a lot of moments no matter they are good or bad have been shared with many people. It has been a great privilege to spend 5 years in the Center of Biosignatures discovery automation at Arizona State University, and its members will always remain dear to me.

My first debt of gratitude must go to my advisors, Dr. Deirdre Meldrum, Dr. Weiwen Zhang and Dr. Shih-hui Chao. They patiently provided the vision, encouragement and advice necessary for me to proceed through the doctoral program and complete my dissertation. Their unflagging encouragement and serving is the best role models to me as a junior member of academia.

Special thanks to my committee member, Dr. Paul Westerhoff, and my special mentor, Dr. Weimin Gao, for their support, guidance and helpful suggestions. Their guidance has served me well and I owe them my heartfelt appreciation. They also encouraged and provided me with support to pursue independent work.

My wife, Shuang Xu, whose love, encouragement and support allowed me to finish this journey. She is the best person out there for me and marrying her is the best achievement I made in the past 5 years. She already has my heart so I will just give her a heartfelt “thanks”. I wish to thank my parents. Their love provided my inspiration and was my driving force. I owe them everything and wish I could show them just how much I love and appreciate them. I also want to thank to my in-laws for their unconditional support. Other important persons I would never forget are my friends in China and US. Thanks for their support and friendship along this journey.

Finally, I would like to thank Arizona State University NEPTUNE fund for their generous support. And all the staff and students at CBDA, I'm grateful to join and be a part of the lab. Thank you for treating me as a friend and helping me to achieve my project. Special thanks to my colleagues, Jieying, Bo, Jia for discussion and consolation. Thanks to Dr. Yanqing Tian and Dr. Fengyu Su, Dr. Qian Mei, Dr. Liqiang Zhang and Dr. Hong Wang and many other friends that have helped me over the years.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

1. OBJECTIVES AND CONTRIBUTION

1.1 Objectives

Ocean environments may be monitored by analyzing a well-suited native microorganism at the single-cell level. A biological tool for environmental monitoring will be constructed and gene expression at the single cell level will be adopted as the detection method. The objectives of the project include: *i*) Identifying a widely spread ocean microorganism to be monitored in both surface and deep ocean waters; thereby, the application will not be constrained to locations where the target microorganism exists. The target microorganism will be determined by clone library and phylogenetic analysis of deep ocean water samples since surface water samples have been intensively studied by other researchers and enough information about the species information has been collected; *ii*) For the first time, tube based two-step RT-qPCR analysis for single cell gene expression will be performed for the target species without preamplification of the single-cell mRNA. This method will help to illuminate the heterogeneity of gene expression and at the same time provide information about stress responses for different nutrient-limited conditions; *iii*) A chip level device will be developed to realize one-step RT-qPCR at the single-cell level for the target species. The chip should be robust with no off-chip operations. Further, the chip must be produced with a simple procedure and materials that are readily accessible in conventional laboratories. With the help of this biological tool, future deployable sensors could be built. Ideally, the chip needs to be compatible with commercially available real-time PCR stations with minor modification/optimization so that more laboratories can run single-cell analyses using this tool.

The ultimate goal of this project can be described by Figure 1. Heterogeneous environmental samples will be collected and properly distributed on a single cell gene expression analysis chip. Single-cell RT-qPCR will be performed on chip and the results will be sent back to the laboratory and compared with baseline information to identify possible environmental perturbation. This thesis will focus on the later portion of the process which mainly focuses on biological tool development and concept validation.

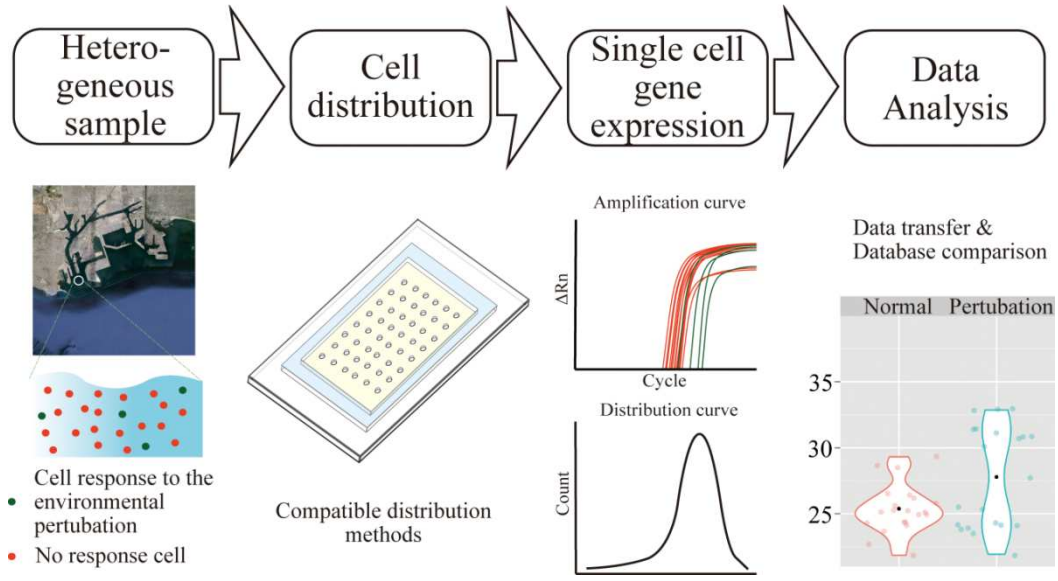

Figure 1. Schematic description of analysis process.

1.2. Scientific contributions

Based on the results of my project, all objectives have been successfully achieved. The scientific contributions of my work include: firstly, phylogenetic and gene expression analysis of photosynthetic cyanobacteria and diatoms in deep ocean samples. Marine surface water photosynthesis microorganisms were observed in the deep ocean samples with moderate activities which were confirmed by RT-qPCR results. The results suggested that our previous understanding of the species distribution in the ocean may

not be accurate. Secondly, tube based two-step RT-qPCR analysis at the single cell level has been applied to model eukaryotic phytoplankton, *Thalassiosira pseudonana* CCMP1335, for the first time. Lognormal distribution which indicated that the geometric average becomes more representative of the whole population than arithmetic average was observed for single cell gene expression. The results confirmed that population level analysis will provide biased information of a population (Lidstrom and Meldrum 2003; Strovas and Lidstrom 2009). At the same time, the results showed *T. pseudonana* stress response pattern to no iron, no nitrogen and no phosphate conditions. This information validates the concept and will be helpful for future sensor system construction. Thirdly, a simple chip level device that can perform robust single cell gene expression analysis by using one-step RT-qPCR has been constructed. The chip can provide new information about environmental stress responses of microbes but can also be used to monitor the effects of unknown environmental perturbations on native ocean species. The significances of this chip are: a) supplying a streamlined protocol which realizes direct cell-to-data processing without cell lysing and nucleic acid purification; b) no special requirement or expertise is required to construct/use this chip, so this device can be used as long as a compatible real time PCR machine is available. Compared with other methods/devices, this device is extremely simple to construct and all the required instruments and materials are easily accessible in conventional laboratories; c) the chip has the same statistical power as other single cell analysis devices, which have higher sensitivity than conventional (population-based) methods. The accuracy of the chip at the single cell level has been proven by single cell results from a tube-based method; d) the

device itself is versatile and can be compatible with different upstream or downstream operations such as cell loading via micromanipulator or dilution-to-extinction, and genetic and/or transcription analysis. This chip not only provides a solid background for future deployable sensors but also provides opportunities for conventional biological and environmental laboratories to perform single cell analysis.

With the results and technologies I learned in the past five years, I have published four peer-reviewed journal papers and one MicroTAS conference paper. One review manuscript is under review and two more manuscripts are under preparation.

1. Gao, Weimin; Shi, Xu; Wu, Jieying; Jin, Yuguang; Zhang, Weiwen; Meldrum, Deirdre R., 2011. Phylogenetic and Gene Expression Analysis of Cyanobacteria and Diatoms in the Twilight Waters of the Temperate Northeast Pacific Ocean. *Microb Ecol* 62(4), 765-775.
2. Shi, Xu; Lin, Liang-I; Chen, Szu-yu; Chao, Shih-hui; Zhang, Weiwen; Meldrum, Deirdre R., 2011. Real-time PCR of single bacterial cells on an array of adhering droplets. *Lab Chip* 11(13), 2276-2281.
3. Shi, Xu; Lin, Liang-I; Gao, Weimin; Chao, Shih-hui; Zhang, Weiwen; Meldrum, Deirdre R., 2011. Single-cell Real-time PCR: direct process from cells to data. *15th International Conference on Miniaturized Systems for Chemistry and Life Sciences*, 362-354.
4. Shi, Xu; Gao, Weimin; Chao, Shih-hui; Zhang, Weiwen; Meldrum, Deirdre R., 2013. Monitoring the Single-Cell Stress Response of the Diatom *Thalassiosira pseudonana* by Quantitative Real-Time Reverse Transcription-PCR. *Applied Environmental Microbiology*, 79 (6), 1850-1858.

5. Shi, Xu; Gao, Weimin; Wang, Jiangxin; Chao, Shih-hui; Zhang, Weiwen, Meldrum, Deirdre R., 2013. Measuring gene expression in single bacterial cells: Recent Advances in methods and micro-devices. *Critical review in biotechnology* (under review).
6. Wang, Jiangxin; Shi, Xu; Johnson, Roger H.; Kelbauskas Laimonas; Weiwen Zhang and Meldrum Deirdre R., Single-cell analysis reveals differential hypoxia response in two human Barrett's esophageal cell lines, *PloS one* 8 (10), e75365.
7. Shi, Xu; Gao, Weimin; Chao, Shih-hui; Zhang, Weiwen; Meldrum, Deirdre R., 2013. Novel Single-Cell Droplet Chip to monitor gene expression under iron limitation condition, in preparation.
8. Chun-Hong Chen; Shi, Xu; Gao, Weimin; Chao, Shih-hui; Meldrum, Deirdre R., 2013. Parallel RNA extraction using magnetic beads and a droplet array, in preparation.

2. INTRODUCTION

Ocean environments provide about 50% of the global primary production (Field et al. 1998), but we know little about it even though we have already explored outer space (Edward F. DeLong et al. 2006). Over the past 2 decades, due to the application of more advanced technologies and tools, such as phylogenetic identification (Pace 1997) and metagenomics analysis (Tyson et al. 2004; Hallam et al. 2006; E. E. Allen et al. 2007; Edward F. DeLong et al. 2006), researchers attained more and more data which enlarged our understanding of the importance and functions of planktonic microorganisms. Meanwhile, due to fossil-fuel combustion (C. Le Quéré et al. 2009), human activities induced hydrodynamic currents (Q. Wang et al. 2004), fertilizer usage (Galloway et al. 2004), industrial activity, (Doney 2010) and so on, the coastal and open-ocean environment have been negatively altered. All these perturbations may impact the normal function of ocean microbes. An efficient way to measure this perturbation at an early stage is urgently required.

Using gene expression tools to measure the environment perturbation will propel the environmental monitoring field forward quickly. Environmental variations would first change the gene expression of organisms, then may cause alteration at the community level (Edward F. DeLong 2009). Normally in the field of microbiology, microbiologists believe that microbial cells growing under the same conditions are a uniform population (Brehm-Stecher and Johnson 2004). However, more recent evidence suggests that even isogenic cells exhibit notable diversity that is an order of magnitude greater than previously thought (Kelly and Rahn 1932; Maloney and Rotman 1973; Siegele and Hu 1997; Lidstrom and Meldrum 2003; Kuang, Biran, and Walt 2004; Becskei, Kaufmann,

and van Oudenaarden 2005; Colman-Lerner et al. 2005; Golding et al. 2005; Le et al. 2005; Pedraza and van Oudenaarden 2005; Rosenfeld et al. 2005; Strovas et al. 2007; Strovas and Lidstrom 2009). Therefore, a surge of researchers have focused on single cell analysis (Walling and Shepard 2011). An unprecedented increase of knowledge about single cells has already altered people's point of view when facing microbiology related issues.

Another important reason to pursue single cell analysis stems from the fact that the majority (>99%) of environmental microbial species cannot be cultured under laboratory conditions (Rajilić-Stojanović, Smidt, and De Vos 2007; S. Giovannoni and Stingl 2007). Therefore, they are not accessible to conventional cultured based gene expression analysis methods.

Single cell level analysis requires higher sensitivity and much more careful sample preparation which is more difficult to perform than population level analysis. Nevertheless, it provides information that population level analysis cannot provide. For example, gene expression patterns among a population (Shi et al. 2013; Bengtsson et al. 2005) can help to identify rare gene expression. In order to monitor the stress conditions that exist in the environment, one microbe that can be widely found and has large representation should be chosen as a target. Based on these considerations, diatom, which is a major group of unicellular phytoplankton (Falkowski et al. 2004; Thamatrakoln et al. 2012) is selected as the target microbe. It has been reported that diatoms contribute up to 40% of the primary productivity in the ocean (Maheswari et al. 2010; Nelson et al. 1995). Using single diatom cells, which are natural inhabitants of the ocean, for monitoring is in

contrast to other approaches which require introducing engineered foreign species to achieve a similar objective (Ripp et al. 2000).

Although single cell level analysis can provide more information than population level analysis, there are still some technical hurdles, such as how to isolate single cells in an effective and efficient way. With the help of advances in microfluidics, single cell level analyses have become more accessible (D. Wang and Bodovitz 2010). Microfluidics technology is especially advantageous to the single cell level analyses for the following reasons: *i*) Individual cells can be precisely trapped, moved, and distributed individually in microscale channels, minimizing contamination (Dorfman et al. 2005) and at the same time decreasing the consumption of chemicals and enzymes (Zare and Kim 2010); *ii*) Isolated individual cells can be easily monitored in microchambers. When a cell is lysed in a sealed microchamber, dilution of the cellular contents is minimized thus increasing the sensitivity of a downstream nucleic acid or protein analysis (Sims and Allbritton 2007); *iii*) Highly parallel, fully automated multi-step operations can be implemented for high-throughput analyses resulting in significant time and cost savings due to fast and highly efficient sample processing. Next is a summary of some single cell isolation technologies and downstream gene expression analysis methods.

2.1. Single cell isolation

Although manipulating single eukaryotic cells has become more and more common for single cell analysis, manipulating small cells like diatoms which are about 5 μm in diameter is significantly more challenging due to the facts that the total volume of one cell is 100−1000 times smaller than that of a typical eukaryotic cell (1-10 fL versus 1 pL)

and hence contains a fewer amounts of analytes. Several methods that can be used to manipulate small cells are summarized below.

2.1.1. Dilution-to-extinction

The conventional dilution-to-extinction method utilizes serial dilution to isolate single cells into test tubes or wells on microtiter plates (Button et al. 1993; Schut et al. 1993; Rappé et al. 2002). Microfluidic devices can apply the same principle to load single cells into microscale reaction chambers (Boedicker et al. 2008). Cell occupancy of the wells follows the *Poisson* distribution and can be manipulated by controlling cell concentration in the bulk media before loading. Because the microchamber volumes are several orders of magnitude smaller than those of conventional analysis vials, the required initial concentration is close to typical cell culture's concentration thus sample dilution is accordingly minimized. For example, if one loads single cells in an array of 1-picoliter chambers, the resulting concentration in the microwells is on the order of 1 cell/pL or 10^9 cell/mL, which is within the range of typical concentrations for bulk cell cultures (Sezonov, Joseleau-Petit, and D'Ari 2007).

Two microfluidic applications have utilized the dilution-to-extinction approach to isolate single cells, either by seeding cells in microfabricated chambers or encapsulating cells in emulsion. An example for the former application was a device developed by Ottesen *et al.* (2006) who isolated bacterial cells randomly from a complex environmental sample and then performed digital PCR to identify new species. They randomly seeded cells from a diluted environmental sample on their device and obtained single-cell occupancy in ~28% of the reaction chambers, while the rest of the chambers contained either multiple cells (6%) or were empty. The other approach is based on the

encapsulation of individual cells in aqueous droplets (Shim et al. 2009; Eun et al. 2011; Guo et al. 2012). Eun *et al.* (2011) used a microfluidic flow-focusing nozzle to generate *Escherichia coli*-containing agarose microdroplets. After the agarose microdroplets solidified, *E. coli* cells were encapsulated in agarose microparticles for downstream incubation and analysis. Zeng *et al.* (2010) randomly seeded *E. coli* cells into droplets containing primer-adhered microspheres and real-time PCR reagents. Lin *et al.* (2009) introduced a new method by generating stationary droplets as reaction chambers. They loaded a diluted suspension of *E. coli* onto an array of oil-covered surface-adhering droplets that were spatially confined by oil through hydrophilic/hydrophobic patterns on the substrate. The number of randomly seeded *E. coli* cells in droplets followed the *Poisson* distribution.

This dilution-to-extinction method does not require complicated single-cell manipulation technologies or devices. As long as the cell concentration in the bulk medium is properly diluted, fast and relatively easy seeding of single cells in microfabricated chambers or droplets can be achieved in a high throughput. Due to its simplicity, this technology has received increasing attention. The major drawbacks of this method are the random nature of the cell occupancy and the low efficiency of obtaining single cell occupancy while reducing the number of wells containing multiple and zero cells. Large numbers of empty compartments result in a waste of chemical reagents, reduced overall throughput, and the need to determine the number of cells in each well to discern reaction chambers containing single cell, multiple cells or empty.

2.1.2. Cell trapping

In contrast to the dilution-to-extinction method, single cell trapping is a deterministic method to isolate single cells. Multiple traps can be implemented in a device, facilitating parallel measurements at a given time. Three types of single-cell trapping methods have been used to isolate small single cells from populations: mechanical, hydrodynamic, and dielectrophoretic. Following is a discussion of the details of these trapping methods.

Mechanical trap: Mechanical cell trapping is achieved by physical obstacles, barriers or side channels/chambers to hold or catch individual cells flowing through microfluidic channels. Microscale U-shaped barriers (D. Di Carlo, Wu, and Lee 2006) have been applied to trap mammalian cells. However, these barriers are inefficient for trapping cells like diatoms due to their small dimensions. Huang *et al.* (2007) used a complicated microfluidic network to trap *Synechococcus* PCC 7942 cells between pneumatic valves and observed significant cell-to-cell heterogeneity in populations under nitrogen-depleted growth condition. Furutani *et al.* (2010) isolated single *Salmonella enterica* cells using an array of microchambers distributed along microchannels. Utilizing the amplification of the DNA of the *invA* gene obtained from single cells trapped in the microchambers, a detection level of <200 cells μL^{-1} of *S. enterica* were achieved with the device.

Hydrodynamic trap: Hydrodynamic trapping is a non-contact cell trapping method. It relies on flow stagnation or microeddies (Lutz, Chen, and Schwartz 2006) to capture cells in flow fields. Compared to eukaryotic cells, applying hydrodynamic traps to small cells poses significant challenges since hydrodynamic forces are typically proportional to the surface areas of cells. However, Tanyeri *et al.* (2010) have

demonstrated hydrodynamic trapping of 100 nm particles which is even smaller than normal bacterial cells. They used hydrodynamic traps to achieve high accuracy of trapping and manipulation of single bacterial cells in a microfluidic device. To trap single cells, their device produced a flow stagnation point in the center of two perpendicularly crossed channels. Indispensable of precise control over the flow to create flow stagnation and eddies is the major constraint of this method. The feedback-based flow control may alleviate this problem albeit at the cost of increased complexity of the system.

Dielectrophoretic trap: A dielectric particle experiences the dielectrophoretic (DEP) force when it is exposed to a non-uniform electric alternative-current (AC) field. Applying DEP forces on small cells can be traced back to the 1980s (Pohl, Kaler, and Pollock 1981). Peitz and Leeuwen first used the DEP force to trap single bacterial cells (Peitz and van Leeuwen 2010). They reported using DEP traps between parallel 10-µm electrodes to capture living *E. coli* K12 cells in a microfluidic channel. Arumugam *et al.* (2007) demonstrated the generation of DEP traps using vertically aligned carbon nanofibers as nanoelectrodes. These nanoscale electrodes generated large DEP forces in a small region, ideal for trapping small single cells. They successfully demonstrated a cheap and convenient way to produce DEP trap arrays with high-throughput. Although DEP can reliably and precisely trap single cells, it requires a tight integration of micro/nanoscale electrodes and driving circuits which increases the complexity of this method.

2.1.3. Micromanipulation

Micromanipulation is a precise method to isolate and manipulate single cells with a relatively low throughput, typically one cell at a time. There have been two types of

micromanipulation devices for cells: mechanical and optical micromanipulation (also known as optical tweezers). Micromanipulation has been applied to single cells since the 1960s (Nossal et al. 1964; Wood 1967). In mechanical micromanipulation, single cells are individually captured from a population and transferred using a micropipette (Anis et al. 2011; Anis, Holl, and Meldrum 2010; Ashida et al. 2010; Gao, Zhang, and Meldrum 2011; Roeder, Wagner, and Rossmanith 2010; Shi et al. 2011; Teramoto et al. 2010; Tsang et al. 2006). The isolated single cells can be subsequently used for different applications such as cultivation or gene expression analysis. In optical micromanipulation (Mirsaidov et al. 2008; H. Zhang and Liu 2008; Altindal, Chattopadhyay, and Wu 2011; Carmon and Feingold 2011), single cells are trapped and manipulated using highly focused laser beams. The foundation of optical traps, also known as optical tweezers, was developed by Ashkin *et al.* in the 1980s (Ashkin and Dziedzic 1987; Ashkin, Dziedzic, and Yamane 1987). They demonstrated optical tweezers for trapping and manipulating single *E. coli* cells in media (Ashkin and Dziedzic 1987; Ashkin, Dziedzic, and Yamane 1987). Block *et al.* (Block, Blair, and Berg 1989, 1) used optical tweezers to measure the mechanical properties of single *E. coli* and *Streptococcus* cells. This method is amenable to the integration with transparent microfluidic devices as long as the device design is compatible with high numerical aperture optics to achieve steep intensity gradients around the target cells in the microchannels (Kühn et al. 2009; Min et al. 2009). Early reports on the integration of optical tweezers and microfluidic devices date back to 2004 (Enger et al. 2004; Munce et al. 2004). In these studies, trapped individual cells were transported by the laser beam to a compartment for subsequent culturing and/or analyses.

Because the trapped cells are not exposed to the ambient environments during optical micromanipulation, sample contamination issues can be minimized.

However, the application of micromanipulation is perhaps subjected to undesired stresses introduced either by mechanical forces or light/thermal damage to the cells (Rasmussen, Oddershede, and Siegumfeldt 2008). In addition, micromanipulation is usually labor-intensive and time consuming with the low throughput representing the major limitation to the method's widespread usage in the research field.

2.1.4. Cell sorting

The purpose of cell sorting is to separate a heterogeneous mixture of biological cells into corresponding sub-populations, typically one cell at a time. In order to distinguish between the sub-populations of cells the use of specific markers or stains and a sensitive detection method are required. Laser Induced Fluorescence (LIF) is recognized as one of the most sensitive and reliable detection methods. Fluorescence-Activated Cell Sorting (FACS) is based upon the detection of laser-induced scattered light and/or fluorescence signals emanating from the cell or fluorescent markers, respectively, and sorting of individual cells according to their scatter/fluorescence signatures. It provides fast, accurate and quantitative recording of fluorescence signals of individual cells as well as physical separation of cell populations of particular interest.

FACS can be used as a very efficient way to separate individual cells. Its application to sort small cells like bacteria started in late 1990s (Fuchs et al. 1996; Yi et al. 1998; Baptista et al. 1999). This technology requires sophisticated devices and trained operators. Potentially inexpensive, chip-level FACS systems have been produced to circumvent these shortcomings. Fu *et al*. (Fu et al. 1999) developed a microfabricated

FACS device and demonstrated its effectiveness in sorting micrometer-sized latex beads and bacterial cells. Compared with the conventional FACS, chip-level FACS devices offer the advantages of: *i*) integration with other chip-level analytical technologies, such as PCR or microarrays, *ii*) incorporation of multiple cell sorters on a single chip for parallel processing, allowing further increased throughput, and *iii*) markedly lower reagent consumption and thus cost-effectiveness. One example is the microfluidic cytometer featuring 384 channels for parallel operation developed by Mckenna *et al.* (Mckenna et al. 2009) for rare-cell screening. Their device was able to perform a genome-wide cDNA screening assay with statistically significant results on positive counts of only several dozen cells in a background of several million negatives. Although this device was not designed for bacterial cells, the principle should be the same for bacterial cell sorting and can potentially be extended to bacterial cell sorting in the near future.

Figure 2 depicts schematic representations of these methods. In general, the selection of these techniques depends on the purpose, available resource and technical requirements of a study. The dilution-to-extinction method is easy to use and is more suitable for cell isolation from a pure culture or of most abundant microbes. The microdevice for this method is simple to design and construct, and it does not require a precise control of liquid manipulation. In the literature, the common loading efficiency of the dilution-to-extinction is about 30% and the capacity depends on the number of chambers. Dilution-to-extinction cell loading can be completed within one minute. Devices that use flowing microdroplets can generate more than 10^3 droplets per second. Therefore, the number of chambers (e.g., microdroplets) is determined by the duration of

droplet generation. Mechanical trapping shares the low level of complexity with the dilution-to-extinction method, while hydrodynamic trapping and DEP trapping require a precise liquid control to achieve accurate and reliable isolation of single cells thus increasing the overall cost and complexity of both techniques. The most reliable techniques for single-cell isolation are micromanipulation and FACS. However, the throughput of the mechanical micromanipulation-based approaches is relatively low, typically about one cell per a few minutes. The major limitation for using a mechanical micromanipulator in combination with a microfluidic device is that microchannels are usually sealed from the ambient environment preventing the pipette tip of the micromanipulator from accessing the samples. A key feature of the micromanipulation method is that it provides researchers with a means to precisely control the cell selection procedure. FACS is typically capable of single-cell separation with throughputs of up to ~10^4 cells/second. Although the principle of separating eukaryotic and prokaryotic cells in FACS is not much different, most of the commercial instruments available currently are not designed for separation of small cells, and therefore would need further optimization (Lomas, Bronk, and van den Engh 2011).

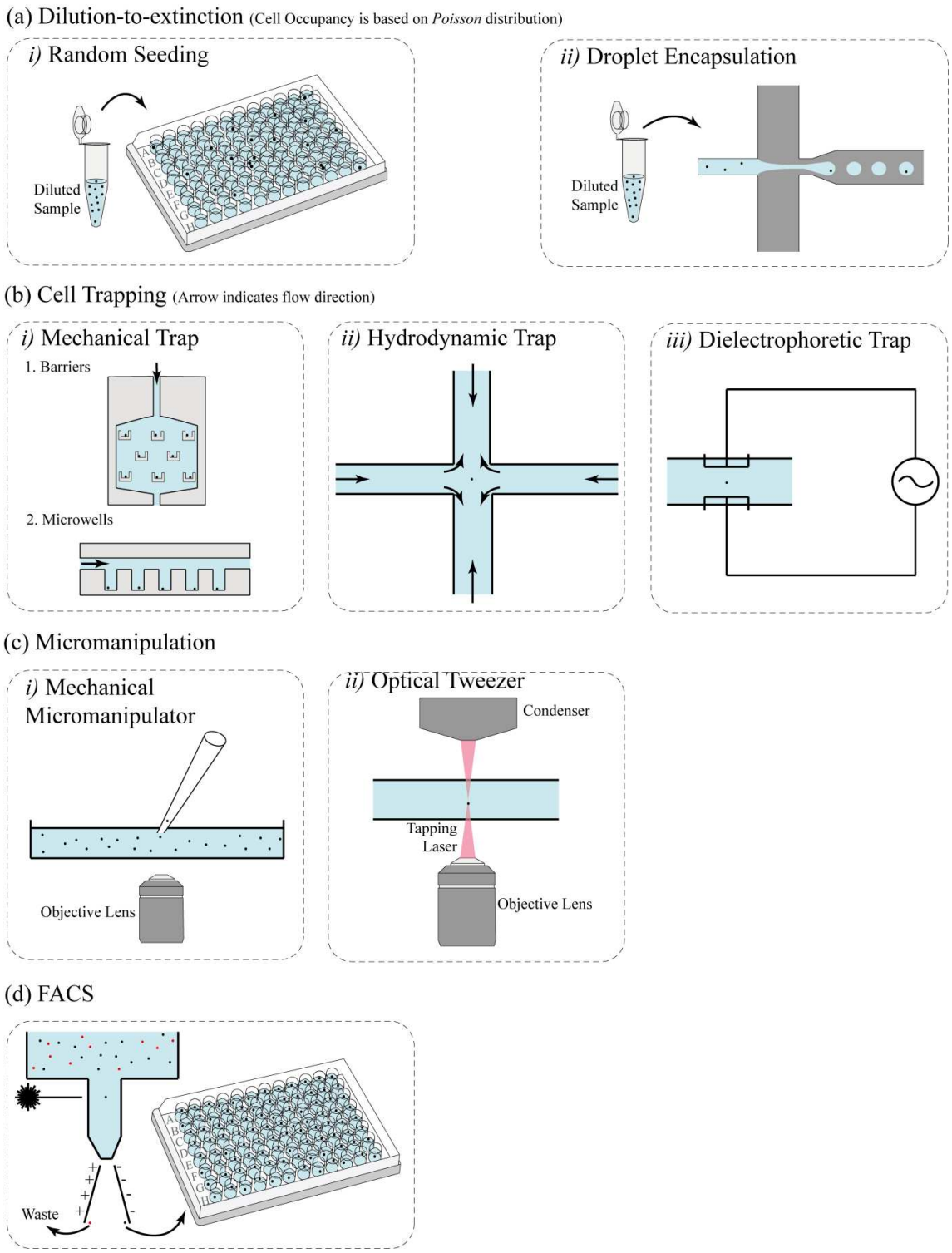

Figure 2. Principle of operation of different methods for small single cell manipulation.

2.2. Single cell gene expression

With the ability to isolate single cells, gene expression at the single cell level is a new avenue of research. Several approaches have been utilized to evaluate gene expression heterogeneity at the single cell level (Stewart and Franklin 2008). The first method is to utilize reporter genes (Chalfie et al. 1994). The simple and sensitive enzymatic assays of e.g. β-galactosidase and luciferase have allowed detailed investigations of gene transcription regulation mechanisms. These reporter systems can be obtained through the construction of the relevant fusions between promoters of interest and the respective reporter genes. However, the main challenge of this approach is that not all species, especially the ones found in natural environments, are amenable to genetic manipulation. The second method is fluorescence *in situ* hybridization (FISH). FISH has been used effectively for assessing the diversity of species in nature. FISH targeting rRNA is a highly useful method for the phylogenetic identification of bacteria (Amann, Ludwig, and Schleifer 1995). However, its accuracy as a quantitative method for determining the expression levels of lowly expressed genes is still questionable. The third method is *in situ* PCR combined with *in situ* reverse transcription *(in situ* RT-qPCR) (Aoi 2002). RT-qPCR was developed to amplify and detect functional genes and their expression levels inside a single cell. This is a very useful approach to characterize the genetic and phylogenetic properties of natural communities at the single-cell level. In the field of environmental microbiology, Hodson *et al.* (1995) developed an *in situ* PCR method for prokaryotic cells (bacteria) and gave examples of its use for the detection of a specific gene *(nahA)* and its transcripts in a single *Pseudomonas* cell within a model microbial community. Since then, the *in situ* RT-qPCR approach has been successfully applied to

detect gene expression in many species, such as *Zac* mRNA in *Salmonella typhimurium* (Tolker-Nielsen, Holmstrøm, and Molin 1997), and *dnaK* in *Methanosarcina mazei* S-6 cells (Lange et al. 2000).

2.2.1. RT-qPCR based gene expression measurements in single cells

Reverse-transcriptase quantitative PCR (RT-qPCR) is the most reliable approach for gene expression analysis in single cells (Kubista et al. 2006; Nolan, Hands, and Bustin 2006). The technology is the same as conventional RT-qPCR for bulk cells except slight modifications are necessary to optimize the performance at the single cell level. Several protocols have already been published for gene expression analysis by RT-qPCR for single mammalian cells (Lindqvist, Vidal-Sanz, and Hallböök 2002; Wacker, Tehel, and Gallagher 2008; K. Taniguchi, Kajiyama, and Kambara 2009). The most advanced protocol was published recently by Taniguchi *et al.* (2009) who used a quantitative PCR method featuring a reusable single-cell cDNA library immobilized on beads for measuring the expression of multiple cDNA targets (from several copies to several hundred thousand copies) in a single mammalian cell.

Advances in gene expression profiling of a small number of cells was witnessed in recent years. For instance, combined with micro-dissection, Lenz *et al.* (2008) captured subsets of cells from the vertical strata within *P. aeruginosa* biofilms and quantified transcripts of mRNA and 16S rRNA using RT-qPCR. So far, few publications have been reported for gene expression measurements in small single cells using the RT-qPCR based method. This is probably due to the technical challenges specific to the analysis of small cells (2-5 µm) as compared with mammalian cells (10-20 µm) and, as a result, lower amounts of any given mRNA molecule. Attempts were made in my laboratory to

overcome these issues by developing two methods based on a combination of SYBR Green and RT-qPCR to directly determine the gene expression levels in small cells like bacterium (Gao, Zhang, and Meldrum 2011). The first method is a single-tube approach which allows the analysis of only one gene from each bacterial cell. The procedure includes single cell picking using a micromanipulator, followed by thermal cell lysing and one-step RT-qPCR. We have optimized the PCR primer design and thermal cycling conditions to avoid the interference from primer dimers during qPCR. Using this procedure, expression levels of the *gadA* gene, a lowly expressed gene that encodes for a glutamate decarboxylase isozyme in single *E. coli* cells, was determined. In addition, an increased expression of *gadA* was observed in all cells exposed to acidic conditions (pH 5), which is consistent with the expected response of the gene reported in the bulk-cell study (Tucker, Tucker, and Conway 2002). The second method features a two-stage protocol that consists of RNA isolation from a single bacterial cell and cDNA synthesis in the first stage, and qPCR in the second stage. After evaluation of different commercial kits for RNA isolation, total RNA isolation and purification from single bacterial cells was achieved for the first time. With optimized conditions for both reverse transcription and qPCR, it is possible to simultaneously determine the expression levels of multiple genes in single bacterial cells. This procedure was applied to study the response to thermal shock in *E. coli* populations with single cell resolution. The reproducible results demonstrated that the method is sensitive enough not only for measuring cellular responses at the single-cell level, but also for revealing gene expression heterogeneity among bacterial cells. Furthermore, our results showed that the two-stage method can reproducibly measure multiple highly expressed genes from a single *E. coli* cell. This

finding provides a foundation for the future development of a high–throughput, lab-on-chip methodology for whole-genome RT-qPCR of single bacterial cells.

2.2.2. Whole-transcriptome based gene expression measurements in single cells

Gene expression profiling for complicated biological traits on a genomic scale depends on recent advances in high-throughput gene expression analysis technologies, such as DNA microarrays, Serial Analysis of Gene Expression (SAGE) or Next Generation Sequencing (NGS). With these techniques we now can quantitatively investigate complex cellular processes systematically (Kitano 2002). Although these fast growing technologies help us to understand the cellular processes from a different point of view, understanding the complexity of cellular processes remain tethered to some technical hurdles, sure as requirement of relatively large quantities of the initial RNA in order to obtain reliable data. For instance, several hundred nanograms to micrograms of total RNA is needed for transcriptome profiling, which is equivalent to a sample size of more than 10,000 eukaryotic cells. To address this issue, several successful attempts have been made in recent years to develop a total transcript amplification (TTA) method for single eukaryotic cells and use either DNA microarray (Kurimoto et al. 2007) or mRNA sequencing (mRNA-Seq) (Tang et al. 2009) technologies to analyze the gene expression levels.

TTA techniques can be divided into four classes based on amplification strategies: *i*) linear amplification methods using T7-mediated transcription (E. Wang et al. 2000; Schneider et al. 2004); *ii*) exponential amplification methods using PCR techniques (Brady and Iscove 1993; Iscove et al. 2002); *iii*) NuGen RiboSPIA amplification processes (Singh et al. 2005); *iv*) SMART-Seq (Ramsköld et al. 2012); and *v*) φ29

polymerase multiple displacement amplification (MDA) of circularized cDNA (Kang et al. 2011). Most of these technologies mentioned above focus on larger eukaryotic cells, but their possible extension to single bacterial cells is discussed below.

DNA microarray researchers have intensively applied the T7-mediated methods and have generated highly reproducible results due to the fact that they are theoretically linear and independent of template sequences (E. Wang et al. 2000). However, the major drawbacks of this method include overlong sample preparation times (1.5-2 days for a single round of TTA), limited sensitivity, and less stable RNA-based products. This method becomes more unreliable when working with less than 10 ng of total RNA, probably because of increased bias and noise arising from multiple rounds of amplification (Wilson et al. 2004; Subkhankulova and Livesey 2006). As a result, T7-mediated methods have not been widely applied for single cell TTA. Meanwhile, most T7-mediated methods require polyA structures for mRNAs then unsuitable for single bacterial global gene expression profiling.

PCR-based TTA strategies were another solution to conduct whole-transcriptome analysis of single cell. Global gene expression profiling of eukaryotic cells has been achieved by using PCR-based approaches with very low total RNA amounts (several picogram range) (Brady and Iscove 1993). The PCR-based methods offer some advantages such as speed, detection sensitivity, and cost. DNA-based PCR products are much more stable than RNA products of T7-mediated methods. The relatively high amount of PCR products allows for multiple-time analysis and the remaining PCR products are stable at low temperature for long time for possible further verification or investigation (Iscove et al. 2002). PCR-based TTA approaches have not yet been widely

adopted, albeit these advantages, for global gene expression analysis. The major reasons are GC-content bias, double stranded products, and the non-linear amplification of the PCR methods (Glanzer and Eberwine 2004). Several research groups developed techniques combining both PCR and T7-mediated approaches to utilize the advantages of both technologies, however, it could not be widely used for bacterial TTA since primer poly(dT) was required for both PCR and T7-mediated methods (Kurimoto et al. 2007).

NuGen Technologies developed an emergence TTA technique named RiboSPIA by applying a chimeric RNA/DNA primer and RNase H and DNA polymerase to produce amplification of several thousand-fold from single-stranded DNA-based amplifiers**.** This technique is relatively fast (typically 6 h per TTA round), can be cooperated with as low as a picogram starting total RNA, and is much more robust because it uses DNA synthesized from RNA (Singh et al. 2005). In principle, there are no technical hurdles to transfer this technique for whole transcriptome studies of single prokaryotic cells.

In a recent editorial highlighted article in nature biotechnology, a new single cell transcriptome technology, SMART-Seq, was claimed as a robust and reproducible method for full length of mRNAs (Ramsköld et al. 2012). However, due to poly(dT) primer is required for this methods, SMART-Seq would not be helpful for single bacteria transcriptome analysis.

The φ29 polymerase multiple displacement amplification (MDA) technique is a versatility technique which can be used for both eukaryotic and prokaryotic cells. The first available single bacterial transcriptome analysis was reported by Kang *et al.* (2011) who used *Burkholderia thailandensis* cells exposed to 0.01% (w/v) of glyphosate, an antibacterial agent, for single-bacterium TTA. The amplified whole transcriptome out of

a single *B. thailandensis* cell was analyzed by means of a DNA microarray. The results showed lower fold-change bias (less than two-fold difference and *Pearson* correlation coefficient $R \sim 0.87–0.89$) and drop-outs (4%–6% of 2842 detectable genes) as compared with the data obtained from non-amplified RNA samples. In addition, Sanger sequencing of 192 clones generated from the TTA product obtained from a single cell, with and without enrichment by eliminating rRNA and tRNA, detected only *B. thailandensis* sequences without contamination. Although the sensitivity and accuracy of the whole transcriptome analysis are in general lower than that of RT-qPCR, it can measure expression levels of several thousand genes simultaneously, a remarkable advantage that has not be replicated by other existing methods. However, the approach is time consuming (about 3 days for a round of TTA). Also, we noted that the estimated total RNA in a single *B. thailandensis* cell is about two picograms, which is several orders of magnitude higher than the estimated total RNA amount of a typical bacterial cell, such as *E. coli* (Gao, Zhang, and Meldrum 2011; Schmid et al. 2010). Therefore, a further evaluation of the method is needed to assess its feasibility for single cell studies.

2.2.3. Imaging-based gene expression measurements in single cells

Powerful methodologies based on reported probes and imaging allow for achieving the spatiotemporal information about expression of specific mRNAs in both intact eukaryotic and prokaryotic cells (Sanjay Tyagi 2009). Because no intrinsically fluorescent RNA motifs exist, *in vivo* imaging of mRNA transcripts is less common than proteins. Instead, fluorescent proteins binding to specific RNA motifs (Bertrand et al. 1998; Calapez et al. 2002; Golding and Cox 2004; Rackham and Brown 2004; Kerppola 2006; Daigle and Ellenberg 2007; Valencia-Burton et al. 2007), sequence-specific oligonucleotide probes

(Cardullo et al. 1988; Morrison, Halder, and Stols 1989; Sixou et al. 1994; Q. Li et al. 2002; Sando and Kool 2002), aptamer tagging (Babendure, Adams, and Tsien 2003; Sando, Narita, and Aoyama 2007), rapid detection of miRNA by a silver nanocluster DNA probe (Yang and Vosch 2011) and RNA mimics of green fluorescent protein (Paige, Wu, and Jaffrey 2011) have been adopted for mRNA imaging. Such efforts, although still in their infancy, have already shed light on the RNA distribution and dynamics in living cells.

The most established method for imaging of the intracellular RNA in live cells is tagging mRNA with a fluorescent protein (i.e., GFP, YFP, and RFP) (Bertrand et al. 1998; Calapez et al. 2002; Golding and Cox 2004; Rackham and Brown 2004; Kerppola 2006; Daigle and Ellenberg 2007; Valencia-Burton et al. 2007). To tag a specific target mRNA, an RNA-binding protein must be fused to GFP and at the same time the 3'-untranslated region of the target mRNA must be tagged with an RNA motif recognized by the RNA-binding protein. Bertrand *et al.* (1998) first introduced the MS2 coat protein-GFP approach for imaging mRNA dynamics in live cells. There were two components of this method. The first is the MS2 coat protein, a phage RNA-binding protein, expressed as a fusion with intact GFP. The second is the target ASH1 mRNA, which is tagged with multiple copies of MS2-binding motifs. When these two components are co-transformed and co-expressed in cells, MS2–GFP fusion proteins bind to their cognate motif on the mRNA and render it fluorescent.

Since then, imaging using GFP as a reporter protein has been applied to different mRNAs in diverse organisms (Golding and Cox 2004). Recently, other RNA motifs, such as λN from bacteriophage λ (Daigle and Ellenberg 2007) and even poly(A)-binding

protein have also been employed to image the dynamics of mRNAs in eukaryotic cells (Calapez et al. 2002). The major challenge in tagging with intact GFP is the need to distinguish bound GFP from unbound GFP, since GFP constructs are always fluorescent. This was recently overcome by adopting the reconstruction of GFP, by splitting GFP into two nonfluorescent fragments. The two fragments are non-fluorescent until a pair of tags attached to each fragment recognize the target mRNA and assemble the two split GFP fragments into a correctly folded and functional protein (Kerppola 2006). MS2 coat protein and zip code-binding protein fused with split GFP fragments (Rackham and Brown 2004), split eIF1A domains fused with N- and C-terminal of GFP fragments (Valencia-Burton et al. 2007), and PUMILIO1 (a unique sequence-specific RNA binding protein) have successfully been demonstrated as applications of the split GFP approach. However, a drawback of the split GFP tagging method is the high affinity of the two protein fragments to each other, making the binding difficult to reverse. This prevents the method from being utilized for imaging of fast dynamic processes (Magliery et al. 2005).

The second approach is based on imaging of endogenous mRNAs using fluorescence resonance energy transfer (FRET) and contact-mediated quenching (Cardullo et al. 1988; Morrison, Halder, and Stols 1989; Sixou et al. 1994; S Tyagi and Kramer 1996; Q. Li et al. 2002; Sando and Kool 2002; Santangelo et al. 2004). Several different probes whose fluorescent properties change upon sequence-specific hybridization have been explored, including competitive hybridization probes (Morrison, Halder, and Stols 1989; Sixou et al. 1994; Q. Li et al. 2002), side-by-side probes (Cardullo et al. 1988), quenched autoligation probes (Sando and Kool 2002), molecular beacon probes (S Tyagi and Kramer 1996), and dual molecular FRET probes (Santangelo

et al. 2004). Probe-based imaging features several distinct advantages: probes detect mRNA in cells directly without the need to engineer target genes and GFP constructs; in addition, the approach can be multiplexed by using spectrally distinguishable fluorophores, and the possibility to sort cells based on gene expression levels but not the only “positive” or “negative” signals (Sanjay Tyagi 2009). The limitation of this method includes lower sensitivity due to these probes have only one fluorophore in each molecule resulting in lower overall signals compared to GFP tags, the need of delivering probes into cells and degradation of probes (Sanjay Tyagi 2009).

The third approach employs tagging of artificial RNA motifs (aptamers) with small nonfluorescent dyes to render fluorescence when combined with specified aptamers (Babendure, Adams, and Tsien 2003; Sando, Narita, and Aoyama 2007). The free dye molecules are nonfluorescent because of the strong dissipation of the excitation energy through vibrational (radiationless) relaxation. Once a selected aptamer binds to the dye molecule restricting its vibrational freedom, the dyes become fluorescent resulting in an increase of the fluorescence signal by more than 2,000 fold (Babendure, Adams, and Tsien 2003). Examples include Hoechst dye variants that are non-fluorescent in the unbound form but show strong fluorescence when bound to pre-selected effective RNA aptamers (Sando, Narita, and Aoyama 2007). The availability of many aptamer-dye combinations allows imaging of multiple mRNA simultaneously. However, the free radicals created by the irradiated dye can destroy the RNA motifs (Grate and Wilson 1999). Taniguchi *et al.* (2010) used a DNA oligomer probe labeled with a single fluorophore to successfully hybridize it with the mRNA on a microfluidic device. No further applications using this method in combination with microfluidic devices have

been reported, and no high-throughput, automated systems are currently available for this methodology.

Other recent methods, such as the silver nanocluster DNA probe for miRNA (Yang and Vosch 2011) and RNA-based variants of green fluorescent protein (Paige, Wu, and Jaffrey 2011), are also expected to contribute to quantitative imaging of multiple mRNAs and small RNAs simultaneously in single cells in the near future.

3. DEEP SEA COMMUNITY STRUCTURE IDENTIFICATION

3.1. Introduction

This section of work has been published in Microbial Ecology (Gao et al. 2011). I would like to thank the first author Dr. Gao and my co-worker Dr. Wu. Only with their help, this portion of work could be achieved.

Photosynthetic microbes (both prokaryotes and eukaryotes) are the most accessible samples in the ocean and intensive work has been done on these surface water species. These tiny microbes have significant effects on the global carbon cycle through photosynthetic fixation of CO_2. Due to the flux of particulate organic carbon (POC), also known as marine 'snow' (Alldredge and Cohen 1987), into the deep sea (Eppley and Peterson 1979), the ocean becomes a natural sink of CO_2 (Dore et al. 2003; C. L. Quéré et al. 2007). It has been reported that global oceanic CO_2 sink may have increased to 118 $\pm$ 19 $\times$ 10^9 tons of carbon from 1800 to 1994 which is equivalent to about 48% of the total fossil-fuel and cement manufacturing emissions (Sabine et al. 2004). Photosynthetic microbes in the ocean may hold the key solution to address this problem.

The average depth of the ocean is about 3,682 meters (Charette and Smith 2010) and can be divided into several zones based on depth, light abundance, and physical and biological conditions. The top 200 meters consists of the epipelagic zone where there is enough light for photosynthesis and thus plants and animals are concentrated in this zone (Gao et al. 2011). While at the bottom of the ocean (from 4000 meters down to the ocean floor), which is called abyssopelagic zone (Ikeda et al. 2007), it is almost entirely dark and no sunlight can reach this depth. Theoretically, no photosynthesis is expected at this

depth since photosynthesis needs enough light energy to initiate O_2 production (Mcallister, Shah, and Strickland 1964).

Recently, more evidence showed that some photosynthetic microbes can also exist in the deeper ocean. For instance, photoautotroph cyanobacteria *Synechococcus* was found in the 800 m deep Adriatic waters during the spring of 2006 using a flow cytometer by Vilibić and Šantić (2008). Zubkov and Burkill (2006) also detected the presence of both *Synechococcus* and *Prochlorococcus* cyanobacteria (10-20 cells per mL) in the aphotic zone down to 300 m depths. Similarly, it was found that diatoms, a major group of eukaryotic algae, possibly exist at about 3,000 meters depth (López-García et al. 2001). Nevertheless, no direct evidence shows those photoautotrophic microbes are actively functioning in that depth. In order to find one ubiquitous species that can be used as a sensor to sense the environmental perturbation, this species cannot be in a dormant status. Gene expression analysis needs to be performed to rule out the dormant status.

The difficulty of deep sea research is those samples cannot be easily cultured in laboratory conditions since the deep sea conditions are hard to regenerate. Because of the vastly improved molecular biology technologies, cultivation independent phylogenetic analysis using ribosomal RNA (rRNA) sequencing has been applied to decipher the community structure of microbes in the ocean (S. J. Giovannoni et al. 1990; López-García et al. 2001; Gao et al. 2011; Stahl et al. 1984). More recently, RT-qPCR (Reverse transcription - quantitative polymerase chain reaction) technology is also used to detect changes at transcription level that correspond to alterations in the environment and to collect information about stress induced response. The RT-qPCR technology has excellent sensitivity, dynamic range, and reproducibility and has become a routine and

robust approach for measuring the expression of genes of interest (Stephen A Bustin et al. 2009; VanGuilder, Vrana, and Freeman 2008).

In this study, community structure and gene expression analyses was performed on the sea water samples collected from sites of 765-790 meters in depth in the northeast Pacific Ocean. The initial 16S rRNA based clone libraries analyses showed that the majority of the archaeal OTUs (Operational Taxonomic Unit) belongs to the uncultured group I *Crenarchaeota*, whereas most of the bacterial OTUs belongs to alpha-, gamma- and delta-*proteobacteria*, consistent with previous analysis of deep sea microbes. In order to further explore the community structure of the deep sea samples, a 23S-rRNA plastid gene cloning library was constructed. The results showed that the majority of this cloning library was occupied by oxygenic photoautotrophic organisms, such as diatoms *Thalassiosira spp*. In addition, RT-qPCR was applied to determine the gene expression for the 23S rRNA plastid gene, which is involved in protein synthesis in both eukaryotic algae and cyanobacteria. The results showed that the microbes here were not in dormant status. The evidence provided by this work has implied that some highly adaptive photoautotrophic organisms could be metabolically functional in the deep ocean so as to be used as a sensor candidate to monitor the environment perturbation in that environment.

3.2. Experiments

3.2.1. Cell recovery and microscopy analysis

Ocean microbes which were retained on a 0.22 µm filter (Millipore, Billerica, MA) were collected from the Pacific Ocean. Microbes were recovered from the filter in the laboratory. A 5 mL syringe mounted with an 18G 1 ½ needle (BD, Franklin Lakes, NJ)

was used to wash the filter. The cells were carefully washed off from the filter membrane using 3 x 5.0 mL of RNALater (Ambion, Austin, TX), which can protect RNA from degradation. The 2.0 mL concentrated cells were equally divided into two microcentrifuge tubes and stored at -20°C. For cell counting, 4 µL of 5 µg/µL 4, 6-diamidino-2-phenylindole (DAPI) (Sigma) and 100 µL of 50% glutaraldehyde was added into the 5 ml cell suspension in phosphate buffer (5 mL), then the stained cells were filtered through an isopore membrane filter (0.22 µm). The isopore membrane was then put on a glass slide, immersion oil added and covered with a cover slip. The slide was observed under epifluorescence microscope (Nikon Eclipse Ti System) using 60× and 40× objectives.

3.2.2. Cloning library construction

The total DNA was extracted and purified from 1.0 mL collected cells using a DNeasy Blood & Tissue Kit (Qiagen, Valencia, CA). Based on the isolated DNA, three DNA fragments were amplified for cloning library construction: 1,400 bp16S rRNA gene from bacteria (FD1, forward primer: AGAGTTTGATCCTGGCTCAG, 1540R, reverse primer: AAGGAGGTGATCCAGCC) (Hwang et al. 2009), 700 bp 16S rRNA gene from Archaea (Ar20F, forward primer: TTCCGGTTGATCCYGCCRG, Arch958R, reverse primer: TCCGGCGTTGAMTCCAATT) (E. F. DeLong 1992) and 500 bp 23S rRNA gene from cyanobacteria and eukaryotic algal chloroplasts (p23SrV-f1, forward primer: GGACAGAAAGACCCTATGAA, p23SrV-r1, reverse primer: TCAGCCTGTTA-TCCCTAGAG) (Sherwood and Presting 2007). The conditions for PCR amplification was initially set up as: (a) bacteria (Hwang et al. 2009): 94°C for 2 min; 30 cycles of 30 s at 94°C, 1 min at 58°C, and 1 min at 72°C; and final cycle at 72°C for 7 min; (b) Archaea

(E. F. DeLong 1992): 94°C for 2 min; 30 cycles of 1.5 min at 95°C, 1.5 min at 55°C, and 1.5 min at 72°C; and final cycle at 72°C for 7 min; and (c) algae and cyanobacteria: 94°C for 2 min; 30 cycles of 1.5 min at 95°C for, 1.5 min at 55°C , and 1.5 min at 72°C; and final cycle at 72°C for 7 min. The expected PCR products were recovered by using a QIAquick Gel DNA Extraction Kit (Qiagen, Valencia, CA).

The PCR products were cloned into a pGEM-T easy vector following the protocol provided by the manufacturer (Promega, Madison, WI). Randomly chosen white clones were cultivated in 96-well plates and used for plasmid isolation and sequenced with an ABI 373 Sequencer using PCR primers. The resultant DNA sequences were subjected to manual editing using Sequence Scanner Software 2.0 (Applied Biosystems, Foster, CA). Sequences were compared with those in GenBank through the NCBI internet service using BLAST 2.2.10 (Altschul et al. 1997). Alignment of sequences was done with online Clustal W (Thompson, Higgins, and Gibson 1994). OTUs were determined based on a 3% divergence cutoff for individual “species” OTU.

3.2.3. Gene expression analysis

Total RNA were extracted from 1.0 mL collected cells using the Trizol Max Bacterial RNA Isolation Kit (Invitrogen, Carlsbad, CA) and purified with RNeasy Mini Kit (Qiagen, Valencia, CA). To determine the gene expression involved in protein synthesis in photoautotrophic microbes, a set of 23S based primers specific to eukaryotic algae and cyanobacteria was adopted from the literature (Sherwood and Presting 2007) and synthesized by Invitrogen Corporation (Carlsbad, CA). Using genomic DNA of *Synechocystis* sp. PCC6803 as template, this 500 bp 23S rRNA gene was amplified for the purpose of generating a standard curve in RT-qPCR. With SYBR Green One-Step

reagents (BioRad, Hercules, CA), RT-qPCR was performed on a Rotor-Gene 6000 (Corbett Life Science, now Qiagen, Valencia, CA). The 20 µL PCR reaction set up was as follows: 10 µL of 2x SYBR Green Rxn mix (BioRad, Hercules, CA), 1.5 µL of forward primer (4 µM), 1.5 µL of reverse primer (4 µM), 4.5 µL of nuclease free-water, 0.5 µL of iScript RT enzyme for One-Step (BioRad, Hercules, CA), and 2 µL of RNA or DNA template. The cDNA synthesis was performed at 50°C for 15 min followed by 95°C for 5 min for inactivation of reverse transcriptase. The PCR cycling program was: 40 cycles of 10 s at 95°C, 56 s at 56°C, 30 s at 72°C. Data analysis was carried out using the software provided by Corbett Life Science (now Qiagen, Valencia, CA).

3.3. Results

3.3.1. Microscopy analysis

These ocean samples were collected at 2-3 m above the sea floor. The epifluorescence microscopic images of the cells after DAPI staining are shown in Figure 3. A large number of microbial cells with size around 1-2 µm was observed, suggesting prokaryotic or small eukaryotic microbes were the dominant species in this environment. To seek evidence for possible photosynthetic microbes in this community, the auto-fluorescence without any dye staining was also checked. Although microbes with auto-fluorescence are rare in the samples, several auto-fluoresced microbes were found out of several hundreds of microscopic image fields examined. The results in Figure 3C show one of the auto-fluorescence images from the possible phototrophic microbes.

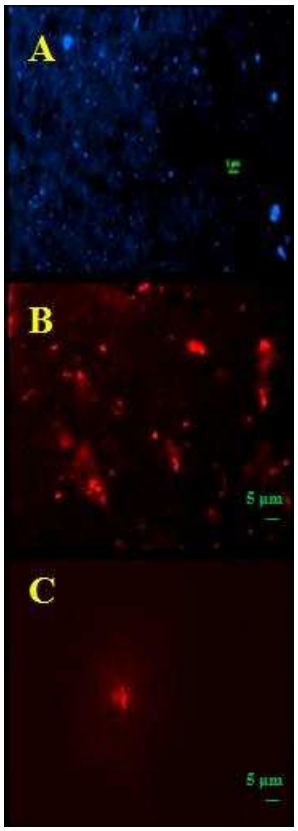

Figure 3. Epifluorescence microscopic images.

3.3.2. Phylogenetic analysis

Total DNA was extracted from 1.0 mL collected cells and a total of around 2600 ng chromosomal DNA was obtained. DNA isolated from the deep-sea water was used to construct three clone libraries for microbial community structure analyses: *i)* 16S rRNA for bacteria, *ii)* 16S rRNA for archaea and *iii)* algae and cyanobacteria specific 23S rRNA libraries.

Archaea: Sequence analysis of 156 random clones revealed 21 different phylotypes based upon 97% sequence similarity. The majority of clones (150 in total) are phylogenetically similar to the uncultured marine group I *Crenarchaeota* (Figure 4), In addition, a small number of clones (6 in total) are phylogenetically affiliated to the

uncultured marine group II *euryarchaeote*, and they represented two previously identified OTUs, sequenced clone UEU78206 from the Santa Barbara Channel (Massana et al. 1997) and DQ300553 from the North Pacific Subtropical Gyre (Edward F. DeLong et al. 2006), respectively.

Bacteria (Figure 5): A preliminary sequencing analysis of 250 random clones revealed significant bacterial diversity present in this site after comparison with the 16S rDNA database from the Ribosomal Database Project (Cole et al. 2005). In general, a majority of the clones (168 in total) from the clone library are phylogenetically associated with the phylum *Proteobacteria*. In addition, forty-two clones (13 OTUs) are phylogenetically affiliated with phylum *Bacteroidetes*. Nineteen clones (4 OTUs) belong to *Actinobacteria*.

Eukaryotic algae and cyanobacteria (Figure 6): Sequencing of 94 random clones revealed that a majority of them are divided between diatoms (61 clones) and cyanobacteria solely belonging to *Synechococcus sp*. (16 clones). Among the diatom group, 45 clones (7 OTUs) are similar to *Thalassiosira spp*. Other closely related diatoms include: *Nitzschia* spp. (11 clones, 7 OTUs), *Odontella spp*. (2 clones, 2 OTUs), and *Phaeodactylum spp*. (3 clones, 2 OTUs). Also, a few clones phylogenetically affiliated to other algae commonly found in the oceans such as *Chlorella sp*. and *Emiliania sp*. (Cattolico et al. 2008) were also identified based on sequence similarity searching.

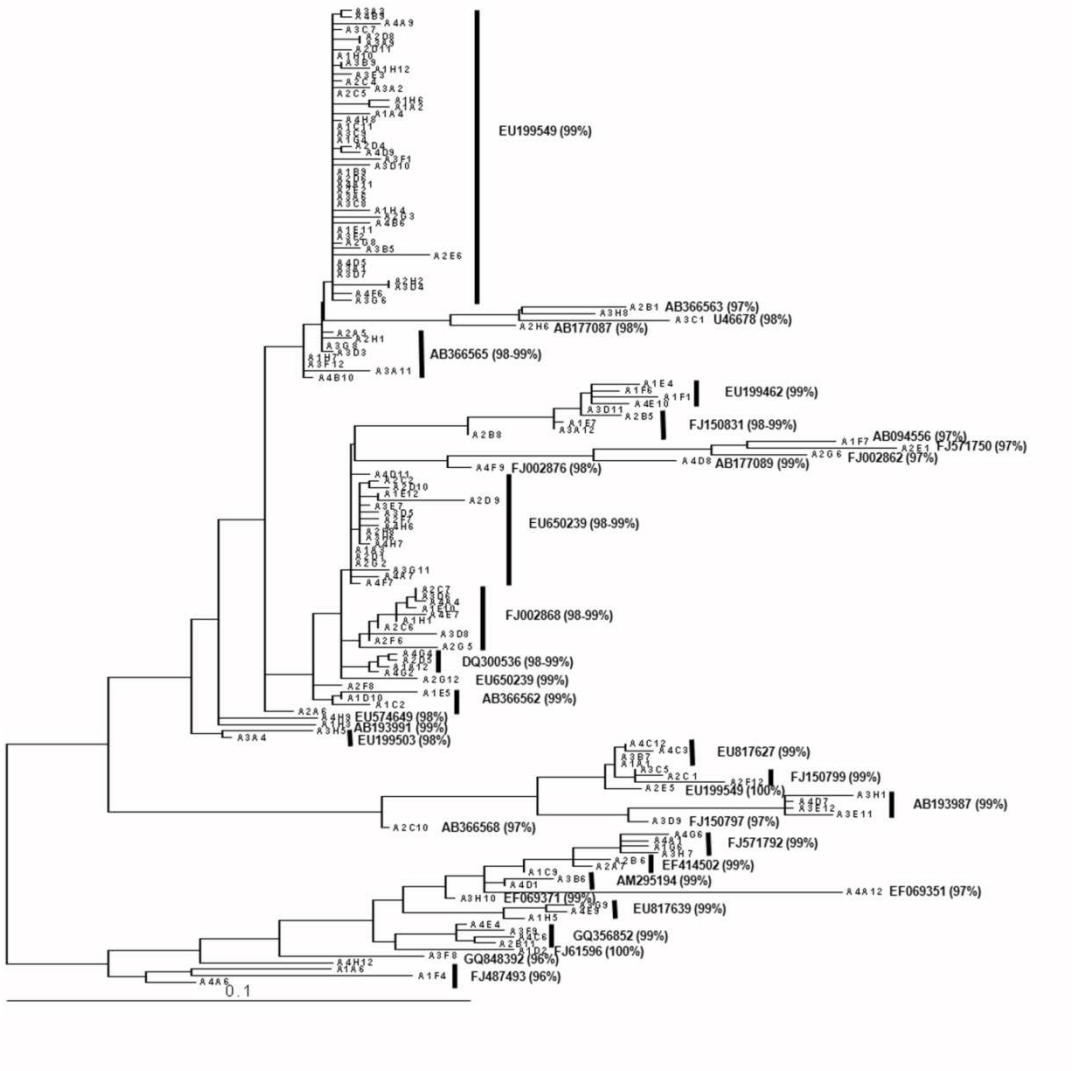

Figure 4. Phylogenetic analysis of archaea belonging to the uncultured marine group I *Crenarchaeota.*

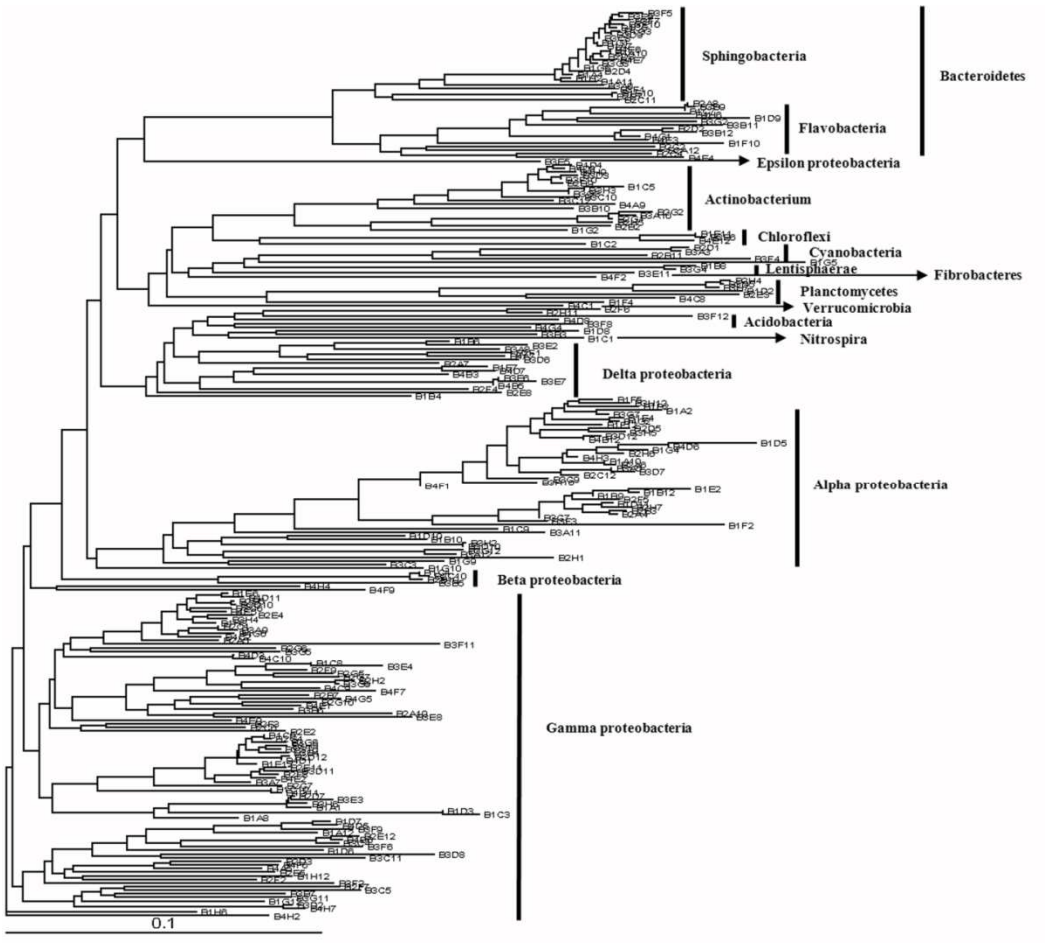

Figure 5. Phylogenetic analysis of bacteria found in the deep ocean water.

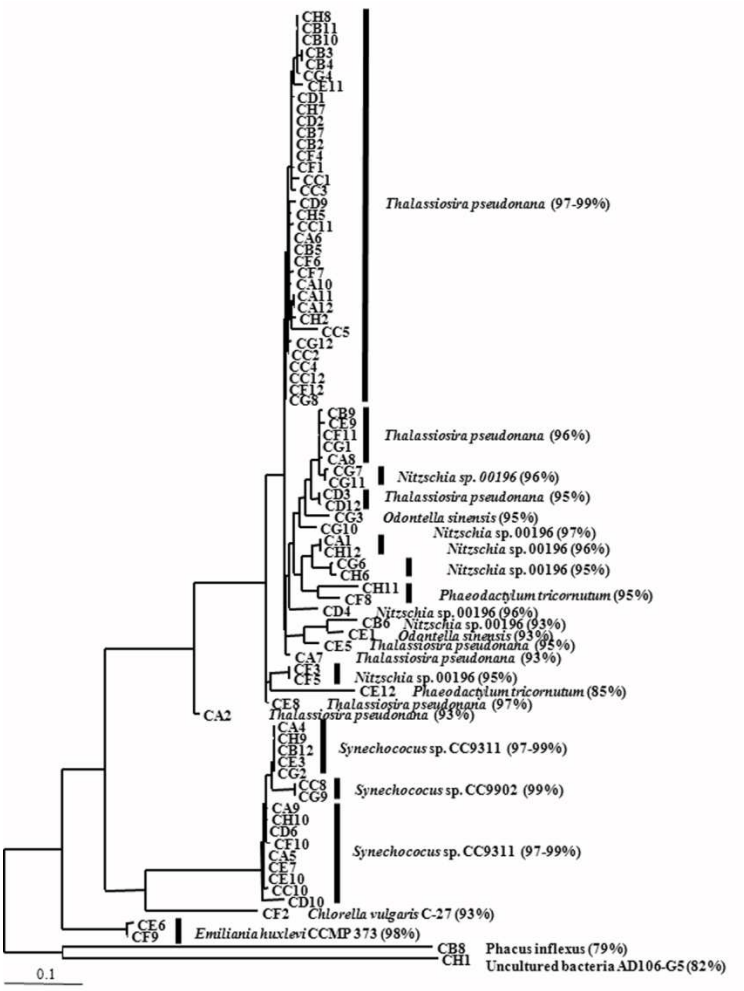

Figure 6. Phylogenetic analysis of the deep ocean photosynthetic eukaryotic algae. and cyanobacteria

The results were interesting since some photoautotrophic species were identified in the deep sea, but the deep sea photoautotrophic community was limited to two dominant species, *Synechococcus sp*. and *Thalassiosira spp.* Surface water contamination was not an issue in this work since no other surface water species was observed in any of the clone libraries. Since those photoautotrophic species also exist in surface water, they may be a good candidate to monitor the environmental perturbation.

3.3.3. Gene expression analysis

In order to use these microbes as sensors, they should not be in a dormant status otherwise they will not response to the environmental change. An immediate question that was raised was whether these photoautotrophic microbes are in a state of active metabolism or not. To seek the answer to this question, gene expression analysis was performed using total RNA isolated from the deep sea samples. The 23S rRNA gene which is involved in protein synthesis was analyzed by RT-qPCR. Figure 7 shows the real-time PCR analysis of the copy number of 23S rRNA gene. The analysis was done separately using DNA or cDNA as a template. The calculated average copy number of the algae and cyanobacteria specific 23S rRNA gene is 5.63×10^4 and 2.04×10^5 per reaction for DNA and cDNA templates, respectively. Considering the dilution factors of each template, the normalized copy number for 23S rRNA from cDNA (or RNA) and DNA is 2.44×10^7 and 4.25×10^6, respectively. Thus, the copy number of 23S rRNA is about 6 times higher than that of 23S rDNA (Figure 7). Meanwhile, considering that the efficiency of reverse transcription cannot be 100%, definitely, the ratio between 23S rRNA and 23S rDNA was underestimated and could be greater than 6. The results demonstrated that the 23S rRNA gene has activity in the deep sea which may indicate that the cells were not in a dormant status.

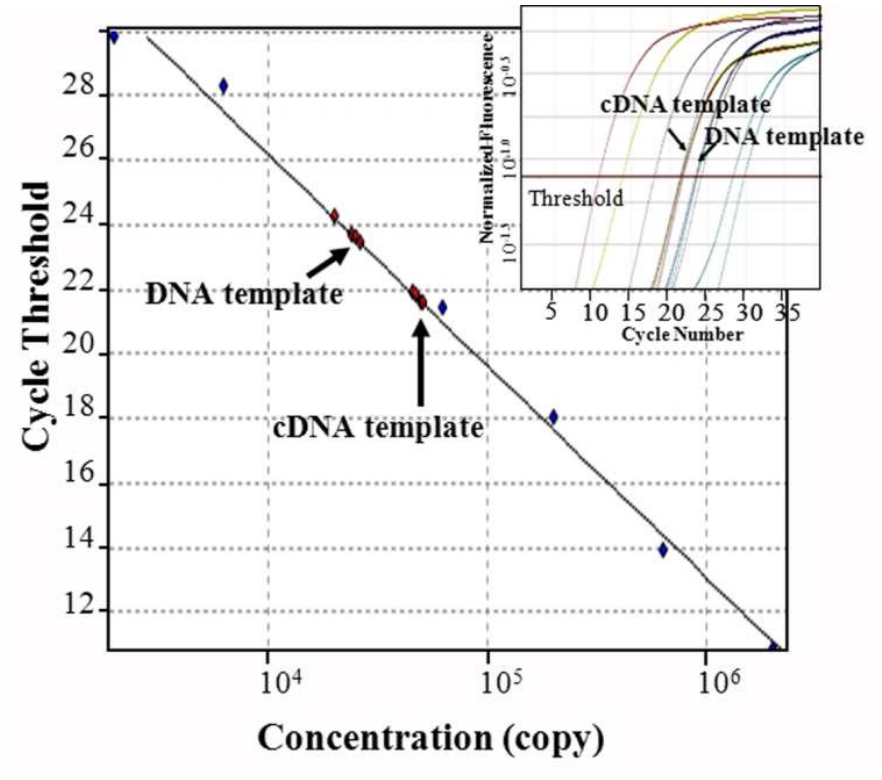

Figure 7. Quantitative RT-PCR analysis of algae and cyanobacteria specific 23S rRNA gene.

3.4. Discussion

It still remains unclear why photoautotrophic microbes, *Synechococcus sp*. and *Thalassiosira spp.*, exist in the deep mesopelagic zone. One plausible explanation is that they sank to this depth either by themselves or sank as a microbial assemblage with larger particles (Alldredge and Cohen 1987). The second possibility is because of the ocean current. However, in order to support those hypotheses, more evidence is still needed to demonstrate that RNA molecules (especially messenger RNA) can be stable through the time period of days or months during the sinking process. Normally the life time of RNA is hours (Gill et al. 2002) which is much shorter than the sinking process. Based on this fact, the RNA should all degrade at this depth; however, RNA was successfully detected

in our samples which may indicate there is another explanation about the existence of these photoautotrophic microbes at this depth.

My assumption is that these photoautotrophic microbes are natural inhabitants at this depth. If these photoautotrophic microorganisms were brought down by force or microbial assemblages associated with large sinking particles as discussed above, other photoautotrophic microorganisms that exist in surface water should also be observed in our samples, but only two dominant species were found. In addition, RT-qPCR results also supported my assumption by showing the activity of RNA at this depth.

Based on my assumption, I selected diatom *Thalassiosira spp*. as my future target, since it is a eukaryotic photosynthetic microorganism which may decrease the potential problems for this single-cell based work. The diameter of this microorganism is around 4-6 μm and it also exists in surface water which makes it an excellent sensor candidate. Meanwhile, diatoms contribute up to 40% of the primary productivity of the ocean (Maheswari et al. 2010; Nelson et al. 1995), if we can understand them better, it will also be very useful for regulating the primary productivity of ocean.

4. CHIP DEVELOPMENT

4.1. Introduction

This section of work has been published in Lab on a Chip (Shi et al. 2011). I would like to thank my co-worker Dr. Gao and my advisors Dr. Chao, Dr. Zhang and Dr. Meldrum. With their suggestions and helpful discussions, this portion of work has been successfully achieved.

Rapid development of microfabrication and microfluidic technologies have enabled development of more and more miniaturized analytical chips capable of performing analysis down to single-cell levels (Zare and Kim 2010; Schmid et al. 2010). The advantages of microfluidic devices for single cell analysis have been addressed in Chapter 2. A lot of researchers have taken advantage of these technologies, for instance, Zhang *et al.* (2011) demonstrated parallel real-time PCR with a sensitivity of about 1000 cDNA copies per 500-nL droplet produced using conventional photolithography. In one recent study, Marcy and colleagues (2007) developed a microfluidic device that performed isolation, amplification and sequencing of individual TM7 cells from a mixed microbial community that inhabits human mouths. The results showed low abundance species which would be easily neglected under traditional approaches. Parallel PCR at the single copy level is another good example of application of microfluidic devices (Musyanovych, Mailänder, and Landfester 2005; Beer et al. 2007; Diehl et al. 2006; Matsubara et al. 2004; Kojima et al. 2005; Nakano et al. 2003). These analyses are highly sensitive, but their analytes which typically are purified DNA are much simpler than analyzing the raw lysate of actual cell. In addition, the process does not involve cell lysing which usually results in more complicated chemical composition and/or fluidic

manipulation. These factors limit direct single-cell PCR at single copy resolution. This research strives to construct a chip that is capable of performing single cell analysis of small phytoplankton which is more challenging than that for mammalian cells due to the small size and their tough cell-wall structure. Recent progress on using microfluidic devices on single-cell PCR have focused on mammalian cells. Applications to other types of single cells are rare. Ottesen *et al*. (2006) used microfluidic digital PCR to amplify and analyze different genes obtained from single bacterial cells gathered from the environment. They used this device to identify bacteria in complex ecosystems and successfully reached the single molecular level resolution based on serial dilution and *Poisson* distribution. Zeng *et al.* (2010) designed an emulsion generating microfluidic device that used small droplets in oil as the reaction chambers. In this experiment, *E. coli* cells or isolated DNA were randomly seeded into the droplets with primer-adhered microspheres and real-time PCR reagents. By measuring the fluorescent emission of the PCR product in droplets using flow cytometry, they demonstrated that single-bacteria-resolution analysis can be achieved. However, all of these devices need complicated microfabrication or/and designs which required specific instruments. Very often these instruments or expertise for fabrication are not readily available for most biological laboratories, which has limited the application of these devices for single cell studies.

In order to finally develop a chip-level device to achieve single cell RT-qPCR, this research developed a chip-level device which involves only inexpensive and easily accessible equipment that is capable of performing single cell qPCR as the first step. This device contains an array of stationary, surface-adhering droplets immersed in oil as real-time PCR chambers. Mineral oil is used to isolate the droplets and prevents the aqueous

solution from evaporating during thermal cycling. The dimensions and locations of the droplets are controlled by hydrophilic patterning on the glass substrate. The volume of each chamber is 5 μL on the current design, but can be smaller and the density of the droplets can be higher if using a customized thermal cycler that can scan dense PCR chamber arrays. In addition, the operation does not require off-chip DNA extraction or purification steps which will diminish the potential effectiveness of downstream real-time PCR analysis, such as inhibition from lysis buffer which is used for off-chip DNA extraction. Although single-cell loading can also be easily achieved by serial dilution (Lin, Chao, and Meldrum 2009), we used a micromanipulator to precisely load one bacterial cell per droplet to validate the sensitivity. Leveraged with a commercially available real-time PCR thermal cycler, it was demonstrated that the device is capable of genetic analysis at the single cell level.

4.2. Experiments

4.2.1. Experiment setup

The current chip is designed to be compatible with an off-the-shelf real-time PCR thermal cycler originally designed to work with conventional PCR tubes/plates. The chip contains an array of surface-adhering droplets on a microscope cover slip, and all droplets are submerged under an open pool of mineral oil confined by a PDMS frame (Figure 8a). Each droplet is isolated by mineral oil as a PCR reaction chamber and to prevent the droplets from evaporation and cross contamination. The chip is optically compatible with the thermal cycler since all components such as mineral oil and droplets are transparent. The locations of the droplets are aligned with the wells of the heating block designed to hold PCR tubes, and the area of a droplet is of the same order as the

top-view area of a PCR tube (Figure 8b). Therefore, the existing fluorescence detection configuration can be directly applied to the use of the chip. A 0.42-mm-thick brass plate was placed under the cover slip during thermal cycling as a mediator to enhance heat transfer for better temperature uniformity on the chip. The cross view of the chip in a typical real-time PCR thermal cycler is shown in Figure 8b.

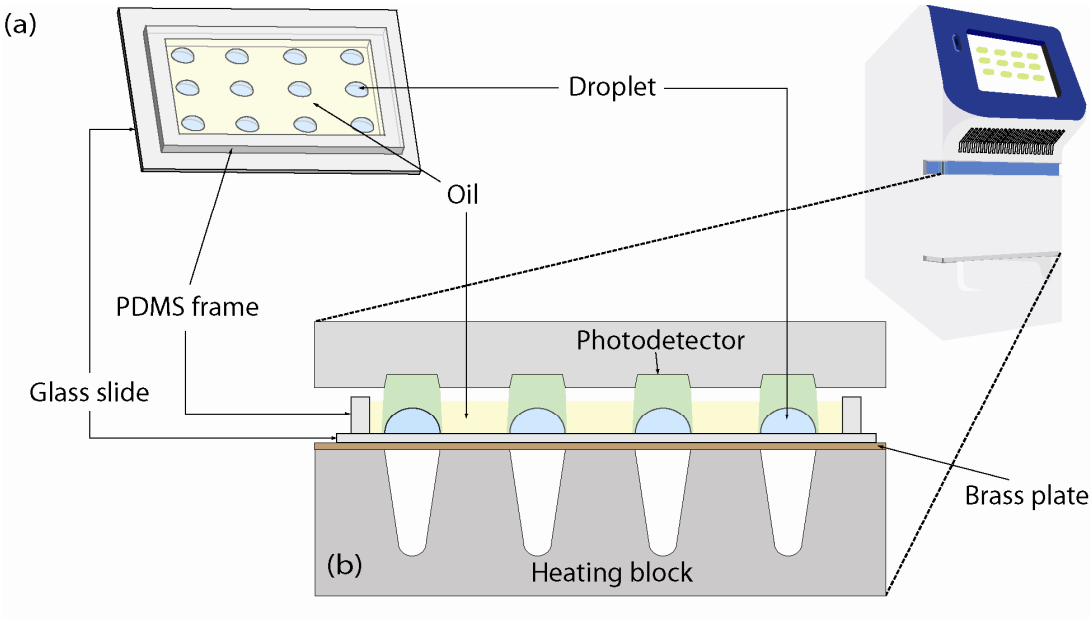

Figure 8. (a) The chip contains an array of surface-adhering droplets submerged in oil. **(b)** The cross section view of the chip placed on a thermal cycler, showing that the droplets are aligned with the wells of the heating block of the thermal cycler (not to scale).

In order to implement single cell analysis, the reaction volume of each chamber was designed to be 5 µL, in contrast to the typical 10-20 µL reaction volume using conventional PCR tubes. Droplets which are smaller than 5 µL tended to be too small for the real-time PCR thermal cycler to detect. The smaller reaction volume increases the local concentration of the template so that the competition with contaminants or endogenously generated background such as primer dimers will be reduced thus

providing more DNA polymerase molecules per template (Musyanovych, Mailänder, and Landfester 2005; Marcy, Ishoey, et al. 2007; Schaerli and Hollfelder 2009).

4.2.2. Chip fabrication

The procedure of producing a droplet array is derived from previous work that isolates single bacteria in a droplet array (Lin, Chao, and Meldrum 2009). The key fabrication process is to make a hydrophilic pattern that confines the aqueous droplets. In this study, I generated such patterns using Microscale Plasma Activated Templating (μPLAT), a technique that employs a stencil to expose air plasma only to designed areas to increase the hydrophilicity of the surface (Chao, Carlson, and Meldrum 2007). The process of making this chip is illustrated in Figure 9. In order to minimize contamination, all components and material used in this study were autoclaved and exposure to UV light for 15 minutes before the experiment.

First, a μPLAT stencil made of a 2-mm-thick PDMS sheet was adhered on a cover slip (Figure 9a). A 3×4 array of 1/8-inch diameter holes was punched, with a 9-mm pitch between centers to align with the real-time PCR thermal cycler. Hence, twelve droplets can fit in a 35mm×50mm cover slip (Fisher Scientific, Pittsburgh, PA). The soft PDMS stencil was then adhered on the cover slip. The assembly was placed in plasma cleaner (Harrick Plasma, Ithaca, NY) for plasma exposure with 6.8 W RF-power for one minute (Figure 9b). The areas exposed to the plasma became more hydrophilic, while the unexposed areas remained unchanged. Then the stencil was removed, leaving an array of hydrophilic circular areas on a more hydrophobic background (Figure 9c). A 2-mm-thick PDMS frame was placed to surround the hydrophilic array to confine a pool of oil (Figure 9d). 800 μL of mineral oil (Sigma M8410) was loaded inside the PDMS frame

(Figure 9e). Finally, twelve 5 µL droplets of PCR mixture were pipetted on each hydrophilic area (Figure 9f). Oil was loaded before the droplets to prevent contamination during the loading process.

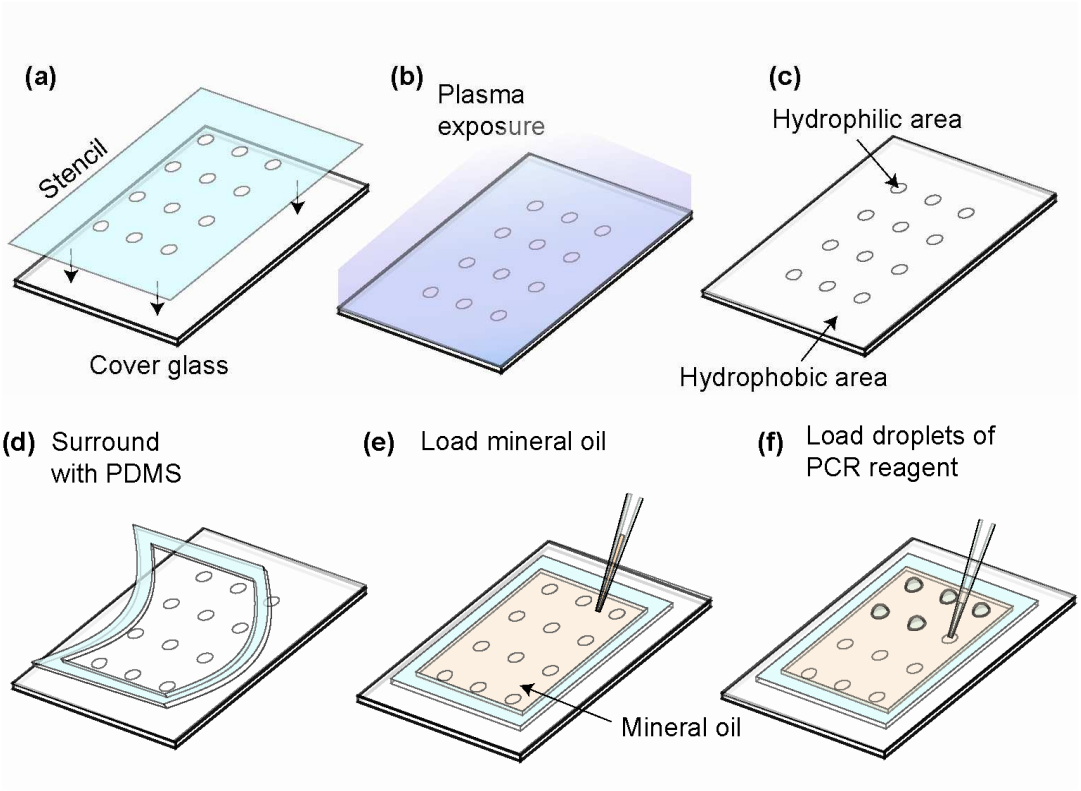

Figure 9. Chip fabrication process.

4.2.3. Strain and Cell culture

In order to prove the single cell sensitivity of the chip, initially the *Synechocystis* PCC 6803 strain was used as a target. The *Synechocystis* PCC 6803 cells were grown at room temperature in BG-11 media (Richaud et al. 2001). The cell density of *Synechocystis* PCC 6803 was measured by a spectrophotometer (Beckman Coulter, Brea, CA), and counted under light microscopy (Nikon, Japan). In general, OD_{730} 1.0 represents approximately 10^8 cells/mL.

4.2.4. Cell loading

Two cell loading strategies were used in this study. The first was serial dilution to load, on average, 1000, 100, and 10 cells per droplet. The second strategy was single-cell loading using a micromanipulator developed in our center (Anis et al. 2011; Anis, Holl, and Meldrum 2010). This micromanipulator uses a piezoelectric actuated diaphragm to dispense/aspirate liquid through a 30-µm capillary at the picoliter level. This device can precisely manipulate cells with small flow rates and therefore gentle shear stresses. The entire loading process was monitored on a microscope, so the loading of single bacteria cell can be visually confirmed.

4.2.5. Real-time PCR

Primers were designed using Primer3 software (Untergasser et al. 2007) and manufactured by Invitrogen (Carlsbad, CA). One set of primers was designed to amplify a 152 bp of 16S rRNA gene of *Synechocystis* PCC 6803: forward primer (CCACGCCTAGTATCCATCGT) and reverse primer (TGTAGCGGTGAAATGC-GTAG). The SYBR GreenER qPCR SuperMix Kit (Invitrogen, Carlsbad, CA) was used for real-time PCR. The PCR reaction mixture contained 2.5 µL qPCR SuperMix, 0.5 µL of each primer with the final concentration of 4 µM, 0.5 µL of 5×BSA, 0.45 µL of DEPC treated water (Ambion, Austin, TX), 0.05 µL of ROX and 0.5 µL of sample in a total volume of 5 µL for each droplet. For a single cell droplet, 0.5 µL of DEPC treated water was loaded and then one single cell was put into the droplet using the micromanipulator. Considering the low fluorescent signal due to the low amount of target, photo bleaching was prevented by blocking the ambient light. The experiments were performed on a

commercially available thermal cycler (StepOne real-time PCR system, Applied Biosystems, Foster, CA).

4.2.6. PCR validation

After real-time PCR, the PCR products were pipetted out one by one and loaded on 1.5% agarose gels (EMD Chemical, Gibbstown, NJ) for electrophoresis analysis. The gels were run under 130 volts for 35 min. The DNA fragments with the expected size were then isolated using QIAquick Gel Extraction Kit (Qiagen, Valencia, CA) and analyzed by sequencing on ABI 3700. In order to confirm the right sequence, online nucleotide blast tools (http://blast.ncbi.nlm.nih.gov/) were used. Over 98% identity was recognized as right amplification. Validation is not necessary for regular utility of the device.

4.3. Results

4.3.1. PCR temperature profile

The efficiency and specificity of PCR are affected by several factors, including cell lysing efficiency, chemical constitution of the PCR system (*i.e.* primer concentration, Mg^{2+} concentration, and SYBR Green concentration, etc.) and annealing temperature (Sipos et al. 2007; Markoulatos, Siafakas, and Moncany 2002). For real-time PCR, the temperature at which fluorescence detection is performed is also a crucial factor.

Thermal lysing was selected over chemical lysing to avoid possible interference with PCR due to chemical lysing (Lu, Schmidt, and Jensen 2005; Lee et al. 2005; Waters et al. 1998). In this study, we found that heating at 94 °C for 10 minutes was enough to fully lyse the *Synechocystis* PCC 6803 cells in the droplets. Based on the sequence analysis and initial tests, 60 °C was found as the optimized annealing temperature, and was selected for rest of the study.

SYBR green dye only binds to double-stranded DNA in a specific temperature range, so the fluorescence emission that indicates the quantity of double-stranded DNA can be detected. The proper detection temperature can be determined through the melt curve analysis. After testing multiple temperatures ranging from 70 °C to 80 °C on chips, the optimized signal detection temperature was determined to be 72 °C, and was selected for the rest of the study.

As shown in Figure 8, the heat from the heating block of the thermal cycler transferred to the droplet through the air in the heating block, the brass plate and the cover slip. Since the droplets did not directly contact the thermal cycler, the temperature of the droplets experienced hysteresis and delay to the setting temperature of the thermal cycler. To solve the issue, initially I put my efforts on minimizing the difference of temperature related to each steps (i.e. cell lysing, annealing and signal detection) between the ideal temperatures and the real temperatures in droplets. The ideal temperature protocol for this study was 15 s at 95 °C for denaturing, 15 s at 60 °C for annealing, 30 s at 72 °C for extension, and 10 s at 72 °C for signal detection. In order to compensate for the offset and hysteresis of the real temperature in the droplets, a calibration was performed by inserting a 0.076 mm-diameter K-type thermocouple (5SC-TT-K-40-36, OMEGA, Stamford, CT) into the center of a droplet to empirically adjust the setting temperatures and corresponding durations of the thermal cycler to fit the actual temperatures in the droplets. A thermocouple reader (50 Series II, Fluke, Everett, WA) was used to record the temperature every five seconds during thermal cycling. Figure 10 shows the temperature profile set on the thermal cycler (black circles) and the compensated temperature profile in the droplets (line with rectangular dots). The

heating/cooling rates were longer than those for conventional in-tube PCR, so the droplets required a longer time to reach a steady-state temperature. To shorten the total duration, we selected not to wait for the temperature to reach steady state of each stage. Instead, the thermal cycling profile was selected such that the droplet temperature was maintained within a ±1 °C range from the desire temperature during each stage.

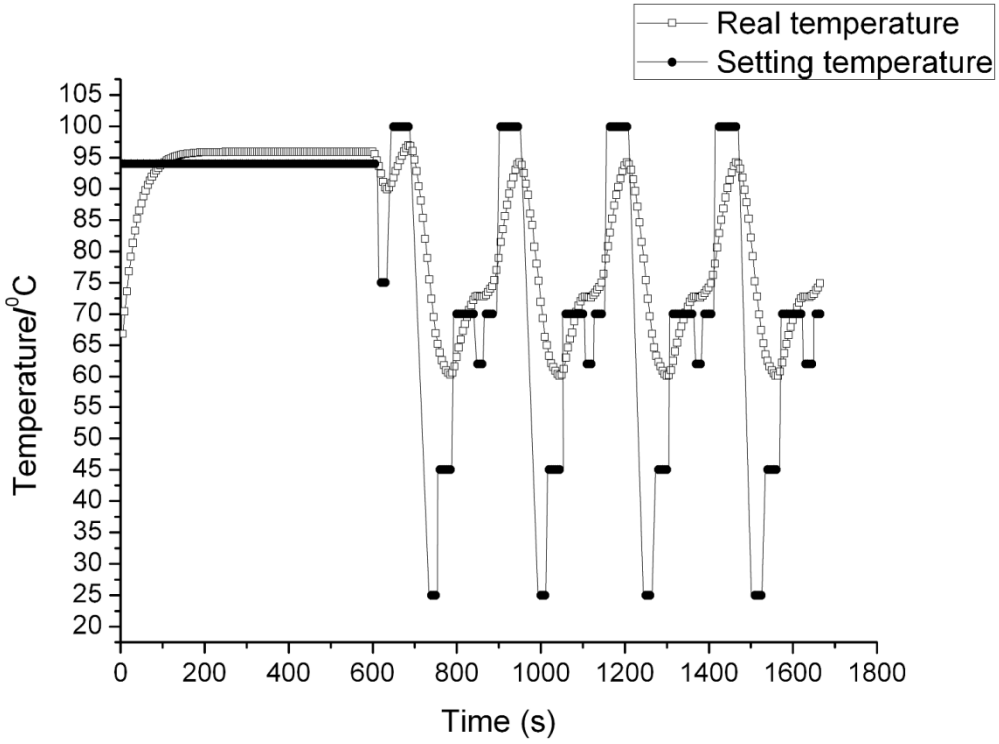

Figure 10. Temperature profile for PCR.

4.3.2. Single bacterial cell analysis

In order to accomplish real-time PCR with single digit template copy in single cells, in addition to the temperature profile, primer design is also a crucial factor (X. Wang and Seed 2003; Pattyn et al. 2003). The criteria for the primer design used in this study is that the size of the amplicon should be around 200 bp, T_m value of the primers should be

around 60 °C and several sets of primer need to be tested before choosing the right primers. Based on these requirements, the primers used in this study were designed by Primer3 (http://frodo.wi.mit.edu/primer3/). BSA was added in the solution to prevent undesired binding of DNA and polymerase to glass surface (Höss and Pääbo 1993; Höss et al. 1992; Kreader 1996; Prakash, Amrein, and Kaler 2007).

The real-time amplification curves of an on-chip PCR experiment with various cell numbers are shown in Figure 11. Four levels of cell numbers on the chip were tested. 1000, 100 and 10 cells per droplet were achieved by serial dilution from bulk cells, while single cells were picked and loaded directly with the micromanipulator. The average C_q values were also shown in the insert of Figure 11. The curves are well clustered for each cell number levels. The differences of the average C_q values between the cell number clusters were 4.2, 2.7 and 2.3, respectively. All amplified DNA was confirmed by sequencing to be the expected products. Although the negative controls were frequently amplified in the experiments, they appeared significantly later than the reactions from single cells. In addition, the T_m values of negative controls are different from that of the template. We also sequenced the amplification products from the negative control, and the BLAST search showed that the products were different, which could be a result of random amplification (data not shown). The possible causes for the amplification of negative controls are: 1) the contamination carried in commercial kits, enzyme and buffer. This contamination has been reported in many studies (Zhou et al. 2007; Panicker, Myers, and Bej 2004); 2) random amplification when the cycle number is high.

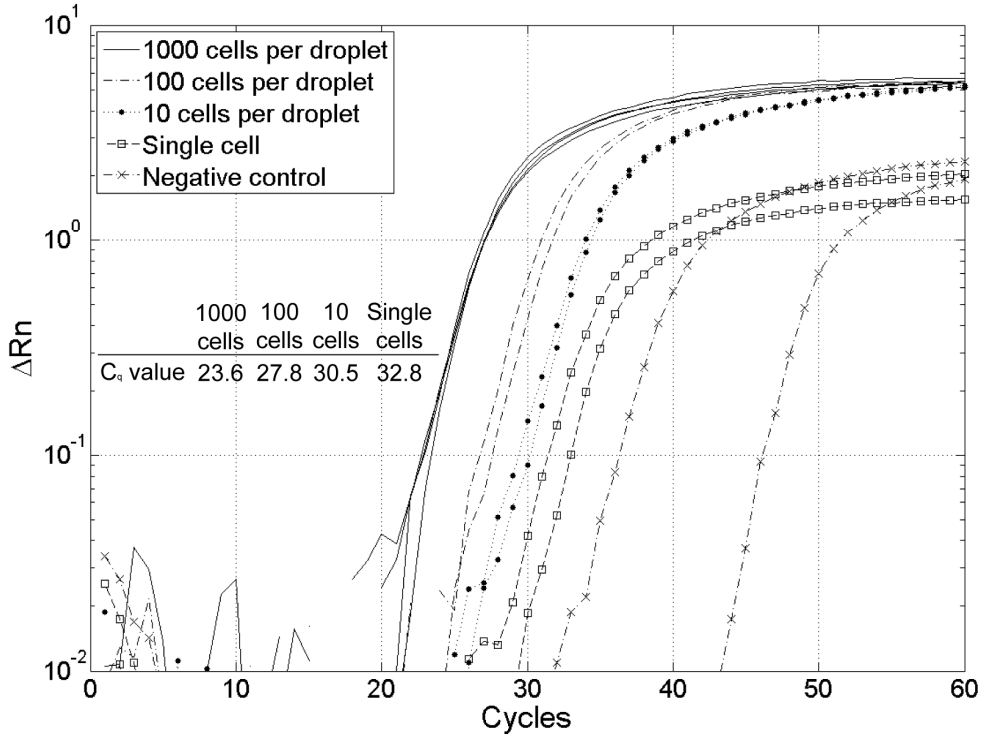

Figure 11. Real-time PCR result at 1000, 100, 10, and single cell levels.

4.3.3. Performance evaluation

In order to evaluate the performance of the chip, we first determined the successful rate of the chip operation. The melt curves were analyzed to define whether the reactions were successful. Briefly, each amplicon should have a specific T_m value, then melt curves with dominant signal at the right T_m value were recognized as a successful amplification. In addition, PCR products were validated through gel analysis and sequencing analysis. Based on these criteria, the overall success rate for the experiments shown in Figure 11 and Figure 12 was over 85% and the single cell level success rate is in the same range.

The means and standard deviations of the C_q values of different 16S rRNA template concentrations are summarized in Figure 12. In these experiments, four different

concentrations of template were loaded on four chips respectively. Each *Synechocystis* PCC 6803 cell contains femtogram-level DNA per cell (Hahn et al. 2000), so 10 pg DNA is equivalent to 1000 *Synechocystis* cells. The error bars of the C_q are larger for lower concentrations, indicating that the reproducibility of the chip-based real-time PCR decreased as the template concentration decreased. One possible reason of the decreased reproducibility could be due to signal detection for the low amount of amplified DNA. Unexpected DNA binding to glass surfaces may be another possible reason for the decreased reproducibility.

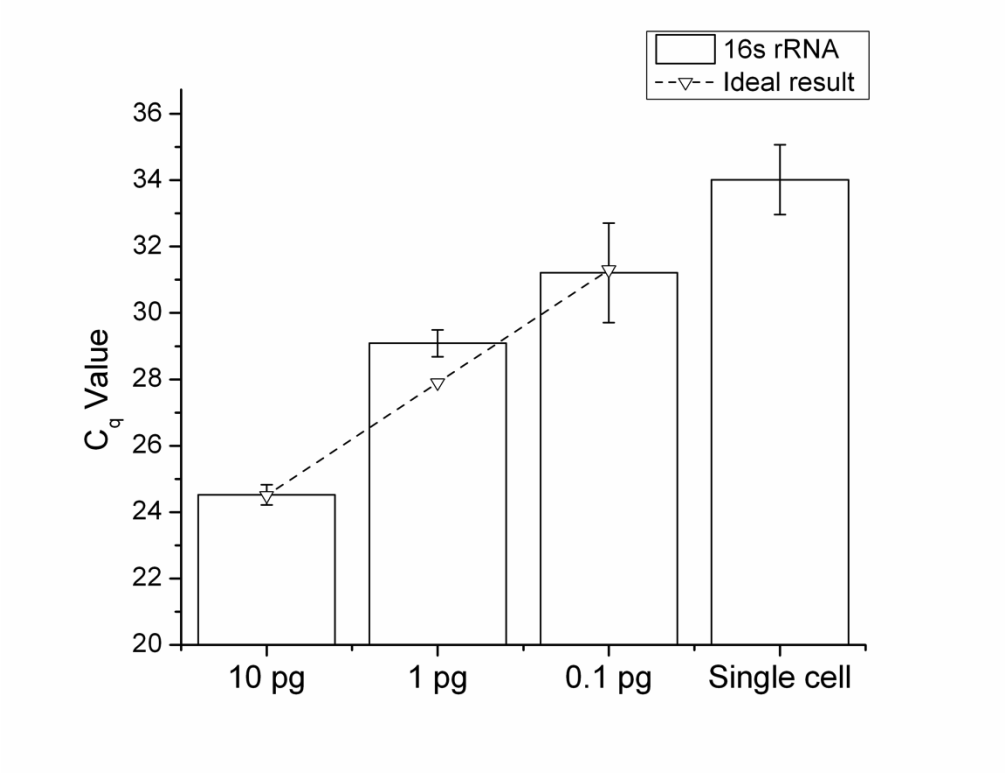

Figure 12. Template concentrations and C_q values for on-chip PCR experiments.

Between each 10x serial dilution, the ideal C_q value difference is around 3.324 (i.e. $\log_2 10$) cycles. Our experimental results showed that the difference between each dilution level is 4.6 and 2.1 cycles between 10 pg to 1.0 pg and 1.0 pg to 0.1 pg dilutions,

respectively, and 6.7 cycles between 10 pg to 0.1 pg dilutions. This result showed that although variations existed for individual experiments, the global trend of the relation between C_q value and template concentration was close to the estimation with the ideal efficiency. We also designed another set of primers to amplify a 198 bp fragment of *rbcL* gene of *Synechocystis* PCC 6803. The C_q difference between 10 pg and 1 pg was 3.2. The result confirmed that the PCR efficiency of our device was robust.

In addition to successful rate and efficiency, sensitivity is another crucial factor for single cell studies since the analyte amount is extremely low (Zare and Kim 2010; Borland et al. 2008; Schmid et al. 2010). One *Synechocystis* PCC 6803 cell has only 2 copies of 16S rRNA gene, according to the NCBI and CyanoBase database (http://www.ncbi.nlm.nih.gov/gene; http://genome.kazusa.or.jp/cyanobase). Therefore, the sensitivity of the presented work approached the single copy level, similar to the sensitivities achieved with 6.25-nL microchambers (Ottesen et al. 2006) and 70-pL droplets (Beer et al. 2007) in previous work. However, because the droplet volume in our device was in the μL-scale, our sensitivity in terms of initial template concentration is much higher. We assert that the one-step operation conserved the small number of templates in the confined droplet volume, and the reduction in liquid transportation also minimized possible contamination which allowed for high thermal circle numbers with acceptable negative control expression. The elimination of DNA extraction and purification did not prevent quantitative analyses at low template concentration.

4.4. Conclusion

In this study, a new design of an easily fabricated multi-chamber real-time PCR chip was demonstrated. The chip was robust and cost-efficient, and the one-step operation does not

require DNA purification. The current chip can analyze twelve single cells in one experiment, constrained only by the commercial thermal cycler. Using this new device, we successfully extended qPCR analysis of gene targets toward the single copy level. Through serial dilution at low template concentration, statistics of C_q values results in the ideal estimation. With a specifically designed thermal cycler for this chip, it will have the capability to further decrease the reaction volume to nL volumes, similar to the other PCR microdevices that use non-adhering droplets as reaction chambers. The application of this device in biological laboratories will provide the needed and convenient tools to perform genetic analysis for single cells and reveal heterogeneity in complex microbial communities. Meanwhile, these results also build a strong foundation for the single cell environmental monitoring device construction in this thesis.

5. TWO-STEP SINGLE-CELL RT-QPCR PROTOCOL DEVELOPMENT

5.1. Introduction

This portion of work has been published in Applied Environmental Microbiology (Shi et al. 2013). I would like to thank Dr. Gao, Dr. Chao, Dr. Zhang and Dr. Meldrum. With their patient supervision and suggestions, I could successfully design the experiment and achieved good results for this research.

For isogenic cell populations, gene expression heterogeneity could arise from the intrinsically stochastic processes of the expression of individual genes. The amplitude of such stochasticity, or noise, in gene expression is controlled by many factors, including transcription rate, regulatory dynamics, and other genetic factors of the cells (e.g. microRNA, transposon etc.) (Banerjee et al. 2004; Colman-Lerner et al. 2005; Pedraza and van Oudenaarden 2005; Rosenfeld et al. 2005; Newman et al. 2006; Strovas et al. 2007). As a result, individual cells in genetically homogeneous populations can contain different copy numbers of messenger RNA (mRNA) molecules, which eventually leads to different numbers of functioning protein molecules. This transcriptional noise, once amplified, could offer the opportunity to generate and sustain heterogeneity at the cellular level in a clonal population. The gene expression heterogeneity suggests that by simply averaging mRNA or proteins from whole populations, crucial information about unique patterns of the gene expression related to specific regions or distinct functional subpopulations may be lost. To gain a deeper insight into the intricacies of cellular diversity and its functional relevance, single cell level analysis needs to be performed. For the purpose of deciphering interesting biology puzzles, conventional tube-based single cell level gene expression analysis was performed targeting *T. pseudonana*, a

typical centric diatom (Armbrust et al. 2004), using single cell RT-qPCR on the basis of previous efforts (J. Zeng et al. 2011; Gao, Zhang, and Meldrum 2011).

It is well known that under adverse conditions such as nutrient-limited or other environmental stresses, microorganisms can trigger protective response mechanisms for survival. Concurrently, many regular physiological activities such as photosynthesis may be repressed under these stresses. Directly monitoring the stress response of microorganisms to their environments could be one way to inspect the health of microorganisms themselves, as well as the environments in which they live. Under such situations, pursuing analysis methods targeting a few or single microbial cells, which are directly recovered from environments without further cultivation, is necessary. Diatoms are a group of unicellular phytoplankton (Falkowski et al. 2004; Thamatrakoln et al. 2012) that are present in wide spread niches, from inland lakes to open oceans (Bennett et al. 2010; Mann and Droop 1996). Because of this, there is no need for introducing foreign species (Ripp et al. 2000) to monitor the environment. Other than using *T. pseudonana* as a sensor, they play significant roles in the global carbon cycle (Maheswari et al. 2010; Nelson et al. 1995) that makes it essential to understand what environmental stresses they are susceptible to and how they respond, in order to maintain the primary productivity in oceans.

Nitrogen is an essential element for living organisms and is required for the biosynthesis of macromolecules such as amino acids. It has been reported that the availability of nitrogen in oceans varies drastically on spatial and temporal scales due to physical and biological processes, and nitrogen has been considered as a major limiting nutrient for primary production in the oceans (Falkowski et al. 2004; Hockin et al. 2012).

Phosphate is another important element involved in many aspects of cellular metabolism, like ATP synthesis. It was reported that photosynthesis was disrupted by low level phosphorus (Bucciarelli and Sunda 2003; Rao, Arulanantham, and Terry 1989). Iron is a key component of Ferredoxin, an iron-sulfur proteins that control electron transfer (Abdel-Ghany et al. 2005), and its limitation and restriction of primary productivity have been reported for some ocean regions (Boyd et al. 2007; Lewandowska and Kosakowska 2004). Because of the short residence time of bioavailable iron (Martin et al. 1994) and the extremely low concentration of iron in the surface water which is only 0.07 nM/kg (Johnson, Gordon, and Coale 1997), phytoplankton growth and primary productivity are restricted in vast high-nutrient, low-chlorophyll (HNLC) regions of the Southern Ocean, the equatorial Pacific and the North Pacific (Boyd et al. 2007; Marchetti et al. 2012).

In this study, in contrast to previously published single-cell analyses on mammalian cells, working with diatoms has its own particular challenges due to their small size (~5 μm diameter) and protective frustules. In this research, the expression of six genes in single *T. pseudonana* cells was quantitatively measured, each with three technical replicates. The single-cell results revealed significant heterogeneity in terms of stress responses within *T. pseudonana* population. This work demonstrated the possibility of applying native habitants as sensors to monitor the environmental stress conditions. Meantime, this study provided the first quantitative gene expression evidence for the response heterogeneity of diatom *T. pseudonana* to environmental stresses.

5.2. Experiments

5.2.1. Cell culture

T. pseudonana (CCMP1335) cells were obtained from the National Center for Marine Algae and Microbiota (NCAM), and were grown in f/2 medium at 24 ± 1°C (Guillard 1975; Guillard 1962) under a constant light condition (30 µmol photons m^{-2} s^{-1} irradiance measured using LiCor (Lincoln, NE)). Cells at middle exponential phase were harvested by centrifugation at 1,500 × *g*, for 5 min at 4°C, and used to inoculate f/2 medium with or without nitrogen ($NaNO_3$, 8.82 × 10^{-4} M), phosphate (NaH_2PO_4, 3.62 x 10^{-5} M) and iron ($FeCl_3 \cdot 6H_2O$, 1.17 × 10^{-5} M) depending on the condition of starvation. Artificial seawater was prepared using chemicals of analytical purity and used instead of filtered nature sea water for f/2 medium and prepared based on the formula of Kester *et al.* (1967).

5.2.2. Sampling and RNA extraction

For bulk-cell based analysis, 1 mL cell culture was collected by centrifugation at 1,500 × *g* for 5 min at 4°C. Hemocytometer 3900 (Hausser Scientific, Horsham, PA) was used to count the cell number directly. RNeasy Mini kit (Qiagen, Valencia, CA) was used to extract RNA from the bulk cells. For single-cell based analysis, a micromanipulator developed in our center (Anis, Holl, and Meldrum 2010; Anis et al. 2011) was used to pick cells from a diluted cell population and load them into individual Eppendorf microtubes. This micromanipulator uses a piezoelectric actuated diaphragm to dispense/aspirate picoliter level liquid through a 30-µm capillary. Owing to its small flow rates, single cells suffer very little shear stress, which minimizes the effects on their gene expression profile. Thirty individual cells from each growth condition were picked. ZR Fungal/Bacterial RNA MicroPrep kit (Zymo Research, Irvine, CA) was used to extract

RNA from single cells and the total RNA was eluted into a final volume of 6 µL in Eppendorf microtubes.

5.2.3. cDNA synthesis

SuperScript VILO cDNA Synthesis kit (Invitrogen, Carlsbad, CA) was used to synthesize cDNA. For cDNA synthesis from bulk-cell RNA, total reaction volume was 20 µL containing 2 µL 10 X SuperScript Enzyme Mix, 4 µL 5× VILO Reaction Mix and 14 µL of eluted RNA. To increase the relative concentration of single-cell mRNA for cDNA synthesis preparation, total reaction volume was decreased to 10 µL which contains 1 µL of specific primer mixture, 1 µL 10 × SuperScript Enzyme Mix, 2 µL 5 × VILO Reaction Mix and 6 µL of eluted RNA. After cDNA synthesis, 10 µL DEPC treated water (Ambion, Austin, TX) was added to make the final volume of 20 µL before they were used as template for quantitative PCR analysis.

5.2.4. Quantitative PCR

Primers for RT-qPCR were designed using Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome). To differentiate PCR products from primer dimers, we selected primers which will generate amplicons with sizes around 170-220 bp (Gao, Zhang, and Meldrum 2011). qPCR was performed using Express SYBR GreenER qPCR SuperMixes Kits (Invitrogen, Carlsbad, CA) on a ABI StepOne Real-Time PCR System for bulk-cell analysis and ABI 7900HT Real-Time PCR System for single cell analysis (Applied Biosystems, Foster, CA), respectively. The temperature of qPCR was 10 min at 95°C for initial hot start, and 40 cycles with conditions as: 15 sec at 95°C for denature, 50 sec at 60°C for annealing and extension and 10 sec at 75°C for signal detection. There was also another melting curve analysis

step which was set to be the default condition based on the real time PCR system. For PCR reactions, 1 µL each primer with the concentration of 4 µM, 5 µL of master mixture, 0.1 µL ROX, 0.9 µL DEPC treated water and 2 µL cDNA were combined. Technical triplicates of PCR analysis were performed for each gene. Reactions without cDNA templates served as negative controls. Expression levels of target genes were normalized against internal control *actin* gene.

5.2.5. Data analysis

To describe the distribution variation of single-cell gene expression levels among cells, nonparametric statistic tests which do not require normal distribution of datasets were applied (Siegel 1957). Kolmogorov-Smirnov and Kruskal-Wallis analysis of variance (ANOVA) tests were used to analyze the relationship between four different groups of RT-qPCR measurements using the OriginPro 8.1 software (OriginLab Corporation, Northampton, MA). Principle component analysis (PCA) was conducted using the SPSS Statistics 20 package (IBM, Armonk, NY) to determine the possible control variances.

5.3. Results

5.3.1. Growth of *T. pseudonana* under stress conditions

T. pseudonana growth was determined by counting the cell number with a hemocytometer directly. Figure 13 showed the growth-time curves of *T. pseudonana* under control and three stress conditions. The results showed that the initial increase in cell numbers over days 1 to 4 were roughly exponential for all conditions although the growth under the no nitrogen and no phosphate conditions were at a relatively low rate. After day 4, the cultures under control and no iron conditions still maintained exponential growth for another 24 h. After day 5, both cultures of the control and no iron conditions

reached their stationary phase, while the cell numbers under the no nitrogen and no phosphate conditions declined. The results showed that all three nutrient-limited conditions caused significant decrease in cell growth, with phosphate and nitrogen affected the most. In these cases, cell number only reached 10-28% of that of the peak cell numbers. The slow growth of *T. pseudonana* under these stress conditions was consistent with previous reports (Bucciarelli and Sunda 2003; Lewandowska and Kosakowska 2004). Cells at the middle exponential phase were collected for RT-qPCR analysis (Figure 13).

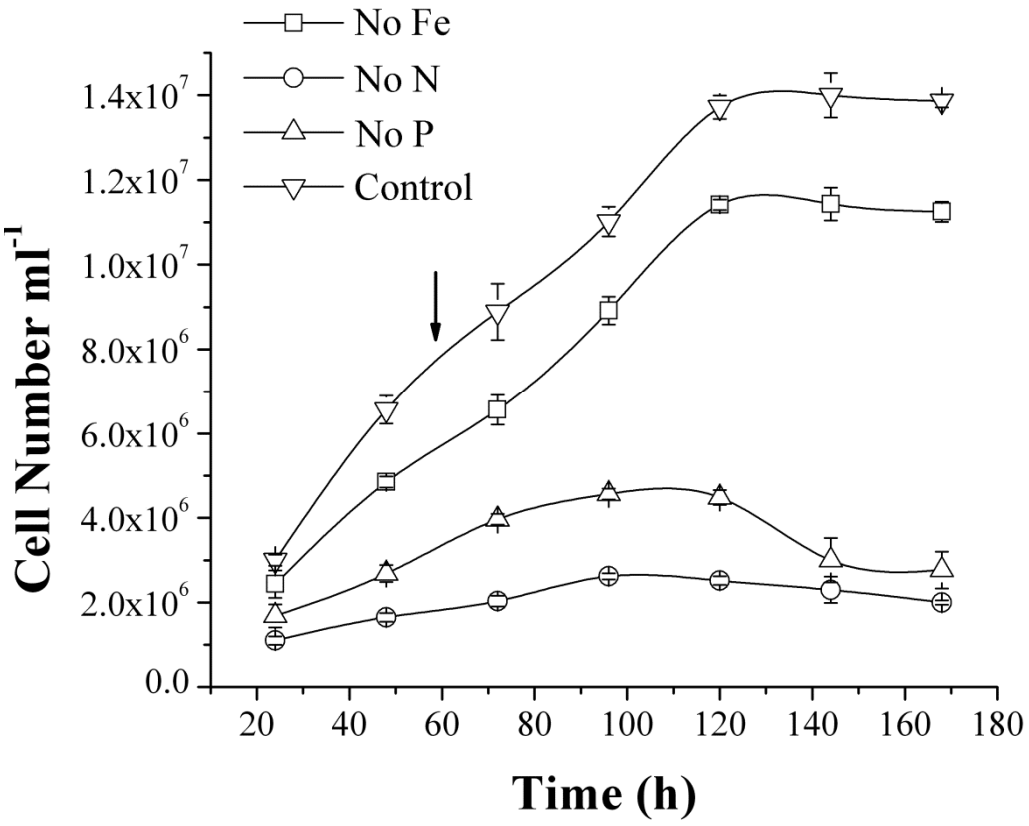

Figure 13. Growth of *T. pseudonana* cells under various conditions. Arrow indicates the sampling time for gene expression analysis.

5.3.2. Primer evaluation

A total of 82 pairs of PCR primers were designed and evaluated for 39 different target genes. Since the major goals of this study were to *i*) evaluate the possibility to use single cells as biosensors, *ii*) to determine the response heterogeneity of *T. pseudonana* to various important environmental factors (*i.e.* nitrogen, phosphate and iron limitation), and *iii*) also to compare the results with those previously obtained at bulk-cell level, the targets genes included some of the genes with demonstrated functions in photosynthesis, iron transportation and stress responses. Although most of the primers (78 out of 82) functioned well with bulk-cell RNA, only one pair of primers each was obtained for nine genes after the evaluation process (Figure 14). A relatively low success rate of primer selection reflected the different performance between bulk-cell based and single-cell based RT-qPCR analyses, and also the difficulty of measuring gene expression at the single-cell level. The successful primer sets and their corresponding gene targets were *psaA*, photosystem I P700 chlorophyll a apoprotein A1 (forward primer: CGGTTCTGCATCTTCAGCATACGGC , reverse primer: GTGCTAAACCAACGGC-ACGACCT); *psaF*, photosystem I reaction center subunit (forward primer: TGTGGCGCAGATGGCTTACCTC, reverse primer: TGCACTCGTACTTACTGCGC-GTA); *psbA*, photosystem II protein D1 (forward primer: CCACATGGCTGGTGTT-GCTGGT, reverse primer: CGACCAAAGTAACCGTGTGCAGCT); *psbC* (forward primer: TCATCTGCACAAGGTCCAACTGGT, reverse primer: AGCAGCACGACG-TTCTTGCCA); *psbC*, photosystem II reaction center protein (forward primer: TCATCTGCACAAGGTCCAACTGGT, reverse primer: AGCAGCACGACGTTCTTG-CCA); *hsp90*, heat shock protein (forward primer: AGGCTCTTACGGCCGGGGCGGA,

reverse primer: AAGACCCGCCAGCCTCGGA-AGCC); *rbcL*, ribulose-bisphosphate carboxylase (forward primer: AGGCTCTTACGGCCGGGGCGGA, reverse primer: TGTAGATAACTTGACGACCTGCGCC); *Actin* (forward primer: CCGTAGTGAA-CGCCTATCGTGGC, reverse primer: CCATCGTCTCGCTGCGGCTG); *Tubulin* (forward primer: GGACGCTACGTTCCTCGTGCC, reverse primer: GCTCTCGGCC-TCCTTCCTCACA); 18S (forward primer: TGCCAGTAGTCATACGCTCGTCTCA, reverse primer: CCTTCCGCGAACAGTCGGGTAT). The primers that functioned well at the bulk-cell level but not at the single-cell level are also provided in the Table 1.

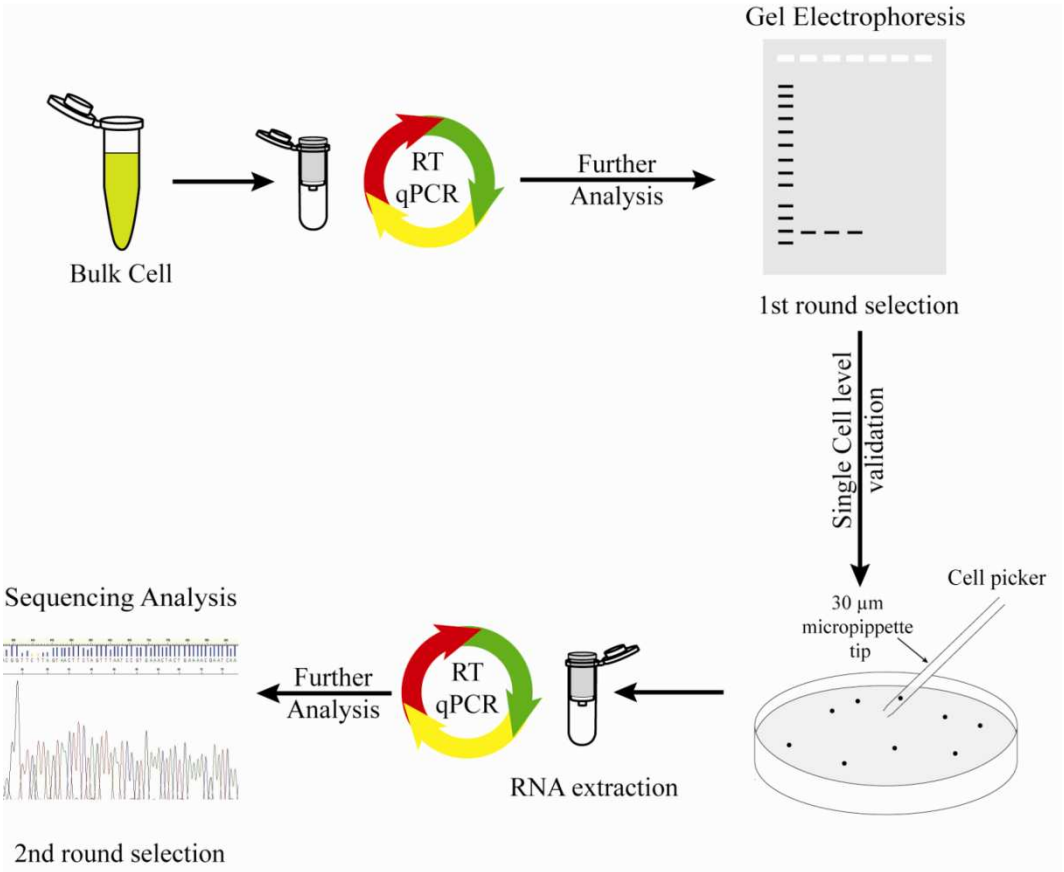

Figure 14. Two rounds of primer selection.

Table 1. Sequences of primers evaluated

Name	Sequence (5’ to 3’)
futA_forward	ACCTCTATGAGCGCTTCACC
futA_reverse	GTCGACCGTCTGGAACAGAT
ftrC_forward	TGACCTCTGTGGTGCTGAAG
ftrC_reverse	GGTTGTCTTCGGTGAGGAAA
phd-forward	CACGACCACACCTCATTCTG
phd-reverse	AAGGTCCGGTGTCAAAAGTG
zep-forward	AGGACGACCCAGAGGAGAAT
zep-reverse	ACCAATGAGGACGACGTTTC
twca$_{TP2}$-forward	ACGGTGACGGTCCCCACGGTAACATC
twca$_{TP2}$-reverse	ACCCACAGCAGAGGCGATATCCTGA
hsp20A_forward	GCCTGGCGGTAGATGTGCCCGGA
hsp20A_reverse	ACCCCATCAGCGAGATGGGCTGTG
p23_forward	CGGTGCAGATGCATGTGATGAAGGCGA
p23_reverse	CCATTCCGCCCATTCCTCCCATTCCCA

hsp70A_forward	GGCCACCAAGGATGCCGGAGCCA
hsp70A_reverse	TGGGTGTCTCCAGCGGTGGCCT
dnaJ1_forward	AGCGTGGGTGGCTGCGTCCGA
dnaJ1_reverse	TGCACACACCAATGGTCCCCCTAGTCC
dnaJ2_forward	ACCCGCCGCCCCAAAGACGGC
dnaJ2_reverse	TCGCCACGCCCGTCGCCAGC
pdz_forward	AGCCTTCGTCTGCGTCTCACCAGCCT
pdz_reverse	AGCGTCGGGAGATCGTCTGATGGGCG
pre_forward	AGGTCTCGTTGCCGCGGTTGCCG
pre_reverse	CGAATCCAGTTCCAACCGCACCCTTCG
smp1_forward	TGCCTCAGGGCGGTCTCGCCA
smp1_reverse	AGCCCTCCACCAGCCTTCAACTCCCT
cat_forward	TGGGGGTGACTGCAGGGGCGA
cat_reverse	AGCCGCCATCGATACCCCACCAGC
Actin 235F	ACCAACTGGGACGACATGGAGAAA

Actin 490R	TGTGGGTAACACCATCTCCCGAAT
Tubulin For	TTCGACCGGATAACTTTG
Tubulin Rev	CGACTAGTCAAAGGAGC
18S rRNA 1F	CTGCCCTATCAGCTTTGG
18S rRNA 1R	CGGCCATGCACCACC
18S rRNA 2F	TTGACTCAACACGGGAAAAC
18S rRNA 2R	ATCCAAAGCTGATAGGGCAG
FRE4F	AAAGTAGGCGACCGCACGGC
FRE4R	GGTGCGAGGGTGAGAGAAGCG
zupT For	TCCTCCTCGTGGTAGCCGCC
zupT Rev	CCTCCCCCAACTCAGCAGCCT
HMA1F	TGGCTTGAGGCACGAGCGAC
HMA1R	CTACCAACCCGTCTGCGGGC
CDF1F	TTCAGGTGGCGAGGTTGCCC
CDF1R	CAGCCGCTGCAATCCCCTGA

CDF2F	TGCAGCCAGCTGGTTTGGGG
CDF2R	TGCTGAAACGGCACAGTGGGT
HemeF	AGTGGGGGCAAGGTGCTGTCT
HemeR	GCCTCGGCTACCAAACGACCA
IscA1F	CGGTGGATGCTCCGGTCTTTCC
IscA1R	ACCCACAGCTCTCCTCGGCA
2Fe-2SF	AGGCATCGGTGAATCAGCCACA
2Fe-2SR	CGGTCGGTACAACGTCTACGCA
CytoF	AACGCAGCCTCCACACTGGC
CytoR	TCGTTCGTTCGCTGCCGTGG
RieF	TTCGCCACTCTTGCCTGCGG
RieR	AGGGCGGTGGTGGATCGCTA
Ctyob5F	ACACTGCAGACTCAGCGTGGA
Ctyob5R	ACGACGACCGTGAACATCGCC
HaemF	TCACGGTGCCAATGCCGGTC

HaemR	ACTGCTCGGGGGAAGTGGCA
SOD1F	AGTGAGCAAGTCGGCAGCGG
SOD1R	TAGCGTGCGGTGCGAGGTTG
SOD2F	GCCACCACCATCGCTACACCC
SOD2R	GCCCGGCCTCCAAAGCATTCA
ChrAF	CGGCGTGTACTGGCACCTCG
ChrAR	GGCGATGCTGCCTCCCAACA
FlavLF	CCGAAGCAGCCGCCGAAGAA
FlavLR	TCGTGCACCTAGAGGAGCTTGTCC
FlavF	GCGGTGGCGACGAGCTACAT
FlavR	CGTGTCGTCCCTGGCTGCAT
psbW_forward	ATCCCACTCCTCCGGACTCTGCATA
psbW_reverse	GGATCCTCCGGCATTGCCACATT
psbC_forward	GCCCAAAGGCCACCAATTTGACAC
psbC_reverse	GCTGCAGGTGGGTGGTTTACTGG

LhcA-forward	GACCTCCCTTGACGCCATGCC
LhcA-reverse	GACCGTTGGTCTCCTCCCACGAT
psaF-forward	GGATGGACGATGGAGTCATGCAGC
psaF-reverse	TCTTGCCACGCTGAAATTGGCCAA
rbcL-forward	ACGGTAGCGCTCACAAGCTGT
rbcL-reverse	TGGGTTACTGGGATGCTGCATACAC
rbL-forward	CGTGCATCTGCTGCAACTGGTG
rbL-reverse	TCGTTTTCACGAGCCCAGTAAGCAA
PPC1_forward	CCGTACCGCCCTTTCCGTGG
PPC1_reverse	CCGGCATACGTCGGAAGCTTGG
ACT1_forward	AAGCGGCTGAGGCTACGTCGAT
ACT1_reverse	GAGGCCATTCCGTCCAATCCACCA
rbcL2_forward	TCATGCGCTGCTGGTTACATCCG
rbcL2_reverse	GTAGATAACTTGACGACCTGCGCCT
ACT1_forward	GATTGTGGCTCCCCCGGAGAGG

ACT1_reverse	TCGAGTCTCCTCAAACCACGAGCC
ACT2_forward	AAGATTGTGGCTCCCCCGGAGAG
ACT2_reverse	TCGAGTCTCCTCAAACCACGAGC
psaF1_forward	GTGGCGCAGATGGCTTACCTCA
psaF1_reverse	TTGCACTCGTACTTACTGCGCGT
psbC2_forward	ATTCGTTCGTGGATGGCTGCACA
psbC2_reverse	CGAGCGTTTCCACTCCACCAAGC
tubulin11_forward	CTGCCGTGCAGGAGACCTGG
tubulin11_reverse	CCTTCCTCGTCGGCAGTTGCAT
tubulin21_forward	CGCCAAGCGTGCCTTTGTGC
tubulin21_reverse	CGCCAACAGCGCGACAGAGT
tubulin22_forward	TGGGCCAGGCCGGTATCCAA
tubulin22_reverse	CGCACACGGTGGGCTCCAAA
actin12_forward	TGGGCCGATCGATCATCCACTGTT
actin12_reverse	CAGGGCAGTTTGCCTCCCGT

actin21_forward	GGCGTCCAACGAGGAAGGCA
actin21_reverse	CGCTGTCGTAATGGCGGGGG
actin22_forward	GGTCTAGCTCCGCCAACCGGA
actin22_reverse	GCATGCTGCCACTGCATCCCT

5.3.3. Enhanced cDNA synthesis by adding target-specific primers

cDNA synthesis typically employs random primers which generate the least bias in the resulting cDNA (Stephen A Bustin and Nolan 2004). However, since we were using total RNA rather than purified mRNA as the starting template, most of the cDNA synthesized through the random primers will be ribosomal RNA-derived cDNA, which could further complicate the single-cell gene expression (Stephen A Bustin and Nolan 2004). To address this issue and to enhance the yield of cDNA derived from target mRNA, primers specific to the target genes were added into the reverse transcription reaction mixture so that more mRNA of the target genes would be converted to cDNA (Ståhlberg et al. 2004). To ensure detection sensitivity and reproducibility of single-cell qPCR, cDNA from each *T. pseudonana* cell was used to detect a maximum of three different genes, each with three technical replicates. In the cDNA synthesis step, 1 µL of primer mixture containing three target gene-specific primers (reverse primers which are complimentary to the mRNA sequence) was added. The final concentration of each target-specific reverse primer in the qPCR reaction is 4 nM. To demonstrate the effects of adding target gene-specific primers on single-cell analysis, we evaluated the single-cell RT-qPCR analysis of

three genes, *psbC*, *actin* and 18S rRNA genes. In this experiment, we diluted the RNA isolated from bulk cells (~10^6 cell/mL) to the level of a single cell which is approximately 50 fg/µL (Schmid et al. 2010). During cDNA synthesis, a primer mixture containing target gene-specific primers was added into 6 replicates, while another 6 replicates contain only random cDNA synthesis primers. The results showed that except for the 18S rRNA gene, addition of the target-specific primers can significantly decrease the C_q values by 2-4 cycles, which is 4-16 times higher yield of target cDNA when compared with control samples for *psbC* and *actin* genes, suggesting the target-specific primers in the cDNA synthesis reaction were able to improve the yield of target cDNA significantly (Figure 15). No effect was observed for the 18S rRNA gene, probably because that it is one of the most abundant genes in the total RNA (Valente et al. 2009). However, even for the 18S rRNA gene, our results showed that addition of target-specific primers can improve the qPCR reproducibility by decreasing the standard deviation of C_q values from 0.2 to 0.1 cycles (Figure 15). The results demonstrated that adding target-specific primers into the cDNA synthesis reaction mixture was a useful approach which can improve the performance of qPCR, especially for the genes with larger C_q values.

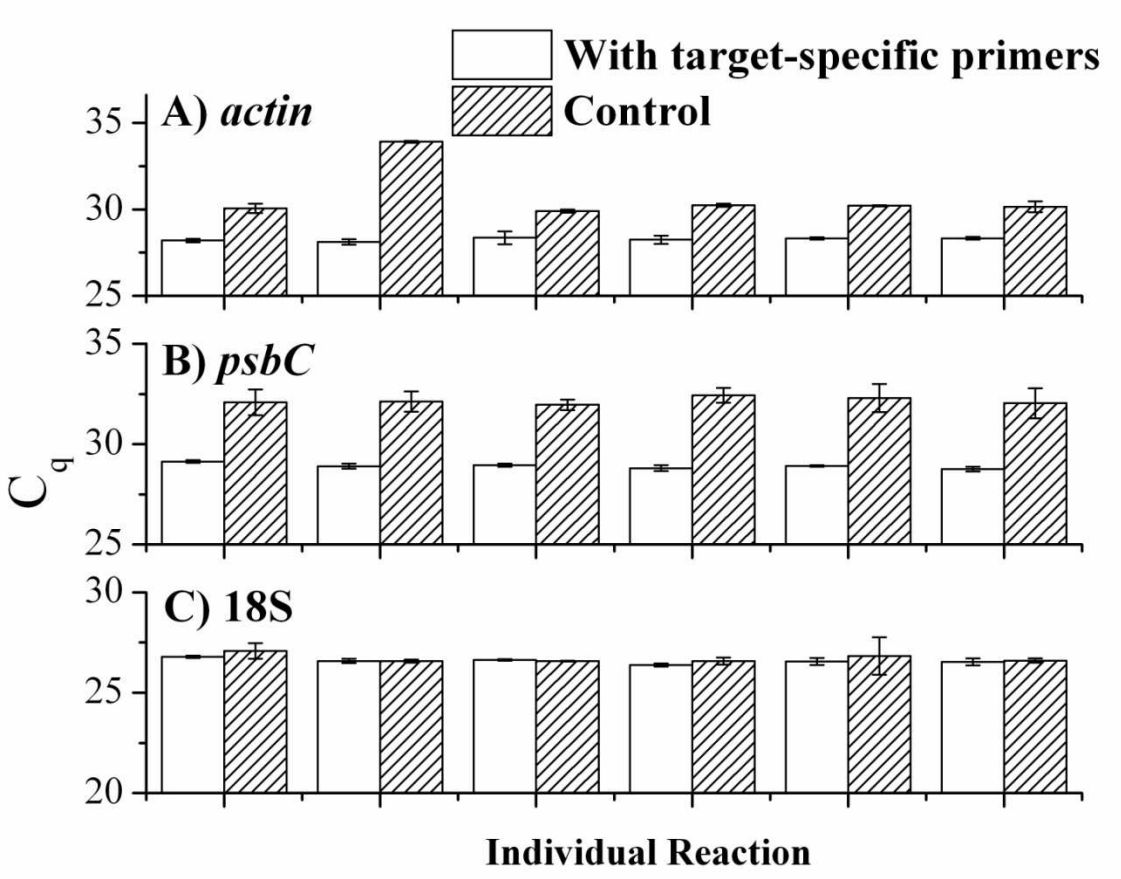

Figure 15. Effects of adding target-specific primers. C_q is quantification cycle, the fractional cycle number where fluorescence increases above the threshold.

5.3.4. Selection of internal reference gene

In order to ensure the gene expression across different conditions or analytical platforms quantitatively comparable, expression measurements need to be normalized against an internal reference gene (Heid et al. 1996). While several internal reference genes have been demonstrated in bulk-cell based RT-qPCR analysis, so far limited information is available regarding the constant expression of these internal reference genes across individual cells (Stephen A Bustin 2002; Stephen A Bustin et al. 2009; Huggett et al. 2005). For single-cell based analysis, relative activities of each target gene against reference gene were acquired by the $\Delta\Delta C_q$ method (Livak and Schmittgen 2001; Pfaffl

2001). Based on previous studies, we selected three genes, *tubulin* gene, 18S rRNA gene and *actin* gene (Goidin et al. 2001; Kim et al. 2003; Nailis et al. 2006) as candidate reference genes for further evaluation. To simplify the selection process, only control and no iron growth conditions were used. A total of 12 cells of control and no iron conditions were picked and subjected to expression determination of the *tubulin* gene, 18S rRNA gene and *actin* gene. The C_q measurements of a total of 24 cells (*i.e.* 12 control and 12 no iron conditions) are presented in Figure 16. The results showed that the standard deviation (StDev) of the C_q values for *tubulin* gene, 18S rRNA gene and *actin* gene among all 24 cells were 0.89, 2.9 and 0.39 cycles, respectively. The *actin* gene had the smallest variance among cells, and was thus selected as an internal control for our further analysis. The result was also consistent with that of Kustka *et al.* (2007) that the expression of the *actin* gene was constitutive under all iron concentrations (Kustka, Allen, and Morel 2007). The results also showed that even for 18S rRNA and *tubulin* genes which were widely used as internal controls in various bulk-cell based RT-qPCR analysis, significant cell-to-cell heterogeneity existed.

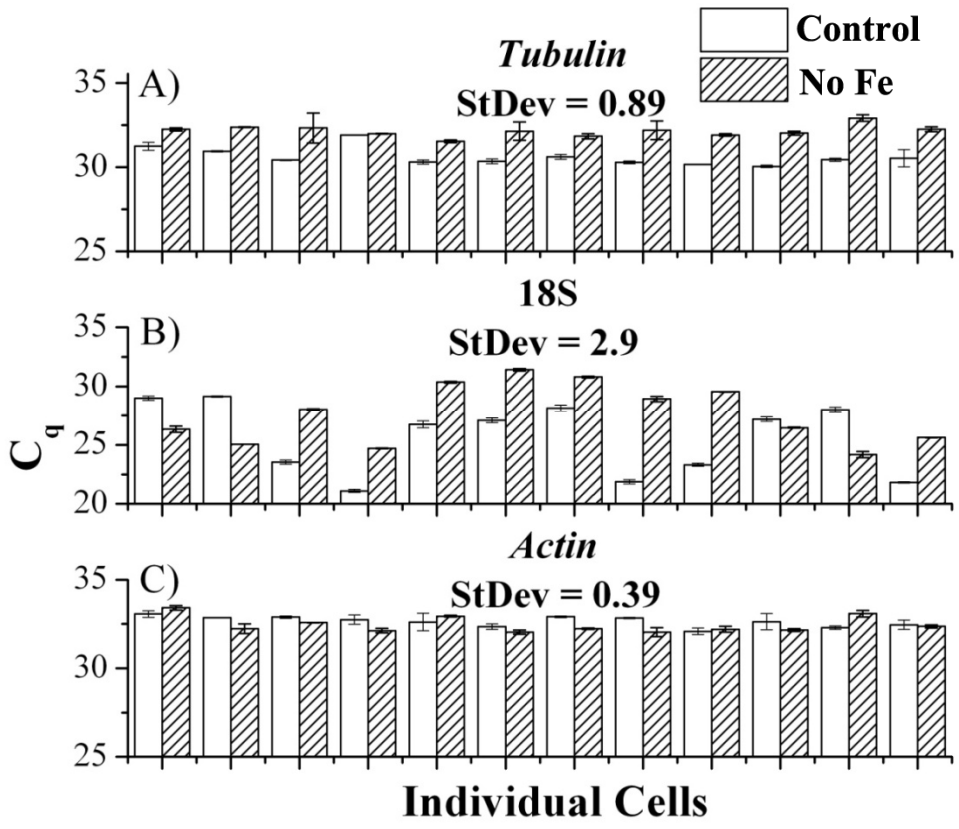

Figure 16. Evaluation of three internal control candidates under control and no iron conditions.

5.3.5. Gene expression under stress conditions

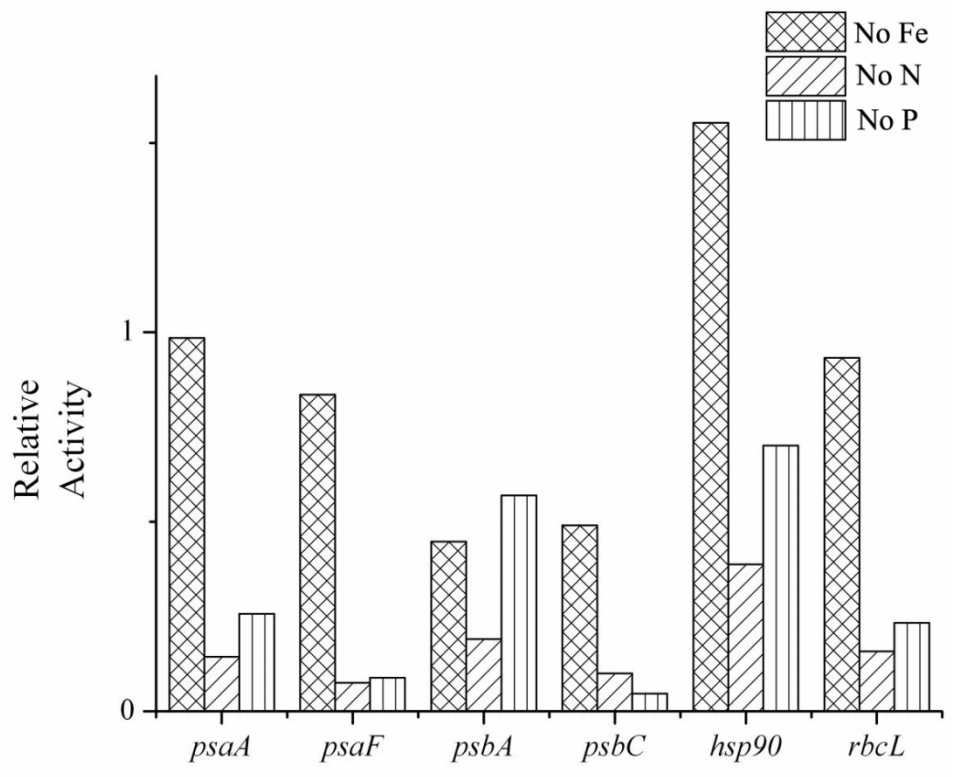

Figure 17. Relative gene expression activity normalized by control growth condition at bulk cell level. (The activity of each gene under control growth condition is equal to one.)

To establish a baseline for single-cell based analysis, we first performed a bulk-cell based RT-qPCR for the selected target genes under three stress conditions. The relative activity of each gene was derived from C_q value which normalized by cell number first and then by the activity of control growth condition. The results showed that except for the *hsp90* gene under no iron condition, all other genes were down-regulated by the stresses (Figure 17). Up-regulation of the *hsp90* gene under no iron condition was also reported by Thamatrakoln *et al.* (2012) who applied a combined genome-wide and targeted comparative transcriptomic analysis with diagnostic biochemistry and *in vivo* cell staining as a platform to identify the suite of genes involved in acclimation to iron

and associated oxidative stress in *T. pseudonana* (Thamatrakoln et al. 2012). In another study, Allen *et al.* (2008) also found that *hsp90* gene was up-regulated under an iron starvation stress condition in a pennate diatom *Phaeodactylum tricornutum*. Both the *psaA* gene encoding photosystem I P700 chlorophyll II apoprotein A1 and the *psaF* gene encoding photosystem I reaction center subunit were down-regulated under three nutrient-limited conditions. Similar results of PS I decrease under iron limitation were also reported by Allen *et al.* (2008). When compared with the *psaA* gene and the *psaF* gene of PS I, the *psbA* gene and the *psbC* gene of PS II were down-regulated more under all nutrient-limited conditions suggesting that photosystem II may be more vulnerable to nutrient-limited conditions than photosystem I, consistent with the results of Mock *et al.* (2008) who analyzed whole-genome expression profiling under several different growth conditions such as no Fe, no N, no Si and high temperature. The *rbcL* gene was down-regulated significantly under nitrogen starvation and no phosphate conditions, but only down-regulated slightly under no iron condition. In a recent study, Allen *et al.* (2008) reported that down-regulation of several proteins, such as phosphoribulokinase (PRK) and two enzymes supplying substrate for RuBisCO will lead to decrease of carbon fluxes toward RuBisCO under Fe stress in *P. tricornutum* (A. E. Allen et al. 2008). In addition, comparison of gene expression patterns showed that although *T. pseudonana* and *P. tricornutum* was divergent ~ 90 million years ago and had vast differences in genome structure (Bowler et al. 2008), but they may still share a similar fundamental response mechanism to iron starvation. Other than these results, Pearson correlation coefficients under different conditions (Tables 2-5) indicated that the *psaA* and *psaF* genes were always negative correlated under different nutrient-limited conditions, suggesting that the

two genes were regulated by a similar mechanism but opposite direction under different nutrient-limited conditions which was rational since both of them belong to photosystem I. However, for the genes *psbA* and *psbC*, no such correlation was found, which possibly suggests that the regulation mechanisms were different between photosystem II and photosystem I.

Table 2. *Pearson* correlation coefficients of no Fe condition

	psaA	*psaF*	*psbA*	*psbC*	*rbcL*	*hsp90*
psaA	1					
psaF	-0.15283	1				
psbA	0.6348*	-0.1411	1			
psbC	-0.01904	0.24029	-0.11868	1		
rbcL	0.23488	-0.32449	0.38733	0.47839	1	
hsp90	0.38389	-0.52269	0.6432	-0.18774	0.32095	1

* Red number indicated significant correlation

Table 3. *Pearson* correlation coefficients of no N condition

	psaA	*psaF*	*psbA*	*psbC*	*rbcL*	*hsp90*
psaA	1					
psaF	-0.59133*	1				
psbA	0.24673	-0.32845	1			
psbC	0.06086	-0.08748	0.06207	1		
rbcL	0.51667	-0.36102	0.10702	0.05714	1	
hsp90	0.69475	-0.63123	0.42323	0.20161	0.48914	1

* Red number indicated significant correlation

Table 4. *Pearson* correlation coefficients of no P condition

	psaA	*psaF*	*psbA*	*psbC*	*rbcL*	*hsp90*
psaA	1					
psaF	-0.46653	1				
psbA	0.3934	0.04633	1			
psbC	-0.18947	0.21417	0.13337	1		
rbcL	0.09046	-0.3235	0.05951	-0.05743	1	
hsp90	0.21243	0.36844	0.18594	-0.29835	0.47688	1

* Red number indicated significant correlation

Table 5. *Pearson* correlation coefficients of control condition

	psaA	*psaF*	*psbA*	*psbC*	*rbcL*	*hsp90*
psaA	1					
psaF	0.31514	1				
psbA	0.8602	0.38443	1			
psbC	-0.35067	-0.05957	-0.28209	1		
rbcL	0.30129	0.28349	0.47	-0.36159	1	
hsp90	-0.23108	-0.11722	-0.22576	-0.09846	0.35078	1

* Red number indicated significant correlation

For single-cell level analysis, $\Delta\Delta C_q$ method was adopted to calculate the relative expression of each gene against the reference *actin* gene. Figure 18 and Figure 19 showed the result of qPCR analysis of 6 genes under control and three stress conditions. For each condition, 30 individual cells were picked and analyzed. Reactions with large variation between technical replicates and/or with multiple peaks observed in the melting curves were considered as failed reactions and were excluded from further analysis. Overall the

success rate of qPCR reactions was approximately 93%. The reproducibility of the qPCR was derived from the StDev of the technical replicates of each cell. Based on our results, *hsp90*, *psbA*, *psbC* and *actin* genes were all with small average StDev values among all samples which were 0.2041 cycles (0.75% of average C_q values), 0.2109 (0.75%), 0.2116 (0.72%) and 0.2148 (0.74%), respectively. For the genes with larger C_q values, although the average StDev values were almost doubled to 0.3847 (1.2%), 0.4048 (1.2%) and 0.422 (1.3%) for *psaF*, *psaA* and *rbcL* genes, respectively, they were still in the relatively low variation rangers. In general, our single-cell level qPCR protocol was robust and able to generate reproducible data.

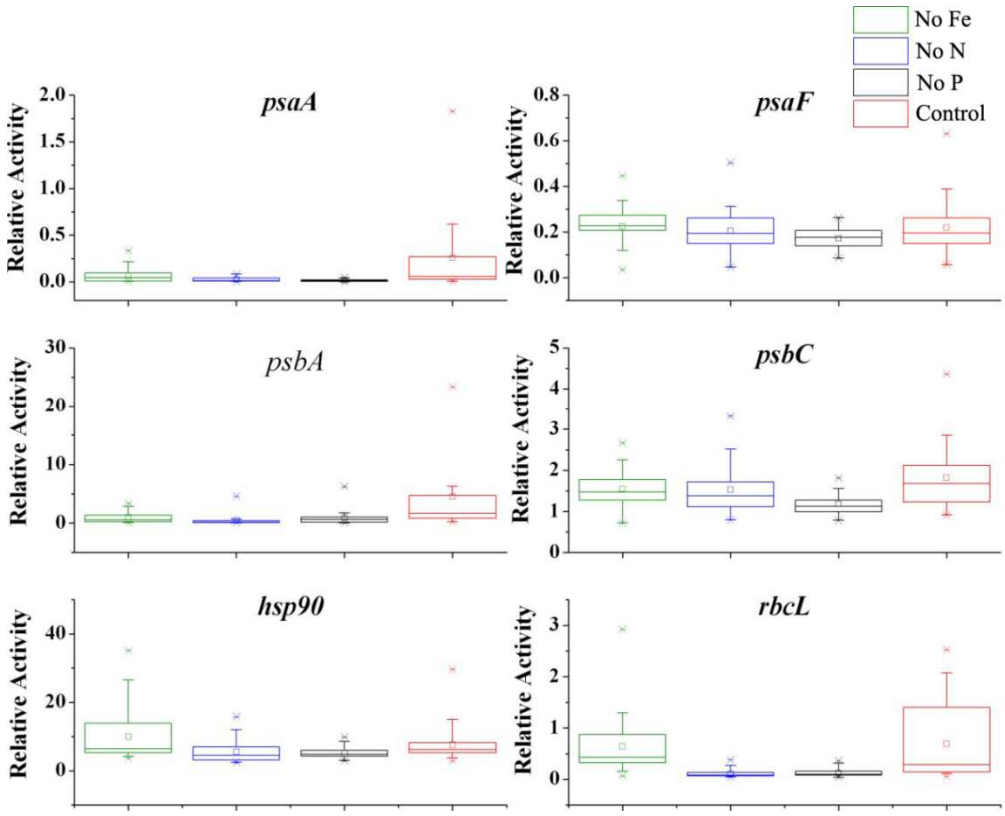

Figure 18. Box chart of expression level of selective genes under different growth conditions.

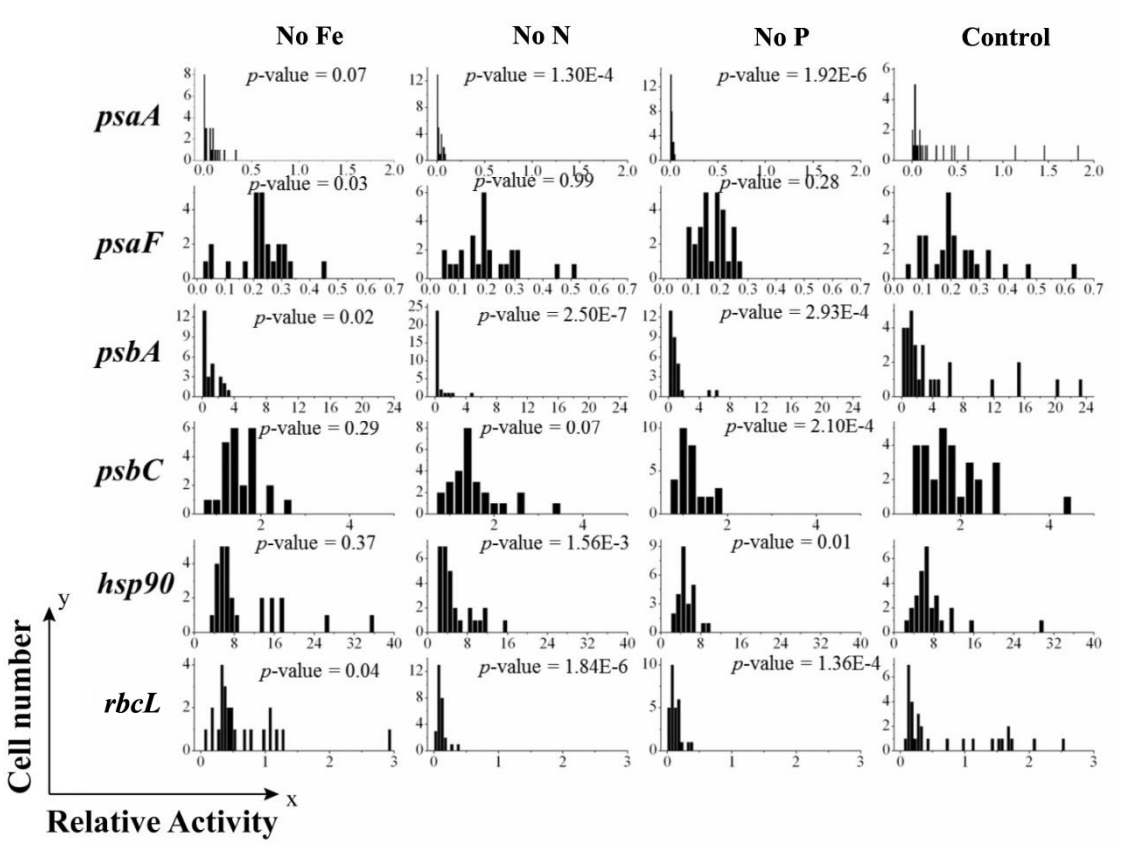

Figure 19. Gene expression distributions of selective genes under four different growth conditions. (*p*-values achieved by using nonparametric two sample Kolmogorov-Smirnov Test between each nutrient-limited and control conditions, $\alpha = 0.05$. X-axis is the relative activity.)

The RT-qPCR results showed that gene expression varied significantly between individual cells, suggesting significant cell-cell heterogeneity existing in the *T. pseudonana* population (Figure 19), consistent with the previous conclusions that stochasticity of transcription contributed significantly to the level of heterogeneity within a clonal population and this heterogeneity cannot be revealed by snap-shot measurements of bulk cells (Bengtsson et al. 2008; Bengtsson et al. 2005; Blake et al. 2003). Comparison of the distribution patterns between conditions can be achieved by Kruskal-Wallis ANOVA test (Schmelz et al. 2003). The results showed that except for the *psaF*

gene, four other genes exhibited independent expression distribution patterns under four growth conditions (Table 6.). The *p*-value of the *psaF* gene was 0.06841 which was close to the cutoff (i.e. < 0.05), indicating that there were still some differences for the *psaF* gene under four conditions.

Table 6. *p*-values of Kruskal-Wallis tests at 95% confidence level

Gene	*p*-value
psaA	8.23229E-15
psaF	0.06841
psbA	0.00255
psbC	9.47652E-5
hsp90	7.70247E-4
rbcL	4.40972E-8

Bulk-cell based analysis showed that the *psaA* gene had a higher expression level in the no phosphate condition than that in the no nitrogen condition (Figure 17). However, a reverse pattern was observed from the single-cell based analysis. A similar pattern between bulk- and single-cell analyses was also observed for the *psaF* gene. For *psbA* genes, the nitrogen depleted condition had the lowest activity among four growth conditions which was only 10% of the control condition. This may be due to the insufficient supply of inorganic nitrogen as found by Kolber *et al.* (1988) that nitrogen limitation could lead to substantial decreases in photosynthetic energy conversion

efficiency and loss of PS II protein D1 which encoded by *psbA* gene. For *psbC* genes, the results from the single cell level were consistent with the results from bulk-cell analysis. Although the bulk cell results indicated that no phosphate had about two times higher activity than the no nitrogen condition for the *hsp90* gene, the single cell level results indicated that the activities were similar to each other. While the up-regulation of the *hsp90* gene under the no iron condition at bulk cell level was confirmed by single cell level results which indicated that low Fe availability indeed triggered stress on *T. pseudonana*. For the *rbcL* gene, the results showed that no iron issued no effect on the *rbcL* expression while no nitrogen affected *rbcL* expression based on both single-cell and bulk-cell based analyses, consistent with previous work in marine diatom *P. tricornutum* (Greene, Geider, and Falkowski 1991).

5.3.6. Principal Components Analysis (PCA) of single cell RT-qPCR data

With the aid of powerful statistical tools, more intrinsic information can be extracted from single-cell based datasets. For instance, besides the independence test based on results of response distributions, principle component analysis (PCA) also can be applied to analyze the relationship between different growth conditions (Figure 20). PCA can provide a simple plot that shows the most important two factors that affect the samples of each growth condition. PCA analysis of *psaF* showed that no nitrogen condition had no significant affect on gene expression in single cells when compared with the control condition since it was close to the control condition in the figure. For *psbA* and *rbcL*, PCA results showed they were separated from each other which meant four growth conditions were distinguished from each other. These results agreed well with the distribution analysis. While for *psaA* and *hsp90* genes, the *p*-value generated from

Kolmogorov-Smirnov Test suggested that there were no significant differences between no iron and control conditions (Figure 19). However, based on the PCA analysis, they had a similar score of component 1 but slightly different score of component 2 of different nutrient-limited conditions, which indicated that expressions of these two genes under nutrient-limited conditions were similar to each other but distinct to the control condition. In addition, for *psbC*, the distribution analysis showed that no iron and control conditions were similar to each other but the PCA results showed that expression of *psbC* gene under no iron and control conditions were not controlled in the same way. Although PCA analysis cannot determine which factors were the most important two factors, it was a good method to visualize the data for larger sample size and to extract the most important properties of the whole sample.

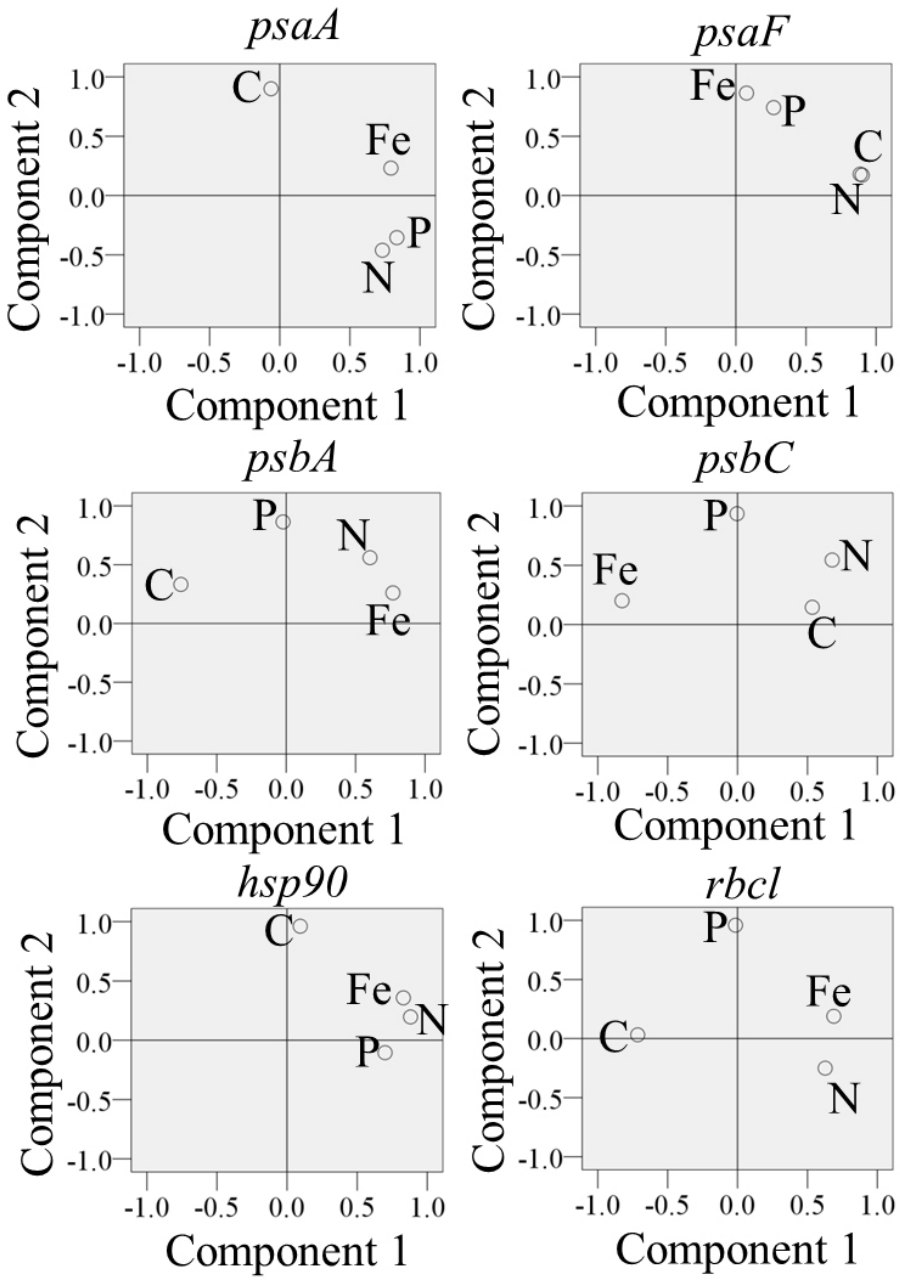

Figure 20. Principle Component Analysis (PCA) of single-cell based analysis of selective genes.

5.4. Discussion

The responses of diatoms to various nutrient-limited conditions have been evaluated at the population level (Mock et al. 2008; Thamatrakoln et al. 2012). However, as planktonic microorganisms, the cell-cell heterogeneity of diatoms in terms of responses

to environmental factors could be significant, and have never been documented. In this study, we made the first attempt to measure the expression of selected genes out of model diatom *T. pseudonana* when they were subject to no nitrogen, no phosphate and no iron conditions. The results showed that significant heterogeneity was found which shed light on potential environmental problems. Opposite expression patterns were found for *psaA*, *psaF*, *psbA* and *hsp90* genes between single-cell based and bulk-cell based analyses. The abnormal cells identified in single cell analysis may be an indicator of potential environmental problems and suggest further investigations would be possibly buried under the average value of bulk cell population analyses.

In order to apply *T. pseudonana* as a sensor by using single cell RT-qPCR, several issues need to be addressed. The first is the sensitivity of the sensor which is equivalent to single cell RT-qPCR sensitivity. Since the sensitivity of our technology can go down to a single cell level, it has the capability to analyze some precious and/or uncultured environmental samples.

The results show that as the copy number of transcripts of a gene decreased, the StDev of RT-qPCR technical replicates increased accordingly. To overcome the issue, small reaction volumes which increase the local template concentration are preferred. In the study, we used 10 µL reaction volumes instead of conventional 20 µL reaction volumes for qPCR. Currently, 10 µL was the smallest volume that we could get consistent and reliable results in the tube/microtiter plate-based qPCR reaction. In order to further decrease the reaction volume, chip-level devices which can decrease the volume to several µL (Shi et al. 2011) even pL level (Beer et al. 2007; Kiss et al. 2008; Lindström and Andersson-Svahn 2010) will be more attractive. In addition, we also

addressed the low starting material issue by increasing the cDNA yield of specific targets through adding target-specific primers. Ståhlberg *et al.* (2004) evaluated 4 different primer strategies which were random hexamers, oligo (dT), gene specific primers and gene-specific primers mixture on five different genes, and the results showed that gene specific primer mixtures had an overall advantage based on the yield and StDev of qPCR results of several different genes (Ståhlberg et al. 2004). In order to simplify the whole process, considering the reverse primer of qPCR is complementary to the mRNA sequence, it may bind to specific mRNA during the cDNA synthesis step which will increase the cDNA yield of specific targets, we added the reverse primer directly rather than using the specially designed specific primers that are complementary to the mRNA sequence as described by Ståhlberg *et al.* (2004). The results showed that adding target-specific primers in the cDNA synthesis step could increase the quality and yield of target cDNA by about 10 fold on average.

The second issue is how to interpret RT-qPCR results in a quantitative way so that the result can be used as an indicator of environmental stress conditions. The use of reference genes is important in order to normalize qPCR results, and much research has been done on the selection of reference genes for various bulk-cell based analyses (Czechowski et al. 2005; Huggett et al. 2005). However, considering gene expression stochasticity in single cells, the reliability of employing these genes for single-cell gene expression is still unclear. In this study, we selected and validated the *actin* gene as an internal reference based on its better performance than other candidate genes, and the expression heterogeneity of the *actin* gene was still observed between individual *T. pseudonana* cells. To fully address the heterogeneity issue, an alternative internal control

strategy, such as using a molecule that is artificially incorporated into the sample as an RNA spike (Bower et al. 2007; Stephen A Bustin and Nolan 2004; Huggett et al. 2005), may be necessary and worth further development.

There are still technical challenges for using microbial gene expression at the small-number-of-cells level as environmental sensors. For example, the targeted microbe is in the mixture with other microbes, yet further manipulation such as cell sorting or cultivation will alter gene expression levels. How to successfully find out the atypical gene expression patterns from a few cells among a larger number of background normal cells is another big challenge. To overcome this, a feasible approach is to perform high-throughput single-cell level analysis to the microbiota, then extract the targeted information using post processing on the acquired data.

Finally, our results demonstrated that with proper selection of gene targets and optimization of RT-qPCR conditions, gene expression measurements at single cell resolution will allow monitoring of the ocean environmental health, possibly at an early stage of potential environmental problems to minimize the cost of environmental remediation. However, currently the technology works well only with highly expressed genes, which limits the selection of gene targets. In the future, further development and optimization of the molecular biology protocol, and integration with a chip-level real-time PCR device (Shi et al. 2011) will generate a chip level sensor instrument for monitoring marine environmental health in a fast and effective way to overcome the remaining technical challenges. At the same time, fundamental microbiology questions about heterogeneity within an isogenic population will be answered as well.

6. ONE-STEP ON-CHIP SINGLE-CELL RT-QPCR

6.1. Introduction

Satellite *in situ* blended ocean chlorophyll records indicate ocean annual primary production declined about 6% from 1980's to 2000's (Gregg et al. 2003). Without further sacrificing the primary production of the ocean, an efficient way to monitor this negative perturbation in the ocean at an early stage which takes advantage of the recent improvement of technologies is urgently required.

Currently, there are several different kinds of sensors that can be used to monitor the ocean environment. For instance, satellite remote monitoring can provide large-scale or even global information (Field et al. 1998; Behrenfeld et al. 2001; Gregg et al. 2003). However, the accessibility of this source is limited and the large-scale imaging typically has an inherent lack of sensitivity. Electrochemical based sensing is another widely used tool (J. Wang 2002). For example, a BOD (biochemical oxygen demand) sensor (Karube et al. 1977; Y. R. Li and Chu 1991; Strand and Carlson 1984; Liu and Mattiasson 2002) has an immobilized biofilm inside the electrode which detects the BOD in a short time which is typically 15-20 mins (BOD_{ST}) instead of conventional BOD_5 which takes five days. The potential problem is the BOD_5 can only be derived from BOD_{ST} under high concentration of fast and easily assimilable compounds which is not the case for ocean environment. There are also other electrochemical based sensors, such as copper sensors (J. Wang et al. 1995), phenolic sensors (J. Wang and Chen 1995) and okadaic acid sensors (Campàs and Marty 2007). All of these sensors are target specific, very sensitive and able to be minimized (J. Wang 2002). Due to these properties, they are good candidates for deployment in ocean. However, there are still challenges that need to be

faced such as long-term stability, related baseline drift and hazardous materials required for the electrodes (J. Wang 2002). Other than that, the electrochemical based sensors normally only target one type/group of substance which decreases the cost-effect of this type of sensors. There is also other type of sensor to monitor the environment *in situ*, such as engineered microorganisms (Ripp et al. 2000). Although the process of constructing an engineered microorganism requires special expertise and is time consuming which limits the application of this method, this method is able to provide the information about the biological impacts which are missing in the aforementioned sensing methods.

Here I introduce a new technology on which a novel environmental sensor network can be built. This real time *in situ* monitoring technology will minimize errors and costs associated with sample transportation and laboratory analyses (J. Wang 2002), will use native inhabitants of the ocean, and will be able to be deployed in the ocean to monitor biological impacts in the environment.

This technology will take advantage of the well known facts that gene expression will be altered when the environmental condition changes (Gasch et al. 2000) and environmental variations will first change the gene expression before altering the whole community (Edward F. DeLong 2009). In order to develop an early warning system, I hypothesize that individual cell gene expression analysis provides a richer reference than the average of a bulk cell population. Our previously developed inexpensive chip level device (Shi et al. 2011) will be combined with tube-based single cell analysis (Shi et al. 2013). In order to monitor the environment of the ocean, one microbe that can be widely found and has broad representation should be chosen as a target. Based on these

considerations, diatom *T. pseudonana*, which belongs to a major group of unicellular phytoplankton, is selected as the target microbe (Falkowski et al. 2004; Thamatrakoln et al. 2012; Shi et al. 2013).

6.2. Experiments

6.2.1. Cell culture

Thalassiosira pseudonana (CCMP1335) cells which belong to diatoms were obtained from the National Center for Marine Algae and Microbiota (NCMA), and were grown in f/2 medium at 24 ± 1°C (Guillard 1975; Guillard 1962) under a constant lighting condition (30 µmol photons m^{-2} s^{-1} irradiance measured using LiCor (Lincoln, NE)). Cells were cultured in normal conditions with no iron. Artificial seawater was prepared using chemicals of analytical purity and prepared based on the formula of Kester *et al.* (1967).

6.2.2. Chip fabrication

The droplet format is based on the thermal cycler used. In this research, the thermal cycler used is ABI StepOne Real-Time PCR System (Applied Biosystems, Foster, CA). Based on the thermal cycler structure of StepOne Real-Time PCR system, a 48 droplet format was applied. The chip fabrication process was the same as described in Section 4.2.2.

6.2.3. One Step reverse transcriptase quantitative PCR (RT-qPCR)

Primers for RT-qPCR were designed using Primer-BLAST. The primer design standard can be found in our previous single cell technologies development paper (Gao, Zhang, and Meldrum 2011; Shi et al. 2013). RT-qPCR was performed using Express One-Step SYBR GreenER Universal Kit (Invitrogen, Carlsbad, CA) on an ABI StepOne Real-Time

PCR System (Applied Biosystems, Foster, CA). PCR reactions contained 0.4 µL of each primer with a concentration of 4 µM, 2 µL of master mixture, 0.04 µL ROX, 0.04 µL SUPERase In, 0.4 µL 5X BSA and 0.08 µL SuperScript III Reverse Transcriptase and 0.64 µL of DEPC treated water (Ambion, Austin, TX). Reactions without cells served as negative controls.

6.2.4. Single cell isolation

Single cells were isolated using the micromanipulator as described in Sections 4.2.4. and 5.2.2.

6.2.5. Data analysis

To describe the distribution variation of single-cell gene expression levels among cells, nonparametric statistic tests which do not require normal distribution of datasets were applied (Siegel 1957). Kolmogorov-Smirnov test was used to analyze the relationship between two different growth conditions of RT-qPCR measurements using the OriginPro 8.1 software (OriginLab Corporation, Northampton, MA).

6.3. Results

6.3.1. On chip one step RT-qPCR optimization

In order to achieve the best performance of the chip, first, three different commercially available one-step RT-qPCR kits were compared: Express One-Step SYBR GreenER Universal Kit (Invitrogen, Carlsbad, CA), SuperScript III Platinum® One-Step qRT-PCR Kit w/ROX (Invitrogen, Carlsbad, CA) and One Step PrimeScrip RT-PCR Kit (Perfect Real Time) (Clontech, Mountain View, CA). Series dilutions of purified RNA were used to test the performance of each kit. Based on the results of Figure 21, SuperScript III Platinum® One-Step qRT-PCR Kit cannot differentiate the difference between negative

control and the lowest concentration of RNA. The PrimeScrip RT-PCR Kit had even worse performance and cannot differentiate between concentrations smaller than 16 pg/µL of RNA.

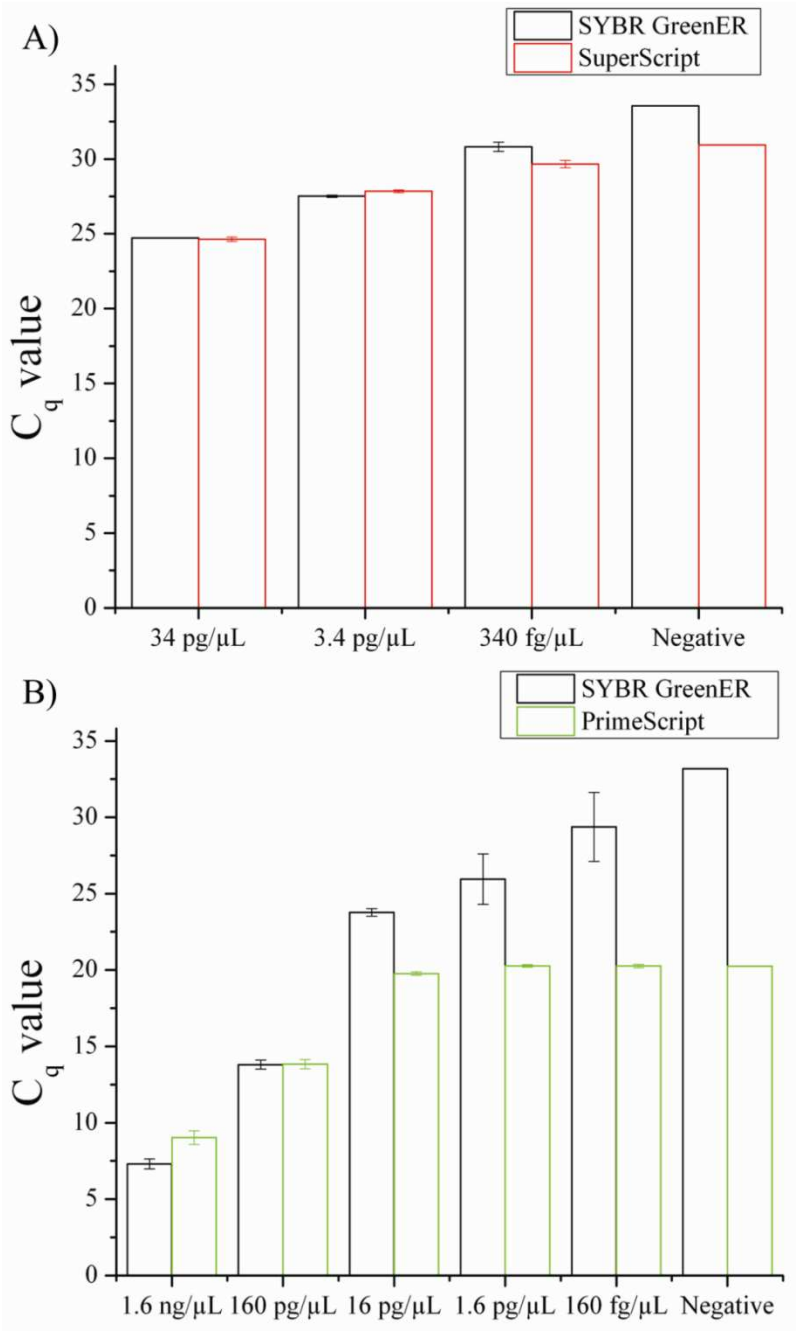

Figure 21. Commercial kits performance comparison.

Secondly, cDNA synthesis time was tested. For the tube-based method, the cDNA synthesis time is about two hours for single cell reactions. Two hours is too long for applying the one-step RT-qPCR technique; hence, the cDNA synthesis time was decreased from 2 hours to 40 mins. Two genes were tested at different concentrations of purified RNA and both of them showed linear response of different concentrations of purified RNA with 40 mins cDNA synthesis time. The results indicated that 40 mins cDNA synthesis time was robust for the one-step RT-qPCR reaction (Figure 22). Further decreasing the cDNA synthesis time decreased the quality of the qPCR results.

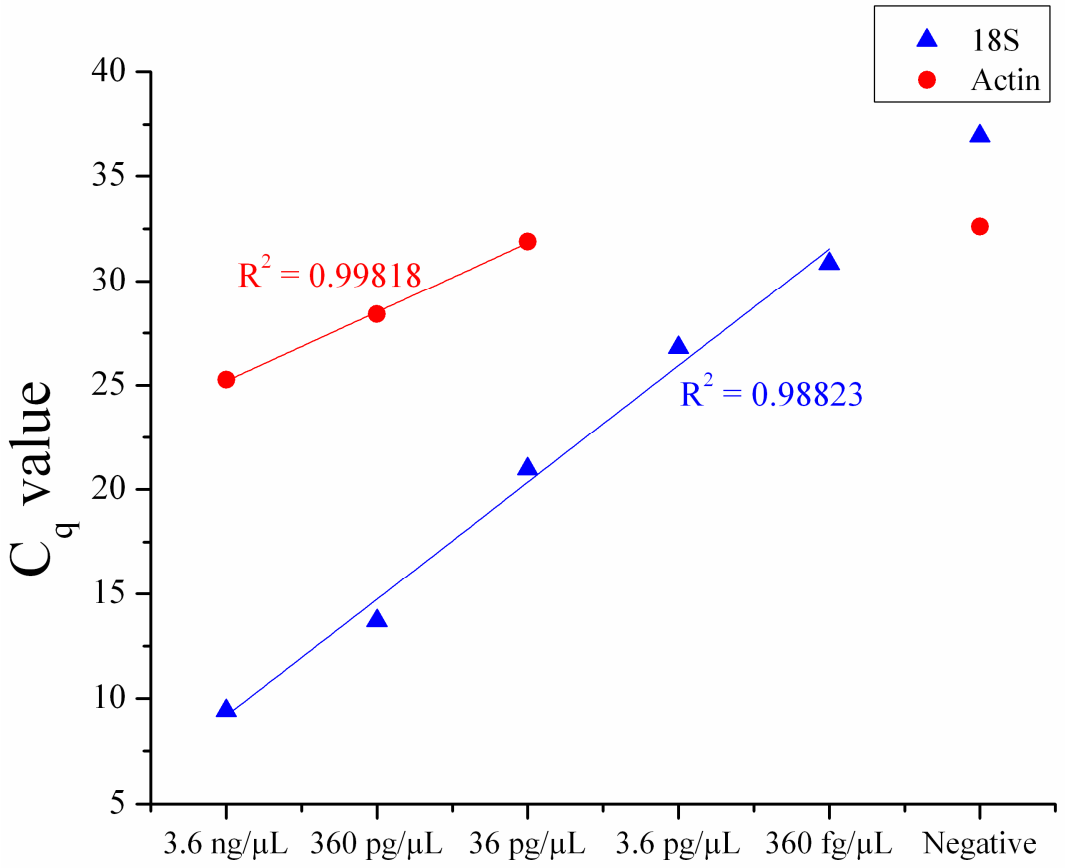

Figure 22. 40 mins cDNA synthesis time test.

Thirdly, in order to directly run reactions in droplets without RNA extraction/purification, the cell lysing method in the droplet needs to be selected carefully. Some methods are compatible with the droplet based format, such as heat lysing (Clark et al. 1993; Hirakata et al. 1998) and Triton X-100 surfactant lysing (Werf, Hartmans, and Tweel 1995; Ren and Schwartz 2000). I tested different temperatures of cell lysing by using heat and found that 49 °C was strong enough to lyse the cell. Two different lysing methods are compared in Figure 23. The results show that heat lysing has significant advantage over Triton X-100 methods of both genes. The possible reason of bad performance of Triton X-100 may be due to the PCR compatibility of this chemical.

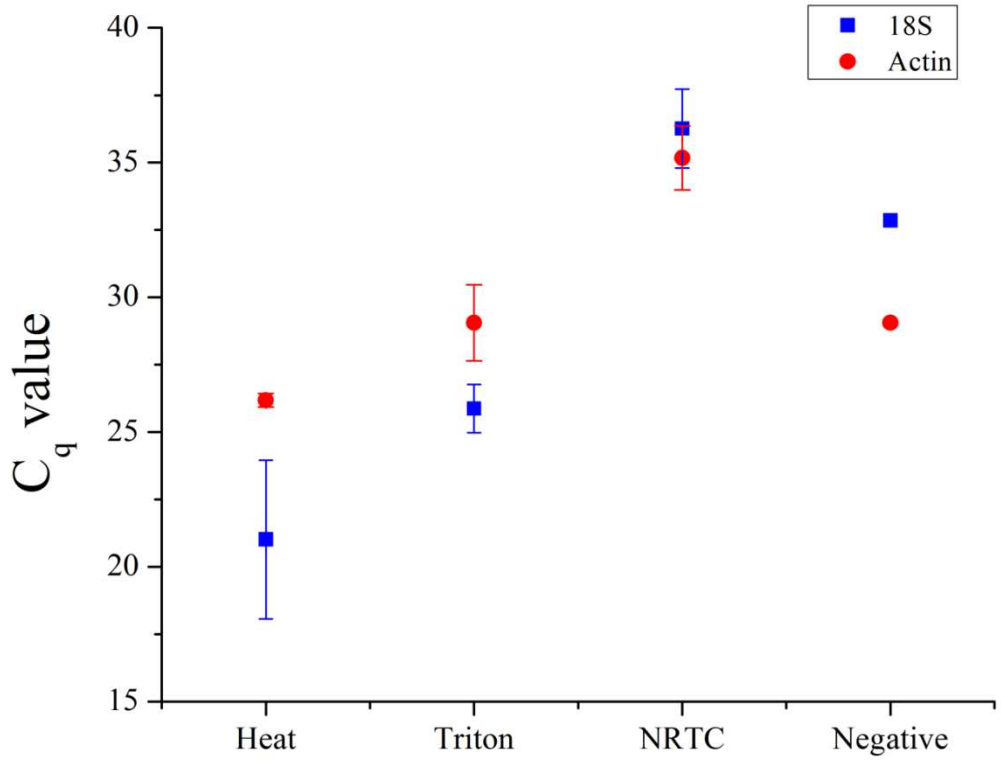

Figure 23. Cell lysing methods comparison. (NRTC stands for no reverse transcriptase control)

6.3.2. Temperature calibration

To perform RT-qPCR in the designed chip level device, a thermal cycler was needed. A customized miniature thermal cycler is the best choice for future sensors which are deployed in the ocean, but at this moment, broadening the potential application scope of this chip is also important. For general use of this single cell sensitivity chip, any thermal cycler can be used. For different thermal cyclers, temperature calibration steps will be needed. The process for temperature calibration was inserting a 0.076 mm-diameter K-type thermocouple (5SC-TT-K-40-36, OMEGA, CT) into the center of a droplet to adjust the setting temperatures and corresponding durations of the StepOne machine to achieve the required temperatures of the droplets. A thermocouple reader (50 Series II, Fluke, WA) was used to read the temperature during thermal cycling. Figure 24 shows the temperature profile set on the thermal cycler (black rectangular dots) and the compensated temperature profile in the droplets (dots). The heating/cooling rates were longer than those for conventional in-tube PCR, but this issue can be solved by using a customized thermal cycler. However, with this temperature calibration step, more biological laboratories can take advantage of this single cell analysis chip as long as a real-time PCR machine with a flat top thermal cycler is available. To further demonstrate that this chip can be widely used in conventional biological laboratories, no customized signal detection part is applied. The signal detection part is solely from the StepOne real time PCR machine.

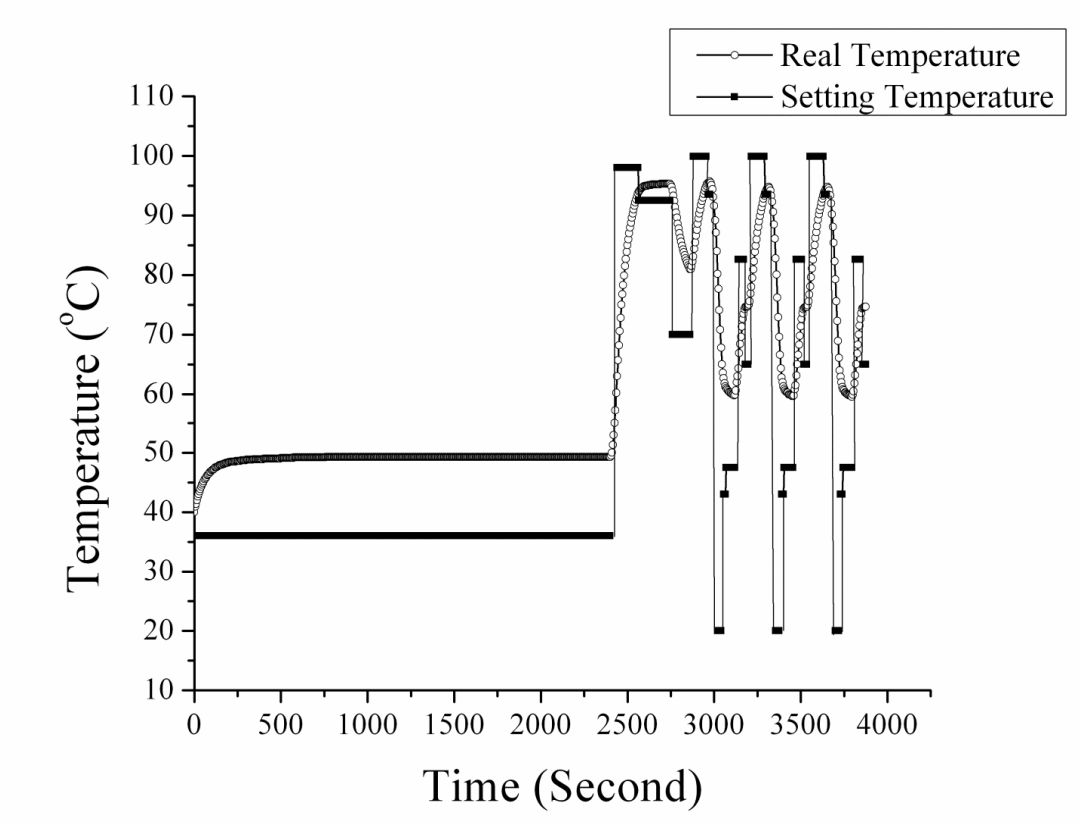

Figure 24. Temperature setting of the thermal cycler.

6.3.3. Chip performance evaluation

Before executing the experiments, the performance of the chip needs to be evaluated. First, different amounts of purified RNA were loaded on the same chip. The results in Figure 25 show clearly that different concentrations of RNA clustered well and only 6 out of 12 negative controls showed amplification signals. Additionally, the signal from the amplified negative controls can be differentiated with the 2 pg reactions. The R^2 is 0.99 which further demonstrated good performance of the chip.

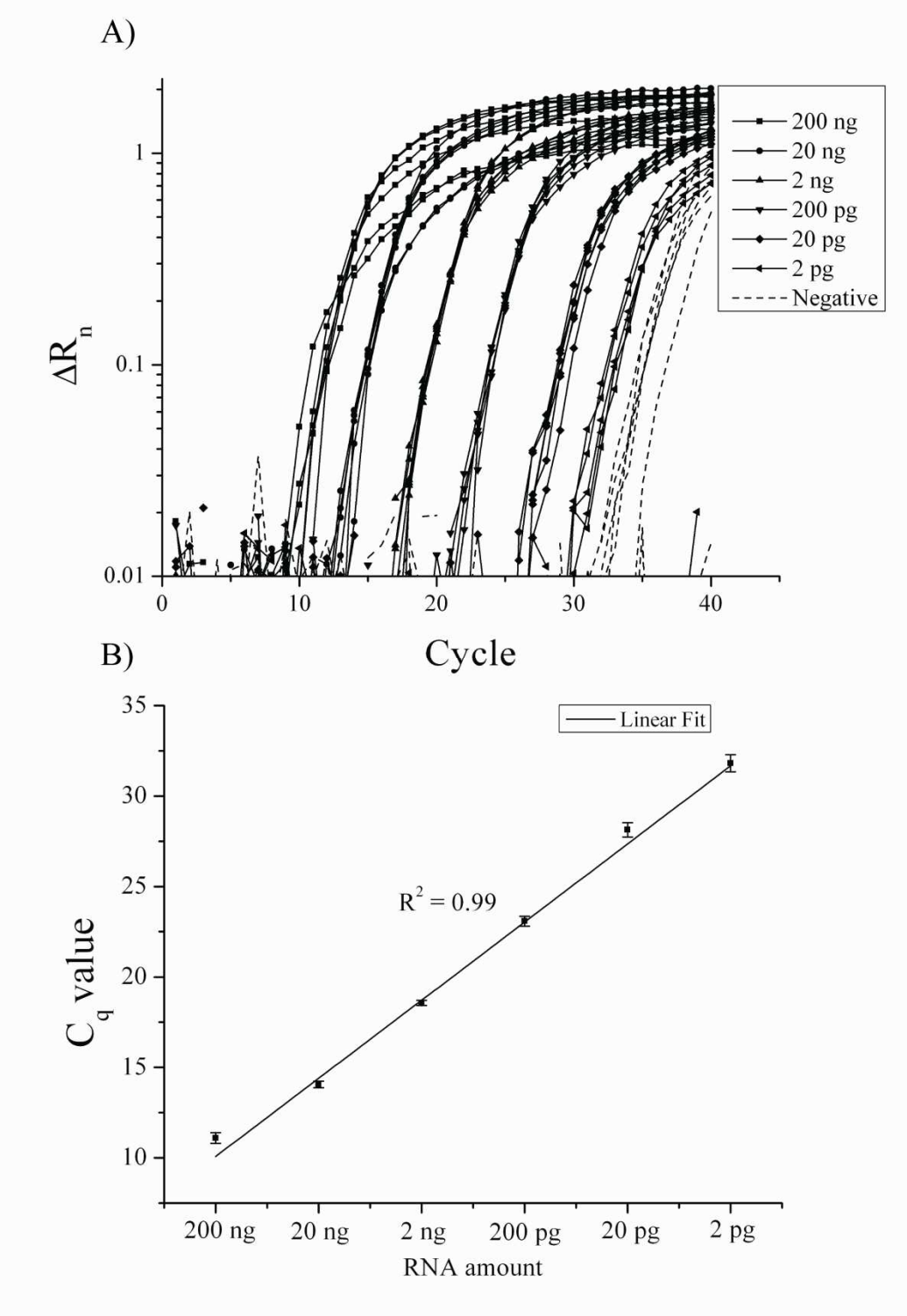

Figure 25. Standard curve analysis.

Other than the standard curve experiment, a good design should have minimized variance over the same chip. This was tested by loading different amounts of purified RNA on the chip. The results in Figure 26 show that the standard deviation over the chip was well controlled. The standard deviation at a 10 pg level was only about 0.3 cycles.

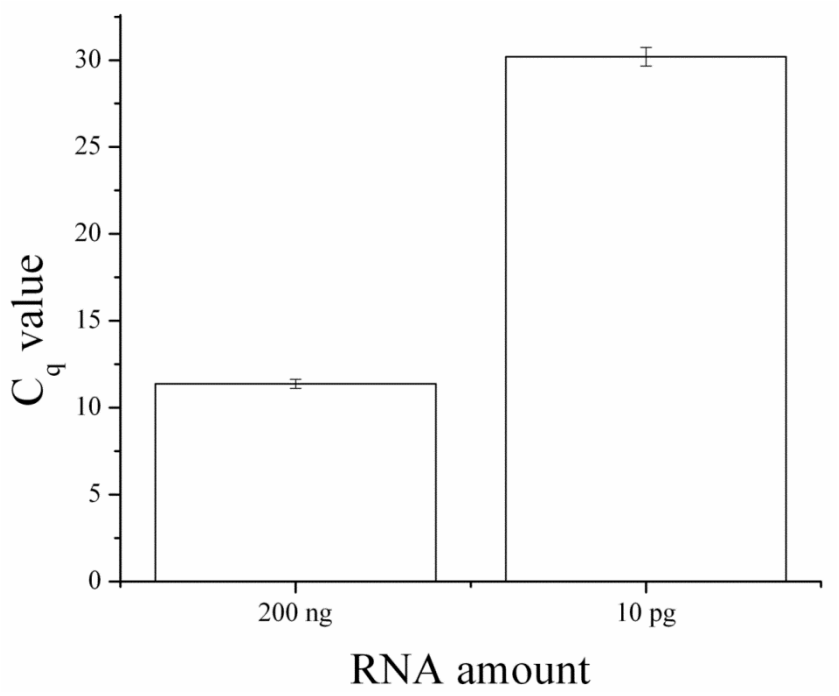

Figure 26. Chip variance test.

Based on the results from Figure 25 and Figure 26, the chips run robustly. Since single-cell resolution is needed, the next question is if single cell analysis can be performed on the chip. First, an estimation of total RNA in a single *T. pseudonana* cell needed to be made. Table 7 provides some general information about *T. pseudonana*. The total RNA in a single *S. cerevisiae* is about 1 pg and in a single mammalian cell is about 20 pg (Schmid et al. 2010). Based on this information, the estimation of total RNA in *T. pseudonana* should be at the several pg level. Considering the results of the performance tests, they were solid demonstrations that single cell resolution can be achieved.

Table 7. General size and genome information

	S. cerevisiae (Schmid et al. 2010)	***Thalassiosira pseudonana***
size	5 µm	5 µm
Genome size	12 Mb	34 Mb
Gene number	5,770	11,242

6.3.4. Single cell gene expression

Single cell gene expression results are shown in Figure 27. For single cell analysis the C_q value is used directly as a representation of the activity of individual cells. No reference gene was used since in each droplet there was only one cell which was visually confirmed. The results showed that variance between two growth conditions cannot be negligible for the genes tested. Single cell reactions with no detected C_q value, C_q value larger than the negative control, or wrong melting curve were removed from the data analysis to ensure the reliability of analysis. Those reactions with no detected C_q value were not necessarily failed and it is highly possible that the transcription number in those reactions were between the detection limit of the chip and zero (Ståhlberg, Kubista, and Åman 2011). Since it was difficult and time consuming to determine the detection limit of an individual gene (Ståhlberg, Kubista, and Åman 2011), those results were removed for convenience. All the negative control reactions either were undetected or can be differentiated by melting curve analysis. After removing all the unreliable reactions, the overall successful reaction rate was over 82%.

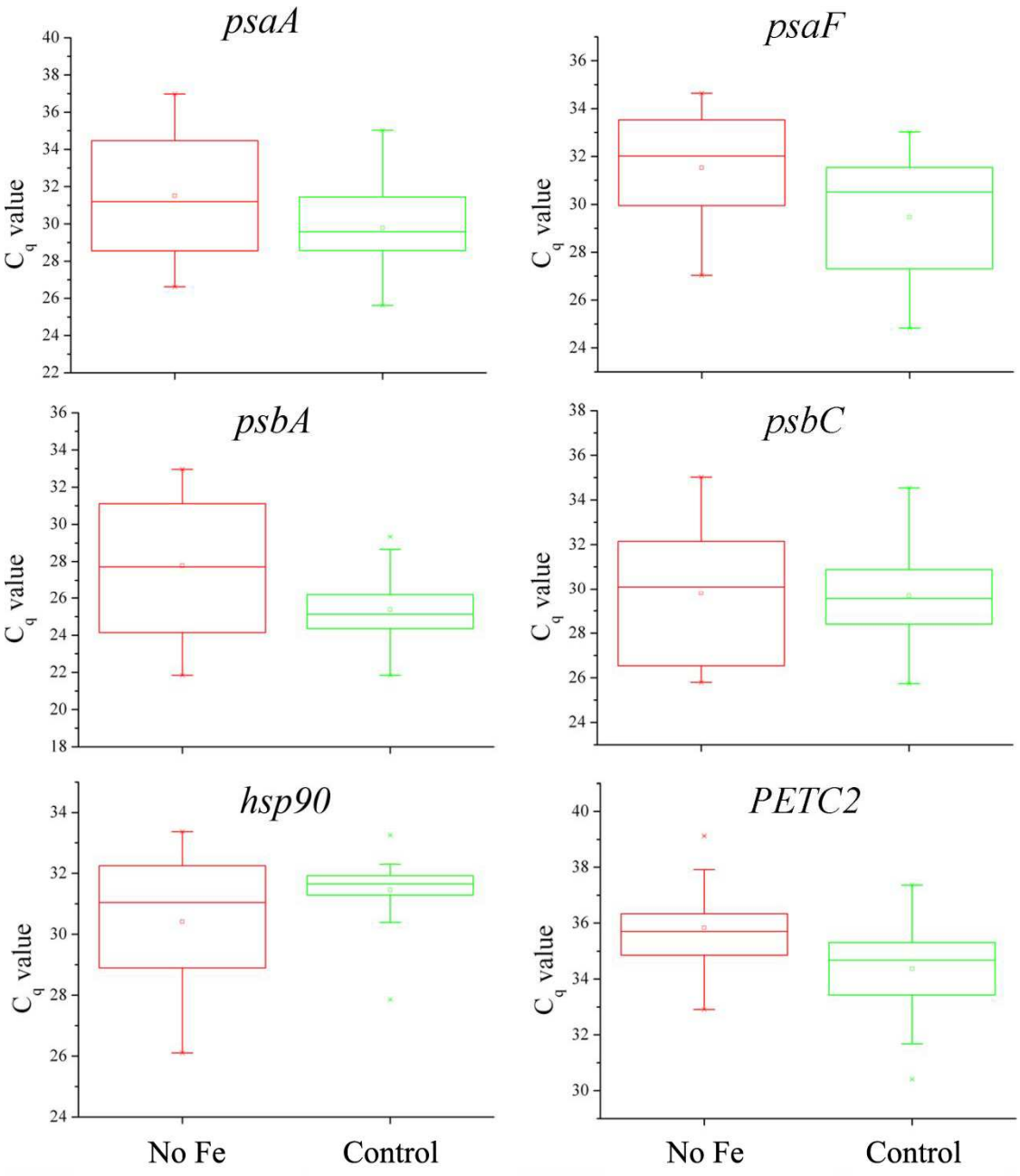

Figure 27. Box chart of single cell gene expression analysis.

The genes selected for this research were based on our previous tube-based single cell analysis results. One more gene *PETC2* was selected to replace the *rbcL* gene since *PETC2* gene codes a b6-f complex iron-sulfur subunit (Maiwald et al. 2003) and is more

closely related to iron transportation. All of the six genes showed lower expression under no Fe conditions except *hsp90*. Up-regulation of *hsp90* indicated that the no iron condition caused a stress on the population which results in higher expression. As aforementioned, *PETC2* codes an iron-sulfur subunit which belongs to a Rieske protein (Yan and Cramer 2003). Down regulation of the Rieske-type protein subunit under iron limitation condition was reported by Allen *et al.* (2008) using microarray data which was a powerful support of our chip. All single cell results showed larger variance among populations and those results were frequently overlooked by population level analysis.

Further investigation of the single cell results reveal that a bimodal distribution can be observed under no iron conditions for *psbA* (Figure 28). This may suggest that the existence of two subpopulations: one with enhanced transcription level and one opposite, resulting in a bimodal distribution (Choi et al. 2008; Longo and Hasty 2006; Adam K White et al. 2011; Bengtsson et al. 2005). The bimodal gene expression has been well documented before. Bengtsson *et al.* (2005) found a bimodal model (Ko 1992) in mouse insulinoma MIN6-cells where one subpopulation has increased activity characterized by a high mean value and another subpopulation has a relatively lower mean value. White *et al.* (2011) also observed such distributions when they analyzed the coregulation of miR-145 and OCT4 in single cells. This expression pattern was not noticed in other genes which indicated that the cell may involve different strategies to regulate gene expression.

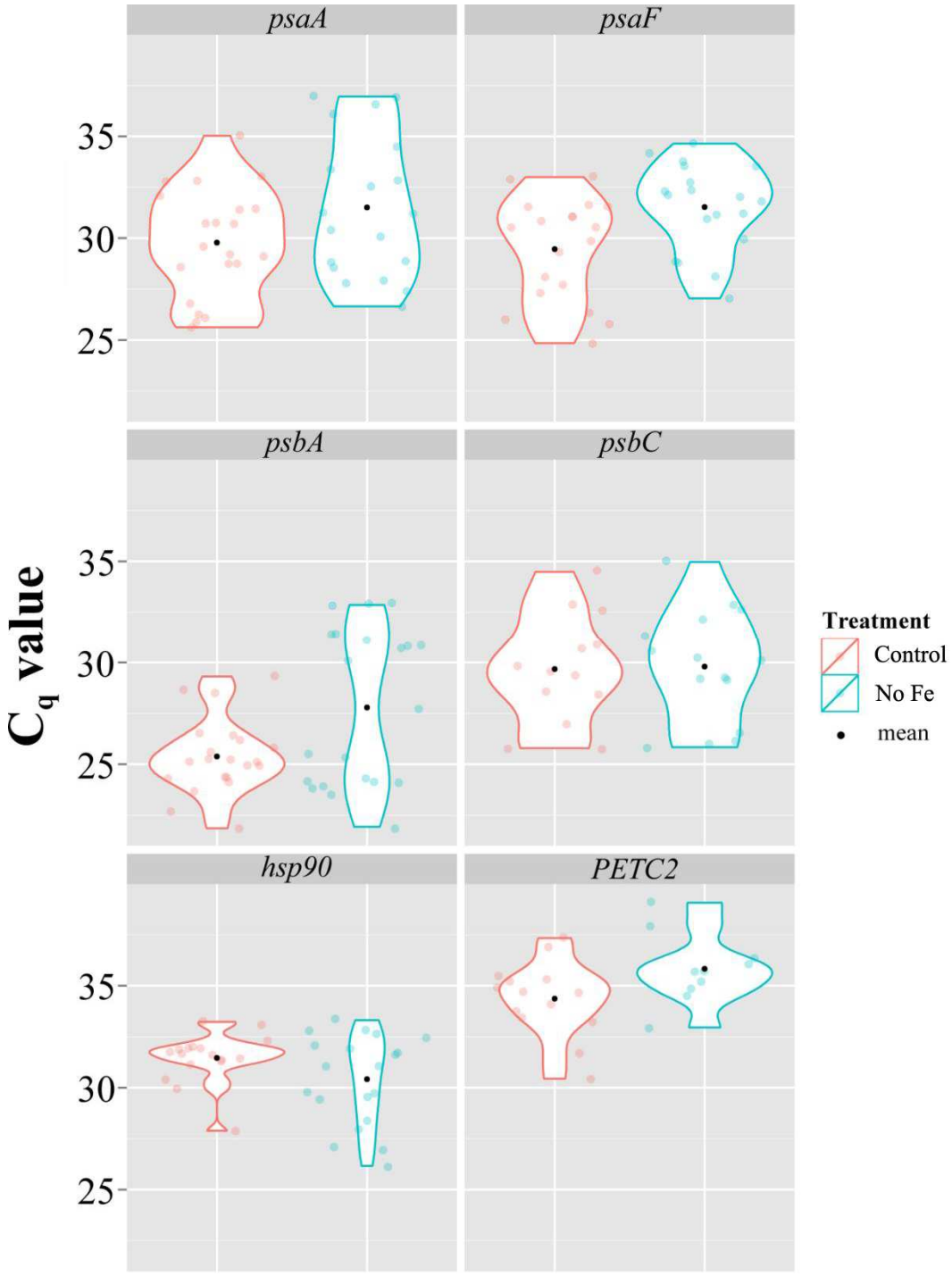

Figure 28. Violin plot of single cell gene expression analysis.

Since single cell level analysis revealed distributions among populations, statistical tools can be involved to analyze the distribution of each sample. Skewness

(Ståhlberg, Kubista, and Åman 2011) is an index that indicates the asymmetry of distribution. A distribution with a longer left tail has negative skewness, and the opposite has positive skewness. If a sample is highly skewed, more cells with expression level far away from the mean value exist which may mean that those so called "outlier" cells are functionally important (Hebenstreit 2012). Those outlier cells cannot be detected by traditional population level analysis. Kurtosis is another index which represents the "peakedness" of a distribution. Sharper and higher peaks will have larger kurtosis and vice versa. Larger kurtosis means the samples are more concentrated. Table 8 shows the skewness and kurtosis of different samples.

Table 8. Skewness and Kurtosis

		psaA	*psaF*	*psbA*	*psbC*	*hsp90*	*PETC2*
skewness	No iron	0.34891	-0.59842	0.00673	0.03252	-0.5463	0.435
	control	0.00618	-0.44573	0.50014	0.16243	-1.5329	-0.531
kurtosis	No iron	-1.2102	-0.41157	-1.78328	-0.7040	-0.8943	0.75557
	control	-0.76336	-1.04537	0.5353	-0.5078	4.03544	0.63393

The results indicate that *psaF* and *hsp90* under no iron condition and *psbA*, *PETC2* under control condition were moderately skewed (skewness within -1 to -0.5 or 0.5 to 1) (Bulmer 1979), and *hsp90* under control condition were highly skewed (skewness less than -1 or larger than 1) (Bulmer 1979). These skewed results demonstrate the importance of single cell analysis since more outlier cells may be found for those

samples and the mean value is less representative of the whole population. For genes *psaA*, *psbA*, *psbC* and *hsp90*, 4 out of 6 kurtoses under control condition had larger values indicating relatively concentrated distributions under these conditions. That means under no iron conditions those genes had broader distribution and the whole population may take advantage of those properties to adapt to the stress condition. No unifying model exists to represent all 6 genes expressions at the single cell level, confirming that gene expression was regulated by multiple elements.

A fit test was also performed at the single cell level for all 6 genes under two growth conditions (Figure 29). The results showed that a lognormal distribution had a good fit for all conditions. This skewed lognormal distribution was consistent with previous reported results (Ståhlberg, Kubista, and Aman 2011; Bengtsson et al. 2005; Shalek et al. 2013; Ståhlberg, Rusnakova, and Kubista 2013). A significant impact of lognormal distribution is that the geometric average becomes more important and more representative of the whole population than arithmetic average (Ståhlberg, Rusnakova, and Kubista 2013). While the population level analysis can only achieve arithmetic average, it again raises the necessity of single cell analysis.

The advantages of our methods are streamlined processing which minimized anthropogenic error and eliminated the possible bias introduced by pre-amplification. The chip design is pretty flexible and can be determined based on the thermal cycler and signal detection module of an existing qPCR system which allows the chip to be widely used by conventional microbiology laboratories. Based on the existing StepOne Real-Time PCR system, the current throughput was limited to 48 reactions per chip, although there is no technical hurdle to increase the throughput significantly with a customized

thermal cycler and signal detection module. However, one limitation of the chip is whole transcriptome analysis cannot be achieved on the chip. Multiplex qPCR can be used to alleviate this problem but still only a limited number of genes can be detected at a time due to spectral overlap of classical fluorophores.

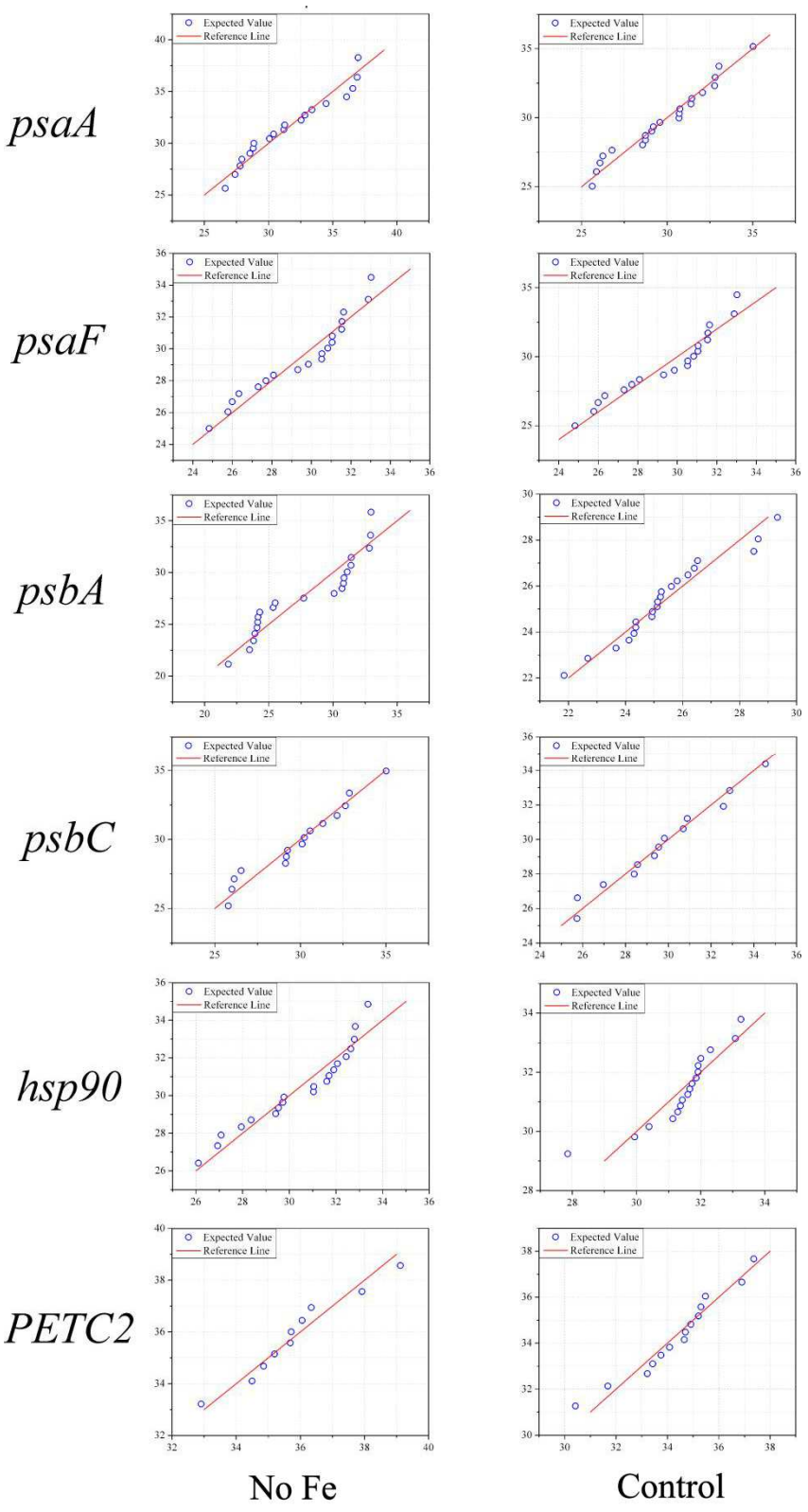

Figure 29. Lognormal distribution fitting.

6.3.5. Power evaluation and sample size estimation

The power of single cell results can be achieved by using OriginPro 8.1 and the formula is:

$$\text{Power} = \text{Prob}(t > t_{(1-\alpha,\nu)}, \nu, \lambda)$$

The results n is sample size, degree of freedom $\nu = 2(n-1)$.

$$\lambda = \frac{\delta}{\sqrt{2/n}} \text{ and } \delta = \frac{|\mu_1 - \mu_2|}{s}$$

μ_1 and μ_2 = mean value of control condition and no iron condition. s = standard deviation of the combined samples. The results can be determined based on probability table.

The calculated power for individual genes is: *psaA*: 69%; *psaF*: 89%; *psbA*: 87%; *psbC*: 12%; *hsp90*: 66%; *PETC2*: 71%. The power was calculated by a 1-side test due to the skewed distribution and at a 0.10 level. The power analysis showed good results which demonstrated that the chip performance was strong enough to represent the information of the population. The lower power for *psbC* was because of the small variance between two populations. In order to distinguish the variance between two populations with small variance, larger sample size is required. Currently, due to some technique issues and biological hurdles, only around 20 cells of each condition can be achieved. A detailed discussion is provided in Appendix A. In order to achieve better power, larger sample sizes need to be accomplished. The required sample size can be evaluated based on the preliminary test of each experiment and it depends on the mean value of each population and standard deviation of the combined population. For my analysis, in order to achieve 90% power, the sample size for each gene will be: *psaA*: 42;

psaF: 21; *psbA*: 25; *psbC*: 6183; *hsp90*: 44; *PETC2*: 23. The results were calculated by OriginPro 8.1 and the formulas are same as power calculation. The exceptional larger sample size for *psbC* was due to the close mean value of two populations under no iron and control conditions. If higher power was needed, larger sample size was needed which can be solved by increasing the throughput. Running of multiple chips for each condition was not a feasible solution because of RNA preservation and chip to chip variance (see Appendix A).

6.3.6. Comparison with conventional in-tube single-cell qPCR results

Tube-based single-cell qPCR has been achieved by us before (Shi et al. 2013) and the results will be used to compare with the chip-based results. Table 9 lists the results of the *p*-value of two growth conditions from two different methods. The *p*-value results have 100% correspondence to each other which indicated the power and reliability of the chip-based experiment.

Table 9. *p*-value of two different methods at 95% confidence level

	tube-based method	chip based method
psaA	0.07	0.37
psaF	0.03	0.03
psbA	0.02	0.02
psbC	0.29	0.98
hsp90	0.37	0.08

Other than the *p*-value, *Pearson* correlation analysis (Lee Rodgers and Nicewander 1988) which calculates the linear correlation between two samples was also applied to the chip-based results. The results are shown in Tables 10a and b.

Table 10a. *Pearson* correlation analysis of no iron condition at the 0.10 level

	psaA	*psaF*	*psbA*	*psbC*	*hsp90*	*PETC2*
psaA	1					
psaF	-0.29735	1				
psbA	0.01336	-0.0553	1			
psbC	0.17019	0.3241	0.30327	1		
hsp90	0.46153*	-0.21957	-0.03133	-0.20126	1	
PETC2	-0.12301	0.30403	-0.10091	-0.19848	0.18193	1

Table 10b. *Pearson* correlation analysis of control condition at the 0.10 level

	psaA	*psaF*	*psbA*	*psbC*	*hsp90*	*PETC2*
psaA	1					
psaF	0.24948	1				
psbA	0.21327	0.12527	1			
psbC	0.01059	0.4479	0.0488	1		
hsp90	0.6084*	0.08453	0.06102	-0.12517	1	
PETC2	-0.03567	0.34246	-0.48679*	0.32337	-0.19856	1

* significant correlation

The correlation analysis showed that the significant negative correlation under control condition between *psbA* and *PETC2* became less significant under no iron condition while the positive correlation between *psaA* and *hsp90* was maintained under both conditions. Compared with previous tube-based results (Tables 2 and 5), there was no correlation between those two methods. Considering that *Pearson* correlation analysis is a statistical analysis tool, the results may influenced by a lot of factors, such as sample size, data normalization method used and so on. The inconsistency of those two results may be reasonable. Although the single cell analysis is becoming more prevalent, how to interpret the results in a reliable and accurate way is far away from mature, so more effort

needs to be put on this topic. This work also provides some possible tools that can be used to analyze single cell results.

7. CONCLUSION AND FUTURE WORK

7.1. Conclusion

The major findings and contributions of my thesis are:

1. The deep sea microbial community structure was analyzed by using culture-free biological tools including clone library construction and phylogenetic analysis. A surface water microorganism, *T. pseudonana*, was found in the deep ocean with moderate activity which was confirmed by RT-qPCR analysis of the ratio of 23S rRNA to 23S rDNA abundance. My assertion is that these photoautotrophic microbes are natural habitants at this depth. If these photoautotrophic microorganisms were brought down by convective forces or microbial assemblages associated with large sinking particles, other photoautotrophic microorganisms extant in surface water should also be observed in the samples. However, only two such species were found. These results provide evidence to indicate that previous information about microorganism distribution in the ocean may not be correct. Because of the natural presence of *T. pseudonana* in diverse marine environments, it is proposed as a candidate organism to serve as a bioreporter in monitoring environmental perturbations in the ocean.
2. For the first time, single cell gene expression for *T. pseudonana* was achieved by tube-based two-step RT-qPCR and a lognormal distribution of single-cell gene expression was observed. The lognormal distribution indicates that the geometric average becomes more important and more representative of the whole population than the arithmetic average. This highly skewed distribution again emphasizes the necessity of single cell gene expression analysis. The single cell analysis also

helped to elucidate the stress response of *T. pseudonana* to different nutrient-limited conditions. Those different responses suggest a solid biological foundation for using single cell gene expression as a detection tool for environmental perturbation.

3. In order to achieve *in situ* monitoring, a chip level device was built. This chip consisted of surface-adhesive droplets covered by oil to prevent evaporation and cross contamination during thermal cycling. The chip was built with materials which are commonly available to conventional laboratories. In particular, the chip production process does not require any specialty in microfabrication. More importantly, no off-chip work, such as cell lysing or RNA/DNA purification, was needed for the chip and it could achieve cell-to-data analysis which minimizes user error. With this chip, single cell one-step RT-qPCR analysis was achieved for *T. pseudonana*. The overall success rate of the on-chip reactions at the single cell level was about 85%. The results of the chip-level single cell analysis were confirmed by previous tube-based two-step single cell analysis results. This chip not only makes it possible to develop a future deployable sensor system in the ocean but it also extends single cell analysis methods to conventional biological and environmental laboratories which commonly lack facilities for or expertise in microfabrication.
4. Although single cell gene expression analysis has been achieved, how to interpret single cell RT-qPCR results in a quantitative and effective way is far from mature. My work tried to use some non-parametric statistical tools to analyze expression data. Such tools may provide more accurate results compared with parametric

statistical tools which are based on normal distribution. The only drawback of non-parametric statistical tools is sample size. Non-parametric statistical tools require larger sample sizes than parametric statistical tools to achieve the same power.

7.2. Future work

Although my research built a solid foundation for developing a deployable environmental sensing device, there are still other problems that need to be solved before application in the field. For instance, sampling is a challenging step for a future deployable device. Since environmental samples contain a large number of other species, a sample sorting step will be needed to reject non-target species. After sorting, the concentration of the targeted species will increase and the specificity of the single cell analyses will depend on the specific primer for RT-qPCR. In addition to this, the current version of chip cannot work well in the ocean environment, since oil may spill out due to ocean waves. Sealing of the whole chip in a chamber may be necessary to prevent the oil spill.

As previously mentioned, non-parametric tools require larger sample sizes in order to analyze the results in a reliable way. A larger sample size requires a larger throughput. Currently, due to some technical hurdles and biological issues, including RNA preservation difficulty and chip-to-chip variance, throughput at the single cell level was limited to 48 reactions per chip in this work. Detailed information can be found in Appendix A. Because of the throughput limitation, currently the power of single cell analysis did not reach more than 90%. But, with those results, I could make better estimation about the sample size for future applications. Based on the sample size test results, most of the genes only required a few more samples to reach 90% power. The

results demonstrated that despite the sample size limitations, the chip still performed with acceptable power. To achieve more reliable results, for example over 95% power, around 2-3 times higher sample size which equivalent to 30-60 cells are needed. In order to overcome those technical hurdles to reach larger sample size, a customized thermal cycler and signal detection module needs to be developed. The thermal cycler should be a miniature system and energy-efficient so it can be directly adopted by the deployable sensor system. For the signal detection module, the sensitivity is the most important factor. In order to achieve single cell resolution, high sensitivity is required. All the customized parts are expected to enable higher throughput which will significantly improve the power of the chip. However, increasing the throughput is not necessarily equivalent to increasing the chip size. The throughput of the chip can be increased by decreasing the reaction volume and maintaining the size of the chip. Decreasing the reaction volume may cause potential issues such as PCR inhibition since no purification step was used in this work. Tests will need to be performed to evaluate the most suitable reaction volume. However, with higher throughput, more reliable single cell level gene expression data can be generated and provide better support for environmental monitoring.

REFERENCES

Abdel-Ghany, Salah E, Hong Ye, Gulnara F Garifullina, Lihong Zhang, Elizabeth A H Pilon-Smits, and Marinus Pilon. 2005. “Iron-sulfur cluster biogenesis in chloroplasts. involvement of the scaffold protein CpIscA.” *Plant Physiology* 138 (1): 161–172.

Alldredge, Alice L., and Yehuda Cohen. 1987. “Can microscale chemical patches persist in the sea? Microelectrode study of marine snow, fecal pellets.” *Science* 235 (4789): 689–691.

Allen, Andrew E., Julie LaRoche, Uma Maheswari, Markus Lommer, Nicolas Schauer, Pascal J. Lopez, Giovanni Finazzi, Alisdair R. Fernie, and Chris Bowler. 2008. “Whole-cell response of the pennate diatom *Phaeodactylum Tricornutum* to iron starvation.” *Proceedings of the National Academy of Sciences* 105 (30): 10438 – 10443.

Allen, Eric E., Gene W. Tyson, Rachel J. Whitaker, John C. Detter, Paul M. Richardson, and Jillian F. Banfield. 2007. “Genome dynamics in a natural archaeal population.” *Proceedings of the National Academy of Sciences* 104 (6): 1883–1888.

Altindal, Tuba, Suddhashil Chattopadhyay, and Xiao-Lun Wu. 2011. “Bacterial chemotaxis in an optical trap.” *PloS One* 6 (4): e18231.

Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schäffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman. 1997. “Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.” *Nucleic Acids Research* 25 (17): 3389–3402.

Amann, R I, W Ludwig, and K H Schleifer. 1995. “Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation.” *Microbiological Reviews* 59 (1): 143–169.

Anis, Yasser, Mark Holl, and Deirdre Meldrum. 2010. “Automated selection and placement of single cells using vision-based feedback control.” *IEEE Transactions on Automation Science and Engineering* 7 (3): 598–606.

Anis, Yasser, Jeffrey Houkal, Mark Holl, Roger Johnson, and Deirdre Meldrum. 2011. “Diaphragm pico-liter pump for single-cell manipulation.” *Biomedical Microdevices* 13 (4): 651–659.

Aoi, Yoshiteru. 2002. “*In situ* identification of microorganisms in biofilm communities.” *Journal of Bioscience and Bioengineering* 94 (6): 552–556.

Armbrust, E. Virginia, John A. Berges, Chris Bowler, Beverley R. Green, Diego Martinez, Nicholas H. Putnam, Shiguo Zhou, et al. 2004. “The genome of the diatom

Thalassiosira Pseudonana: Ecology, evolution, and metabolism." *Science* 306 (5693): 79–86.

Arumugam, Prabhu U, Hua Chen, Alan M Cassell, and Jun Li. 2007. "Dielectrophoretic trapping of single bacteria at carbon nanofiber nanoelectrode arrays." *The Journal of Physical Chemistry. A* 111 (49): 12772–12777.

Ashida, Naoaki, Satoshi Ishii, Sadakazu Hayano, Kanako Tago, Takashi Tsuji, Yoshitaka Yoshimura, Shigeto Otsuka, and Keishi Senoo. 2010. "Isolation of functional single cells from environments using a micromanipulator: Application to study denitrifying bacteria." *Applied Microbiology and Biotechnology* 85 (4): 1211–1217.

Ashkin, A, and J M Dziedzic. 1987. "Optical trapping and manipulation of viruses and bacteria." *Science* 235 (4795): 1517–1520.

Ashkin, A, J M Dziedzic, and T Yamane. 1987. "Optical trapping and manipulation of single cells using infrared laser beams." *Nature* 330 (6150): 769–771.

Babendure, Jeremy R, Stephen R Adams, and Roger Y Tsien. 2003. "Aptamers switch on fluorescence of triphenylmethane dyes." *Journal of the American Chemical Society* 125 (48): 14716–14717.

Banerjee, B, S Balasubramanian, G Ananthakrishna, T V Ramakrishnan, and G V Shivashankar. 2004. "Tracking operator state fluctuations in gene expression in single cells." *Biophysical Journal* 86 (5): 3052–3059.

Baptista, M, P Rodrigues, F Depardieu, P Courvalin, and M Arthur. 1999. "Single-cell analysis of glycopeptide resistance gene expression in teicoplanin-resistant mutants of a VanB-type *Enterococcus Faecalis*." *Molecular Microbiology* 32 (1): 17–28.

Becskei, Attila, Benjamin B Kaufmann, and Alexander van Oudenaarden. 2005. "Contributions of low molecule number and chromosomal positioning to stochastic gene expression." *Nature Genetics* 37 (9): 937–944.

Beer, N Reginald, Benjamin J Hindson, Elizabeth K Wheeler, Sara B Hall, Klint A Rose, Ian M Kennedy, and Bill W Colston. 2007. "On-chip, real-time, single-copy polymerase chain reaction in picoliter droplets." *Analytical Chemistry* 79 (22): 8471–8475.

Behrenfeld, Michael J., James T. Randerson, Charles R. McClain, Gene C. Feldman, Sietse O. Los, Compton J. Tucker, Paul G. Falkowski, et al. 2001. "Biospheric primary production during an ENSO transition." *Science* 291 (5513): 2594–2597.

Bengtsson, Martin, Martin Hemberg, Patrik Rorsman, and Anders Ståhlberg. 2008. "Quantification of mRNA in single cells and modeling of RT-qPCR induced noise." *Bmc Molecular Biology* 9.

Bengtsson, Martin, Anders Ståhlberg, Patrik Rorsman, and Mikael Kubista. 2005. “Gene expression profiling in single cells from the pancreatic islets of langerhans reveals lognormal distribution of mRNA Levels.” *Genome Research* 15 (10): 1388–1392.

Bennett, Joseph R., Brian F. Cumming, Brian K. Ginn, and John P. Smol. 2010. “Broad-scale environmental response and niche conservatism in lacustrine diatom communities.” *Global Ecology and Biogeography* 19 (5): 724–732.

Bertrand, E, P Chartrand, M Schaefer, S M Shenoy, R H Singer, and R M Long. 1998. “Localization of *ASH1* mRNA particles in living yeast.” *Molecular Cell* 2 (4): 437–445.

Blake, W. J., M. Kaern, C. R. Cantor, and J. J. Collins. 2003. “Noise in eukaryotic gene expression.” *Nature* 422 (6932): 633–637.

Block, S M, D F Blair, and H C Berg. 1989. “Compliance of bacterial flagella measured with optical tweezers.” *Nature* 338 (6215): 514–518.

Boedicker, JQ, L Li, TR Kline, and RF Ismagilov. 2008. “Detecting bacteria and determining their susceptibility to antibiotics by stochastic confinement in nanoliter droplets using plug-based microfluidics.” *Lab Chip* 8 (8): 1265–1272.

Borland, Laura M, Sumith Kottegoda, K Scott Phillips, and Nancy L Allbritton. 2008. “Chemical analysis of single cells.” *Annual Review of Analytical Chemistry (Palo Alto, Calif.)* 1: 191–227.

Bower, Neil Ivan, Ralf Joachim Moser, Jonathan Robert Hill, and Sigrid Arabella Lehnert. 2007. “Universal reference method for real-time PCR gene expression analysis of preimplantation embryos.” *BioTechniques* 42 (2): 199–206.

Bowler, Chris, Andrew E. Allen, Jonathan H. Badger, Jane Grimwood, Kamel Jabbari, Alan Kuo, Uma Maheswari, et al. 2008. “The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes.” *Nature* 456 (7219): 239–244.

Boyd, P W, T Jickells, C S Law, S Blain, E A Boyle, K O Buesseler, K H Coale, et al. 2007. “Mesoscale iron enrichment experiments 1993-2005: Synthesis and future directions.” *Science* 315 (5812): 612–617.

Brady, G, and N N Iscove. 1993. “Construction of cDNA libraries from single cells.” *Methods in Enzymology* 225: 611–623.

Brehm-Stecher, Byron F, and Eric A Johnson. 2004. “Single-cell microbiology: Tools, technologies, and applications.” *Microbiology and Molecular Biology Reviews: MMBR* 68 (3): 538–559.

Brouzes, Eric, Martina Medkova, Neal Savenelli, Dave Marran, Mariusz Twardowski, J. Brian Hutchison, Jonathan M. Rothberg, Darren R. Link, Norbert Perrimon, and Michael L. Samuels. 2009. “Droplet microfluidic technology for single-cell high-throughput screening.” *Proceedings of the National Academy of Sciences* 106 (34): 14195–14200.

Bucciarelli, Eva, and William G. Sunda. 2003. “Influence of CO2, nitrate, phosphate, and silicate limitation on intracellular dimethylsulfoniopropionate in batch cultures of the coastal diatom *Thalassiosira Pseudonana*.” *Limnology and Oceanography* 48 (6): 2256–2265.

Bulmer, M. G. 1979. *Principles of Statistics*. Courier Dover Publications.

Bustin, S A. 2002. “Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): Trends and problems.” *Journal of Molecular Endocrinology* 29 (1): 23–39.

Bustin, Stephen A, Vladimir Benes, Jeremy A Garson, Jan Hellemans, Jim Huggett, Mikael Kubista, Reinhold Mueller, et al. 2009. “The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments.” *Clinical Chemistry* 55 (4): 611–622.

Bustin, Stephen A, and Tania Nolan. 2004. “Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction.” *Journal of Biomolecular Techniques: JBT* 15 (3): 155–166.

Button, D K, F Schut, P Quang, R Martin, and B R Robertson. 1993. “Viability and isolation of marine bacteria by dilution culture: Theory, procedures, and initial results.” *Applied and Environmental Microbiology* 59 (3): 881–891.

Calapez, Alexandre, Henrique M Pereira, Angelo Calado, José Braga, José Rino, Célia Carvalho, João Paulo Tavanez, Elmar Wahle, Agostinho C Rosa, and Maria Carmo-Fonseca. 2002. “The intranuclear mobility of messenger RNA binding proteins is ATP dependent and temperature sensitive.” *The Journal of Cell Biology* 159 (5): 795–805.

Campàs, Mònica, and Jean-Louis Marty. 2007. “Enzyme sensor for the electrochemical detection of the marine toxin okadaic acid.” *Analytica Chimica Acta* 605 (1): 87–93.

Cardullo, R A, S Agrawal, C Flores, P C Zamecnik, and D E Wolf. 1988. “Detection of nucleic acid hybridization by nonradiative fluorescence resonance energy transfer.” *Proceedings of the National Academy of Sciences of the United States of America* 85 (23): 8790–8794.

Carlo, Dino Di, and Luke P. Lee. 2006. “Dynamic single-cell analysis for quantitative biology.” *Analytical Chemistry* 78 (23): 7918–7925.

Carmon, G, and M Feingold. 2011. “Rotation of single bacterial cells relative to the optical axis using optical tweezers.” *Optics Letters* 36 (1): 40–42.

Cattolico, Rose A., Michael A. Jacobs, Yang Zhou, Jean Chang, Melinda Duplessis, Terry Lybrand, John McKay, Han C. Ong, Elizabeth Sims, and Gabrielle Rocap. 2008. “Chloroplast genome sequencing analysis of *Heterosigma Akashiwo* CCMP452 (West Atlantic) and NIES293 (West Pacific) strains.” *BMC Genomics* 9 (1): 211.

Chalfie, M., Y. Tu, G. Euskirchen, W. W. Ward, and D. C. Prasher. 1994. “Green fluorescent protein as a marker for gene expression.” *Science* 263 (5148): 802–805.

Chao, Shih-hui, Robert Carlson, and Deirdre R Meldrum. 2007. “Rapid fabrication of microchannels using microscale plasma activated templating (microPLAT) generated water molds.” *Lab on a Chip* 7 (5): 641–643.

Charette, Matthew A., and Walter H. F. Smith. 2010. “The volume of earth’s ocean.” *Oceanography* 23 (2): 112–114.

Charlson, Robert J., James E. Lovelock, Meinrat O. Andreae, and Stephen G. Warren. 1987. “Oceanic phytoplankton, atmospheric sulphur, cloud albedo and climate.” *Nature* 326 (6114): 655–661.

Choi, Paul J., Long Cai, Kirsten Frieda, and X. Sunney Xie. 2008. “A stochastic single-molecule event triggers phenotype switching of a bacterial cell.” *Science* 322 (5900): 442–446.

Clark, N. C., R. C. Cooksey, B. C. Hill, J. M. Swenson, and F. C. Tenover. 1993. “Characterization of glycopeptide-resistant enterococci from U.S. hospitals.” *Antimicrobial Agents and Chemotherapy* 37 (11): 2311–2317.

Cole, J R, B Chai, R J Farris, Q Wang, S A Kulam, D M McGarrell, G M Garrity, and J M Tiedje. 2005. “The ribosomal database project (RDP-II): Sequences and tools for high-throughput rRNA analysis.” *Nucleic Acids Research* 33 (Database issue): D294–296.

Colman-Lerner, Alejandro, Andrew Gordon, Eduard Serra, Tina Chin, Orna Resnekov, Drew Endy, C Gustavo Pesce, and Roger Brent. 2005. “Regulated cell-to-cell variation in a cell-fate decision system.” *Nature* 437 (7059): 699–706.

Czechowski, Tomasz, Mark Stitt, Thomas Altmann, Michael K Udvardi, and Wolf-Rüdiger Scheible. 2005. “Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*.” *Plant Physiology* 139 (1): 5–17.

Daigle, Nathalie, and Jan Ellenberg. 2007. “LambdaN-GFP: An RNA reporter system for live-cell imaging.” *Nature Methods* 4 (8): 633–636.

DeLong, E. F. 1992. “Archaea in coastal marine environments.” *Proceedings of the National Academy of Sciences* 89 (12): 5685–5689.

DeLong, Edward F. 2009. “The microbial ocean from genomes to biomes.” *Nature* 459 (7244): 200–206.

DeLong, Edward F., Christina M. Preston, Tracy Mincer, Virginia Rich, Steven J. Hallam, Niels-Ulrik Frigaard, Asuncion Martinez, et al. 2006. “Community genomics among stratified microbial assemblages in the ocean’s interior.” *Science* 311 (5760): 496–503.

Di Carlo, Dino, Liz Y Wu, and Luke P Lee. 2006. “Dynamic single cell culture array.” *Lab on a Chip* 6 (11): 1445–1449.

Diehl, Frank, Meng Li, Yiping He, Kenneth W Kinzler, Bert Vogelstein, and Devin Dressman. 2006. “BEAMing: single-molecule PCR on microparticles in water-in-oil emulsions.” *Nature Methods* 3 (7): 551–559.

Dixon, A K, P J Richardson, R D Pinnock, and K Lee. 2000. “Gene-expression analysis at the single-cell level.” *Trends in Pharmacological Sciences* 21 (2): 65–70.

Doney, Scott C. 2010. “The growing human footprint on coastal and open-ocean biogeochemistry.” *Science* 328 (5985): 1512–1516.

Dore, John E., Roger Lukas, Daniel W. Sadler, and David M. Karl. 2003. “Climate-driven changes to the atmospheric CO2 sink in the subtropical North Pacific Ocean.” *Nature* 424 (6950): 754–757.

Dorfman, Kevin D, Max Chabert, Jean-Hugues Codarbox, Gilles Rousseau, Patricia de Cremoux, and Jean-Louis Viovy. 2005. “Contamination-free continuous flow microfluidic polymerase chain reaction for quantitative and clinical applications.” *Analytical Chemistry* 77 (11): 3700–3704.

Elowitz, M. B., A. J. Levine, E. D. Siggia, and P. S. Swain. 2002. “Stochastic gene expression in a single cell.” *Science* 297 (5584): 1183–1186.

Enger, Jonas, Mattias Goksör, Kerstin Ramser, Petter Hagberg, and Dag Hanstorp. 2004. “Optical tweezers applied to a microfluidic system.” *Lab on a Chip* 4 (3): 196.

Eppley, Richard W., and Bruce J. Peterson. 1979. “Particulate organic matter flux and planktonic new production in the deep ocean.” *Nature* 282 (5740): 677–680.

Eun, Ye-Jin, Andrew S Utada, Matthew F Copeland, Shoji Takeuchi, and Douglas B Weibel. 2011. “Encapsulating bacteria in agarose microparticles using microfluidics for high-throughput cell analysis and isolation.” *ACS Chemical Biology* 6 (3): 260–266.

Falkowski, Paul G, Miriam E Katz, Andrew H Knoll, Antonietta Quigg, John A Raven, Oscar Schofield, and F J R Taylor. 2004. “The evolution of modern eukaryotic phytoplankton.” *Science* 305 (5682): 354–360.

Field, Christopher B., Michael J. Behrenfeld, James T. Randerson, and Paul Falkowski. 1998. “Primary production of the biosphere: Integrating terrestrial and oceanic components.” *Science* 281 (5374): 237–240.

Fu, A Y, C Spence, A Scherer, F H Arnold, and S R Quake. 1999. “A microfabricated fluorescence-activated cell sorter.” *Nature Biotechnology* 17 (11): 1109–1111.

Fuchs, P, W Weichel, S Dübel, F Breitling, and M Little. 1996. “Separation of *E. Coli* expressing functional cell-wall bound antibody fragments by FACS.” *Immunotechnology: An International Journal of Immunological Engineering* 2 (2): 97–102.

Furutani, Shunsuke, Hidenori Nagai, Yuzuru Takamura, and Izumi Kubo. 2010. “Compact Disk (CD)-shaped device for single cell isolation and PCR of a specific gene in the isolated cell.” *Analytical and Bioanalytical Chemistry* 398 (7-8): 2997–3004.

Galloway, J. N., F. J. Dentener, D. G. Capone, E. W. Boyer, R. W. Howarth, S. P. Seitzinger, G. P. Asner, et al. 2004. “Nitrogen cycles: Past, present, and future.” *Biogeochemistry* 70 (2): 153–226.

Gao, Weimin, Xu Shi, Jieying Wu, Yuguang Jin, Weiwen Zhang, and Deirdre R Meldrum. 2011. “Phylogenetic and gene expression analysis of cyanobacteria and diatoms in the twilight waters of the temperate Northeast Pacific Ocean.” *Microbial Ecology* 62 (4):765-775.

Gao, Weimin, Weiwen Zhang, and Deirdre R Meldrum. 2011. “RT-qPCR based quantitative analysis of gene expression in single bacterial cells.” *Journal of Microbiological Methods* 85 (3) (June): 221–227.

Gasch, Audrey P., Paul T. Spellman, Camilla M. Kao, Orna Carmel-Harel, Michael B. Eisen, Gisela Storz, David Botstein, and Patrick O. Brown. 2000. “Genomic expression programs in the response of yeast cells to environmental changes.” *Molecular Biology of the Cell* 11 (12): 4241–4257.

Gill, Ryan T., Eva Katsoulakis, William Schmitt, Gaspar Taroncher-Oldenburg, Jatin Misra, and Gregory Stephanopoulos. 2002. “Genome-wide dynamic transcriptional profiling of the light-to-dark transition in *Synechocystis Sp.* Strain PCC 6803.” *Journal of Bacteriology* 184 (13): 3671–3681.

Giovannoni, Stephen J., Theresa B. Britschgi, Craig L. Moyer, and Katharine G. Field. 1990. “Genetic diversity in Sargasso sea bacterioplankton.” *Nature* 345 (6270): 60–63.

Giovannoni, Stephen, and Ulrich Stingl. 2007. “The importance of culturing bacterioplankton in the ‘Omics’ age.” *Nature Reviews Microbiology* 5 (10): 820–826.

Glanzer, J G, and J H Eberwine. 2004. “Expression profiling of small cellular samples in cancer: Less is more.” *British Journal of Cancer* 90 (6): 1111–1114.

Goidin, D, A Mamessier, M J Staquet, D Schmitt, and O Berthier-Vergnes. 2001. “Ribosomal 18S RNA prevails over glyceraldehyde-3-phosphate dehydrogenase and beta-actin genes as internal standard for quantitative comparison of mRNA levels in invasive and noninvasive human melanoma cell subpopulations.” *Analytical Biochemistry* 295 (1): 17–21.

Golding, Ido, and Edward C Cox. 2004. “RNA dynamics in live *Escherichia Coli* cells.” *Proceedings of the National Academy of Sciences of the United States of America* 101 (31): 11310–11315.

Golding, Ido, Johan Paulsson, Scott M Zawilski, and Edward C Cox. 2005. “Real-time kinetics of gene activity in individual bacteria.” *Cell* 123 (6): 1025–1036.

Grate, D, and C Wilson. 1999. “Laser-mediated, site-specific inactivation of RNA transcripts.” *Proceedings of the National Academy of Sciences of the United States of America* 96 (11): 6131–6136.

Greene, Richard M., Richard J. Geider, and Paul G. Falkowski. 1991. “Effect of iron limitation on photosynthesis in a marine diatom.” *Limnology and Oceanography* 36 (8): 1772–1782.

Gregg, Watson W., Margarita E. Conkright, Paul Ginoux, John E. O’Reilly, and Nancy W. Casey. 2003. “Ocean primary production and climate: Global decadal changes.” *Geophysical Research Letters* 30 (15): 1809–1812.

Guillard, R. R. L. 1962. “Studies of marine planktonic diatoms.I.*Cyclotella Nana Hustedt* and *Detonula Confervacea Cleve*.” *Can.J.Microbiol.* 8: 229–239.

Guillard, R. R. L. 1975. “Culture of phytoplankton for feeding marine invertebrates.” *Culture of Marine Invertebrates*: 29–60.

Guo, Mira T, Assaf Rotem, John A Heyman, and David A Weitz. 2012. “Droplet microfluidics for high-throughput biological assays.” *Lab on a Chip* 12 (12): 2146–2155.

Hahn, S, X Y Zhong, C Troeger, R Burgemeister, K Gloning, and W Holzgreve. 2000. "Current applications of single-cell PCR." *Cellular and Molecular Life Sciences: CMLS* 57 (1): 96–105.

Hallam, Steven J., Konstantinos T. Konstantinidis, Nik Putnam, Christa Schleper, Yoh-ichi Watanabe, Junichi Sugahara, Christina Preston, José de la Torre, Paul M. Richardson, and Edward F. DeLong. 2006. "Genomic analysis of the uncultivated marine crenarchaeote *Cenarchaeum Symbiosum*." *Proceedings of the National Academy of Sciences* 103 (48): 18296–18301.

Hebenstreit, Daniel. 2012. "Methods, challenges and potentials of single cell RNA-seq." *Biology* 1 (3): 658–667.

Heid, C A, J Stevens, K J Livak, and P M Williams. 1996. "Real time quantitative PCR." *Genome Research* 6 (10): 986–994.

Hein, Mette, and Kaj Sand-Jensen. 1997. "CO_2 increases oceanic primary production." *Nature* 388 (6642): 526–527.

Hirakata, Yoichi, Koichi Izumikawa, Toshiyuki Yamaguchi, Hiromu Takemura, Hironori Tanaka, Ryoji Yoshida, Junichi Matsuda, et al. 1998. "Rapid detection and evaluation of clinical characteristics of emerging multiple-drug-resistant Gram-negative rods Carrying the *Metallo-β-Lactamase Genebla IMP*." *Antimicrobial Agents and Chemotherapy* 42 (8): 2006–2011.

Hockin, Nicola Louise, Thomas Mock, Francis Mulholland, Stanislav Kopriva, and Gill Malin. 2012. "The response of diatom central carbon metabolism to nitrogen starvation is different from that of green algae and higher Plants." *Plant Physiology* 158 (1): 299–312.

Hodson, R E, W A Dustman, R P Garg, and M A Moran. 1995. "*In situ* PCR for visualization of microscale distribution of specific genes and gene products in prokaryotic communities." *Applied and Environmental Microbiology* 61 (11): 4074–4082.

Höss, M, M Kohn, S Pääbo, F Knauer, and W Schröder. 1992. "Excrement analysis by PCR." *Nature* 359 (6392): 199–199.

Höss, M, and S Pääbo. 1993. "DNA extraction from pleistocene bones by a silica-based purification method." *Nucleic Acids Research* 21 (16): 3913–3914.

Huang, Bo, Hongkai Wu, Devaki Bhaya, Arthur Grossman, Sebastien Granier, Brian K Kobilka, and Richard N Zare. 2007. "Counting low-copy number proteins in a singlecell." *Science* 315 (5808): 81–84.

Huggett, J, K Dheda, S Bustin, and A Zumla. 2005. “Real-time RT-PCR normalisation; Strategies and considerations.” *Genes and Immunity* 6 (4): 279–284.

Hwang, Chiachi, Weimin Wu, Terry J. Gentry, Jack Carley, Gail A. Corbin, Sue L. Carroll, David B. Watson, et al. 2009. “Bacterial community succession during *in situ* uranium bioremediation: Spatial similarities along controlled flow paths.” *The ISME Journal* 3 (1): 47–64.

Ikeda, Tsutomu, Fumikazu Sano, Atsushi Yamaguchi, and Takashi Matsuishi. 2007. “RNA: DNA ratios of calanoid copepods from the epipelagic through abyssopelagic zones of the North Pacific Ocean.” *Aquatic Biology* 1 (2): 99–108.

Iscove, Norman N, Mary Barbara, Marie Gu, Meredith Gibson, Carolyn Modi, and Neil Winegarden. 2002. “Representation is faithfully preserved in global cDNA amplified exponentially from sub-picogram quantities of mRNA.” *Nature Biotechnology* 20 (9): 940–943.

Johnson, Kenneth S., R.Michael Gordon, and Kenneth H. Coale. 1997. “What controls dissolved iron concentrations in the world ocean?” *Marine Chemistry* 57 (3–4): 137–161.

Kang, Yun, Michael H Norris, Jan Zarzycki-Siek, William C Nierman, Stuart P Donachie, and Tung T Hoang. 2011. “Transcript amplification from single bacterium for transcriptome analysis.” *Genome Research* 21 (6): 925–935.

Karube, Isao, Tadashi Matsunaga, Satoshi Mitsuda, and Shuichi Suzuki. 1977. “Microbial electrode BOD Sensors.” *Biotechnology and Bioengineering* 19 (10): 1535–1547.

Kelly, C D, and O Rahn. 1932. “The growth rate of individual bacterial cells.” *Journal of Bacteriology* 23 (2): 147–153.

Kerppola, Tom K. 2006. “Visualization of molecular interactions by fluorescence complementation.” *Nature Reviews. Molecular Cell Biology* 7 (6): 449–456.

Kester, Dana R., Iver W. Duedall, Donald N. Connors, and Ricardo M. Pytkowicz. 1967. “Preparation of artificial seawater.” *Limnology and Oceanography* 12 (1): 176–179.

Kim, Bo-Ra, Hee-Young Nam, Soo-Un Kim, Su-Il Kim, and Yung-Jin Chang. 2003. “Normalization of reverse transcription quantitative-PCR with housekeeping genes in rice.” *Biotechnology Letters* 25 (21): 1869–1872.

Kiss, Margaret Macris, Lori Ortoleva-Donnelly, N Reginald Beer, Jason Warner, Christopher G Bailey, Bill W Colston, Jonathon M Rothberg, Darren R Link, and John H Leamon. 2008. “High-throughput quantitative polymerase chain reaction in picoliter droplets.” *Analytical Chemistry* 80 (23): 8975–8981.

Kitano, Hiroaki. 2002. “Computational systems biology.” *Nature* 420 (6912): 206–210.

Ko, Minoru S. H. 1992. “Problems and paradigms: Induction mechanism of a single gene molecule: Stochastic or deterministic?” *BioEssays* 14 (5): 341–346.

Kojima, Takaaki, Yoshiaki Takei, Miharu Ohtsuka, Yasuaki Kawarasaki, Tsuneo Yamane, and Hideo Nakano. 2005. “PCR amplification from single DNA molecules on magnetic beads in emulsion: Application for High-throughput screening of transcription factor targets.” *Nucleic Acids Research* 33 (17): e150.

Kolber, Zbigniew, Jonathan Zehr, and Paul Falkowski. 1988. “Effects of growth irradiance and nitrogen limitation on photosynthetic energy conversion in photosystem II 1.” *Plant Physiology* 88 (3): 923–929.

Kreader, CA. 1996. “Relief of amplification inhibition in PCR with bovine serum Albumin or T4 Gene 32 Protein.” *Appl. Environ. Microbiol.* 62 (3): 1102–1106.

Kuang, Yina, Israel Biran, and David R Walt. 2004. “Simultaneously monitoring gene expression kinetics and genetic noise in single cells by optical well arrays.” *Analytical Chemistry* 76 (21): 6282–6286.

Kubista, Mikael, José Manuel Andrade, Martin Bengtsson, Amin Forootan, Jiri Jonák, Kristina Lind, Radek Sindelka, et al. 2006. “The real-time polymerase chain reaction.” *Molecular Aspects of Medicine* 27 (2-3): 95–125.

Kühn, S, P Measor, E J Lunt, B S Phillips, D W Deamer, A R Hawkins, and H Schmidt. 2009. “Loss-based optical trap for on-chip particle analysis.” *Lab on a Chip* 9 (15): 2212–2216.

Kurimoto, Kazuki, Yukihiro Yabuta, Yasuhide Ohinata, and Mitinori Saitou. 2007. “Global single-cell cDNA amplification to provide a template for representative high-density oligonucleotide microarray analysis.” *Nature Protocols* 2 (3): 739–752.

Kustka, Adam B., Andrew E. Allen, and François M. M. Morel. 2007. “Sequence analysis and transcriptional regulation of rion acquisition genes in two marine diatoms.” *Journal of Phycology* 43 (4): 715–729.

Lange, M, T Tolker-Nielsen, S Molin, and B K Ahring. 2000. “*In situ* reverse transcription-PCR for monitoring gene expression in individual *Methanosarcina Mazei S-6* cells.” *Applied and Environmental Microbiology* 66 (5): 1796–1800.

Le Quéré, Corinne, Michael R. Raupach, Josep G. Canadell, Gregg Marland et Al, Corinne Le Quéré et Al, Corinne Le Quéré et Al, Michael R. Raupach, et al. 2009. “Trends in the sources and sinks of carbon dioxide.” *Nature Geoscience* 2 (12): 831–836.

Le, Thuc T, Sébastien Harlepp, Calin C Guet, Kimberly Dittmar, Thierry Emonet, Tao Pan, and Philippe Cluzel. 2005. “Real-time RNA profiling within a single bacterium.” *Proceedings of the National Academy of Sciences of the United States of America* 102 (26): 9160–9164.

Lee, CY, GB Lee, JL Lin, FC Huang, and CS Liao. 2005. “Integrated microfluidic systems for cell lysis, mixing/pumping and DNA amplification.” *Journal of Micromechanics and Microengineering* 15 (6): 1215–1223.

Lenz, Ailyn P, Kerry S Williamson, Betsey Pitts, Philip S Stewart, and Michael J Franklin. 2008. “Localized gene expression in *Pseudomonas Aeruginosa* biofilms.” *Applied and Environmental Microbiology* 74 (14): 4463–4471.

Levsky, Jeffrey M., Shailesh M. Shenoy, Rossanna C. Pezo, and Robert H. Singer. 2002. “Single-cell gene expression profiling.” *Science* 297 (5582): 836–840.

Lewandowska, Jolanta, and Alicja Kosakowska. 2004. “Effect of iron limitation on cells of the diatom *Cyclotella Meneghiniana Kützing*.” *Global Biogeochemical Cycles* 46 (2): 269–287.

Li, Qingge, Guoyan Luan, Qiuping Guo, and Jixuan Liang. 2002. “A new class of homogeneous nucleic acid probes based on specific displacement hybridization.” *Nucleic Acids Research* 30 (2): E5.

Li, You-Rong, and Ju Chu. 1991. “Study of BOD Microbial Sensors for Waste Water Treatment Control.” *Applied Biochemistry and Biotechnology* 28-29 (1): 855–863.

Lidstrom, Mary E, and Deirdre R Meldrum. 2003. “Life-on-a-chip.” *Nature Reviews. Microbiology* 1 (2): 158–164.

Lin, Liang-I, Shih-Hui Chao, and Deirdre R Meldrum. 2009. “Practical, microfabrication-free device for single-cell isolation.” *PloS One* 4 (8): e6710.

Lindqvist, Niclas, Manuel Vidal-Sanz, and Finn Hallböök. 2002. “Single cell RT-PCR analysis of tyrosine kinase receptor expression in adult rat retinal ganglion cells isolated by retinal sandwiching.” *Brain Research. Brain Research Protocols* 10 (2): 75–83.

Lindström, Sara, and Helene Andersson-Svahn. 2010. “Overview of single-cell analyses: microdevices and applications.” *Lab Chip* 10(24): 3363–3372.

Liu, Jing, and Bo Mattiasson. 2002. “Microbial BOD sensors for wastewater analysis.” *Water Research* 36 (15): 3786–3802.

Livak, K J, and T D Schmittgen. 2001. “Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method.” *Methods*25 (4): 402–408.

Lomas, Michael W, Deborah A Bronk, and Ger van den Engh. 2011. “Use of flow cytometry to measure biogeochemical rates and processes in the ocean.” *Annual Review of Marine Science* 3: 537–566.

Longhurst, Alan, Shubha Sathyendranath, Trevor Platt, and Carla Caverhill. 1995. “An estimate of global primary production in the ocean from satellite radiometer data.” *Journal of Plankton Research* 17 (6): 1245–1271.

Longo, Diane, and Jeff Hasty. 2006. “Dynamics of single-cell gene expression.” *Molecular Systems Biology* 2: 64.

López-García, Purificación, Francisco Rodríguez-Valera, Carlos Pedrós-Alió, and David Moreira. 2001. “Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton.” *Nature* 409 (6820): 603–607.

Lu, Hang, Martin A Schmidt, and Klavs F Jensen. 2005. “A microfluidic electroporation device for cell lysis.” *Lab Chip* 5 (1): 23–29.

Lutz, Barry R, Jian Chen, and Daniel T Schwartz. 2006. “Hydrodynamic tweezers: 1. noncontact trapping of single cells using steady streaming microeddies.” *Analytical Chemistry* 78 (15): 5429–5435.

Magliery, Thomas J, Christopher G M Wilson, Weilan Pan, Dennis Mishler, Indraneel Ghosh, Andrew D Hamilton, and Lynne Regan. 2005. “Detecting protein-protein interactions with a green fluorescent protein fragment reassembly trap: Scope and mechanism.” *Journal of the American Chemical Society* 127 (1): 146–157.

Maheswari, Uma, Kamel Jabbari, Jean-Louis Petit, Betina M Porcel, Andrew E Allen, Jean-Paul Cadoret, Alessandra De Martino, et al. 2010. “Digital expression profiling of novel diatom transcripts provides insight into their biological functions.” *Genome Biology* 11 (8): R85.

Maiwald, Daniela, Angela Dietzmann, Peter Jahns, Paolo Pesaresi, Pierre Joliot, Anne Joliot, Joshua Z. Levin, Francesco Salamini, and Dario Leister. 2003. “Knock-out of the genes coding for the rieske protein and the ATP-synthase δ-subunit of *Arabidopsis*. Effects on photosynthesis, thylakoid protein composition, and nuclear chloroplast gene expression.” *Plant Physiology* 133 (1): 191–202.

Maloney, P C, and B Rotman. 1973. “Distribution of suboptimally induces -D-galactosidase in *Escherichia Coli*. The enzyme content of individual Cells.” *Journal of Molecular Biology* 73 (1): 77–91.

Mann, D., and S. Droop. 1996. “3. Biodiversity, biogeography and conservation of diatoms.” *Hydrobiologia* 336 (1): 19–32.

Marchetti, Adrian, David M Schruth, Colleen A Durkin, Micaela S Parker, Robin B Kodner, Chris T Berthiaume, Rhonda Morales, Andrew E Allen, and E Virginia Armbrust. 2012. “Comparative metatranscriptomics identifies molecular bases for the physiological responses of phytoplankton to varying iron availability.” *Proceedings of the National Academy of Sciences of the United States of America* 109 (6): E317–325.

Marcus, Joshua S., W. French Anderson, and Stephen R. Quake. 2006. “Microfluidic single-cell mRNA isolation and analysis.” *Analytical Chemistry* 78 (9): 3084–3089.

Marcy, Yann, Thomas Ishoey, Roger S Lasken, Timothy B Stockwell, Brian P Walenz, Aaron L Halpern, Karen Y Beeson, Susanne M D Goldberg, and Stephen R Quake. 2007. “Nanoliter reactors improve multiple displacement amplification of genomes from single cells.” *PLoS Genetics* 3 (9): 1702–1708.

Marcy, Yann, Cleber Ouverney, Elisabeth M Bik, Tina Lösekann, Natalia Ivanova, Hector Garcia Martin, Ernest Szeto, et al. 2007. “Dissecting biological ‘dark matter’ with single-cell genetic analysis of rare and uncultivated TM7 microbes from the human mouth.” *Proceedings of the National Academy of Sciences of the United States of America* 104 (29): 11889–11894.

Markoulatos, P, N Siafakas, and M Moncany. 2002. “Multiplex polymerase chain reaction: A practical approach.” *Journal of Clinical Laboratory Analysis* 16 (1): 47–51.

Martin, Jh, Kh Coale, Ks Johnson, Se Fitzwater, Rm Gordon, Sj Tanner, Cn Hunter, et al. 1994. “Testing the iron hypothesis in ecosystems of the equatorial Pacific-Ocean.” *Nature* 371 (6493): 123–129.

Massana, R, A E Murray, C M Preston, and E F DeLong. 1997. “Vertical distribution and phylogenetic characterization of marine planktonic archaea in the Santa Barbara Channel.” *Applied and Environmental Microbiology* 63 (1): 50–56.

Matsubara, Yasutaka, Kagan Kerman, Masaaki Kobayashi, Shouhei Yamamura, Yasutaka Morita, Yuzuru Takamura, and Eiichi Tamiya. 2004. “On-chip nanoliter-volume multiplex TaqMan polymerase chain reaction from a single copy based on counting fluorescence released microchambers.” *Analytical Chemistry* 76 (21): 6434–6439.

Mcallister, Cd, N. Shah, and Jdh Strickland. 1964. “Marine phytoplankton photosynthesis as a function of light intensity - a comparison of methods.” *Journal of the Fisheries Research Board of Canada* 21 (1): 159–181.

Mckenna, Brian K, A A Selim, F Richard Bringhurst, and Daniel J Ehrlich. 2009. “384-channel parallel microfluidic cytometer for rare-cell screening.” *Lab Chip* 9 (2): 305–310.

Min, Taejin L, Patrick J Mears, Lon M Chubiz, Christopher V Rao, Ido Golding, and Yann R Chemla. 2009. “High-resolution, long-term characterization of bacterial motility using optical tweezers.” *Nature Methods* 6 (11): 831–835.

Mirsaidov, Utkur, Winston Timp, Kaethe Timp, Mustafa Mir, Paul Matsudaira, and Gregory Timp. 2008. “Optimal optical trap for bacterial viability.” *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics* 78 (2): 021910.

Mock, Thomas, Manoj Pratim Samanta, Vaughn Iverson, Chris Berthiaume, Matthew Robison, Karie Holtermann, Colleen Durkin, et al. 2008. “Whole-genome expression profiling of the marine diatom *Thalassiosira Pseudonana* identifies genes involved in silicon bioprocesses.” *Proceedings of the National Academy of Sciences of the United States of America* 105 (5): 1579–1584.

Morrison, L E, T C Halder, and L M Stols. 1989. “Solution-phase detection of polynucleotides using interacting fluorescent labels and competitive hybridization.” *Analytical Biochemistry* 183 (2): 231–244.

Munce, Nigel R, Jianzhao Li, Peter R Herman, and Lothar Lilge. 2004. “Microfabricated system for parallel single-cell capillary electrophoresis.” *Analytical Chemistry* 76 (17): 4983–4989.

Musat, Niculina, Hannah Halm, Bärbel Winterholler, Peter Hoppe, Sandro Peduzzi, Francois Hillion, Francois Horreard, Rudolf Amann, Bo B Jørgensen, and Marcel M M Kuypers. 2008. “A single-cell view on the ecophysiology of anaerobic phototrophic bacteria.” *Proceedings of the National Academy of Sciences of the United States of America* 105 (46): 17861–17866.

Musyanovych, Anna, Volker Mailänder, and Katharina Landfester. 2005. “Miniemulsion droplets as single molecule nanoreactors for polymerase chain reaction.” *Biomacromolecules* 6 (4): 1824–1828.

Nailis, Heleen, Tom Coenye, Filip Van Nieuwerburgh, Dieter Deforce, and Hans J Nelis. 2006. “Development and evaluation of different normalization strategies for gene expression studies in *Candida Albicans* biofilms by Real-time PCR.” *BMC Molecular Biology* 7: 25.

Nakano, Michihiko, Jun Komatsu, Shun-ichi Matsuura, Kazunori Takashima, Shinji Katsura, and Akira Mizuno. 2003. “Single-molecule PCR using water-in-oil emulsion.” *Journal of Biotechnology* 102 (2): 117–124.

Nelson, Dm, P. Treguer, Ma Brzezinski, A. Leynaert, and B. Queguiner. 1995. “Production and dissolution of biogenic silica in the ocean - Revised global estimates, comparison with regional data and relationship to biogenic sedimentation.” *Global Biogeochemical Cycles* 9 (3): 359–372.

Newman, John R S, Sina Ghaemmaghami, Jan Ihmels, David K Breslow, Matthew Noble, Joseph L DeRisi, and Jonathan S Weissman. 2006. “Single-cell proteomic analysis of *S. Cerevisiae* reveals the architecture of biological noise.” *Nature* 441 (7095): 840–846.

Nolan, Tania, Rebecca E Hands, and Stephen A Bustin. 2006. “Quantification of mRNA using real-time RT-PCR.” *Nature Protocols* 1 (3): 1559–1582.

Nossal, G J, A Szenberg, G L Ada, and C M Austin. 1964. “Single cell studies on 19S antibody production.” *The Journal of Experimental Medicine* 119: 485–502.

Ottesen, Elizabeth A, Jong Wook Hong, Stephen R Quake, and Jared R Leadbetter. 2006. “Microfluidic digital PCR enables multigene analysis of individual environmental bacteria.” *Science* 314 (5804): 1464–1467.

Pace, Norman R. 1997. “A molecular view of microbial diversity and the biosphere.” *Science* 276 (5313): 734–740.

Paige, Jeremy S, Karen Y Wu, and Samie R Jaffrey. 2011. “RNA mimics of green fluorescent protein.” *Science* 333 (6042): 642–646.

Panicker, Gitika, Michael L Myers, and Asim K Bej. 2004. “Rapid detection of *Vibrio Vulnificus* in shellfish and gulf of Mexico water by real-time PCR.” *Applied and Environmental Microbiology* 70 (1): 498–507.

Pattyn, Filip, Frank Speleman, Anne De Paepe, and Jo Vandesompele. 2003. “RTPrimerDB: The real-Time PCR primer and probe database.” *Nucleic Acids Research* 31 (1): 122 –123.

Pedraza, Juan M, and Alexander van Oudenaarden. 2005. “Noise propagation in gene networks.” *Science* 307 (5717): 1965–1969.

Peitz, Ingmar, and Rien van Leeuwen. 2010. “Single-cell bacteria growth monitoring by automated DEP-facilitated image analysis.” *Lab Chip* 10 (21): 2944–2951.

Pfaffl, M W. 2001. “A new mathematical model for relative quantification in real-time RT-PCR.” *Nucleic Acids Research* 29 (9): e45.

Pohl, H. A., Karan Kaler, and Kent Pollock. 1981. “The continuous positive and negative dielectrophoresis of microorganisms.” *Journal of Biological Physics* 9 (2): 67–86.

Prakash, A. Ranjit, Matthias Amrein, and Karan V. I. S. Kaler. 2007. “Characteristics and impact of Taq enzyme adsorption on surfaces in microfluidic devices.” *Microfluidics and Nanofluidics* 4 (4): 295–305.

Quéré, Corinne Le, Christian Rödenbeck, Erik T. Buitenhuis, Thomas J. Conway, Ray Langenfelds, Antony Gomez, Casper Labuschagne, et al. 2007. “Saturation of the Southern Ocean CO_2 sink due to recent climate change.” *Science* 316 (5832): 1735–1738.

Rackham, Oliver, and Chris M Brown. 2004. “Visualization of RNA-protein interactions in living cells: FMRP and IMP1 interact on mRNAs.” *The EMBO Journal* 23 (16): 3346–3355.

Rajilić-Stojanović, Mirjana, Hauke Smidt, and Willem M De Vos. 2007. “Diversity of the human gastrointestinal tract microbiota revisited.” *Environmental Microbiology* 9 (9): 2125–2136.

Ramsköld, Daniel, Shujun Luo, Yu-Chieh Wang, Robin Li, Qiaolin Deng, Omid R. Faridani, Gregory A. Daniels, et al. 2012. “Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells.” *Nature Biotechnology* 30 (8): 777–782.

Rao, I M, A R Arulanantham, and N Terry. 1989. “Leaf phosphate status, photosynthesis and carbon partitioning in sugar beet: II. Diurnal changes in sugar phosphates, adenylates, and nicotinamide nucleotides.” *Plant Physiology* 90 (3): 820–826.

Rappé, Michael S, Stephanie A Connon, Kevin L Vergin, and Stephen J Giovannoni. 2002. “Cultivation of the ubiquitous SAR11 marine bacterioplankton clade.” *Nature* 418 (6898): 630–633.

Rappé, Michael S, and Stephen J Giovannoni. 2003. “The uncultured microbial majority.” *Annual Review of Microbiology* 57: 369–394.

Rasmussen, M. B., L. B. Oddershede, and H. Siegumfeldt. 2008. “Optical tweezers cause physiological damage to *Escherichia Coli* and *Listeria* bacteria.” *Applied and Environmental Microbiology* 74 (8): 2441–2446.

Ren, Xiang-Dong, and Martin Alexander Schwartz. 2000. “Determination of GTP loading on rho.” *Methods in Enzymology*, 325:264–272.

Richaud, Catherine, Gerald Zabulon, Annette Joder, and Jean-Claude Thomas. 2001. “Nitrogen or sulfur starvation differentially affects phycobilisome degradation and expression of the *nblA* gene in *Synechocystis* strain PCC 6803.” *J. Bacteriol.* 183 (10): 2989–2994.

Ripp, Steven, David E. Nivens, Yeonghee Ahn, Claudia Werner, John Jarrell, James P. Easter, Chris D. Cox, Robert S. Burlage, and Gary S. Sayler. 2000. “Controlled field release of a bioluminescent genetically engineered microorganism for bioremediation process monitoring and control.” *Environmental Science & Technology* 34 (5): 846–853.

Roeder, Barbara, Martin Wagner, and Peter Rossmanith. 2010. “Autonomous growth of isolated single *Listeria Monocytogenes* and *Salmonella Enterica Serovar Typhimurium* cells in the absence of growth factors and intercellular contact.” *Applied and Environmental Microbiology* 76 (8): 2600–2606.

Rosenfeld, Nitzan, Jonathan W. Young, Uri Alon, Peter S. Swain, and Michael B. Elowitz. 2005. “Gene regulation at the single-cell level.” *Science* 307 (5717): 1962–1965.

Sabine, Christopher L., Richard A. Feely, Nicolas Gruber, Robert M. Key, Kitack Lee, John L. Bullister, Rik Wanninkhof, et al. 2004. “The oceanic sink for anthropogenic CO_2.” *Science* 305 (5682): 367–371.

Sando, Shinsuke, and Eric T Kool. 2002. “Imaging of RNA in bacteria with self-ligating quenched probes.” *Journal of the American Chemical Society* 124 (33): 9686–9687.

Sando, Shinsuke, Atsushi Narita, and Yasuhiro Aoyama. 2007. “Light-up Hoechst-DNA aptamer pair: Generation of an aptamer-selective fluorophore from a conventional DNA-staining dye.” *Chembiochem: A European Journal of Chemical Biology* 8 (15): 1795–1803.

Santangelo, Philip J, Brent Nix, Andrew Tsourkas, and Gang Bao. 2004. “Dual FRET molecular beacons for mRNA detection in living cells.” *Nucleic Acids Research* 32 (6): e57.

Schaerli, Yolanda, and Florian Hollfelder. 2009. “The potential of microfluidic water-in-oil droplets in experimental biology.” *Molecular bioSystems* 5 (12): 1392–1404.

Schmelz, M, R Schmidt, C Weidner, Marita Hilliges, H E Torebjork, and H O Handwerker. 2003. “Chemical response pattern of different classes of C-nociceptors to pruritogens and algogens.” *Journal of Neurophysiology* 89 (5): 2441–2448.

Schmid, Andreas, Hendrik Kortmann, Petra S Dittrich, and Lars M Blank. 2010. “Chemical and biological single cell analysis.” *Current Opinion in Biotechnology* 21 (1): 12–20.

Schneider, Jörg, Andreas Buness, Wolfgang Huber, Joachim Volz, Petra Kioschis, Mathias Hafner, Annemarie Poustka, and Holger Sültmann. 2004. “Systematic analysis of T7 RNA polymerase based in vitro linear RNA amplification for use in microarray experiments.” *BMC Genomics* 5 (1): 29.

Schut, F, E J de Vries, J C Gottschal, B R Robertson, W Harder, R A Prins, and D K Button. 1993. “Isolation of typical marine bacteria by dilution culture: Growth, maintenance, and characteristics of isolates under laboratory conditions.” *Applied and Environmental Microbiology* 59 (7): 2150–2160.

Sezonov, Guennadi, Danièle Joseleau-Petit, and Richard D’Ari. 2007. “*Escherichia Coli* physiology in Luria-Bertani broth.” *Journal of Bacteriology* 189 (23): 8746–8749.

Shalek, Alex K., Rahul Satija, Xian Adiconis, Rona S. Gertner, Jellert T. Gaublomme, Raktima Raychowdhury, Schragi Schwartz, et al. 2013. “Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells.” *Nature* advance online publication.

Sherwood, Alison R., and Gernot G. Presting. 2007. “Universal primers amplify a 23S rDNA plastid marker in eukaryotic algae and cyanobacteria.” *Journal of Phycology* 43 (3): 605–608.

Shi, Xu, Weimin Gao, Shih-hui Chao, Weiwen Zhang, and Deirdre R. Meldrum. 2013. “Monitoring the single-cell stress response of the diatom *Thalassiosira Pseudonana* by quantitative real-time reverse transcription-PCR.” *Applied and Environmental Microbiology* 79 (6): 1850–1858.

Shi, Xu, Liang-I Lin, Szu-Yu Chen, Shih-Hui Chao, Weiwen Zhang, and Deirdre R Meldrum. 2011. “Real-time PCR of single bacterial cells on an array of adhering droplets.” *Lab Chip* 11 (13): 2276–2281.

Shim, Jung-uk, Luis F Olguin, Graeme Whyte, Duncan Scott, Ann Babtie, Chris Abell, Wilhelm T S Huck, and Florian Hollfelder. 2009. “Simultaneous determination of gene expression and enzymatic activity in individual bacterial cells in microdroplet compartments.” *Journal of the American Chemical Society* 131 (42): 15251–15256.

Siegel, Sidney. 1957. “Nonparametric statistics.” *The American Statistician* 11 (3): 13–19.

Siegele, D A, and J C Hu. 1997. “Gene expression from plasmids containing the *araBAD* promoter at subsaturating inducer concentrations represents mixed populations.” *Proceedings of the National Academy of Sciences of the United States of America* 94 (15): 8168–8172.

Sims, Christopher E, and Nancy L Allbritton. 2007. “Analysis of single mammalian cells on-chip.” *Lab Chip* 7 (4): 423–440.

Singh, Ruchira, Rajanikanth J Maganti, Sairam V Jabba, Martin Wang, Glenn Deng, Joe Don Heath, Nurith Kurn, and Philine Wangemann. 2005. “Microarray-based comparison of three amplification methods for nanogram amounts of total RNA.” *American Journal of Physiology. Cell Physiology* 288 (5): C1179–1189.

Sipos, R, AJ Szekely, M Palatinszky, S Revesz, K Marialigeti, and M Nikolausz. 2007. “Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targetting bacterial community analysis.” *FEMS Microbiology Ecology* 60 (2): 341–350.

Sixou, S, F C Szoka Jr, G A Green, B Giusti, G Zon, and D J Chin. 1994. “Intracellular oligonucleotide hybridization detected by fluorescence resonance Eenergy transfer (FRET).” *Nucleic Acids Research* 22 (4): 662–668.

Stahl, David A., David J. Lane, Gary J. Olsen, and Norman R. Pace. 1984. “Analysis of hydrothermal vent-associated symbionts by ribosomal RNA sequences.” *Science* 224 (4647): 409–411.

Ståhlberg, Anders, Joakim Håkansson, Xiaojie Xian, Henrik Semb, and Mikael Kubista. 2004. “Properties of the reverse transcription reaction in mRNA quantification.” *Clinical Chemistry* 50 (3): 509–515.

Ståhlberg, Anders, Mikael Kubista, and Pierre Åman. 2011. “Single-cell gene-expression profiling and its potential diagnostic applications.” *Expert Review of Molecular Diagnostics* 11 (7): 735–740.

Ståhlberg, Anders, Vendula Rusnakova, and Mikael Kubista. 2013. “The added value of single-cell gene expression profiling.” *Briefings in Functional Genomics* 12 (2): 81–89.

Stewart, Philip S, and Michael J Franklin. 2008. “Physiological heterogeneity in biofilms.” *Nature Reviews. Microbiology* 6 (3): 199–210.

Strand, Stuart E., and Dale A. Carlson. 1984. “Rapid BOD measurement for municipal wastewater samples using a biofilm electrode.” *Journal (Water Pollution Control Federation)* 56 (5): 464–467.

Strovas, Tim J, and Mary E Lidstrom. 2009. “Population heterogeneity in *Methylobacterium Extorquens AM1*.” *Microbiology* 155 (6): 2040–2048.

Strovas, Tim J, Linda M Sauter, Xiaofeng Guo, and Mary E Lidstrom. 2007. “Cell-to-cell heterogeneity in growth rate and gene expression in M*ethylobacterium Extorquens AM1*.” *Journal of Bacteriology* 189 (19): 7127–7133.

Subkhankulova, Tatiana, and Frederick J Livesey. 2006. “Comparative evaluation of linear and exponential amplification techniques for expression profiling at the single-cell level.” *Genome Biology* 7 (3): R18.

Tang, Fuchou, Catalin Barbacioru, Yangzhou Wang, Ellen Nordman, Clarence Lee, Nanlan Xu, Xiaohui Wang, et al. 2009. “mRNA-Seq whole-transcriptome analysis of a single cell.” *Nature Methods* 6 (5): 377–382.

Taniguchi, Kiyomi, Tomoharu Kajiyama, and Hideki Kambara. 2009. “Quantitative analysis of gene expression in a single cell by qPCR.” *Nature Methods* 6 (7): 503–506.

Taniguchi, Yuichi, Paul J Choi, Gene-Wei Li, Huiyi Chen, Mohan Babu, Jeremy Hearn, Andrew Emili, and X Sunney Xie. 2010. “Quantifying *E. Coli* proteome and transcriptome with single-molecule sensitivity in single cells.” *Science* 329 (5991): 533–538.

Tanyeri, Melikhan, Eric M Johnson-Chavarria, and Charles M Schroeder. 2010. “Hydrodynamic trap for single particles and cells.” *Applied Physics Letters* 96 (22): 224101.

Teramoto, Jun, Yoko Yamanishi, El-Shimy H Magdy, Akiko Hasegawa, Ayako Kori, Masahiro Nakajima, Fumihito Arai, Toshio Fukuda, and Akira Ishihama. 2010. “Single live-bacterial cell assay of promoter activity and regulation.” *Genes to Cells: Devoted to Molecular & Cellular Mechanisms* 15 (11): 1111–1122.

Thamatrakoln, Kimberlee, Olga Korenovska, A Kalani Niheu, and Kay D Bidle. 2012. “Whole-genome expression analysis reveals a role for death-related genes in stress acclimation of the diatom *Thalassiosira Pseudonana*.” *Environmental Microbiology* 14 (1): 67–81.

Thompson, J D, D G Higgins, and T J Gibson. 1994. “CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.” *Nucleic Acids Research* 22 (22): 4673–4680.

Tolker-Nielsen, T, K Holmstrøm, and S Molin. 1997. “Visualization of specific gene expression in individual *Salmonella Typhimurium* cells by *in situ* PCR.” *Applied and Environmental Microbiology* 63 (11): 4196–4203.

Tsang, Peter H, Guanglai Li, Yves V Brun, L Ben Freund, and Jay X Tang. 2006. “Adhesion of single bacterial cells in the micronewton range.” *Proceedings of the National Academy of Sciences of the United States of America* 103 (15): 5764–5768.

Tucker, Don L., Nancy Tucker, and Tyrrell Conway. 2002. “Gene expression profiling of the pH response in *Escherichia Coli*.” *Journal of Bacteriology* 184 (23): 6551–6558.

Tyagi, S, and F R Kramer. 1996. “Molecular beacons: Probes that fluoresce upon hybridization.” *Nature Biotechnology* 14 (3): 303–308.

Tyagi, Sanjay. 2009. “Imaging intracellular RNA distribution and dynamics in living cells.” *Nature Methods* 6 (5): 331–338.

Tyson, Gene W., Jarrod Chapman, Philip Hugenholtz, Eric E. Allen, Rachna J. Ram, Paul M. Richardson, Victor V. Solovyev, Edward M. Rubin, Daniel S. Rokhsar, and Jillian F. Banfield. 2004. “Community structure and metabolism through reconstruction of microbial genomes from the environment.” *Nature* 428 (6978): 37–43.

Untergasser, Andreas, Harm Nijveen, Xiangyu Rao, Ton Bisseling, René Geurts, and Jack A M Leunissen. 2007. “Primer3Plus, an enhanced web interface to Primer3.” *Nucleic Acids Research* 35 (Web Server issue): W71–74.

Valencia-Burton, Maria, Ron M McCullough, Charles R Cantor, and Natalia E Broude. 2007. “RNA visualization in live bacterial cells using fluorescent protein complementation.” *Nature Methods* 4 (5): 421–427.

Valente, Valeria, Silvia A Teixeira, Luciano Neder, Oswaldo K Okamoto, Sueli M Oba-Shinjo, Suely K N Marie, Carlos A Scrideli, Maria L Paçó-Larson, and Carlos G Carlotti Jr. 2009. “Selection of suitable housekeeping genes for expression analysis in glioblastoma using quantitative RT-PCR.” *BMC Molecular Biology* 10: 17.

VanGuilder, Heather D, Kent E Vrana, and Willard M Freeman. 2008. “Twenty-five years of quantitative PCR for gene expression analysis.” *BioTechniques* 44 (5): 619–626.

Vilibić, Ivica, and Danijela Šantić. 2008. “Deep water ventilation traced by *Synechococcus Cyanobacteria*.” *Ocean Dynamics* 58 (2): 119–125.

Wacker, Michael J, Michelle M Tehel, and Philip M Gallagher. 2008. “Technique for quantitative RT-PCR analysis directly from single muscle fibers.” *Journal of Applied Physiology* 105 (1): 308–315.

Walling, Maureen A, and Jason R E Shepard. 2011. “Cellular heterogeneity and live cell arrays.” *Chemical Society Reviews* 40 (7): 4049–4076.

Wang, Daojing, and Steven Bodovitz. 2010. “Single cell analysis: The new frontier in ‘Omics’.” *Trends in Biotechnology* 28 (6): 281–290.

Wang, E, L D Miller, G A Ohnmacht, E T Liu, and F M Marincola. 2000. “High-fidelity mRNA amplification for gene profiling.” *Nature Biotechnology* 18 (4): 457–459.

Wang, Joseph. 2002. “Real-Time electrochemical monitoring: Toward green analytical chemistry.” *Accounts of Chemical Research* 35 (9): 811–816.

Wang, Joseph, and Qiang Chen. 1995. “Remote electrochemical biosensor for field monitoring of phenolic compounds.” *Analytica Chimica Acta* 312 (1): 39–44.

Wang, Joseph, Nancy Foster, Saulius Armalis, David Larson, Alberto Zirino, and Khris Olsen. 1995. “Remote stripping electrode for *in situ* monitoring of labile copper in the marine environment.” *Analytica Chimica Acta* 310 (2): 223–231.

Wang, Qianrui, Daekeun Kim, Dionysios D. Dionysiou, George A. Sorial, and Dennis Timberlake. 2004. “Sources and remediation for mercury contamination in aquatic systems—a literature review.” *Environmental Pollution* 131 (2): 323–336.

Wang, Xiaowei, and Brian Seed. 2003. “A PCR primer bank for quantitative gene expression analysis.” *Nucleic Acids Research* 31 (24): e154.

Waters, L C, S C Jacobson, N Kroutchinina, J Khandurina, R S Foote, and J M Ramsey. 1998. “Microchip device for cell lysis, multiplex PCR amplification, and electrophoretic sizing.” *Analytical Chemistry* 70 (1): 158–162.

Werf, M. J. van der, S. Hartmans, and W. J. J. van den Tweel. 1995. “Permeabilization and lysis of *Pseudomonas Pseudoalcaligenes* cells by Triton X-100 for efficient production of dmalate.” *Applied Microbiology and Biotechnology* 43 (4): 590–594.

Wheeler, Aaron R., William R. Throndset, Rebecca J. Whelan, Andrew M. Leach, Richard N. Zare, Yish Hann Liao, Kevin Farrell, Ian D. Manger, and Antoine Daridon. 2003. “Microfluidic device for single-cell analysis.” *Analytical Chemistry* 75 (14): 3581–3586.

White, A. K., K. A. Heyries, C. Doolin, M. VanInsberghe, and C. L. Hansen. 2013. “High-throughput microfluidic single-cell digital polymerase chain reaction.” *Analytical Chemistry* 85 (15): 7182–7190.

White, Adam K, Michael VanInsberghe, Oleh I Petriv, Mani Hamidi, Darek Sikorski, Marco A Marra, James Piret, Samuel Aparicio, and Carl L Hansen. 2011. “High-throughput microfluidic single-cell RT-qPCR.” *Proceedings of the National Academy of Sciences of the United States of America* 108 (34): 13999–14004.

Wilson, Claire L, Stuart D Pepper, Yvonne Hey, and Crispin J Miller. 2004. “Amplification protocols introduce systematic but reproducible errors into gene expression studies.” *BioTechniques* 36 (3): 498–506.

Wood, T H. 1967. “Genetic recombination in *Escherichia Coli*: Clone heterogeneity and the kinetics of segregation.” *Science* 157 (786): 319–321.

Yan, Jiusheng, and William A. Cramer. 2003. “Functional insensitivity of the cytochrome B6 f complex to structure changes in the hinge region of the rieske iron-sulfur protein.” *Journal of Biological Chemistry* 278 (23): 20925–20933.

Yang, Seong Wook, and Tom Vosch. 2011. “Rapid detection of microRNA by a silver nanocluster DNA probe.” *Analytical Chemistry* 83 (18): 6935–6939.

Yi, W C, S Hsiao, J H Liu, P C Soo, Y T Horng, W C Tsai, H C Lai, et al. 1998. “Use of fluorescein labelled antibody and fluorescence activated cell sorter for rapid identification of *Mycobacterium* species.” *Biochemical and Biophysical Research Communications* 250 (2): 403–408.

Zare, Richard N., and Samuel Kim. 2010. “Microfluidic platforms for single-cell analysis.” *Annual Review of Biomedical Engineering* 12 (1): 187–201.

Zeng, Jia, Jiangxin Wang, Weimin Gao, Aida Mohammadreza, Laimonas Kelbauskas, Weiwen Zhang, Roger H Johnson, and Deirdre R Meldrum. 2011. “Quantitative single-cell gene expression measurements of multiple genes in response to hypoxia treatment.” *Analytical and Bioanalytical Chemistry* 401 (1): 3–13.

Zeng, Yong, Richard Novak, Joe Shuga, Martyn T Smith, and Richard A Mathies. 2010. “High-performance single cell genetic analysis using microfluidic emulsion generator arrays.” *Analytical Chemistry* 82 (8): 3183–3190.

Zhang, Hu, and Kuo-Kang Liu. 2008. “Optical tweezers for single cells.” *Journal of the Royal Society, Interface / the Royal Society* 5 (24): 671–690.

Zhang, Yunxia, Ying Zhu, Bo Yao, and Qun Fang. 2011. “Nanolitre droplet array for real time reverse transcription polymerase chain reaction.” *Lab Chip* 11 (8): 1545–1549.

Zhou, S, Z Hou, N Li, and Q Qin. 2007. “Development of a SYBR Green I real-time PCR for quantitative detection of *Vibrio Alginolyticus* in seawater and seafood.” *Journal of Applied Microbiology* 103 (5): 1897–1906.

Zubkov, Mikhail V., and Peter H. Burkill. 2006. “Syringe pumped high speed flow cytometry of oceanic phytoplankton.” *Cytometry Part A* 69A (9): 1010–1019.

APPENDIX A

SAMPLE SIZE ISSUE

Multiple approaches had been explored to increase the sample size of each nutrition condition. Because the chip created in this research has only 48 reactions per chip, increasing sample sizes requires the use of multiple chips. The relatively small reaction number cannot provide higher than 90% power of single cell level analyses. Running multiple chips for each condition can alleviate the low throughput issue. However, running multiple chips for each condition was not feasible under current experimental conditions. The reasons are summarized below.

The first reason was that there was not a method for preserving the state of the RNA of each cell while waiting for the chip to be available for another run. Due to the difficulty of single cell loading, the whole processing time for each chip is between 2 to 3 hours. The experiment run time is about 4 hours. The whole process adds up to 6 to 7 hours per chip during which time the RNA of cells waiting to be analyzed has changed. The RNA preservation effectiveness was evaluated for RNAprotect (Table 11).

Table 11. RNAprotect RNA preservation effectiveness

	18S	*Actin*
Control	14.06	30.42719
RNAprotect	17.38	31.46477

The results showed that the RNA profile changed after a couple hours' storage in RNAprotect solution. RNALater had similar performance and RNALater is a high salt solution without purification which may cause inhibition of downstream qPCR.

The second reason is the chip-to-chip variance. Although much effort has been spent on optimization of the performance, chip-to-chip variance was still observed

(Figure 30). A lot of factors can affect the results of RT-qPCR reactions, for example, temperature and reverse transcription efficiency. Most of these factors cannot be easily controlled. As shown in Figure 30, a variance exists for the same concentration run on a different chip. The largest difference was about 2.2 cycles for the mean value.

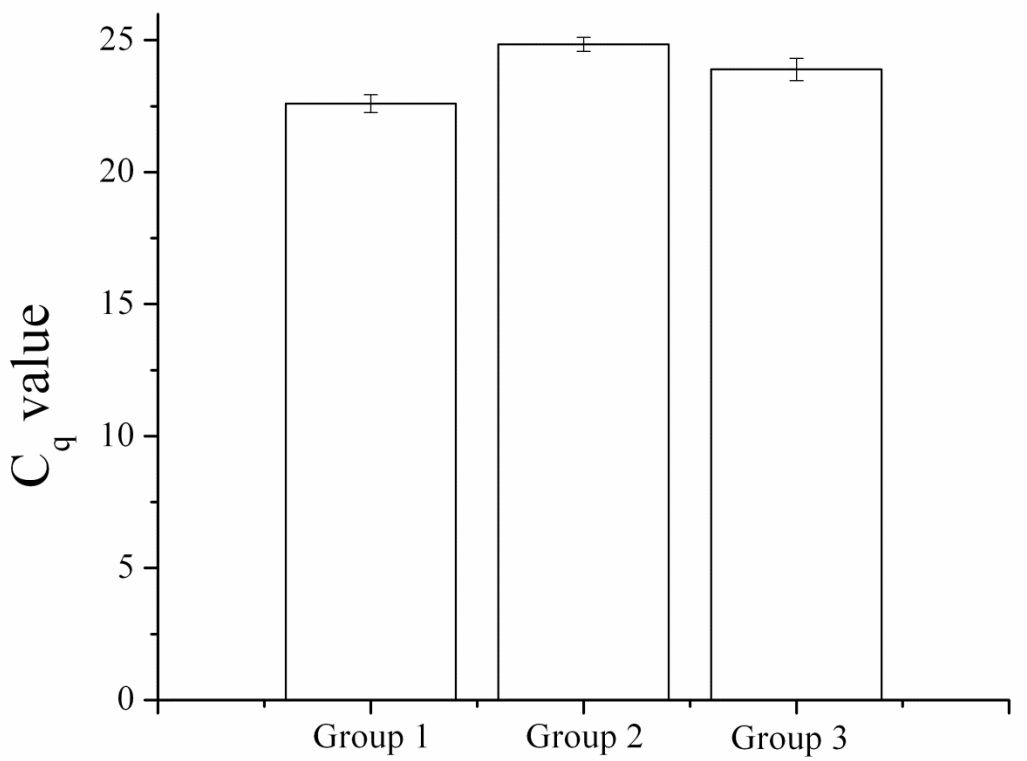

Figure 30. Chip-to-chip variance at pg level

If a known droplet volume is used to normalize the results, the chip-to-chip variance can be minimized (Figure 31). The variance is smaller but cannot be totally removed. In this case, the minimum variance is about one cycle. A one cycle difference is equivalent to about two times difference of the activity which is not good enough for single cell analysis. Combining different chips in an experiment may produce biased results.

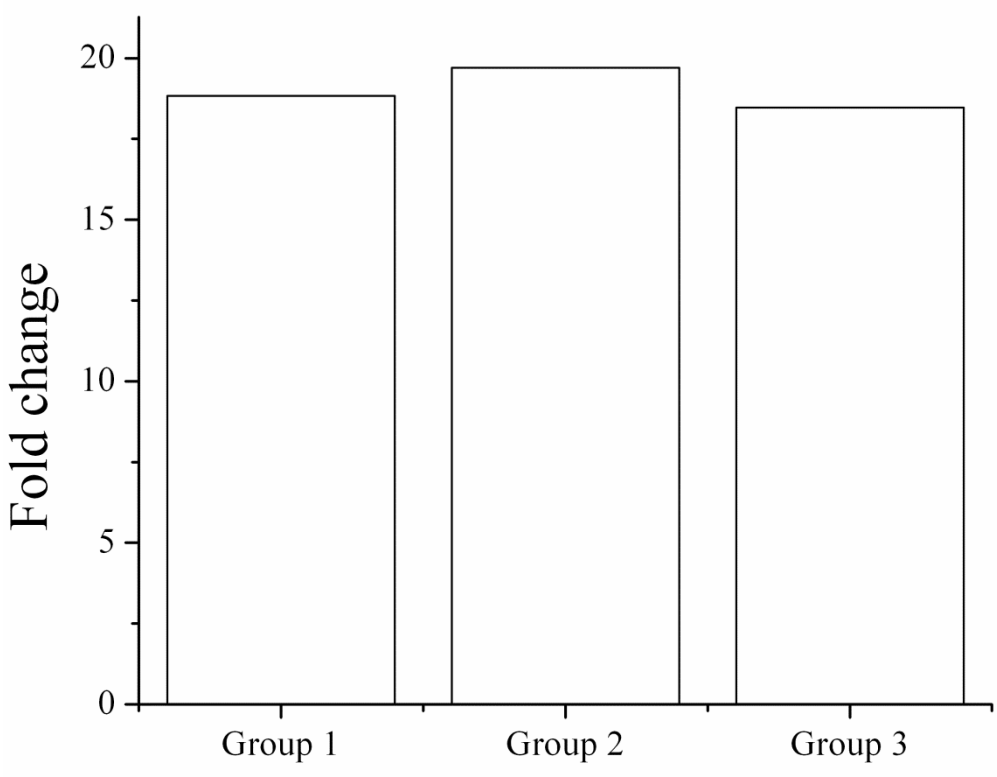

Figure 31. Normalized results at pg level.

In general, there was no reliable method to overcome the RNA preservation issue and chip-to-chip variance. Based on these limitations, only reactions run on the same chip can be used to achieve consistent results at the single cell level. The most reliable solution to increase the power of the chip is to increase the throughput of the chip and to build customized parts for the thermal cycling and signal detection.