METABOLIC NETWORK CONSTRUCTION BASED ON THE GENOME OF THE MARINE DIATOM *THALASSIOSIRA PSEUDONANA* AND THE ANALYSIS OF GENOME-WIDE TRANSCRIPTOME DATA TO INVESTIGATE TRIACYLGLYCERIDE ACCUMULATION

A Thesis

Presented to

The Faculty of Moss Landing Marine Labs

San José State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

By

Karen R. Parker

December 2013

The Designated Thesis Committee Approves the Thesis Titled

METABOLIC NETWORK CONSTRUCTION BASED ON THE GENOME OF THE MARINE DIATOM *THALASSIOSIRA PSEUDONANA* AND THE ANALYSIS OF GENOME-WIDE TRANSCRIPTOME DATA TO INVESTIGATE TRIACYLGLYCERIDE ACCUMULATION

By

Karen R. Parker

APPPROVED FOR MOSS LANDING MARINE LABS

SAN JOSÉ STATE UNIVERSITY

December 2013

Dr. Nick Welschmeyer	Moss Landing Marine Labs
Dr. Jason Smith	Moss Landing Marine Labs
Dr. Jon Geller	Moss Landing Marine Labs
Dr. Ginger Armbrust	University of Washington

ABSTRACT

METABOLIC NETWORK CONSTRUCTION BASED ON THE GENOME OF THE MARINE DIATOM *THALASSIOSIRA PSEUDONANA* AND THE ANALYSIS OF GENOME-WIDE TRANSCRIPTOME DATA TO INVESTIGATE TRIACYLGLYCERIDE ACCUMULATION

By Karen R. Parker

Algal lipids called triacylglycerols (TAGs) are essential compounds in marine ecology and are also used as the basis for commercial production of algal biofuels and omega-3 nutraceuticals. The goal of this thesis was to elucidate the biochemical processes associated with algal lipid accumulation in diatoms under different environmental conditions. The approach made use of the Mock et al. whole genome tiling microarray gene expression database for *Thalassiosira pseudonana* and gene annotations derived from the whole genome sequencing work of Armbrust et al. The analytical approach used bioinformatic tools, including Stanford Research Institute's (SRI) BioCyc Pathway tool and the Joint Genome Institute's (JGI) Integrated Microbial Genomes/Expert Review (IMG/ER) tool. This analysis has resulted in building the first whole genome BioCyc pathway model for *T. pseudonana* that includes over three hundred metabolic pathways. The analysis of the Mock transcriptome data in combination with the pathway model illustrated not only the activity of fatty acid and lipid pathways but also the interplay of other pathways that affect the accumulation of TAGs under different environmental conditions. This in silico analytical approach also revealed evidence for un-annotated gene functions and lack of the key regulatory protein PII pointing to the distinctive metabolic features of marine diatoms.