Development of a Promising Methanotrophic Bacterium as an Industrial Biocatalyst

Alexey Gilman

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2017

Reading Committee:

Mary Lidstrom, Chair

David Beck, Co-Chair

François Baneyx

Jesse Zalatan

Program Authorized to Offer Degree:

Chemical Engineering

University of Washington

University of Washington

Abstract

Development of a Promising Methanotrophic Bacterium as an Industrial Biocatalyst

Alexey Gilman

Chair of the Supervisory Committee:
Professor Mary Lidstrom
Chemical Engineering

Methane, the major component of natural gas, is a cheap and abundant source of energy. A significant amount of energy is wasted in the form of natural gas flaring. This mostly occurs in remote locations where transportation or conversion of methane is not economical. Methanotrophic bacteria are ubiquitous in nature, and have the unique capability of utilizing methane as the sole source of carbon and energy. These organisms have high potential in biotechnological application for converting methane into liquid fuels at the relatively small scale required for such remote locations. Presently however, knowledge gaps in methanotrophy stand in the way of developing an economically viable process for the biological catalysis of methane. In this work, we have focused the research efforts toward an industrially promising methanotrophic bacterium in 3 parts: (i) Baseline strain parameter characterization under varying growth conditions. (ii) Investigation of a hypothesized metabolism for excretion of high titers of organic acids. (iii) Exploratory transcriptomics data-analysis study involving copper (Cu), a key metal in methane oxidation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to acknowledge my lab-mate and friend Aaron Puri with whom I shared a research bay for the entirety of my 5 years in the Lidstrom lab. I am in high admiration for Aaron's commitment and dedication to science. As a very knowledgeable biologist, Aaron was influential in my acquiring the passing for biology.

Next, I would like to acknowledge my co-advisor David Beck. Dave has been a significant influence on my professional interests and the direction where I would like to go post graduation. The depth and breadth of Dave's technical expertise is a reminder to me to continue to explore new subjects, while to strive to improve already familiar topics. All this, combined with Dave's gregarious demeanor and a passion for connecting with people will remain influential for years to come.

Lastly and most notably, I would like to acknowledge my advisor Mary Lidstrom for the incredible amount of patience and support throughout my Ph.D. work. It is now obvious to me that I was extremely fortunate and privileged to have joined Mary's lab. I am still at a loss as to what it means to be Mary's last Ph.D. student and whether it is different from the other 30 students mentored by Mary. I do know however, that I will do my best to emulate and employ the generosity, compassion, and endless resilience that I have witnessed in the Lidstrom lab.

DEDICATION

I dedicate the personal growth and learning that were a part of my 5-year graduate school experience to the great persons who continue to inspire and motivate me:

"Imagination is more important than knowledge" – Albert Einstein

"The only purpose in life is to serve humanity" – Leo Tolstoy

"I actually don't have a bad hairline" – Donald Trump

Chapter 1. INTRODUCTION

1.1 BACKGROUND

Natural gas is a cheap and abundant form of energy in the United States. Advancements in drilling technologies like hydraulic fracturing have made previously untapped natural gas reserves readily mineable and accessible. Unfortunately, a large amount of energy is also wasted in the form of flared natural gas. Oil drilling operations flare natural gas that is typically found along with petroleum if the supply is not adequately large. Flaring typically takes place in remote locations where it is uneconomical to build transportation infrastructure for a relatively low quantity of natural gas. Conversion to higher-order hydrocarbon molecules is also impractical with presently available technology due to the need of a very large scale of operation to make the capital expense economically feasible [1]. Wasted energy in the form of flared natural gas is a major topic in the field of chemical engineering. One possible way to utilize this energy is via biological catalysis. This approach has several advantages on the small scale. Enzymatic catalysis operates at low-cost ambient conditions, biological or enzymatic catalysis is highly efficient, and small reactor units are easily transportable to remote locations.

1.2 METHANOTROPHIC BACTERIA

Methanotrophic bacteria have the unique capability of assimilating methane, the major component of natural gas, into complex carbon molecules. Methanotrophs build all the essential components required for life such as sugars, lipids, DNA, and proteins entirely from an organic single carbon molecule: methane. Enzymatic methane oxidation has high potential for industrial

conversion of natural gas into commodity products. Typical biocatalytic processes such as fermentation of corn starch into ethanol takes place on relative small scales in comparison to petroleum processing. This feature is an advantage for processes in remote locations where methane is typically being flared [1].

Nevertheless, the knowledge gaps in methanotrophy are a major inhibitor of technological advancement and process development. To date, the function of methane oxidation has never been transferred successfully to a non-methanotrophic host, despite the existing knowledge of methane oxidation genes and the central assimilation pathways [2]. With improved understanding of methanotrophy in combination with advancements in biotechnology, the biocatalysis of natural gas into commodity products is a realizable objective.

1.3 Thesis Objective

The objective of this thesis is to advance the understanding of methanotrophy for the application of biocatalysis of natural gas into liquid fuel and commodity products.

1.4 Thesis Roadmap

This thesis involves three interrelated efforts, which are briefly summarized below.

1.4.1 Characterize an industrially promising bacterium in a lab scale bioreactor

The industrially promising bacterium *Methylomicrobium buryatense* was characterized in a lab scale bioreactor under different limitation conditions and different substrates. This work provided baseline information for our ARPA-E project, and set the stage for two partners (NREL and LanzaTech) to develop production prototypes. When I joined the laboratory, I had little background in biology or bioengineering. At that time, the lab was just beginning a large ARPA-E project that involved our group, a national lab (NREL), and two company partners (LanzaTech

and Johnson Matthey). I began my labwork by assembling an experimental system that included a lab-scale bioreactor connected to a gas chromatograph for continuous headspace gas concentration measurements. I designed and set up this system, and carried out modifications. The gas chromatograph was calibrated at a continuous flow-rate to match experimental conditions, and the GC inlet had to be controlled via a valve and mass flow meter for consistency. With this setup, I developed a protocol to maintain a steady state culture of *M buryatense*, and characterized growth parameters and gas uptake rates for a range of growth conditions.

1.4.2 Investigate the excretion of organic acids in an obligate aerobic methane oxidizing bacteria

A previously hypothesized metabolism involving excretion of organic acids in aerobic methanotrophic bacteria was investigated in a steady-state bioreactor culture. The ability of methanotrophs to excrete organic acids may serve a starting point for the production of methane-derived commodity products. . It had previously been hypothesized that this mode of metabolism might represent a type of fermentation. We attempted to induce this condition by severely limiting the culture for O_2 and monitoring the concentration of excreted organic products and H_2 concentration in the headspace. We discovered that under these conditions, excreted acids did not increase compared to a fast growth O_2–limited condition, and the H_2 production was determined to be a byproduct of N_2 fixation. My work helped determine the physiological state involved in organic acid excretion, and I learned a great deal of biochemistry and the fundamentals of methanotrophic metabolism.

1.4.3 Explore co-regulated gene expression in a time-course transcriptomics data-set during a metabolic switch experiment

A large transcriptomics dataset consisting of over 40 samples from 12 different bioreactor experiments was analyzed. As part of the bioreactor work I carried out, other lab members collected samples for RNAseq analysis. For this last portion of my thesis, I conducted a computational transcriptomics study utilizing the collected RNAseq dataset. The main focus of this analysis was to identify co-expressed genes during a copper-induced metabolic switch experiment. The identified genes were used to form a novel testable hypothesis relating to a key metal in methanotrophy: copper. In this work, I have learned a whole new ecosystem involving computational data-analysis research. In addition to learning the basic coding language of python and the statistical packages that are available in python (pandas for data manipulation, scipy for scientific computing, matplotlib for data visualization, and scikit learn for statistical learning analysis), I also gained experience with the BASH command line interface, git version control, and good code writing practices.

Chapter 2. BIOLOGICAL CONVERSION OF NATURAL GAS TO LIQUID FUEL (BIO-GTL)

2.1 BACKGROUND

In 2012 a total global flaring of natural gas was estimated at 143 (±14) billion cubic meters, or 3.5% of global production [3]. Natural gas is flared mostly in remote locations and/or where the quantity of available gas is too small to be economically transported or converted into liquid fuel with presently available technology (Fischer-Tropsch). Natural gas that fits this description has been called “stranded natural gas”, and presents an economic opportunity if a small-scale (low CAPEX) and portable technology can be developed for conversion of natural gas to liquid fuel. Methane, the major component of natural gas is currently abundant and cheaper than oil per unit of energy. The low cost is due to recent advancements in hydraulic fracturing technologies and relative abundance of global shale gas [4, 5]. The development of a small-scale cost-effective technology for the conversion of methane to liquid fuel is therefore, a promising and impactful field of research [1].

Methanotrophic bacteria evolved to utilize methane as the sole carbon and energy source. Biological assimilation of methane takes place at ambient conditions, and starts with oxidation of methane to methanol by the enzyme methane monooxygenase (MMO). MMO is the only known way to oxidize methane to methanol at ambient temperature and pressure [6]. Using a methanotrophic organism to convert methane into desirable fuels/products would significantly reduce operational process expense (OPEX) compared to conventional solid-state catalysis, and may have an advantage at the small scale [1]. It is attractive to exploit what these organisms have naturally evolved to do: assimilate methane into higher order molecules.

The overall objective of the multi-institution project that this work was carried out under was to develop a proof of concept technology for the conversion of natural gas to liquid fuel via a microbial biocatalyst. This project investigated the feasibility of a biological gas to liquid-fuel conversion (bio-GTL) with the presently available technology and an industrially promising methanotrophic bacterium, *Methylomicrobium buryatense.*

2.1.1 Process overview

The process of bio-GTL is envisioned as being composed of several steps: (1) A high cell density continuous culture is grown aerobically on methane; (2) Biomass is harvested and cells are lysed for lipid extraction. (3) The extracted membrane lipids are catalytically upgraded into alkanes (diesel equivalent). This project was a collaboration between two companies, a national lab, and a research university (UW), with each member responsible for research and development of a portion of the overall project [Figure 2.1].

LanzaTech is an Illinois-based company that specializes in growing high cell density bacterial culture on gaseous substrates. Their role was to grow *M. buryatense* to high cell density culture in a lab scale bioreactor. A team at the National Renewable Energy Lab (NREL) was responsible for lipid extraction and characterization. They have prior experience from working with algae and were able to transfer their techniques to microbial lipid extraction. The last step in the process was to catalytically convert the phospholipids and free fatty acids into lipid derived alkanes – a diesel equivalent. Catalytic upgrading of methanotroph-derived lipids was investigated by Johnson Matthey, a UK based company. The role of University of Washington in the project was to provide the methanotrophic strain, develop genetic tools for metabolic engineering, to modify the strain to increase cell lipid content, and to characterize the baseline process parameters for this strain as a starting point for further metabolic engineering.

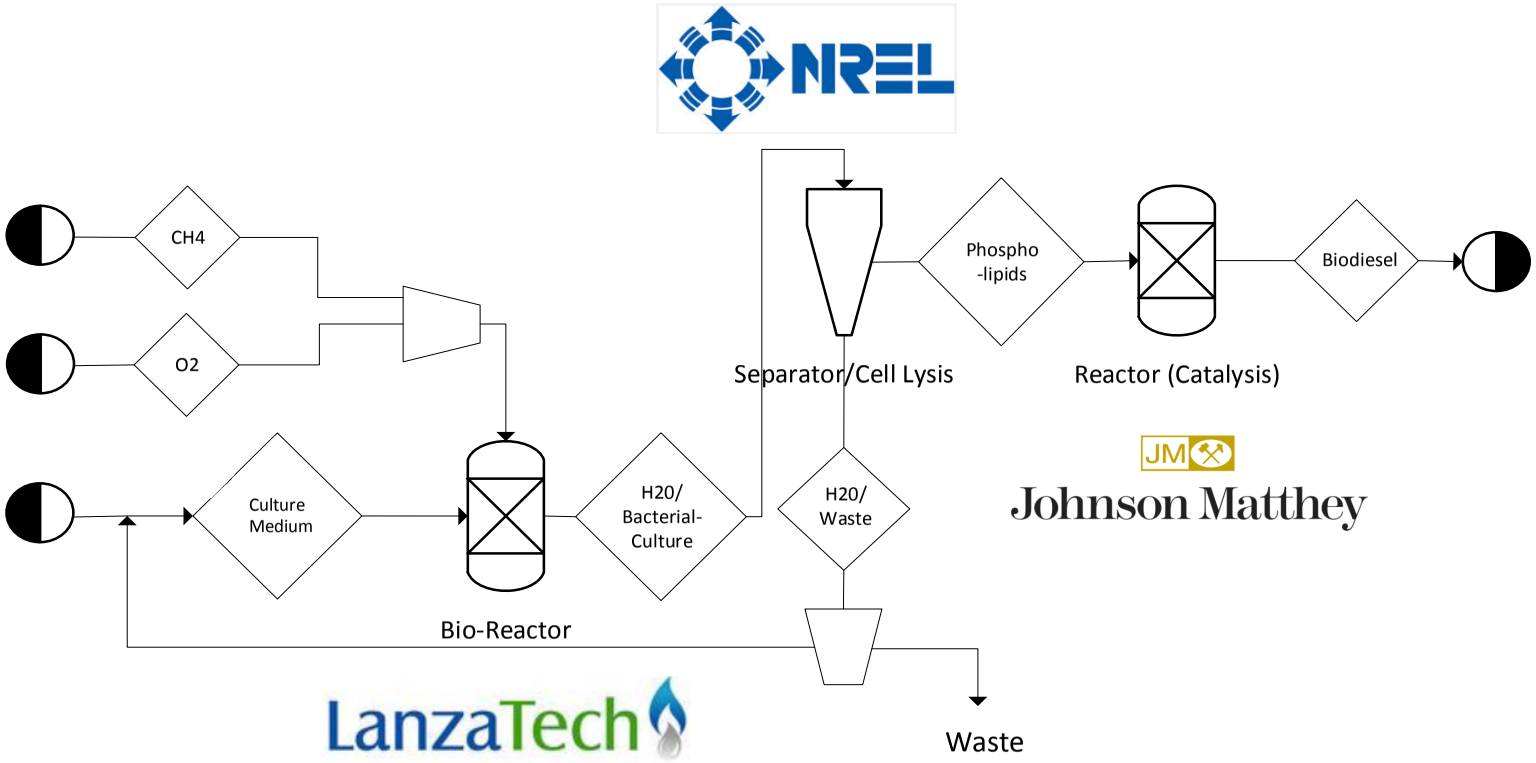

Figure 2.1. Hypothetical process for biological gas-to-liquid conversion (bio-GTL).

Collaborating partners responsible for unit development are pictured: LanzaTech – High cell density culture, bioreactor design; NREL – cell lysis, lipid extraction; Johnson Matthey – catalytic upgrading to alkanes.

2.1.2 Industrially promising strain

Methylomicrobium buryatense is a haloalkaliphile that was isolated from a soda lake in Russia [7]. It grows optimally at pH 9.5, salinity of 0.75% wt. NaCl, and a temperature of 30° C. The soda lake from which it was isolated fluctuates widely in temperature, salinity, and pH due to seasonal changes in temperature, precipitation, and evaporation rates. This methanotroph has evolved to be unusually robust, and can survive a broad range of conditions. The combination of these factors, and a fast growth rate (T_d = 2.9 hrs), make commercial-scale contamination unlikely and make *M. buryatense* a promising industrial organism. Methane oxidation begins in the cell membrane (pMMO) and periplasm (methanol dehydrogenase) where two oxidation steps catalyze the conversion of methane into methanol and then to formaldehyde [8]. From there, formaldehyde is assimilated either by the ribulose monophosphate cycle (RuMP) for type I, or via the serine cycle for type II methanotrophs [Figure 2.2][9].

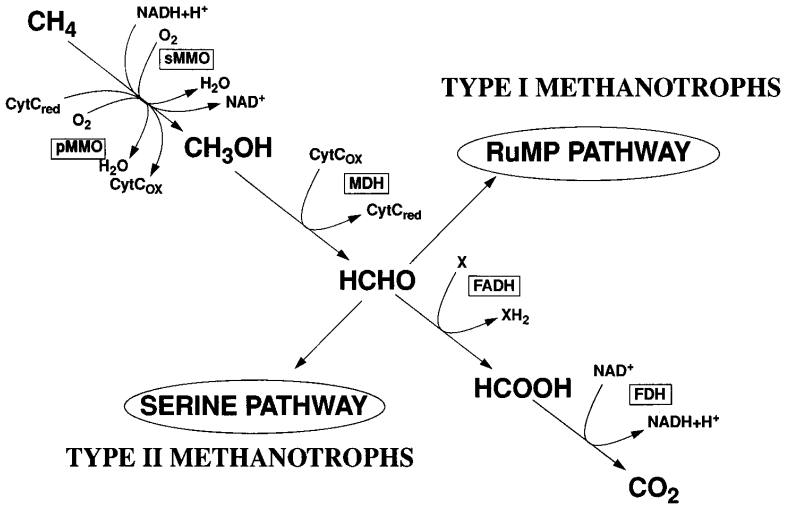

Figure 2.2. Pathways for aerobic oxidation of methane and formaldehyde assimilation.

Type I methanotrophs belong to the class of gammaproteobacteria and type II belong to alphaproteobacteria [9].

M. buryatense has two methane catalyzing enzymes: a membrane associated particulate methane monooxygenase (pMMO), and a soluble methane monooxygenase (sMMO). pMMO serves as the predominant enzyme in the presence of Cu, and is associated with the high cell lipid content due to the "stacked" intracytoplasmic membrane discs [Figure 2.3][7, 10].

2.1 Results and Discussion

My role in this project was to measure the baseline performance parameters of in a bench scale bioreactor, and assess the response of lipid content to different growth conditions. The culture was grown under four different conditions for comparison: two as unrestricted fed-batch culture (grown on methane and methanol), and another two as a continuous chemostat culture under

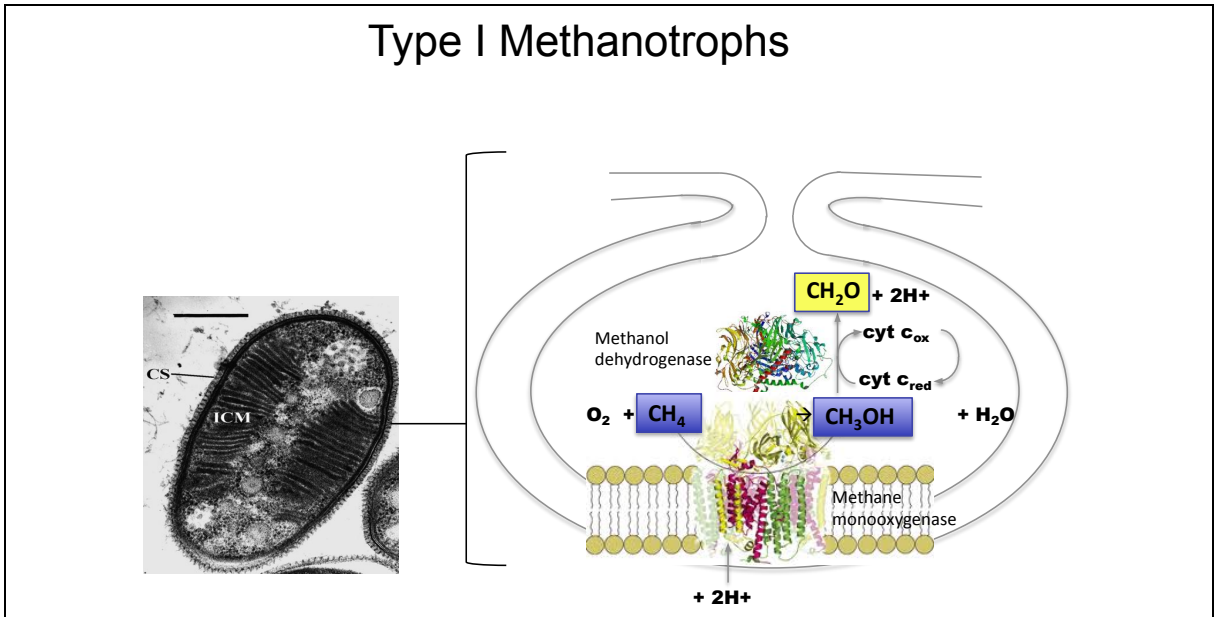

Figure 2.3. Membrane bound particulate methane monooxygenase (pMMO).

A membrane bound enzyme that oxidizes methane to methanol. Methanol is oxidized to formaldehyde in the periplasm by methanol dehydrogenase (MDH) [8]. Formaldehyde is transported into the cytoplasm where it is assimilated into biomass. Intracytoplasmic membranes contain stacks of vesicular discs, which contribute to the high cell lipid content [7].

either O_2 or methane limitation. Each condition was completed in duplicate, and the following parameters were measured: growth rate, cell dry weight (CDW), methane and O_2 uptake rates, lipid content (reported as fatty acid methyl esters, FAME, and assumed to be generated mainly from membrane phospholipids [7, 11]), yield calculation as carbon conversion efficiency (CCE), glycogen as %CDW, and excreted organic acids (formate, acetate, and lactate) [Table 2.1]. In addition to these parameters, RNA samples were collected at the end of each experiment for transcriptomics analysis (Chapter 4). The objective for transcriptomics analysis at the varying conditions is to study the metabolism and identify targets for genetic modification in order to maximize lipid productivity. For these experiments, a premixed gas mixture of methane and O_2 that is below or above the explosive range was supplied. The three gas mixtures used for these experiments are: Unlimited growth - 10%CH_4, 5%O_2, 85%N_2; O_2 limited - 20% CH_4, 5%O_2, 75%N_2, Methane limited – 2.5% CH_4, 97.5% Air. The bioreactor out-gas line is split and a portion of the out-gas was directed to a gas chromatograph for automatic sampling, and the rest is directed to a vent. The vent line is fitted with a needle valve in order to control the quantity of gas that is directed to the GC [Figure 2.4]. With this setup, the flow rate to the GC can be maintained constant (identical to the gas flow-rate that was used to calibrate the instrument) even if the total gas flow rate to the bioreactor is changed between experiments. Using the known inlet and outlet gas concentrations, and a conservation of species balance over N_2, the gas uptake rates were calculated. The cell biomass was measured at comparable optical density (OD) to determine the conversion factor between OD and cell dry weight concentration. The total biomass carbon and nitrogen contents were measured to be (w/w) 45.0 ± 1.3% and 9.9 ± 1.0 %, respectively. From the biomass carbon content, the cell dry weight conversion, the dilution rate,

and the uptake measurements, the biomass yield coefficient (reported here as carbon conversion efficiency, g-C biomass/g-C substrate) is determined [Table 2.2].

An unrestricted fed-batch culture was grown on methane, and a second set was completed in batch mode with methanol as the sole carbon source [Figure 2.5]. The methane grown culture was fed a gas mixture of 10%CH_4, and 5% O_2 (balance N_2), at a 0.1vvm. The experiment was terminated when the dissolved O_2 probe reached 1% of saturation, and unrestricted conditions confirmed by exponential growth. The culture grown on methanol was fed with a constant supply of air, and the DO probe reading remained above 85% of saturation. *M. buryatense* grew faster on methane compared to methanol; μ_{max} = 0.239 and 0.224 h^{-1} on methane compared with μ_{max} = 0.169 and 0.173 h^{-1} on methanol. This is an interesting observation, since the culture grown on methanol does not include the additional methane oxidation step in carbon assimilation.

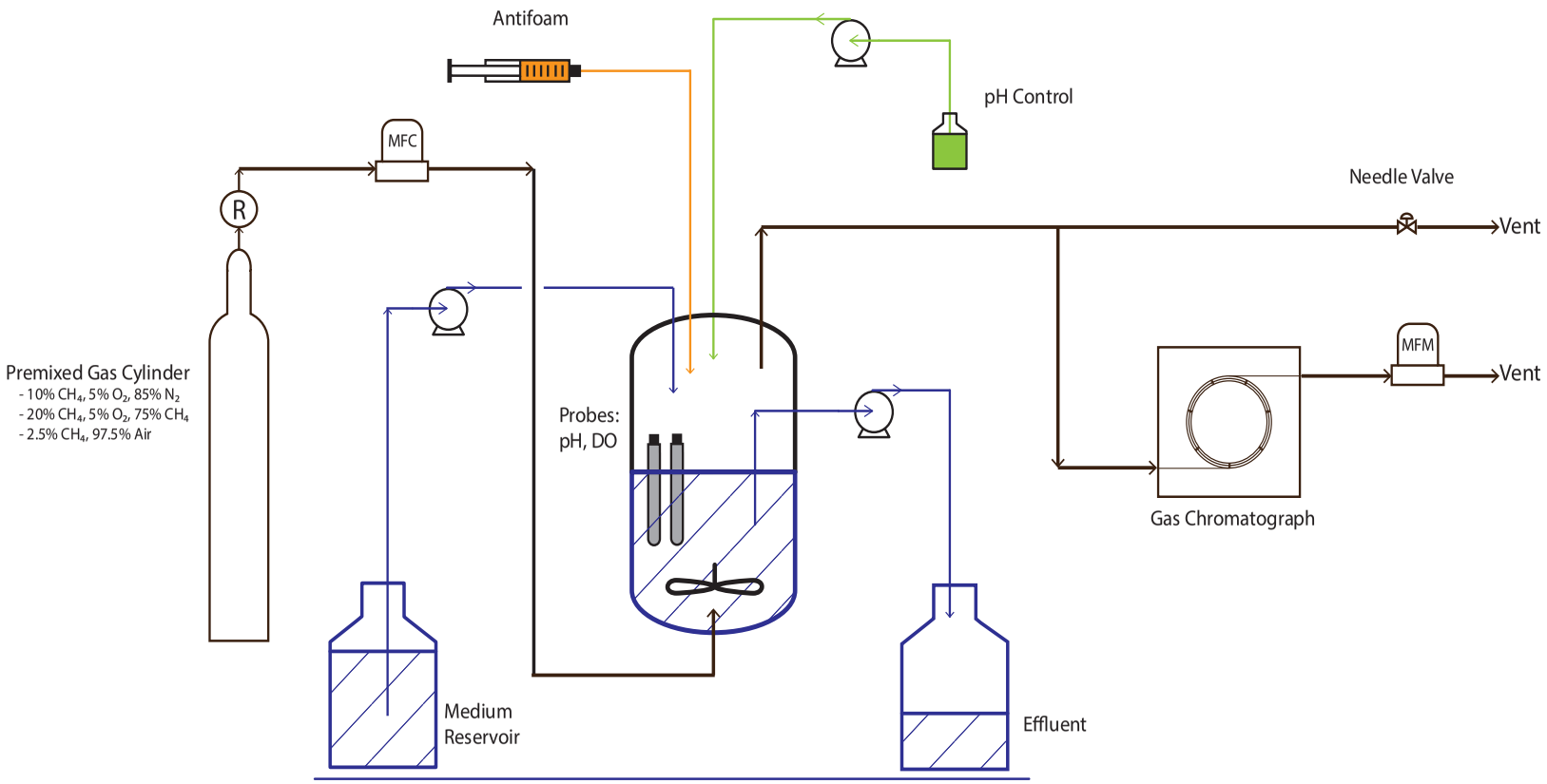

Figure 2.4. Experimental bioreactor setup.

Premixed gas cylinders are used to supply methane and O_2 outside the explosive mixture range. Out-gas line is directed to a gas chromatograph to measure gas uptake rates.

2.1.1 Unrestricted fed-batch growth

Methane-grown culture contained 8.2 and 8.5% of CDW as extracted fatty acids, which is similar to values previously reported for *Methylococcus capsulatus* (Bath) [11]. The fatty acid composition was similar to that reported previously for a related *M. buryatense* strain [7] and other reported methanotrophic bacteria [12, 13], being dominated by C16 fatty acids. When grown on methanol, *M. buryatense* expressed lower lipid concentrations of 5.1 and 6.0% of CDW. As the main methane oxidation machinery (pMMO) is housed within the cell membranes, decreased extracted fatty acid content of the cells is in agreement with decreased lipid content [14]. Under methanol growth the culture also excreted unusually large quantities of formate (as high as 13,486 μmol $gCDW^{-1}$) and accumulated large amounts of glycogen (42.8 ± 17.5% CDW).

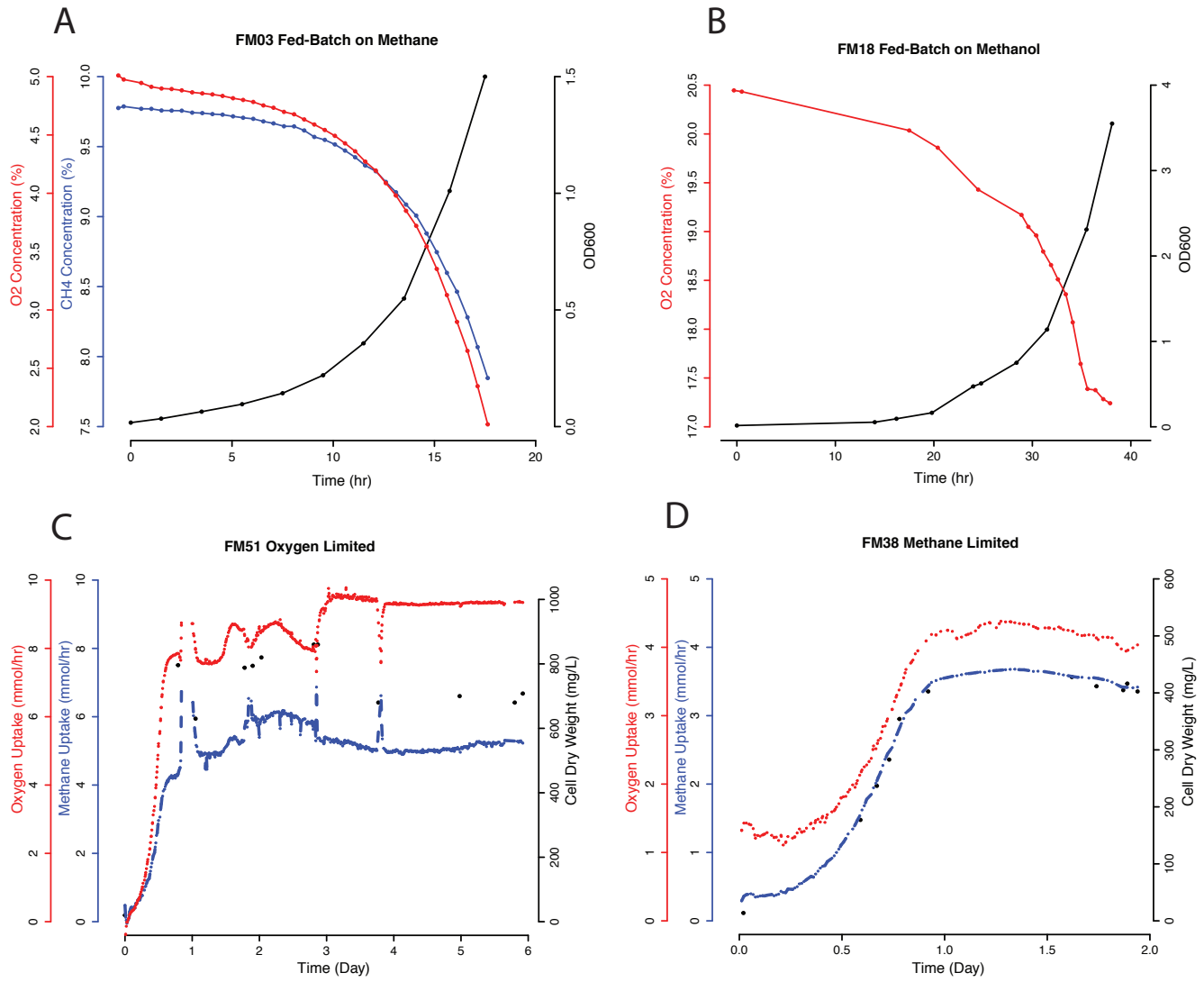

Figure 2.5. Bioreactor experiments growth profiles.

A) Unrestricted growth on methane. B) Unrestricted growth on methanol. C) O_2 limited steady-state. D) Methane limited steady-state.

Both the high formate and glycogen have been reported previously for the related strain *M. buryatense* 5B and attributed to unbalanced growth [14, 15]. The high formate excretion may also account for the slower growth rate when grown on methanol. Glycogen represents a carbon sink that may not be available to generate targeted products and may not be a desirable trait for a

commercial strain. Excreted acetate and lactate were low under both conditions, ranging between 55 and 123 μmol $gCDW^{-1}$. For methane-grown cells, the total excreted carbon in formate, acetate, and lactate accounted for 1.8 and 2.9% of the carbon in biomass, consistent with previous results from a different *Methylomicrobium* species [16].

2.1.2 Methane limitation

Stable cultures with cell densities of 0.45 and 0.46 g CDW L^{-1} [Figure 2.5] were achieved with the dilution rates of 0.122 and 0.126 hr^{-1}, respectively. The dissolved O_2 probe maintained a reading of 92% of saturation with air under these conditions. Methane limitation resulted in a total fatty acid concentration of 10.2 and 10.5% of CDW, higher than in the fed-batch cultures. The O_2/CH_4 uptake ratio of 1.6 for both replicates is in the range reported for other gammaproteobacterial methanotrophs for methane-limited cultures (1.41-2.15) [17-19]. Additionally, the calculated CCE values of 54 and 61% are in the upper range (42 – 67%) of those previously reported for methane-limited culture [17-19]. Excreted formate, acetate, and lactate were shown to be low: 1.1 to 1.5% of the biomass carbon. Formate concentration did not exceed 459 μmol $gCDW^{-1}$ and excreted acetate or lactate concentrations did not exceed 85 μmol $gCDW^{-1}$ [Table 2.2]. Likewise, a moderate amount of glycogen was produced under both methane and O_2 - limiting conditions, 6.0 – 13.2 % CDW, higher than the fed-batch culture [Table 2.2].

Table 2.1. Growth parameters for fed-batch culture grown on methane and methanol.

Experiment ID		FM03	FM21	FM18	FM23
Carbon Source		Methane	Methane	Methanol	Methanol
Total FAME (% Cell Dry Weight)		8.2	8.5	6.0	5.1
Specific Growth Rate (h^{-1})		0.239	0.224	0.169	0.173
Cell Density ($gCDW\ L^{-1}$)		0.34	0.31	0.81	0.54
Oxygen Uptake ($mmol\ h^{-1}$)		7.6	7.6	9.1	5.1
Methane Uptake ($mmol\ h^{-1}$)		5.8	6.1	N/A	N/A
Specific Oxygen Uptake ($mmol\ gCDW^{-1}\ h^{-1}$)		22.4	24.7	11.3	9.4
Specific Methane Uptake ($mmol\ gCDW^{-1}\ h^{-1}$)		17.1	19.8	N/A	N/A
O_2/CH_4 Uptake Ratio		1.3	1.2	N/A	N/A
Excreted Products	Formate ($\mu mol\ gCDW^{-1}$)	400	672	11,758	13,486
	Acetate ($\mu mol\ gCDW^{-1}$)	96	123	55	56
	Lactate ($\mu mol\ gCDW^{-1}$)	26	62	16	31
Glycogen Content (% CDW)		N/A	2.7 +/- 1.1	N/A	42.8 +/- 17.5
Inlet Gas Composition		10%CH_4, 5%O_2, 75%N_2	10%CH_4, 5%O_2, 75%N_2	0.5% CH_3OH^*, Air	0.5% CH_3OH^*, Air

Fed-batch gas uptake rates are at the time of harvest. Secreted products are measured at the time of harvest and represent accumulated concentration.

FAME: Fatty Acid Methyl Ester.

*Initial methanol concentration. No additional methanol is added during the experiment.

Table 2.2. Growth parameters for steady-state culture grown under methane or methanol limitation.

Experiment ID	FM50	FM51	FM38	FM39
Limiting Condition	Oxygen Limited	Oxygen Limited	Methane Limited	Methane Limited
Total FAME (% Cell Dry Weight)	10.5	10.7	10.2	10.5
Dilution Rate (h^{-1})	0.119	0.122	0.122	0.126
Cell Density (gCDW L^{-1})	0.79	0.77	0.45	0.46
Dissolved Oxygen (% of saturation)	1	1	92	92
Oxygen Uptake (mmol h^{-1})	8.8	9.5	6.2	5.9
Methane Uptake (mmol h^{-1})	7.7	8.2	3.8	3.6
Specific Oxygen Uptake (mmol $gCDW^{-1}$ h^{-1})	11.1	12.3	13.7	12.8
Specific Methane Uptake (mmol $gCDW^{-1}$ h^{-1})	9.7	10.6	8.4	7.8
O_2/CH_4 Uptake Ratio	1.1	1.2	1.6	1.6
Carbon Conversion Efficiency (%)	46	43	54	61
Excreted Products: Formate (µmol $gCDW^{-1}$)	286	301	255	459
Excreted Products: Acetate (µmol $gCDW^{-1}$)	36	85	32	28
Excreted Products: Lactate (µmol $gCDW^{-1}$)	24	28	47	21
Glycogen Content (% CDW)	7.2 ± 1.9	13.1 ± 4.0	6.0 ± 0.5	8.1 ± 2.5
Inlet Gas Composition	20% CH_4, 5%O_2, 75%N_2	20% CH_4, 5%O_2, 75%N_2	2.5% CH_4, 97.5% Air	2.5% CH_4, 97.5% Air

FAME: Fatty Acid Methyl Ester.

2.1.3 O_2 limitation

The dilution rate for O_2 limitation was set to match methane limitation conditions within experimental error. The dissolved O_2 probe maintained a reading of ~1% of saturation during continuous growth. O_2-limited growth resulted in a fatty acid content (10.5 and 10.7% of CDW) similar to that of the methane-limited culture. The CCEs of 46 and 43% [Table 2.2] are lower than methane limitation, but within the upper range of the values previously reported for other gammaproteobacterial methanotrophs under O_2-limitation (36-48%)[17, 18, 20]. The O_2/CH_4 ratios of 1.1 and 1.2 are lower than for the methane-limited cultures and slightly lower than for fed-batch cultures, as might be expected under O_2-limitation. Since every molecule of methane utilized requires one molecule of O_2 for the pMMO reaction, the low ratio indicates a metabolism in which very little of the ATP required for growth is generated from NADH oxidation via oxidative phosphorylation[21]. Further experiments will be required to determine how the metabolic networks change under these conditions, but it is clear that significant metabolic differences occur dependent on the culture conditions.

2.1.4 Medium optimization work

In support of medium optimization work for LanzaTech's high cell density goals, and to provide data for a patent application [22], two experiments were completed relating to medium Cu concentration. Copper is the metal center for pMMO[6]. It also regulates pMMO expression and the expression of a number of other enzymes [23]. It is a key medium component in methanotrophy, and likely has an effect on the regulation of phospholipid synthesis.

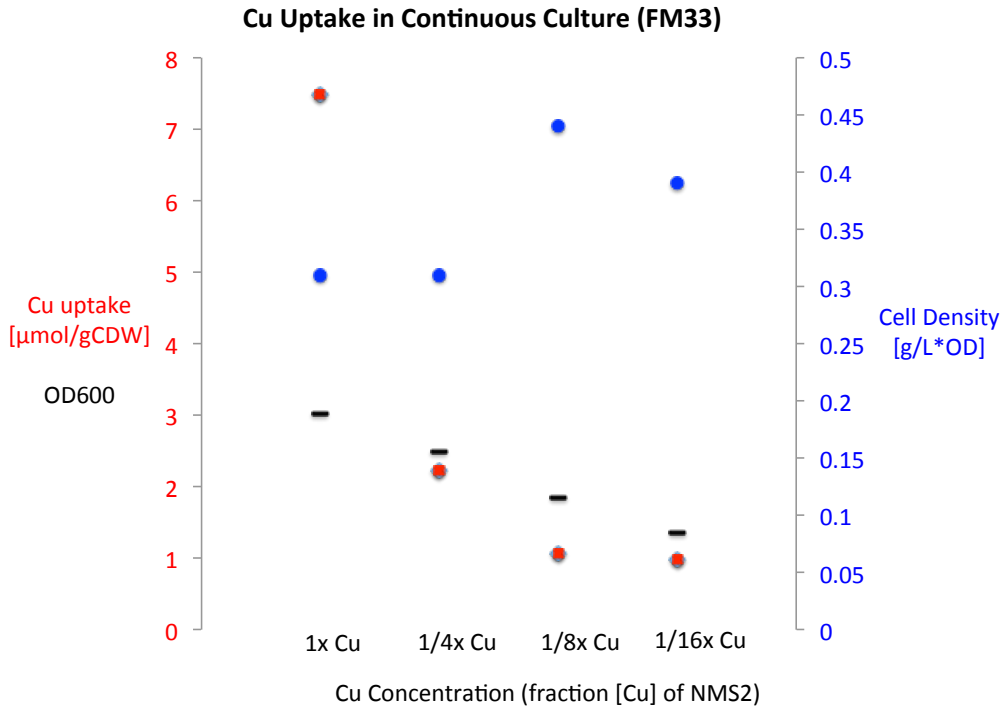

Figure 2.6. Copper medium optimization.

Cell density response to decreased Cu concentration in a chemostat culture.

Two bioreactor experiments were carried out to study *M. buryatense* response to Cu concentration in the medium. The first experiment was completed in a steady-state chemostat culture where the medium Cu concentration was decreased in a stepwise fashion over multiple days [Figure 2.6]. The dilution rate and gas feed rate remained the same. During each Cu concentration, a biomass sample had been filtered in order to determine the cell dry weight at each of the Cu limiting concentration. This is done to ensure that the specific uptake calculation remains accurate even if there are changes in the cell physiology. The cell density response to reduced Cu concentration confirms limitation, and it is assumed that the concentration of Cu remaining in solution is negligible [24]. The resulting minimum specific uptake was calculated to

be 1μmol/gCDW [Figure 2.6]. The second experiment served a dual purpose. The primary goal was to complete a -/+ Cu transition and take time-course RNA samples for transcriptomics analysis [Figure 2.7]. The objective was to measure cell lipid concentration response to Cu concentration in the medium. The transcriptomics data will be studied in order to gain a better understanding of the coupling between pMMO and phospholipid synthesis, in attempts to identify genetic targets to maximize cell lipid content. It is observed that a 2x increase in Cu results in higher lipid content of ~12% CDW, but 4x Cu concentration becomes toxic and results in a culture crash. The transcriptomics data collected during this experiment are discussed in Chapter 4 of this document.

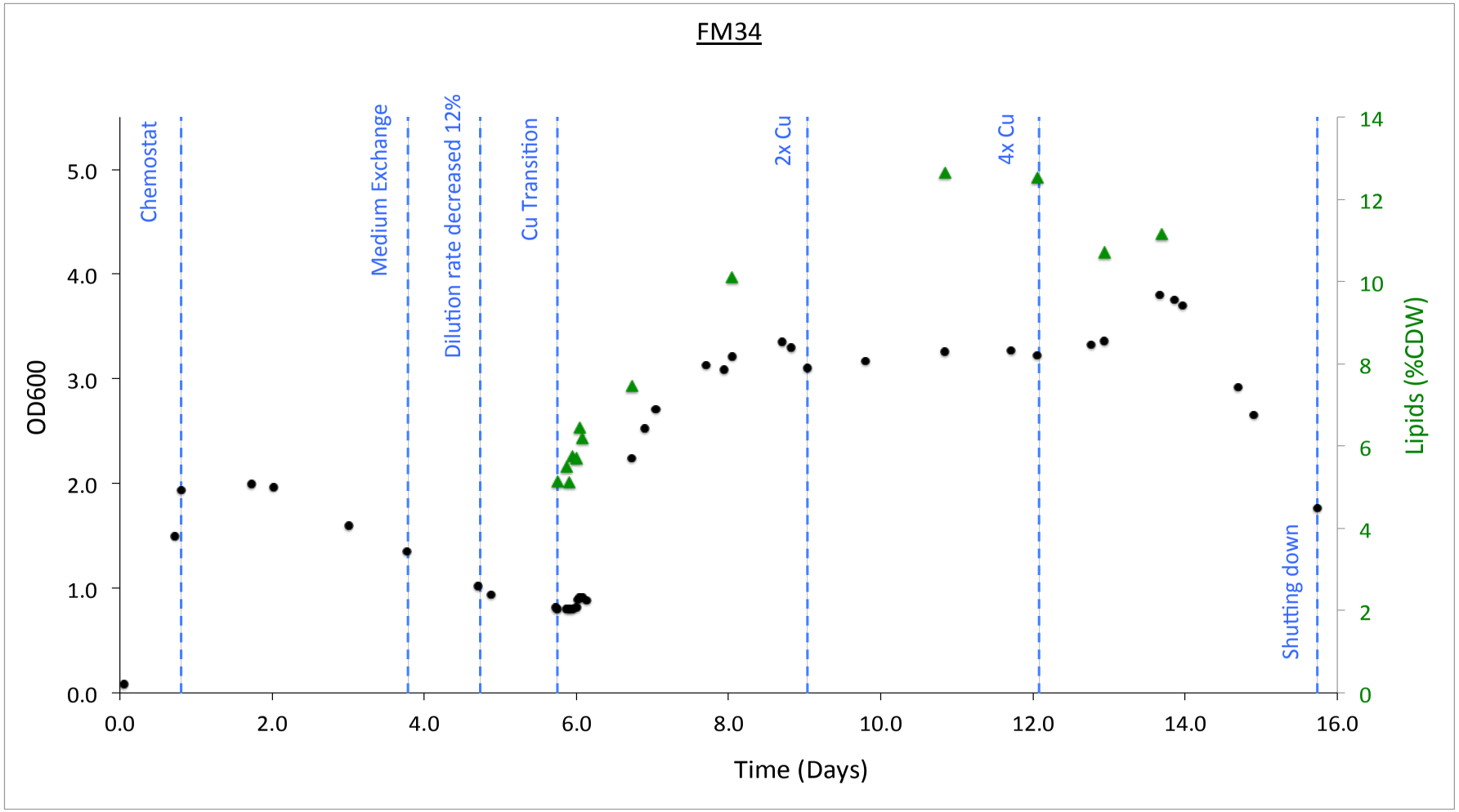

Figure 2.7. FM34 Bioreactor growth profile: -Cu/+Cu time-course switch.

Samples for lipid membrane concentration and transcriptomics were taken in a time-course fashion after the switch. Relative lipid concentration and cell density was observed to increase post Cu transition.

2.2 SUMMARY AND CONCLUSION

M. buryatense shows attractive features as an industrial strain. The culture grown under high Cu concentration exhibited the highest cell membrane content of the tested conditions. Both growth rate and carbon conversion efficiency are near the top of reported methanotroph values. The datasets presented here provide foundational information for future strain and process development - such comprehensive datasets have not previously been published for any methanotroph. Both O_2-limited and methane-limited conditions resulted in a significant glycogen content. Since a glycogen-negative mutant of *M. buryatense* is now available [25], it may be possible to redirect more carbon into non-glycogen biomass via this strain. Future process work should focus on further increasing cell lipid content as well as high cell density growth characterization. Cell densities as high as 14 g CDW L^{-1} have been published for a methanotroph [26], suggesting that with appropriate medium and bioreactor configurations, industrial-scale processes can be developed.

Flux analysis and transcriptomics datasets analysis are needed to understand the metabolic differences occurring in each of these growth conditions. The differences noted suggest a metabolic flexibility in *M. buryatense* that should be advantageous for future bioprocess strain manipulation. In this project, *Methylomicrobium buryatense* was investigated as an industrial bio-catalyst of methane. A proof-of-concept technology for the biological conversion of methane to liquid-fuel was demonstrated. This work has established baseline process estimates for the industrial use of methanotrophs. With the continued study of methanotrophy and the rapid advances in biotechnology tools, novel strain improvement have promise for yielding a marketable technology for bio-GTL.

Chapter 3. EXCRETION OF ORGANIC ACIDS IN AEROBIC METHANE OXIDIZING BACTERIA

3.1 BACKGROUND

In nature, lake sediments and soils are some of the commonly found environments where aerobic methane oxidation takes place [27]. In these natural environments, the dissolved O_2 concentration decreases with depth in the sediment as methane concentrations increase due to decomposition of biomass. Recent work has shown that some aerobic methanotrophs may have special adaptations to survive at low O_2 concentrations by mechanisms such as nitrate reduction, or excretion of organics acids in a proposed hybrid fermentation-type metabolism[16, 28]. Some O_2 is always required in order to complete the first oxidation step from methane to methanol by methane monooxygenase (MMO). In a metabolism where O_2 is used solely by MMO and the electron transport chain is inactive, the O_2-methane uptake ratio is 1:1 (3.1) [29]. The balanced reaction for enzymatic oxidation of methane by MMO:

$$\mathrm{CH_4} + O_2 + 2H^+ + 2e^- \xrightarrow{MMO} CH_3OH + H_2O \qquad (3.1)$$

Methanotrophs assimilate methane-derived carbon into biomass from formaldehyde as describe in [Figure 2.2].

The first evidence for a specialized metabolism involving excretion of organic acids was observed in *Methylomicrobium alcaliphilum* (20Z), a type I methanotroph from the gammaproteobacteria class[16]. When grown under low O_2 conditions, *M. alcaliphilum* excreted significant amounts of fermentation products (10 μmol/gCDW to 1.8 mmol/gCDW): formate, acetate, lactate, and H_2 in the flask headspace [Figure 3.1] [16]. In addition, the metabolic

pathways found in *M. alcaliphilum* are predicted to support a mixed acid type of fermentation metabolism, although as noted above, O_2 is required for oxidation of methane to methanol by the MMO. For biotechnology applications, the ability to convert methane to excreted products by switching to a fermentation-like condition would be a highly attractive process, and provides the possibility to "drop-in" modules designed for *E. coli* under similar process conditions [2]. The same fermentation-related genes are present in the genome of *M. buryatense*, suggesting it should carry out a similar type of metabolism, but nothing was known at the time about the metabolism in *M. buryatense.*

We set out to study the excretion of organic acids in methanotrophic bacteria when grown under low O_2 conditions. The model organism for this study is *Methylomicrobium buryatense*, a bacterium that is closely related to *M. alcaliphilum*, but which has a host of recently developed genetic tools available [25, 30]. The objective of this project was to characterize *M. buryatense* in a mode that resulted in high organic acid and H_2 excretion in a steady-state bioreactor culture and compare results to a respiring culture and the work described in Chapter 2[31]. Rates of methane and O_2 uptake were measured, and samples were collected to measure excreted products (H_2, formate, acetate, lactate, succinate), and RNAseq samples for transcriptomics analysis.

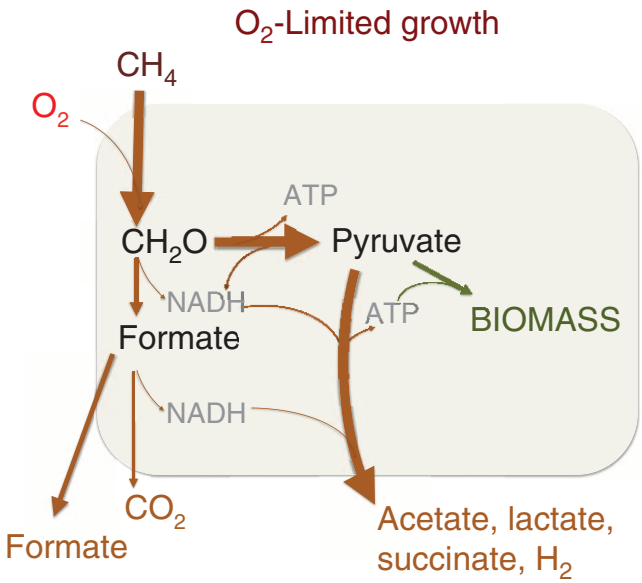

Figure 3.1. Proposed fermentation-type metabolism in aerobic methanotrophic bacteria.

During methanotrophic fermentation metabolism, it is hypothesized that O_2 is only utilized for methane oxidation and not as a terminal electron accepter in oxidative phosphorylation [32].

The acquired dataset can be utilized to constrain a previously developed metabolic flux model [21] in order to elucidate the pathways of the metabolic network and to make additional predictions. Gas uptake data was used to estimate O_2 :methane uptake ratio in order to determine if this is a true fermentation metabolism or if electrons pass through the electron transport chain to O_2. The acquired transcriptomics data will be used to gain clues about the condition involving excreted products by comparison gene expression values to respiration conditions (see chapter 4). My role in this work was to identify a chemostat culture condition where increased production of the proposed fermentation products takes place and complete a successful fermentation-respiration switchover experiment.

3.2 Results and Discussion

A bench scale bioreactor system was used for this study as described in [Figure 2.4], where the bioreactor out-gas composition is periodically sampled by a gas chromatograph. In a continuous chemostat culture grown on gaseous substrate the dilution rate (which in traditional glucose grown culture, controls the substrate feed rate), is decoupled from the gas feed rate. Therefore, in a search for conditions that generate high levels of proposed fermentation products, changing the dilution rate while keeping the gas delivery rate constant will also control the culture cell density (Appendix B). In order to identify a potential fermentation condition, the following experimental parameters were modified: gas flow rate, gas mixture composition, and dilution rate (growth rate). For safety reasons a similar set of available gas blends as described in [Figure 2.4] were used (plus: air; 100%CH_4; 20%CH_4/N_2 blend) except that these blends were mixed in varying ratios to increase experimental flexibility. The gas mixtures were kept outside the explosive range at all times. Two different mass flow controllers were used (20 SCCM and 90 SCCM).

3.2.1 Transition experiments (runs FM64 & FM69)

Since the gas chromatograph can measure H_2 in real-time, this was initially used as an indicator of the proposed fermentation metabolism. After a series of runs with various conditions and altering dilution rates, inlet gas composition, and gas flow rates; a condition was found where a steady-state culture generated a continuous detectable concentration of H_2 [Figure 3.2]. A stable production of H_2 was achieved using a premixed gas concentration containing 20%CH_4, 5% O_2 and 75%N_2 and a slow dilution rate of 0.03 hr^{-1} (24hr doubling time). The dissolved O_2 concentration (measured by optical DO probe) remained 0.07 mg/L (~1% of saturation) throughout the experiment. Supernatant samples were taken during the experiment to measure excreted products. Concentrations of excreted formate were variable. A high concentration,

comparable to previously reported values, [16], was detected during day 8 of the experiment (where +50% nutrients were spiked in the medium) but very low concentrations were observed just a few days prior when grown on standard nitrate mineral salt medium (NMS2)[17]. A moderate amount of formate was measured on day 15 with faster dilution rate and also on standard NMS2 medium. The number of measurements is cost -constrained by our need to use NMR analysis (due to the high salt and pH medium). Key experimental parameters and results are summarized in [Table 3.1]. The O_2/CH_4 uptake ratio of 1.3 and 1.4 during the two conditions suggests that a significant amount of respiration still occurs (see equation 3.1).

Table 3.1. FM64. Parameters summary table for varying dilution rates.

Similar dissolved O_2 concentration and gas uptake (highlighted) indicate limitation for gas substrate. For organic acids, maximum values are shown.

Experiment ID **Description**	FM#64 **Slow Dilution**	FM#64 **Fast Dilution**
OD	6.8	4.6
Biomass (gCDW)	1.564	1.058
Dilution Rate (hr^-1)	0.029	0.056
Doubling Time (hr)	24	12
Dissolved Oxygen (mg/L)	0.07	0.07
Hydrogen in headspace (ppm)*	14	6.5
Oxygen Uptake (mmol/hr)	7.2	7.6
Methane Uptake (mmol/hr)	5.3	5.9
Specific Oxygen Uptake (mmol/(hr*gCDW)	4.6	7.2
Specific Methane Uptake (mml/(hr*gCDW)	3.4	5.6
Oxygen/Methane Uptake Ratio	1.4	1.3
Secreted formate (mico Molar)	2830	480
Secreted acetate (micro Molar)	1852	64
Secreted lactate (micro Molar)	Not detected	Not detected
Secreted formate (mico Molar/g CDW)	1810	453.7
Secreted acetate (micro Molar/gCDW)	1184	60.5
Secreted lactate (micro Molar/gCDW)	N/A	N/A

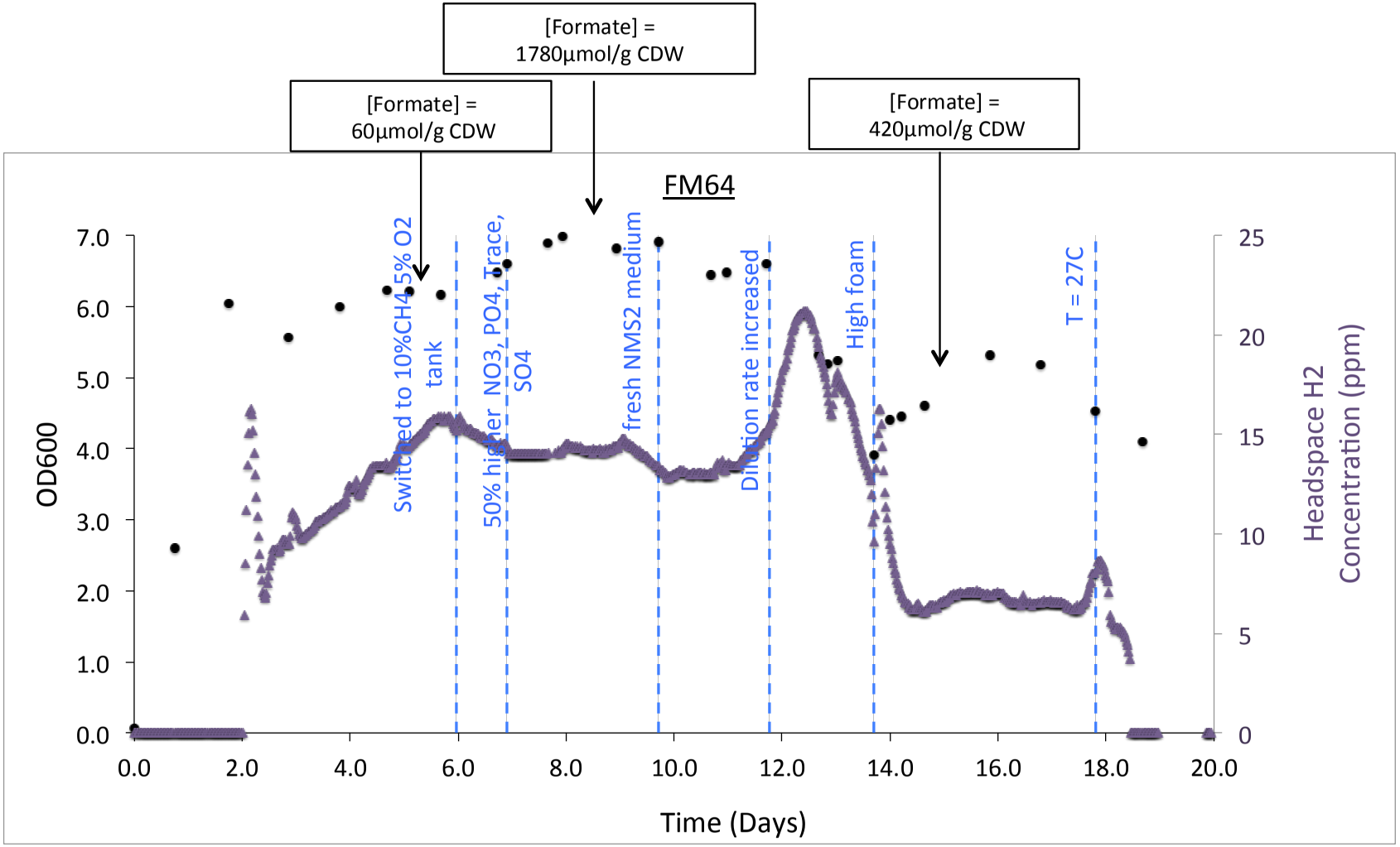

Figure 3.2. FM64 Bioreactor growth profile: fermentation condition testing.

Continuous H_2 production with varying dilution rates and nutrient concentrations is observed, with varying excreted formate concentration.

These results suggest that a component in the medium might influence the production of H_2 and excretion of organic acids. In run FM64, a condition was identified where hydrogen gas was continuously produced (suggestive of fermentation) and constant gas uptake rate and cell density response to dilution rate change confirmed gas-feed limitation (Appendix B). However, cell density response to nutrient increase, and varying concentration of excreted formate suggested that the culture might be under nutrient stress and this could be related to regulation of excreted acids. This hypothesis was tested using the slower growth condition and a switch to respiration to acquire transcriptomics datasets before and after the switch. This transition experiment (run FM69) was completed using the slow growth rate condition described for run FM64. The transition was induced by changing the inlet gas composition with a mixture of 20sccm of 100% CH_4 and 90sccm air, resulting in a mix composed of 18%CH_4, 17%O_2, and 64%N_2. During the transition experiment, continuous production of H_2 proved to be difficult to reproduce, and spikes of H_2 were observed, especially after the medium bottle was changed [Figure 3.4]. Similar to run FM64, the excreted organic acids during the transition experiment remained low with an exception of a high peak in formate on day 28 [Figure A.1]. The high formate concentration peak of ~1700µmol/gCDW is similar to that observed in run 64 and to previously reported values [16]. These data suggest that the excretion of formate is not constant as would be expected from a stable continuous culture and the cause of the spike in formate may be related to a metabolic perturbation such as nutrient stress. The headspace H_2 concentration dropped below detection limit post transition to respiration as expected.

Bacteriohemerythrin is a prokaryotic homologue of hemerythrin, a non-heme O_2 transporter protein commonly found in marine invertebrates [33]. The expression of bacteriohemerythrin is O_2-regulated and has been shown to improve in vitro pMMO activity,

presumably by providing extra O_2 [34]. During the O_2-limited slow growth rate conditions, bacteriohemerythrin was expressed ~10x higher than O_2-limited fast growth chemostat culture and unrestricted fed-batch culture (conditions as described in Chapter 2) [Figure 3.3], further confirming that the cell metabolism is responding to severe O_2 limitation. The high expression of bacteriohemerytherin during low O_2 condition of FM69 confirms a cellular stress for oxygen and as expected, transition to respiration conditions reduced bacteriohemerythrin expression to baseline level [Figure 3.3].

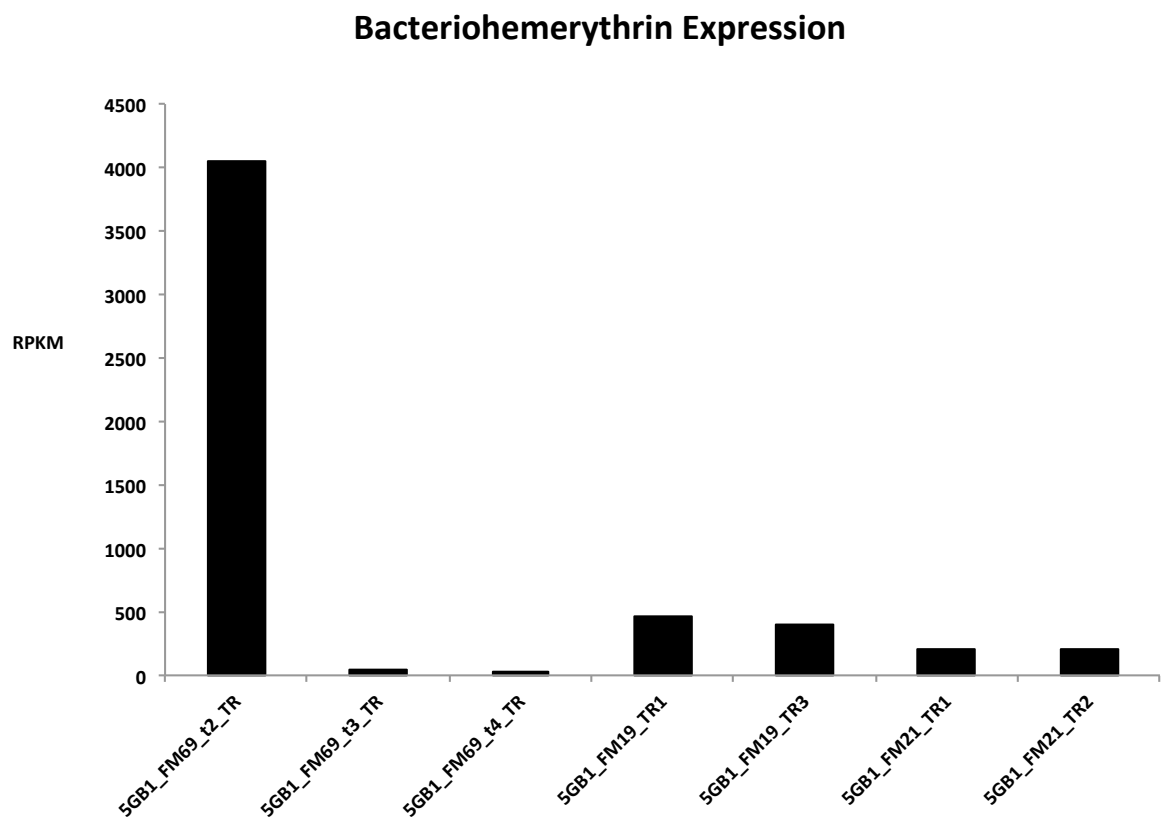

Figure 3.3. Bacteriohemerythrin expression under O_2 limitation.

Expression during slow dilution fermentation (FM69_t2), slow dilution respiration (FM69_t3, FM69_t4), O_2 - limited fast growth (FM19), and unrestricted fed-batch growth (FM21).

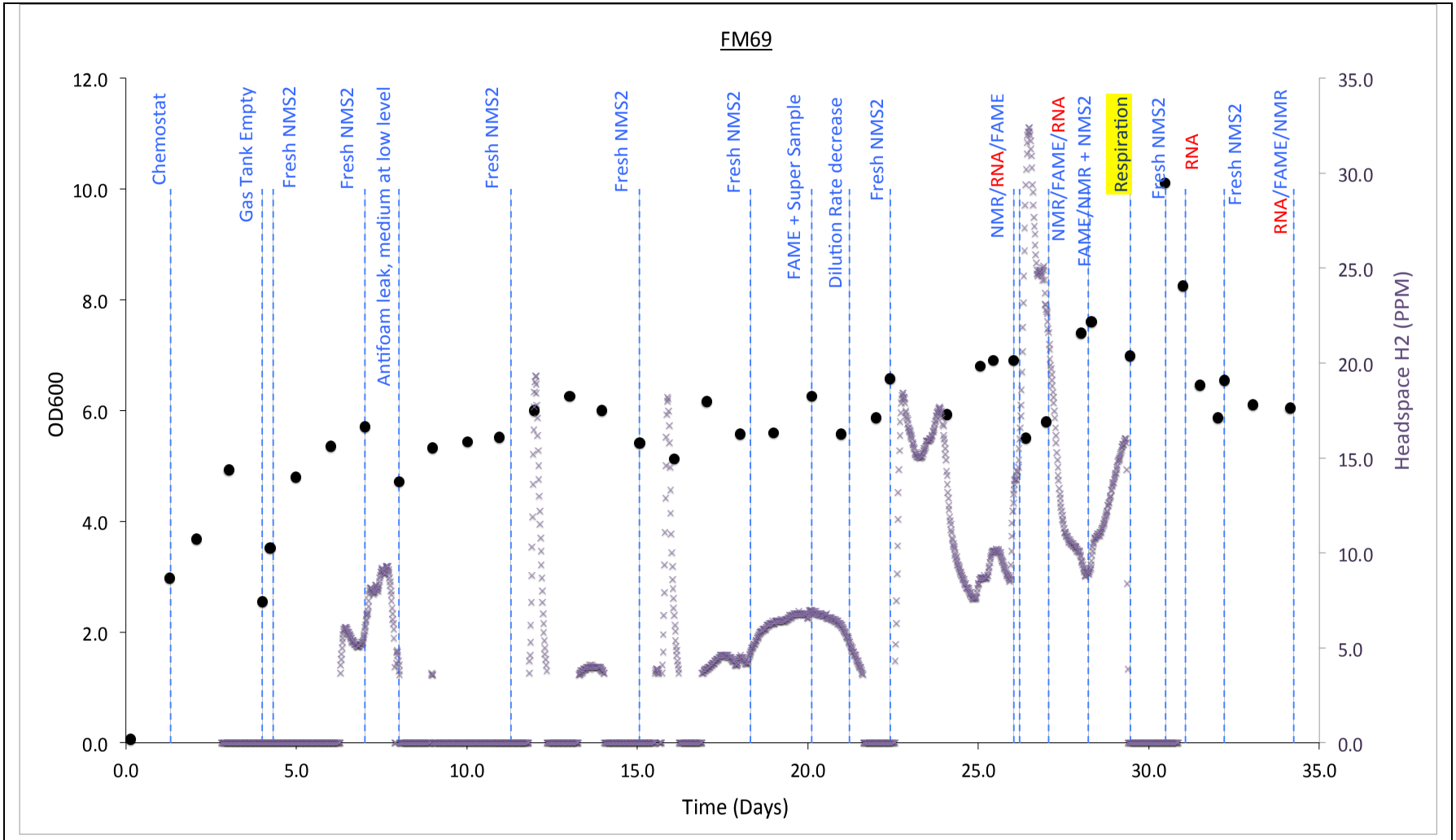

Figure 3.4. FM69 Bioreactor growth profile: switch from low O_2 to high O_2 condition.

Intermittent and variable H_2 production observed during low O_2 conditions. H_2 is not detected after switch to excess O_2.

However, under these conditions, nitrogen fixation genes (*nif*) along with genes involved in metal scavenging are also highly expressed. While bacteriohemerythrin expression confirms severe O_2 limitation, the activation of nitrogenase also suggests nitrogen starvation. More importantly, nitrogenase is known to produce H_2 during fixation of N_2 [35]. Hydrogen evolution through the NADH pathway (hydrogenase), is driven by a necessity to reoxidize NADH when NADH pools are high[36]:

$$\mathrm{NADH} + \mathrm{H}^+ \longrightarrow \mathrm{H}_2 + \mathrm{NAD}^+ \tag{3.2}$$

Biological reduction of dinitrogen to ammonia by nitrogenases occurs when both nitrogen sources and O_2 are low, and electrons are provided via ferredoxin [37]:

$$\mathrm{N}_2 + 8e^- + 8\mathrm{H}^+ \longrightarrow 2\mathrm{NH}_3 + \mathrm{H}_2 \tag{3.3}$$

The induction of these genes strongly suggested these high cell density cultures, although clearly O_2-limited, were also undergoing nutrient starvation in the form of nitrate- and metal-stress.

3.2.2 High nutrient condition (runs FM80 & FM81)

The next set of experiments was carried out in order to remove the confounding factors of nitrate and metal stress. The culture was grown in steady-state conditions similar to runs FM69 and FM64 (slow dilution rate) where continuous production of H_2 is detected. The medium was gradually exchanged to higher concentrations of nitrate and trace elements, and samples were taken periodically. The quantity of transcripts of *nifU*, *corB*, and *hugZ* (nitrogenase-related genes and two metal scavenging proteins) were measured by qRT-PCR and normalized to previous O_2-limited transcriptomics data [Figure 3.5]. The results show that 3x the regular amount of nitrate is required under this condition to repress nitrogenase expression. Similarly, 2x the regular amount of trace elements is required to repress the metal scavenging genes.

Two interesting and impactful observations were made during these experiments. Firstly, H_2 production begins to decline shortly after 2x NO_3^- is added for FM80 and after 3x NO_3 is

added in FM81 [Appendix A]. The decline of H_2 production confirms that the detected H_2 is related to either nutrient or trace-metals limitation and not severe O_2-limitation. The measured excreted products during these experiments also remained low. With the exception of the two very high spikes in formate concentration in FM64 and FM69, all other measurements remained at baseline levels. These sets of experiments show that conditions previously reported to achieve high organic acid and H_2 excretion in *M. alcaliphilum* [16] did not achieve a similar metabolic shift in *M. buryatense.* Note that *M. alcaliphilum* does not contain nitrogenase, and therefore the H_2 production observed in that strain cannot be due to nitrogenase and nitrogen-limiation.

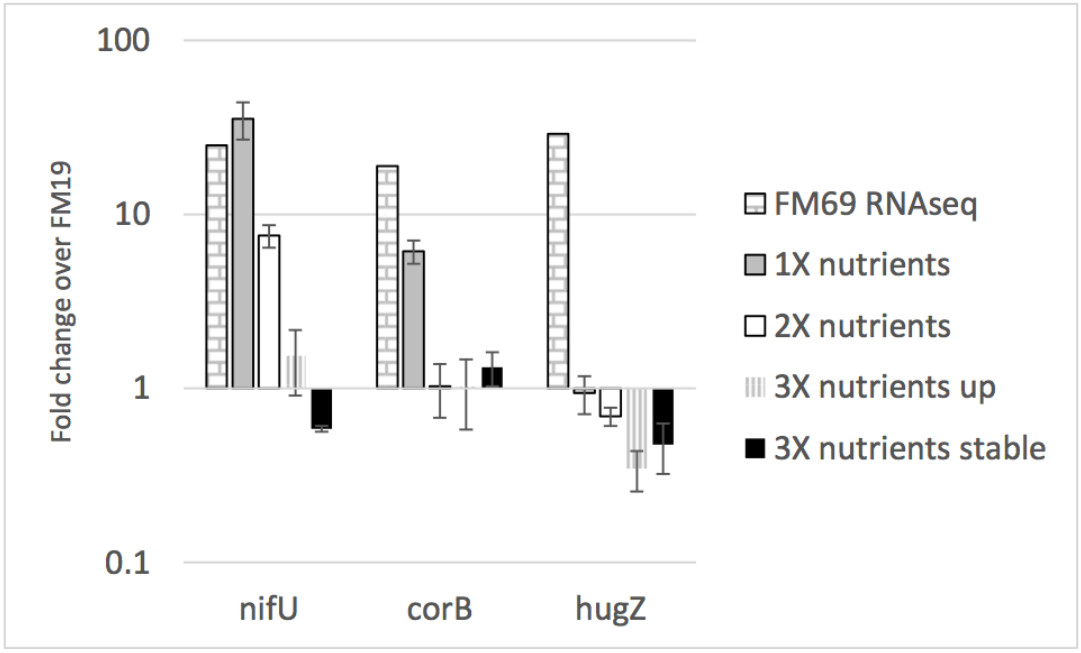

Figure 3.5. qRT-PCR for *nifU*(nitrogen fixation), *corB* (copper repressible gene), and *hugZ* (Fe acquisition).

FM69 (severe O_2 limitation slow growth rate condition) compared to FM81 (severe O_2 limitation slow growth rate condition with 1x, 2x or 3x nutrient concentration). Gene expression is normalized to O_2 -limited fast growth rate culture FM19 (Chapter 2). *nifU* is a gene involved in nitrogen fixation [35], and *hugZ* is involved with Fe acquisition siderophores [38]. [Samples processed and analyzed by Frances Chu].

3.1 SUMMARY AND CONCLUSION

In this study we have tested a variety of conditions to create a steady-state culture for high organic acid excretion by limiting for O_2 under slow growth rate and high cell density. The high cell density growth resulted in a metabolic response of high expression of bacteriohemerythrin (~10x low cell density O_2-limited growth), confirming the O_2 limitation stress. Under this condition, a detectable and sustained H_2 production was observed – the expected signal for this metabolic mode. However, it was later confirmed that H_2 production was a result of other nutrient limitation, presumably occurring as a result of the high cell density, and implicated N_2 fixation and nitrate starvation as the causative factors in H_2 production.

The measured O_2:CH_4 uptake ratios are similar to those observed in Chapter 2 for respiring cultures. If the cultures were truly fermenting, the expected ratio would be 1:1, but it was 1:1.3 and 1:1.4, suggesting that the cultures are respiring. Others in the lab will use the values obtained here to constrain metabolic models and predict ATP and NADH fluxes for specific reactions, to further assess this severely oxygen limited metabolism.

Other work in the lab involving RNAseq analysis confirms the conclusion of this study, that *M. buryatense* does not respond to severe O_2-limitation by switching to a metabolic mode involving high organic acid excretion like *M. alcaliphilum*. Instead, our combined results demonstrate that the cultures exhibit a response to severe O_2 limitation that indicates a rising NADH pool and alterations in the metabolic network to mitigate pool increase. Likely, the spikes of formate excretion observed in my bioreactor experiments reflect fluctuations in the NADH pool. When the NADH pool is high, it is expected that formate will be excreted, rather than oxidized to CO_2 to generate NADH (Figure 3.1). The information obtained in this project is important in increasing our understanding of the metabolic network in *M. buryatense*, and

suggests that achieving the goal of production-level excretion of organic acids in this strain will necessitate further metabolic network engineering. For instance, it will likely require manipulation of flux to key intermediates such as pyruvate and AcCoA, and changes to NADH-producing and consuming reactions to balance NADH pools.

Chapter 4. COMPARATIVE TRANSCRIPTOMICS DURING TIME-COURSE COPPER TRANSITION EXPERIMENT

4.1 BACKGROUND

Cu is a key metal in methanotrophy. It is the metal center of the primary methane oxidation apparatus found in most methanotrophs. It also acts as a regulatory signal for transitioning to a less efficient methane oxidation machinery when Cu is not available. The mechanism of Cu uptake and metabolic regulation of the two known methane oxidation systems are not well understood despite significant research efforts.

As part of the work described in the previous two chapters, transcriptomics data-sets were obtained from many of the bioreactor experiments. 40+ transcriptomics samples (including technical replicates) were collected and RNAseq datasets were generated. In order to analyze this dataset, we implemented data-science and bioinformatics techniques, and focused on biological questions regarding copper (Cu). The goal is to increase our understanding of the interplay between genetics, Cu uptake, and the effect of Cu on metabolic regulation. We identified gene expression patterns that are unique to a time-course metabolic transition experiment, and formulated a novel testable hypothesis relating to important knowledge gaps in methanotrophy and Cu-regulated genes.

4.1.1 sMMO and pMMO

All aerobic enzymatic methane oxidation is carried out by methane monooxygenase (MMO) [29]. Two distinct and unrelated types of MMO exist that have evolved independently to accomplish the same task: the cytoplasmic soluble methane monoxygenase (sMMO), and the

membrane associated particulate methane monooxygenase (pMMO) [Figure 4.1]. These enzymes have distinct structures, and different active metal centers; however, methane oxidation takes place via the same overall reaction (see reaction 3.1). Methane is oxidized to methanol by the accompanying reduction of dioxygen molecule into water. Both forms of MMO are monooxygenases – both react one oxygen atom of O_2 to complete the oxidation step from methane to methanol, and the other oxygen atom is converted to H_2O. A critical component of this study has to do with the different metal centers of the two MMOs. pMMO uses copper as the active metal center and sMMO contains iron as its active metal center. pMMO is the more efficient enzyme of the two but is active only if copper is bioavailable [6].

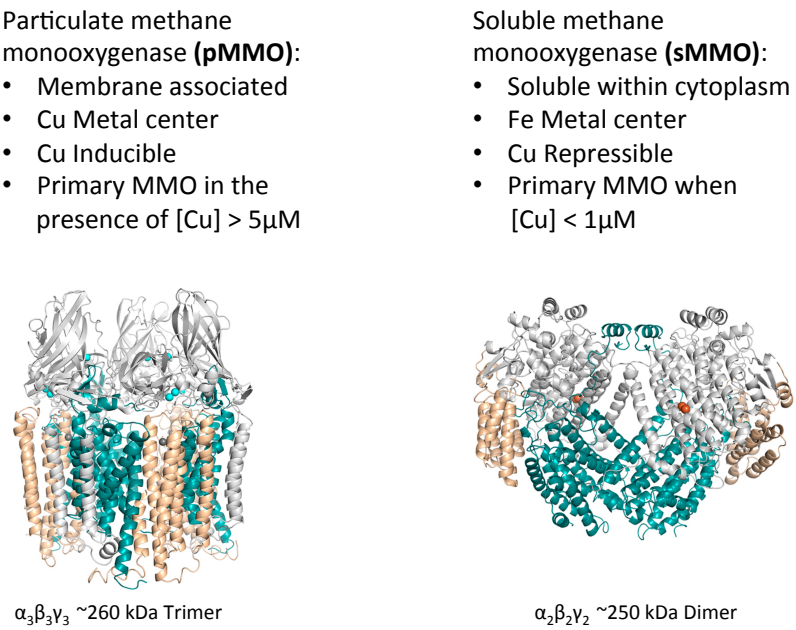

Figure 4.1. The two forms of methane monooxygenase: pMMO and sMMO

The switch between the two methane oxidation systems is regulated by Cu bioavailability [6].

Most aerobic methanotrophs have pMMO, while a subset have both pMMO and sMMO. A few bacteria have been discovered to only have sMMO in their genome [39]. Methanotrophs

that possess both pMMO and sMMO have a unique capability of switching between sMMO and pMMO based on the availability of copper. These bacteria will oxidize methane by pMMO if soluble copper concentrations are above 5μM, but will switch to sMMO at lower Cu concentrations [6]. As a result, a number of copper switch experiments have been completed in efforts to study methane oxidation [23, 40-44]. *Methylomicrobium buryatense* has both sMMO and pMMO. We have carried out a Cu-induced metabolic switch in a steady-state bioreactor culture and collected time-course transcriptomics samples post the Cu transition. The goal of this study is to explore co-regulated genes during the Cu-induced metabolic transition and to identify genes that might be involved in Cu uptake and metabolism.

4.1.2 Cu switch time-course experiment

Some impactful -/+ Cu comparison and switch experiments have been completed in the past. These experiments include transcriptomics and proteomics of various methanotroph species [23, 40], however, this is the first reported Cu-switch time-course transcriptomics study for any methanotroph. This experiment was described in Chapter 2, and involved growing *M. buryatense* culture in a steady-state chemostat culture with no added Cu in the growth medium. The metabolic switch was initiated by exchanging to the standard Cu-containing growth medium, and RNA samples were collected at time points for a duration of 8 hours after the switch [Figure 2.7]. The addition of Cu into the bioreactor induced a higher rate of cell growth and a new sustained steady-state cell concentration. The chemostat cell density was declining before the addition of Cu despite a dilution rate below 80% of maximum that is predicted as a sustaining dilution rate by chemostat theory [45].

In addition to the RNAseq, cell biomass samples were also collected in order to measure the concentration of membrane lipids within the cell. As expected, membrane lipid content

increased after the addition of Cu. This effect has been observed in other methanotrophs, and pMMO is known to be membrane bound and characteristic of methanotrophs grown in the presence of Cu. This experiment was completed as part of the bio-GTL project described earlier in this document.

4.2 Computational Data Analysis

Transcriptomics samples from a total of 12 bioreactor experiments (including the Cu-switch time-course) were processed and analyzed [Table 4.1]. RNA samples were collected according to standard procedure as previously described [16] (RNA samples collected and processed by Frances Chu). RNA sequencing was completed by 2nd gen Illumina sequencing. cDNA read fragments were mapped to the *M. buryatense* genome using accepted standard technique - burrows wheeler alignment [46, 47]. The resulting data-set was then normalized for gene length and total counts using the transcripts per million method (TPM)[48, 49]. The normalized read counts (TPM) were the starting point for the computational data analysis.

Table 4.1. Summary table of bioreactor experiments with available transcriptomics data

Experiment ID	Limiting Nutrient	Growth Rate (per day)	Note
FM12	CH_4	3.0	CH_4 limited chemostat
FM14	CH_4	3.0	CH_4 limited chemostat
FM18	N/A	4.1	Unrestricted growth on methanol
FM19	O_2	3.5	O_2 limited chemostat
FM20	N/A	5.2	Unrestricted growth on methane
FM21	N/A	5.4	Unrestricted growth on methane
FM22	O_2	4.2	O_2 limited chemostat
FM34	Cu	2.7	Cu switch time-course
FM40	Cu	2.9	Cu switch time-course
FM69	O_2 +	0.7	"Fermentation" to respiration switch
FM80	O_2 +	0.7	Severe O2 limitation , incremental nutrient increase
FM81	O_2 +	0.7	Severe O2 limitation , incremental nutrient increase

4.2.1 Time-course data challenges

The statistics relating to time-course data are notoriously difficult and there is no current consensus as to the standard method for time-course transcriptomics data [50]. However, several different statistical packages have been developed for the analysis of time-course transcriptomics [51-54]. We attempted using the MAseqPro2[51] statistical package for this time-course data-set, however, it was quickly discovered to be incompatible due to lack of biological replicate samples. A different approach had to be implemented in order to identify genes with similar expression profiles across the time-course.

4.2.2 Distance metric across n dimensional sample space

Similarly to finding the geometric distance between two vectors, one can calculate the difference in the expression profile for the same gene across a time course with n samples. In this case, each dimension is the expression value of each time point sample:

$$\text{distance} = \sqrt{(a_0 - b_0)^2 + (a_1 - b_1)^2 + \cdots + (a_n - b_n)^2} \quad (4.1)$$

In the above equation, a and b are expression values of the same gene in a control and treatment parallel time-course experiments. In an experimental setup like this, the difference in expression profiles (distance) is readily determined and the genes with the highest differential expression between treatment and control are readily identified. However, this analysis requires two parallel time-course experiments with a control (absence of Cu) and a treatment (addition of Cu at start of time-course). Our experiment was not completed in parallel with a treatment and a control, but only one treatment group - the metabolic switch time-course. In order to identify groups of co-regulated genes that stand out during the transition experiment, we incorporated all the transcriptomics datasets from prior experiments, which include a wide range of conditions [Table 4.1]. The total transcriptomics dataset is then (i) Log_2 fold transformed and normalized to a baseline condition (fed-batch on methanol), (ii) converted to a euclidean (geometric) distance table between every pair of genes for the n-dimensional (every sample as a dimension) vector, and lastly (iii) separated into clusters of genes with similar expression vector distances using the k-means clustering algorithm.

4.2.3 Log_2 fold transform and normalization

Differences in gene expression between different treatment groups may span several orders of magnitude and high variance. A log_2 fold transformation and normalization to a standard

condition helps reduce the variance between samples and is considered standard practice for transcriptomics data sets. A logarithmic transformation and ratio allows for the comparison of order of magnitude of differential expression rather than the absolute and is also more intuitive to visualize and interpret.

4.2.4 Euclidean distance metric table

A pairwise distance metric table is computed prior to clustering; the euclidean distance of gene expression between every pair genes. The table contains each gene expression profile as a vector with expression value of each sample as a dimension. The k-means algorithm then identifies k number of clusters that have expression distances of similar values.

4.2.5 K-means clustering

k-means is an unsupervised machine learning algorithm. It is one of the first and widely used clustering algorithm [55]. K-means works by randomly placing k number of seed centers (clusters) within the dimensions and range of the desired data-set. Each data point in the sample-space is then assigned to the closest cluster center. In an iterative fashion, the seed point is then moved to the center of all the data points that were assigned to that cluster. The data points are then resigned to the new closest cluster centers after the move. This iteration of moving cluster centers to the data-point centroids, reassigning every data-point to the new cluster center, and then again moving the cluster center to the centroid of the newly formed cluster is continued until the clusters centers converge to a set position. The k-means algorithm is very effective in identifying clusters of dense data points in a multi-dimensional sample space. The major challenge of using k-means is that the user has to specify the k-value a-priori. For example, if k=1, then the entire dataset falls into a single large cluster. The other extreme is that if k =

number of data points, then each data point is its own distinct cluster. It is difficult to pick a k value if little is known about the data. Heuristics methods are available in order to attempt to elucidate the number of clusters within unknown data. Two common methods are the elbow analysis and Silhouette analysis. Nonetheless, k-means should be ideally used with a data-set with some known properties and a good estimate of the number of clusters within the data-set.

4.2.6 Elbow analysis

The elbow method is a way of determining the k value for k-means clustering analysis. This is a visual method that involves computing the k-means clustering for a range of k-values. The average sum of square error (SSE) is calculated for every k value in the selected range. For increasing k-value, it is expected that the SSE will decrease as the result is approaching closer to the scenario where each point belong to its own cluster (SSE=0). When plotting the sum of square error of every cluster against the range of k-values, the “elbow” may occur where the plot has an abrupt change of slope. This is the k-value to be selected for clustering. In our analysis, however, the “elbow” point is not readily identified and the elbow method did not prove to be useful [Figure 4.2].

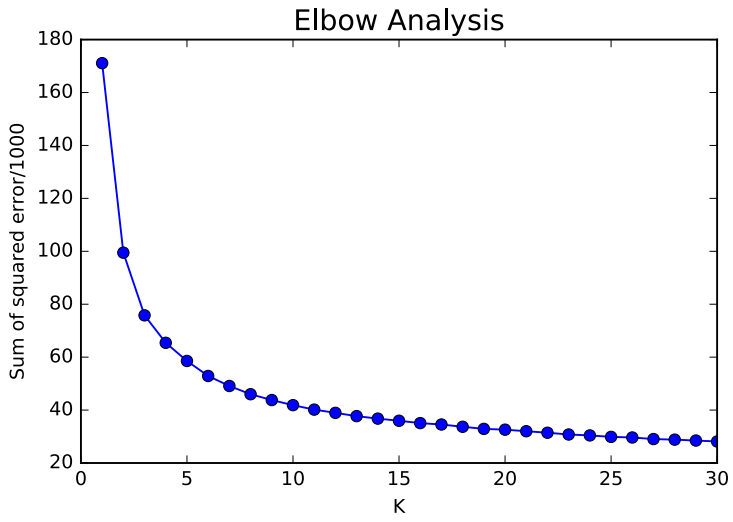

Figure 4.2. Elbow analysis plot for range of k values.

The elbow method is a very generic heuristic approach to estimate number of clusters in an unknown dataset.

4.2.7 Silhouette analysis

The silhouette analysis is a measure of both cohesion of points within their cluster and separation from the next nearest cluster. The silhouette coefficient is determined by first calculating the average distance of one data-point to all other points belonging to the same cluster, this value is denoted as A. The second step is to calculate the average distance of that same point to all the points in the nearest neighboring cluster, this is the separation value B. The silhouette coefficient of a point is the difference between A and B, divided by the maximum of A or B. If the cohesion within clustered data points is high, A is small. If the separation is high, then B is large. The silhouette coefficient approaches 1 if the overall clusters are well separated with tightly packed cohesive points within each cluster. For the dataset in this study, the silhouette analysis did not prove to be useful at suggesting a k-value [Figure 4.3]. For our dataset, silhouette analysis was skewed toward a smaller k-values which would not be useful for this situation. This approach was abandoned.

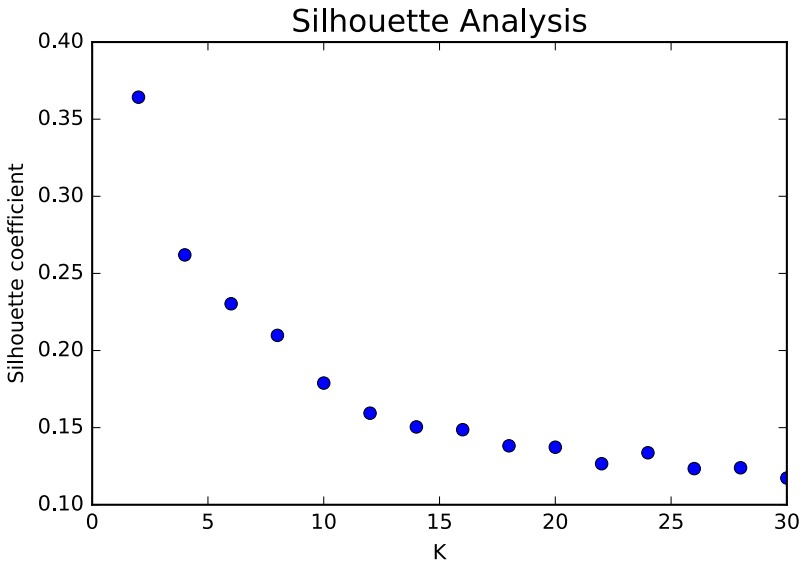

Figure 4.3. Silhouette analysis plot for range of k values.

The Silhouette method is a very generic heuristic approach to estimate number of clusters in an unknown dataset.

4.2.8 Biological data

An advantage for this dataset is that some key information about the data was known, and this was used as a starting point. Since sMMO is known to be Cu-repressible, a list of sMMO related genes was generated and these were confirmed to be highly expressed at the start of the Cu-switch experiment and then repressed throughout the time-course [Table 4.2][Figure 4.4]. This approach involves computing the k-means clustering analysis for varying k-value until a k is found for which the list of the known Cu-repressible genes form an individual cluster. The analysis started with a high k (k=50), and incrementally decreased until the desired cluster was identified. For this case, it is preferred to start with more clusters and merge them together for each decreasing k value, rather than start with a few noisy clusters and split them apart with increasing k. The desired cluster was obtained when k=20 [Figure 4.5].

Table 4.2. List of Cu-repressible sMMO related genes.

Gene Label	Gene Annotation
MBURv2_130043	60 kDa chaperonin 3
MBURv2_130044	conserved protein of unknown function
MBURv2_130045	conserved protein of unknown function
MBURv2_130046	Methane monooxygenase component C
MBURv2_130047	Soluble methane monooxygenase component D (modular protein)
MBURv2_130048	Methane monooxygenase component A gamma chain
MBURv2_130049	Methane monooxygenase regulatory protein B
MBURv2_130050	Methane monooxygenase component A beta chain
MBURv2_130051	Methane monooxygenase component A alpha chain

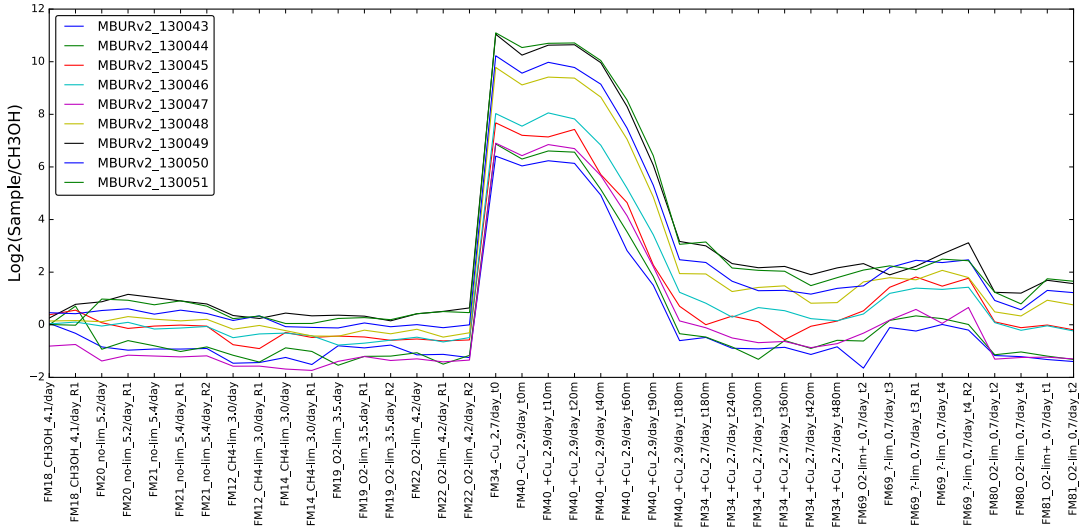

Figure 4.4. Expression profile of Table 4.2 genes across available transcriptomics data.

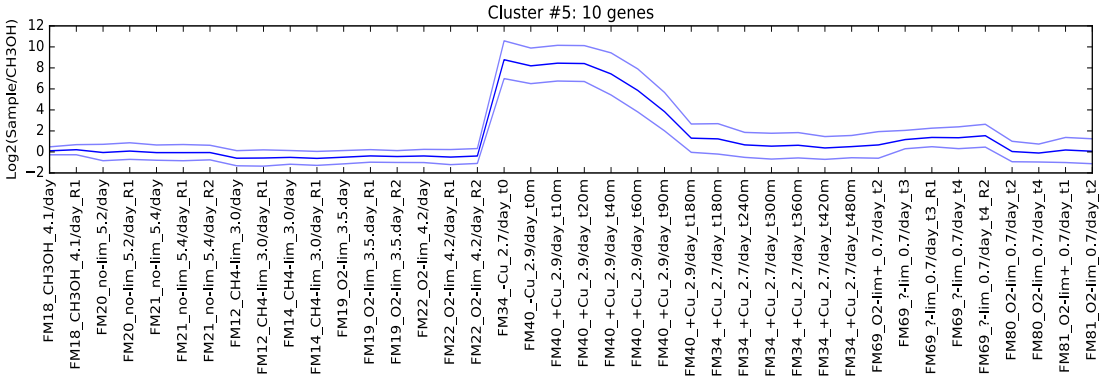

Figure 4.5. k-means clustering identified genes from Table 4.2 as a single cluster for k=20.

4.3 Results and Discussion.

The result of this work is the analysis of the 20 clusters with the intent of identifying differential gene expression during the Cu transition time-course samples. The complete list of figures of the average expression profiles of the 20 identified clusters is found in Appendix C.

4.3.1 Cluster 8 and cluster 18 – genes of interest

From the 20 clusters, there are two that stand out in particular during the time-course samples. These are clusters #8 (6 genes), and #18 (18 genes) [Figure 4.6][Figure 4.7]. Both show high expression at time point zero (no cupper), and consistent repression throughout the time course. The expression profile is very similar in the two clusters, but the expression magnitude differs. The set of genes from clusters 8 and 18 are listed in [Table 4.3]. It is important to note that besides the time-course transition, other sample points stand out with high expression – this is FM69 time points #2 [Figure A.1] and both samples from FM81 [Figure A.4]. FM69 is severe O_2 limitation to respiration transition. Clusters 8 &18 genes show high expression in the time-point prior to the high O_2 respiration switch. A possible explanation is that under this condition of high O_2 stress, the metabolism is less efficient at Cu uptake, in addition to Cu depletion due to relatively high cell density. FM81 is a replicate experiment of FM80, a O_2 limited growth condition with incremental nutrient increase. It is not clear why these genes are highly expressed in FM81 but not during FM80 experiment.

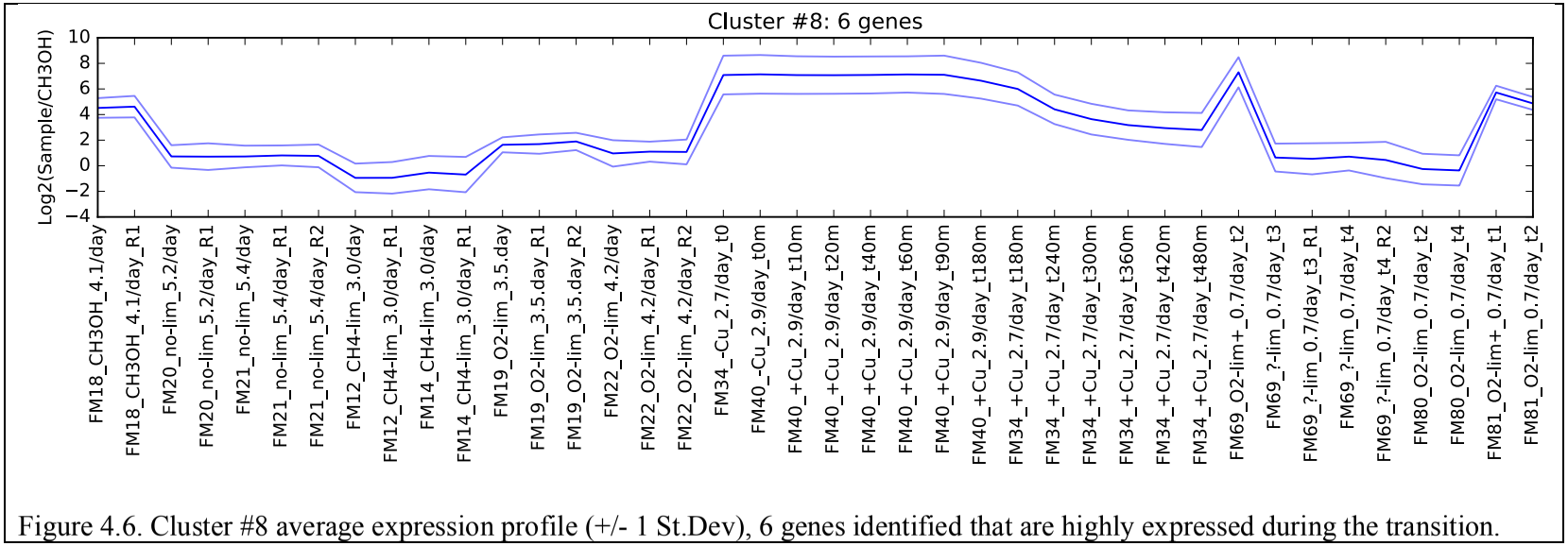

Figure 4.6. Cluster #8 average expression profile (+/- 1 St.Dev), 6 genes identified that are highly expressed during the transition.

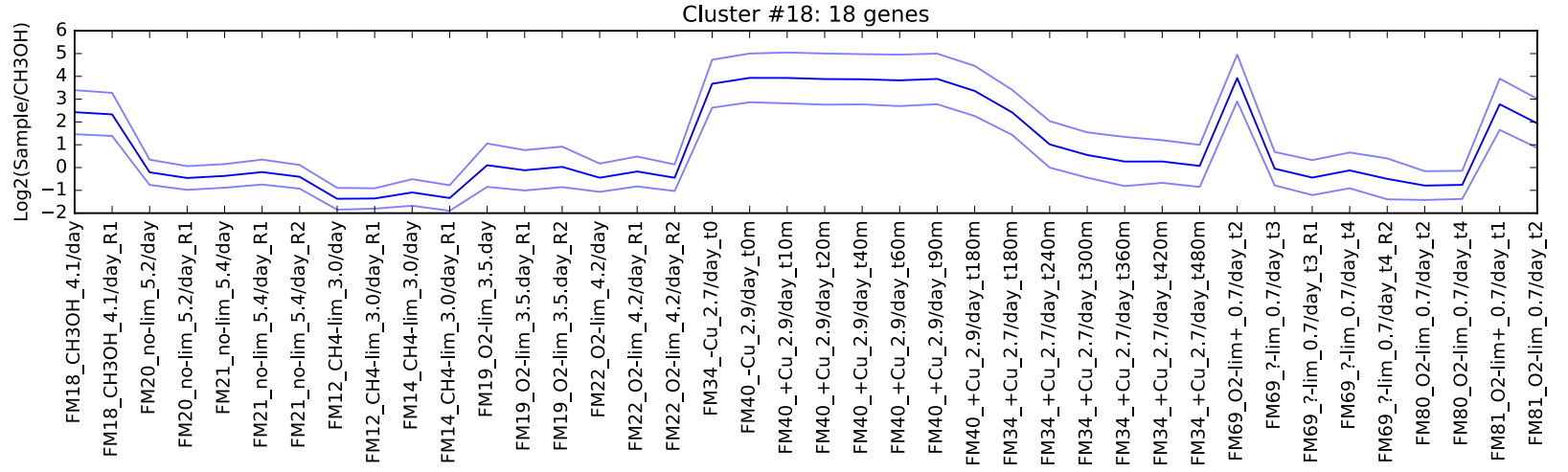

Figure 4.7. Cluster #8 average expression profile (+/- 1 St.Dev), 18 genes identified that are highly expressed during the transition.

Table 4.3. List of genes from cluster #'s 8 & 18 (24 genes).

Label	Annotation	Length (nucelotides)	Cluster
MBURv2_200002	conserved protein of unknown function	911	8
MBURv2_210001	protein of unknown function	149	8
MBURv2_210002	conserved exported protein of unknown function	395	8
MBURv2_210003	conserved exported protein of unknown function	569	8
MBURv2_210004	conserved protein of unknown function	1397	18
MBURv2_210005	conserved protein of unknown function	653	8
MBURv2_210006	**General secretion pathway protein E**	1808	18
MBURv2_210007	**putative Bacterial type II secretion system protein**	1217	18
MBURv2_210008	conserved exported protein of unknown function	464	18
MBURv2_210009	conserved exported protein of unknown function	2381	18
MBURv2_210010	conserved protein of unknown function	1250	18
MBURv2_210011	exported protein of unknown function	2678	18
MBURv2_210012	conserved protein of unknown function	1484	18
MBURv2_210013	conserved protein of unknown function	701	18
MBURv2_210014	**putative methanol dehydrogenase regulatory protein**	905	18
MBURv2_210015	conserved membrane protein of unknown function	974	18
MBURv2_210016	**Transglutaminase domain protein**	1949	18
MBURv2_250053	**TonB-dependent receptor domain protein**	2168	18
MBURv2_60380	**Copper-repressible polypeptide**	707	8
MBURv2_60381	**CorB**	2135	18
MBURv2_160425	exported protein of unknown function	1364	18
MBURv2_160426	conserved exported protein of unknown function	2576	18
MBURv2_160023	conserved protein of unknown function	230	18
MBURv2_130844	**RNA polymerase, sigma-24 subunit, ECF subfamily**	503	18

4.3.2 Gene of interest MBURv2_200002

Clusters 8 & 18 contain 24 genes of interest that will be further studied. Of these 24 genes, one stands out in particular: MBURv2_200002. It is a gene of unknown function and length of 911 nucleotides. Several features flag this gene as interesting: (i) It is the highest expressed gene in the absence of Cu and during the first several time points after Cu induction. This is highly unusual; all transcriptomics samples collected outside the Cu-switch show MMO as consistently highest expressed [Table 4.4]. (ii) MBURv2_200002 is found on its own contig in the genome. This suggests that it may be flanked by transposable elements – a unique feature [Figure 4.9]. (iii) The PHYRE2 protein structure prediction tool [56] was used predict the structure of the

protein encoded by this gene. A transmembrane helix domain is predicted with 99.7% confidence, suggesting a membrane-bound protein [Figure 4.10]. (iv) MBURv2_200002 is conserved across a number of methanotrophs that require sodium ions for growth, and with the exception of *M. luteus*, were isolated from marine or hypersaline environments [Table 4.5]. It does not appear to be ubiquitous in methanotrophy, yet the fact that it is conserved in other halophilic methanotrophs may be indicative of important function.

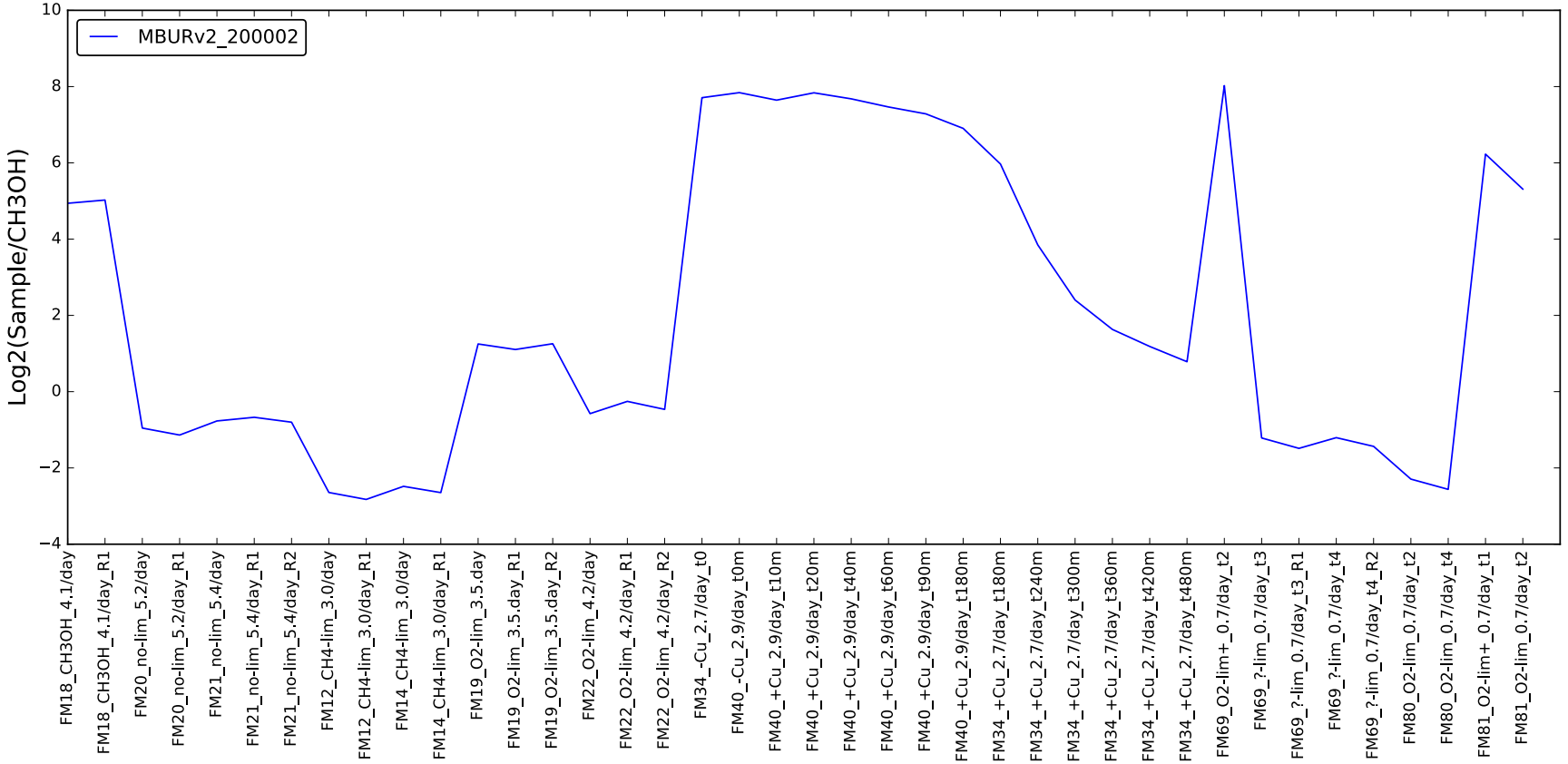

Figure 4.8. Expression profile of gene MBURv2_200002 across all transcriptomics samples.

Table 4.4. Top 3 highest expressed genes during different growth conditions.

rank	Unrestricted (CH4)	Unrestricted (CH3OH)	O2 limited	Cu Switch t0	Cu Switch t20	Cu Switch t90
1	pmoC	pmoC	pmoC	MBURv2_200002	MBURv2_200002	pmoC
2	pmoB	pmoB	pmoB	mmoX	pmoC	pmoB
3	pmoA	tRNA	pmoA	pmoC	mmoX	MBURv2_200002

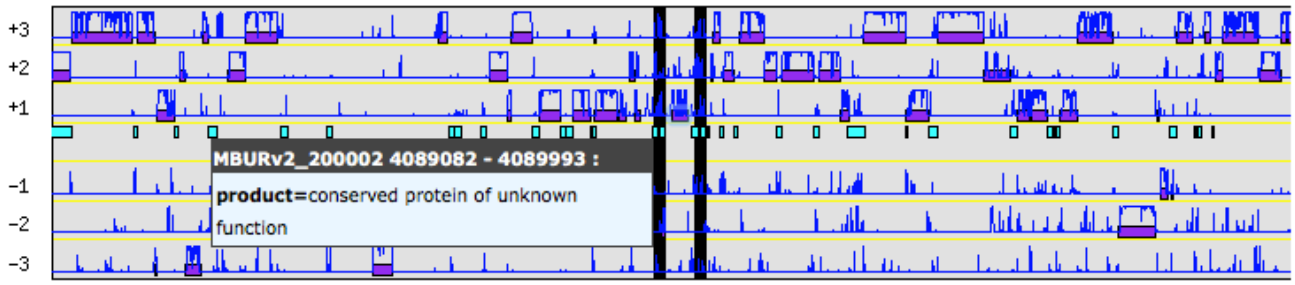

Figure 4.9. Segment of *M. buryatense* genome with the gene of interest (MBURv2_200002), found on a separate contig.

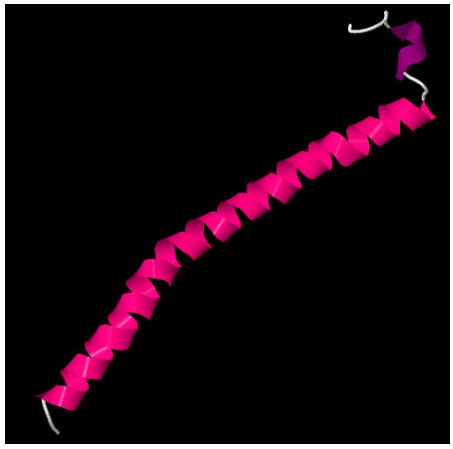

Figure 4.10. Phyre2 structure predictions of a portion of gene: MBURv2_200002.

4.3.3 Genes from cluster #s 8 & 18

The remaining genes of clusters 8 and 18, while not as unique as MBURv2_200002, offer additional clues about the co-regulated functionality of these genes. The expression profile for

the genes within these two clusters is similar, and it is likely that these genes encode interacting proteins that complete a specific function. Some interesting observations from gene labels and gene annotations include [Table 4.3]: (i) the sequentially labeled genes MBURv2_210001 – MBURv2_210016. Most are of unknown function, however, the annotation includes general secretion pathway proteins. (ii) MBURv2_210001 – MBURv2_210016 are located within 30 kilo-base pairs from pMMO and appear to be conserved in the same methanotrophic species as MBURv2_200002 [Table 4.6][Table 4.7]. (iii) MBURv2_250053 is a cluster 18 gene that is annotated as TonB-dependent receptor protein. TonB receptors are a family of outer membrane proteins that are part of a complex required for the transport of large compounds involved in metal uptake, vitamin B12 uptake, etc. The most common example is the transport of siderophores in Fe uptake [57]. The high expression of a TonB-related gene would be consistent with involvement in metal uptake.

Table 4.5. BLAST search result for MBURv2_200002

Description	Max Score	Total Score	Query Cover	E value	Ident
hypothetical protein [Methylobacter luteus]	248	248	99%	7.00E-78	47%
hypothetical protein [Methylobacter marinus]	231	231	99%	3.00E-71	45%
hypothetical protein [Methylobacter sp. BBA5.1]	225	225	97%	3.00E-69	45%
hypothetical protein [Methylobacter whittenburyi]	217	217	97%	9.00E-66	44%
hypothetical protein [Methyloprofundus sedimenti]	129	129	90%	2.00E-31	35%
hypothetical protein [Methylomicrobium alcaliphilum]	108	108	80%	8.00E-24	36%
hypothetical protein [Methylohalobius crimeensis]	93.2	93.2	89%	2.00E-18	29%
hypothetical protein [Methylomarinum vadi]	61.2	61.2	29%	4.00E-07	56%

Table 4.6. BLAST search result for MBURv2_210004

Description	Max Score	Total Score	Query Cover	E value	Ident
hypothetical protein [Methylomicrobium alcaliphilum]	943	943	100%	0.00E+00	98%
hypothetical protein [Methyloprofundus sedimenti]	448	448	98%	8.00E-152	46%
MULTISPECIES: hypothetical protein [Methylobacter]	373	373	98%	2.00E-122	41%
hypothetical protein [Methylobacter luteus]	360	360	98%	3.00E-117	39%
hypothetical protein [Methylomarinum vadi]	350	350	98%	3.00E-113	40%
hypothetical protein [Methylohalobius crimeensis]	258	258	97%	1.00E-77	31%

Table 4.7. BLAST search result for MBURv2_210010. Autotransporter superfamily detected

Description	Max Score	Total Score	Query Cover	E value	Ident
hypothetical protein [Methylomicrobium alcaliphilum]	672	672	80%	0.00E+00	99%
hypothetical protein [Methylobacter sp. BBA5.1]	158	158	78%	5.00E-41	31%
hypothetical protein [Methylohalobius crimeensis]	157	157	76%	9.00E-41	29%
MULTISPECIES: hypothetical protein [Methylobacter]	156	156	78%	3.00E-40	31%
hypothetical protein [Methylobacter luteus]	155	155	80%	7.00E-40	30%
hypothetical protein [Desulfuromonas sp. TF]	145	145	60%	1.00E-36	30%

4.3.4 Methanobactin: known system for Cu acquisition in methanotrophs

Methanobactin is siderophore-like Cu chelater, a low molecular mass (<1,200 Da) ribosomal peptide that has a very high binding affinity to Cu (amongst the highest of known natural products) [Figure 4.11] [58, 59]. It was first discovered in the membrane bound protein fraction during a proteomics study that focused on pMMO [60]. It is expressed under low Cu conditions and is part of an operon of several genes that catalyze the post translational modification of the precursor methanobactin peptide.

Methanobactin has been intensively studied for the last decade and is proposed for biomedical applications in the treatment of Wilsons disease. There are several research groups that study methanobactin, however, only 16 species of bacteria (including many non methanotrophs) have been shown to encode methanobactin genes [61]. The majority of methanotroph genomes, including that of *M. buryatense,* do not contain a recognizable methanobactin gene or the other genes in the methanobactin cluster. Since Cu is an important component of methane oxidation, an outstanding question is what is the Cu uptake system in methanotrophs that do not possess methanobactin?

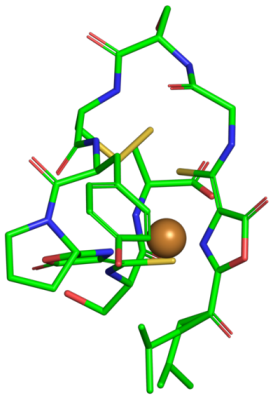

Figure 4.11. Cu binding peptide methanobactin from *Methylosinus trichosporium* OB3b. Image from Protein Data Bank of Europe.

4.1 Summary and Conclusion

The transcriptomics data analysis presented in this work reveals clusters of genes that have similar expression profiles across the Cu-induced metabolic switch time-course samples, as well as the other transcriptomics datasets. The expression of these genes is repressed in the time-point samples following Cu addition. Most of these genes are of unknown function and high expression in the absence of Cu. There is evidence for membrane bound genes, exporter genes,

secretion system genes, and a TonB receptor gene. All this evidence taken together is consistent with the hypothesis that these genes are involved in a novel Cu uptake system in *Methylomicrobium buryatense*.

4.1.1 Proposed hypothesis

The proposed hypothesis for clusters 8 & 18: genes from [Table 4.3] are responsible for Cu uptake and transfer to pMMO, with MBURv2_20002 being of particularly high relevance in this system.

4.1.2 Experimental plan for proposed hypothesis

If the proposed hypothesis is correct and the identified cluster of genes is related to Cu uptake mechanism, then a growth defect would be expected in a modified strain that is missing this set of genes and grown at low Cu condition. However, since *M. buryatense* has the capability of switching to sMMO during Cu deficiency, the growth phenotype will be missed unless the sMMO operon is also deleted from the genome. Fortunately, an sMMO negative mutant of *M. buryatense* is already available in the Lidstrom lab, and only the deletion of the identified cluster needs to be performed. A growth curve of the wild type strain, the sMMO knockout strain, and the double (sMMO and identified cluster) knockout across a range of Cu concentrations will be completed in shake-flasks across the following range of Cu concentrations (μM): [0.1, 0.5, 1, 2, 8]. Wild type strain should have a slight or no growth defect across all Cu concentrations due to the ability to switch to sMMO. A growth defect is expected for the ΔsMMO strain at a low Cu concentration due to its inability to switch to sMMO for methane oxidation. Lastly, a growth defect is expected at a higher Cu concentration for the double knockout mutant (ΔsMMO,

Δcluster) due to the hypothesized inability to scavenge for Cu and or switch to sMMO when Cu is not available.

Chapter 5. OVERVIEW AND CLOSING REMARKS

Methylomicrobium buryatense is a promising strain for industrial biocatalysis of methane. It is fast growing, it is highly robust to a range of process conditions, and the optimal growth condition is not susceptible to contamination. We have completed a comprehensive characterization of *M. buryatense* in a bench scale bioreactor under fed-batch and several limiting chemostat conditions. This dataset will serve as a starting point for further strain modification and techno-economic feasibility analysis for industrial processes.

Furthermore, *M. buryatense* was evaluated for a proposed fermentation-like metabolism. It was hypothesized that high concentration of H_2 and organic acids are excreted by some aerobic methanotrophs under low oxygen conditions. This is highly attractive from the perspective of industrial application for the use of culture process conditions to initiate the excretion of methane-derived products. However, after characterizing severely oxygen limited continuous culture of *M. buryatense,* the excretion of organic acids remained variable suggesting a fluctuating NADH and likely nutrient stress. The production of hydrogen was demonstrated to be a byproduct of nitrogen fixation, and the oxygen/methane uptake ratio of 1.3 confirms active respiration. The measured metabolic parameters and collected transcriptomics data will be used as a constraint for a metabolic flux model for further studies and metabolic engineering efforts.

Lastly an exploratory transcriptomics data-analysis was carried out from the numerous RNAseq samples collected from bioreactor experiments. Copper is a key metal in enzymatic oxidation of methane, and this study was focused on gene expression response to available copper concentration. A machine learning approach was used to cluster co-regulated gene expression during a copper induced metabolic time-course transition experiment. The result of

this study reveals two clusters of genes with unusually high expression during the time-course dataset. Most of the identified genes are very highly expressed prior to the metabolic switch following a gradual decline in expression levels post transition. Interestingly, several genes within the identified cluster share homology with a type II secretion system, an auto-transporter family protein, and a TonB dependent transporter. These clues taken together are suggestive of the identification of a novel Cu scavenging system. Further tests are still needed to confirm this hypothesis, however, if proven true, this discovery will be a step closer toward understanding the complete system involved in microbial methane oxidation and a step closer to implementing methanotrophy for commercial biocatalysis of methane.

BIBLIOGRAPHY

1. Clomburg JM, Crumbley AM, Gonzalez R: **Industrial biomanufacturing: The future of chemical production.** *Science* 2017, **355**.
2. Kalyuzhnaya MG, Puri AW, Lidstrom ME: **Metabolic engineering in methanotrophic bacteria.** *Metab Eng* 2015, **29:**142-152.
3. Elvidge CD, Zhizhin M, Baugh K, Hsu FC, Ghosh T: **Methods for Global Survey of Natural Gas Flaring from Visible Infrared Imaging Radiometer Suite Data.** *Energies* 2016, **9**.
4. **Canada's Liquefied Natural Gas Opportunity**
5. **Technically Recoverable Shale Oil and Shale Gas Resources: An Assessment of 137 Shale Formations in 41 Countries Outside the United States.** U.S. Energy Information Administration 2013.
6. Sirajuddin S, Rosenzweig AC: **Enzymatic oxidation of methane.** *Biochemistry* 2015, **54:**2283-2294.
7. Kaluzhnaya M, Khmelenina V, Eshinimaev B, Suzina N, Nikitin D, Solonin A, Lin J-L, McDonald I, Murrell C, Trotsenko Y: **Taxonomic characterization of new alkaliphilic and alkalitolerant methanotrophs from soda lakes of the Southeastern Transbaikal region and description of Methylomicrobium buryatense sp. nov.** *Systematic and applied microbiology* 2001, **24:**166-176.
8. Semrau JD, DiSpirito AA, Yoon S: **Methanotrophs and copper.** *FEMS Microbiology Reviews* 2010, **34:**496-531.
9. Hanson RS, Hanson TE: **Methanotrophic bacteria.** *Microbiol Rev* 1996, **60:**439-471.
10. Semrau JD, DiSpirito AA, Yoon S: **Methanotrophs and copper.** *FEMS Microbiol Rev* 2010, **34:**496-531.
11. Peltola P, Priha P, Laakso S: **Effect of copper on membrane lipids and on methane monooxygenase activity of Methylococcus capsulatus (Bath).** *Archives of microbiology* 1993, **159:**521-525.
12. Jahnke LL, Nichols PD: **Methyl sterol and cyclopropane fatty acid composition of Methylococcus capsulatus grown at low oxygen tensions.** *Journal of Bacteriology* 1986, **167:**238-242.
13. Fang J, Barcelona MJ, Semrau JD: **Characterization of methanotrophic bacteria on the basis of intact phospholipid profiles.** *FEMS Microbiology Letters* 2000, **189:**67-72.
14. Lieberman RL, Rosenzweig AC: **The quest for the particulate methane monooxygenase active site.** *Dalton Transactions* 2005**:**3390-3396.
15. Eshinimaev BT, Khmelenina VN, Sakharovskii VG, Suzina NE, Trotsenko YA: **Physiological, Biochemical, and Cytological Characteristics of a Haloalkalitolerant Methanotroph Grown on Methanol.** *Microbiology* 2002, **71:**512-518.
16. Kalyuzhnaya MG, Yang S, Rozova ON, Smalley NE, Clubb J, Lamb A, Gowda GA, Raftery D, Fu Y, Bringel F, et al: **Highly efficient methane biocatalysis revealed in a methanotrophic bacterium.** *Nat Commun* 2013, **4:**2785.
17. Leak DJ, Dalton H: **Growth yields of methanotrophs.** *Applied Microbiology and Biotechnology* 1986, **23:**470-476.

18. Morinaga Y, Yamanaka S, Yoshimura M, Takinami K, Hirose Y: **Methane metabolism of the obligate methane-utilizing bacterium, Methylomonas flagellata, in methane-limited and oxygen-limited chemostat culture.** *Agricultural and Biological Chemistry* 1979, **43:**2453-2458.
19. Harwood JH, Pirt SJ: **Quantitative Aspects of Growth of the Methane Oxidizing Bacterium Methylococcus capsulatus on Methane in Shake Flask and Continuous Chemostat Culture.** *Journal of Applied Bacteriology* 1972, **35:**597-607.
20. Sheehan BT, Johnson MJ: **Production of Bacterial Cells from Methane.** *Applied Microbiology* 1971, **21:**511-515.
21. de la Torre A, Metivier A, Chu F, Laurens LM, Beck DA, Pienkos PT, Lidstrom ME, Kalyuzhnaya MG: **Genome-scale metabolic reconstructions and theoretical investigation of methane conversion in Methylomicrobium buryatense strain 5G(B1).** *Microb Cell Fact* 2015, **14:**188.
22. Torre Adl, Metivier A, Chu F, Laurens L, Beck DAC, Pienkos PT, Lidstrom ME, Kalyuzhnaya MG: **Genome-scale metabolic reconstructions and theoretical investigation of methane conversion in methylomicrobium buryatense strain 5GB1.** *Microbial Cell Factories* 2015, **xxx:**xxx.
23. Kao WC, Chen YR, Yi EC, Lee H, Tian Q, Wu KM, Tsai SF, Yu SS, Chen YJ, Aebersold R, Chan SI: **Quantitative proteomic analysis of metabolic regulation by copper ions in Methylococcus capsulatus (Bath).** *J Biol Chem* 2004, **279:**51554-51560.
24. El Ghazouani A, Basle A, Firbank SJ, Knapp CW, Gray J, Graham DW, Dennison C: **Copper-Binding Properties and Structures of Methanobactins from Methylosinus trichosporium OB3b.** *Inorganic Chemistry* 2011, **50:**1378-1391.
25. Puri AW, Owen S, Chu F, Chavkin T, Beck DAC, Kalyuzhnaya MG, Lidstrom ME: **Genetic tools for the industrially promising methanotroph Methylomicrobium buryatense.** *Applied and Environmental Microbiology* 2014**:**AEM. 03795-03714.
26. Han B, Su T, Wu H, Gou Z, Xing X-H, Jiang H, Chen Y, Li X, Murrell JC: **Paraffin oil as a "methane vector" for rapid and high cell density cultivation of Methylosinus trichosporium OB3b.** *Applied Microbiology and Biotechnology* 2009, **83:**669-677.
27. Auman AJ, Stolyar S, Costello AM, Lidstrom ME: **Molecular characterization of methanotrophic isolates from freshwater lake sediment.** *Applied and Environmental Microbiology* 2000, **66:**5259-5266.
28. Kits KD, Klotz MG, Stein LY: **Methane oxidation coupled to nitrate reduction under hypoxia by the Gammaproteobacterium Methylomonas denitrificans, sp. nov. type strain FJG1.** *Environ Microbiol* 2015, **17:**3219-3232.
29. Blanchette CD, Knipe JM, Stolaroff JK, DeOtte JR, Oakdale JS, Maiti A, Lenhardt JM, Sirajuddin S, Rosenzweig AC, Baker SE: **Printable enzyme-embedded materials for methane to methanol conversion.** *Nat Commun* 2016, **7:**11900.
30. Yan X, Chu F, Puri AW, Fu Y, Lidstrom ME: **Electroporation-Based Genetic Manipulation in Type I Methanotrophs.** *Appl Environ Microbiol* 2016, **82:**2062-2069.
31. Gilman A, Laurens LM, Puri AW, Chu F, Pienkos PT, Lidstrom ME: **Bioreactor performance parameters for an industrially-promising methanotroph Methylomicrobium buryatense 5GB1.** *Microb Cell Fact* 2015, **14:**182.
32. Kalyuzhnaya MG, Yang S, Rozova ON, Smalley NE, Clubb J, Lamb A, Gowda GAN, Raftery D, Fu Y, Bringel F: **Highly efficient methane biocatalysis revealed in a methanotrophic bacterium.** *Nature communications* 2013, **4**.

33. Kao WC, Wang VC, Huang YC, Yu SS, Chang TC, Chan SI: **Isolation, purification and characterization of hemerythrin from Methylococcus capsulatus (Bath).** *J Inorg Biochem* 2008, **102:**1607-1614.
34. Chen KH, Wu HH, Ke SF, Rao YT, Tu CM, Chen YP, Kuei KH, Chen YS, Wang VC, Kao WC, Chan SI: **Bacteriohemerythrin bolsters the activity of the particulate methane monooxygenase (pMMO) in Methylococcus capsulatus (Bath).** *J Inorg Biochem* 2012, **111:**10-17.
35. Das D, Veziroğlu TN: **Hydrogen production by biological processes: a survey of literature.** *International Journal of Hydrogen Energy* 2001, **26:**13-28.
36. Tanisho S, Kuromoto M, Kadokura N: **Effect of CO2 removal on hydrogen production by fermentation.** *International Journal of Hydrogen Energy* 1998, **23:**559-563.
37. Dixon R, Kahn D: **Genetic regulation of biological nitrogen fixation.** *Nat Rev Micro* 2004, **2:**621-631.
38. Guo Y, Guo G, Mao X, Zhang W, Xiao J, Tong W, Liu T, Xiao B, Liu X, Feng Y, Zou Q: **Functional identification of HugZ, a heme oxygenase from Helicobacter pylori.** *BMC Microbiol* 2008, **8:**226.
39. Dedysh SN, Knief C, Dunfield PF: **Methylocella species are facultatively methanotrophic.** *J Bacteriol* 2005, **187:**4665-4670.
40. Larsen O, Karlsen OA: **Transcriptomic profiling of Methylococcus capsulatus (Bath) during growth with two different methane monooxygenases.** *Microbiologyopen* 2016, **5:**254-267.
41. Murrell JC, McDonald IR, Gilbert B: **Regulation of expression of methane monooxygenases by copper ions.** *Trends Microbiol* 2000, **8:**221-225.
42. Nielsen AK, Gerdes K, Degn H, Murrell JC: **Regulation of bacterial methane oxidation: transcription of the soluble methane mono-oxygenase operon of Methylococcus capsulatus (Bath) is repressed by copper ions.** *Microbiology* 1996, **142 (Pt 5):**1289-1296.
43. Nielsen AK, Gerdes K, Murrell JC: **Copper-dependent reciprocal transcriptional regulation of methane monooxygenase genes in Methylococcus capsulatus and Methylosinus trichosporium.** *Mol Microbiol* 1997, **25:**399-409.
44. Stafford GP, Scanlan J, McDonald IR, Murrell JC: **rpoN, mmoR and mmoG, genes involved in regulating the expression of soluble methane monooxygenase in Methylosinus trichosporium OB3b.** *Microbiology* 2003, **149:**1771-1784.
45. Schill N, van Gulik WM, Voisard D, von Stockar U: **Continuous cultures limited by a gaseous substrate: Development of a simple, unstructured mathematical model and experimental verification with Methanobacterium thermoautotrophicum.** *Biotechnol Bioeng* 1996, **51:**645-658.
46. Li H, Durbin R: **Fast and accurate short read alignment with Burrows-Wheeler transform.** *Bioinformatics* 2009, **25:**1754-1760.
47. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Proc GPD: **The Sequence Alignment/Map format and SAMtools.** *Bioinformatics* 2009, **25:**2078-2079.
48. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, Szczesniak MW, Gaffney DJ, Elo LL, Zhang X, Mortazavi A: **A survey of best practices for RNA-seq data analysis.** *Genome Biol* 2016, **17:**13.

49. Castro-Melchor M, Le H, Hu WS: **Transcriptome data analysis for cell culture processes.** *Adv Biochem Eng Biotechnol* 2012, **127:**27-70.
50. Spies D, Ciaudo C: **Dynamics in Transcriptomics: Advancements in RNA-seq Time Course and Downstream Analysis.** *Comput Struct Biotechnol J* 2015, **13:**469-477.
51. Nueda MJ, Tarazona S, Conesa A: **Next maSigPro: updating maSigPro bioconductor package for RNA-seq time series.** *Bioinformatics* 2014, **30:**2598-2602.
52. Anders S, Huber W: **Differential expression analysis for sequence count data.** *Genome Biol* 2010, **11:**R106.
53. Jo K, Kwon HB, Kim S: **Time-series RNA-seq analysis package (TRAP) and its application to the analysis of rice, Oryza sativa L. ssp. Japonica, upon drought stress.** *Methods* 2014, **67:**364-372.
54. Wise A, Bar-Joseph Z: **SMARTS: reconstructing disease response networks from multiple individuals using time series gene expression data.** *Bioinformatics* 2015, **31:**1250-1257.
55. Jain AK: **Data clustering: 50 years beyond K-means.** *Pattern Recognition Letters* 2010, **31:**651-666.
56. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ: **The Phyre2 web portal for protein modeling, prediction and analysis.** *Nat Protoc* 2015, **10:**845-858.
57. Noinaj N, Guillier M, Barnard TJ, Buchanan SK: **TonB-dependent transporters: regulation, structure, and function.** *Annu Rev Microbiol* 2010, **64:**43-60.
58. DiSpirito AA, Semrau JD, Murrell JC, Gallagher WH, Dennison C, Vuilleumier S: **Methanobactin and the Link between Copper and Bacterial Methane Oxidation.** *Microbiology and Molecular Biology Reviews* 2016, **80:**387-409.
59. Dassama LMK, Kenney GE, Ro SY, Zielazinski EL, Rosenzweig AC: **Methanobactin transport machinery.** *Proceedings of the National Academy of Sciences of the United States of America* 2016, **113:**13027-13032.
60. Zahn JA, DiSpirito AA: **Membrane-associated methane monooxygenase from Methylococcus capsulatus (Bath) (vol 178, pg 1018, 1996).** *Journal of Bacteriology* 1996, **178:**2726-2726.
61. Kenney GE, Rosenzweig AC: **Genome mining for methanobactins.** *BMC Biol* 2013, **11:**17.

APPENDIX A: BIOREACTOR GROWTH PROFILE

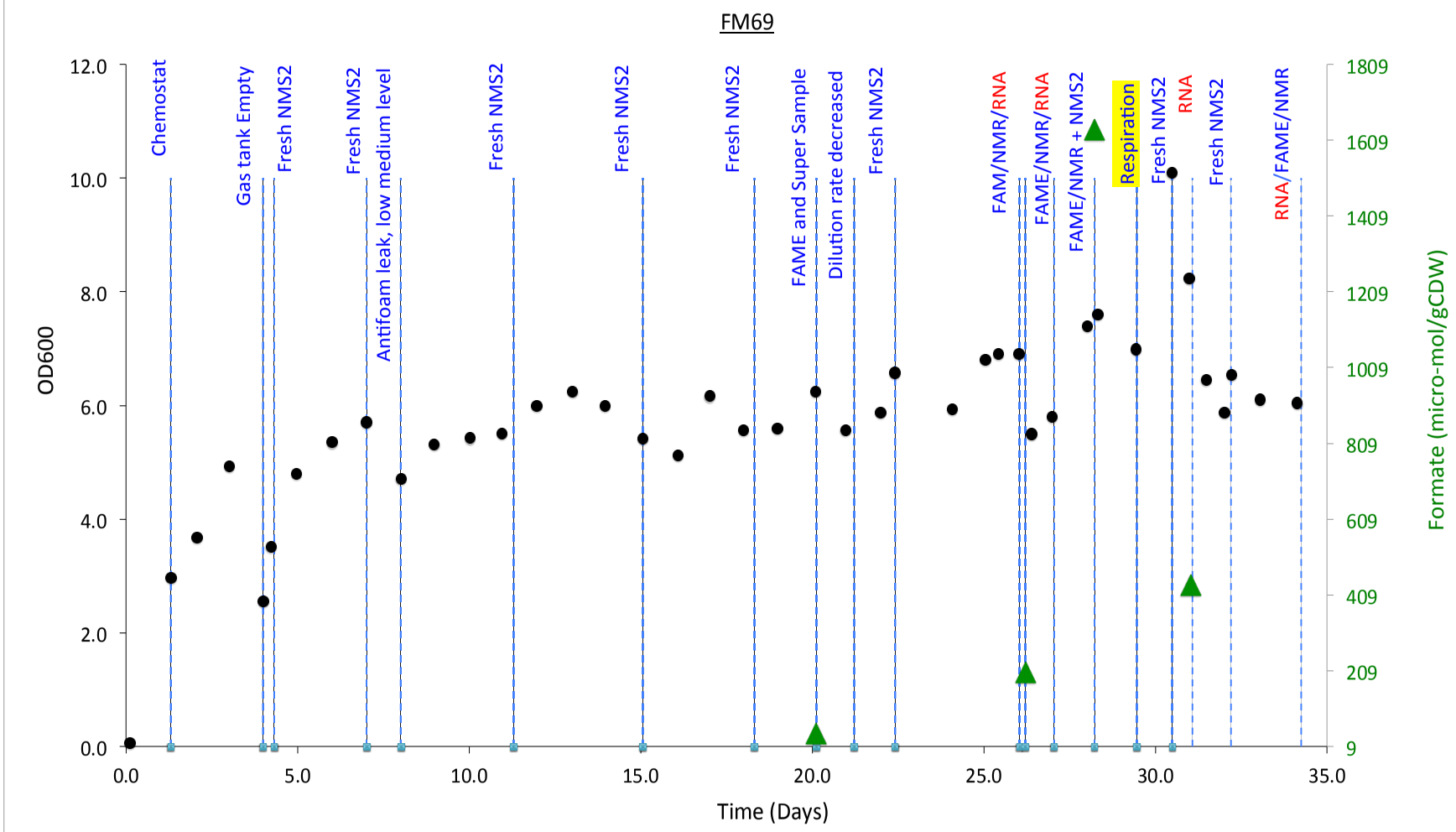

Figure A.1. FM69 Bioreactor growth profile: excreted formate concentration during switch from low O_2 to high O_2 condition.

Excreted formate concentration appears to fluctuate between the few sampled time points.

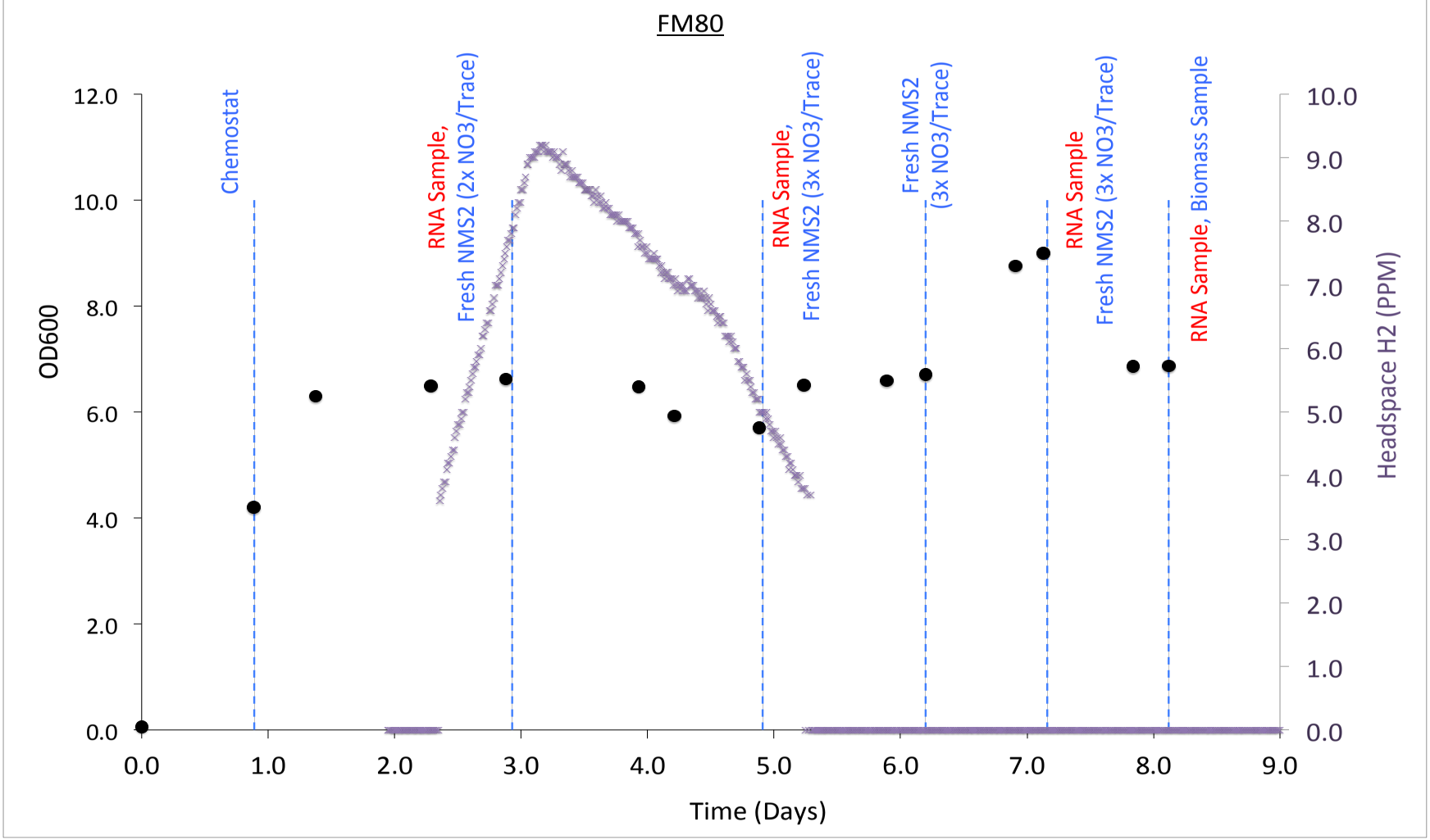

Figure A.2. FM80 Bioreactor growth profile: headspace H_2 concentration in a fermentation experiment with stepwise increase in nutrient concentration.

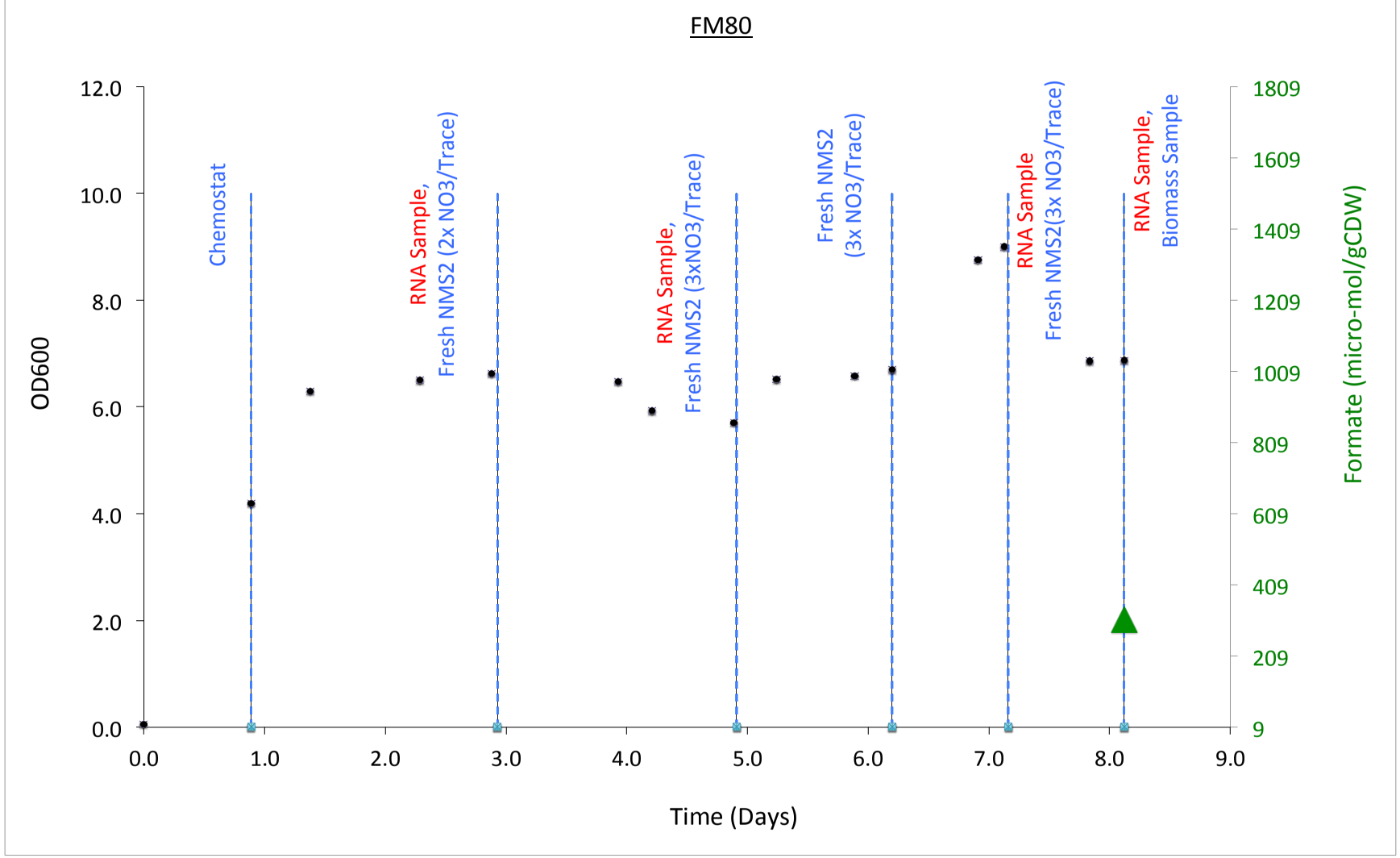

Figure A.3. FM80 Bioreactor growth profile: excreted formate concentration in fermentation experiment with stepwise increase in nutrient concentration.

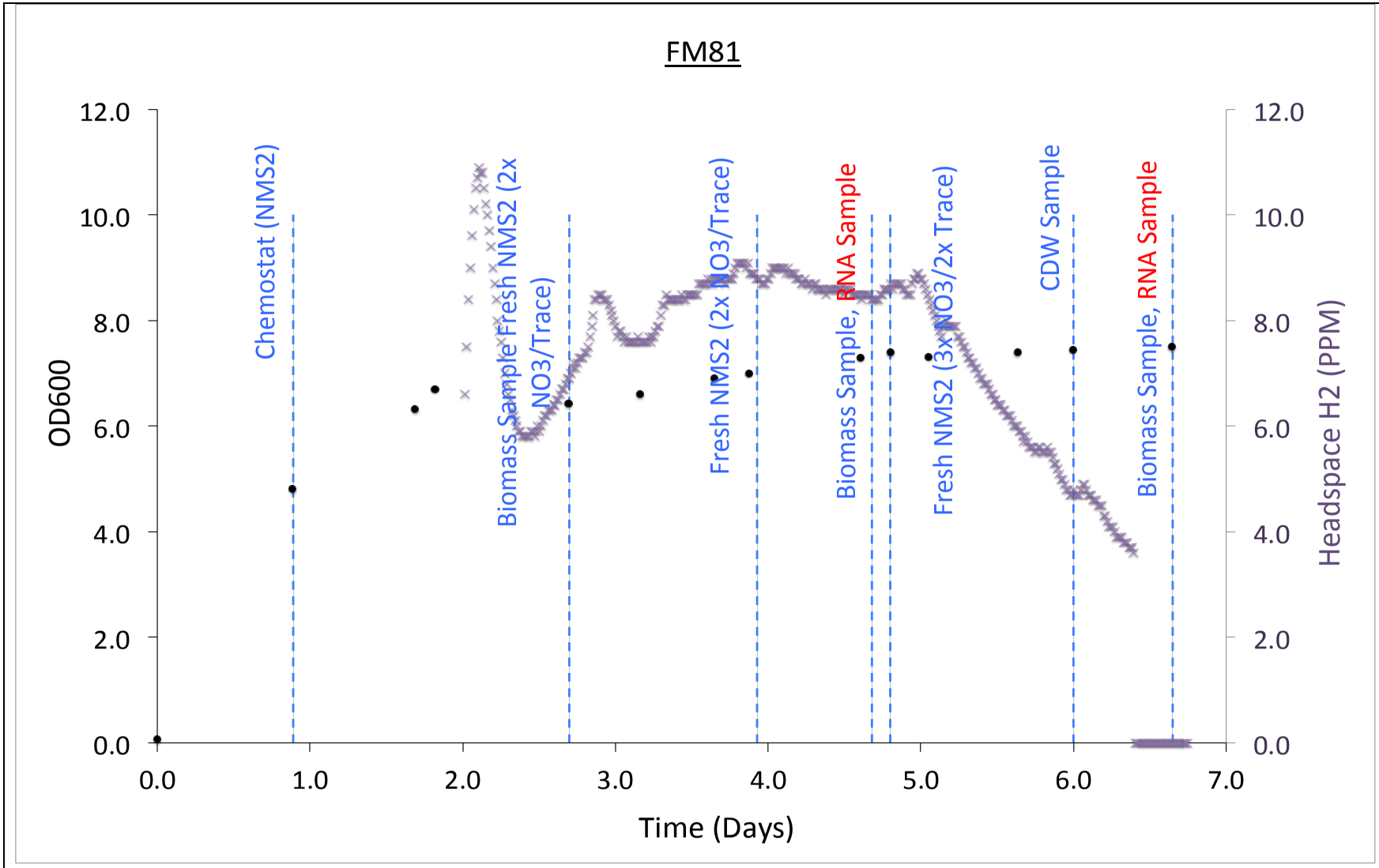

Figure A.4. FM81 Bioreactor growth profile: headspace H_2 concentration in a fermentation experiment with stepwise increase in nutrient concentration.

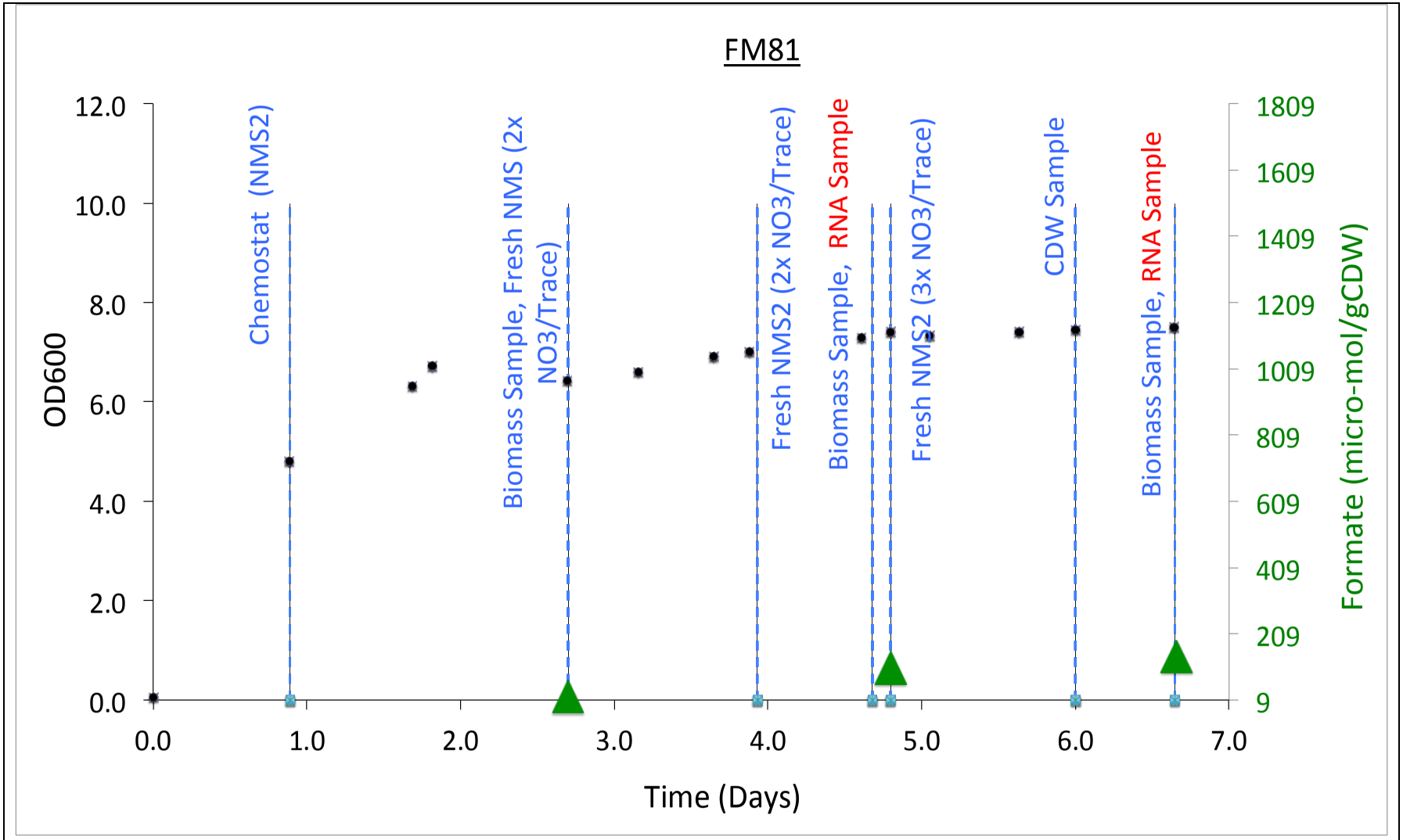

Figure A.5. FM81 Bioreactor growth profile: excreted formate concentration in fermentation experiment with stepwise increase in nutrient concentration.

APPENDIX B: CHEMOSTAT THEORY FOR GAS SUBSTRATE LIMITATION

A review of chemostat theory for growth on limiting gas substrates as published in [45].

Chemostat growth where the culture is limited by a gas substrate has several properties that are drastically different from culture that is limited by a nutrient in the inlet medium stream. The primary difference arises from the delivery of the limiting gas substrate. The gas that is sparged into a CSTR is completely independent from the medium flow rate and is decoupled from the dilution rate/growth rate of the culture. When a chemostat culture is limited for a gas substrate: (i) the gas uptake rate remained constant over a large range of dilution rates; and (ii) the biomass concentration decreases with increasing dilution rate. Gas conversion remains the same (with the same gas feed) because increased dilution rate does not increase dissolved gas concentration in the medium, and the limiting gas substrate remains at the same low level as during slow dilution. Therefore, the culture responds with a lower cell density with increased dilution rate because less biomass can be supported by a constant limiting substrate when the culture is at faster growth rate.

The gas-substrate limiting culture is governed by substrate saturation kinetics similar to Monod growth equation:

$$r_G = \left[q_G^{max} \cdot \frac{C_G}{C_G + K_G} \right] X$$

Where the rate of gas uptake [g-gas L^{-1} hr^{-1}] is proportional to the maximum specific gas uptake of the organism [g-gas g-biomass^{-1} hr^{-1}], the concentration of gas in the medium (with saturation constant, $K_{G)}$, and cell density. Analogous to Monod model:

$$\mu = \mu^{max} \frac{[S]}{K_S + [S]}$$

As well as Michaelis-Menten kinetics:

$$\nu = \nu^{max} \frac{[S]}{K_M + [S]}$$

The mass transfer of gas substrate into the liquid phase is the rate limiting step and is proportional to the gradient (or driving force) of gas concentration in the medium, with the difference of gas saturation concentration at the gas bubble interface and the gas concentration in bulk medium:

$$GTR = K_L a (C_G^* - C_G)$$

Where $K_L a$ is the volumetric mass transfer coefficient [hr^{-1}]. Mass balance for gas substrate and biomass, where it is assumed that the dissolved gas leaving in the medium stream is negligible compared to the gassing rate:

$$\frac{dC_G}{dt} = GTR - r_G - \cancel{DC_G}$$

$$\frac{dX}{dt} = r_X - DX$$

Where the rate of biomass production r_x is the product of the rate of gas uptake r_G and the biomass yield coefficient $Y_{X/G}$. Schill et al. subtract a maintenance term υ_m from the total rate of gas uptake in order to account for substrate that is spent on maintenance rather than biomass production.

$$r_x = Y_{x/G} \cdot (r_G - \nu_m)$$

However, for the chemostat experiments presented here, the yield coefficient (or carbon conversion efficiency) is calculated using the gas uptake rate data and a cell dry weight measurement. Therefore, the quantity of substrate used for maintenance is already accounted for in the yield coefficient and the maintenance term is not necessary. The combined equations that model biomass concentration X as a function of D with the maintenance term omitted are:

$$\boxed{X = \frac{Y_{X/G} K_L a (C_G^* - C_G)}{D}}$$

Where,

$$C_G = \frac{Y_{X/G} K_G}{Y_{X/G} q^{max} - D}$$

The same equations with maintenance term as described in Schill et al.

$$X = \frac{K_L a \left[C_G^* - \frac{(D + Y_{X/G} \cdot m) K_G}{\mu^{max} - D} \right] Y_{X/G}}{D + Y_{X/G} \cdot m} \qquad \mu^{max} = Y_{X/G}(q_G^{max} - m)$$

In contrast the classical model for chemostat culture cell density as a function of dilution rate is similarly derived as:

$$\boxed{[X] = Y_{X/S}\left([S]_{in} - \frac{DK_s}{\mu_{max} - D} \right)}$$

When comparing the cell density profile as a function of dilution rate for gas and liquid substrates two main differences arise (Schill et al. 1996):

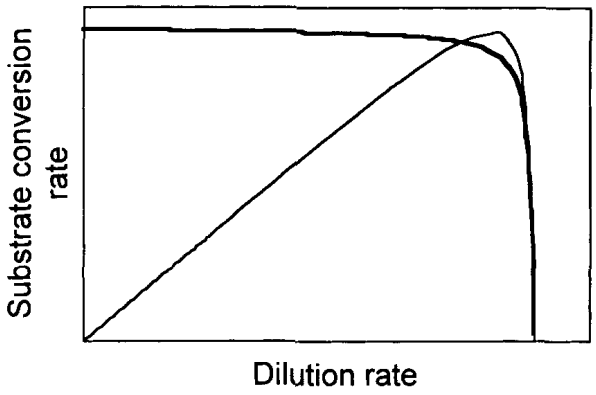

Figure 8. Comparison of the profile of r_D vs. D between a continuous culture limited by a gaseous substrate (thick line) and a continuous culture limited by a liquid substrate (thin line).

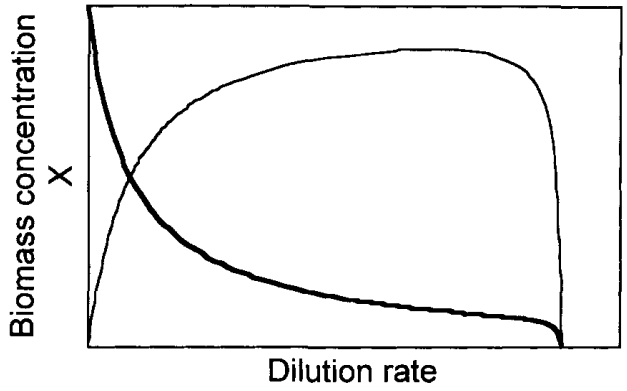

Figure 10. Comparison of the profile of biomass concentration X vs. D between a continuous culture limited by a gaseous substrate (thick line) and a continuous culture limited by a liquid substrate (thin line).

(i) Gas substrate conversion rate remains constant with dilution rate when gas substrate is limiting. (ii) Biomass concentration decreases with increasing dilution rate when the gas substrate is limiting.

Nomenclature:
r_G - uptake rate of gas substrate [g $L^{-1}hr^{-1}$].
q_G^{max} - maximum specific conversion rate gas substrate [g-substrate g-biomass^{-1}hr^{-1}].
C_G - dissolved concentration of gas substrate in bulk medium [g L^{-1}].
X - biomass dry weight concentration [g L^{-1}].
GTR - gas transfer rate [g L^{-1} hr^{-1}].
K_La - volumetric mass transfer coefficient [hr^{-1}].
C_G^* - dissolved gas concentration at saturation [g L^{-1}].
D - dilution rate [hr^{-1}].
r_x - molar biomass generation rate [g L^{-1} hr^{-1}]
$Y_{X/G}$ - molar biomass yield coefficient [g-biomass g-substrate^{-1}]
ν_m - volumetric reaction rate required for maintenance [g L^{-1} hr^{-1}]
m - maintenance coefficient [g-substrate g-biomass^{-1} of X hr^{-1}]

APPENDIX C: GENE EXPRESSION PROFILES FOR ALL CLUSTERS

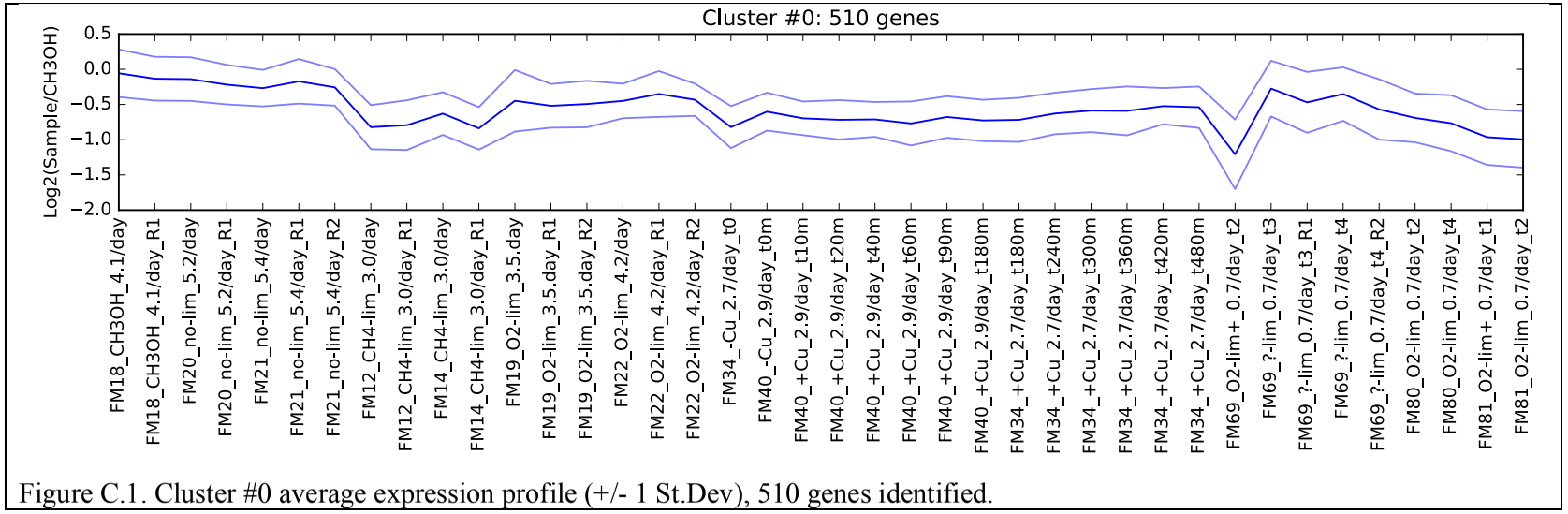

Figure C.1. Cluster #0 average expression profile (+/- 1 St.Dev), 510 genes identified.

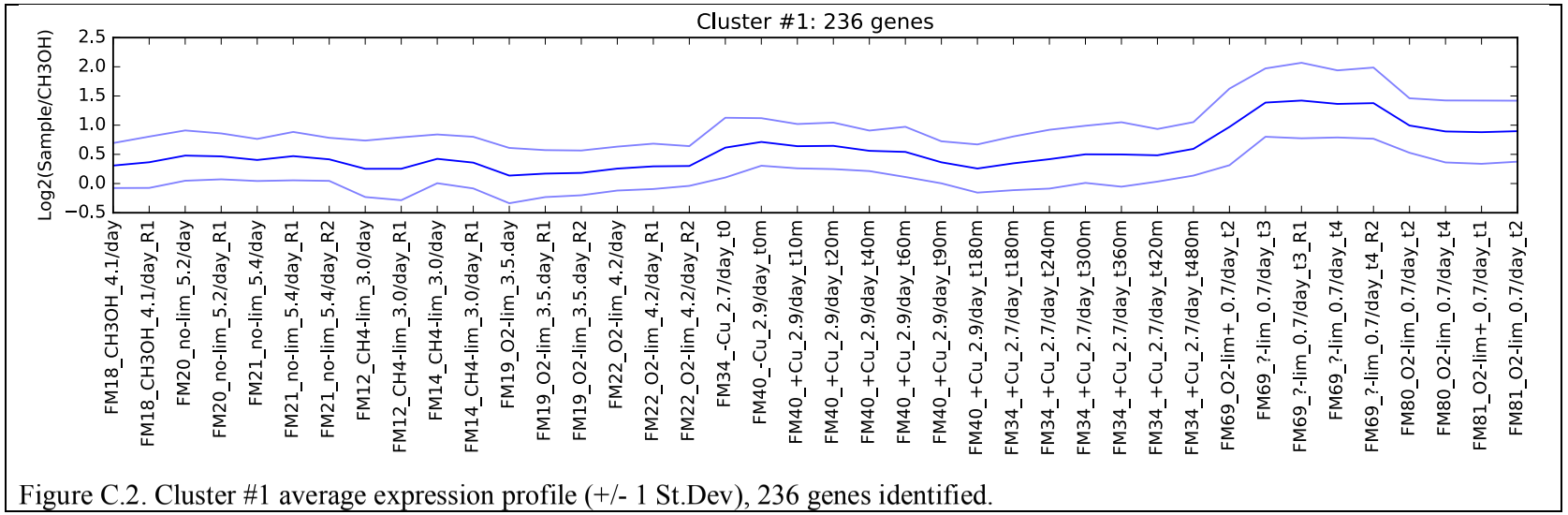

Figure C.2. Cluster #1 average expression profile (+/- 1 St.Dev), 236 genes identified.

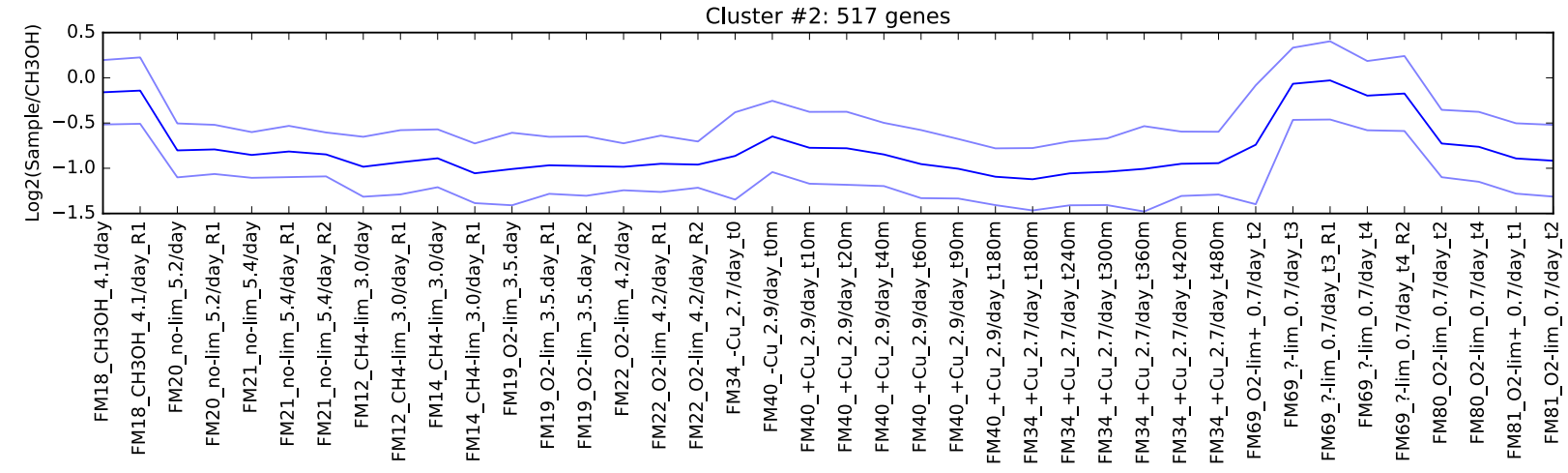

Figure C.3. Cluster #2 average expression profile (+/- 1 St.Dev), 517 genes identified.

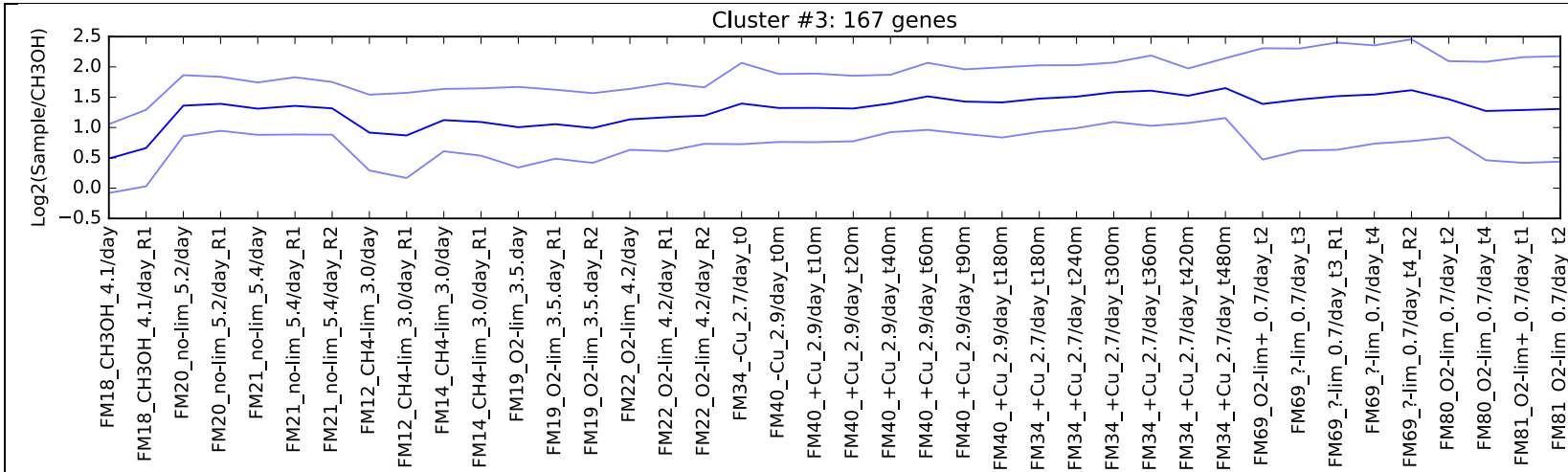

Figure C.4. Cluster #3 average expression profile (+/- 1 St.Dev), 167 genes identified.

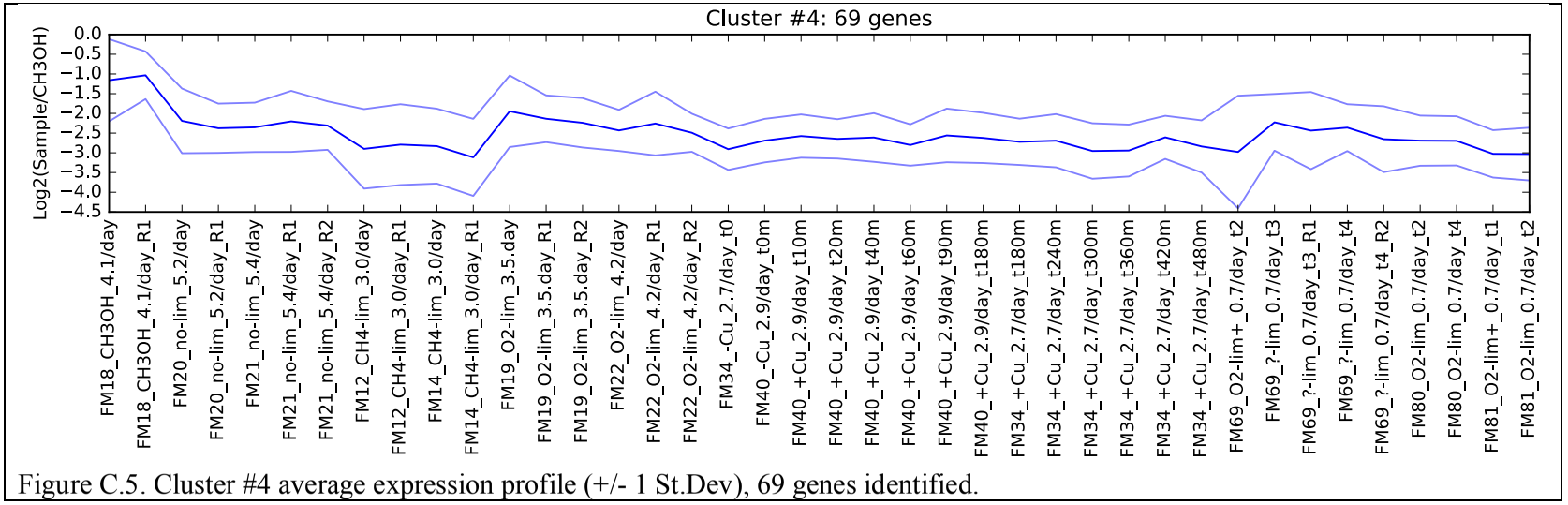

Figure C.5. Cluster #4 average expression profile (+/- 1 St.Dev), 69 genes identified.

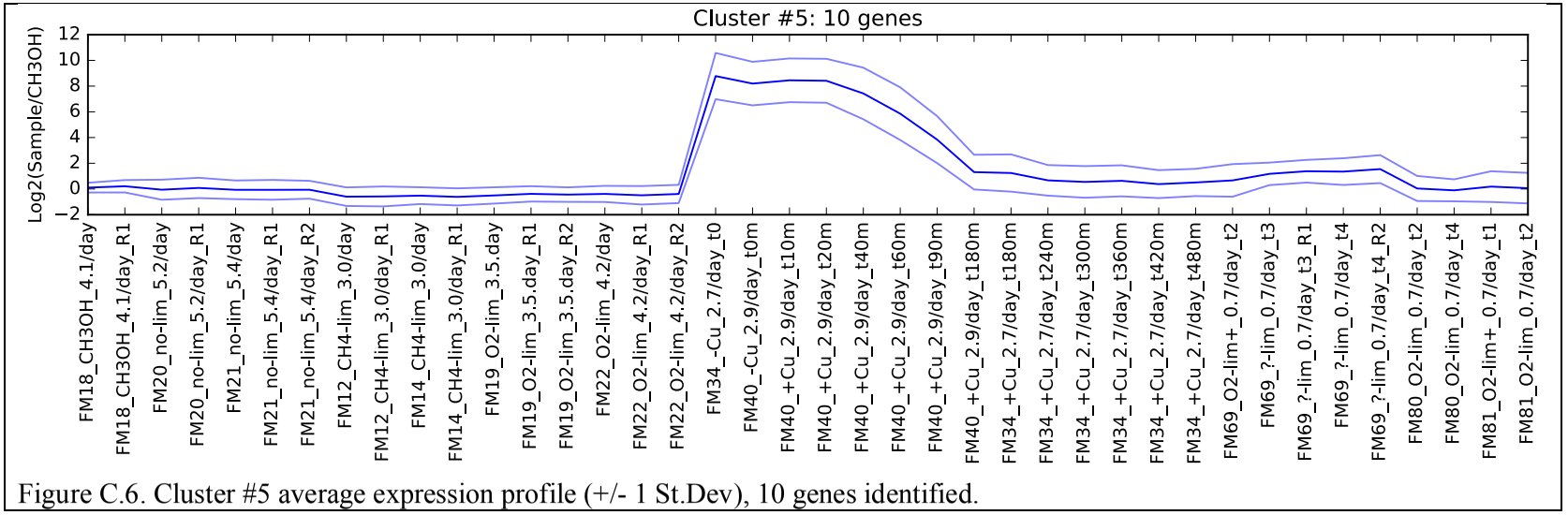

Figure C.6. Cluster #5 average expression profile (+/- 1 St.Dev), 10 genes identified.

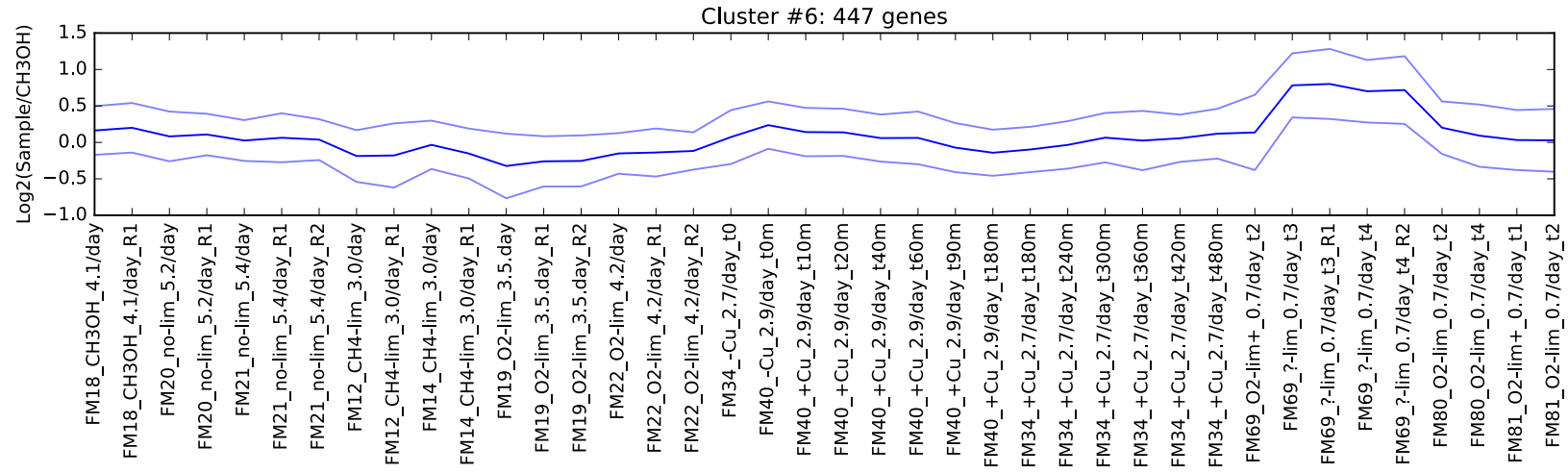

Figure C.7. Cluster #6 average expression profile (+/- 1 St.Dev), 447 genes identified.

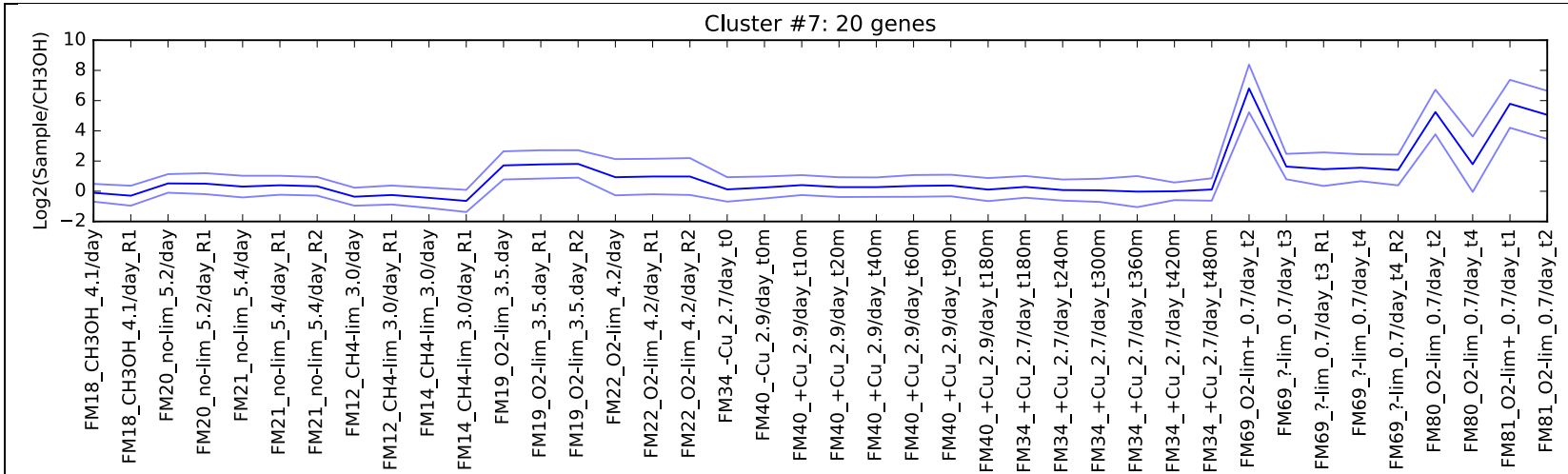

Figure C.8. Cluster #7 average expression profile (+/- 1 St.Dev), 20 genes identified.

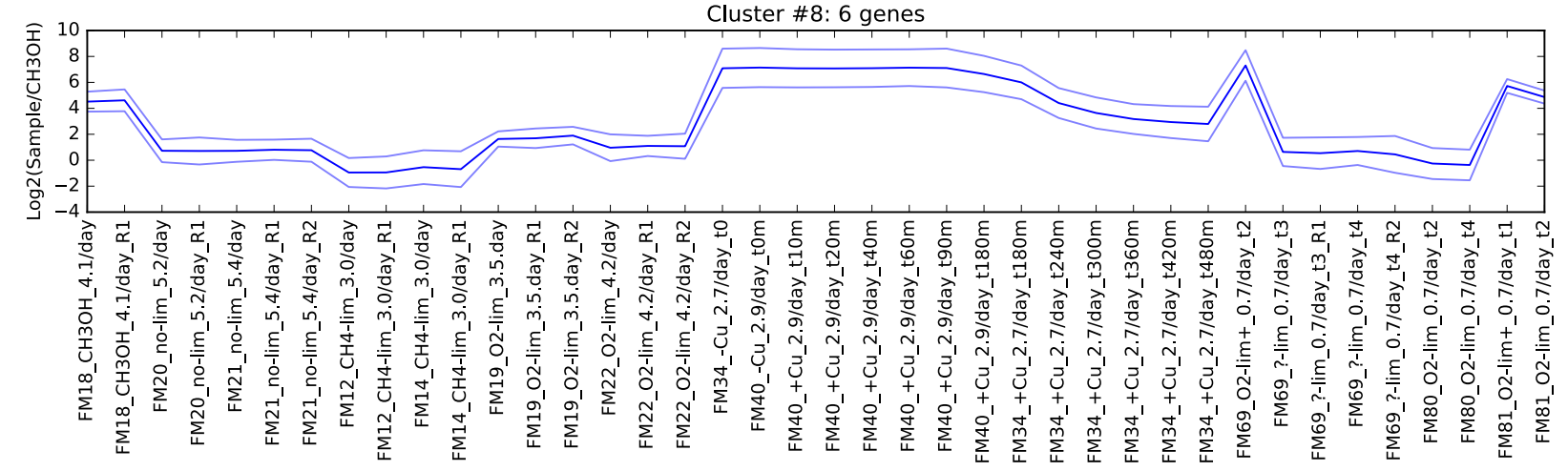

Figure C.9. Cluster #8 average expression profile (+/- 1 St.Dev), 6 genes identified.

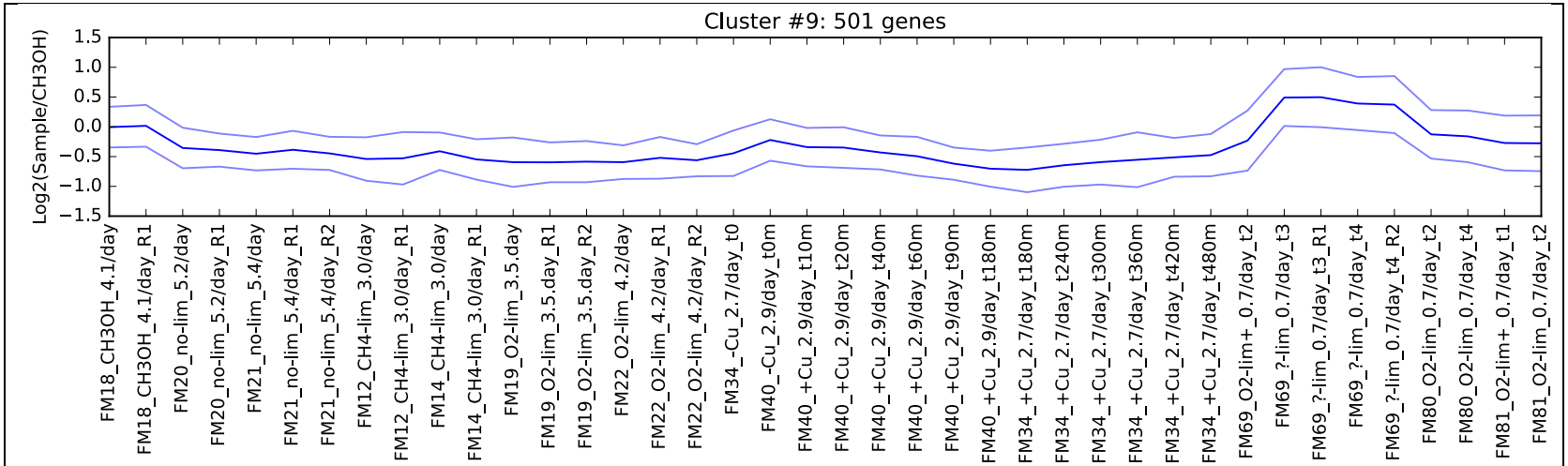

Figure C.10. Cluster #9 average expression profile (+/- 1 St.Dev), 501 genes identified.

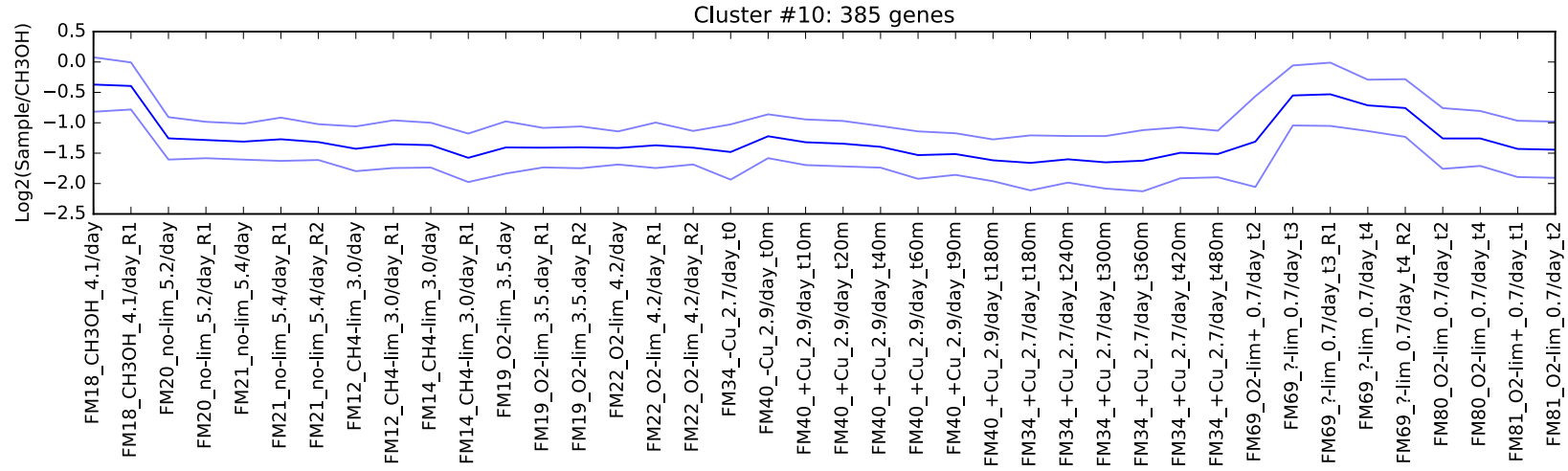

Figure C.11. Cluster #10 average expression profile (+/- 1 St.Dev), 385 genes identified.

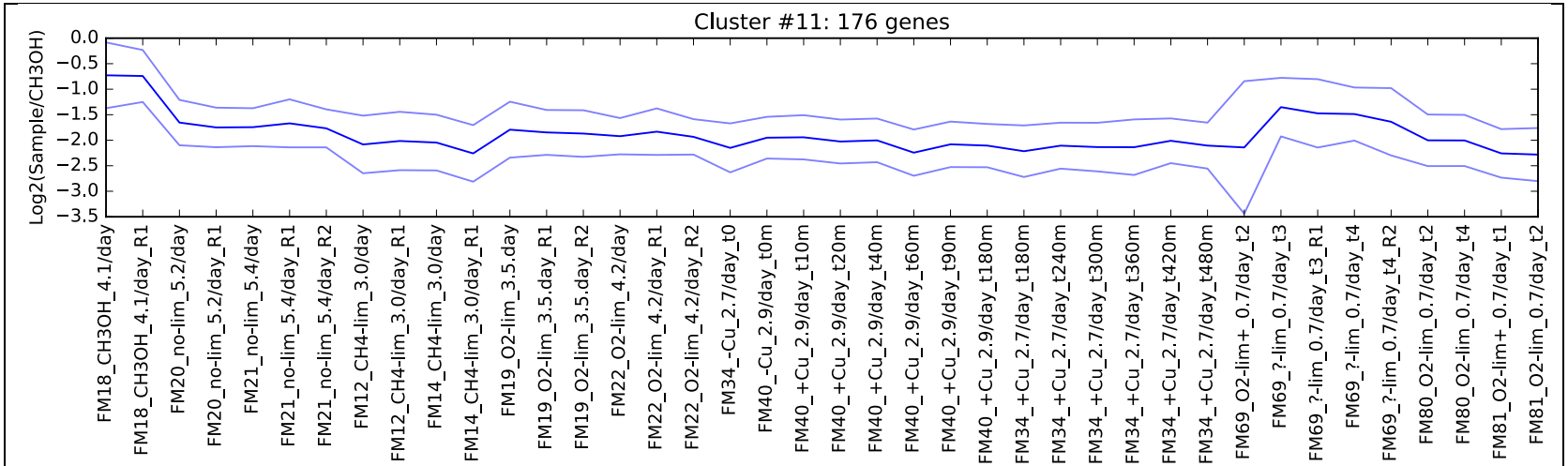

Figure C.12. Cluster #11 average expression profile (+/- 1 St.Dev), 176 genes identified.

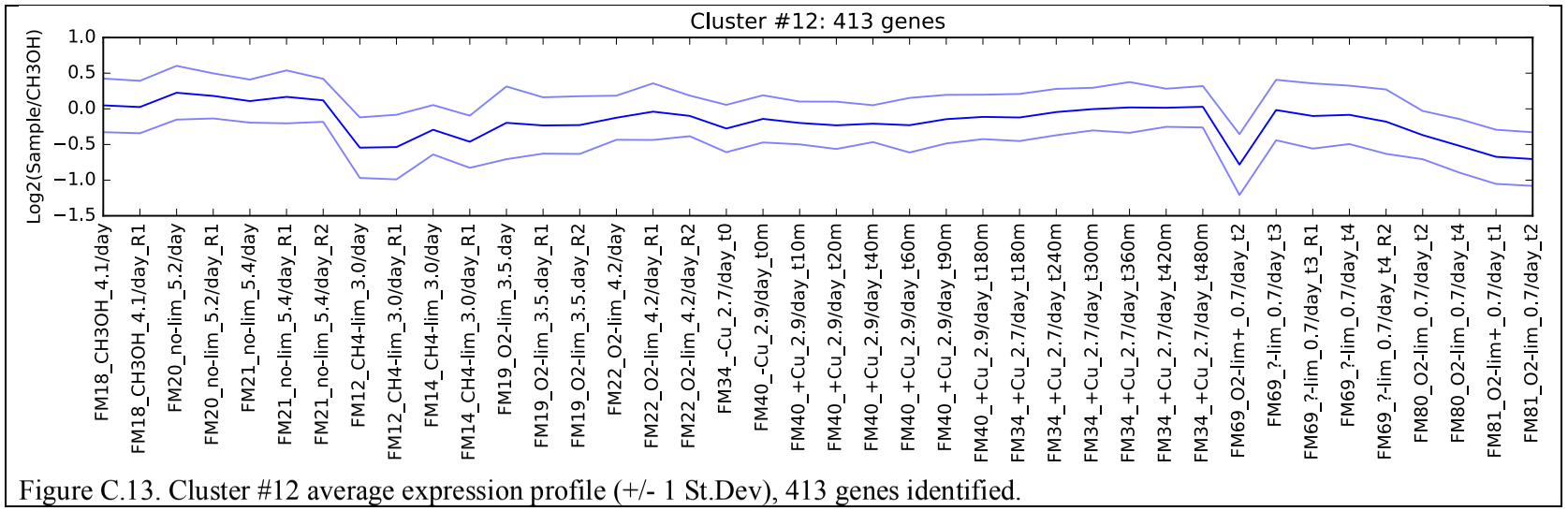

Figure C.13. Cluster #12 average expression profile (+/- 1 St.Dev), 413 genes identified.

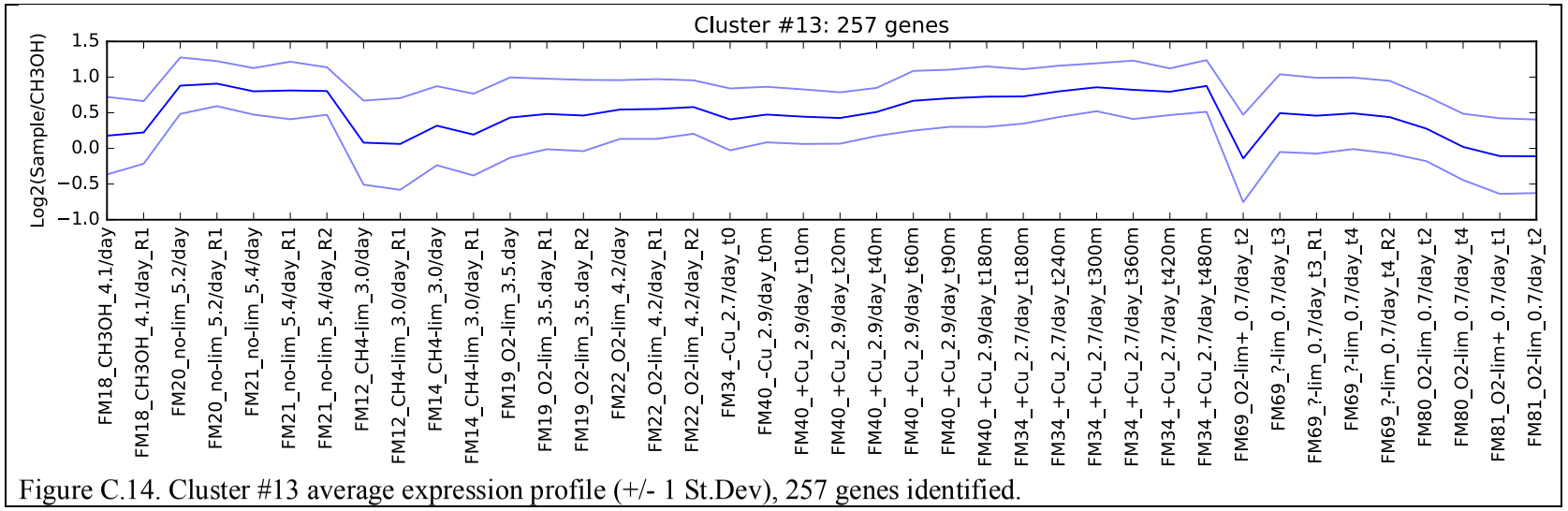

Figure C.14. Cluster #13 average expression profile (+/- 1 St.Dev), 257 genes identified.

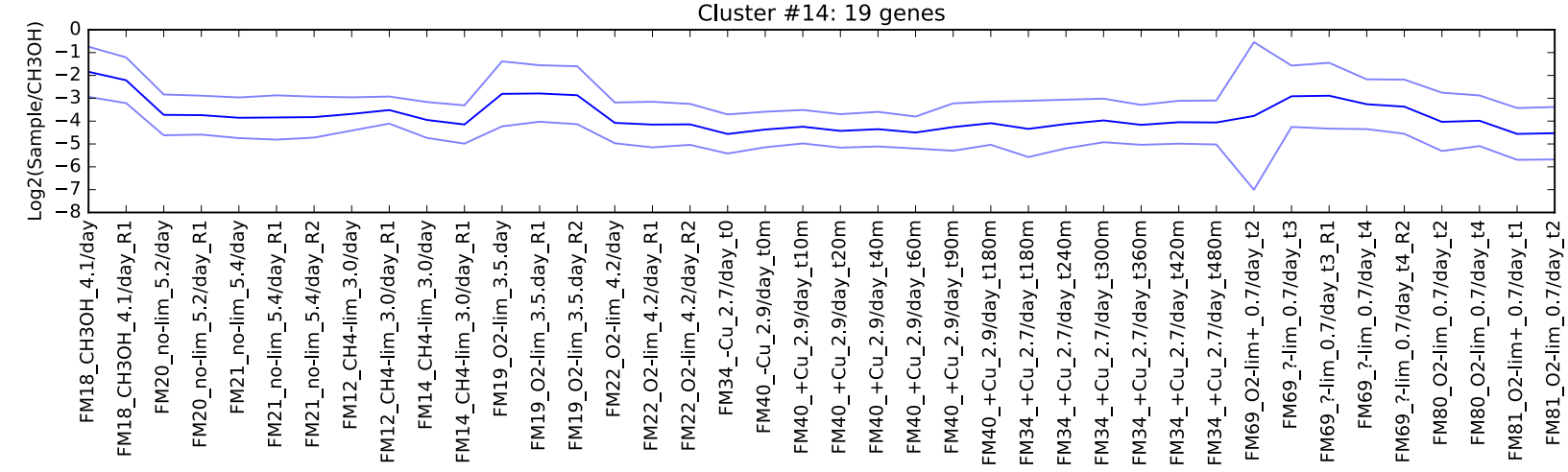

Figure C.15. Cluster #14 average expression profile (+/- 1 St.Dev), 19 genes identified.

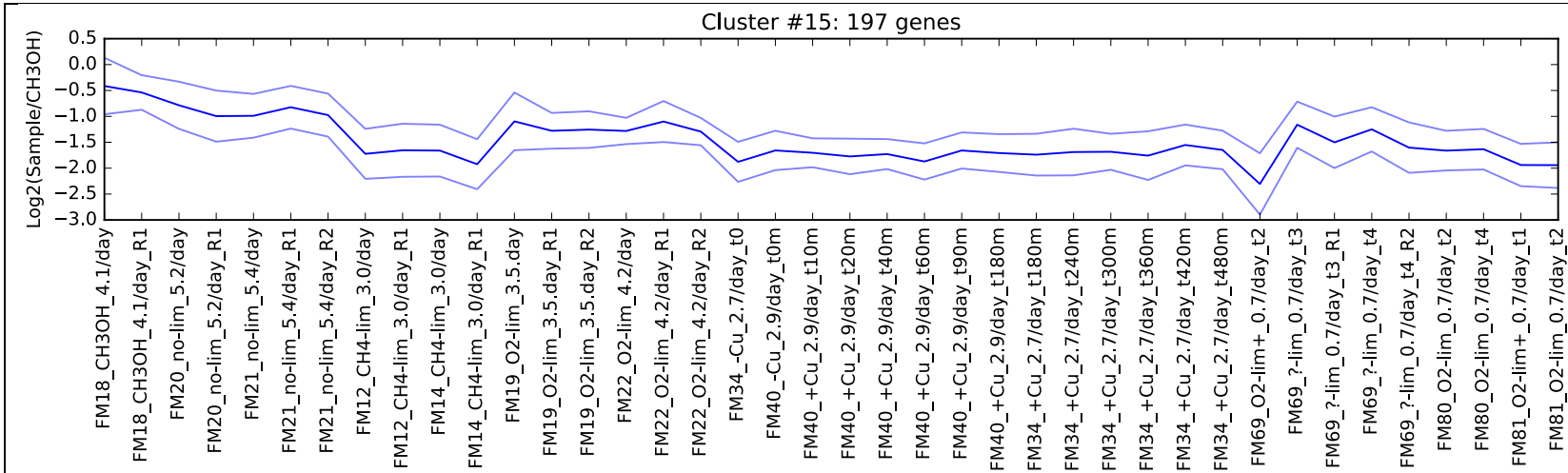

Figure C.16. Cluster #15 average expression profile (+/- 1 St.Dev), 197 genes identified.

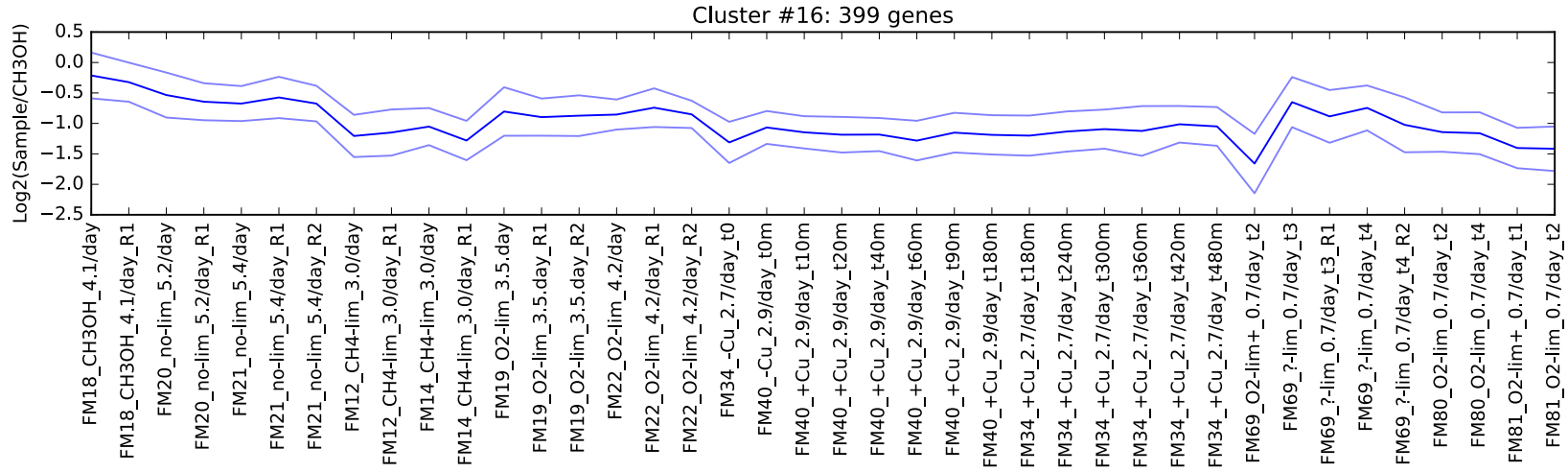

Figure C.17. Cluster #16 average expression profile (+/- 1 St.Dev), 399 genes identified.

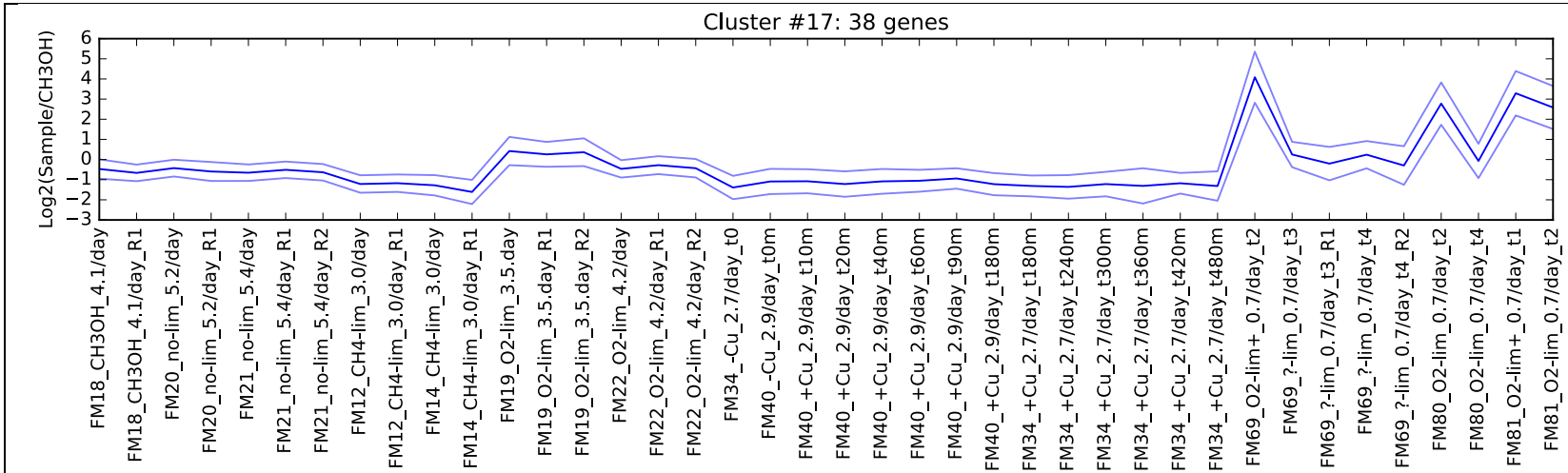

Figure C.18. Cluster #17 average expression profile (+/- 1 St.Dev), 38 genes identified.

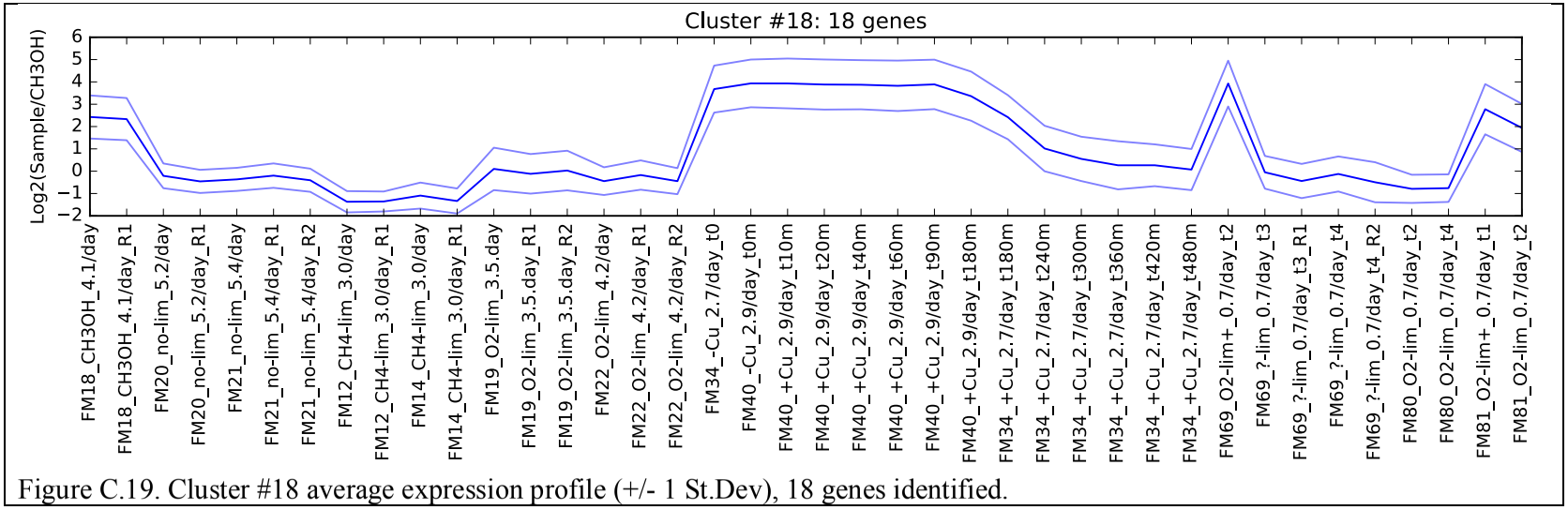

Figure C.19. Cluster #18 average expression profile (+/- 1 St.Dev), 18 genes identified.

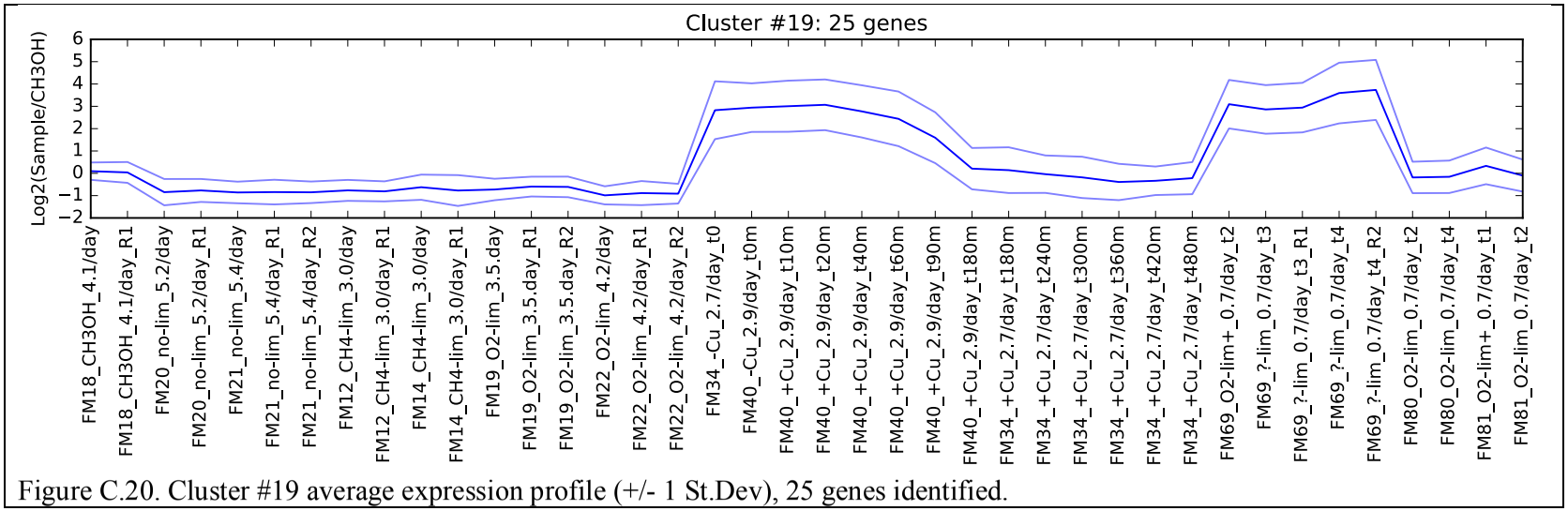

Figure C.20. Cluster #19 average expression profile (+/- 1 St.Dev), 25 genes identified.

VITA

After a series of turbulent relocations, Alexey Gilman has spent his teenage years growing up in Portland, Oregon. He attended Parkrose High School in NE Portland where he first discovered an interest for science in his chemistry and physics classes. Alexey studied chemistry with an option in chemical engineering at Oregon State University and graduated with the distinction of summa cum laude. Alexey worked as a process engineer in Portland for two years prior to attending University of Washington for graduate school.