

3-23-2010

# A Multi-Faceted Diagnostic Approach to Lung Infections in Patients with Cystic Fibrosis

Melissa S. Doud

*Florida International University*, mdoud001@fiu.edu

**DOI:** 10.25148/etd.FI10041609

Follow this and additional works at: <https://digitalcommons.fiu.edu/etd>

---

## Recommended Citation

Doud, Melissa S., "A Multi-Faceted Diagnostic Approach to Lung Infections in Patients with Cystic Fibrosis" (2010). *FIU Electronic Theses and Dissertations*. 166.

<https://digitalcommons.fiu.edu/etd/166>

This work is brought to you for free and open access by the University Graduate School at FIU Digital Commons. It has been accepted for inclusion in FIU Electronic Theses and Dissertations by an authorized administrator of FIU Digital Commons. For more information, please contact [dcc@fiu.edu](mailto:dcc@fiu.edu).

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

A MULTI-FACETED DIAGNOSTIC APPROACH TO LUNG INFECTIONS IN  
PATIENTS WITH CYSTIC FIBROSIS

A dissertation submitted in partial fulfillment of the

requirements for the degree of

DOCTOR OF PHILOSOPHY

in

BIOLOGY

by

Melissa S. Doud

2010

To: Dean Kenneth Furton  
College of Arts and Sciences

This dissertation, written by Melissa S. Doud, and entitled A Multi-Faceted Diagnostic Approach to Lung Infections in Patients with Cystic Fibrosis, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

---

Alejandro Barbieri

---

Kenneth Furton

---

Michael Light

---

John Makemson

---

Kalai Mathee, Major Professor

Date of Defense: March 23, 2010

The dissertation of Melissa S. Doud is approved.

---

Dean Kenneth Furton  
College of Arts and Sciences

---

Interim Dean Kevin O'Shea  
University Graduate School

Florida International University, 2010

## DEDICATION

This dissertation is dedicated to the many cystic fibrosis patients that I have met during my research. Meeting you inspired me to continue my master's project and enroll as a PhD student. May your lives be improved by the information presented in this work. This dissertation is also dedicated to my family. My family has supported me throughout my schooling. They are always there for me. Little to nothing should ever come before family.

## ACKNOWLEDGMENTS

I would like to thank my mentor Kalai Mathee for taking me under her wing and showing me every facet of academia. Her lessons in and out of the lab have greatly affected how I think about science, my future career, and life. I am of course grateful to be a member of the Mathee lab. Members of the lab have come and gone but all of them supported me in my endeavors. I am thankful Lisa Schneper joined the lab. Lisa took over many of my responsibilities and rescued my sanity. Despite her tremendous work load, she was always willing to give research advice. Natalie Maricic has been with me throughout the PhD program. Natalie has been a good friend and a hard working lab manager. I thank Natalie for running my errands, troubleshooting science with me, listening to me vent, and doing non-science things with me. It pains me to think what FIU would have been like without Natalie. I would like to acknowledge Michael Light for allowing me to use his clinic to meet with CF patients, working with me on the IRBs, and signing countless forms. I am thankful for my financial support provided by FIU MBRS-RISE (NIH/NIGMS R25 GM61347), two semesters of teaching assistantship provided by the Department of Biological Sciences, the Cystic Fibrosis Foundation Student Training Grant (DoudOH08) and summer BRI grants. To Rory, thanks for working on my table of contents, driving to Miami, dealing with a disorganized house, and most importantly giving me the support I needed to get through this last semester. I also want to thank my family for loving me.

ABSTRACT OF THE DISSERTATION

A MULTI-FACETED DIAGNOSTIC APPROACH TO LUNG INFECTIONS IN  
PATIENTS WITH CYSTIC FIBROSIS

by

Melissa S. Doud

Florida International University, 2010

Miami, Florida

Professor Kalai Mathee, Major Professor

One in 3,000 people in the US are born with cystic fibrosis (CF), a genetic disorder affecting the reproductive system, pancreas, and lungs. Lung disease caused by chronic bacterial and fungal infections is the leading cause of morbidity and mortality in CF. Identities of the microbes are traditionally determined by culturing followed by phenotypic and biochemical assays. It was first thought that the bacterial infections were caused by a select handful of bacteria such as *S. aureus*, *H. influenzae*, *B. cenocepacia*, and *P. aeruginosa*. With the advent of PCR and molecular techniques, the polymicrobial nature of the CF lung became evident. The CF lung contains numerous bacteria and the communities are diverse and unique to each patient. The total complexity of the bacterial infections is still being determined. In addition, only a few members of the fungal communities have been identified. Much of the fungal community composition is still a mystery. This dissertation addresses this gap in knowledge. A snap shot of CF sputa bacterial community was obtained using the length heterogeneity-PCR community profiling technique. The profiles show that south Florida CF patients have a unique, diverse, and dynamic bacterial community which changes over time. The identities of

the bacteria and fungi present were determined using the state-of-the-art 454 sequencing. Sequencing results show that the CF lung microbiome contains commonly cultured pathogenic bacteria, organisms considered a part of the healthy core biome, and novel organisms. Understanding the dynamic changes of these identified microbes will ultimately lead to better therapeutical interventions. Early detection is key in reducing the lung damage caused by chronic infections. Thus, there is a need for accurate and sensitive diagnostic tests. This issue was addressed by designing a bacterial diagnostic tool targeted towards CF pathogens using SPR. By identifying the organisms associated with the CF lung and understanding their community interactions, patients can receive better treatment and live longer.

## TABLE OF CONTENTS

CHAPTER	PAGE
Chapter 1: Introduction	1
1.1 Preface	2
1.2 Cystic fibrosis	2
1.2.1 CF sputum bacterial flora	5
1.2.2 CF sputum fungal flora	6
1.3 Common assays used for clinical identification of bacteria and their limitations	8
1.3.1 Culturing and commercial phenotypic identification assays	8
1.3.2 Accuracy of commercial identification systems	13
1.3.2.1 Bacterial identification	13
1.3.2.2 Fungal identification	14
1.3.3 Misidentification through culturing	15
1.4 Identification through mass spectrometry	17
1.4.1 Mass spectrometry (MS)	17
1.4.2 MALDI-TOF MS	18
1.4.3 Electrospray ionization (ESI)	22
1.5 Detection through antibody-antigen interactions	25
1.5.1 The antibody response	25
1.5.2 Enzyme-linked immunosorbant assay (ELISA)	27
1.5.3 Surface plasmon resonance (SPR)	31
1.6 Molecular analysis of isolates	32
1.6.1 Molecular markers	32
1.6.2 Ribosomal RNA (rRNA)	33
1.6.3 16S ribosomal RNA (rRNA)	35
1.6.4 Fungal internal transcribed spacer regions (ITS)	37
1.6.5 Polymerase chain reaction (PCR)	38
1.6.6 Multiplex PCR	38
1.7 Molecular tools for community studies	39
1.7.1 Terminal restriction fragment length (T-RFLP) analysis	40
1.7.2 Amplicon length heterogeneity (LH)	42
1.8 Sequencing	46
1.8.1 Sanger sequencing	46
1.8.2 Pyrosequencing	49
1.8.3 454 sequencing	52
1.8.4 Illumina	53
1.8.5 SOLiD	56
1.9 Metagenomic statistics and bioinformatics	58
1.9.1 Statistical analysis based on ecological indices	58
1.9.2 Statistical analysis based on abundance models	59
1.9.3 Comparative metagenomics	61
1.10 Hypothesis and aims	63



Chapter 2: Combination of 16S rRNA Variable Regions Provide a Detailed Analysis of Bacterial Community Dynamics in the Lungs of Cystic Fibrosis Patients	65
2.1 Abstract	66
2.2 Introduction	67
2.3 Material and methods	70
2.3.1 Sample collection	70
2.3.2 Control bacteria isolates	70
2.3.3 LH-PCR	71
2.3.4 LH-PCR analysis	72
2.3.5 Electropherogram analysis	72
2.3.6 Statistical analysis	73
2.3.7 Presumptive identity analysis	73
2.4 Results	74
2.4.1 PCR amplification of 16S rRNA genes for metagenome profiling	74
2.4.2 Comparison of LH eubacterial profiles within CF center	77
2.4.3 Patient CF flora changes over time	80
2.4.4 Antibiotics drive CF community	84
2.4.5 Presumptive identity analysis	87
2.5 Discussion	90
2.5.1 The V1 region of the 16S rRNA gene provides a more detailed look at the complex bacterial community in the CF lung	90
2.5.2 CF patients harbor diverse bacterial communities	91
2.5.3 Challenges of identifying organisms in complex communities	93
2.5.4 The eubacterial communities in CF lungs are dynamic	95
2.6 Acknowledgements	99
Chapter 3: Cystic Fibrosis Lung Eubacteriome and Mycobiome as Revealed by 454 Sequencing	110
3.1 Abstract	111
3.2 Introduction	112
3.3 Materials and methods	115
3.3.1 Ethics statement	115
3.3.2 Study participants	115
3.3.3 DNA extraction	117
3.3.4 Metagenome analysis	117
3.3.5 Data analysis	118
3.4 Results	120
3.4.1 Participant demographics	120
3.4.2 Identification of bacterial taxa in CF samples	120
3.4.3 Relationship between gender and ethnicity and the eubacteriome	123
3.4.4 Healthy oral eubacteriome	123
3.4.5 Comparison of CF and healthy eubacteriomes	126
3.4.6 Identification of fungal taxa in CF samples	127
3.4.7 Relationship between gender and ethnicity and the mycobiome	130

3.4.8	Healthy oral mycobiome	131
3.4.9	Comparison of CF and healthy microbiomes	132
3.5	Discussion	136
3.5.1	MTPS shows greater diversity in the CF bacterial community as compared to clonal sequencing	136
3.5.2	Common and emerging pathogenic bacteria are present in CF lungs	137
3.5.3	Oral microbiota may play a role in disease progression	138
3.5.4	Rarely cultured bacteria are readily identified using molecular methods	140
3.5.5	Unusual organisms detected in CF sputa	141
3.5.6	Identifications above genus level are difficult to interpret	144
3.5.7	Gender and ethnicity of the host may not affect the bacterial community	145
3.5.8	Healthy and CF bacterial communities are statistically different despite shared bacteria	145
3.5.8.1	CF pathogens mostly absent in healthy eubacteriomes	146
3.5.8.2	CF and healthy samples share bacteria	147
3.5.8.3	Bacterial diversity decreased in CF samples	148
3.5.9	Fungi are more prevalent in CF sputa	148
3.5.10	Pathogenic fungi undetected by clinicians	149
3.5.11	Environmental fungi abundant in CF sputa	150
3.5.12	Gender and ethnicity may not affect fungal community composition	150
3.5.13	Healthy oral cavity contains a diverse mycobiome	151
3.5.14	CF sputa and oral rinses contain unique mycobiomes	151
3.5.14.1	Pathogenic fungi were present in healthy and CF subjects	152
3.5.14.2	Environmental fungi abundant in both CF sputa and healthy oral samples but no species are in common	153
3.5.15	Conclusion	153
3.6	Acknowledgements	154
Chapter 4: SPR Detection of Bacteria Associated with the Cystic Fibrosis Lung		158
4.1	Abstract	159
4.2	Introduction	160
4.3	Materials and methods	170
4.3.1	Bacterial isolates	170
4.3.2	Determination of bacterial concentration	171
4.3.3	Reagents and instrumentation	171
4.3.4	Preparation of gold sensor chip	172
4.3.5	Immobilization of antibodies	172
4.3.6	Detection of <i>S. aureus</i> and <i>P. aeruginosa</i>	174
4.3.7	Determination of the sensitivity for the <i>P. aeruginosa</i> biosensor	174
4.3.8	Determination of the specificity for the <i>P. aeruginosa</i> sensor	175
4.4	Results	176
4.4.1	Self assembly of <i>S. aureus</i> immunosensor	176

4.4.2	Detection of <i>S. aureus</i>	176
4.4.3	Detection of <i>P. aeruginosa</i>	178
4.4.4	<i>P. aeruginosa</i> immunosensor used for sensitivity tests	178
4.4.5	The effect of BSA on non-selective binding using the <i>P. aeruginosa</i> immunosensor	181
4.5	Discussion	183
4.5.1	SPR immunosensor detected <i>S. aureus</i>	183
4.5.2	<i>P. aeruginosa</i> immunosensor detected solutions containing <i>P. aeruginosa</i>	184
4.5.3	BSA decreases the sensitivity of the <i>P. aeruginosa</i> sensor	186
4.5.4	Non-BSA chip shows selective binding of bacterial cells with cognate antibodies	187
4.5.5	Optimization of immunosensors	187
4.6	Acknowledgments	190
Chapter 5: Summary and General Discussion		191
5.1	Overview	192
5.2	Bacterial community profiling in the CF lung	193
5.3	High throughput 454 sequencing identifies bacteria and fungi associated with the CF lung	200
5.4	SPR as a bacterial detection tool	208
5.5	Conclusions	212
BIBLIOGRAPHY		215
VITA		254

## LIST OF TABLES

TABLE	PAGE
2.1 Amplicons seen in the V1 and V1_V2 regions for samples	75
2.2 Amplicon contributions for the Patients UMX and UMY over a five-month time period	85
3.1 Patient demographics	116
3.2 Frequency distribution (%) of bacteria present with at least 1 % abundance in the CF sputa eubacteriome	122
3.3 Frequency distribution (%) of bacteria present with at least 1% abundance in the healthy mouth eubacteriome	125
3.4 Bacteria sequenced in CF sputa and/or mouth rinse from healthy individuals	126
3.5 Frequency distribution (%) of fungi present with at least 1% abundance in the CF sputa mycobiome	129
3.6 Frequency distribution (%) of fungi present with at least 1% abundance in the healthy oral mycobiome	133
3.7 Fungi sequenced in CF sputa and/or mouth rinse from healthy individuals	134
Supplemental Table 3.8 Frequency distribution (%) of fungal species present with at least 1% abundance in the CF sputa mycobiome	155
Supplemental Table 3.9 Frequency distribution (%) of fungal species present with at least 1% abundance in the healthy oral rinses mycobiome	156

## LIST OF FIGURES

FIGURE	PAGE
1.1 Cystic fibrosis transmembrane conductance regulator (CFTR) Channel	4
1.2 Respiratory infections at a given age	5
1.3 Methods of bacterial and fungal diagnosis	9
1.4 CHROMagar plates	10
1.5 CHROMagar <i>Candida</i> plate	12
1.6 MALDI-TOF MS spectra for six different bacterial species	21
1.7 Electrospray ionization	24
1.8 IgG antibody	26
1.9 Sandwich ELISA	28
1.10 Surface plasmon resonance	31
1.11 Ribosome subunits	34
1.12 16S rRNA secondary structure map	35
1.13 Schematic representation of the variable and conserved regions of the 16S rRNA genes, using <i>Escherichia coli</i> genome as a reference	36
1.14 Fungal rRNA genes	37
1.15 Terminal restriction fragment length polymorphism (T-RFLP)	41
1.16 Sanger sequencing	47
1.17 Pyrosequencing reaction	50
1.18 Overview of 454 sequencing	54
1.19 Illumina sequencing technology	55
1.20 SOLiD™ sequencing technology	57

2.1	Hypervariable V1_V2 amplicon data graphs	76
2.2	Hypervariable V1 amplicon data	78
2.3	Similarity based clustering between LH profiles of patients	79
2.4	Similarity based clustering between LH profiles over time in one patient	82
2.5	Similarity between LH profiles over time in one patient	83
Supplemental Figure 2.6	Hypervariable region V1_V2 amplicon data graphs	100
Supplemental Figure 2.7	Hypervariable region V1 amplicon data graphs	104
3.1	Bacteria taxon present $\geq 1\%$ abundance in CF sputum samples	121
3.2	Relationship between the host and the eubacteriome	124
3.3	PCO analysis of healthy mouth rinse and CF eubacteriomes	127
3.4	Fungi genera present in $\geq 1\%$ abundance in CF sputum samples	128
3.5	Relationship between the host and the fungal community	131
3.6	Relationship between CF sputa and healthy mycobiomes	135
4.1	Surface plasmon resonance phenomenon	164
4.2	Self-assembled monolary (SAM) formation on a gold coated chip	165
4.3	Schematic diagram of SPR instrument	166
4.4	Antibody orientation	167
4.5	The SPR response	168
4.6	A representative sensograph	173
4.6	Detection of <i>S. aureus</i> using an immunosensor	177
4.7	<i>P. aeruginosa</i> detection using an immunosensor	179
4.8	Sensitivity of the <i>P. aeruginosa</i> immunosensor	180
4.9	Specificity of <i>P. aeruginosa</i> immunosensor	182

## LIST OF ABBREVIATIONS

TERM/UNIT OF MEASUREMENT	SYMBOL/ABBREVIATION
Adenosine triphosphate	ATP
Amplicon length heterogeneity – polymerase chain reaction	LH
Bovine serum albumin	BSA
<i>Burkholderia cenocepacia</i> selective agar	BCSA
Bronchoalveolar lavage	BAL
Colony forming units	cfu
Cystic fibrosis	CF
Cystic fibrosis transmembrane conductance regulator	CFTR
Degree Celsius	°C
Deoxyribonucleic acid	DNA
Deoxyribonucleotide triphosphate	dNTP
Diethylpyrocarbonate	DEPC
Electrospray ionization	ESI
Enzyme-linked immunosorbent assay	ELISA
<i>et alia</i>	<i>et al.</i>
Hypervariable region 1	V1
Hypervariable region 1 and 2	V1_V2
Institutional review board	IRB
Internal transcriber spacer	ITS
Lab-on-chip	LOC
Luria Bertani	LB

Magnesium chloride	MgCl <sub>2</sub>
Mass spectrometry	MS
Matrix-associated laser desorption/ionization	MALDI
Microgram	µg
Milliliter	mL
MultiTag <sup>TM</sup> pyrosequencing	MTPS
N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride	EDC
N-hydroxysuccinimide	NHS
Non-metric multidimensional scaling	MDS
Overnight	ON
Percent	%
Polymerase chain reaction	PCR
Phosphate buffer saline	PBS
Principle coordinate analysis	PCO
<i>Pseudomonas</i> isolation agar	PIA
Pyrophosphate	pp <sub>i</sub>
Ribonucleic acid	RNA
Ribosomal RNA	rRNA
Self-assembled monolayer	SAM
Surface plasmon resonance	SPR
Terminal restriction fragment length polymorphism	T-RFLP
Time-of-flight	TOF



## **Chapter 1**

### **Introduction**

Sections have been published in Human Genomics

Volume 3 pgs 246-256, 2009

Copyright © Henry Stewart Publications

Melissa Doud, Erliang Zeng, Lisa Schneper, Giri Narasimhan, and Kalai Mathee

## 1.1 Preface

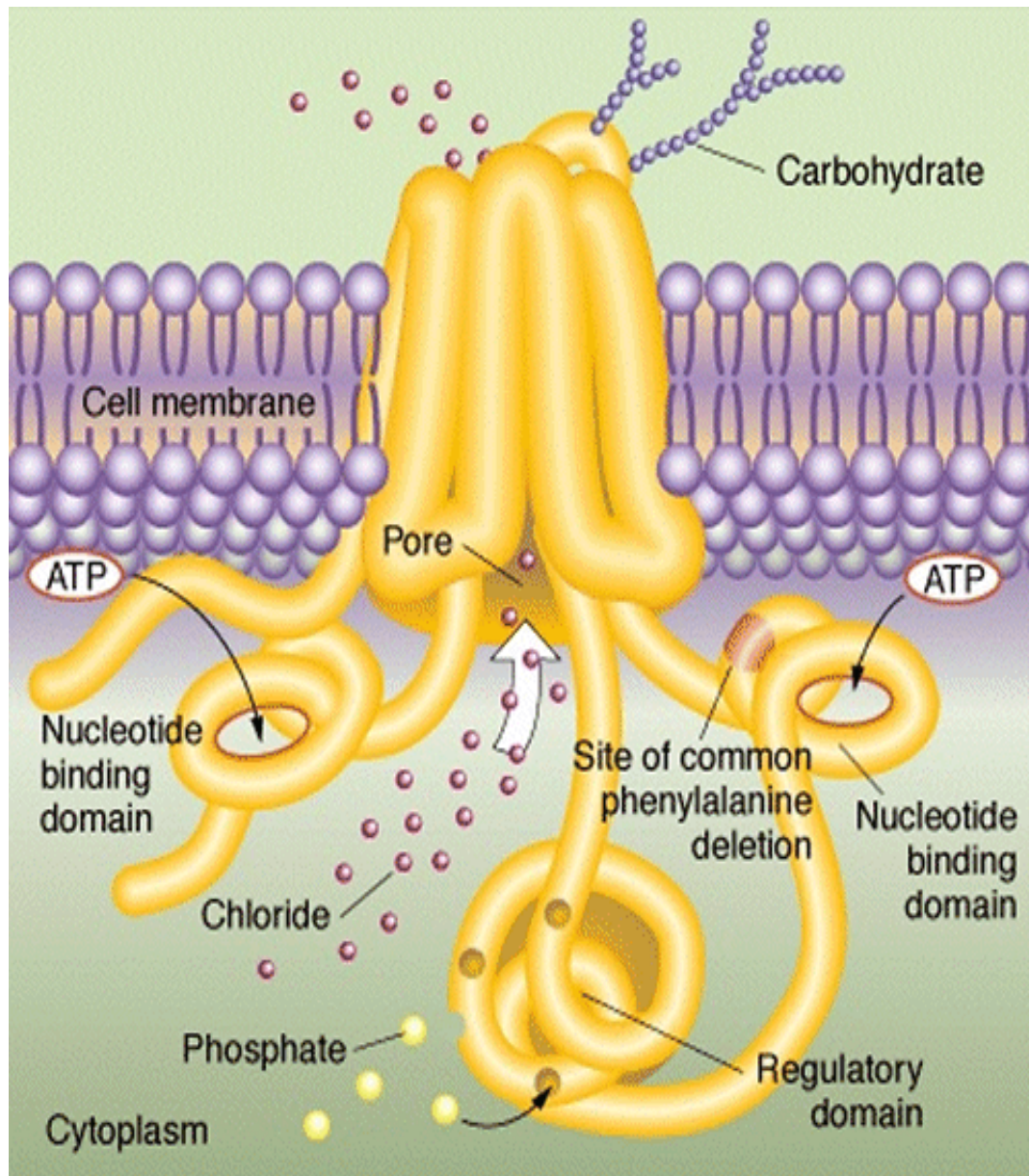
Cystic fibrosis (CF) is a genetic disorder that affects the exocrine glands of the human body (Davis, 1987). Sweat glands, lungs, and pancreas are the most affected but problems can arise in other systems such as the reproductive system (Davis, 1987). The mucus found in the lungs, pancreas, and reproductive system is thicker and stickier in people who have CF and this leads to chronic lung infection, digestion problems and infertility, respectively (Andersen, 1938; Johansen *et al.*, 1968; Oppenheimer & Esterly, 1969; Oppenheimer *et al.*, 1970). The leading cause of morbidity and mortality in CF patients is due to lung deterioration as the result of the chronic infections (Buzzetti *et al.*, 2009; Patient-registry, 2008). Diagnosis and treatment of the pulmonary infections is key to the overall health and age of survival for the CF individual. The following introduction discusses the microbial community found in the lungs of CF patients and the techniques that identify these microbes.

## 1.2 Cystic fibrosis

Cystic fibrosis (CF) is a fatal inherited disease primarily affecting Caucasians. In the United States (US), 3500 children are born with the disease each year (Patient-registry, 2008). The gene responsible for CF encodes a protein called the CF transmembrane conductance regulator (CFTR) (Kerem *et al.*, 1989; Riordan *et al.*, 1989; Rommens *et al.*, 1989). The CFTR is a secretory epithelial cyclic-AMP-activated chloride channel; mutations in the *cftr* gene lead to decreased fluid secretion and dehydration of epithelial surfaces (Figure 1.1) (Anderson *et al.*, 1991). The change in

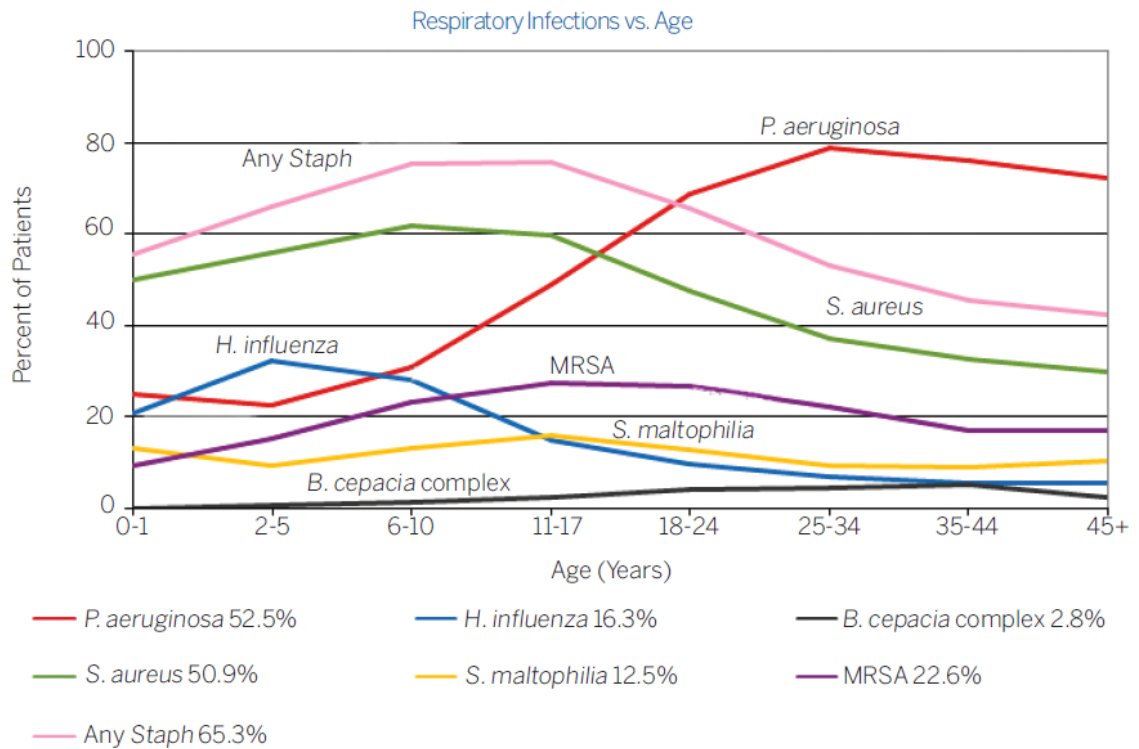
mucus production and consistency affects various systems of the human body. The thick mucus can plug the pancreatic ducts which lead to pancreatic enzyme insufficiency and thus vitamin deficiency and malnutrition (Andersen, 1938; Johansen *et al.*, 1968; Oppenheimer & Esterly, 1969). The thick mucus also causes reproductive complications (Kaplan *et al.*, 1968; Kotloff *et al.*, 1992; Oppenheimer & Esterly, 1969; Oppenheimer & Esterly, 1970). Although the anatomy of the female reproductive tract is normal, pregnancy can be difficult as a result of the presence of the viscous mucus preventing the sperm from fertilizing the egg (Kopito *et al.*, 1973; Oppenheimer *et al.*, 1970). Over 90 % of males are born with congenital bilateral absence of the vas deferens, meaning that a man can produce sperm but due to the lack of vas deferens the sperm cannot leave the testicles (Landing *et al.*, 1969; Mercier *et al.*, 1995; Oppenheimer & Esterly, 1969).

The airways in a CF individual are also negatively affected by the defective CFTR protein. Oversecretion of thick mucus in the airway leads to congestion of the respiratory tract and increased susceptibility to chronic broncho-pulmonary infection, which is the major cause of morbidity and mortality among patients with CF (Buzzetti *et al.*, 2009). To retard the rate of decreasing lung function, bacterial infections are treated with antibiotics, however these must be tailored to the particular infection, which is often polymicrobial. For example, anti-pseudomonal drugs are often ineffective against *Burkholderia cenocepacia* infections due to the bacteria's resistance to the drug (McGowan Jr, 2006). Thus, it is important to correctly identify the infecting pathogens, bacteria and fungi, in order to prescribe an appropriate treatment and prolong the life of the individual.



**Figure 1.1:** Cystic fibrosis transmembrane conductance regulator (CFTR) channel. Schematic representation of proposed CFTR structure. CFTR is made up of five domains: two nucleotide-binding domains that bind and hydrolyze ATP and a regulatory domain comprised of two membrane-spanning domains which form the chloride ion channel, (<http://www.cfgenethrapy.org.uk/>).

**1.2.1 CF sputum bacterial flora.** Initially because of culturing techniques, it was thought that only a few bacterial species were present in the CF lungs. *Staphylococcus aureus*, *Haemophilus influenzae* and *Pseudomonas aeruginosa* are the primary pathogens cultured from CF patients (Whittier, 2001). *Staphylococcus aureus*, the predominant pathogen in children, is succeeded by *H. influenzae* during early childhood and *P. aeruginosa* becomes the predominant pathogen during adolescence, reaching a prevalence rate of 80% in adults (Beringer & Appleman, 2000; Patient-registry, 2008) (Figure 1.2).



**Figure 1.2:** Respiratory infections at a given age. Patients throughout their life are affected with numerous bacteria. Each line represents a specific bacterial infection over the lifespan of the patient. The height of the line indicates the percentage of patients that have a particular bacterial infection at that specific age (Patient-registry, 2008).

However, there has been the emergence of other opportunistic pathogens such as *Burkholderia cenocepacia*, *Alcaligenes xylosoxidans*, *Ralstonia pickettii*, *Burkholderia gladioli*, *Stenotrophomonas maltophilia*, and *Mycobacterium* species (Pierre-Audigier *et al.*, 2005; Whittier, 2001). The occurrence of the more recent emerging organisms increases with advancing age and severity of lung disease (Beringer & Appleman, 2000; Lewin *et al.*, 1990; Patient-registry, 2008; Pierre-Audigier *et al.*, 2005). The identification of these emerging pathogens suggests the presence of a complex bacterial community present in the CF lung. The true polymicrobial nature of the lung was revealed with the development of DNA based diagnostic tools. These methods have identified *Streptococcus mitis*, *Streptococcus pneumoniae*, *Prevotella melaninogenica*, *Veionella* sp., *Granulicatella para-adiacens* and *Exiguobacterium* in CF samples (Bittar *et al.*, 2008; Harris *et al.*, 2007). Potential novel pathogens *Lysobacter*, *Coxiellaceae*, and *Rickettsiales* have also been found (Harris *et al.*, 2007). The diversity of the CF lung community continues to increase and change as more samples from CF lungs are analyzed.

**1.2.2 CF sputum fungal flora.** Invasive mold infections are very common in patients suffering from chronic bronchitis. Thus, it was not surprising when the fungal infections in CF patients were recognized as early as 1967 by the presence of precipitating antibody to *Aspergillus fumigatus* (Mearns *et al.*, 1967). In fact, *A. fumigatus* is one of the most common fungal pathogens found in the CF lung. In one study, 65 patients underwent a lung transplant and it was noted that 52% of these patients were determined to being infected with *A. fumigatus* before surgery

(Paradowski, 1997). *Candida albicans* is the other common fungal pathogens (Bakare *et al.*, 2003; Cimon *et al.*, 2001). Recently, *Candida dubliniensis*, a close relative of *C. albicans*, has been detected in these patients as well (Kim *et al.*, 2003). Occasionally, other fungi such as *Scedosporium apiospermum*, *Scedosporium prolificans*, *Aspergillus terreus*, *Acrophialophora fusispora*, *Exophiala dermatitidis* and *Penicillium emersonii* have been shown to infect CF patients (Bakare *et al.*, 2003; Cimon *et al.*, 1999; Cimon *et al.*, 2003; Cimon *et al.*, 2005; Defontaine *et al.*, 2002; Hennequin *et al.*, 1997; Horr e *et al.*, 2004).

Recent statistics have shown that there is an increase in the prevalence of fungi being detected in the CF patient which may be a consequence from aggressive antibacterial therapy (Nagano *et al.*, 2007). Yet, the prevalence of fungi present in the CF lung is still unknown. For instance, the prevalence of *Aspergillus* infection ranges from nine to 57 % of CF individuals (Bakare *et al.*, 2003). In addition, the fungal community may be more diverse than culturing results have shown. Depending on the organism present, fungal infections can lead to health problems ranging from allergic bronchopulmonary aspergillosis (ABPA) to post-transplant fungaemia (Fahy *et al.*, 1991; Kubak, 2002; Mastella *et al.*, 2000; Nagano *et al.*, 2007) It is also possible for a fungal infection to become chronic and lead to bronchitis (Nagano *et al.*, 2007).

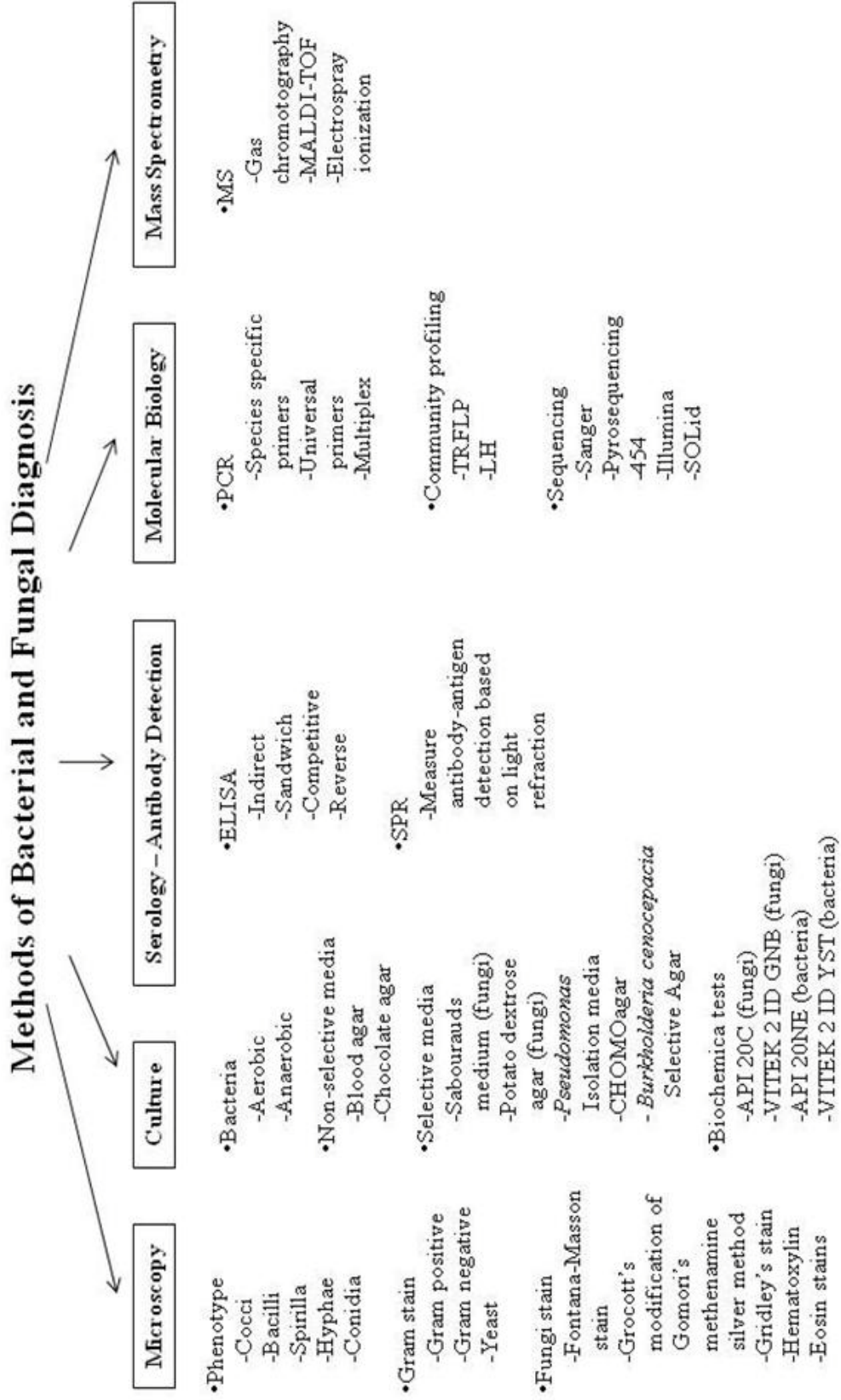
### **1.3 Common assays used for clinical identification of bacteria and their limitations**

Cystic fibrosis patients that are experiencing an exacerbation typically produce a sputum sample that can be analyzed for the presence of pathogenic bacteria and fungi. Most often diagnostic clinics use culturing and commercial kits to identify common pathogenic organisms that may be present in the sample (Figure 1.3). A few of these assays and their limitations are discussed in further detail below.

**1.3.1 Culturing and commercial phenotypic identification assays.** Culturing for microbial identification is commonly used in laboratories handling CF sputum. Using various media, sputum is swabbed onto plates and then the organism is allowed to grow for a period of time. Based on the phenotype of the organism and the media it is growing on, an identification of the bacteria or fungi is made. Bacteria can be cultured using various media including: Luria Bertani (LB), blood agar, chocolate agar, *Pseudomonas* Isolation Agar (PIA), *Burkholderia cenocepacia* Selective Agar (BCSA) and MacConkey agar (Murray & Baron, 2007). The phenotype of the bacteria can be examined using microscopy. In addition, a sample of the bacteria can undergo Gram Staining to determine if the bacteria is Gram-positive or Gram-negative (Gram, 1884). This information in conjunction with phenotypic results is used to identify organisms present in the sample.



**Figure 1.3:** Methods of bacterial and fungal diagnosis. Flowchart demonstrating numerous ways organisms can be identified.



CHROMagar, a specialized media can identify some bacteria based on the colony color; thus, circumventing any need for microscopy (Figure 1.4). CHROMagar plates contain a specific substrate that is hydrolyzed by a specific organism which results in the bacterial colony being colored or surrounded by a diffusible color (Merlino *et al.*, 1996). There are plates can identify various species of *Salmonella*, *Staphylococcus aureus* including methicillin resistant strains, Group B *Streptococcus*, *Enterococcus*, various strains of *Escherchia coli*, *Listeria monocytogenes*, and various species of *Vibrio* (Alonso *et al.*, 1996; Gaillot *et al.*, 1999; Hegde *et al.*, 2007; Merlino *et al.*, 1996; Pape *et al.*, 2006; Su *et al.*, 2005).



**Figure 1.4:** CHROMagar plates. Seven bacteria were grown on a plate. The color of the colonies indicates the organism plated ([www.chromagar.com](http://www.chromagar.com)).

Besides visual phenotypes, the biochemical tests are performed. Currently, multiple biochemical tests can be performed at the same time. For example, the API 20 NE system (BioMérieux, Marcy l'Etoile, France) is a strip that contain up to 20 biochemical tests. There are 16 identification products with comprised of different

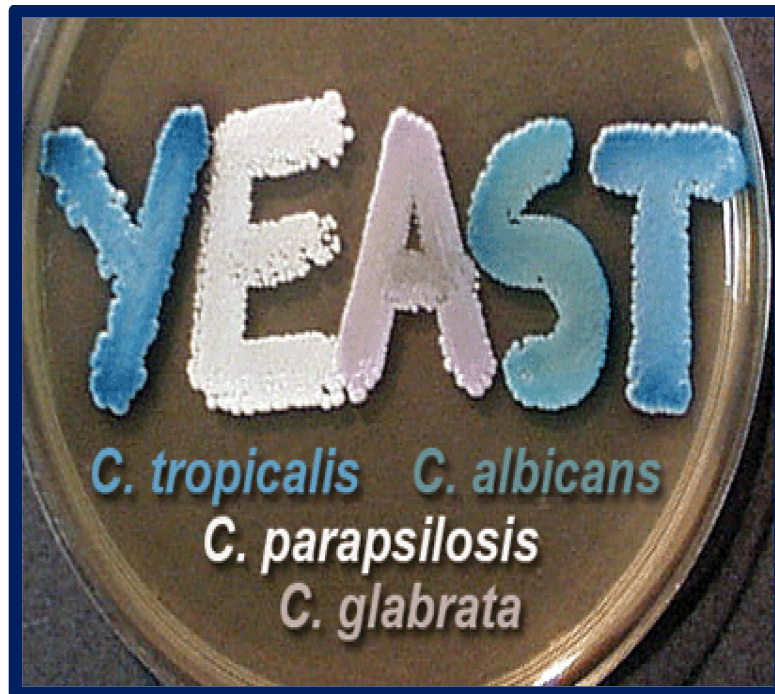
biochemical tests that should identify up to 550 different species. These products identify bacteria in two to 24 hours. Another advanced product that identifies bacteria is the VITEK 2 by BioMérieux. The fully automated BioMérieux system uses different cards containing biochemical assays. Fluorescence is used to detect the outcome of the assay and then based on the results of the assays bacteria are identified. The VITEK2 ID-GNB card can identify bacteria based on 41 biochemical tests and 20 preformed enzymes (O'Hara & Miller, 2003).

Fungal identification tools are similar to the methods used in bacterial detection. Culturing, phenotypic determination using microscopy and Gram staining can all be used in the identification process. Fungi can be cultured on blood agar, chocolate agar, eosine methylene blue agar, Sabouraud glucose agar, and potato dextrose agar (Murray & Baron, 2007). Often the CF sputum sample contains *Pseudomonas aeruginosa* which is a bacterium that can grow on the media listed above and thus impede the growth and identification of the fungi. Therefore, antibiotics can be added to the medium to stop bacterial contamination (Bakare *et al.*, 2003). Culturing fungi can take as little as three days but many species need weeks to grow. After the organism has been isolated, microscopy is used to examine the fungi's hyphae. The phenotype can be used to provide an identification (Bakare *et al.*, 2003).

Gram staining can also be used for yeast such as *Candida* (Gram negative) to make the organism more readily observed microscopically (Subhash, 2004). *Aspergillus* and other molds only retain a small amount of color and thus are poorly stained. Many other fungi can be visualized and differentiated using other stains such as Fontana-Masson stain, Grocott's modification of Gomori's methenamine silver method, Gridley's

stain, hematoxylin and eosin stains (Dixon & Polak-Wyss, 1991; Huppert *et al.*, 1978; Saeed & Hay, 1981; Schwarz, 1982; Wood & Russel-Bell, 1983).

Biochemical assays can also aid in identification. These methods may include CHROMagar plates which contain a substrate that is hydrolyzed by a specific fungus. The product from the reaction stains the fungus and based on the color the *Candida* species can be identified (Figure 1.5) (Odds & Bernaerts, 1994). For example, CHROMagar *Candida* plates uses a chromogenic  $\beta$ -glucosaminidase substrate, which when metabolized results in green colonies of *C. albicans*, steel blue colonies of *C. tropicalis*, and fuzzy rose-colored colonies of *C. krusei* (Baumgartner *et al.*, 1996; Cooke *et al.*, 2002) (CHROMagar, Paris, France). Currently, no other CHROMagar plates have been developed for other species of fungi.



**Figure 1.5:** CHROMagar *Candida* plate. Plate shows the growth of four different *Candida* species on CHROMagar *Candida* media ([www.mycology.adelaide.edu.au](http://www.mycology.adelaide.edu.au)).

Multiple phenotypic commercial assays exist for fungal identification. For example, BioMérieux (Marcy l'Etoile, France) makes strips such as the API 20C AUX that contain 18 biochemical tests which aid in fungal identification (Gundes *et al.*, 2001). In this assay, a liquid suspension is made from a fungal isolate that has been growing for 48-72 hours. The suspension is placed onto the strip and based on the results of the individual chemical tests, an identification is made (Gundes *et al.*, 2001). The automatic identification system VITEK 2 (BioMérieux) also works with fungi by using the ID-YST card which consists of 47 biochemical tests and is capable of identifying 51 taxa in 15 hours (Graf *et al.*, 2000).

**1.3.2 Accuracy of commercial identification systems.** Multiple studies have been performed to determine the accuracy and specificity of the commercially available identification system for both bacteria and fungi (Bosshard *et al.*, 2006; Brisse *et al.*, 2002; O'Hara & Miller, 2003; van Pelt *et al.*, 1999).

**1.3.2.1 Bacterial identification.** Three studies have examined the accuracy of the VITEK2 ID-GNB card (BioMérieux, Marcy l'Etoile, France). In one study, the commercial system accurately identified 93.0 % of 414 clinically relevant enteric bacteria strains (O'Hara & Miller, 2003). This study also correctly identified 95.1 % of the 103 glucose-fermenting and nonfermenting nonenteric strains (O'Hara & Miller, 2003).

Another study compared the bacterial identification systems, API 20 NE (BioMérieux) and the VITEK2 ID-GNB card (BioMérieux), with sequencing data from the same 107 isolates (Bosshard *et al.*, 2006). The API system was only able to identify 54 % of the isolates to species level and 7 % to the genus level. Thirty-nine percent or 42 isolates were not identified at all and of 26 % (11 isolates) belonged to a species that was not in the API database (Bosshard *et al.*, 2006). The VITEK 2 system had similar results where as 53 % of the isolates were identified to species level, one percent to genus level, and 46 % (49 isolates) were unidentifiable. Of the 49 isolates, 39 belonged to a species not represented in the VITEK 2 database (Bosshard *et al.*, 2006). A third study compared the VITEK with the BD Phoenix (BD Diagnostic System, Sparks, MD) to determine the accuracy in identifying isolates of the *B. cenocepacia* complex, a known CF pathogen (Brisse *et al.*, 2002). The study determines an overall correct ID rate of 50 % for both systems. The authors conclude by strongly recommending the use of molecular techniques to identify members of the *B. cenocepacia* complex (Brisse *et al.*, 2002).

**1.3.2.2 Fungal identification.** Commercial fungal identification systems have also been evaluated. The API 20C AUX strip (BioMérieux, Marcy l'Etoile, France) was determined to be 87 % accurate in a study that used 116 strains representing 11 species of *Candida* and one species of *Trichosporon* (Gundes *et al.*, 2001). This increased by 12 % when the strains not equivocally identified were analyzed with an additional step. Only one strain was misidentified. In a separate study, the VITEK 2

ID-YST card (BioMérieux) was tested with 241 strains representing 21 species (Graf *et al.*, 2000). Of these strains, 211 were identified to the species level. Eleven other strains were identified conclusively after an additional assay was performed. Of the remaining 19 strains, 10 strains were not conclusively identified even with additional tests, four fungi were misidentified, and five were unidentifiable (Graf *et al.*, 2000). The most common *Candida* species isolated in clinic laboratories were correctly identified 93.8 % of the time. With the addition of a second assay, 97.5 % of *Candida* species were identified correctly. These two studies concluded that both BioMérieux systems can identify high percentage of clinically relevant fungi that have been isolated from a patient sample.

**1.3.3 Misidentification through culturing.** Commercial phenotypic tests are indeed a rapid cost efficient way of identifying pathogens but there are potential issues that lead to a misdiagnosis (Bosshard *et al.*, 2006). These systems rely on a reference database which can be limiting because the sample may contain organisms that are not in the database (Bosshard *et al.*, 2006; Ramani *et al.*, 1998). Error may also occur because of strain variation, that is, not all strains within a species possess the same phenotypes (Brisse *et al.*, 2002; Wigley & Burton, 1999). Patients have been diagnosed with *Burkholderia cenocepacia* on the basis of results obtained from culturing techniques but with more stringent molecular techniques, it was shown to be *Burkholderia gladioli* or *Stenotrophomonas maltophilia* (McDowell *et al.*, 2001; McMenamain *et al.*, 2000). Identification of common CF pathogen, *Pseudomonas*

*aeruginosa*, can be difficult. The CF-derived isolates of this bacterium acquire many mutations resulting in a huge phenotypic diversity from loss of pigment production, mucoidy, and change in lipopolysaccharide (Spilker *et al.*, 2004). Occasionally, *P. aeruginosa* is identified as another bacterium. In one study, 770 strains of *B. cenocepacia* were reanalyzed using a polyphasic approach that included selective media and biochemical tests. Only 682 strains were confirmed to be *B. cenocepacia*. The remaining 88 samples were misidentified of which 11 of them were *P. aeruginosa* (Shelly *et al.*, 2000). The misidentified strains were inaccurately identified using a variety of commercial assays including the VITEK and API 20NE (BioMérieux, Marcy l'Etoile, France).

Multiple studies have shown that fungal misidentification is also a problem (Balajee *et al.*, 2006; Rowen *et al.*, 1999). For example, a common CF fungus, *Aspergillus fumigatus* is identified based on macroscopic and microscopic features. Through the use of molecular techniques, *A. fumigatus* identified isolates were shown to be *Aspergillus lentulus* and *Aspergillus udagawae* (Balajee *et al.*, 2006). These organisms vary greatly in their resistance to certain drugs and thus misidentification may contribute to failure in treatment certain patients (Balajee *et al.*, 2006). Strains of *Candida* are often confused with one another because of similar phenotypes (Coleman *et al.*, 1997). Some strains also possess similar biochemical activities which can cause inaccurate results when using commercial systems like the API 20C thus leading to misidentification (Elie *et al.*, 1998; Rowen *et al.*, 1999).



All of the above bacterial and fungal identification systems rely on culturing which therein is a problem. It has been estimated that less than 1 % of eubacteria in the environment can be cultured (Roszak & Colwell, 1987; Ward *et al.*, 1990a; Ward *et al.*, 1990b). Therefore, the bacteria present in a sample may not grow in a laboratory setting and thus can never be identified using culture based phenotypic assays. Fungal organisms can also be difficult to grow and some species can take up to 5 weeks to culture (Elie *et al.*, 1998; Murray & Baron, 2007). Misidentification problems can be reduced or completely eliminated by using non-phenotypic methods like mass spectrometry and genotype-based molecular identification methods (Elie *et al.*, 1998; Spilker *et al.*, 2004).

#### **1.4 Identification through mass spectrometry**

There are many ways to identify bacteria and fungi that do not rely on phenotypic assays. Some of these techniques originated from the chemistry field and were later adapted to study microorganisms. One of the most widely studied techniques relies on mass spectrometry.

**1.4.1 Mass spectrometry (MS).** Since the 1970's mass spectrometry has been used to identify bacteria (Abel *et al.*, 1963; Fox, 2006). Mass spectrometry determines the elemental composition of a sample by identifying the mass to charge ratio of the compound (Aston, 1919; Dempster, 1918; Gross, 2004). The first mass spectrometry based bacterial identification technique used MS coupled with gas chromatography (GC)

to study the fatty acid composition of a bacterium (Abel *et al.*, 1963; Fox, 2006). In this technique, the polymers such as phospholipids are first broken into monomers by hydrolysis or saponification. The monomers are then chemically converted through derivatization by methylation in order to inhibit ionic or hydrogen-bonding interactions which allow the compounds to be volatile and separated through the column in the gas chromatograph. The isolated monomer which is in the gas phase then enters the mass spectrometer and is ionized (Gross, 2004). The mass to charge ratio for the monomer is calculated and with the generated retention time from the gas chromatograph, a fatty acid mixture profile is determined for the bacteria being analyzed. The profile is characteristic to each bacterium and thus provides identification of a sample (Fox, 2006).

In the 1980s and 1990s, MS technology evolved and soft ionization MS was developed (Fenn *et al.*, 1989; Fox, 2006; Karas & Hillenkamp, 1988). In soft ionization MS, compounds do not have to be separated into individual components before undergoing MS or if separation is needed then it is done using a liquid phase (Gross, 2004). Soft ionization MS has been coupled to numerous instruments and continues to be used in bacterial identification (Fox, 2006). The roles and limitations of matrix-assisted laser desorption/ionization – time of flight mass spectrometry (MALDI-TOF MS) and electrospray ionization (ESI-MS) in bacterial identification will be discussed in more detail.

**1.4.2 MALDI-TOF MS.** In this technique, the sample being analyzed is spotted, with a matrix, on a metal plate and allowed to air dry. The matrix is typically a solution containing 3,5-dimethoxy-4-hydroxycinnamic acid (sinapinic acid),  $\alpha$ -cyano-4-

hydroxycinnamic acid (alpha-cyano or alpha-matrix) or 2,5-dihydroxybenzoic acid (DHB) and water, ethanol, or acetonitrile (Gross, 2004). Once the solvent has evaporated, a crystallized matrix with the sample is left on the plate. The MALDI plate is placed in the TOF MS. The matrix spot is struck with a laser beam which absorbs the light and transfers some of the charge to the sample of interest. The ions from the sample drift through a tube to the MS detector. The smaller the ion, the faster it will traverse the tube and be detected. The TOF MS spectrum produced indicates the mass to charge ratio for the sample and the time in which the ions took to reach to the detector (Gross, 2004).

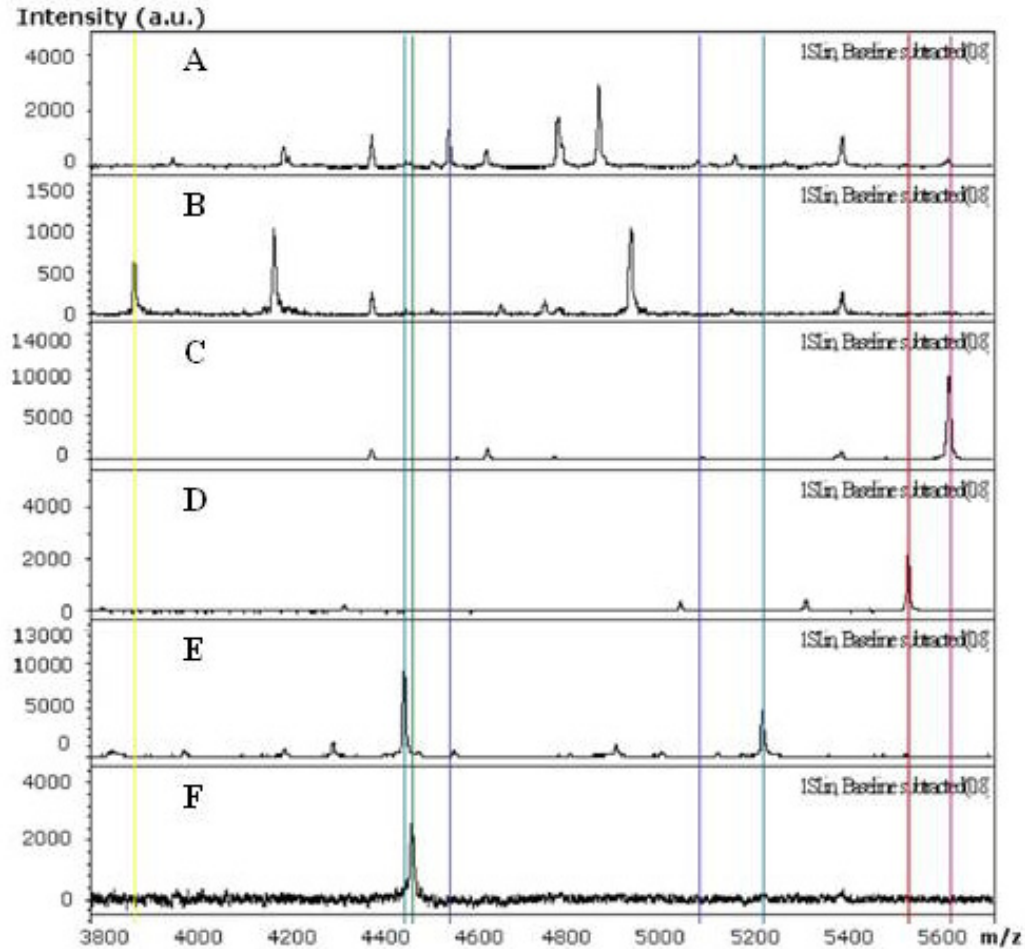
The MALDI-TOF MS spectra has been used to look at the DNA, proteins, bacterial toxins, bacteria DNA/RNA and bacterial metabolites (Jackson, 2001). Most notably here is the use of proteins and whole cells to differentiate between bacteria species. Bacteria were identified first by using MALDI-TOF MS to generate a total protein profile from a total protein fraction from lysed bacteria (Cain *et al.*, 1994; Jackson, 2001). Another group, Holland *et al.* took bacteria colonies from an agar plate and dried them on the MALDI plate with a solvent. The protein profile did not contain all proteins that would be seen on the whole protein fraction spectra yet enough ions were present to still ascertain and identification (Holland *et al.*, 1996; Jackson, 2001). Reproducibility of these spectra from whole cells, both from colonies and from liquid culture, plagues this analysis (Jackson, 2001). However, the technique was refined and eventually identification of whole cell bacteria became a more reliable tool.

To make MALDI-TOF MS more applicable as a tool for clinical diagnosis, a database of protein profiles for individual bacteria was generated (Keys *et al.*, 2004). Over 3500 spectra with multiple spectra for a species were produced. The ability to

identify an unknown profile using the database ranges from 33 to 100 % (Hsieh *et al.*, 2008; Keys *et al.*, 2004). As a result of the strain variation in protein profiles, some species were not correctly identified. This error decreased as more strains were added to the database (Keys *et al.*, 2004).

The database software may have also caused low identification rates (Hsieh *et al.*, 2008). This problem was addressed using bioinformatics programs like Support Vector Machines and other algorithms (Hsieh *et al.*, 2008). By using clustering methods, identification based on the protein profile of a bacterial colony greatly increased. In addition, the programs were able to identify components of bacteria mixtures from six human pathogenic bacteria (*Staphylococcus aureus*, *Streptococcus* serogroup B, *Escherichia coli*, *Klebsiella pneumoniae*, *Salmonella enterica* serogroup B, and *Pseudomonas aeruginosa*) (Figure 1.6) (Hsieh *et al.*, 2008). The MALDI-TOF MS spectra has been determined for many CF pathogens including members of the *B. cenocepacia* complex (Degand *et al.*, 2008; Keys *et al.*, 2004).

The MALDI-TOF MS technique has also been used to identify fungi. In one study, 92.5 % of 247 clinical fungal isolates were correctly identified using this technique (Marklein *et al.*, 2009). The strains in this study included *Candida*, *Cryptococcus*, *Saccharomyces*, *Trichosporon*, *Geotrichum*, *Pichia*, and *Blastoschizomyces* spp. Some fungi were not identified because of the lack of a reference strain in the current database. Updating the database ensured that the remaining fungi were identified based on the protein profile (Marklein *et al.*, 2009).



**Figure 1.6:** MALDI-TOF MS spectra for six different bacterial species. The representative MALDI-TOF MS spectra of each of the six bacterial species (Panels A–F) representing *Escherichia coli* (Panel A, blue), *Klebsiella pneumoniae* (Panel B, yellow), *Salmonella enterica* (Panel C, purple), *Staphylococcus aureus* (Panel D, red), *Pseudomonas aeruginosa* (Panel E, light blue), and *Streptococcus* serogroup B (Panel F, green), respectively are displayed. The enlarged region with the  $m/z$  ranging from 3800 to 5600 is shown. Species-specific  $m/z$  values across spectra are indicated by different colors as noted above in arbitrary units. Image adapted from (Hsieh *et al.*, 2008).

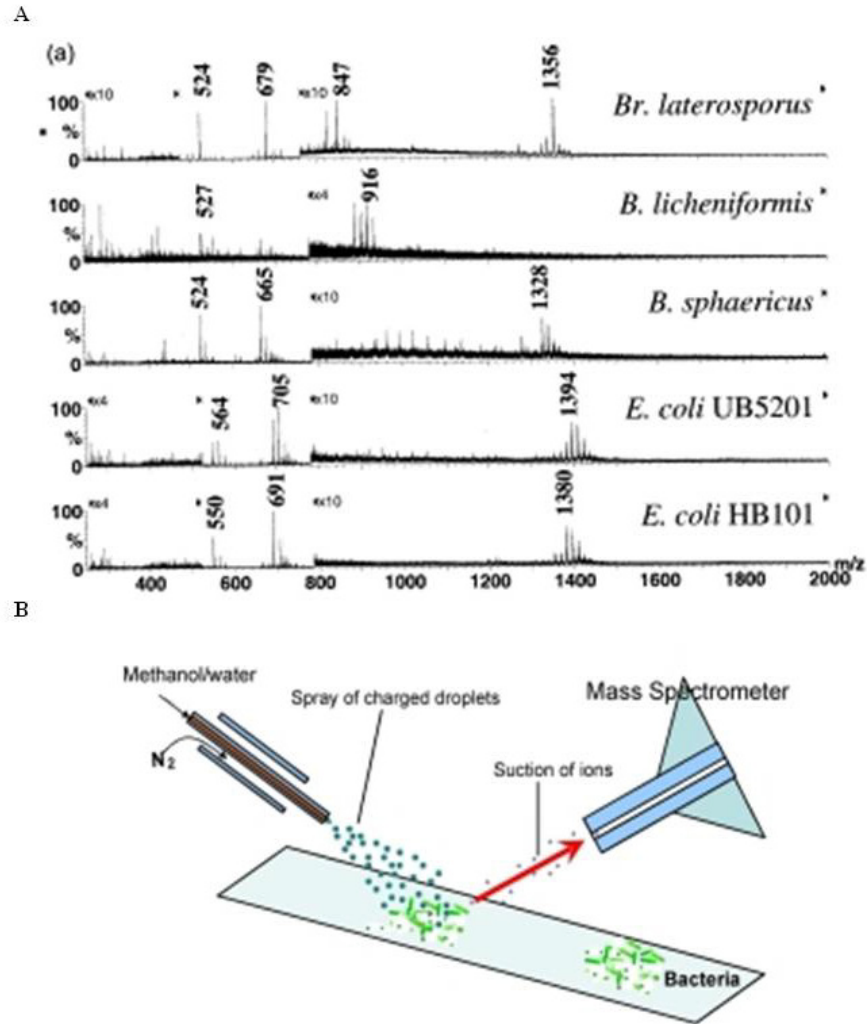
Though MALDI-TOF MS can only use cultured organisms, it has its own merit in that it is rapid and involves minimal sample preparation (Degand *et al.*, 2008). In addition, the technique may have the potential to be cost efficient when used in a clinical setting (Carbonnelle *et al.*, 2007; Tang & Stratton, 2006). Recently, a routine clinical laboratory identified 1600 strains using MALDI-TOF MS (Seng *et al.*, 2009). Identification was attained in two to two and half hours. The cost of identification was estimated to be between 22 and 32 % of the cost of traditional phenotypic assays (Seng *et al.*, 2009). Once the equipment is purchased, it is only a few cents to process a sample because of the minimal consumables that are needed and the lack of a special operator (Carbonnelle *et al.*, 2007). These studies highlight the role MALDI-TOF MS may play in future diagnostic laboratories.

**1.4.3 Electrospray ionization (ESI).** ESI-MS is another technique that can identify bacteria but unlike MALDI-TOF MS, the sample being analyzed can be in liquid form (Fenn *et al.*, 1989; Vaidyanathan *et al.*, 2002). In ESI-MS, the analyte that is either liquid or dissolved in a liquid solvent is dispersed by an electrospray, an instrument that uses electricity to disperse the liquid into fine aerosol. The aerosol enters the vacuum stage of the MS where it is further heated evaporating any residual solvent and the remaining ions are detected by the MS which determines the mass to charge ratio (Griffiths *et al.*, 2001). The technique has been used to characterize bacteria by analyzing specific cell fractions of the bacteria or by studying phospholipids, lipopolysaccharides, lipooligosaccharides, muramic acids, and proteins (Vaidyanathan *et al.*, 2002). In 1999, Goodacre *et al.* was the first to use ESI-MS to study intact whole

bacteria. In this study, three strains of *Bacillus cereus* and three strains of *Escherichia coli* were subjected to ESI-MS and produced mass spectra (Figure 1.7, Panel A) (Goodacre *et al.*, 1999). Each spectra contained fragments that correlated to fragments produced from bacterial phospholipids. In addition, spectra were also specific for the organism analyzed (Goodacre *et al.*, 1999). In a related study, *E. coli*, *Bacillus sphaericus*, *Bacillus licheniformis*, and *Brevibacillus laterosporus* were analyzed with flow injection ESI-MS. Again, the mass spectra were produced for each bacterium and the results were reproducible (Vaidyanathan *et al.*, 2001).

In recent years, ESI-MS technology has been modified so that the analyte can be in a solid form (Dole *et al.*, 1968; Fenn *et al.*, 1989). This technique, desorption ESI-MS (DESI-MS), ionizes an analyte that is placed on a solid surface using an electrospray (Figure 1.7, Panel B). The sample flows into the MS as it is ionized. Some analyte will stay on the surface longer and thus the ions are more separated before entering the MS. This method can also produce bacteria mass spectra and aid in bacterial identification and detection (Mohammed *et al.*, 2007). Both ESI-MS and DESI-MS have been used to analyze bioactivity of fungi but at this time it appears that specific detection of fungi using these techniques is limited (Pankiewicz *et al.*, 2006).

Many instruments have been coupled to MS and used to study various biomarkers of bacteria and fungi, enzyme pathways, and overall mass spectra of organisms. Mass spectrometry techniques can be rapid, efficient, and reproducible but they are rarely used in a clinical setting (Fox, 2006; Seng *et al.*, 2009). These techniques may begin to cross over to the clinical laboratory when there are cost-efficient instrumentation and reagents, developed reference databases, and validated standard operation protocols (Nassif, 2009).



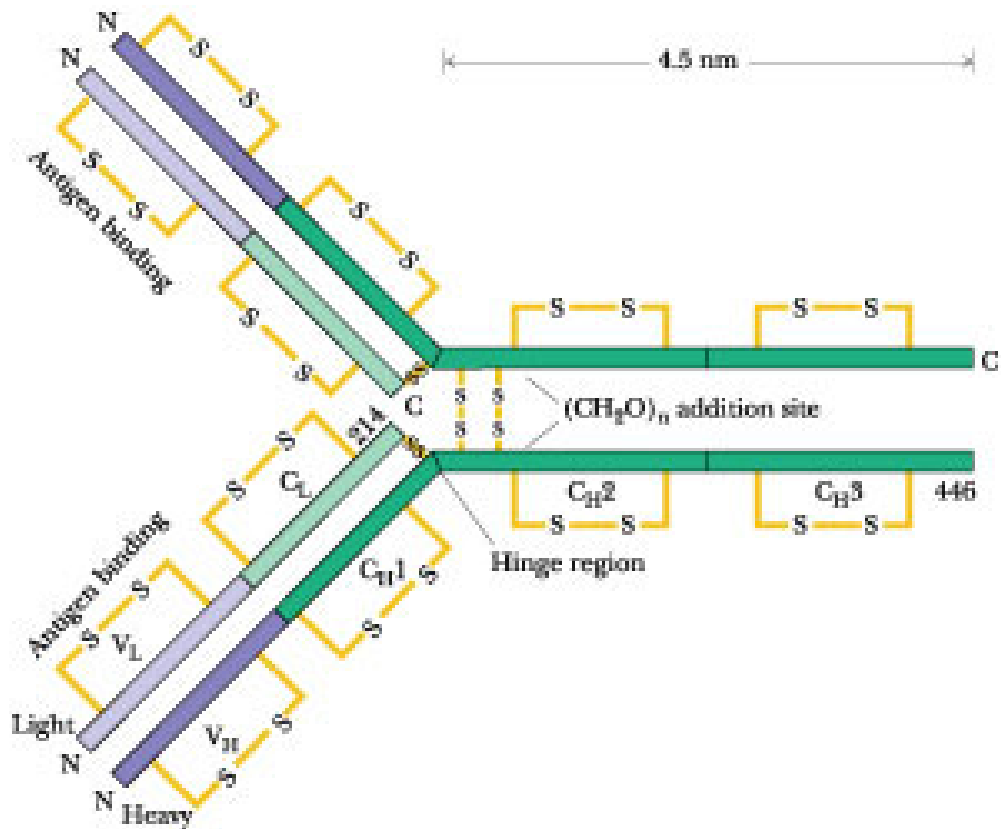
**Figure 1.7:** Electrospray ionization. Panel A shows the ESI spectra for *Bacillus sphaericus*, *Bacillus licheniformis*, *Brevibaacillus laterosporus*, and two strains of *Escherichia coli* (Goodacre *et al.*, 1999). Panel B depicts the desorption electrospray ionization method in which bacteria in an ambient environment can be identified using mass spectrometry. Water is sprayed into an electric field making hydronium ions which will transfer the extra proton to the bacterial sample causing the bacteria to be ionized. The ions are then vacuumed into the mass spectrometer where the mass to charge ( $m/z$ ) ratio is determined for the ions (Song *et al.*, 2007).



## 1.5 Detection through antibody-antigen interactions

Detection of microorganisms through antibody-antigen interaction is a main stay in most clinical microbiology laboratories. Bacteria and fungi detection systems have been developed based on the natural human immune response to these organisms. The next session gives a brief introduction to antibody response followed by two techniques, enzyme-linked immunosorbent assay (ELISA) and surface plasmon resonance (SPR) that heavily depend on the host immune response.

**1.5.1 The antibody response.** Antibodies are glycoproteins that are produced in response to the presence of foreign molecules (like bacteria and fungi) in the body (Edelman *et al.*, 1961; Edelman & Poulik, 1961; Edelman & Benacerraf, 1962; Nisonoff *et al.*, 1960; Porter, 1959). In general, antibodies are in the shape of a Y (Figure 1.8) and are composed of two identical copies of a polypeptide chain (heavy chain) and two identical copies of another polypeptide (light chain) (Berggard & Edelman, 1963; Nisonoff *et al.*, 1960). Humans have five classes of antibodies, IgG, IgM, IgA, IgE, and IgD which vary in the number of Y units and the types of heavy-chain polypeptide they contain (Ceppellini *et al.*, 1964). For example, IgG antibodies have one Y unit (monomer) which has three protein domains; the Fc domain is the stalk of the Y and the two Fab domains are the branches of the Y (Ceppellini *et al.*, 1964). At the top of the branch or the Fab domain is the antigen binding site (Figure 1.8). This site is composed of the variable regions of the heavy and the light chains and will interact with the foreign body (Ceppellini *et al.*, 1964). The hinge region can vary in structure giving rise to



**Figure 1.8:** IgG antibody. Two identical L chains are joined with two identical H chains. Each light (L) chain is held to a heavy (H) chain via an interchain disulfide bond. The variable regions of the four polypeptides lie at the ends of the arms of the Y-shaped molecule. These regions are responsible for the antigen recognition function of the antibody molecules. The hinge region varies based on the isotype of IgG present (Garrett & Grisham, 2010).

subtypes or isotypes within IgG molecules (Figure 1.8). IgG has four isotypes and they provide the majority of the antibody-based immunity against pathogens in the human (Jefferis *et al.*, 1985; Klein *et al.*, 1977). Like IgG, IgD, and IgE exist as monomers. IgD has one isotype and it functions as an antigen receptor on B cells (Pernis *et al.*, 1966;

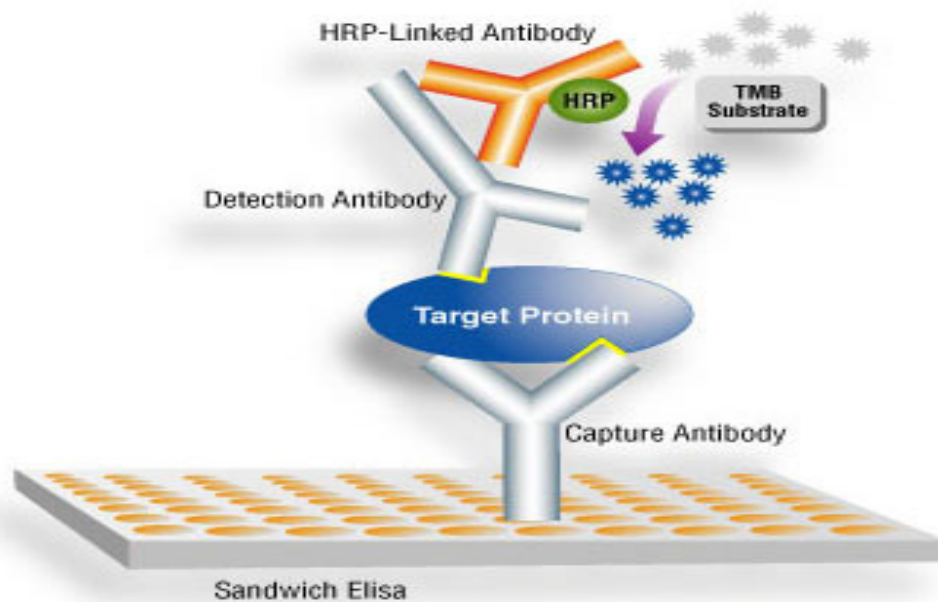
Takahashi *et al.*, 1982). The only isotype of IgE is involved in the immune response to allergies (Stanworth, 1969). The two isotypes of IgA exist as dimers and are found in mucosal areas such as the respiratory tract to help prevent colonization by pathogens (Heremans *et al.*, 1959). The last antibody, IgM, is a pentamer (Tomasi, 1973). The only isotype of IgM is expressed on the surface of B cells and is the primary antibody produced at the first encounter to an antigen (Sanders *et al.*, 1987).

All immunoglobulins have a site that interacts with antigens. Interactions can only occur if the antigen is close enough to the antibody Fab domain. This domain of 110-130 amino acids gives specificity to antibody binding (Novotný *et al.*, 1983). If the proper antigen is present it will form an antibody-antigen complex held together by many weak interactions including hydrogen bonds and van der Waals forces. Once the antibody has bound the antigen, other immune cells will help remove the antigen from the human (Gowans, 1971; Nossal & Ada, 1971).

**1.5.2 Enzyme-linked immunosorbent assay (ELISA).** The presence of an antigen such as bacteria or fungi or an antibody in a human sample can be detected using the same theory as the immune response mechanism described above. Antibody-antigen interactions can be specific. By detecting this interaction, the causative agent of an infection can be determined. To detect an antigen using ELISA, the sample being analyzed will first be fixed to a polystyrene microtiter plate. The antibody, for the organism that is being detected, is linked to an enzyme and then added to the plate. If the proper antigen is present, the enzyme-linked antibody will bind. Unbound antibodies are washed away with a mild detergent. Then, an enzymatic substrate is added to the plate.

Any enzyme present will react with the substrate producing a visible signal; thus indicating the presence of a certain bacteria or fungi in the sample. This reaction can be chromogenic or fluorogenic (Engvall & Perlmann, 1971; van Weemen & Schuurs, 1971; Voller *et al.*, 1978).

The ELISA technique can also be performed by immobilizing antibodies to the plate. A variation of this method is the sandwich ELISA (Figure 1.9) (Engvall & Perlmann, 1972; Miles & Hales, 1968; Voller *et al.*, 1978). In this assay, two antibodies



**Figure 1.9:** Sandwich ELISA. An antigen is detected using an assay in which horseradish peroxidase (HRP) (linked to the antibody that binds the antigen) oxidizes tetramethylbenzidine (TMB) to produce a blue color ([www.newenglandbiolabs.de](http://www.newenglandbiolabs.de)).

are used. The first antibody called capture antibodies are attached to a plate. The sample is added and if the specific antigen is present it will bind. After the remaining sample is washed away, a primary antibody is added which will also bind to the antigen present. Next, an enzyme-linked antibody is added which will bind to the primary antibody. During the last step, an enzyme is added which is converted to a detectable form (color or fluorescence). This signal tells the user if a bacteria or fungi are present in the sample and the intensity of the signal can indicate the quantity (Voller *et al.*, 1978).

Indirect ELISA is similar to the sandwich method except that first capture antibody is serum that contains an unknown concentration of the antibody (Engvall & Perlmann, 1972). The secondary antibody affixed to an enzyme is added and the sandwich ELISA method is followed (Engvall & Perlmann, 1972). A major disadvantage of the indirect ELISA is that the method of antigen immobilization is non-specific. Any amount of analyte present in the serum can stick to the plate. Thus indirect ELISA can only give determine if an antigen is present but not the abundance (Engvall & Perlmann, 1972).

Competitive ELISA uses the theory of competitive binding. In this method, unlabeled antibody is incubated with the antigen (Engvall & Perlmann, 1971). The complexes that form are added to an antigen coated well. The plate is washed and unbound antibody is removed. The antibody competes with the free antigen to bind to the plate fixed antigen. Then, the secondary antibody with the attached fluorophore is added and the signal is measured (Engvall & Perlmann, 1971). For competitive ELISA, the higher the original antigen concentration, the weaker the eventual signal. This type of

ELISA is well suited in situations in which the antigen is a low concentration or when using crude or impure samples (Jordan, 2005).

Reverse ELISA uses a solid phase made up of an immunosorbent polystyrene rod with 4-12 protruding fingerlike objects (ogives). The entire device is immersed in a test tube containing the collected sample. Then, the ogives are dipped into wells of a standard microplate. These wells have been pre-filled with ELISA reagents. The ogives are washed by a buffer in one well. The rods are then moved and the ogives are immersed in the next well which contains enzyme labeled antibodies. After conjugation, the ogives are incubated in a chromogenous sample. The develop color is read by spectrophotometer (Sakaguchi *et al.*, 1989). The ogives can be sensitized to the reagents and thus be able to detect more than one antigen at a time (Emmerich *et al.*, 2006). In addition, sample sizes can be optimized to achieve high levels of antigen detection (Emmerich *et al.*, 2006)

The ELISA technique has been performed on numerous types of samples for numerous organisms including *P. aeruginosa* and *A. fumigates* in CF sera (Brett *et al.*, 1988; Knutsen *et al.*, 1994). An ELISA test has been developed to diagnose chronic *P. aeruginosa* infections by fixing antigens from 17 serotypes of the bacteria to a plate (Pedersen *et al.*, 1987). The presence of serum IgG to the bacteria was then detected indicating the patient had *P. aeruginosa* infection (Pedersen *et al.*, 1987). Serum IgG antibodies to *Candida albicans* and *Aspergillus fumigatus* have also been detected using indirect ELISA (Przyklenk *et al.*, 1987). The ELISA method is not used to detect organisms directly from CF sputum but bacterial isolates and CF sera can be used with ELISA detection systems.

**1.5.3 Surface plasmon resonance (SPR).** Interactions between two surfaces, such as antibodies and antigens, can also be identified using SPR. This technique relies on the phenomenon of surface plasmons which are electron oscillations that exist at the interface between any two materials (Figure 1.10) (Otto, 1968; Wood, 1902, 1912). The interaction between two molecules can be measured by exciting the surface plasmons with light and then measuring the degree of light refraction (Kretschmann & Raether, 1968; Otto, 1968). SPR was originally developed for chemical sensing. In 1983, SPR began to be utilized for biosensing (Kretschmann & Raether, 1968).

One specific type of SPR biosensors is an immunosensor which detects interactions between antibodies fixed on a gold sensor and antigens (Geddes *et al.*, 1994; Mayo & Hallock, 1989). The interactions cause charge-density oscillations that occur at the interface of the two media; a SPR sensor measures the angle, wavelength, and the intensity of each oscillation. These factors vary when the oscillations at the interface between the two media change. For example, an antibody attached to the gold surface will affect the oscillations (Dudak & Boyaci, 2009; Geddes *et al.*, 1994). Most often the angular modulation (the change in degree of light refraction from the gold sensor) is measured. This measurement change when antigens bind to the antibody that is fixed to the gold (Dudak & Boyaci, 2009).

Immunosensors have been used to identify *E. coli* O157:H7 (Fratamico *et al.*, 1998; Meeusen *et al.*, 2005; Oh *et al.*, 2003; Su & Li, 2005; Subramanian *et al.*, 2006a; Taylor *et al.*, 2005), *Vibrio cholera* (Jyoung *et al.*, 2006), *Salmonella enteritis* (Bokken *et al.*, 2003), *Salmonella typhimurium* (Oh *et al.*, 2004a), *Salmonella paratyphi* (Oh *et al.*, 2004b), *Listeria monocytogenes* (Hearty *et al.*, 2006; Koubova *et al.*, 2001; Leonard *et*

*al.*, 2005), *Yersinia enterocolitica* (Oh *et al.*, 2005), *Campylobacter jejuni* (Taylor *et al.*, 2006), and *Staphylococcus aureus* (Subramanian *et al.*, 2006b).

The sensitivity is not as good as many other detection systems that are marketed; depending on the specific design of the SPR sensor, the detecting limits were between  $10^2$  to  $10^7$  cfu/ml (Dudak & Boyaci, 2009). The sensitivity of the assay can be changed by modifying the mechanism in which the antibodies bind to the gold and the antigens bind to the gold as well as specific instrument parameters (Dudak & Boyaci, 2009). Surface plasmon resonance has been used to detect some isolated CF pathogens such as *S. aureus* but has not been used to detect bacteria directly from sputum or bronchiolavage fluid (Subramanian *et al.*, 2006b).

## **1.6 Molecular analysis of isolates**

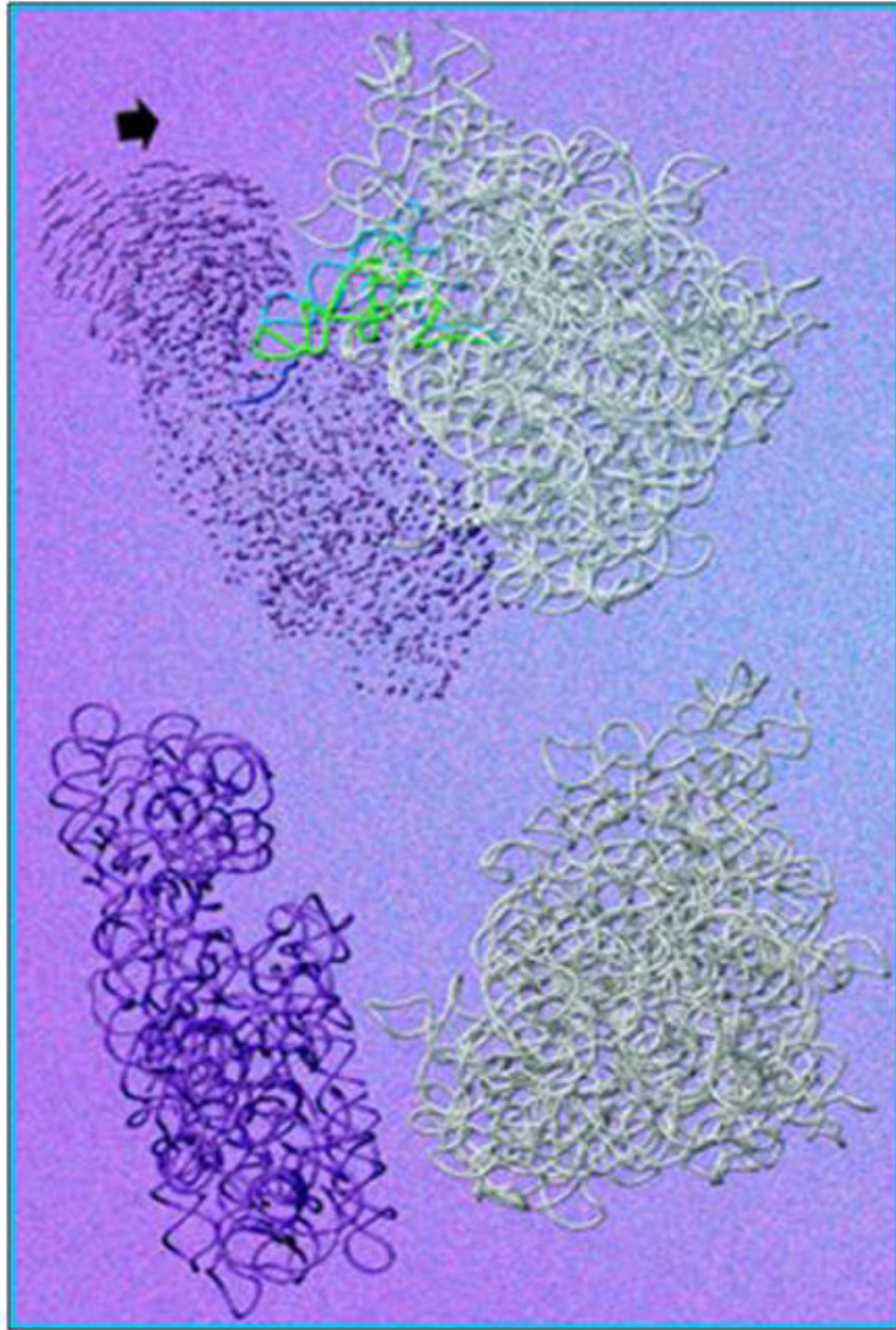
Technologies based on DNA are increasingly used as diagnostic tools. Most often bacteria and fungi are identified using ribosomal RNA. Specifically, molecular markers such as 16S ribosomal RNA (16S rRNA) and fungal internal transcribed spacer regions (ITS) contain information that can lead to the identification of pure isolates of many types of bacteria and fungi, respectively.

**1.6.1 Molecular markers.** Not all regions of DNA can be used to identify an organism. The regions that are used for speciation can be referred to as molecular markers and these regions share a few characteristics. First, markers are molecular chronometers meaning the region contains a DNA sequences that changes randomly in time but overall have a constant rate of evolution. This clocklike behavior can be seen in



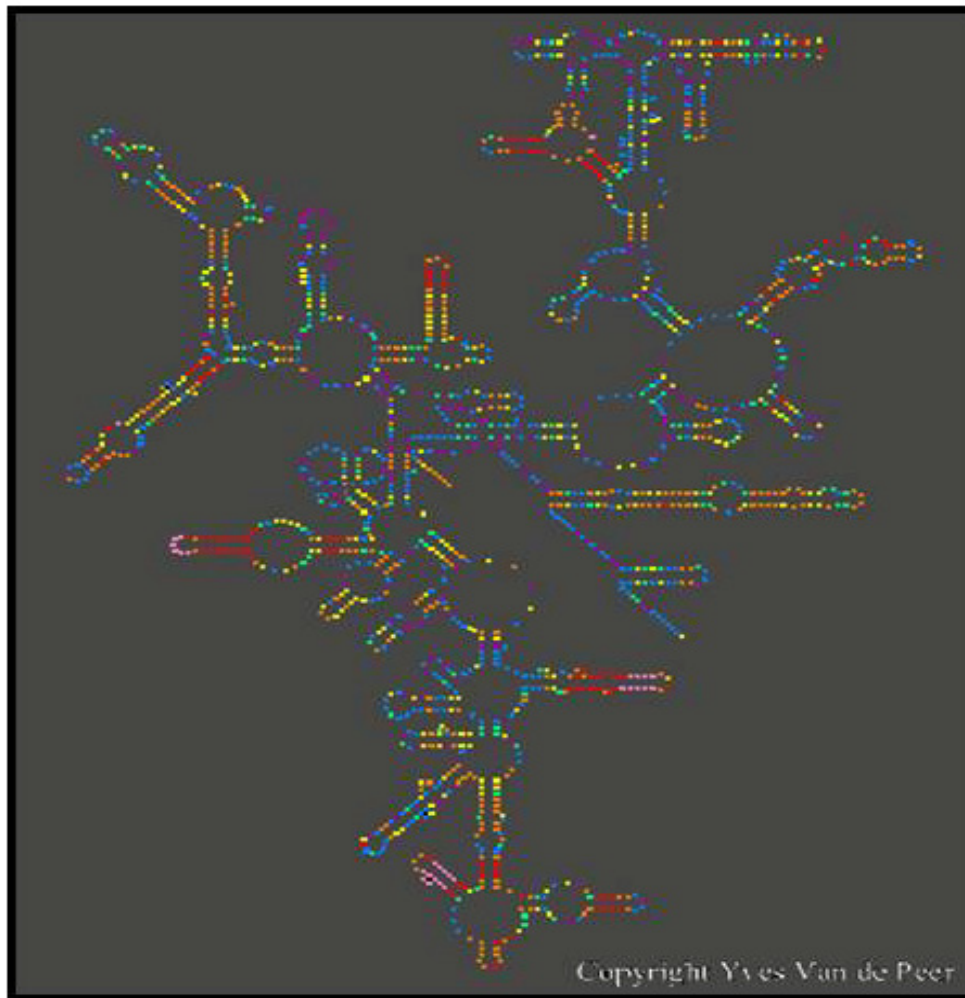
genes that have a high degree of functionally constancy (Woese, 1987). Molecular chronometers should also have a large phylogenetic range and therefore be present in most to all organisms (Olsen & Woese, 1993). Lastly, the molecular marker needs to have a large size with regions that are evolutionarily independent (Woese, 1987). One family of genes fills all of these requirements to become the “ultimate molecular chronometers”, the rRNA genes (Olsen & Woese, 1993). The rRNAs have a constant function, are found in all organisms, and the sequences change at different positions and at different rates (Olsen & Woese, 1993).

**1.6.2 Ribosomal RNA (rRNA).** Ribosomal RNA is directly associated with ribosomes (ribonucleoproteins) that play a role in protein synthesis (Lewin, 2002; Nomura & Erdmann, 1970; Nomura *et al.*, 1977). The basic form of the ribosome is conserved throughout organisms although changes in overall proportions and size of RNA and protein can be seen (Nomura & Erdmann, 1970; Nomura *et al.*, 1977). All ribosomes in a cell are identical and in prokaryotes they consist of a small (30S) subunit that is composed of 16S rRNA and 21 proteins and the large (50S) subunit (Figure 1.11). The larger unit contains 23S rRNA, 5S rRNA, and 31 proteins (Lewin, 2002; Nomura & Erdmann, 1970; Nomura *et al.*, 1977). In eukaryotes, ribosomes consist of a small (40S) subunit that is composed of 18S rRNA. The large subunit (60S) consists of the 5S, 5.8S, and the 28S subunits (Hillis & Dixon, 1991).

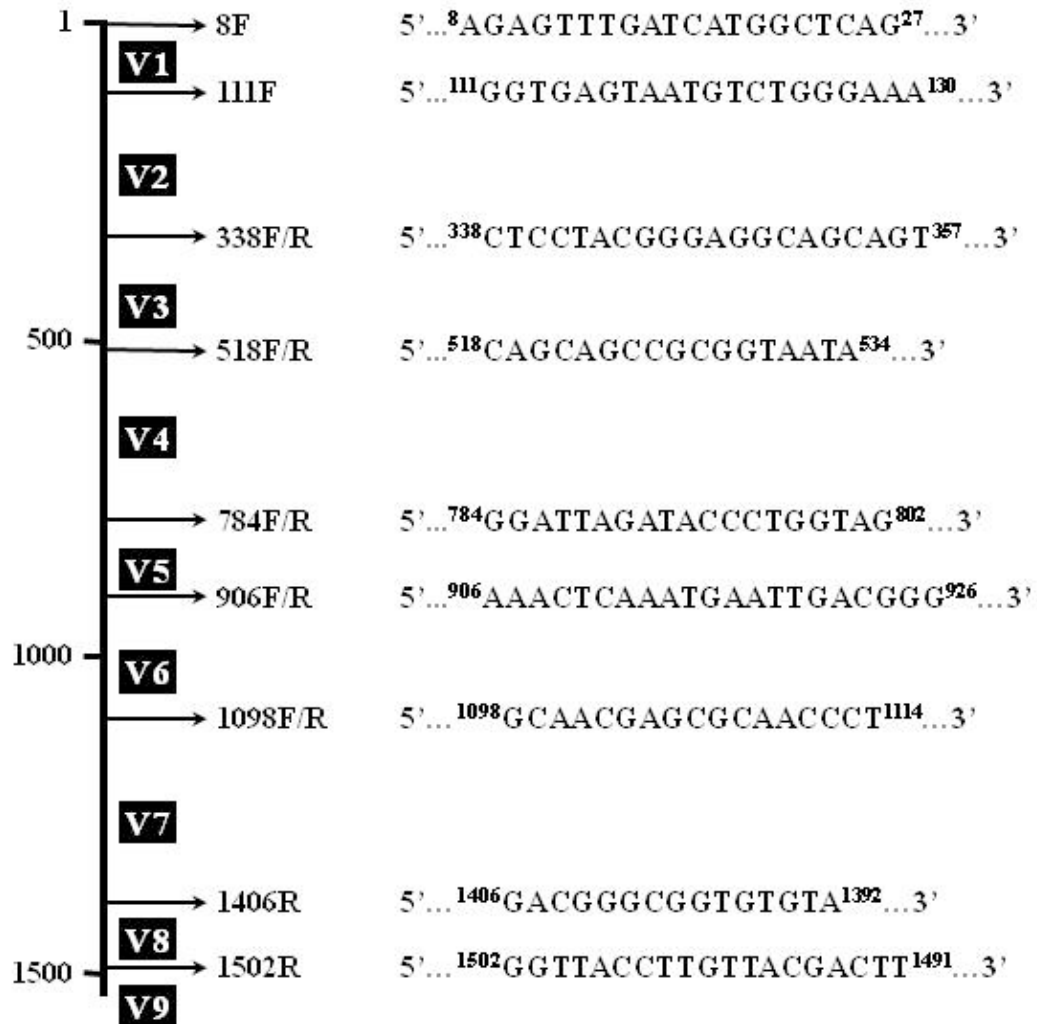


**Figure 1.11:** Ribosome subunits. The purple 30S subunit joins the gray 50S subunit to form the 70S ribosome (Yarris, 2003).

**1.6.3 16S ribosomal RNA (rRNA).** The 16S ribosomal subunit is most commonly used in studies of prokaryotic phylogeny (Klappenbach *et al.*, 2001). This gene is composed of alternating evolutionary conserved and variable regions (Figure 1.12 and 1.13) (van de Peer *et al.*, 1996). The conserved regions are involved in functions that are shared across the species, as such any mutation in this region would affect its ability to translate (Klappenbach *et al.*, 2001).

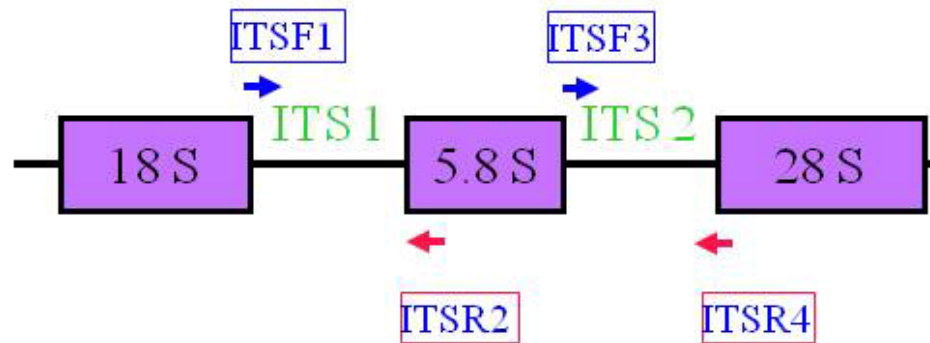


**Figure 1.12:** 16S rRNA secondary structure map. Purple indicates absolutely conserved regions, blue represents regions with a high amount of conservation. Red areas are the most variable (van de Peer *et al.*, 1996).



**Figure 1.13:** Schematic representation of the variable and conserved regions of the 16S rRNA genes, using *Escherichia coli* genome as a reference. The diagram illustrates the approximate positions of some of the variable regions of the gene via base number. The variable regions are interspersed with conserved regions. LH-PCR primer sequences are included (Doud *et al.*, 2009).

**1.6.4 Fungal internal transcribed spacer (ITS) regions.** Similarly, molecular diagnosis of the 18S rRNA (the fungal equivalent of 16S rRNA) can be used to identify fungi (Atkins & Clark, 2004; Melchers *et al.*, 1994). However, very often closely related species in a genus are not easily differentiated. To resolve closely related species, the ITS regions between the 18S and the 5.8S rRNA genes and between the 5.8S and 28S rRNA genes, referred to as *ITS1* and *ITS2*, respectively, can be used (Figure 1.14) (Leaw *et al.*, 2006). The contiguous ITS regions have also been successfully used to differentiate between the *Aspergillus* species, such as *Aspergillus flavus*, *Aspergillus fumigatus*, and *Aspergillus terreus*, (Henry *et al.*, 2000) and other medically important yeasts (Chen *et al.*, 2000; Chen *et al.*, 2001).



**Figure 1.14:** Fungal rRNA genes. ITS1 region is located between the 18S and 5.8S rRNA genes and can be amplified using the ITSF1 and the ITS2R primers. The ITS2 region is between the 5.8S and the 28S rRNA genes and can be amplified using the ITSF3 and ITSR4 primers. All primers bind to conserved regions within the rRNA genes and the reverse primers can be labeled with fluorophores for downstream applications.

**1.6.5 Polymerase chain reaction (PCR).** The PCR technique is often used to isolate and amplify the bacterial and fungal molecular markers for use in downstream applications (Erlich, 1991; Mullis *et al.*, 1986). This method heralded a new era for the detection and identification of various microorganisms in any sample. Thus, many techniques that study microorganisms are molecular-based using both universal and species-specific primers to select molecular markers (Liesack & Stackebrandt, 1992). Polymerase chain reaction can be used to identify the presence of one specific organism using one pair of specific primers. There have been numerous PCR assays developed to identify the common CF pathogens *P. aeruginosa* (da Silva Filho *et al.*, 1999), *H. influenzae* (Falla *et al.*, 1994), *Staphylococcus aureus* (Stepán *et al.*, 2001), and *Aspergillus fumigatus* (Reddy *et al.*, 1993).

**1.6.6 Multiplex PCR.** By combining pairs of primers and optimizing the PCR reaction, multiple organisms can be detected in the same reaction. Multiple PCR assays have been developed for many different combinations of CF pathogen including but not limited to: *P. aeruginosa*, *Burkholderia cenocepacia*, and *Stenotrophomonas maltophilia* (da Silva Filho *et al.*, 2004); *Streptococcus pneumoniae*, *H. influenzae*, *Mycoplasma pneumonia*, and *Chlamydophila pneumonia* (Stralin *et al.*, 2005); and *Aspergillus fumigatus*, *Aspergillus flavus*, *Aspergillus niger*, and *Aspergillus terreus* (Logotheti *et al.*, 2009). By using multiplex PCR, multiple organisms can be identified in a small time frame (Henegariu *et al.*, 1997). Establishing a multiplex PCR assay can be difficult due to uneven or lack of amplification for certain targets and difficulties reproducing the

results. Thus, optimization of PCR protocols and carefully planned PCR primers are crucial in the successfulness of a multiplex reaction (Henegariu *et al.*, 1997).

### **1.7 Molecular tools for community studies**

The aforementioned techniques more readily identify an organism if it has been isolated from a sample through culturing. Though some techniques such as multiplex PCR can identify multiple organisms in a sample, the entire diversity of microbial communities in complex samples such as water, soil, and CF sputa cannot be determined. Community profiling techniques such as terminal restriction fragment length polymorphism – polymerase chain reaction (T-RFLP) and amplicon length heterogeneity – PCR (LH-PCR) are molecular tools that identify organisms in mixed samples without the need to culture or isolate from the sample.

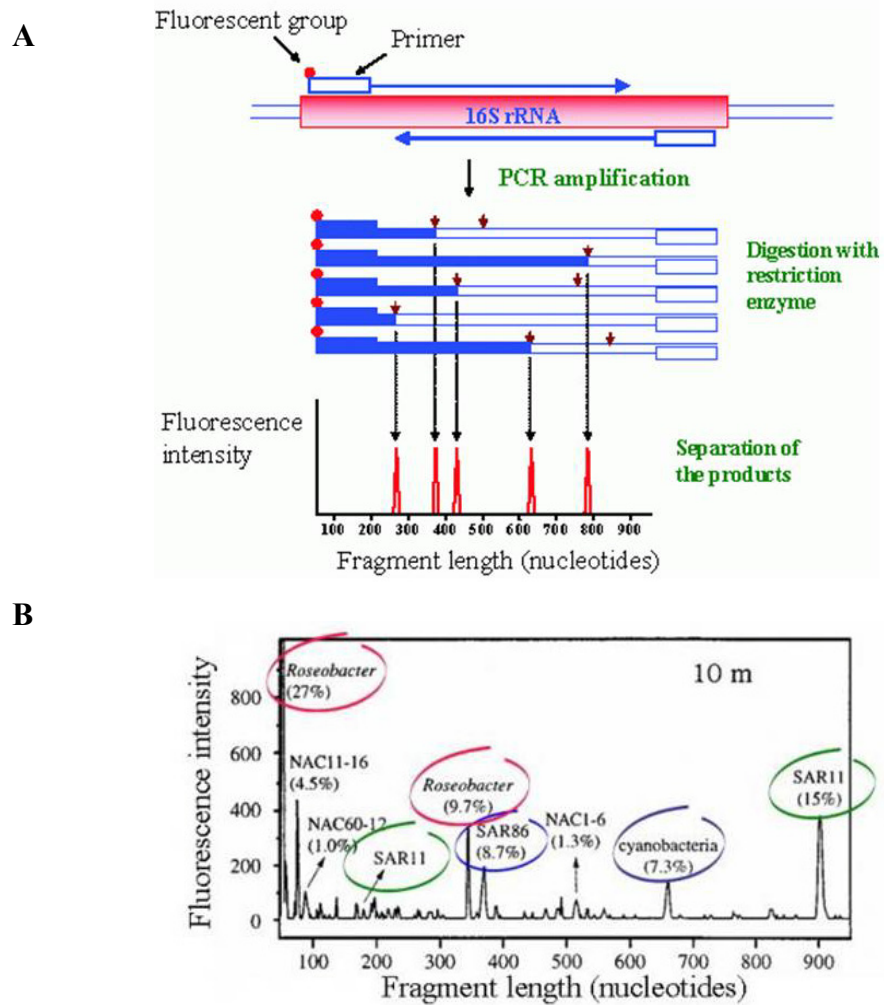
Community profiling techniques use variation within the molecular marker as described earlier to determine the composition of the community within a sample. For bacteria, the lengths and sequences of one or more of the nine variable regions of 16S rRNA can lead to the identification of bacteria present in a complex community. Similarly, the 18S or the ITS regions can be used to identify members in a fungal community. The two techniques, T-RFLP and LH-PCR generate PCR products of different lengths which can provide a community profile and can putatively identify individual organisms at the strain, species or genus level (Klappenbach *et al.*, 2001). (Klappenbach *et al.*, 2001; Suzuki *et al.*, 1998). In addition, the sequence of the fragments can be determined using high throughput sequencing which enables rapid sequencing of the amplicons (bacterial and fungal, with the use of appropriate primers)

which would lead to the rapid understanding of the community structure of any complex niche.

**1.7.1 Terminal restriction fragment analysis (T-RFLP) analysis.** This PCR based technique relies on the inherent variation of the sequence of a molecular marker (Dunbar *et al.*, 2001) and is the most widely used method in identifying phylogenetic specificity in bacterial communities (Hiraishi *et al.*, 2000). The amplicons are generated using one primer that is fluorescently end-labeled, followed by restriction enzyme digestion, and detection of the terminal restriction fragment by an automated DNA sequencer or capillary electrophoresis (Figure 1.15, Panel A) (Hiraishi *et al.*, 2000). The resulting output consists of a microbial profile where each detected length is that of specific fragments from the digested PCR product. Each length represents one or more bacteria that have the same terminal restriction fragment length (Figure 1.15, Panel B). The profiles can be used for community differentiation, identification of specific organisms in populations, and comparison of the relative phylotype richness and community structure (Dunbar *et al.*, 2001).

This method has been successful in differentiation of bacterial and fungal communities present in many environments including marine samples, soil samples, and sputum samples from CF patients (Dickie & FitzJohn, 2007; Hiraishi *et al.*, 2000; Mills *et al.*, 2003; Rogers *et al.*, 2003). Rogers *et al.* analyzed T-RFLP amplicons of CF patient sputa and bronchoscopy samples using a computer program called MapSort





**Figure 1.15:** Terminal restriction fragment length polymorphism (T-RFLP). Panel A illustrates the T-RFLP process starting with PCR, followed by restriction digest and ending with product separation (<http://nodens.ceab.csic.es/t-rfpred/>). Panel B is a T-RFLP of a mixed DNA sample. Each peak represents at least one bacteria (Gonzalez *et al.*, 2000).

(Wisconsin Package version 10.3; Accelrys) that has a database which contains restriction patterns and lengths of fragments generated for known 16S rRNA bacterial sequences (Rogers *et al.*, 2003). The analysis inferred the presence of *P. aeruginosa*, *Burkholderia cenocepacia*, *Staphylococcus aureus*, and *H. influenzae* in the CF samples (Rogers *et al.*, 2003). T-RFLP analysis of the fungal community in the CF lung does not appear to have been performed yet.

The T-RFLP method is fast and data can be easily replicated for statistical analysis. The major disadvantage of T-RFLP is that many bacteria produce similar fragment sizes and thus not all peaks in the profiles are species-specific. Some peaks may even represent more than one genus (Dunbar *et al.*, 2001; Rogers *et al.*, 2003). There are also inherent problems using restriction enzymes, such as incomplete digestion, which can produce DNA fragments that do not correlate to the correct bacterium (Mills *et al.*, 2003). Therefore to achieve better identification of the organism, further analysis like sequencing of the 16S rRNA gene must be performed.

**1.7.2 Amplicon length heterogeneity (LH).** The LH technique analyzes microbial populations based on the lengths of generated PCR products produced from the hypervariable regions of the 16S rRNA or the fungal ITS regions (Bernhard *et al.*, 2005; Martin & Rygiewicz, 2005; Moreno *et al.*, 2006; Ritchie *et al.*, 2000; Suzuki *et al.*, 1997; Yang *et al.*, 2006). Profiles from each region can be produced and used to understand the community. These profiles represent the minimum diversity of bacteria present within the eubacterial or fungal community. The profiles contain peaks at specific amplicon lengths, which is representative of the number of nucleotides in the hypervariable region

between the conserved regions. The peak heights are representative of the relative abundance of amplicons for that particular length present in the community.

The organisms present in a community are determined by using a database which contains the expected amplicon sizes for bacteria and fungi for given a certain set of primers (Doud *et al.*, 2009). The fragment lengths in the sample profile are compared against the database to identify the putative organisms. A profile resulting from this analysis infers the presence of certain organisms and the definitive absence of others. In cases where the amplicon length is not species-specific, it is often genus-specific (Doud *et al.*, 2009).

The LH profiles can also be used to compare community profiles from multiple samples. Previous research using LH profiles has shown that the compositions of bacterial communities are highly specific to the environment in which they are found (Mills *et al.*, 2003; Moreno *et al.*, 2006). Changes in the community's niche can drastically influence bacteria and thus add specificity to the profile of a bacterial community showing that the overall bacterial community has many unique features from sample to sample (Mills *et al.*, 2003; Moreno *et al.*, 2006).

The main advantages of LH-PCR are that it rapidly surveys relative gene frequencies within complex mixtures of DNA, is reproducible, requires small sample sizes, and can be performed simultaneously with many samples (Suzuki *et al.*, 1998). The LH profiles provide information about the members of the entire bacterial/fungal community (not just specific isolates) and their relative abundance. This data allows one to make taxonomic inferences and sample comparisons (Suzuki *et al.*, 1998). A major disadvantage of this technique is that one amplicon in the profile can represent more than

one bacterium or fungus (Manter & Vivanco, 2007). Therefore, identification at the species level cannot be guaranteed. This is also true with many length-based molecular techniques such as T-RFLP. However, the fragments are discrete ‘units’ of information that can be used for comparative analyses (Dunbar *et al.*, 2001). Analysis of different combinations of the 16S rRNA variable regions and ITS regions can increase the power of microbial detection and sample discrimination and lead to more definitive identification (Yang *et al.*, 2006).

The LH-PCR method was first used in several ecological research projects to compare microbial communities between samples and identify members within one community (Mills *et al.*, 2003; Mills *et al.*, 2006; Moreno *et al.*, 2006; Yang *et al.*, 2006). Multiples types of soil and water samples have been profiled using LH and various regions of the 16S rRNA gene (Mills *et al.*, 2003; Moreno *et al.*, 2006; Ritchie *et al.*, 2000; Suzuki *et al.*, 1998). LH has also been used to study bacterial communities present in the CF lung. Fourteen CF sputum samples were analyzed using LH-PCR for the presence of eubacteria (Rogers *et al.*, 2003). The raw data generated from the Genetic Analyzer was first processed with corresponding software such as GeneImageIR v.3.56 (Scanalytics, Fairfax, Va.) (Rogers *et al.*, 2003). To presumptively identify the bacteria present in the CF samples, the fragment lengths were compared with a database of theoretical fragment lengths constructed using GAP (Wisconsin Package version 10.3; Accelrys, Inc., San Diego, CA) (Rogers *et al.*, 2003). For example, *P. aeruginosa* was presumptively identified in all 14 CF samples, five of which were confirmed by cloning and sequencing (Rogers *et al.*, 2003).

To expand the use of LH-PCR in clinical settings, Bjerketorp *et al.* combined it with a lab-on-chip (LOC) system, which is used for sizing and quantifying DNA, to analyze samples containing mixtures of known human gut microbes (Bjerketorp *et al.*, 2008). An Agilent 2100 Bioanalyzer, a bench-top instrument that uses microfluidics-based separation, was used to detect the LH fragments. This modified method allows LH-PCR to be more affordable and faster and thus more convenient and suitable for clinical and diagnostic situations (Bjerketorp *et al.*, 2008). To test this system, samples containing mixtures of human gut microbes and known human gut bacteria isolates were analyzed using both LOC and a capillary electrophoresis based genetic analyzer. The latter method had a higher resolution and was thus able to resolve more peaks or fragments from one another. It is important to clearly separate PCR fragments as that LH identification is based on the lengths of PCR products. Single base pair length differences are known to occur between species and even at the genus level. Though the level of resolution for the LOC LH-PCR technique is limiting, the technique is rapid, economical and easier to analyze than the traditional system. Future modifications may improve the resolution making it more useful for clinical diagnosis (Bjerketorp *et al.*, 2008).

Amplicon length heterogeneity studies of fungal communities from CF lungs have not been performed but the lengths of ITS 1, ITS 2 and ITS 1-4 have been determined via sequencing for many fungi that are known to colonize the CF lung such as various species of *Candida* and *Aspergillus* (Henry *et al.*, 2000; Leaw *et al.*, 2006).

Community profiling techniques such as T-RFLP and LH-PCR can be used to study both bacterial and fungal communities. The profiles generated provide a snapshot

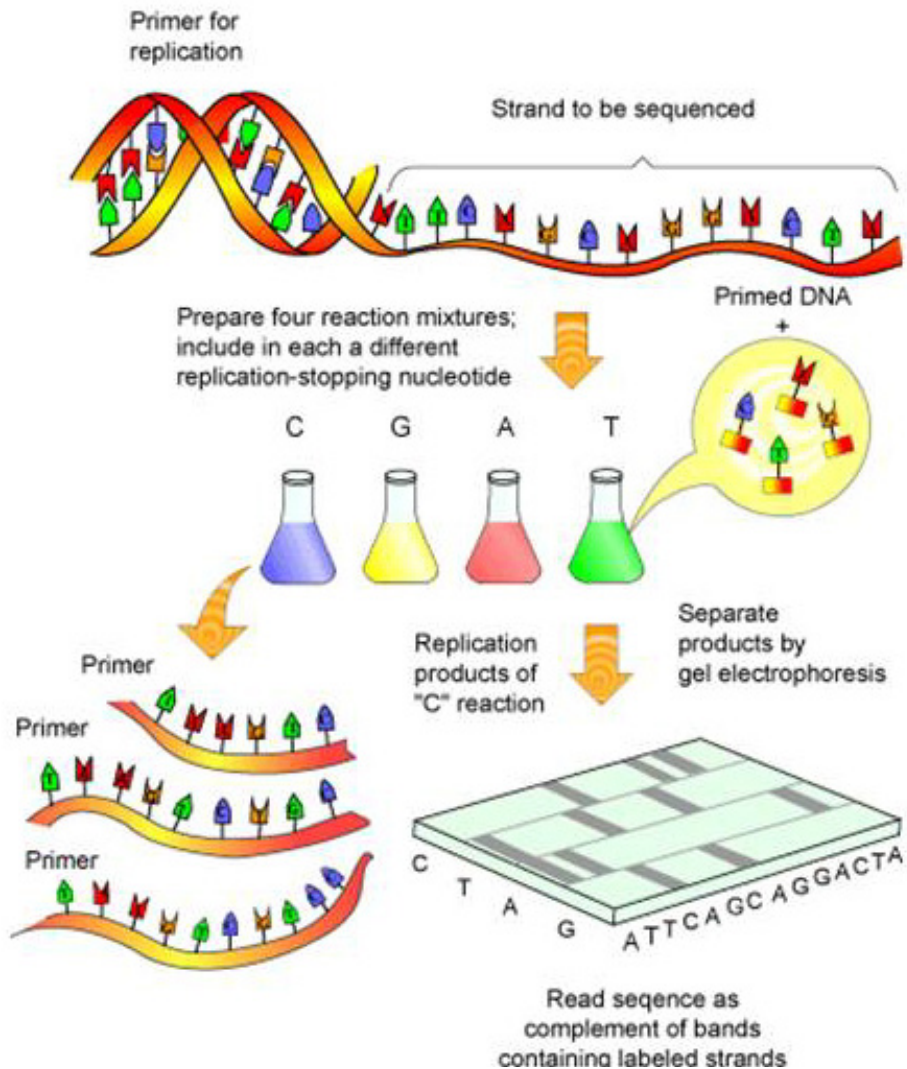
of the microorganism community as a whole. The identity of community members may be determined with the aid of reference database; but, due to shared amplicon lengths for certain organisms, some community members will not be identified. Through the use of sequencing, these organisms can be identified.

## **1.8 Sequencing**

Many organisms may share the amplicon length for a certain molecular marker but they rarely share the same sequence for the same given marker. Sequencing of the entire or partial 16S rRNA gene or 18S gene and the associated ITS regions is imperative in identifying a bacterium or fungus with near certainty.

**1.8.1 Sanger sequencing.** The most common method of sequencing is the Sanger Method developed in 1977 (Sanger *et al.*, 1977). This method uses a single strand of DNA. The DNA is divided into four reactions, each reaction containing four standard deoxynucleotides (dNTPs), one dideoxynucleotide (ddNTP), and a polymerase (Figure 1.16). The complementary strand of DNA is made with the dNTPs and when a ddNTP is incorporated into the growing strand, the chain extension will terminate. These different sized newly synthesized fragments are then separated by size and visualized (using radioactivity) to determine the sequence of the DNA (Sanger *et al.*, 1977). This technique can be modified so that all sequencing reactions occur in one tube through the use of fluorescently labeled dideoxynucleotides. The fragments are separated by size and sequence is determined by the fluorescence. Once the sequences are generated they are compared to known 16S rRNA, 18S, or ITS sequences stored in the Ribosomal Database

Project II (Maidak *et al.*, 1999), Greengenes (DeSantis *et al.*, 2006), and GenBank (Benson *et al.*, 2008) to identify organisms in any sample including the CF lung (Bosshard *et al.*, 2006; Harris *et al.*, 2007).



**Figure 1.16:** Sanger sequencing. Single stranded DNA is amplified using fluorescently labeled ddNTPs and unlabelled dNTPs. Incorporation of a ddNTP causes the chain lengthening to stop, leaving the fragment fluorescently labeled, which are separated by polyacrylamide gel electrophoresis (Kae, 2009).

Sanger sequencing is able to identify organisms if the correct marker is being sequenced. For some bacteria, partial sequencing of the 16S rRNA gene would lead to identification, for others the entire gene would need to be analyzed. The same could be said for fungi; that is, only the entire 18S rRNA gene sequence can be used to distinguish between some fungi. Some bacteria or fungi can only be identified using the more Sequencing isolates can be performed in a timely manner and the data produced is quite easy to analyze especially with the use of commercial sequencing kits (Vliegen *et al.*, 2006).

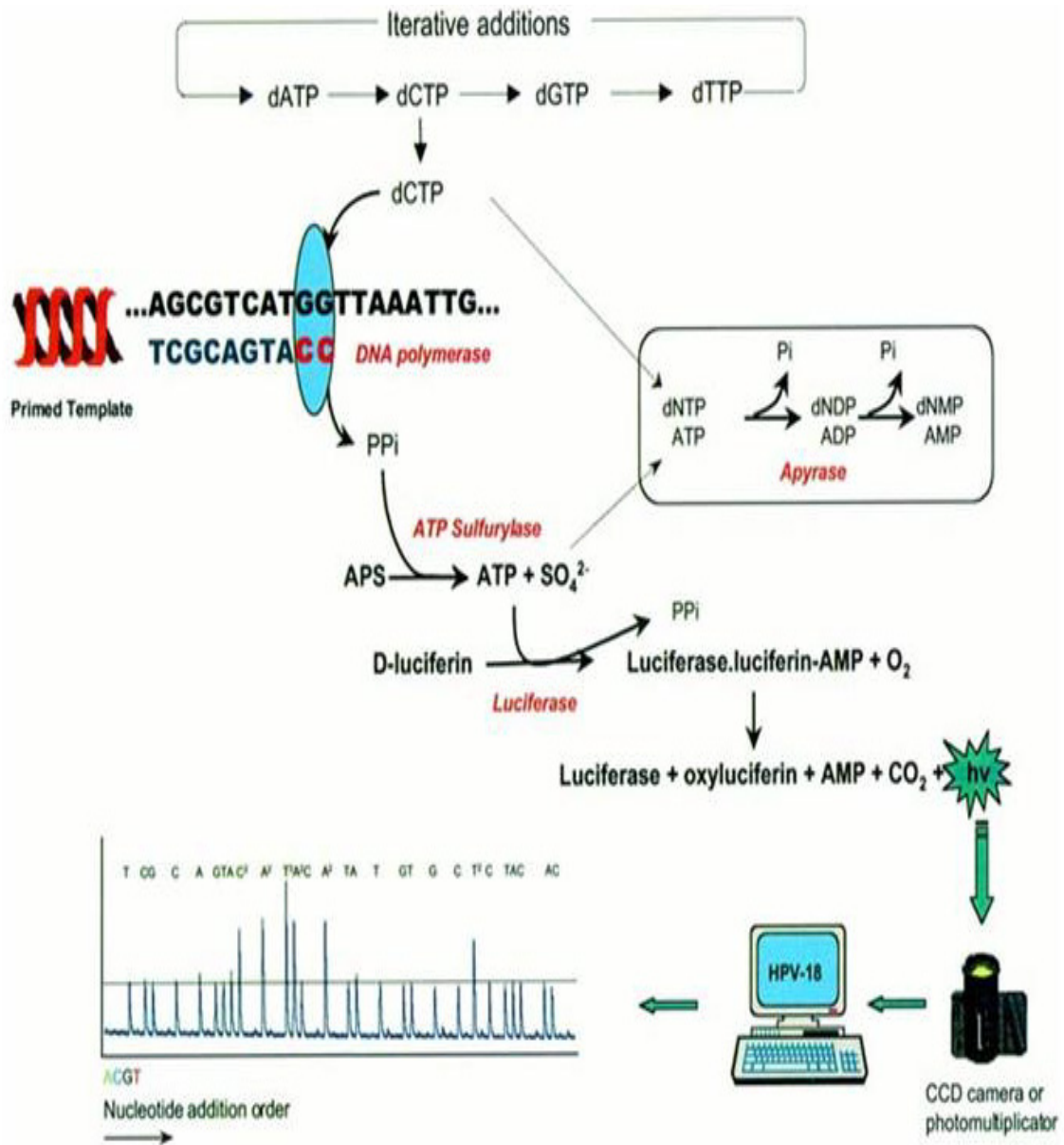
Sanger sequencing has been used to identify the organisms present in CF children bronchiolavage (BAL) samples (Harris *et al.*, 2007). As that the total DNA extracted is a mixture of DNA (fungal, bacterial, and human), the bacterial DNA was isolated using RFLP-PCR and the resulting PCR products were separated using bacterial cloning methods. The individual RFLP fragments were then sequenced and bacteria such as *P. aeruginosa*, *Staphylococcus aureus*, *Stenotrophomonas maltophilia*, and *H. influenzae* were found to be present in the BAL sample (Harris *et al.*, 2007). Novel pathogens from the genus *Lysobacter*, *Coxiellaceae*, and *Rickettsiales* were also sequenced (Harris *et al.*, 2007).

Fungi have also been identified in CF samples through PCR and Sanger sequencing (Bouchara *et al.*, 2009). By sequencing the entire PCR product from the ITS region, *A. fumigatus* and *C. albicans* were identified in CF samples (Bouchara *et al.*, 2009). It is more common to use specie specific PCR to detect fungi in CF samples because few fungal isolates have been sequenced (Williamson *et al.*, 2001).



Sanger sequencing is capable of providing the identification of isolates, but the high cost of reagents and labor may be too expensive for wide-spread clinical use (Patel, 2001). This method can also be used to identify members of a complex community, but it requires additional protocols, reagents, and time. With traditional sequencing methods, cloning must be performed to isolate individual 16S rRNA genes or fungal ITS regions which were amplified by PCR. The PCR product from a complex community will contain multiple copies of the same gene and further screening such as restriction digest must be performed to ensure that the same PCR product is not repetitively sequenced thereby wasting time, reagents, and money. The cloning and screening process also increases the likelihood that certain members of the community are never sequenced and the community composition is not entirely determined. Though possible, traditional sequencing is a time consuming, arduous method when addressing microorganism community based questions.

**1.8.2 Pyrosequencing.** New developments in sequencing technologies are revolutionizing the way microbial communities are being studied (Mardis, 2008; Schuster, 2008). Pyrosequencing relies on a process referred to as sequencing-by-synthesis, a technique that allows for real-time monitoring of DNA synthesis (Ahmadian *et al.*, 2006; Ronaghi *et al.*, 1998). Pyrosequencing is based upon the principle that pyrophosphate (PPi) is released when the DNA polymerase adds a nucleotide to the growing complementary strand (Figure 1.17). The PPi is converted to adenosine triphosphate (ATP) that is used as a substrate in a chemical reaction that results in visible



**Figure 1.17:** Pyrosequencing reaction. Schematic representation of the automated pyrosequencing (sequence-by-synthesis) method. The addition of a nucleotide to the growing strand releases pyrophosphate (PP<sub>i</sub>) which is converted to ATP by ATP sulfurylase. The ATP then catalyzes a luciferase mediated conversion of luciferin to oxyluciferin that produces light and is measured by a camera (Gharizadeh *et al.*, 2001).

light emission. The detectable amount of light produced is relative to the amount of synthesis (Ronaghi *et al.*, 1998). The dNTPs are added sequentially and light is only emitted if the dNTP added to the reaction is the next one to be incorporated to the growing strand. Like the Sanger method, pyrosequencing can only sequence individual PCR products and thus must be used in conjunction with cloning to study microbial communities (Ronaghi *et al.*, 1998). The method can produce a large amount of sequencing reads in a small amount of time. Thus, it is more efficient to sequence a large clone library using pyrosequencing rather than the Sanger method.

Pyrosequencing techniques that allow faster sequencing at a lower cost, are opening doors for many laboratories to use sequence data for microbial identification. Pyrosequencing of the first and third variable region of the 16S rRNA gene has been used to identify bacterial isolates of *Helicobacter pylori* from secretions, indwelling catheters and prosthetic devices, urine, blood and faecal specimens (Jonasson *et al.*, 2002). In addition, *Acinetobacter* and *Haemophilus* isolates from gastric biopsies were sequenced for the same regions (Monstein *et al.*, 2001). Most recently, clinical isolates of mycobacteria was identified using the V1 region (Heller *et al.*, 2008). A year later, antibiotic resistance of mycobacteria could be determined by pyrosequencing the resistance-determining regions of three genes and looking for specific mutations (Bravo *et al.*, 2009).

Importantly, pyrosequencing surpassed traditional methods of detection in a clinical setting by identifying 90 % of the isolates at least at the genus level (Luna *et al.*, 2007). The remaining 10 % could not be identified due to the short sequencing reads, a clear drawback of pyrosequencing (Luna *et al.*, 2007). Pyrosequencing may help

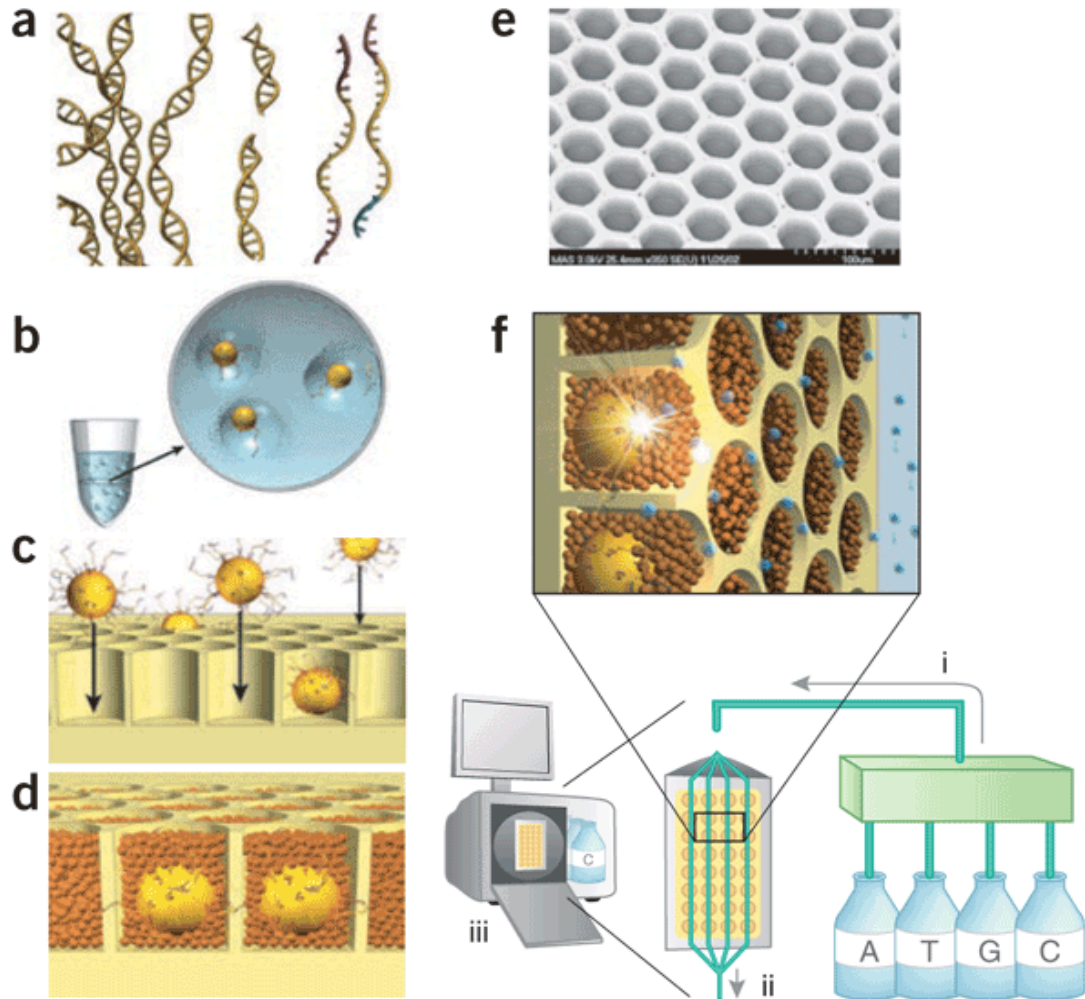
bacterial identification in samples that do not lend themselves to polyphasic approaches, such as phenotypic assays (Heller *et al.*, 2008; Tuohy *et al.*, 2005). This technique has also been shown to clearly distinguish between multiple species of *Mycobacterium*. Pyrosequencing can also identify fungi including different species and strains of *Candida*, a known CF pathogen (Boyanton *et al.*, 2008).

With the additional step of cloning, pyrosequencing can be used to identify organisms in a complex sample. One study used cloning followed by pyrosequencing to identify bacteria present in 13 adult CF sputa samples (Kolak *et al.*, 2003). Sequences were determined by the amplification of the V3 and V6 region of the 16S rRNA gene. Known CF pathogens were sequenced and the results concurred with the culture identifications (Kolak *et al.*, 2003). Sanger sequencing and pyrosequencing were also compared in a later study. Both techniques were used to analyze the same sputum sample from a 13 month old CF girl (Armougom *et al.*, 2009). Both techniques identified many of the same bacteria but pyrosequencing showed greater species richness and a more reliable estimate on bacteria abundance. The authors concluded that pyrosequencing is reliable and with newer technologies, the short reads will no longer be an issue (Armougom *et al.*, 2009). The fungal community in CF samples has yet to be analyzed with pyrosequencing.

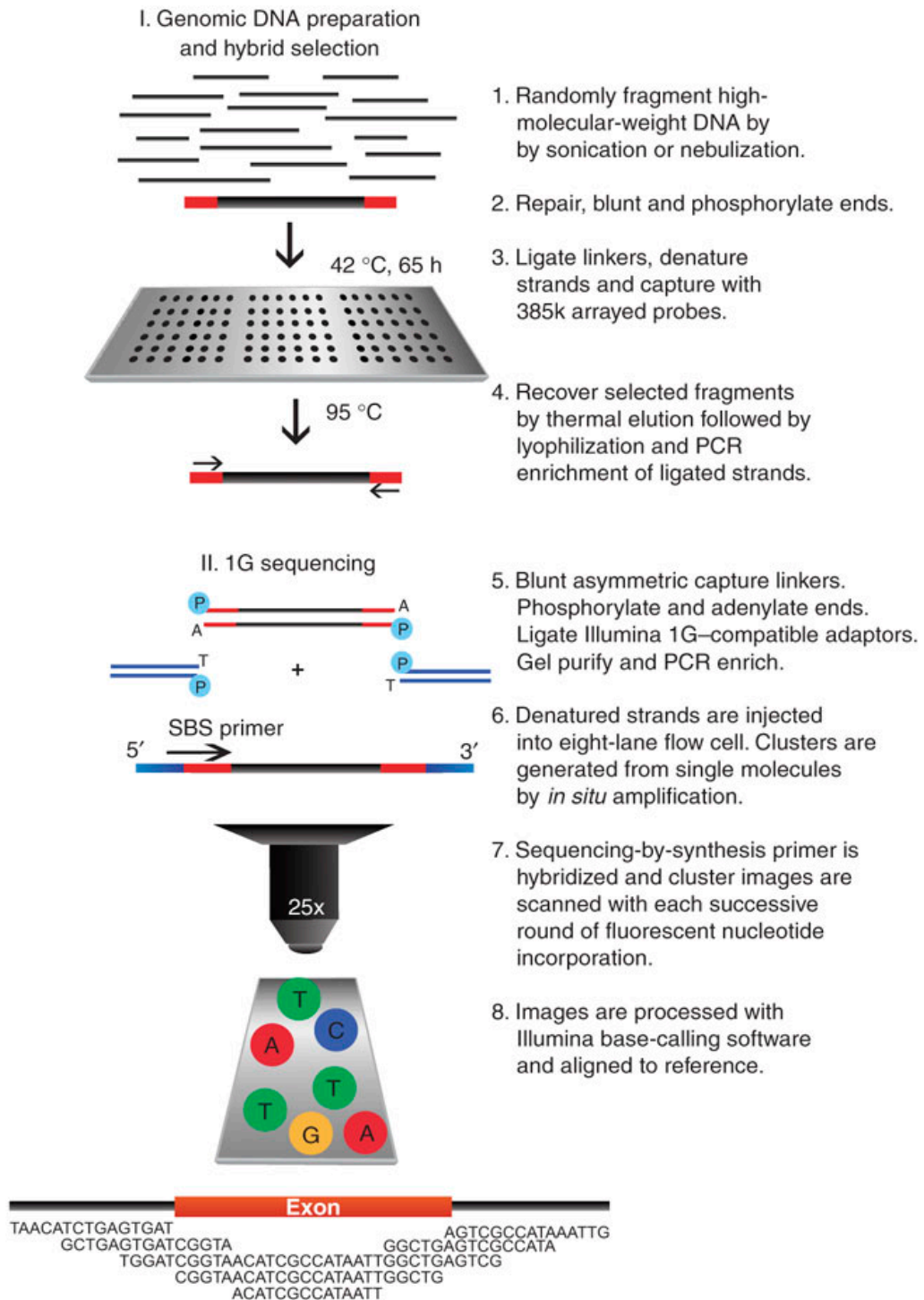
**1.8.3 454 sequencing.** Deep sequencing technologies are capable of sequencing a whole genome in a matter of days. There are multiple platforms that are capable of performing high throughput, parallel sequencing of a single isolate or mixed samples. Previously cloning was necessary to provide clean sequence reads from mixed DNA

samples. To circumvent the need for cloning, Roche's 454 sequencing, which performs many pyrophosphate-sequencing reactions in parallel, was developed (Figure 1.18) (Margulies *et al.*, 2005). The 454 sequencing combines an emulsion-based method that isolates and amplifies DNA fragments *in vitro* with an instrument that performs pyrosequencing in picoliter-sized wells (Margulies *et al.*, 2005). The reactions are resolved on a Genome Sequencer FLX which reads 200-300 bases and in one run can read up to 400,000 bases (Patrick, 2007). This method has been used to study the microbial diversity of the deep sea (Sogin *et al.*, 2006) and the metagenome found in honey bees, which led to the discovery of a possible causative agent of colony collapse disorder (Cox-Foster *et al.*, 2007). Members of the human oral bacterial community have been identified using 454 sequencing (Zaura *et al.*, 2009). Fungal communities have also been sequenced in soil and the human mouth (Ghannoum *et al.*, 2010; Lauber *et al.*, 2009).

**1.8.4 Illumina.** Another platform that is capable of high throughput, parallel sequencing is the Illumina Genome Analyzer (previously called Solexa). Illumina sequencing technology amplifies randomly fragmented genomic DNA that are attached to a planar, optically transparent surface which creates an ultra-high density sequencing flow cell with hundreds of millions of clusters, each containing ~1,000 copies of the same template (Adam, 2001; Balasubramanian *et al.*, 2002; Mardis, 2008). These templates are sequenced using a robust four-color DNA sequencing by synthesis technology that employs reversible terminators with removable fluorescent dyes (Figure 1.19) (Adam, 2001; Balasubramanian *et al.*, 2002; Mardis, 2008).



**Figure 1.18:** Overview of 454 sequencing. (a) Genomic DNA is isolated, fragmented, ligated to adapters and separated into single strands (ss). (b) One fragment per bead is placed in a PCR-reaction-mixture-in-oil emulsion and amplified. (c) Beads carrying ssDNA templates are deposited into wells. (d) Smaller beads carrying immobilized enzymes required for pyrophosphate sequencing reaction are deposited. (e) Scanning electron micrograph of a portion of a fiber-optic slide. (f) The 454 sequencing instrument consists of: a fluidic assembly (i), a flow cell (ii), a CCD camera-based imaging assembly (iii), and a computer (iii) (Rothberg & Leamon, 2008).



**Figure 1.19.** Illumina sequencing technology. Schematic diagram of technique with a brief explanation (Hodges *et al.*, 2007).

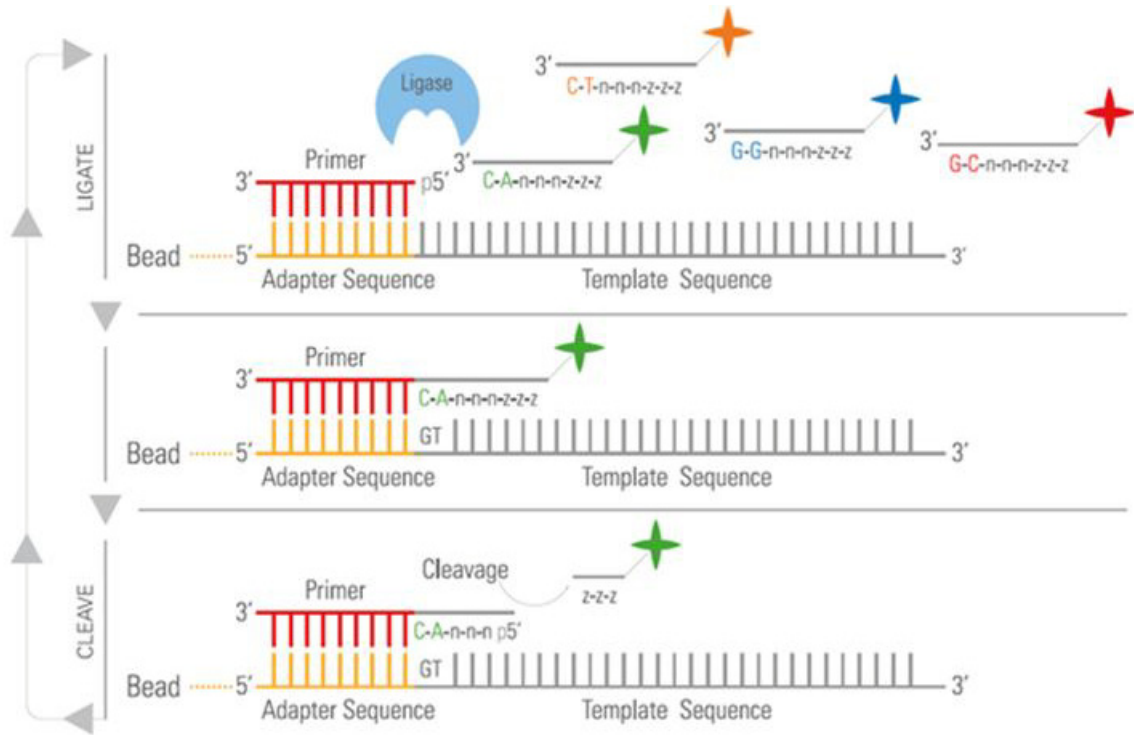
Illumina sequencing has a high degree of accuracy due to the base-by-base sequencing which enables sequencing through homopolymers and repetitive sequences and eliminates sequence-context specific errors (Balasubramanian *et al.*, 2002; Mardis, 2008). This platform has been used to sequence the human oral bacterial community, where hundreds of millions of 50 to 100 bp reads were obtained revealing the abundance of bacterial species present in the mouth (Lazarevic *et al.*, 2009; Mardis, 2008). No fungal communities have been sequence using this technology.

**1.8.5 SOLiD.** Another next generation sequencing platform is the Applied Biosystem SOLiD<sup>TM</sup> (sequencing by oligonucleotide ligation and detection) system (McKernan *et al.*, 2009; Valouev *et al.*, 2008). As with 454 sequencing, a DNA library is made which then undergoes emulsion PCR. The products are render single stranded and fixed on a bead and deposited into a well. A set of four fluorescently labeled di-base probes are added to each well where they compete for ligation to the DNA being sequenced (sequencing by oligonucleotide ligation). A DNA polymerase is not used for sequencing but rather a DNA ligase that attaches the labeled di-base probe to the growing strand (Figure 1.20). On the basis of the fluorescence produced by the molecule, the identity of the nucleotide at each position can be determined (McKernan *et al.*, 2009; Valouev *et al.*, 2008). At this time SOLiD<sup>TM</sup> cannot be used for metagenomic sequencing but technological advances may change this in the future.

Illumina and 454 sequencing are both capable of sequencing mixed DNA samples and within days the microorganisms present in a sample can be determined. At this time



the bacterial and fungal community found in the CF sputum has not been analyzed with either methods.



**Figure 1.20.** SOLiD™ sequencing technology. Schematic diagram of sequencing by ligation. The single stranded template DNA is ligated to a bead via an adapter sequence. The primer for sequencing hybridizes to the adapter on the templated beads. A set of four fluorescently labeled di-base probes compete for ligation to the sequencing primer. The probe with the complementary base will attach to the primer using a DNA ligase. The fluorescent molecule will produce a detectable signal. The fluor will be cleaved from the probe leaving a phosphate at the end of the growing strand that will be used in the next cycle of ligation. Multiple cycles of ligation, detection and cleavage are performed (Hodges *et al.*, 2007). SOLiD™ sequencing technology description and figure was obtained from the manufacturer (Applied Biosystems, Foster City, CA).

## 1.9 Metagenomic statistics and bioinformatics

Identifying members of a community through LH-PCR or 454 sequencing only addresses one question, “Who belongs to the community?” The significance of these identities is determined using multiple statistical tests and bioinformatic programs. Statistical analysis using 454 data is a field that is just beginning to be developed. Analyzing the large amount of sequencing data obtained from 454 uses some mathematical analyses and bioinformatic programs which are based on previous methods that are used in LH, TRFLP, and other community profiling techniques, while other analysis methods are created anew.

**1.9.1 Statistical analysis based on ecological indices.** The earlier approaches used to study microbial diversity and community dynamics include computing measures derived from ecological indices such as species richness and dominance or evenness indices (Hill *et al.*, 2003). Traditional indices include the *richness* (S), the *Shannon information index* (H), and the *evenness* (E) derived from it, and are defined as follows in Equations (1), (2) and (3), respectively:

$$S = \# \text{ of peaks of in each sample,} \quad (1)$$

$$H = -\sum_i p_i (\ln p_i), \quad (2)$$

where  $p_i$  is the ratio of individual peak height to the sum total of the heights of all the peaks in the LH profile, and

$$E = H/H_{max}, \quad (3)$$

where  $H_{\max} = \ln(S)$ . Note the traditional diversity indices are based on the clear definition of an ecological description of an individual species but these definitions have been modified for presumptive identification of LH/T-RFLP profiles by replacing the definition of an individual species with that of individual peaks in LH/T-RFLP profiles.

These calculations identify the richness (how many different types of organisms are present in a community), the evenness (how these organisms are distributed throughout the community), and the Shannon index (the diversity of the community based on number of species and their abundance in the community). These indices provide the basic framework to understanding a particular niche.

**1.9.2 Statistical analysis based on abundance models.** Even with the availability of the numerous diversity indices, analyzing microbial diversity and communities merely using ecological indices has its shortcomings (Mills *et al.*, 2006). Although each index represents an attempt to distill diversity information into a single quantity, each one ends up measuring specific aspects of diversity. Diversity indices vary in their sensitivity to different abundance classes. Species abundance models are considered to be more sophisticated tools to investigate diversity because they examine the distribution of abundances in a population.

Statistical models used for species abundance of microbial communities include log series distribution, log-normal distribution, broken stick model, and overlapping niche model (Curtis *et al.*, 2002; Hill *et al.*, 2003). The most frequently used statistical model for species abundance distributions is the log-normal distribution. In log-normal

communities, the null model for the bacterial species abundance is a lognormal distribution as follows:

$$S(R) = \left( S_T / [\sigma(2\pi)]^{0.5} \right) e^{[-R^2/2\sigma^2]}$$

where  $S(R)$  is the number of species that contain  $R$  individuals,  $S_T$  is the total number of species in the community, and  $\sigma^2$  is the variance of the distribution. The parameters  $S_T$  and  $\sigma^2$  can be estimated from a sample of measured species abundance data by using statistical techniques such as the method of moments or least squares analysis (Curtis *et al.*, 2002). This distribution is seen when a species is persistent within a community (Magurran & Henderson, 2003). The log series distribution comes into effect when a species is not persistent in a community but only occasionally distributed (Magurran & Henderson, 2003). Occasionally, a species can transition between the two types of distribution (Magurran & Henderson, 2003). The broken stick model describes the relative abundance of species by random breaking of a theoretical line which represents the resources of the environment. The length of each broken line or stick represents the abundance for a particular species (McArthur, 1957). This definition can be rephrased to mean that “a group of  $n$  species of equal competitive ability simultaneously occupying the total niche and jostling each other to determine niche boundaries” (Tokeshi, 1993). McArthur also proposed the overlapping niche which refers to species that share a part of the same niche as another species (McArthur, 1957). Each of these statistical models can be used to describe bacterial or fungal communities when species abundance is of interest.

**1.9.3 Comparative metagenomics.** Ecological indices and abundance models can be informative when studying one population at a time or comparing between two species. Comparing communities as a whole can also provide clues as to how communities change over time or how a community is affected by its environment or outside factors. For instance, by comparing all sequencing or LH data, it can be determined if there are statistically differences between the bacterial or fungal community from a healthy sputum sample versus a CF sample. In addition, the factors that cause these distinctions can be identified.

Comparative metagenomics can be studied using statistics or bioinformatics. Most statistical comparisons start with a distance matrix which determines the similarity of one community to another community. Jaccard's coefficient, Hellinger distance, Pearson chi-square test, Euclidean distance matrices, and Bray-Curtis/Sorenson coefficient are often used to determine the relationship between two samples or two populations (Beran, 1977; Bray & Curtis, 1957; De Leuw & Heiser, 1982; Jaccard, 1901; Pearson, 1900). These equations use the presence/absence and/or abundance of all the organisms in a sample to determine the relatedness between two samples (Beran, 1977; Bray & Curtis, 1957; De Leuw & Heiser, 1982; Jaccard, 1901; Pearson, 1900). The resulting distance/similarity matrices can then be further analyzed using multivariate statistics.

Multivariate statistics can be used to identify what data links two samples (either members in a sample or external environmental factors). These analyses can also be used to determine if samples within a population are similar to one another (Rudi *et al.*, 2007). An example of multivariate statistical analysis is the clustering of microbial communities

in soil using the unweighted pair-group method which uses the arithmetic averages (UPGMA) algorithm on the data derived from the distance metric data (such as the Jaccards or Hellinger or Pearson distances) (Blackwood *et al.*, 2003; Dunbar *et al.*, 2000; Griffiths *et al.*, 2000). Such methods have been used to support claims that certain relationships between communities can be discerned, that the groupings are natural, and that outliers can be identified (Clarke *et al.*, 2006). Other statistics such as principle component analysis, and multi-dimensional scaling will be discussed in detail in Chapters 2 and 3.

Phylogenetics can also be used to compare among samples and populations. The evolution and/or relatedness of a community can be determined using the phylogenetic tree that is present in the sample. The bioinformatic program, UniFrac can use phylogenetic trees to determine if the microbial community in a sample is similar to another sample factors (Lozupone & Knight, 2005; Lozupone *et al.*, 2006; Lozupone *et al.*, 2007). If there are differences between the communities, then the program will determine if a particular lineage of the tree is responsible for the difference. The UniFrac program can also cluster communities based on environmental factors (Lozupone & Knight, 2005; Lozupone *et al.*, 2006; Lozupone *et al.*, 2007). Megan 3 (metagenome analysis software) is another bioinformatic program that can compare communities by using the original sequencing data and identifying taxa (Huson *et al.*, 2007). Both of these programs can handle large amount of data which is key when analysis 454 sequencing reads. The output for both programs can be used in collaboration with statistics to identify correlations between communities.

With rigorous statistical analysis and bioinformatics, it is possible to differentiate between significant differences and random events. Thus, leading one to understand how a microbial community is affected by its members and environment.

### **1.10 Hypothesis and aims**

The CF lower respiratory tract contains a complex community of fungi and bacteria which cause chronic lung infections. Because of the long term effects on the lungs by microbial infection, a better understanding of the microbes present and the factors that drive evolution of the microbial community will enable the development of improved therapies.

This project proposes to characterize the bacterial and fungal communities found in the CF lungs using the molecular techniques, amplicon length heterogeneity–polymerase chain reaction (LH-PCR) and 454 sequencing. In addition, this study proposes to develop a novel detection assay for CF pathogens. The hypothesis of this project is that there is both bacteria and fungi present in the CF lung and the dynamics of the community is dependent on the host and antimicrobial treatment. The hypothesis will be tested by accomplishing the following three aims:

***Aim I:*** Analyze the metagenomic DNA from the CF sputa samples using the V1 and V1\_V2 regions of the bacterial 16S rRNA gene using LH-PCR community profiling. Community dynamics will be studied using these profiles. The working hypothesis is the CF lung contains numerous bacteria and selective pressures from the host and antibiotic therapy cause the community to change over time.

***Aim II:*** Analyze the metagenomic DNA from the CF sputa samples using the V1\_V2 region of the bacteria 16S rRNA gene and the fungal ITS1 region using 454 sequencing. The CF microbiomes will be compared with healthy oral rinse microbiomes. The working hypothesis is the CF microbiome will contain organisms that are associated with healthy individuals as well as pathogens unique to the CF lung infection.

***Aim III:*** Detect bacteria from CF sputa samples using SPR technology. The working hypothesis is CF pathogens can be detected based on antibody-antigen reactions measured by SPR.



## **Chapter 2**

### **“Combination of 16S rRNA Variable Regions Provide a Detailed Analysis of Bacterial Community Dynamics in the Lungs of Cystic Fibrosis Patients”**

As published in Human Genomics

Volume 4, Issue 3 pgs 147-169, 2009

Copyright © Henry Stewart Publications

Melissa S. Doud, Michael Light, Gisela Gonzalez, Giri Narasimhan, and Kalai Mathee

## 2.1 Abstract

Chronic bronchopulmonary bacterial infections remain the most common cause of morbidity and mortality among patients with cystic fibrosis (CF). Recent community sequencing work has now shown that the bacterial community in the CF lung is polymicrobial. Identifying bacteria in the CF lung through sequencing can be costly and is not practical for many laboratories. Molecular techniques such as terminal restriction fragment length polymorphism or amplicon length heterogeneity - polymerase chain reaction (LH-PCR) can provide many laboratories the ability to study CF bacterial communities without costly sequencing. The aim of this study was to determine if the use of LH-PCR with multiple hypervariable regions of 16S rRNA gene could be used to identify organisms found in sputum DNA. This work also determined if LH-PCR could be used to observe the dynamics of the lung infection over a period of time. Nineteen samples were analyzed with the V1 and the V1\_V2 region of the 16S rRNA gene. Based on the amplicon size present in the V1\_V2 region, *P. aeruginosa* was confirmed to be in all 19 samples obtained from the patients. The V1 region provided a higher power of discrimination between bacterial profiles of patients. Both regions were able to identify trends in bacterial population over a period of time. LH profiles showed that the CF lung community is dynamic and changes in the community may in part be driven by the patient's antibiotic treatment. LH-PCR is a tool that is well suited for studying bacterial communities and their dynamics.

**Key Words:** LH-PCR, 16S rRNA, V1, V1\_V2, cystic fibrosis, AmpliQué, Cystic fibrosis microbiome, metagenome

## 2.2 Introduction

Cystic fibrosis (CF) is an autorecessive disease affecting one in 3500 Caucasian live births in the U.S (Patient-registry, 2008). CF results from a mutation in the gene that encodes the cystic fibrosis transmembrane conductance regulator (CFTR) protein (Riordan *et al.*, 1989; Wainwright *et al.*, 1985). A defect in the CFTR protein leads to a malfunctioning cyclic-AMP-activated chloride channel in secretory epithelia (Kunzelmann, 1999). This defect in the lung leads to the inability to secrete chloride and to the excess re-absorption of sodium (Kunzelmann, 1999). Thus, there is decreased fluid secretion and the mucus becomes immobilized and adheres to the epithelial cells. Overproduction of mucus in the airway results in congestion of the respiratory tract and increases susceptibility to bronchopulmonary infection. CF patients often suffer from infections with *Staphylococcus aureus*, *Haemophilus influenzae*, *Pseudomonas aeruginosa*, and *Burkholderia cenocepacia* (Lyczak *et al.*, 2002). These chronic infections by highly adapted lung microbes cause inflammation and eventually lung damage; therefore, lung disease is the major cause of morbidity and mortality among these patients (Hull, 2003; Mathee *et al.*, 1999).

It has also been estimated that less than 1% of eubacteria in the environment can be cultured (Ward *et al.*, 1990a; Ward *et al.*, 1990b). Thus, these identification methods will fail to detect all pathogens that might be causing the lung infection. Fortunately, with the advent of molecular techniques, culturing for identification purposes can be circumvented (van Belkum *et al.*, 2000). Recent molecular studies using terminal restriction fragment length analysis and sequencing have shown that the lung community is quite complex. *Achromobacter (Alcaligenes) xylooxidans*, *Rhodotorula mucilaginosa*,

*Abiotrophia* spp., *Bacteroides gracilis*, *Eubacterium brachy*, *Mycobacterium mucilaginosus*, *Mycoplasma salivarium*, *Porphyromonas salivae*, *Ralstonia* spp., *Staphylococcus hominis*, *Streptococcus anginosus*, *Treponema vincentii*, *Veillonella* spp., *Burkholderia gladioli*, *Stenotrophomonas maltophilia* and *Pandoraea atypical* have all been demonstrated to be members of this polymicrobial infection (Bittar *et al.*, 2008; Burns *et al.*, 1998; Coenye *et al.*, 2001; Coenye *et al.*, 2002; Coenye *et al.*, 2003; Coenye *et al.*, 2005; Harris *et al.*, 2007; Rogers *et al.*, 2003).

Amplicon length heterogeneity – polymerase chain reaction (LH-PCR) is another molecular method that has been used to study various diverse microbial communities (Bernhard *et al.*, 2005; Suzuki *et al.*, 1998) including those of CF sputum (Rogers *et al.*, 2003). LH-PCR utilizes the 16S rRNA molecular marker to analyze microbial populations (Ritchie *et al.*, 2000; Suzuki *et al.*, 1998). The sequence and length of the variable regions within the 16S rRNA gene are often used to determine phylogeny (Coenye & Vandamme, 2003). It should be noted that some bacteria have multiple copies of the 16S rRNA gene which can show intragenomic heterogeneity in sequence and gene length (Coenye & Vandamme, 2003). This variation is unlikely to have an effect on classification at the genus and species level and the marker is routinely used for identification (Case *et al.*, 2007). Identification based on length variation in the hypervariable regions of the gene is exploited by LH-PCR. In this technique, forward and reverse primers bind to the conserved regions of the 16S rRNA gene and then amplify the varying lengths seen in the hypervariable regions. LH-PCR has proven to be a robust, time efficient and reproducible method (Mills *et al.*, 2003; Moreno *et al.*, 2006; Suzuki *et al.*, 1998). The main advantages of LH-PCR are that it surveys relative gene

frequencies within complex mixtures of DNA, is reproducible, requires small sample sizes, and can be performed with many samples simultaneously (Moreno *et al.*, 2006; Suzuki *et al.*, 1998). Furthermore, some of the size classes emerging from LH-PCR analyses can be related at the genus level to archived database sequences (Suzuki *et al.*, 1998). Overall, the attributes of LH-PCR make it useful for quick assessment of the diversity of microbial communities for comparative purposes (Suzuki *et al.*, 1998).

Prior studies have shown that LH-PCR and the hypervariable regions one and two (V1\_V2) of the 16S rRNA gene can be used to detect eubacteria in sputum from CF patients (Miller, 2003; Rogers *et al.*, 2003). The technique had limited efficacy when determining overall diversity and attaining organism identification and it was concluded that the small differences in amplicon lengths generated from the organisms was at fault (Rogers *et al.*, 2003). In this study, community profiles of CF sputa were obtained using two regions of the 16S rRNA gene (hypervariable region 1 [V1] and [V1\_V2]). Bacterial identification based on amplicon length was attempted using a computer program, AmpliQué, which determined the theoretical V1 and V1\_V2 fragment lengths for known organisms. Community profiles and identification were analyzed using each region separately and combined to determine if a combinatorial approach to LH-PCR (i) could provide a more detailed analysis of the bacterial community in the CF lung and (ii) could be used to observe the dynamics of this community over a period of time. This work has shown that the V1 region provides a more detailed community profile which leads to a higher power of discrimination between profiles than the V1\_V2 region. The LH analyses provided snapshots of the CF bacterial communities. However, the analyses are limited by their ability to identify community members with the V1 or V1\_V2 regions

separately or combined. In addition, through the comparisons of V1\_V2 or V1 LH profiles over time, it was established that the CF lung community is not only diverse but also dynamic.

## **2.3 Material and methods**

**2.3.1 Sample collection.** Nineteen sputum samples were obtained from CF patients attending the University of Miami adult CF clinic (Miami, FL). Samples are referred to by UM followed by a letter. The study was carried out in accordance with the Declaration of Helsinki (2000) of the World Medical Association and was approved by the appropriate institutional review board (FIU IRB approval # 033004-02). Appropriate consent was obtained from human subjects. All 19 expectorated sputum samples were frozen at -20 °C prior to DNA extraction. Sputum was liquefied using the method of Reischl *et al* and the metagenomic DNA was extracted and cleaned using the GeneClean Spin Kit (Qbiogene, Irvine, CA) as per manufacturer's protocol (Reischl *et al.*, 1994). The eluted DNA was quantified and diluted to 10 ng/μL and refrigerated at 4°C until further use.

**2.3.2 Control bacteria isolates.** Standard DNA manipulations were used to extract chromosomal DNA from CF-related pathogens; *P. aeruginosa* strains PAO1 and PA14 (burn wound isolates) and four additional *P. aeruginosa* strains isolated from CF patients (Klockgether *et al.*, 2007), *B. cenocepacia* (ATCC BAA-246; CF isolate), *Staphylococcus aureus* (ATCC 12600; pleural fluid isolate), and *Stenotrophomonas*

*maltophilia* (ATCC 17672; sputum isolate) were used as controls in the LH-PCR experiments.

**2.3.3 LH-PCR.** The metagenomic DNA derived from the extracted sputum samples were amplified in triplicate using eubacterial primers for the 16S rRNA gene. The V1\_V2 region of the 16S rRNA gene were amplified using 27F-6FAM forward primer (5'– 6FAM – AGA GTT TGA TCM TGG – 3') and the 355R reverse primer (5' – GCT GCC TCC CGT AGG AGT -3') (Doud *et al.*, 2009). The V1 region was amplified using fluorescent primer P1F-6FAM (5' – 6FAM – GCG GCG TGC CTA ATA CAT GC – 3') and reverse primer P2R (5' - TTC CCC ACG CGT TAC TCA CC -3') (Cocolin *et al.*, 2001; Doud *et al.*, 2009) All forward primers are fluorescently labeled with 6-FAM<sup>TM</sup> (Integrated DNA Technologies, Shokie, IL). The final V1\_V2 PCR reaction mixture was composed of: 1X PCR buffer (supplied with the enzyme by Applied Biosystems, Foster City, CA), 2.5 mM MgCl<sub>2</sub>, 0.25 mM dNTPs, 0.5 μM forward and reverse primers, 0.1% bovine serum albumin, 0.025 U of AmpliTaq Gold LD DNA polymerase<sup>TM</sup> (Applied Biosystems, Foster City, CA), and diethylpyrocarbonate (DEPC) water to make up the final volume of 20 μL. A half a nanogram per microliter of DNA template was added per reaction. The following parameters were used to amplify the selected fragments on a MJ Research Peltier Thermal Cycler 200 (Waltham, MA): initial denaturation at 95°C for 11 minutes, 25 cycles of denaturation at 95°C, annealing at 55°C and extension at 72°C, each for 1 minute and a final elongation at 72°C for 10 minutes. A negative control containing water was amplified for every PCR master mix to check

for contamination. Agarose gel electrophoresis was run to confirm the success of amplification.

**2.3.4 LH-PCR analysis.** LH-PCR analysis was performed by the Forensic DNA Profiling Facility (Florida International University, Miami, FL). The PCR products for the three replicate reactions were detected using the ABI Prism 310 Genetic Analyzer (Applied Biosystems, Foster City, CA). A formamide-size standard mix was prepared using a 96:1 ratio of Hi-Di<sup>TM</sup> (highly deionized) formamide and GeneScan<sup>TM</sup> 500 ROX<sup>TM</sup> internal standard (Applied Biosystems). Each PCR product was denatured by adding 0.50  $\mu$ l to 9.5  $\mu$ l of the formamide-size standard mix for 2 minutes at 95°C and snap cooling for 5 minutes on ice. Each sample was run for 28 minutes on the ABI Prism 310. Fragments were separated by capillary electrophoresis using polymer POP-4, matrix DS-30\_6FAM\_HEX\_NED\_ROX and filter D (Applied Biosystems).

**2.3.5 Electropherogram analysis.** The output was collected and analyzed using GeneMapper<sup>®</sup> Software Version 3.7 (Applied Biosystems) and Microsoft Excel (Microsoft, Redmond, WA). The ABI Prism<sup>TM</sup> Genotyper software analysis parameters were set to the local Southern Size calling, no correction, and the minimum noise threshold was set at 70 fluorescent units (Mills *et al.*, 2003). For V1\_V2, amplicons were called if between 300 and 400 base pairs (bp). Amplicons for V1 region were between 60 and 120 bp.



**2.3.6 Statistical analysis.** Data were imported into MS Excel (Microsoft, Redmond, WA), to determine mean relative ratios for each triplicate PCR reaction. The normalized data for each sample were exported into PRIMER 5 statistical software for further analyses (PRIMER E Ltd., Plymouth Marine Laboratory, United Kingdom). Bray – Curtis similarity index and non metric multi-dimensional scaling (MDS) was used to compare sample data between patients and patient profiles over time (Hartmann *et al.*, 2005; Rees *et al.*, 2004).

**2.3.7 Presumptive identity analysis.** The GeneScan™ 500 ROX™ internal standard was used to determine amplicon sizes in base pairs. Several approaches were used to presumptively identify an organism in the CF community based upon the length of the amplicon generated. Chromosomal DNA from isolates of *Burkholderia cenocepacia*, *Staphylococcus aureus*, *Stenotrophomonas maltophilia*, and *Pseudomonas aeruginosa* were used to determine the experimental amplicon length for the V1\_V2 and V1 regions. Next, *in silico* analysis was performed in which the expected amplicon length was determined based on where the primers of interest would theoretically bind to the published sequence. Lastly, the lengths of the expected amplicons for all bacteria were determined using AmpliQué, a newly designed bioinformatics program (Gonzalez *et al.*, 2009). The Perl scripts are available at <http://biorg.cs.fiu.edu/AmpliQue/>. AmpliQué uses the 16S rRNA sequences downloaded from the Ribosomal Database Project II (Release 9.60). It takes a forward and a reverse primer (both of which may be potentially degenerate) as input. The stringency of the primer binding can be set in the program to control the exactness of the primer binding site within the 16S rRNA sequences. The

stringency was first set to 100 % (E = 10) to identify the organisms that have exactly the same conserved sequences as the primer and subsequently lowered to E = 5,000 which allowed a larger database to be created as the conserved regions of the 16S rRNA gene are not 100 % identical in sequence. The resulting output shows the identity of all the 16S rRNA sequences that will produce an amplicon with a particular pair of primers, along with the length(s) of the amplicons. Thus, one can query the database using any primer to determine the theoretical fragment lengths that would be generated from an organism when analyzed. In addition, a list of all bacteria that produce a specific fragment length can be generated. One amplicon length may represent more than one bacterium and thus only a presumptive identity of that amplicon can be achieved.

## **2.4 Results**

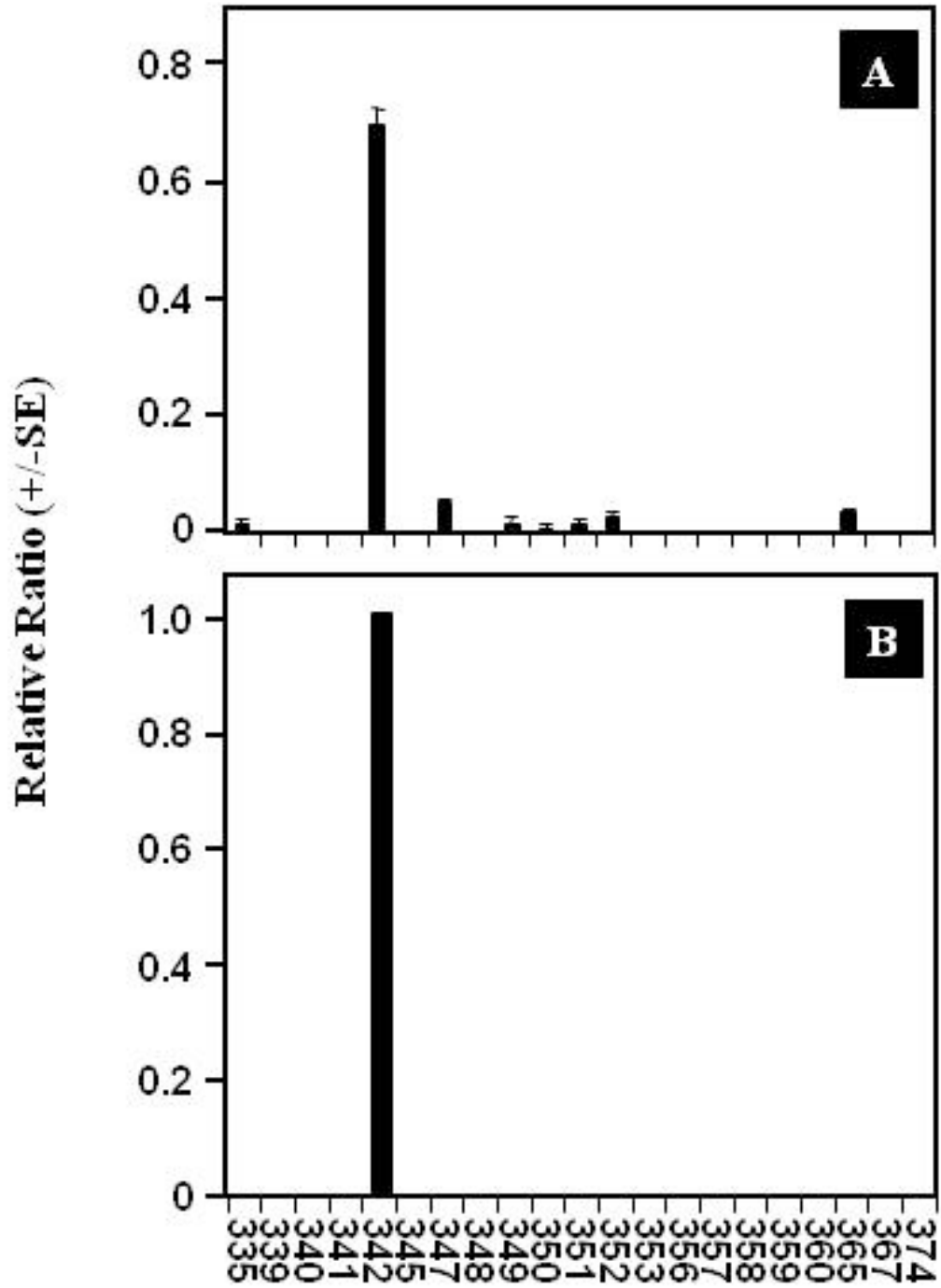
### **2.4.1 PCR amplification of 16S rRNA genes for metagenome profiling.**

Variable regions V1 and V1\_V2 were successfully amplified using universal eubacterial primers from the metagenomic DNA samples (See Methods). As expected, the amplicon sizes from the 19 samples ranged from 335 to 365 bp for the V1\_V2 region (Table 2.1 and Supplemental Figure 2.6). Sample UMC had the most complex V1\_V2 profile with eight amplicons (Figure 2.1A and Table 2.1). Samples UMF, UMG, UMH, UMJ, UMK, UMP, and UMT had the least complex LH profiles with only one amplicon present at 342 bp (Figure 2.1B and Table 2.1). The remaining 11 samples produced LH profiles containing multiple amplicons (Table 2.1, Supplemental Figure 2.6). Amplicon 342 was observed in all 19 samples (Table 2.1). Out of these 19 samples, 16 samples (84.2 %)

**Table 2.1: Amplicons seen in the V1 and V1\_V2 regions for samples**

Sample	V1_V2 (%)*	Total # V1_V2	V1	Total # V1
UMA	342, <b>349 (39.0)</b> , 350	3	60, 62, 63, 65, 68, 69, 70, 72, 73, 75, 76, 77, 78, 81, 82, 83, <b>86 (17.3)</b> , 90, 91, 92, 93, 94, 96, 98, 100, 101, 108	27
UMB	<b>342 (61.7)</b> , 347, 365	3	62, 63, 72, 73, 75, 76, 77, 78, 81, 82, <b>88 (15.7)</b> , 91, 93, 94, 98, 100, 102, 103, 112	19
UMC	335, <b>342 (79.0)</b> , 347, 349, 350, 351, 352, 365	8	61, 62, 68, 69, 70, 72, 73, 75, 76, 77, 78, 81, 82, 84, 86, <b>88 (12.0)</b> , 90, 91, 92, 93, 94, 96, 98, 100, 102, 104	26
UMD	<b>342 (93.7)</b> , 351, 352, 365	4	<b>64 (14.4)</b> , 70, 73, 77, 82, 86, 88, 98, 103, 112	10
UME	<b>342 (99.0)</b> , 348	2	60, 62, 64, 66, 67, 68, 69, 71, <b>72 (12.3)</b> , 73, 74, 75, 76, 77, 78, 80, 81, 86, 88, 91, 93, 94, 98, 100, 102, 103, 112, 113, 115, 120	30
UMF	<b>342 (100)</b>	1	60, 61, 62, 65, 68, 70, <b>72 (12.6%)</b> , 73, 75, 77, 78, 85, 91, 93, 94, 98, 103, 112, 113, 115	20
UMG	<b>342 (100)</b>	1	60, 61, 62, 68, <b>72 (3.6)</b> , 73, 75, 77, 78, 86, 91, 93, 98, 103, 112, 113, 115, 119, 120	19
UMH	<b>342 (100)</b>	1	60, 61, 62, 65, 68, 70, 73, 75, 76, <b>77 (3.9)</b> , 78, 80, 81, 84, 85, 86, 88, 90, 91, 93, 94, 96, 98, 101, 103, 111, 112, 113, 115	29
UMJ	<b>342 (100)</b>	1	60, 61, 62, 63, 65, 68, 70, 72, 73, 75, 77, 78, 80, 81, 84, 86, 88, 91, 92, 93, 94, 98, <b>102 (6.6)</b> , 112, 113, 115, 120	27
UMK	<b>342 (100)</b>	1	60, 61, 62, 65, 67, 68, 70, 72, 73, 75, 77, 78, 81, 85, 86, 90, 91, 93, 94, 98, 100, <b>101 (7.3)</b> , 103, 111, 112, 113, 120	27
UML	342, 350, <b>360 (55.7)</b> , 365	4	60, 61, 62, 63, 64, 65, 67, 68, 70, 71, 72, 73, 76, 78, 81, 85, 88, 91, <b>94 (31.3)</b> , 98, 100, 102, 103, 112, 113, 115, 119, 120	28
UMM	<b>342 (98.7)</b> , 351, 365	3	61, 62, 65, 68, 70, 73, 75, 76, 77, 78, 81, 82, 84, 86, 88, 91, 94, 98, <b>101 (4.8)</b> , 103, 111, 113, 115, 119, 120	25
UMN	339, <b>342 (92.7)</b> , 347, 352	4	60, 61, 62, 65, 68, <b>70 (10.2)</b> , 73, 77, 78, 81, 85, 86, 88, 90, 91, 92, 94, 98, 100, 103, 108, 112, 113, 115	24
UMO	<b>342 (74.0)</b> , 350, 352	3	60, 62, 65, 73, 77, 78, 80, 81, 85, 91, 98, <b>102 (3.1)</b> , 103, 108, 111, 112, 113, 115, 119, 120	20
UMP	<b>342 (100)</b>	1	60, 61, 62, 65, 77, 78, 81, 98, <b>102 (5.7)</b> , 112, 115, 120	12
UMQ	<b>342 (41.3)</b> , 348, 350, 365	4	60, 61, 62, 63, 64, 65, 71, 72, 76, 78, 81, 84, 85, 87, 91, 94, 98, <b>102 (5.3)</b> , 103, 112, 115	21
UMR	342, <b>360 (63.7)</b>	2	60, 62, 63, 65, 73, 77, 78, 80, 81, 82, 85, <b>94 (13.9)</b> , 98, 102, 108, 112, 115, 120	18
UMS	339, <b>342 (87.7)</b> , 347	3	60, 62, 63, 66, 72, 73, 74, 77, 78, 80, 81, 85, 88, 91, 94, 98, <b>102 (9.9)</b> , 103, 111, 112, 113, 115, 119, 120	24
UMT	<b>342 (100)</b>	1	60, 62, 63, 65, 71, 72, 73, 76, 78, 81, 85, 92, 94, 98, <b>102 (13.2)</b> , 111, 112, 115, 120	19

\*Amplicons in bold are the most abundant amplicon in the profile for that particular sample. The number in parenthesis indicates the percentage abundance of the dominant amplicon for that region.



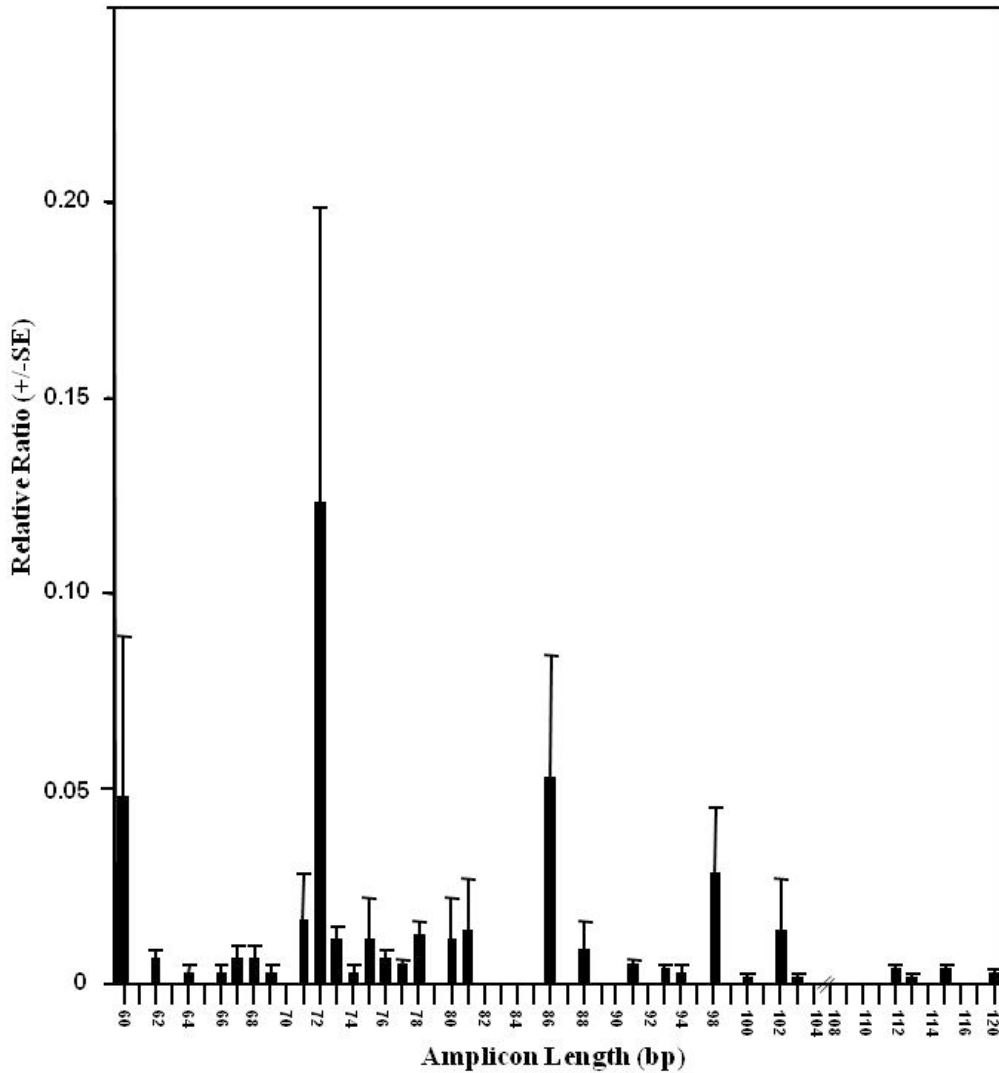
**Figure 2.1:** Hypervariable V1\_V2 amplicon data graphs. Panel A represents the V1\_V2 profile from patient UMC. Panel B depicts the profile obtained from patients UMF, UMG, UMH, UMJ, UMK, UMP, and UMT using the V1\_V2 region.

had amplicon 342 as the dominant member of the profile. The most abundant amplicons in UMA, UML, and UMR were 349, 360, and 360, respectively (Table 2.1).

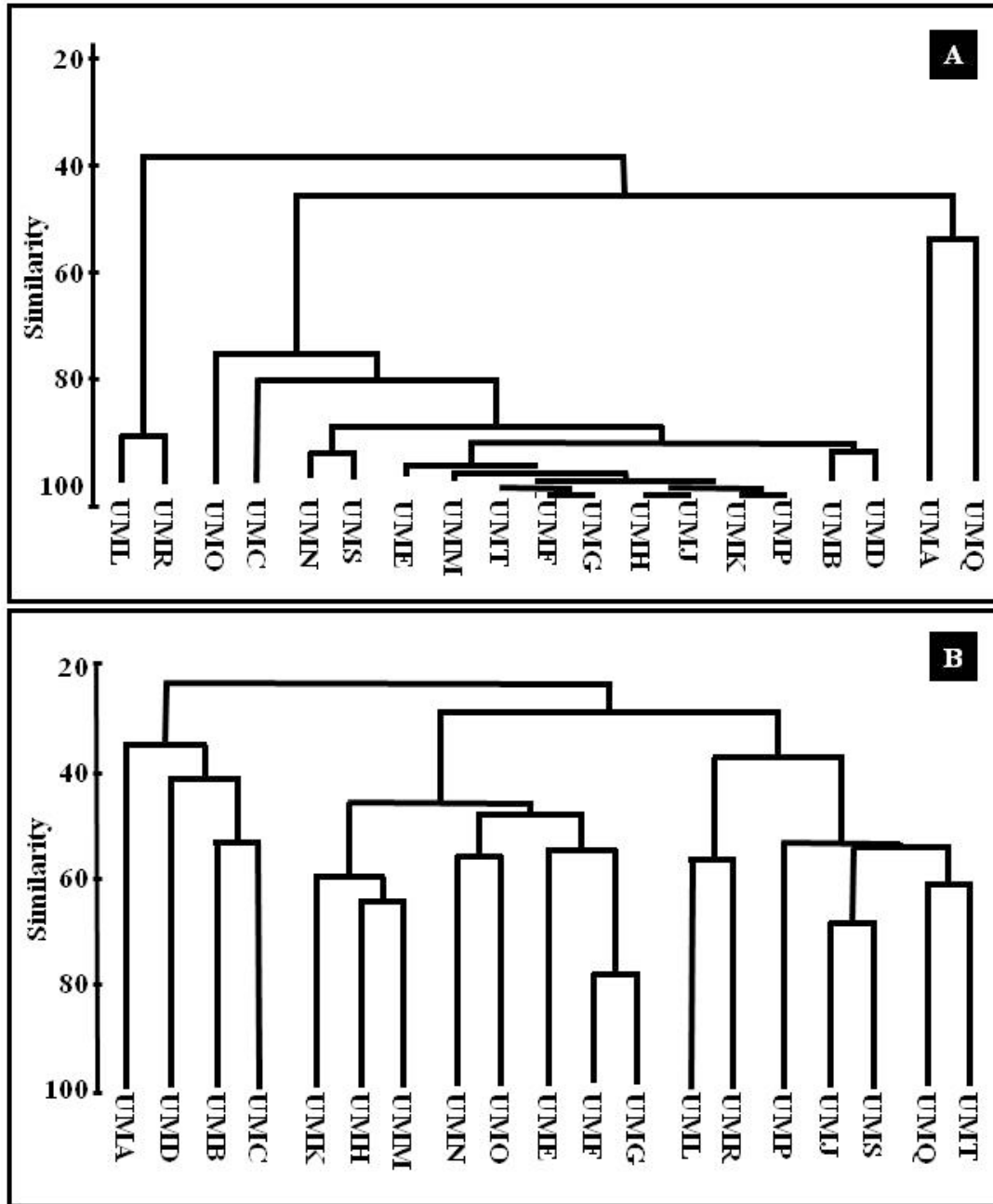
As predicted, V1 profiles ranged from 60 to 120 bp in length (Table 2.1 and Supplemental Figure 2.7). Unlike V1\_V2 profiles, all 19 samples produced more diverse profiles containing multiple amplicons in each profile (Table 2.1, Supplemental Figure 2.7). The most complex profile contained 30 amplicons for UME (Figure 2.2, Table 2.1). The least diverse profile with 10 amplicons belonged to UMD (Table 2.1, Supplemental Figure 2.7). All the samples contained a 98 bp amplicon, though it was not the most abundant amplicon (Table 2.1). The 86 bp amplicon was present in 10 out of the 19 profiles and was the most dominant peak in UMA (17%) (Table 2.1). The 84 bp amplicon was found in five profiles. Seventeen out of 19 samples contained a 73 bp amplicon.

**2.4.2 Comparison of LH eubacterial profiles within CF center.** The profiles for each sample were compared to each other to determine if the bacterial communities were similar between patients that were treated at the same CF center. The normalized abundance data from the V1\_V2 region was analyzed using the Bray-Curtis similarity index. The similarity between samples was shown in a dendograph using group linkage clustering. The UMF profile has 100 % similarity to UMG, UMH, UMJ, UMK, UMP, and UMT as they all produced 100 % abundance for amplicon 342 (Table 2.1 and Figure 2.3A). The UMR profile had 36.0 % or less similarity with 13 of the samples, but had a 91.0 % similarity with UML (Figure 2.3A). The UMR and UML profiles had two amplicons in common. The V1 region was able to discriminate between all samples; no

two profiles were exactly the same (Table 2.1 and Figure 2.3B). The highest similarity between two profiles was 79.2 % between UMF and UMG due to the presence of many shared amplicons (Figure 2.3B). The most dissimilar profile was between UMD and UMQ with 6.2 % similarity. Four out of the 10 amplicons were common between UMD and UMQ (Table 2.1).



**Figure 2.2:** Hypervariable V1 amplicon data. Graph depicts the V1 profile from patient UME, which was the most diverse profile found in this study.



**Figure 2.3:** Similarity based clustering between LH profiles of patients. Panels A and B are dendrograms indicating the relatedness of LH profiles of patients generated from the V1\_V2 and V1 regions, respectively. Bray – Curtis similarity index and group linked clustering were used to determine similarity of bacterial communities.

The similarity of the patients' bacterial population with both the V1\_V2 and the V1 data were also compared. The combination of regions decreased the degree of similarity between V1\_V2 profiles yet the LH profiles were most dissimilar with the V1 region analysis (data not shown).

**2.4.3 Patient CF flora changes over time.** The lung flora changes over time in the CF lung. It is also known that the bacterial community composition in the lung at any one given time is varied due to the compartmentalized structure of the organ (Harrison, 2007). When working with expectorated sputum, it was assumed that the bacterial community within a given patient may vary from sample to sample. In order to study the dynamics of lung flora over a period of time, sampling bias was addressed. That is, any changes seen in community profiles over time may potentially be an artifact from the sampling method. A control experiment was performed to understand the effect of sampling.

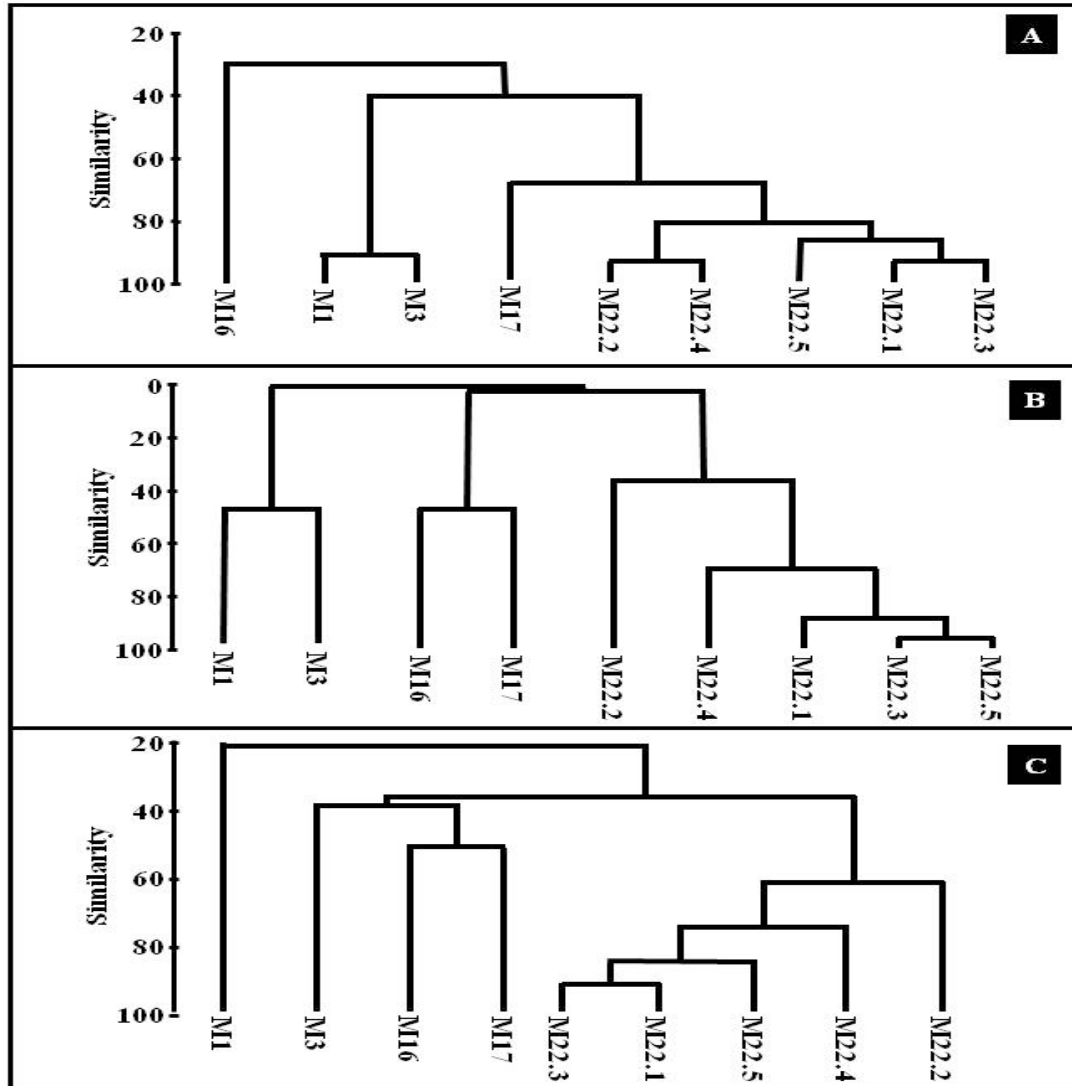
Sputa samples from patient UML were analyzed over the course of two years (Month 1 to Month 22). The samples were referred to as M0, M3, M16, M17, M22.1, M22.2, M22.3, M22.4, and M22.5. Five samples were taken on the same day every three hours for twelve hours (M22.1 - M22.5). The sampling scheme was designed to account for the potential variability contributed by the lung structure. All samples were analyzed using the V1 and V1\_V2 regions. Bray-Curtis similarity index was used to determine the variation between the community profiles of the patient samples. The relatedness of bacterial communities was shown with dendographs created from the calculated similarity (Figure 2.4). In addition, the relationship between profiles was graphically



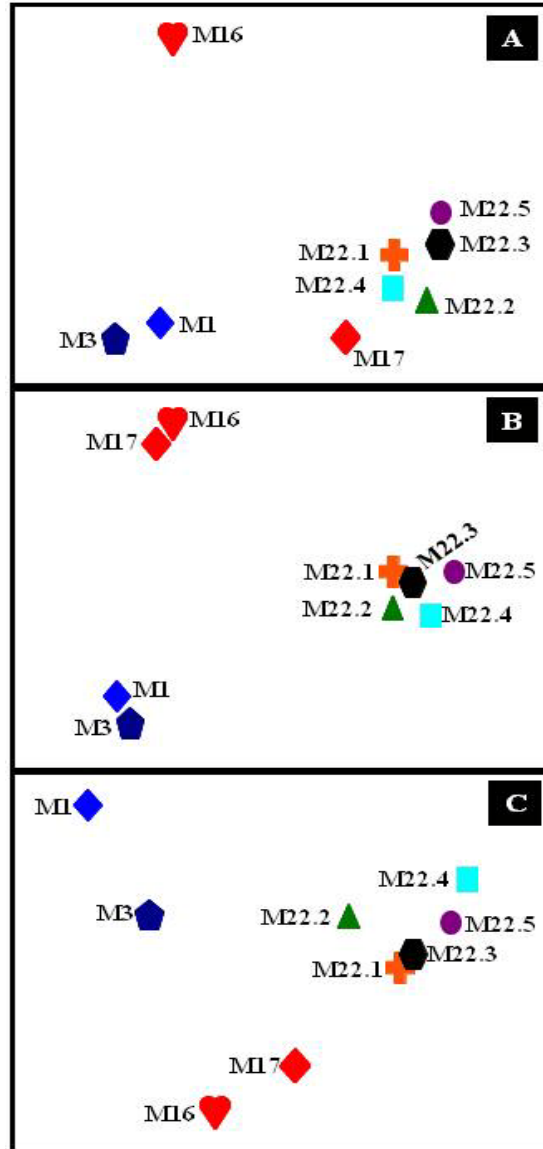
depicted using non-metric MDS which represents the samples as points in a low-dimensional space so that the relative distances of the symbols is correlated with the same rank order determined by the Bray-Curtis similarity index (Figure 2.5). These analyses visually showed that samples taken on the same day were more similar than LH profiles from earlier time points (Figure 2.4 and 2.5). For the V1\_V2 region, the samples in M22 ranged from 70.6 % to 93.4 % in similarity. The first three samples M1, M3, and M16 had less than 42.9 % similarity to any profile produced from the hourly samples. The LH profile M17 and the five profiles in M22 are related with a range of 52.8 % to 78.3 % similarity. M22.4 was more similar to M17 than the other same-day samples (Figure 2.4A).

The V1 profiles for these samples were also analyzed and compared. Overall, the similarities between samplings were much lower when using the V1 versus the V1\_V2 region (Figure 2.4B and 2.5B). The highest profile similarities were seen within M22 samples; as high as 96.0 % were seen between M22.3 and M22.5. M1 and M3 had a similarity of 48.1 %; M16 and M17 were 47.1 % similar. Some of the samplings taken months prior to the daily sampling in M22 were completely different (Figure 2.4B). The MDS analysis clearly showed samples taken on the same day (M22) clustered (Figure 2.5B).

Combining LH data from both the V1 and the V1\_V2 regions further indicated the overall similarity of bacterial community within one day versus over a period of time (Figure 2.4C and 2.5C). The additive effect of the regions created a tighter clustering of the same-day samples and increased distances between samples taken in different months (Figure 2.4C).



**Figure 2.4:** Similarity based clustering between LH profiles over time in one patient. Panels A and B are dendrograms indicating the relatedness of LH profiles of patients generated from the V1\_V2 and V1 regions, respectively. Panel C describes the community similarity when both V1\_V2 and V1 data are combined. Bray – Curtis similarity index and group linked clustering was used to determine similarity of bacterial communities.



**Figure 2.5:** Similarity between LH profiles over time in one patient. Panels A and B are MDS figures which represent the profiles generated from one patient over a period of time using the V1\_V2 and V1 regions, respectively. Panel C is the MDS figure representing the combined V1\_V2 and V1 data. Blue diamond, blue pentagon, red heart, red diamond, orange cross, green triangle, black hexagon, aqua square, and purple circle represent samples from Month 1, Month 3, Month 16, Month 17, Month 22.1- 22.5, respectively.

**2.4.4 Antibiotics drive CF community.** After the initial time study with Patient UML, two more patients were followed for a period of time to determine the effect of antibiotics on the lung microbial flora. Azithromycin, bactrim, tobramycin and colistin were administered at different time points for each patient. Azithromycin, which is most often used to treat Gram-negative bacteria, is a macrolide antibiotic which interferes with protein synthesis preventing growth of bacteria (Retsema *et al.*, 1987). Bactrim is a combination of sulfamethoxazole and trimethoprim and it interferes with DNA synthesis in bacteria that cause respiratory infections, especially in *Stenotrophomonas maltophilia* (Hughes & Russell, 1982; Muder *et al.*, 1996). Tobramycin, an aminoglycoside, prevents translation in bacteria. The inhaled form of the drug targets *P. aeruginosa* in the lungs (Ramsey *et al.*, 1999). Colistin or polymyxin E is effective against most Gram negative bacilli and is often used against multidrug resistant strains of *P. aeruginosa* (Beringer, 2001). The drug regime and sputa samples from two patients referred to as patient UMX and patient UMY were analyzed every other month for five months. The three time points are referred to as M0, M3, and M5. The amplicons of size 60 - 120 bp and 300 – 400 bp were obtained from the V1 and the V1\_V2 region, respectively (Table 2.2).

Patient UMX was under triple combination therapy with azithromycin, bactrim, and tobramycin at least two weeks prior and up to the sampling at month 0 (M0) and month 3 (M3). Two weeks prior to month 5 (M5), the patient stopped taking tobramycin but continued the treatment with azyithromycin and bactrim. The number of amplicons for the V1\_V2 and V1 LH profiles varied for the three time points (Table 2.2). The dominant 342 bp amplicon was present in each time point (93.7 %, 100 % and 74 % at M0, M3, and M5, respectively) for the V1\_V2 region (Table 2.2).

**Table 2.2:** Amplicon contributions for Patients UMX and UMY over a five-month time period

Patient UMX				Patient UMY			
Amplicon (bp)*	M0 <sup>a</sup> (%)	M3 <sup>a</sup> (%)	M5 <sup>b</sup> (%)	Amplicon (bp)	M0 <sup>c</sup> (%)	M3 <sup>d</sup> (%)	M5 <sup>d</sup> (%)
Unique Amplicons							
352	P (2.34)		P (13.00)				
351	P (1.67)						
348				348		P (1.00)	P (33.88)
347				347	P (35.33)		
120		P (1.88)	P (5.88)	120			
119		P (1.88)	P (3.53)	119			
113		P (3.77)	P (2.35)	113			
111			P (3.53)	111			
108			P (18.82)	108			
100				100	P (4.66)		
94				94	P (0.99)	P (0.81)	P (1.85)
87				87			P (2.47)
84				84			P (1.23)
76				76	P (1.84)	P (1.63)	P (4.94)
74				74		P (0.81)	
71				71		P (4.07)	P (6.79)
70	P (4.36)			70			
69				69		P (0.81)	
67				67		P (0.81)	
66				66		P (0.81)	
65				65			P (3.70)
63				63	P (3.25)		P (3.09)
Shared Amplicons							
342	P (93.67)	P (100)	P (74.00)	342	P (61.67)	P (99.00)	P (34.16)
365	P (2.00)			365	P (3.00)		P (2.20)
350			P (13.00)	350			P (29.75)
115		P (3.77)	P (2.35)	115		P (0.81)	P (2.47)
112	P (16.50)	P (3.77)	P (4.71)	112	P (10.31)	P (0.81)	P (3.09)

Patient UMX				Patient UMY			
Amplicon (bp)*	M0 <sup>a</sup> (%)	M3 <sup>a</sup> (%)	M5 <sup>b</sup> (%)	Amplicon (bp)	M0 <sup>c</sup> (%)	M3 <sup>d</sup> (%)	M5 <sup>d</sup> (%)
Shared Amplicons							
103	P (4.85)	P (9.43)	P (5.88)	103	P (2.68)		P (1.23)
102			P (10.59)	102	P (8.62)	P (3.25)	P (29.63)
98		P (5.66)	P (8.24)	98	P (6.92)	P (6.50)	P (3.70)
93		P (3.77)		93	P (0.99)	P (0.81)	
91		P (3.77)	P (4.71)	91	P (3.53)	P (1.63)	P (2.47)
88	P (12.13)			88	P (19.92)	P (1.63)	
86		P (1.88)		86		P (13.01)	
85			P (7.06)	85			P (6.79)
82	P (14.07)			82	P (2.97)		
81			P (1.18)	81	P (2.12)	P (3.25)	P (3.70)
80			P (1.18)	80		P (2.44)	
78		P (3.77)	P (1.18)	78	P (1.69)	P (3.25)	P (1.85)
77	P (8.73)	P (3.77)	P (5.88)	77	P (4.24)	P (1.63)	
75		P (1.88)		75	P (2.26)	P (2.44)	
73	P (8.73)	P (3.77)	P (3.53)	73	P (2.26)	P (3.25)	
72		P (20.75)		72	P (1.27)	P (30.08)	P (11.73)
68		P (3.77)		68		P (0.81)	
64	P (21.35)			64		P (0.81)	P (1.23)
62		P (3.77)	P (4.71)	62	P (18.79)	P (1.63)	P (2.47)
61		P (3.77)		61			P (1.23)
60		P (15.09)	P (4.71)	60		P (12.20)	P (3.70)

\*V1\_V2 amplicons range from 300 - 400 whereas V1 amplicons range from 60 -120 bp.

M0, M3 and M5 refers to collection of samples at time 0, 3 and 5 months

P indicates presence of amplicon with its percentage abundance in the individual profile for that specific region

<sup>a</sup>Patient UMX was treated with Azithromycin, Bactrim, & Tobramycin

<sup>b</sup>Patient UMX was treated with Bactrim

<sup>c</sup>Patient UMY was treated with Levaquin

<sup>d</sup>Patient UMY was treated with Colistin

The dominant amplicons in V1 for the three time points, M0, M3 and M5 were 64 (21.4 %), 72 (20.8 %) and 108 (18.8 %) bp, respectively. For the V1 profile, amplicons 73, 77, 103, and 112 bp were present at all time points. The amplicons 64, 70, 82, 88, 351 and 365 bp were present only during M0. The 352 bp amplicon was found during M0 was absent in M3 and then present again at M5. Amplicons with 61, 68, 72, 75, 86, and 93 bp were only present during M3. Amplicons with 60, 62, 78, 91, 98, 113, 115, 119, and 120 bp appeared at M3 and were also detected at M5. M5 unique amplicons had 80, 81, 85, 102, 108, 111, and 350 bp.

Patient UMY was taking levaquin two weeks prior to the M0 sampling. At least two weeks prior to the second sampling, M3, the therapy was changed to colistin. The LH profiles (number and abundance) varied over time (Table 2.2). The 342 bp amplicon in the V1\_V2 region was the most abundant with 61.7 % at M0 and 99.0 % at M3. However, this amplicon was drastically decreased at M5 (34.2 %) (Table 2.2). The V1 amplicons with 62, 72, 76, 78, 81, 91, 94, 98, 102, and 112 bp were found at all three time points. The amplicons with 63, 103, and 365 bp were present in M0 and M5 but absent in M3. Amplicons unique to M0 were 82, 100, and 347 bp. Amplicons present at M0 and M3 but absent in M5 were 73, 75, 77, 88 and 93 bp. M3 and M5 shared amplicons with 60, 64, 71, and 115 bp. The six unique amplicons present in M5 had 61, 65, 84, 85, 87, and 350 bp (Table 2.2).

**2.4.5 Presumptive identity analysis.** Based on amplicon lengths for both the V1\_V2 and V1 regions, separately or combined, a presumptive identity for an organism can be determined. In attempt to achieve identification of an amplicon, chromosomal

DNA from several CF related pathogens were amplified using V1\_V2 primers and V1 primers. All strains of *Pseudomonas aeruginosa* used in this study amplified a 342 and a 83 bp amplicon for the V1\_V2 and V1 regions, respectively. *Stenotrophomonas maltophilia* amplified a 350 and a 87 bp amplicon for the V1\_V2 and V1 regions, respectively. A 353 bp amplicon was generated for *Staphylococcus aureus* for the V1\_V2 region. However, V1 primers amplified two peaks at 77 and 88 bp for *S. aureus*. *Burkholderia cenocepacia* produced a 339 bp and a 85 bp amplicon for V1\_V2 and V1, respectively.

Many LH profiles did contain amplicons that correspond to these bacterial isolates' fragment sizes. Nineteen LH V1\_V2 profiles may have contained the pathogen *P. aeruginosa* but only one sample contained the corresponding V1 amplicon (83 bp) (Table 2.1). Five samples had the 350 bp amplicon which may be from *Stenotrophomonas maltophilia* but only one sample had the corresponding V1 amplicon. None of the samples had *Staphylococcus aureus* specific V1\_V2 fragment of 353 bp. However, the V1 specific fragments (77/88) were present in 16 samples; nine of these samples contained both fragments. The 339 and 85 bp which may correspond to *B. cenocepacia* were present in two and ten samples, respectively. Two samples contained both amplicons.

The theoretical amplicon lengths for bacteria were determined manually using *in silico* analysis and AmpliQué and then compared to the LH profiles of the strains. The manual *in silico* analysis determined the V1\_V2 fragment length for one specific strain of a CF pathogen. These results agreed with the experimental data. The AmpliQué program was developed to determine the V1\_V2 fragment lengths for all bacterial strains



which had a 16S rRNA sequence in RDPII. The data generated from the program did agree with the experimental data for some strains yet showed different fragments for other strains. For example, in the database some strains of *Stenotrophomonas maltophilia* produce the same 350 bp fragment as the LH result. In addition, many more fragment lengths (304-344 and 346-351 bp) were identified for various strains of *S. maltophilia*. Various isolates of *P. aeruginosa* produced theoretical fragments ranging from 337-345 bp. Only two fragments, 347 and 348, were produced for this region based on *H. influenzae* sequences. AmpliQué determined different isolates of *Staphylococcus aureus* have a V1\_V2 length of 340-344 and 346-351 bp. Based on *B. cenocepacia* sequences, any of the fragments 327, 335, 337, 338, 340, 344, or 346 bp could be generated using the V1\_V2 primers. Most bacterial species produced multiple hits for the V1\_V2 region (data not shown).

AmpliQué was unable to determine the fragment sizes for many CF related pathogens including *P. aeruginosa*. *In silico* analysis was able to determine the fragment lengths of CF pathogens as that the site of primer binding could be detected manually. Again, fragment lengths generated from sequences of individual strains varied from the actual LH results for the specific control strain. Four different isolates of *P. aeruginosa* produced an 86 bp fragment for the V1 region while experimentally the region was 83 bp long. Some strains of *Stenotrophomonas maltophilia* produced the same 87 bp fragment as the control strain while others were determined to have a fragment of 85 or 88 bp. *Staphylococcus aureus* strains were 91 bp long in that region but the test strain did not produce this fragment. *B. cenocepacia* was three bp longer (88 bp) than the LH isolate's profile (85 bp).

## 2.5 Discussion

LH analysis of 16S rRNA genes has been used to detect both known and novel organisms that may be present in many complex microbial communities including sputa from CF patients (Miller, 2003; Mills *et al.*, 2003; Rogers *et al.*, 2003; Suzuki *et al.*, 1998). In these previous studies, only one region of the 16S rRNA was used to identify bacteria present in the community at one time point (Miller, 2003; Rogers *et al.*, 2003). Identification based on one region is difficult and it is hypothesized that the use of multiple regions might provide more information about the community. In this study, the V1 and V1\_V2 regions (separately and combined) were analyzed with LH-PCR to determine the diversity and dynamics of the CF lung. Identification of the most abundant amplicon lengths in the LH profiles was attempted through *in silico* analysis and AmpliQué, a newly designed bioinformatics program. Lastly, the dynamics of the eubacterial community in the sputum over a period of time was further studied using data generated from the LH profiles.

**2.5.1 The V1 region of the 16S rRNA gene provides a more detailed look at the complex bacterial community in the CF lung.** The 16S rRNA gene contains nine variable regions. In this study, the V1 region and the V1\_V2 regions were used to produce bacterial LH profiles of the CF sputa samples (Table 2.1). The V1\_V2 profiles were identical for seven patients and hence were unable to discriminate between patients. The 342 bp amplicon produced a high relative fluorescence in the electropherograms which may have caused the less dominant amplicons to be below the threshold limit and inadvertently ignored (Table 2.1). PCR amplification bias may have also decreased the

detection of other bacteria that were less abundant than the organism(s) represented by the 342 bp amplicon. Therefore, it appeared that seven CF samples had identical bacterial communities. Statistically, the V1\_V2 profiles cannot discriminate between patients at the UM center (Figure 2.3A). The V1 LH profiles were more complex and statically discriminated between samples to a higher degree than the V1\_V2 region (Figure 2.3). The V1 profiles which were unique to each patient contained multiple amplicons per sample (Table 2.1).

The use of multiple regions to increase discrimination between samples was not beneficial. The identical communities shown in the V1\_V2 profiles negated the discriminatory power of the V1 region when analyzed in conjunction with each other. The discriminatory strength of the V1 region as compared to the V1\_V2 region was also evident in a soil community LH study (Moreno *et al.*, 2006). Prior LH analysis on CF sputa used only the V1\_V2 region and more information may have been gained by using the V1 region (Rogers *et al.*, 2003). Using the profiles from the two regions together did not differentiate profiles to the same extent as the V1 region. The combinatorial approach may be of more use in amplicon identification or observing community dynamics.

**2.5.2 CF patients harbor diverse bacterial communities.** The 19 sputum samples obtained from South Florida patients produced unique profiles. No two LH profiles for the V1 region were the same (Table 2.1 and Figure 2.3B). The V1\_V2 region profiles all contained the amplicon length of 342 bp, which presumptively could be identified as *Pseudomonas* species, *Burkholderia* species (but not *B. cenocepacia*), or

*Ralstonia* species. The 342 bp amplicon is indicative of *P. aeruginosa* since control CF isolates produced the same fragment. In addition, the patients were clinically diagnosed with *P. aeruginosa*. The high relative fluorescence of the 342 bp amplicon also indicated a severe infection which would be in agreement with culturing results. The intense fluorescence of the fragment in seven samples may have skewed the community profile causing it to appear as if there was only one bacteria present. The V1 region for these seven samples show some shared amplicons, but largely they had very diverse microbial communities. Though some amplicons (presence not abundance) were common to patients, no two overall profiles were identical to each other (Table 2.1 and Figure 2.3B).

It was thought that there are only a few pathogens commonly found in the lungs of CF patients due to clinical diagnostic procedures (Govan & Nelson, 1993). Molecular techniques have revealed that many more organisms plague the CF lung which is in agreement with our findings (Harris *et al.*, 2007; Rogers *et al.*, 2003). Many of the samples in this study produced LH profiles that contained numerous amplicons, some in low abundance. These less dominant amplicons may represent bacteria that are not routinely cultured and are potentially responsible for some of the differences seen in disease manifestation between CF patients with the same genetic mutations. Recent Sanger sequencing projects have identified uncultured bacteria in the CF lung (Bittar *et al.*, 2008; Harris *et al.*, 2007) Deeper sequencing techniques such as 454 could be used to identify other less abundant and potentially uncultured organisms (Margulies *et al.*, 2005). In order to understand the significance of these organisms within the CF lung, proper statistical tools for comparative metagenomics need to be developed.

**2.5.3 Challenges of identifying organisms in complex communities.** One of the downsides to the LH technique is that the organism found in a sample cannot be conclusively identified. We hypothesized that organisms could be presumptively identified based on the V1\_V2 and V1 amplicon lengths that would be generated from LH-PCR. In this study, the fragment lengths for both regions were experimentally determined for one strain of four known CF pathogens. In addition, *in silico* analysis and AmpliQué were used to determine the expected amplicon length of a variable region for any given organism. This information was correlated with the LH profile in an attempt to identify bacteria present in the CF sputum. The bioinformatics approach highlighted the presence of intraspecies length variation present in both regions. The high degree of variation seen in different isolates drastically decreased the ability to identify a bacteria based on the length of one hypervariable region. It was previously known that one amplicon length could represent multiple species; thereby allowing only a presumptive identification. Based on the absence of a peak, a specific genus or species could be ruled out from being a member of the community. The AmpliQué output clearly demonstrates that one V1\_V2 fragment could represent a multitude of species or genera. The wide range of fragment lengths generated for different isolates of the species may be accurate or it may be due to problems related to the database used in the program. The database contains sequences which are determined by various users. It is possible that some of the length heterogeneity at the isolate level may arise from poor sequencing reads, incorrect trimming of sequences, and old taxonomy references (some organism names have changed). Database issues will change as more genome wide sequencing is performed.

As it stands currently, identification is nearly impossible based solely on amplicon length from the V1\_V2 region.

The V1 region, which was more informative when looking at the community profile as a whole, proved to be more problematic when analyzing individual peaks. Chromosomal DNA of *P. aeruginosa*, *Staphylococcus aureus*, *Stenotrophomonas maltophilia*, and *B. cenocepacia* were amplified using the V1 primers and the fragment lengths were detected. *In silico* analysis was used to determine fragment lengths for the CF associated lung pathogens including those listed above (though not always the same isolate). The initial analysis indicated the presence of length heterogeneity within species. AmpliQué was used to determine the degree of strain variation present in the V1 region. The primer binding parameter was initially set to 100 % stringency which resulted in a very small database of organisms and their corresponding fragment lengths. Many CF pathogens including *P. aeruginosa* were not present at this high stringency. By lowering the stringency, AmpliQué generated more results yet no information was given for *P. aeruginosa*. The V1 region primers (P1F and P2R) experimentally amplified *P. aeruginosa* PAO1. To determine the putative primer binding site, the 16S rRNA sequence for PAO1 was manually examined. Both the forward and reverse primer sequences show some degeneracy. Lowering the stringency parameter was unable to correct for this degeneracy. This issue can be resolved by determining the binding motif of the primers using programs such as iterative enhancement of motifs (IEM) (Zeng *et al.*, 2007a) and then using that motif in AmpliQué. Although P1F is a common V1 primer used extensively in the lab, other less degenerate V1 forward primers such as 27F (which was used to amplify V1\_V2) could be used when attempting to identify

organisms based on amplicon length. Selection of primers for LH-PCR is critical to ensure that all bacteria are being amplified and to increase the success of identification based on fragment length. The widely used primers, 27F and 355R, are based on *E. coli* sequence and should be thought of as generalized primers and not universal primers (Wang & Qian, 2009). Primers that have been developed with a bioinformatics approach may prove to be more useful in future LH studies (Wang & Qian, 2009).

Due to the limitations of the V1 AmpliQué database, identification using two regions simultaneously could not be performed. Therefore, identification based on two regions has yet to be proven. Further modification of AmpliQué and/or the use of different primers or variable regions may eventually lead to LH-PCR being used as an identification technique. At this time, LH-PCR can still be useful when looking at a bacterial community as a whole.

**2.5.4 The eubacterial communities in CF lungs are dynamic.** The LH-PCR profiles provided a view of the CF bacterial community as a whole. It is known that in soil and water samples, the members and their abundances change and these changes can be driven by external factors (Fierer & Jackson, 2006; Shade *et al.*, 2008). Thus, it is hypothesized that the bacterial community within the CF lung also changes based on external factors. To understand the changes in the community, LH-PCR was used to profile three patients over a period of time. Samples from patient UML were used to determine the bacterial community variation over a short period of time (3 hours) and over a long period of time (two years). UMY and UMX were used to study how antibiotics may affect the lung community.

Sputum samples from patient UML were analyzed over the course of two years (M0, M3, M16, M17, M22.1, M22.2, M22.3, M22.4, and M22.5). Five samples were taken on the same day every three hours for twelve hours (M22.1 - M22.5). For both the V1 and the V1\_V2 LH profiles, the overall bacterial community was more similar within the one day then across multiple time periods over two years (Figure 2.4 and 2.5). The LH profiles from the samples taken in one day were not identical. These variations may arise from transient bacteria that were present in the oral cavity at the time of sampling such as *Lactobacillus* from yogurt. Analyzing the samples with both regions further demonstrated that bacterial communities change over time. By combining regions, the data indicated increased similarity of the samples taken in one day which was a result of the more abundant organisms, not the less dominant amplicons in the profile that may represent the transient bacteria in the oral community. The dissimilarity between the samples taken months apart was increased due to the presence of the more dominant possibly disease-causing organisms (Figure 2.4C and 2.5C). This LH analysis is the first to show that colonization of the CF lung is dynamic and these changes can be tracked.

To date the dynamics of the bacterial community in the CF lung in response to external drivers such as the presence of antibiotics has yet to be studied. Thus, a short term study was performed in an attempt to understand how the overall community in the lung changes when under attack by antibiotics. To begin, CF sputum samples from Patient UMX and Patient UMY were analyzed every other month for five months. During this time, patients were taking a combination of antibiotics.

During the five-month period, Patient UMX was taking azithromycin, bactrim and tobramycin. Shortly before the third time point, the patient stopped taking tobramycin.



The patient's profiles contain a large number of V1 amplicons whose presence and abundance change throughout the time course (Table 2.2). For V1\_V2, Patient UMX's profiles consisted predominately of the 342 bp amplicon at all time points (Table 2.2). This fragment was presumptively identified as *P. aeruginosa* which concurred with the patient's clinical diagnosis. Interestingly, UMX had been prescribed tobramycin (which targets *P. aeruginosa*) for the majority of the sampling period, yet the 342 bp amplicon was always present in large abundance (74 – 100 %). LH analysis with both the V1 and V1\_V2 regions indicate a dynamic environment within the lungs of UMX. These changes seen over the time period are likely to be an effect of the antibiotic regime.

Patient UMY's profile was more dynamic than Patient UMX. Patient UMY's sputum showed a high level of abundance for the 342 bp amplicon between the first two samplings followed by a drastic decrease in the abundance of the 342 bp amplicon in the profile from the third sampling (Table 2.2). This change in profile may correspond to the patient's antibiotic regime; specifically, colistin replacing levaquin. This may have caused the pathogen associated with that amplicon to disappear. As the abundance of the 342 bp amplicon decreased, a 348 bp amplicon emerged. It is possible that as one pathogen was cleared, it was replaced with another bacterium that fit that niche (Boni & Feldman, 2005). The decreasing 342 bp amplicon may represent *P. aeruginosa*, which had been cultured from the patient. Levaquin and colistin are both used to target against this pathogen. There are studies that have demonstrated levaquin being less effective in treating *P. aeruginosa* than the first generation fluoroquinolones (Phillips *et al.*, 2000). Antibiotic resistant strains of *P. aeruginosa* are often treated with colistin (Canton *et al.*, 2005; Scheld, 2003) . It could be hypothesized that Patient UMY was

infected with an antibiotic resistant strain of *P. aeruginosa* which may have been susceptible to the stronger antibiotic. The patient's lung infection was then dominated by a bacterium represented by the 348 bp amplicon, which may belong to *H. influenzae* or *Stenotrophomonas maltophilia*. Changes in the bacterial community were also seen in the V1 profiles (Table 2.2). Due to the limitations of AmpliQué, it is not clear if similar changes in abundance were seen in peaks that may represent *P. aeruginosa*.

The microbial community was more dynamic in UMY (a 60-year-old Caucasian man) than UMX (a 23 year-old Hispanic woman). The number of unique amplicons in UMX and UMY were eight and 14, respectively. UMX had five (73, 77, 103, 112, and 342 bp) and UMY had 11 (62, 72, 76, 78, 81, 91, 94, 98, 102, 112, and 342 bp) amplicons that were present at all three time points. These constant amplicons may come from chronic colonizers that no longer respond to the treatment regimen. The uniqueness, diversity and selective abundance of certain amplicons could be attributed to the age, gender and ethnicity. The patients' respective lung flora may have adapted to the patient's immune response and the presence of antibiotics differently. The variations in profiles due to age and antibiotic treatment have been previously demonstrated (Beringer & Appleman, 2000; Lyczak *et al.*, 2002).

Clearly, LH can be used to analyze changes in a microbial community. Although this study used a small sample size over a relatively short period of time, changes in the presence and abundance of amplicons in the LH profiles could be seen. To accurately identify the drivers of the community, a more comprehensive long term study needs to be performed. Armed with the patient's antibiotic information and long term sampling, it

should be possible to determine the effectiveness of a drug against pathogens in the CF lung. This could eventually lead to more effective treatments.

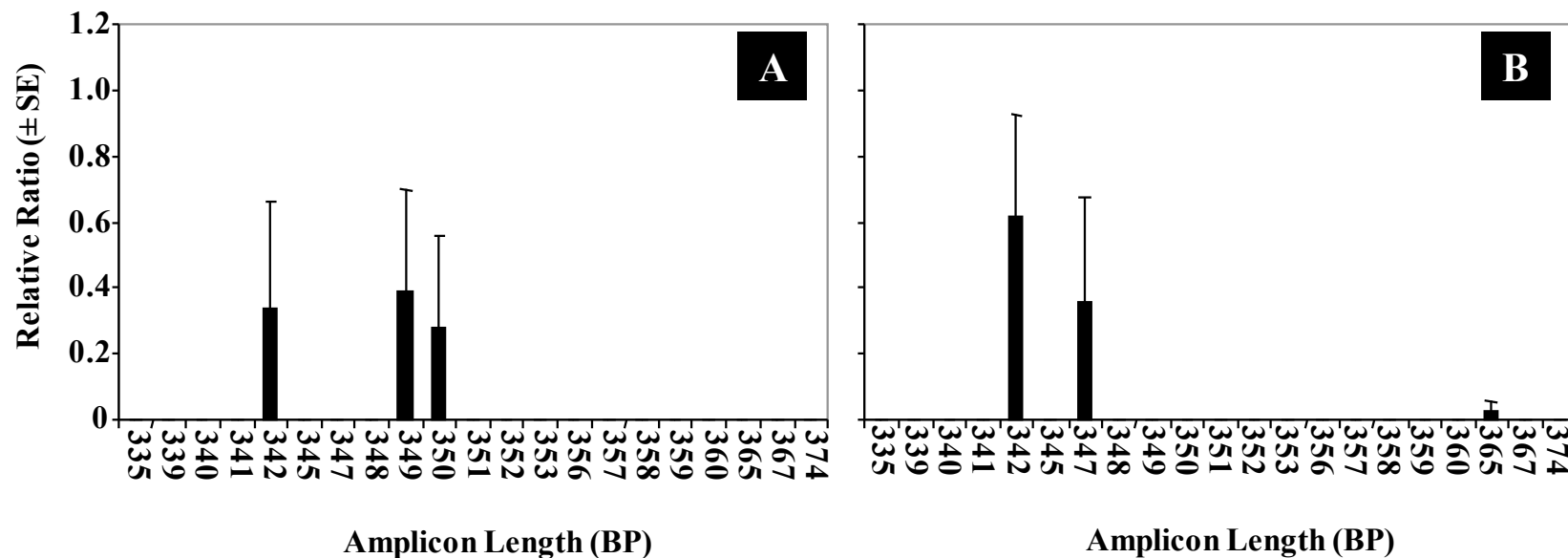
In conclusion, LH-PCR analysis detected the dynamic and complex flora that is present in South Florida CF patients. Although LH-PCR could not be used to identify bacteria in this study, it still proved to be a useful community profiling technique. LH-PCR of the V1\_V2 and to a greater extent the V1 region of the 16S rRNA gene can be used to compare microbial communities between samples. Interpreting LH profiles may provide insight into the evolution of a microbial community and identify the factors that drive these changes. Technologies such as pyrosequencing will likely be used to identify the members present in the bacterial community. However, LH-PCR is still an accessible tool that can be implemented to study microbial communities.

## **2.6 Acknowledgements**

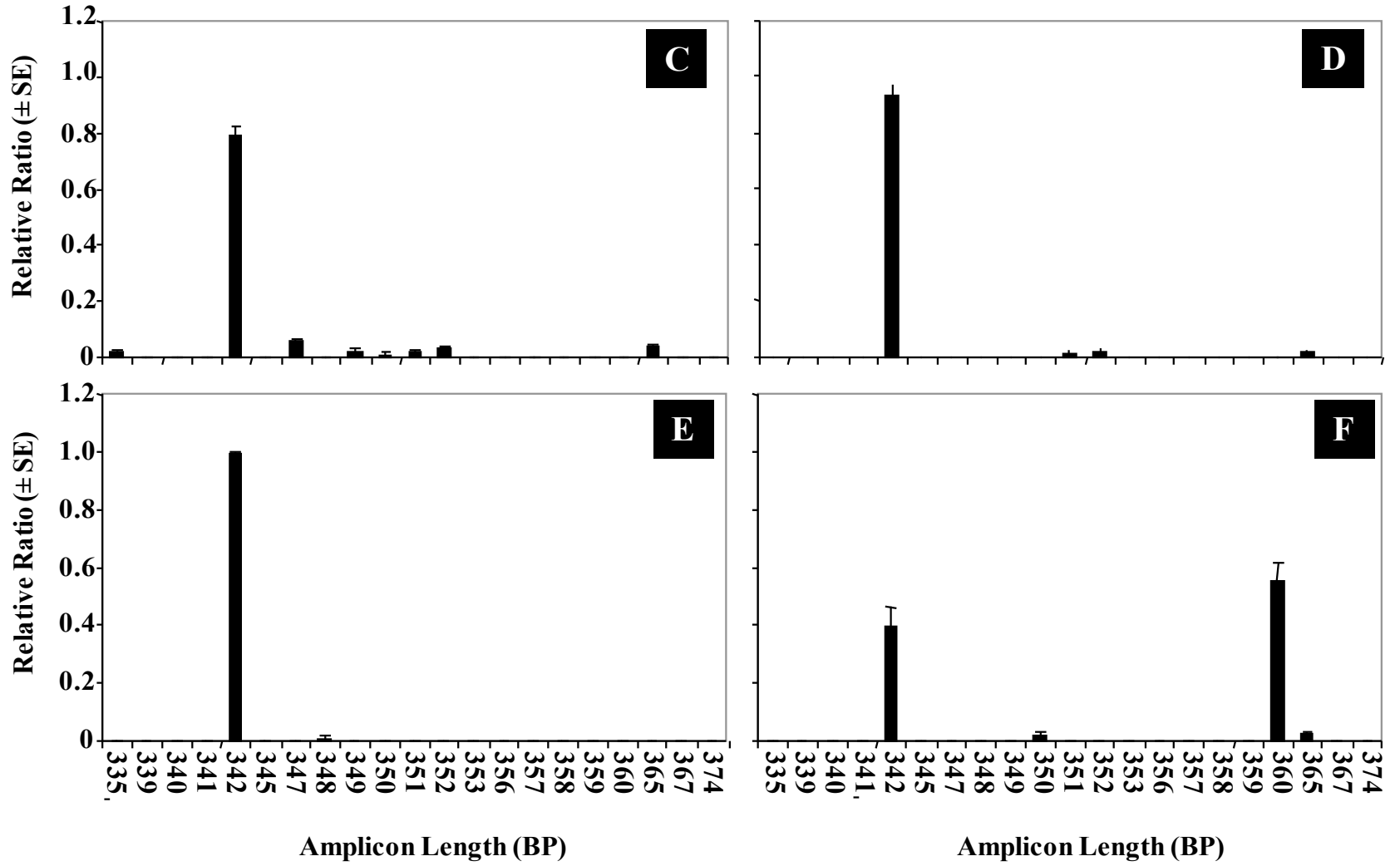
This work has been supported by Florida International Teaching Assistantship (MD) and MBRS Research Initiative for Scientific Enhancement (RISE) Program NIH/NIGMS R25 GM061347 (MD). We would like to thank DeEtta Mills for her many conversations that helped MD troubleshoot the project over the years, Sasha Miller for training MD in the lab and Lilliana Moreno for help with the statistical analysis. We are grateful to Robert Sautter for developing the macros used in the analysis. We also thank Burkhard Tümmler for use of his *P. aeruginosa* strains and Lisa Schneper for the final reading of the manuscript. All LH-PCR samples were analyzed at the FIU Forensic DNA Profiling Facility. Lastly, we are indebted to all South Floridian CF patients who were willing to enroll in our study.

**Supplemental Figure 2.6:** Hypervariable region V1\_V2 amplicon data graphs. Panels A – E represents patients UMA – UME. Panels F – I refers to patients UML, UMM, UMN, UMO, respectively. Panels J, K, and L represents patients UMQ, UMR, and UMS, respectively. Panel M represents profiles from patients UMF, UMG, UMH, UMJ, UMK, and UMT.

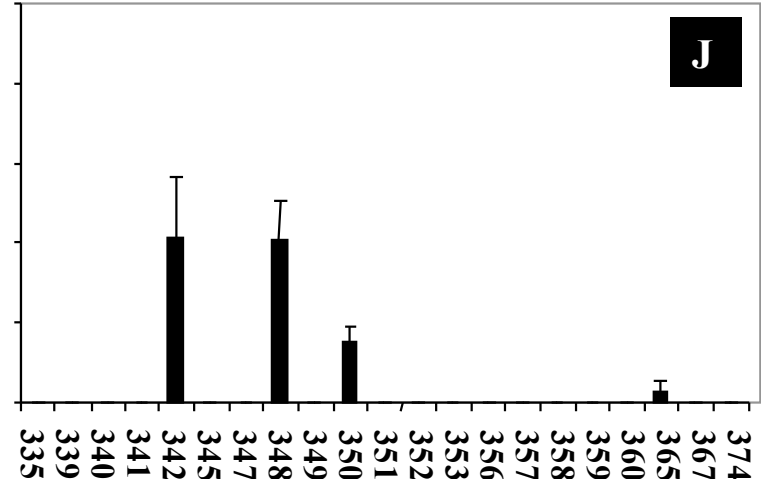
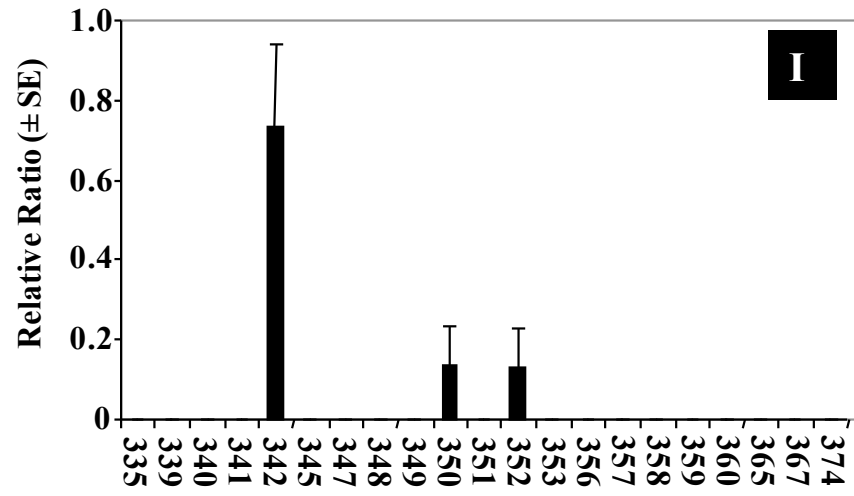
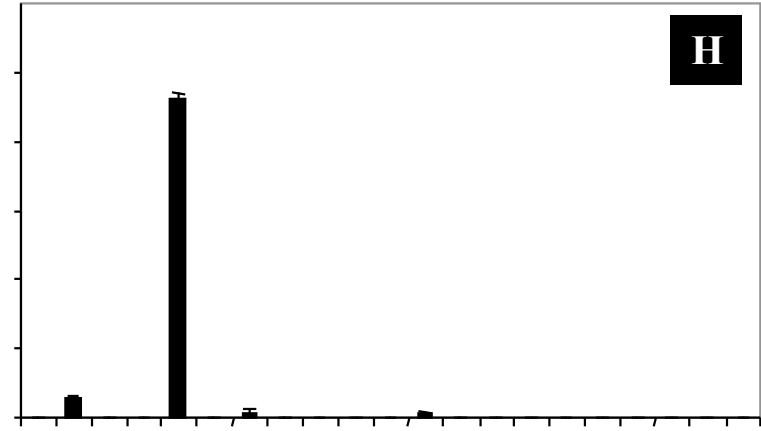
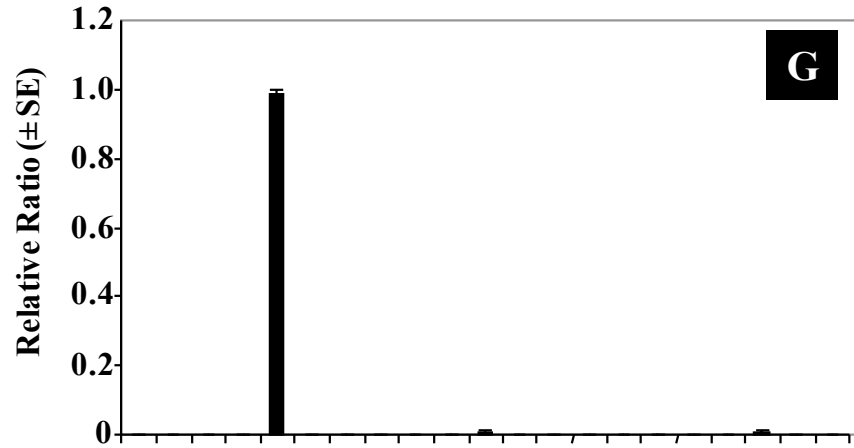
100



Supplemental Figure 2.6



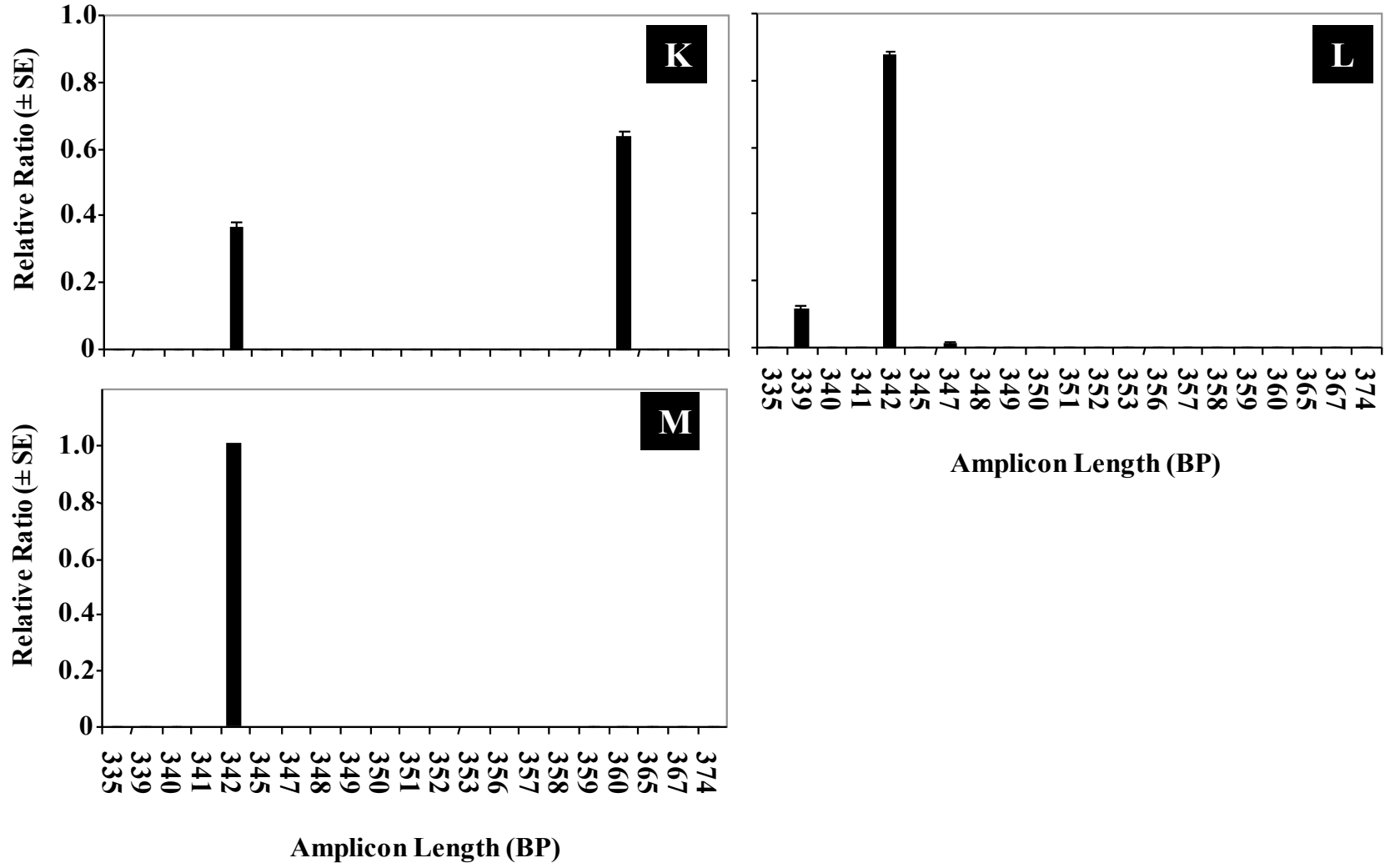
Supplemental Figure 2.6



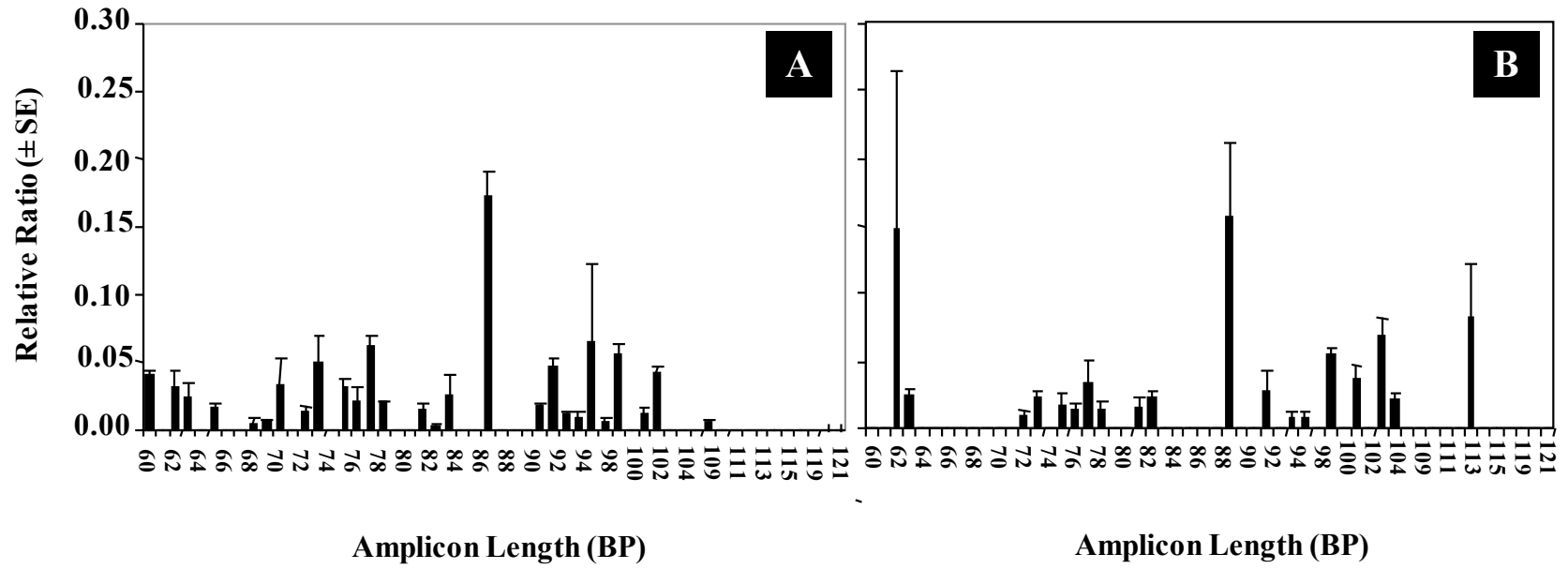
Amplicon Length (BP)

Amplicon Length (BP)

Supplemental Figure 2.6

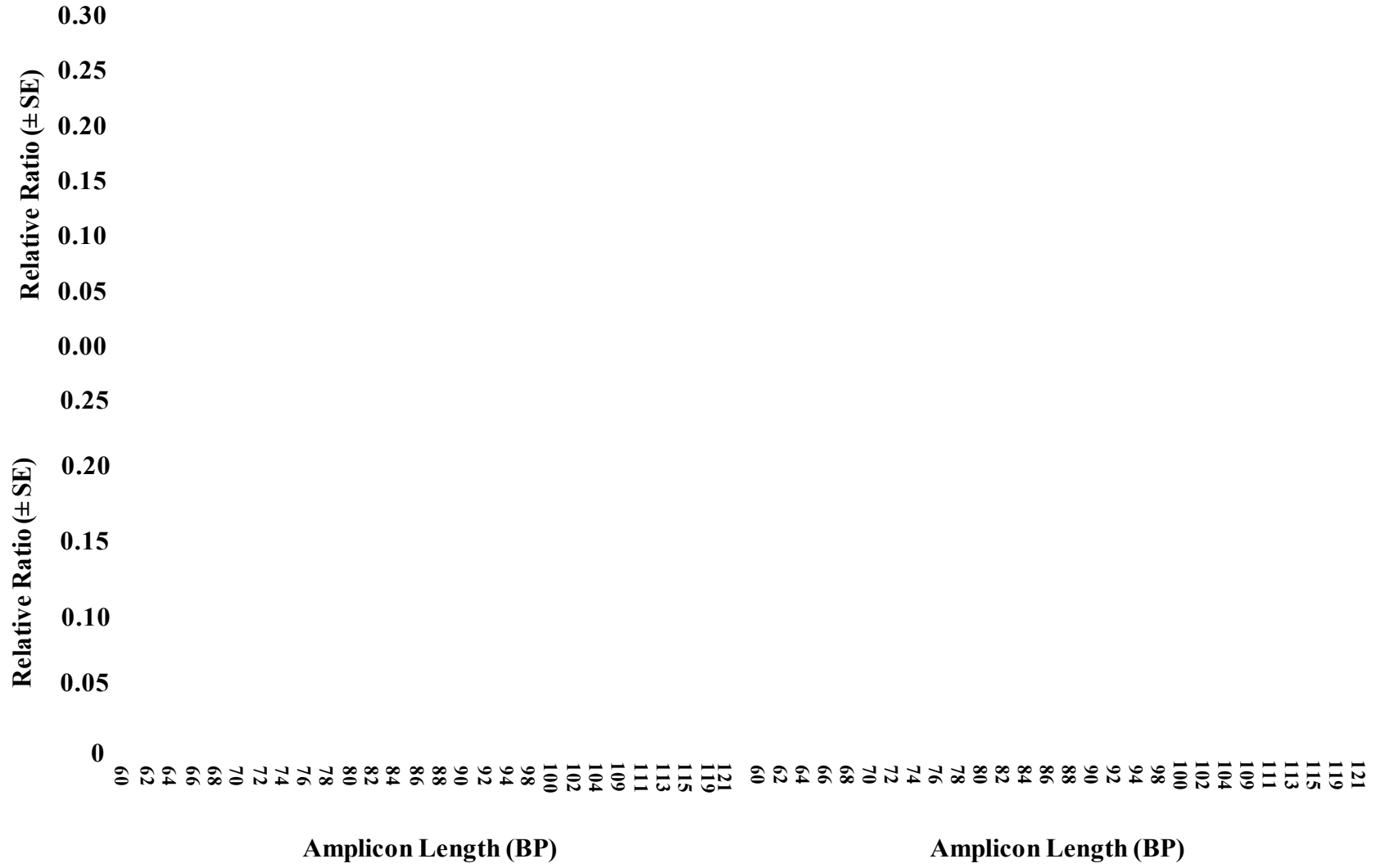


**Supplemental Figure 2.7:** Hypervariable region V1 amplicon data graphs. A-H are the graphs for patients UMA-UMH, respectively. I-S refers to patients UMJ-UMT, respectively.





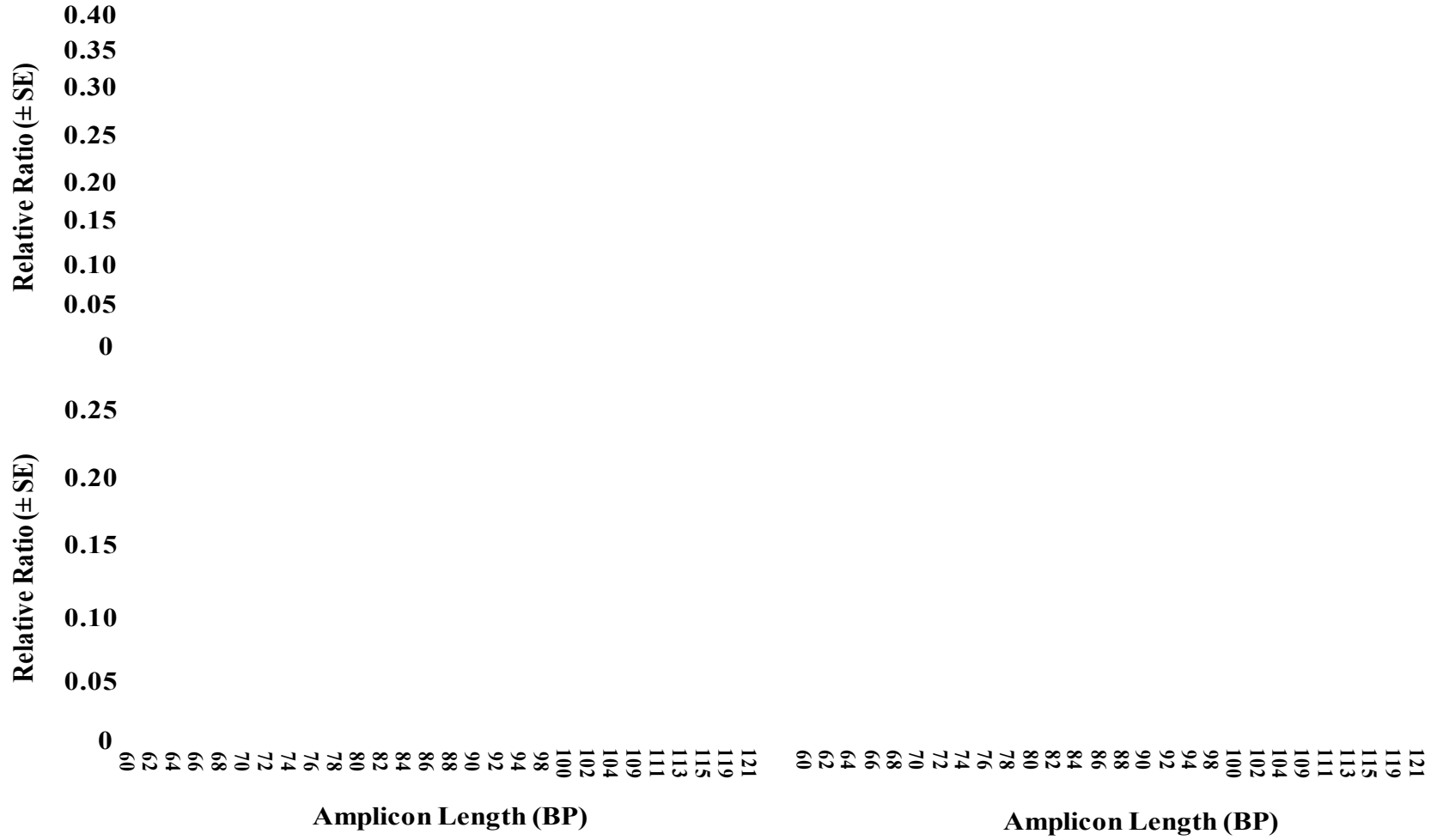
Supplemental Figure 2.7



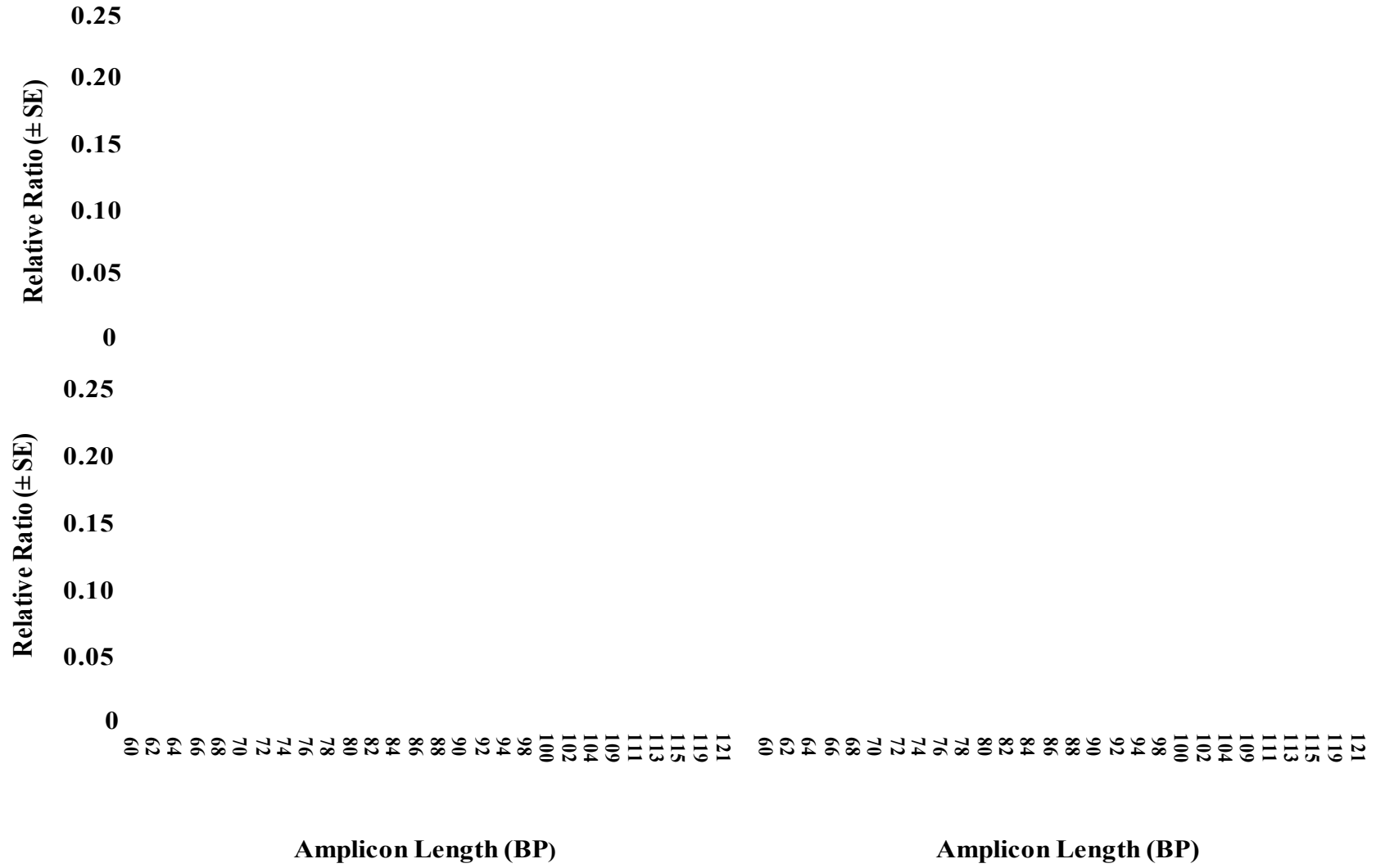
Supplemental Figure 2.7



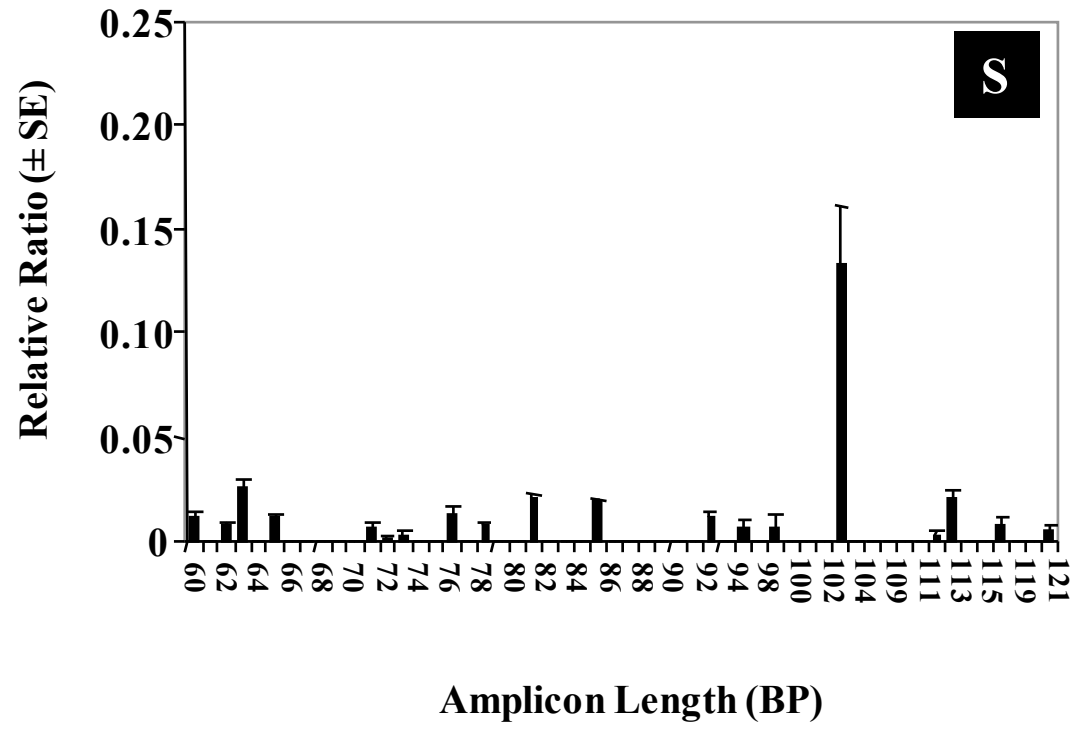
Supplemental Figure 2.7



Supplemental Figure 2.7



Supplemental Figure 2.7



## **Chapter 3**

### **“Cystic Fibrosis Lung Eubacteriome and Mycobiome as Revealed by 454 Sequencing”**

### 3.1 ABSTRACT

Community profiling and Sanger sequencing projects have been used to study the polymicrobial nature of the cystic fibrosis (CF) lung. These studies have shown that the bacterial communities are diverse and unique to each patient. Additionally, species specific PCR and culturing have demonstrated the presence of fungi in the CF lungs. However, the diversity of the fungal community remains to be elucidated. In this study, MultiTag Pyrosequencing<sup>TM</sup> (454) was used to resolve the bacterial and fungal community in sputa from 18 south Florida CF patients. Sequencing results corroborated the polymicrobial nature of the lungs. CF sputa contained DNA from known pathogenic genera, *Pseudomonas*, *Staphylococcus*, *Haemophilus*, *Stenotrophomonas*, and *Burkholderia*. A large percentage of sequenced genera were associated with oral microbiota which may play a role in the progression of lung disease. Potentially novel bacteria, *Chryseomonas* and *Flavimonas* were also identified. In addition, the CF mycobiome dataset contained 33 genera. Many of the species were associated with the environment. Though patients were not being treated for a fungal infection, DNA from pathogenic species of *Aspergillus*, *Exophiala*, *Cryptococcus*, *Wangiella*, *Candida*, and *Pichia* were sequenced from the sputa. The CF microbiomes were unique to each patient and dissimilar to the microbiomes present healthy oral rinses.

**Key words:** cystic fibrosis, 454 sequencing, microbiome, mycobiome, eubacteriome, bacterial community, fungal community, MultiTag Sequencing

**Abbreviations:** CF, cystic fibrosis; BAL, bronchoalveolar lavage; MTPS, MultiTag Pyrosequencing; MDS, non-metric multidimensional scaling; PCO, principle coordinate analysis; T-RFLP, terminal restriction fragment length polymorphism

### 3.2 Introduction

The major cause of morbidity and mortality in CF patients is lung disease as a result from chronic infections (Buzzetti *et al.*, 2009; Patient-registry, 2008). These infections are most often caused by bacteria, many of which are resistant to antibiotics. *Pseudomonas aeruginosa*, *Burkholderia cenocepacia*, *Staphylococcus aureus*, and *Haemophilus influenzae* are routinely cultured from CF sputa (Patient-registry, 2008). Recently, two Sanger sequencing projects have identified the presence of numerous bacteria, many of which were rarely if at all cultured from CF sputa or bronchiolavage samples (Bittar *et al.*, 2008; Harris *et al.*, 2007). Based on sequence analysis, known pathogens, aerobic and anaerobic species were detected in the CF samples. Emerging pathogens such as *Stenotrophomonas maltophilia*, *Serratia marcescens*, and *Achromobacter xylosoxidans* were also identified (Bittar *et al.*, 2008; Harris *et al.*, 2007). In addition, organisms not readily associated with CF, such as *Moxaxella cattharalis*, *Granulicatella elegans*, *Inquilinus limosus*, *Granulicatella adiacens*, *Rothia mucilaginosa*, *Dialister pneumosintes*, *Dolosigranulum pigrum*, *Lysobacter*, and *Coxiellaceae* were identified from either bronchoalveolar or sputum samples from CF patients (Bittar *et al.*, 2008; Harris *et al.*, 2007; Rogers *et al.*, 2003). Oral bacteria, *Prevotella*, *Veillonella*, *Porphyomonas*, and *Streptococcus* were present in numerous samples (Bittar *et al.*, 2008; Harris *et al.*, 2007). Another method, 454 sequencing, (ultra-



deep, high throughput, parallel-sequencing), was used to analyze the lung bacterial community in one adolescent CF girl (Armougom *et al.*, 2009). In this study, 1,000 bp of the 16S rRNA gene present in the metagenomic DNA was amplified. The PCR products were analyzed using 454 sequencing (Armougom *et al.*, 2009). The most commonly identified organisms were *M. catarrhalis*, *D. pigrum*, *G. adiacens*, *H. influenzae* and *Streptococcus mitis*. Sequencing techniques have identified numerous bacteria in the CF lung eubacteriome that have previously gone undetected with culturing methods

CF individuals can also suffer from fungal infections. *Candida albicans* and *Aspergillus fumigatus* are the most commonly diagnosed fungal pathogens (Bakare *et al.*, 2003; Cimon *et al.*, 2001; Mearns *et al.*, 1967). Other fungi such as *Scedosporium apiospermum*, *Scedosporium prolificans*, *Aspergillus terreus*, *Acrophialophora fuispora*, *Exophiala dermatitidis* and *Penicillium emersonii* have been shown to infect CF patients (Bakare *et al.*, 2003; Cimon *et al.*, 1999; Cimon *et al.*, 2003; Cimon *et al.*, 2005; Defontaine *et al.*, 2002; Hennequin *et al.*, 1997; Horré *et al.*, 2004). Recent statistics have shown an increase in the prevalence of fungi in the lung which may have resulted as a consequence of aggressive antibacterial therapy (Nagano *et al.*, 2007).

The prevalence of fungal infections in the CF population varies from one study to another. Using culturing and microscopy analysis, *A. fumigatus* and *C. albicans* were detected in 45.7 % and 75. 5%, respectively of the 369 samples obtained from 94 CF patients over a period of time. In another study, 24.5 % of 201 patients were culture positive for *A. fumigatus* (Burns *et al.*, 1998). Detection of fungal infections is crucial as the infection can lead to complications such as allergic bronchopulmonary aspergillosis (ABPA) and post-transplant fungaemia (Fahy *et al.*, 1991; Kubak, 2002; Mastella *et al.*,

2000; Nagano *et al.*, 2007; Stevens *et al.*, 2003) It is also possible for a fungal infection to become chronic which leads to bronchitis (Nagano *et al.*, 2007).

The total composition of the fungal community (mycobiome) within the CF lung has not been well explored and many of the fungi were identified using culturing and microscopy. Some fungi, such as *Scediosporim* and *Pseudallescheria* are difficult to identify with these methods (Horré *et al.*, 2009; Williamson *et al.*, 2001). In those cases, molecular analysis using fungal ribosomal rRNA genes and the associated internal spacer regions (ITS) are employed (Bouchara *et al.*, 2009; Hinrikson *et al.*, 2005; Kumeda & Asao, 1996; Schmidt & Rath, 2003; Williamson *et al.*, 2001; Zeng *et al.*, 2007b). Most PCR-based assays only identify one organism. An oligonucleotide array can identify 20 fungi simultaneously in CF samples but the panel may not include all fungi present in the mycobiome (Bouchara *et al.*, 2009). Therefore, the CF lung mycobiome may be better characterized with sequencing technique, such as 454 sequencing.

The 454 sequencing technique has been used to used to analyze the fungal mycobiome in the healthy human mouth (Ghannoum *et al.*, 2010). The technique has also identified members of the bacterial communities present in the healthy oral cavity, gut, and vagina (Aas *et al.*, 2005; Keijser *et al.*, 2008; Spear *et al.*, 2008; Turnbaugh *et al.*, 2009; Zaura *et al.*, 2009). The method has also been used to characterize the lung eubacteriome from one CF individual (Armougom *et al.*, 2009). The 454 sequencing technique has been shown to provide a more accurate estimation of diversity and a more reliable estimate of the relative abundance of bacterial species (Armougom *et al.*, 2009; Quince *et al.*, 2009; Sogin *et al.*, 2006).

Sanger sequencing projects have shown the CF lung eubacteriome to be large and phylogenetically diverse. Eubacteriomes are unique to each individual and the differences in microbiota may be influenced by host's factors such as gender, ethnicity, and *cfr* mutation. Therefore, it can be hypothesized south Florida CF patients eubacteriomes will differ and may harbor novel bacteria. It is also likely the lung mycobionome is more diverse than previously determined by culturing and species-specific PCR. This study proposes to resolve the eubacteriome and mycobionome associated with CF sputum from 19 south Florida CF patients by using MultiTag Pyrosequencing™. This analysis will provide a more accurate estimation of diversity for the CF lung microbiome. In addition, the CF sputa microbiomes will be compared with healthy oral rinse microbiomes.

### **3.3 Materials and methods**

**3.3.1 Ethics statement.** Written informed consent was obtained from all participants in this study. Recruitment of study participants was performed according to protocol (number 040407-01) approved by the Human Subjects Institutional Review Board (IRB) of Florida International University, Miami, Fl.

**3.3.2 Study participants.** Sputum samples were obtained from 19 adult CF individuals. The individuals were patients at the University of Miami adult cystic fibrosis care center. Summary demographic information of the study participants was provided in Table 3.1. All CF patients were of the Caucasian race as defined by the United States Census Bureau. Ethnicities, Hispanic and not Hispanic, were included in the

demographics (Table 3.1). Inclusion criteria were >18 years of age, diagnosed to have CF, and proficient in English. Samples were taken during routine visits or during exacerbations. All participants gave an expectorated sputum sample which was immediately kept frozen until analysis.

**Table 3.1:** Patient demographics. Patient number, gender and ethnicity of the study subjects are listed.

<b>Patient</b>	<b>Gender</b>	<b>Ethnicity*</b>
CF1	Male	NH
CF2	Female	NH
CF3	Male	H
CF4	Male	H
CF5	Female	NH
CF6	Male	NH
CF7	Female	H
CF8	Male	H
CF9	Female	NH
CF10	Female	H
CF11	Female	NH
CF12	Male	NH
CF13	Male	NH
CF14	Male	NH
CF15	Female	NH
CF16	Male	NH
CF17	Male	H
CF18	Female	H
CF19	Male	H

\*NH refers to non-Hispanic  
H refers to Hispanic

**3.3.3 DNA extraction.** The sputum samples were liquefied using a freeze/boil method (Reischl *et al.*, 1994). Each sample was separated into three aliquots which were used for metagenomic DNA extraction. The DNA was extracted and cleaned using the FastDNA<sup>®</sup> Spin Kit (Qbiogene, Irvine, CA) according to the manufacture's protocol. Briefly, liquefied sputa were homogenized and cells were mechanically lysed using the FastPrep<sup>®</sup> instrument (Qbiogene) and the lysing matrix. The DNA was cleaned using the GeneClean<sup>®</sup> procedure included in the kit. The DNA obtained from the three aliquots was combined to ensure that the total DNA present in each sputum sample would be analyzed. The eluted DNA was quantified and diluted to 100 ng/μL and refrigerated at 4°C until further use.

**3.3.4 Metagenome analysis.** The eubacteriome and mycobiome analysis of the sputum samples were performed using the MultiTag Pyrosequencing<sup>™</sup> (MTPS) technique (Gillevet, 2006). Bacteria were identified by sequencing the hypervariable region 1 and 2 (V1\_V2) of the 16S rRNA gene using the forward primer 27F (5'- AGA GTT TGA TCM TGG CTC AG-3') and the reverse primer 355R (5'- GCT GCC TCC CGT AGG AGT-3') (Figure 1.13) (Doud *et al.*, 2009; Suzuki *et al.*, 1998). Fungi were identified by sequencing the ITS 1 region using the forward ITS1 primer (5'-TCC GTA GGT GAA CCT GCG G-3') and the reverse primer ITS2 (5'-GCT GCG TTC TTC ATC GAT GC-3') (Figure 1.14) (White *et al.*, 1990).

MTPS technique is a variation of 454 pyrosequencing in which a molecular barcode is incorporated into the fusion primer (Gillevet, 2006). In general, the fusion primers, referred to as Adapter A/B consist of a 20 base sequencing primer, a four base

sequence key (used to calibrate base calling during pyrosequencing), and a PCR primer. In MTPS, Adapter A also contains a 7 base “barcode”. This barcode is used to sort the sequencing reads for downstream analyses. For this experiment, a set of 42 emulsion PCR fusion primers were designed to amplify the microbial DNA. The fungal Adapter A and B were designed to have the ITS1 and ITS2 sequences, respectively. Similarly, the bacterial Adapter A and B contained the 27F and 355R primer, respectively. Therefore, each DNA sample was amplified with sets of uniquely barcoded fungal and bacterial forward and reverse primers. All sequences were read from the A adapter side (forward PCR primer) using the GS-FLX pyrosequencer (454 Life Sciences-Roche, Branford, CT). The incorporated nucleotides were detected by the generation of a bioluminescent signal that is dependent on a luciferase mediated reaction. The intensity of the signal was proportional to the number of incorporated nucleotides. The signals were detected and analyzed by the instrument in real time and a pyrogram composed of a series of peaks reflecting the DNA sequence was generated.

**3.3.5 Data analysis.** A custom PERL script was developed to sort the sequencing reads into bacterial and fungal bins per each patient. The Ribosomal Database Project II (Release 9) Bayesian classifier was used to obtain a classification for each sequence (Maidak *et al.*, 1999). The annotations for each sequence were downloaded, and the PERL script sorted the output into taxa per sample. Using the PERL script, the fungal ITS sequencing data was automatically compared against GenBank (Benson *et al.*, 2008) and identity of the sequence was determined at a 98 % cutoff. The classifications per sequence were again sorted into taxa per sample. The

final output of the PERL script for both the bacterial and fungal data was a table (one table per community) showing the taxon in each sample and its associated abundance determined as number of taxon sequence reads divided by total number of sequenced read per sample

The bacterial data were imported into Primer 5 (Primer-E, Ltd, Ivybridge, United Kingdom) and square root transformed to reduce the influence of the more common organisms (Grant & Ogilvie, 2003; Mills *et al.*, 2003). To determine similarity between bacterial communities, a similarity matrix was calculated using the Bray-Curtis similarity index (Bray & Curtis, 1957; Doud *et al.*, 2009; Mills *et al.*, 2006; Moreno *et al.*, 2006; Rees *et al.*, 2004). Principal coordinate analysis (PCO) was performed on the Bray Curtis distance matrix generated from the CF communities using Multivariate Statistical Package, MVSP (Kovach, Wales, UK; (Ghannoum *et al.*, 2010). Graphically, the PCO analysis plots each community as a rotation of data points in multidimensional space so that the axis with the greatest variance is the first principle component axis (Ghannoum *et al.*, 2010). This analysis will group or separate the samples in the dataset revealing which samples are more similar in the dataset. PCO analyses were used to identify similarities between bacterial communities, bacterial community and age/gender, and between CF sputa and healthy oral rinse eubacteriomes.

Similarities between fungal communities and host's age/gender were determined using the same method. Healthy fungal communities from oral rinses was also compared to the CF sputa mycobiomes (Ghannoum *et al.*, 2010). Presence/absence of taxa were used to construct a Sørensen similarity matrix between the two communities (Sorensen,

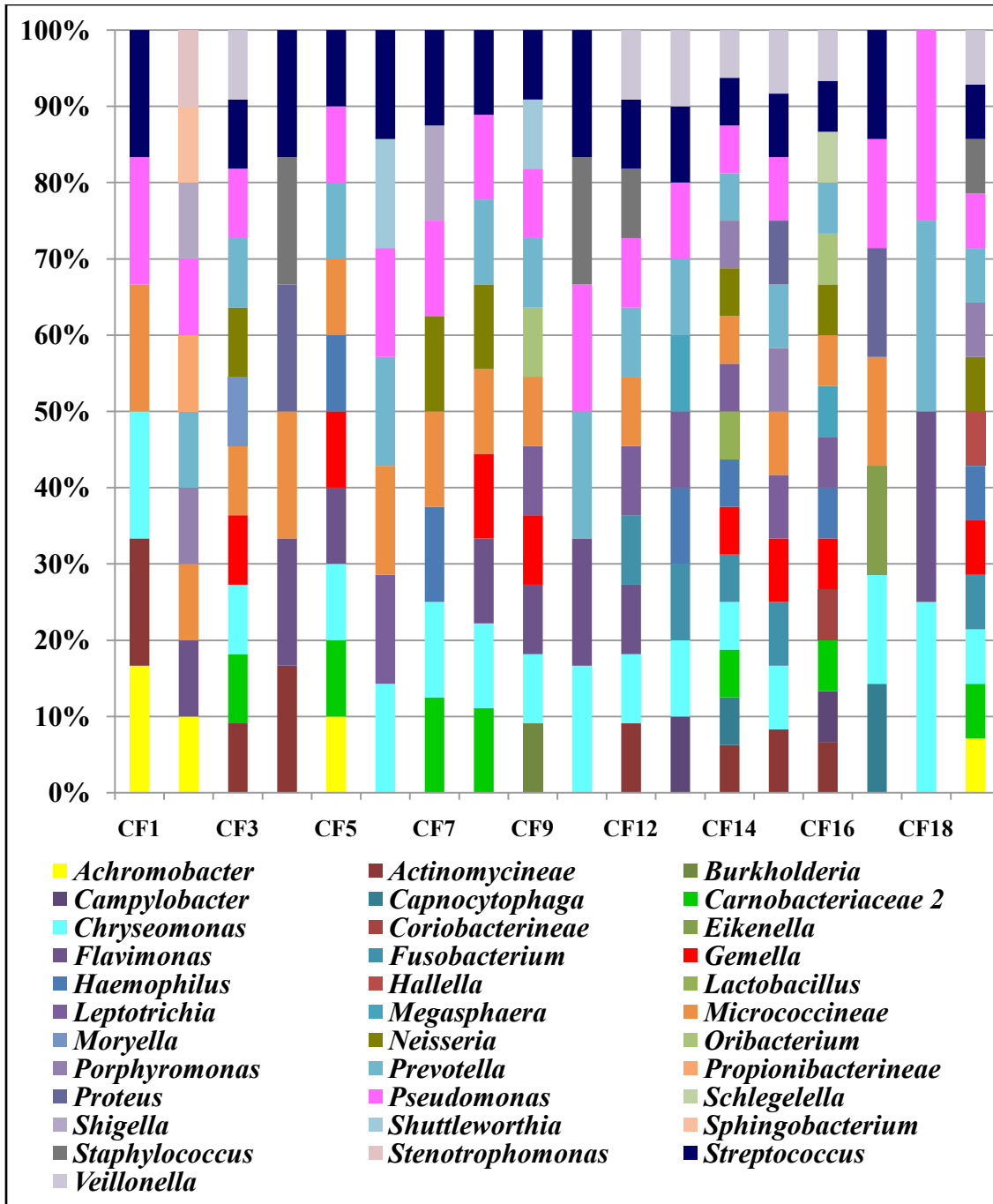
1948). Unweight pair group average clustering was performed on the matrix and results were plotted on a dendograph.

### 3.4 Results

**3.4.1 Participant demographics.** The CF samples were obtained from 11 males and eight females. Of these individuals, 11 were Caucasian of non-Hispanic origin and eight were Caucasian of Hispanic origin (Table 3.1).

**3.4.2 Identification of bacterial taxa in CF samples.** Of the 19 samples, 18 samples produced on average 2,809 MTPS reads. Only 11 reads were obtain from CF\_11. Thus, the sample from a female of non-Hispanic origin was removed from all bacterial analyses. Using a cutoff  $\geq 1\%$  of abundance, a total of 34 taxa were identified from the remaining 18 CF samples (Figure 3.1 and Table 3.2). Sputum samples contained between four and 16 different taxa with the average of  $9.6 \pm 3.3$  genera. Sixteen samples contained sequences belonging to *Pseudomonas* (absent from CF4 and CF16) and *Streptococcus* (absent from CF2 and CF18). *Chryseomonas* was present in 15 samples (absent from CF2, CF4, and CF16). Fourteen samples contained sequences from *Micrococcineae* (absent from CF10, CF13, CF18, and CF19) and *Prevotella* (absent from CF1, CF4, CF7, and CF17). Ten taxa (*Burkholderia*, *Coriobacteriaceae* 2, *Elkenella*, *Haella*, *Moryella*, *Propionbacterineae*, *Schlegelella*, *Spingobacterium*, and *Stenotrophomonas*) were not present in multiple samples. The most abundant taxon in 14 samples was *Chryseomonas*. CF18 was more abundant with





**Figure 3.1:** Bacteria taxon present  $\geq 1\%$  abundance in CF sputum samples. Each color represents a different genus. The height of the color represents the abundance of the reads for that genus in comparison to the total number of sequencing reads for the sample.

**Table 3.2:** Frequency distribution (%) of bacteria present with at least 1 % abundance in the CF sputa eubacteriome

Organism	CF1	CF2	CF3	CF4	CF5	CF6	CF7	CF8	CF9	CF10	CF12	CF13	CF14	CF15	CF16	CF17	CF18	CF19	#
<i>Achromobacter</i>	10.42	4.24			6.05													5.41	4
<i>Actinomycineae</i>	4.90		1.98	1.13							3.48		2.70	2.16	4.31				7
<i>Burkholderia</i>									23.47										1
<i>Campylobacter</i>												1.03			1.22				2
<i>Capnocytophaga</i>													1.09			5.03			2
<i>Carnobacteriaceae</i>			4.60		2.13		1.75	1.51					2.25		1.65			1.44	7
<i>Chryseomonas</i>	67.03		15.73		23.00	25.91	58.39	52.11	14.29	42.93	27.45	19.74	6.45	48.68		32.53	44.08	38.76	15
<i>Coriobacterineae</i>															1.76				1
<i>Eikenella</i>																1.72			1
<i>Flavimonas</i>		23.18		1.39	1.18			3.07	1.59	1.40	1.02						1.52		8
<i>Fusobacterium</i>											1.30	2.15	1.91	3.84				2.30	5
<i>Gemella</i>			2.69		1.35			1.16	1.13				2.18	1.32	1.29			2.74	8
<i>Haemophilus</i>					1.13		1.72					1.37	9.62		3.57			1.97	6
<i>Hallella</i>																		1.12	1
<i>Lactobacillus</i>													1.74						1
<i>Leptotrichia</i>						15.17			19.67		1.14	9.09	7.34	1.91	1.14				7
<i>Megasphaera</i>												1.12			1.53				2
<i>Micrococcineae</i>	5.02	24.37	14.55	3.55	30.31	4.12	1.45	3.59	1.42		1.43		16.62	1.79	2.51	1.46			14
<i>Moryella</i>			2.44																1
<i>Neisseria</i>			2.23				2.85	1.85					4.37		8.20			1.06	6
<i>Oribacterium</i>									1.19						1.45				2
<i>Porphyromonas</i>		15.76											6.18	3.34				6.53	4
<i>Prevotella</i>		2.38	9.34		3.83	21.02		1.74	22.05	2.44	5.91	13.67	2.90	6.61	18.35		7.23	2.09	14
<i>Propionibacterineae</i>		1.85																	1
<i>Proteus</i>				47.58										2.24		32.79			3
<i>Pseudomonas</i>	4.04	2.38	6.32		12.33	15.50	15.63	17.95	3.74	29.09	1.45	6.88	3.82	12.00		18.23	45.07	12.51	16
<i>Schlegelella</i>															7.37				1
<i>Shigella</i>		1.46					1.16												2
<i>Shuttleworthia</i>						1.74			2.04										2
<i>Sphingobacterium</i>		2.12																	1
<i>Staphylococcus</i>				35.38						14.63	30.72							4.09	4
<i>Stenotrophomonas</i>		7.55																	1
<i>Streptococcus</i>	6.62		34.09	4.53	14.16	9.50	9.22	10.02	4.93	4.88	14.82	30.71	23.34	6.08	26.63	4.12		6.47	16
<i>Veillonella</i>			2.66								4.08	7.63	1.91	2.21	7.80			7.68	7
Count*	6	10	11	6	10	7	8	9	11	6	11	10	16	12	15	7	4	14	

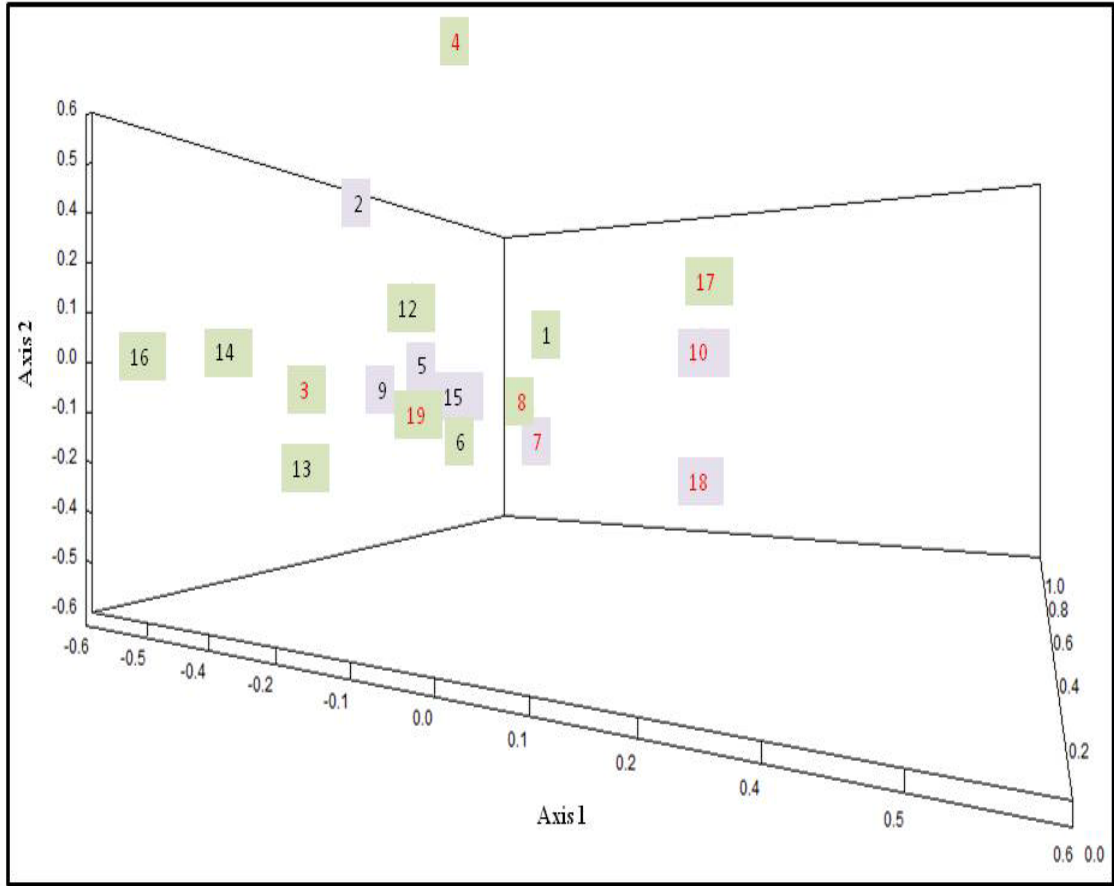
# is the total samples that contained that specific taxon

\* is the total number of taxa present per sample

hits belonging to the *Pseudomonas* genus. *Micrococcineae*, *Proteus* and *Prevotella* were the most sequenced organism in CF2, CF4 and CF15, respectively (Table 3.2).

**3.4.3 Relationship between gender, ethnicity and the eubacteriome.** A PCO analysis was performed on the Bray-Curtis similarity matrix constructed from the CF eubacteriomes to determine if there was an association between the communities and the patient gender and/or ethnicity (Figure 3.2). No clustering was present on the PCO plot regardless of bacterial community profile, patient gender, or ethnicity.

**3.4.4 Healthy oral eubacteriome.** Oral rinse samples were obtained from 25 healthy individuals (data obtained from collaborator). The bacterial community for each sample was determined using MTPS. The data was processed in the same manner as the CF data. Only taxa with greater than 1 % abundance were included in the analysis (Table 3.3). The 25 samples had between eight and 18 bacterial taxa with the average of  $12.84 \pm 2.37$  taxa. *Prevotella*, *Streptococcus*, and *Veillonella* were present in all samples. *Micrococcineae* was present in all except for H9. Twenty-three taxa were not present in multiple samples i.e. each taxon was only sequence from one sample, and not necessarily the same sample. The most abundant organism in 12 samples was *Streptococcus*. *Neisseria* and *Prevotella* were most abundant in six (H8, H11, H12, H21, H23, and H25) and five (H2, H3, H6, H18, H24) samples, respectively. *Porphyromonas*, *Lepotrichia*, and *Haemophilus* and were found to be most abundant in one sample (H9, H10, and H13, respectively).



**Figure 3.2:** Relationship between the host and the eubacteriome. Principal coordinate analysis (PCO) was used to represent the relationship between the bacterial communities as represented by the Bray-Curtis similarity matrix. Lilac squares are samples from females. Males are represented by green squares. Numbers in black indicate samples from non-Hispanic subjects while red text represents Hispanic samples.

**Table 3.3:** Frequency distribution (%) of bacteria present with at least 1 % abundance in the healthy mouth eubacteriome

Organism	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	H16	H17	H18	H19	H20	H21	H22	H23	H24	H25	#	
<i>Achromobacter</i>																										1	
<i>Actinobacillus</i>						3.61										4.77		1.03		1.61				1.89		5	
<i>Actinomycineae</i>	1.86	2.09	2.71				2.87	2.45		5.67		1.53					2.16	4.15					1.75	3.71		11	
<i>Aggregatibacter</i>						1.13										1.88				1.38						3	
<i>Alkaliflexus</i>								1.17																		1	
<i>Aquitalea</i>																							1.26			1	
<i>Bacteroides</i>					1.91		1.33		1.32			1.93				1.39	1.89						1.59		3.71	8	
<i>Bergeyella</i>																				3.58						1	
<i>Bulleidia</i>												1.58														1	
<i>Burkholderia</i>																							1.18			1	
<i>Campylobacter</i>	1.55	1.69	1.26							2.06								2.30							1.02	6	
<i>Capnocytophaga</i>						1.39				2.97									2.18	2.12						4	
<i>Carnobacteriaceae 2</i>	2.01	1.69		1.67	3.19		1.53		1.28				1.64	2.27		2.32			3.67	1.14			1.06		1.18	13	
<i>Coriobacterineae</i>																								1.65		1	
<i>Corynebacterineae</i>										2.14																1	
<i>Delftia</i>																							2.02			1	
<i>Eikenella</i>																										1	
<i>Elizabethkingia</i>																				3.22						1	
<i>Empedobacter</i>							1.83																			1	
<i>Enhydrobacter</i>																1.30										1	
<i>Filifactor</i>									4.22																	1	
<i>Fusobacterium</i>	2.79	6.23	4.50		6.76	7.38	3.26	4.58	7.76	5.14	9.21	1.27	4.74			1.75	2.39	5.67	2.90	3.42	5.49		3.22	3.18	4.34	21	
<i>Gemella</i>				5.06	2.06		1.88							2.62	4.91	5.30	1.45			2.67		4.54	1.90		1.81	11	
<i>Haemophilus</i>	8.15	12.98		9.75	11.13	9.15	5.24	6.11	1.77	5.89	6.54	12.44	24.40	16.64	7.87	17.16	12.58	6.60	7.37	10.18	8.42	2.10	8.50	8.14	14.26	24	
<i>Hallella</i>														1.51												1.02	2
<i>Kingella</i>																				1.22						1	
<i>Kordia</i>												1.18	1.41					1.52	1.22					1.48	1.05	6	
<i>Leptotrichia</i>	10.42	2.46	1.79		1.03	1.17		1.15	4.63	11.52	3.02	3.55					1.01	6.20					1.06		1.36	14	
<i>Megasphaera</i>		1.21																								1	
<i>Micrococineae</i>	3.20	10.15	2.98	7.09	1.13	1.54	11.32	6.82		2.40	7.68	7.14	10.61	13.44	9.27	4.81	1.01	13.53	11.80	2.71	2.97	2.69	1.90	8.26	2.32	24	
<i>Mycoplasma</i>			1.39						1.96																	2	
<i>Neisseria</i>	15.02	11.91	2.12	22.90	21.81	17.36		27.81	11.34	10.40	20.21	24.53	11.31	1.45	11.81	13.94	3.27		22.03	24.53	33.07	6.22	21.78		15.13	22	
<i>Oribacterium</i>	1.03	1.25		1.04				2.17	1.15		2.06	1.63							3.37						1.69	9	
<i>Pasteurella</i>														1.75												1	
<i>Pelomonas</i>																										1	
<i>Phocoenobacter</i>													1.17	1.11												2	
<i>Porphyromonas</i>	1.24	4.22	16.94	1.15	11.86	3.95		5.34	22.49	6.38	2.23	3.02	3.24	1.16		4.28	7.89		2.98	3.18	8.23	1.26	7.32		8.44	21	
<i>Prevotella</i>	10.01	14.52	19.99	9.55	4.85	19.09	14.98	6.11	15.22	6.34	16.54	16.78	13.80	16.17	5.76	2.16	11.06	22.33	4.96	3.85	8.05	1.76	13.58	26.42	9.86	25	
<i>Propionibacterineae</i>																							9.83			1	
<i>Proteiniphilum</i>					1.32											1.14										1.36	3
<i>Ralstonia</i>																							1.01			1	
<i>Schlegelella</i>		1.72						3.16	1.73	1.54	3.52	1.23			2.84				4.35							8	
<i>Sebadella</i>										5.37																	1
<i>Selenomonas</i>											1.05																1
<i>Sneathia</i>											1.09																1
<i>Sphingomonas</i>																							1.34			1	
<i>Streptococcus</i>	18.63	9.93	15.02	28.48	23.53	14.61	38.51	19.79	6.37	10.47	8.77	6.75	16.10	21.41	38.00	30.29	30.56	15.63	28.14	28.46	23.06	26.39	17.68	25.17	14.56	25	
<i>Tannerella</i>									1.32		1.19																2
<i>Treponema</i>			4.04						2.41																		2
<i>Veillonella</i>	6.50	8.98	12.57	4.54	3.48	4.10	7.46	7.20	2.00	7.70	3.32	8.72	3.47	12.10	7.11	1.51	9.92	10.16	1.87	1.22	5.30	3.19	8.16	12.08	5.21	25	
<b>Count*</b>	13	15	12	10	13	12	12	12	16	18	12	14	11	12	8	15	13	12	12	15	8	16	14	11	15		

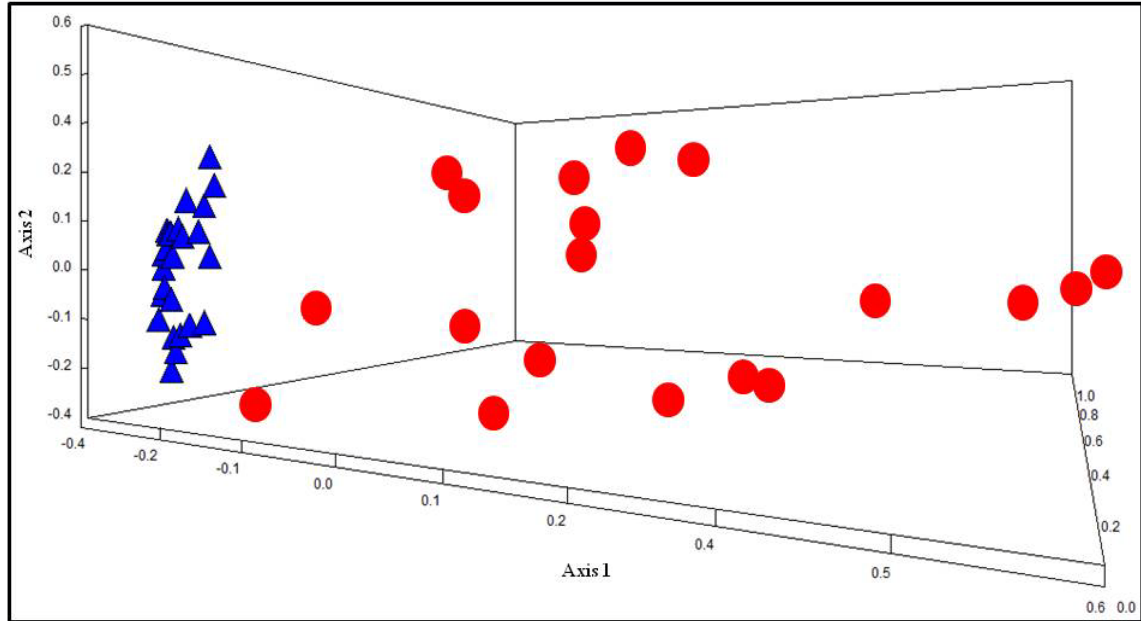
# is the total samples that contained that specific taxon

\* is the total number of taxa present per sample

**3.4.5 Comparison of CF and healthy eubacteriomes.** Twelve taxa were present in CF samples but not in healthy samples (Table 3.4). Common CF pathogens *Pseudomonas*, *Staphylococcus*, and *Stenotrophomonas* were only sequenced in the CF samples. In addition, *Chryseomonas*, *Eikenella*, *Flavimonas*, *Lactobacillus*, *Moryella*, *Proteus*, *Shigella*, *Shuttleworthia*, and *Sphingobacterium* were exclusively present in CF samples. Twenty-seven taxa were unique to healthy oral samples (Table 3.4). A similarity index of the two dataset was calculated using the Bray-Curtis similarity index on the square root transformed abundance data. The similarity matrix was used in the PCO analysis. The PCO graph showed the data points (blue triangles) representing the healthy oral rinse microbiomes clustered closer to each other than the CF microbiomes (red circles) (Figure 3.3).

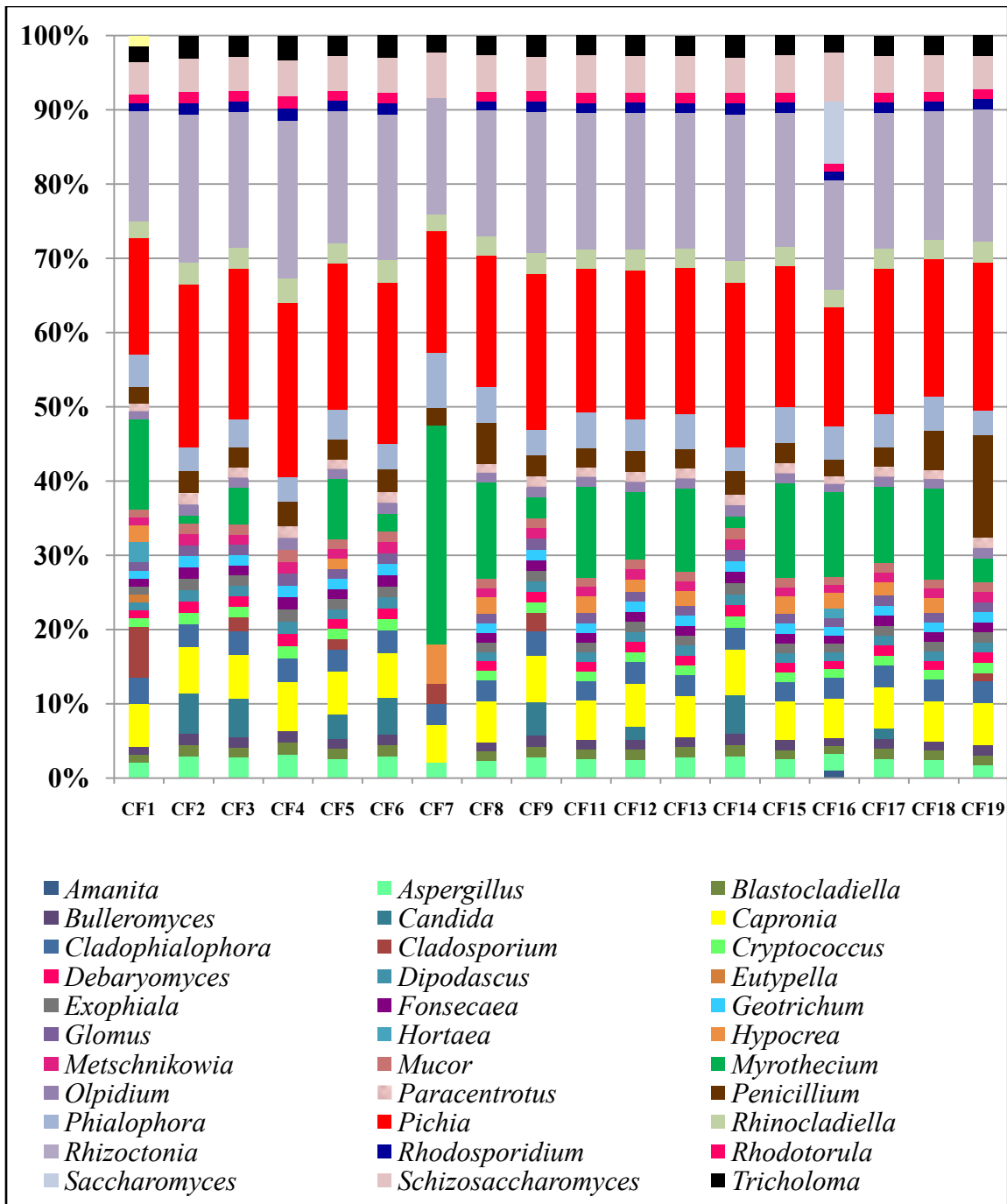
**Table 3.4:** Bacteria sequenced in CF sputa and/or mouth rinse from healthy individuals

Only in CF	Only in Healthy		Present in Both	
<i>Chryseomonas</i>	<i>Actinobacillus</i>	<i>Kordia</i>	<i>Achromobacter</i>	<i>Neisseria</i>
<i>Eikenella</i>	<i>Aggregatibacter</i>	<i>Mycoplasma</i>	<i>Actinomycineae</i>	<i>Oribacterium</i>
<i>Flavimonas</i>	<i>Alkaliflexus</i>	<i>Pasteurella</i>	<i>Burkholderia</i>	<i>Porphyromonas</i>
<i>Lactobacillus</i>	<i>Aquitalea</i>	<i>Pelomonas</i>	<i>Campylobacter</i>	<i>Prevotella</i>
<i>Moryella</i>	<i>Bacteroides</i>	<i>Phocoenobacter</i>	<i>Capnocytophaga</i>	<i>Propionibacterineae</i>
<i>Proteus</i>	<i>Bergeyella</i>	<i>Proteiniphilum</i>	<i>Carnobacteriaceae 2</i>	<i>Schlegelella</i>
<i>Pseudomonas</i>	<i>Bulleidia</i>	<i>Ralstonia</i>	<i>Coriobacterineae</i>	<i>Streptococcus</i>
<i>Shigella</i>	<i>Corynebacterineae</i>	<i>Sebaldella</i>	<i>Fusobacterium</i>	<i>Veillonella</i>
<i>Shuttleworthia</i>	<i>Delftia</i>	<i>Selenomonas</i>	<i>Gemella</i>	
<i>Sphingobacterium</i>	<i>Elizabethkingia</i>	<i>Sneathia</i>	<i>Haemophilus</i>	
<i>Staphylococcus</i>	<i>Empedobacter</i>	<i>Sphingomonas</i>	<i>Hallella</i>	
<i>Stenotrophomonas</i>	<i>Enhydrobacter</i>	<i>Tannerella</i>	<i>Leptotrichia</i>	
	<i>Filifactor</i>	<i>Treponema</i>	<i>Megasphaera</i>	
	<i>Kingella</i>		<i>Micrococcineae</i>	



**Figure 3.3:** PCO analysis of healthy mouth rinse and CF eubacteriomes. Bray-Curtis similarity matrix and PCO was used to compare the sequenced eubacteriomes from healthy mouth rinse samples (blue triangles) and CF sputa (red circles).

**3.4.6 Identification of fungal taxa in CF samples.** Of the 19 samples, 18 samples produced between 5,467 and 45,178 reads with the average read count of 21,976. Only 3453 reads were obtained from CF\_10, a female Hispanic patient. The number of sequencing reads is low in comparison to the rest of the samples and thus CF\_10 was removed from all fungal analyses. Using a cutoff of  $\geq 1\%$  of abundance, 33 taxa were identified from the remaining 18 CF samples (Figure 3.4 and Table 3.5). Sputum samples contained between 13 and 30 different taxa with an average of  $26 \pm 3$  (Table 3.5). Sample CF1 was the most diverse community with 30 taxa. Twenty-nine and 28 taxa were present CF16 and CF5, respectively. Twenty-seven taxa were sequenced in



**Figure 3.4:** Fungi genera present in  $\geq 1\%$  abundance in CF sputum samples. Each color represents a different genus. The height of the color represents the abundance of the reads for that genus in comparison to the total number of reads for the sample.



**Table 3.5:** Frequency distribution (%) of fungi present with at least 1 % abundance in the CF sputa mycobiome

Organism	CF1	CF2	CF3	CF4	CF5	CF6	CF7	CF8	CF9	CF11	CF12	CF13	CF14	CF15	CF16	CF17	CF18	CF19	#
<i>Amanita</i>															1.06				1
<i>Aspergillus</i>	2.10	2.93	2.67	3.13	2.60	2.89	1.78	2.29	2.76	2.59	2.46	2.80	2.92	2.51	2.16	2.58	2.32	1.69	18
<i>Blastocladiella</i>	1.05	1.44	1.33	1.59	1.31	1.44		1.18	1.38	1.28	1.31	1.29	1.46	1.28	1.06	1.29	1.19	1.30	17
<i>Bulleromyces</i>	1.07	1.48	1.35	1.60	1.32	1.46		1.23	1.39	1.30	1.33	1.31	1.47	1.30	1.12	1.32	1.21	1.32	17
<i>Candida</i>		5.42	4.99		3.31	4.84			4.47		1.78		5.26			1.39			8
<i>Capronia</i>	5.68	6.07	5.82	6.44	5.59	5.94	4.24	5.26	6.03	5.25	5.58	5.41	5.97	5.21	5.18	5.48	5.20	5.52	18
<i>Cladophialophora</i>	3.52	3.08	3.09	3.22	2.95	3.01	2.34	2.78	3.23	2.62	2.90	2.78	3.00	2.60	2.91	2.83	2.73	2.86	18
<i>Cladosporium</i>	6.85		1.80		1.43		2.31		2.39									1.07	6
<i>Cryptococcus</i>	1.07	1.49	1.36	1.61	1.32	1.46		1.21	1.39	1.30	1.33	1.31	1.48	1.29	1.12	1.31	1.21	1.32	17
<i>Debaryomyces</i>	1.08	1.49	1.36	1.61	1.32	1.46		1.23	1.40	1.30	1.34	1.31	1.48	1.30	1.12	1.32	1.21	1.32	17
<i>Dipodascus</i>	1.07	1.49	1.36	1.61	1.32	1.46		1.23	1.40	1.30	1.33	1.31	1.49	1.30	1.12	1.32	1.21	1.32	17
<i>Eutypella</i>	1.03																		1
<i>Exophiala</i>	1.08	1.49	1.36	1.61	1.32	1.46		1.23	1.40	1.31	1.34	1.31	1.49	1.30	1.13	1.32	1.23	1.33	17
<i>Fonsecaea</i>	1.08	1.49	1.36	1.61	1.32	1.46		1.23	1.40	1.31	1.34	1.31	1.49	1.30	1.13	1.32	1.23	1.33	17
<i>Geotrichum</i>	1.08	1.49	1.36	1.61	1.32	1.46		1.23	1.40	1.31	1.34	1.31	1.49	1.30	1.13	1.32	1.23	1.33	17
<i>Glomus</i>	1.08	1.50	1.36	1.61	1.32	1.46		1.23	1.40	1.31	1.34	1.31	1.49	1.30	1.13	1.32	1.23	1.33	17
<i>Hortaea</i>	2.76														1.32				2
<i>Hypocrea</i>	2.17				1.42		4.49	2.27		2.22	1.64	1.98		2.30	2.09	1.81	1.98		11
<i>Metschnikowia</i>	1.04	1.44	1.32	1.55	1.29	1.44		1.15	1.36	1.29	1.31	1.28	1.46	1.25	1.06	1.27	1.18	1.29	17
<i>Mucor</i>	1.04	1.44	1.32	1.55	1.29	1.44		1.15	1.37	1.28	1.31	1.29	1.46	1.27	1.06	1.28	1.18	1.29	17
<i>Myrothecium</i>	12.11	1.01	4.75	0.00	7.94	2.37	24.79	12.62	2.68	12.23	8.95	10.89	1.54	12.70	11.32	10.00	11.62	3.10	17
<i>Olpidium</i>	1.08	1.49	1.36	1.61	1.32	1.46		1.23	1.40	1.31	1.34	1.31	1.49	1.30	1.13	1.32	1.23	1.33	17
<i>Penicillium</i>	2.13	2.98	2.71	3.21	2.64	2.92	1.89	5.28	2.78	2.61	2.73	2.62	3.06	2.75	2.23	2.63	5.03	13.46	18
<i>Phialophora</i>	4.31	3.17	3.56	3.35	4.01	3.36	6.31	4.71	3.26	4.82	4.27	4.59	3.24	4.83	4.35	4.43	4.40	3.20	18
<i>Pichia</i>	15.61	21.49	19.81	23.21	19.35	21.46	13.73	17.12	20.46	19.28	19.71	19.25	21.88	18.83	15.99	19.10	17.61	19.42	18
<i>Rhinocladiella</i>	2.15	2.99	2.72	3.22	2.65	2.93	1.90	2.46	2.79	2.62	2.68	2.62	2.97	2.60	2.27	2.64	2.46	2.66	18
<i>Rhizoctonia</i>	14.82	19.53	17.80	21.03	17.65	19.27	13.20	16.33	18.44	18.45	18.11	17.90	19.48	17.97	14.67	18.05	16.52	17.33	18
<i>Rhodosporidium</i>	1.08	1.49	1.36	1.61	1.32	1.46		1.23	1.40	1.31	1.34	1.31	1.49	1.30	1.13	1.32	1.23	1.33	17
<i>Rhodotorula</i>	1.08	1.49	1.36	1.61	1.32	1.46		1.23	1.40	1.31	1.34	1.31	1.49	1.30	1.13	1.32	1.23	1.33	17
<i>Saccharomyces</i>															8.29				1
<i>Schizosaccharomyces</i>	4.32	4.48	4.51	4.83	4.68	4.61	5.09	4.83	4.43	5.05	4.85	4.92	4.59	5.04	6.51	4.86	4.78	4.27	18
<i>Tricholoma</i>	2.15	2.99	2.72	3.22	2.65	2.93	1.90	2.46	2.79	2.64	2.68	2.62	2.97	2.60	2.27	2.64	2.46	2.66	18
<i>Wangiella</i>	1.38																		1
Count*	30	26	27	24	28	26	13	26	27	26	27	26	26	26	29	27	26	26	

# is the total samples that contained that specific taxon

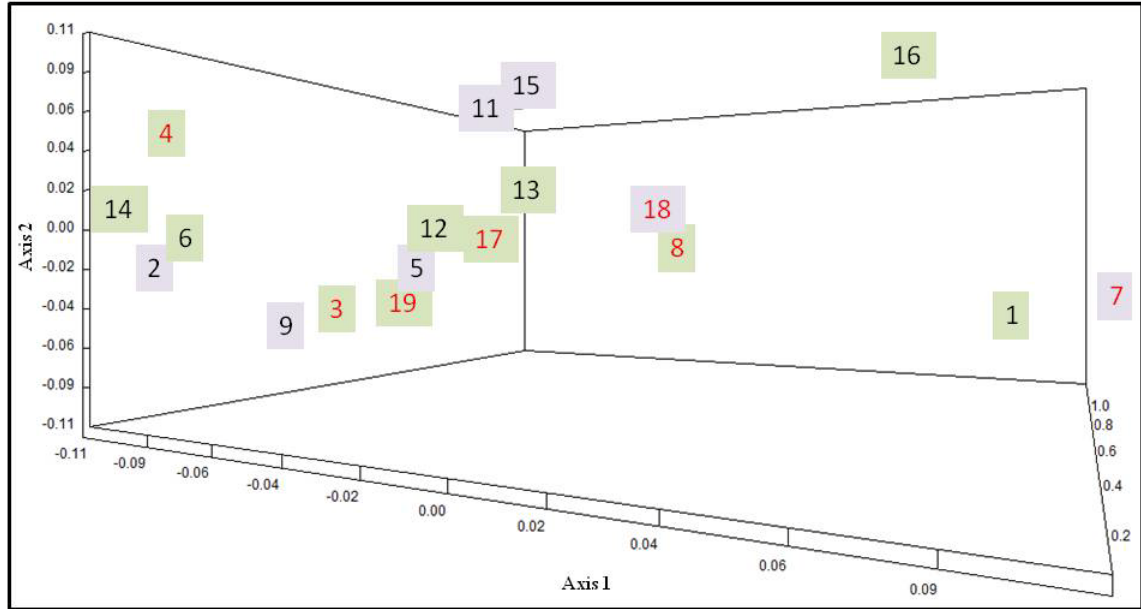
\* is the total number of taxa present per sample

four samples (CF3, CF9, CF12, and CF17). Twenty-six taxa were present in nine samples (CF2, CF6, CF8, CF11, CF13, CF14, CF15, CF18, and CF19). Twenty-four and 13 taxa were identified from samples CF4 and CF7, respectively.

Ten genera were present in all 18 samples (*Aspergillus*, *Capronia*, *Penicillium*, *Phialophora*, *Pichia*, *Rhinocladiella*, *Rhizoctonia*, *Schizosaccharomyces*, and *Tricholoma*). Fifteen genera were also common in 17 samples (*Blastocladiella*, *Bulleromyces*, *Cryptococcus*, *Debaryomyces*, *Dipodascus*, *Exophiala*, *Fonsecaea*, *Geotrichum*, *Glomus*, *Metschnikowia*, *Mucor*, *Myrothecium*, *Rhodosporidium*, and *Rhodotorula*). Sequences for *Amanita* and *Saccharomyces* were sequenced in sample CF16. *Eutypella* and *Wangiella* were present in CF1. The species for each genera are listed in Supplemental Table 3.8.

In regards to abundance, the most sequenced genus in 17 samples was *Pichia* with sequences from *P. anomala*, *P. membranaefaciens*, and *Pichia* sp. CF7 was the exception with a 24.8 % abundance for the genus *Myrothecium*. The second most abundant genera in all 17 samples was *Rhizoctonia* (Table 3.5).

**3.4.7 Relationship between gender and ethnicity and the mycobiome.** A Bray-Curtis similarity matrix was constructed from the fungal abundance table. The results were plotted using PCO to identify relationships between the communities and the gender and/or ethnicity of the subjects (Figure 3.5). No clustering of community based on gender and/or ethnicity was seen.



**Figure 3.5:** Relationship between the host and the fungal community. Principal coordinate analysis was used to represent the relationship between the fungal communities as represented by the Bray-Curtis similarity matrix. Lilac squares are samples from females. Males are represented by green squares. Numbers in black indicate samples from non-Hispanic subjects while red text represent Hispanic samples.

**3.4.8 Healthy oral mycobiome.** Fungal DNA was sequenced from 20 oral rinse samples from healthy individuals. The data were analyzed separately from CF sequences and have been published (Ghannoum *et al.*, 2010). With author's permission, the summary of results and tables showing the genera and species ( $\geq 1\%$  of the community) identified in the oral rinse samples have been included (Table 3.6 and Supplemental Table 3.9). All samples contained organisms identified as non-culturable; this group was included in all diversity accounts. Between three (H14) and 16 (H2) genera were

sequenced from the healthy dataset with the average diversity of  $10.85 \pm 3.07$ . Fifteen genera were present in two samples (H10 and H16). Thirteen genera were sequenced from H1, H8, and H19. Twelve genera were identified in three samples (H6, H7, and H18). Eleven genera were sequenced H13 and H17. Three samples had ten genera (H5, H9, and H11) while H3 was the only samples with nine fungal taxa. Eight genera were present four samples (H4, H12, H15, and H18) (Ghannoum *et al.*, 2010).

Excluding the non-culturable group, the most commonly sequenced genus was *Candida* (15/20) followed by *Cladosporium* (13/20), *Auereobasidium* (10/20) and the family *Saccharomycetales* (10/20) (Table 3.6). Thirty-nine fungal taxa were not present in more than one sample (Ghannoum *et al.*, 2010). The unculturable group was the most abundant in 12 samples. *Cladosporim* was the most abundant sequence in H11, H13, and H15. *Schizosaccharomyces*, *Aspergillus*, *Saccharomyces*, *Nigrospora*, and *Ophiostoma* were the most abundant taxon in H1, H8, H9, H12, and H20, respectively (Table 3.6). The species present for each genera were listed in Supplemental Table 3.9.

**3.4.9 Comparison of CF and healthy mycobiomes.** To compare the communities, a Sørensen similarity matrix was calculated based on the presence/absence of organisms sequenced in the two communities (Sorensen, 1948). The matrix was used for unweight pair group average clustering. The results, represented on a dendograph, showed no healthy mycobiomes clustered with the CF mycobiomes (Figure 3.6). There were 22 genera that were sequenced in CF sample and not found in the healthy samples (Table 3.7). In addition, there were 63 genera unique to the healthy samples. Eleven genera but only one organism, *C. albicans*, were present in both datasets.

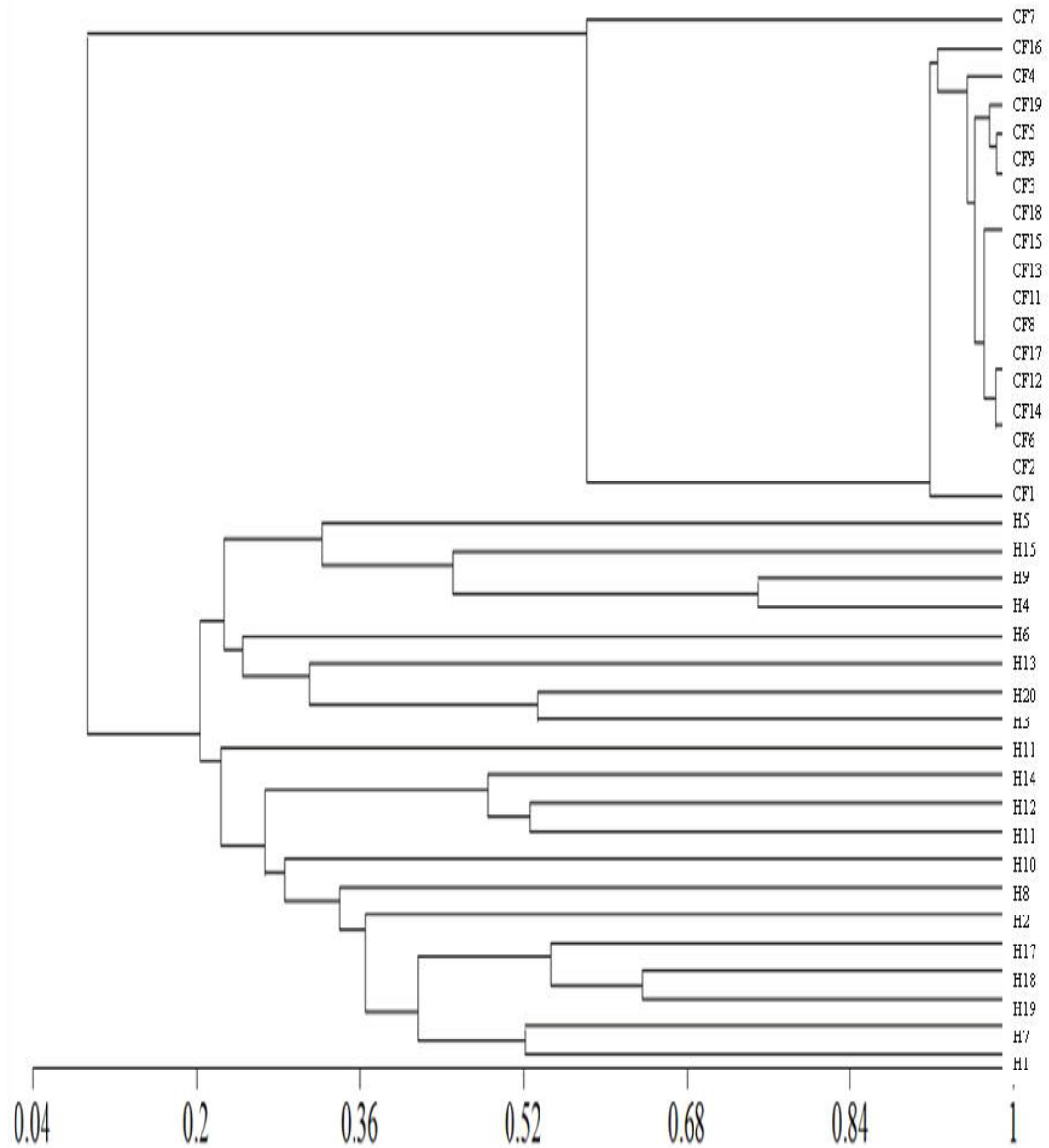
**Table 3.6:** Frequency distribution (%) of fungi present with at least 1 % abundance in the healthy oral microbiome

Organism	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	H16	H17	H18	H19	H20	#
<i>Acremonium</i>		1.17																			1
<i>Alternaria</i>	3.10	1.84		1.03	1.57				2.72							3.32	2.82		1.94		8
<i>Amanita</i>															5.04						1
Ascomycete										3.00		2.26									2
<i>Aspergillus</i>	8.50	3.90					8.79	34.75									12.07	2.77	4.03		7
<i>Atractylodes</i>	6.03						7.23														2
<i>Aureobasidium</i>			6.49		4.70			2.35		1.94		11.29		11.76		3.02	5.96	4.52	6.91		10
<i>Avena</i>															5.04						1
<i>Bagnisiella</i>							7.99														1
<i>Campylobacter</i>										2.47											1
<i>Candida</i>	1.44	5.47	12.05	11.24	0.00	11.81	14.39	1.05	2.27	14.64				14.98		2.97	2.41	1.18	2.58	2.69	15
<i>Chaetomium</i>												1.24									1
<i>Cladosporium</i>	1.55	2.47					7.39	5.41		3.30	56.18	3.49	31.83		42.02	3.95	14.08	15.70	2.42		13
<i>Claviceps</i>										1.20											1
<i>Claviceps</i>								2.46													1
<i>Conopsea</i>								3.40													1
<i>Corynespora</i>		2.42																			1
<i>Cryptococcus</i>				1.06	3.79				2.45							5.55					4
<i>Cystofilobasidium</i>																3.66					1
<i>Davidiella</i>											1.06										1
Dothideomycete		10.72							1.45	4.11											3
Dothioraceae			3.01		1.73					1.94		5.34		4.20		1.61	2.77	1.29	2.34		9
<i>Emericella</i>					1.73											1.01					2
<i>Epicoccum</i>															1.26						1
<i>Eurotium</i>	10.97	1.30					4.74			1.44								17.80			5
Filobasidiales			1.85																		1
<i>Flammulina</i>											8.83										1
<i>Fusarium</i>		4.08					2.18		2.00	2.20			3.70			5.26					6
<i>Gibberella</i>				1.66		1.65			3.91							17.90					4
<i>Gigaspora</i>					1.15																1
<i>Glomus</i>	5.40				13.14		2.02		1.82							7.15					5
<i>Guignardia</i>															5.04						1
<i>Hormonema</i>													1.03					1.18			2
Hypocreales		1.03																	5.16		2
<i>Issatchenkia</i>						1.65															1
<i>Lasiodiplodia</i>	2.87																				1
<i>Lecania</i>																				4.45	1
<i>Lewia</i>				1.10					2.36												2
<i>Macrophomina</i>		3.54																			1
<i>Nectria</i>														1.03							1
<i>Nigrospora</i>										1.24		48.37									2
<i>Ochroconis</i>	3.45																7.65				2
<i>Ophiostoma</i>																				41.33	1
<i>Orpinomyces</i>							2.09														1
<i>Paecilomyces</i>		1.08																	1.61		2
<i>Penicillium</i>	2.81		12.82			1.92															3
<i>Peziza</i>										3.71											1
<i>Phaeosphaeria</i>						1.51															1
<i>Phoma</i>					2.55					1.72											2
<i>Pichia</i>						1.10													1.13		2
Pleosporaceae	8.39									1.15											2
<i>Pongo</i>																	4.03				1
<i>Pueraria</i>										1.44											1
<i>Pyrenophora</i>										2.11											1
<i>Rhizopogon</i>			6.10																		1
<i>Saccharomyces</i>				10.02					49.05			2.47						1.18	3.39	29.39	6
Saccharomycetales			5.64	13.65	1.11	12.36	5.99			8.04			1.85	18.09			1.41		7.26		10
<i>Schizophyllum</i>			14.75													2.57				2.22	3
<i>Schizosaccharomyces</i>	20.10					1.10															2
<i>Sebacina</i>		1.84																			1
<i>Sordariomycete</i>		1.70														1.32					2
<i>Stachybotrys</i>						6.18		1.34		1.20											3
<i>Stemphylium</i>																			6.45		1
<i>Subulispora</i>												2.47									1
<i>Teratosphaeria</i>							1.24	1.87		17.14	1.12	17.66		16.81	1.09	10.87	5.37				9
<i>Torulaspora</i>													2.05								1
Tremellales								1.87													1
<i>Trichaptum</i>						2.34															1
Trichocomaceae						2.61	2.80												2.27		3
<i>Trichosporon</i>										1.10						4.12					2
<i>Verticillium</i>																1.37					1
<i>Wallemia</i>							7.31														1
<i>Xylariales</i>		7.71										31.27									2
<i>Zygosaccharomyces</i>								3.96				0.00									1
Non-culturable	16.37	34.44	34.98	52.08	57.11	53.02	29.47	25.61	20.80	44.69	1.94	8.32	19.92	62.26	6.72	31.85	39.03	34.34	53.39	6.09	20
Count*	13	16	9	8	10	12	12	13	10	15	10	8	11	3	8	15	11	12	13	8	

# is the total samples that contained that specific taxon  
 \* is the total number of taxa present per sample

**Table 3.7:** Fungi sequenced in CF sputa and/or mouth rinse from healthy individuals

<b>Only in CF</b>	<b>Only in Healthy</b>		<b>Present in Both</b>
<i>Blastocladiella</i>	<i>Acremonium</i>	<i>Nectria</i>	<i>Amanita</i>
<i>Bulleromyces</i>	<i>Alternaria</i>	<i>Nigrospora</i>	<i>Aspergillus</i>
<i>Capronia</i>	<i>Ascomycete</i>	<i>Ochroconis</i>	<i>Candida</i>
<i>Cladophialophora</i>	<i>Atractylodes</i>	<i>Ophiostoma</i>	<i>Cladosporium</i>
<i>Debaryomyces</i>	<i>Aureobasidium</i>	<i>Orpinomyces</i>	<i>Cryptococcus</i>
<i>Dipodascus</i>	<i>Avena</i>	<i>Paecilomyces</i>	<i>Glomus</i>
<i>Eutypella</i>	<i>Bagnisiella</i>	<i>Peziza</i>	<i>Hypocreales</i>
<i>Exophiala</i>	<i>Camplobacter</i>	<i>Phaeosphaeria</i>	<i>Penicillium</i>
<i>Fonsecaea</i>	<i>Chaetomium</i>	<i>Phoma</i>	<i>Pichia</i>
<i>Geotrichum</i>	<i>Claviceps</i>	<i>Pleosporaceae</i>	<i>Saccharomyces</i>
<i>Hortaea</i>	<i>Clavispora</i>	<i>Pongo</i>	<i>Schizosaccharomyces</i>
<i>Metschnikowia</i>	<i>Conoplea</i>	<i>Pueraria</i>	
<i>Mucor</i>	<i>Corynespora</i>	<i>Pyrenophora</i>	
<i>Myrothecium</i>	<i>Cystofilobasidium</i>	<i>Rhizopogon</i>	
<i>Olpidium</i>	<i>Davidiella</i>	<i>Saccharomycetales</i>	
<i>Phialophora</i>	<i>Dothideomycete</i>	<i>Schizophyllum</i>	
<i>Rhinocladiella</i>	<i>Dothioraceae</i>	<i>Sebacina</i>	
<i>Rhizoctonia</i>	<i>Emericella</i>	<i>Sordariomycete</i>	
<i>Rhodosporidium</i>	<i>Epicoccum</i>	<i>Stachybotrys</i>	
<i>Rhodotorula</i>	<i>Eurotium</i>	<i>Stemphylium</i>	
<i>Tricholoma</i>	<i>Filobasidiales</i>	<i>Subulispora</i>	
<i>Wangiella</i>	<i>Flammulina</i>	<i>Teratosphaeria</i>	
	<i>Fusarium</i>	<i>Torulaspora</i>	
	<i>Gibberella</i>	<i>Tremellales</i>	
	<i>Gigaspora</i>	<i>Trichaptum</i>	
	<i>Guignardia</i>	<i>Trichocomaceae</i>	
	<i>Hormonema</i>	<i>Trichosporon</i>	
	<i>Issatchenkia</i>	<i>Verticillium</i>	
	<i>Lasiodiplodia</i>	<i>Wallemia</i>	
	<i>Lecania</i>	<i>Xylariales</i>	
	<i>Lewia</i>	<i>Zygosaccharomyces</i>	
	<i>Macrophomina</i>		



**Figure 3.6:** Relationship between CF sputa and healthy oral rinse mycobiomes. A Sørensen similarity matrix based on presence/absence of fungi was used with unweighted pair group average clustering to determine the similarity between CF and healthy oral fungal communities. The length of the branch connecting two samples indicates the dissimilarity between the samples.

### 3.5 Discussion

The leading cause of morbidity and mortality in CF is lung disease is a result of chronic infections. Through the use of molecular techniques, the CF lung was shown to contain numerous bacteria. Sanger sequencing projects have identified both known and novel pathogens revealing a complex bacterial community (Bittar *et al.*, 2008; Harris *et al.*, 2007). Though fungal infections occur in the CF lung, the composition of the community has not been determined. This study further characterizes the eubacteriome and mycobiome by identifying organisms using MTPS which has been shown to provide a more accurate estimation of diversity. In addition, the CF microbiomes from 18 sputa samples were compared to the communities in healthy oral samples.

**3.5.1 MTPS shows greater diversity in the CF bacterial community as compared to clonal sequencing.** The sputa samples were sequenced for bacteria using the first two hypervariable regions of the 16S rRNA gene. After the data was processed, every sample was revealed to contain anywhere from four to 16 genera with an average of  $9.6 \pm 3.3$  genera. The diversity per patient was greater than the previous cloning/sequencing study by Bittar *et al.* (one to 14 with an average  $7.2 \pm 3.9$  species) (Bittar *et al.*, 2008). The sequences in this study were identified to the genus level or higher. It is likely that multiple species were present for some genera which would further increase the diversity of organisms in the lung. As with other 454 sequencing studies, the estimation of diversity within the CF sputa community was higher than other sequencing and community profiling methods have shown (Armougom *et al.*, 2009; Quince *et al.*, 2009; Sogin *et al.*, 2006).



The number of genera associated with CF dataset was similar to the earlier clonal sequencing studies. In total, sequences were identified from 34 genera. In comparison, the Sanger sequencing study using sputa by Bittar *et al.* identified 53 species from 32 genera (Bittar *et al.*, 2008). The other study by Harris *et al.* sequenced DNA belonging to 65 species from 35 defined genera and one unidentified organism (Harris *et al.*, 2007). All three studies highlighted the bacterial diversity within the CF lung.

### **3.5.2 Common and emerging pathogenic bacteria are present in CF lungs.**

CF samples are routinely culture positive for *Burkholderia*, *Haemophilus*, *Pseudomonas*, and *Staphylococcus*. Therefore, it was not surprising to identify these pathogens from the samples (Figure 3.1 and Table 3.2). Of these organisms, *Pseudomonas* was present in 16 of the 18 samples. This was expected as *P. aeruginosa* is quite common in adult patients (Patient-registry, 2008). *Haemophilus* and *Staphylococcus* were present in some samples. This agrees with epidemiology reports which show that less than 10 % and less than 40 % of patients at the age of 35 have culturable *Haemophilus* and *S. aureus* infections, respectively (Patient-registry, 2008). Though *Burkholderia cenocepacia* is associated with CF, it was surprising to identify *Burkholderia* sequences in one sample at a 23.4 % abundance. This individual has never been culture positive for any members of the *Burkholderia cenocepacia* complex. *Burkholderia* infections are extremely difficult to treat and are considered a poor indication of the prognosis for the patient (Jones *et al.*, 2004). The identity of the species should be determined using other techniques to ensure the patient is undergoing proper treatment.

There are other pathogens that are cultured less frequently from CF sputa (De Baets *et al.*, 2007; Demko *et al.*, 1998; Goss *et al.*, 2004; Lambiase *et al.*, 2009; Mentasti *et al.*, 2008; Ojeda-Vargas *et al.*, 1990; Reina *et al.*, 1992; Spicuzza *et al.*, 2009). The less prevalent pathogens *Achromobacter*, *Stenotrophomonas*, *Proteus*, and *Sphingobacterium* were also sequenced from the samples. Of these genera, *Proteus* was the only organism to dominate a community (CF4). *Pseudomonas* was not present in the sample which may suggest they share similar resources and *Proteus* outcompeted *Pseudomonas* for the niche (Figure 3.1 and Table 3.2).

**3.5.3 Oral microbiota may play a role in disease progression.** Some sequences obtained belong to genera that are known to be found in the oral cavity. In this dataset, *Actinomycineae*, *Capnocytophaga*, *Lactobacillus*, *Leptotrichia*, *Fusobacterium*, *Neisseria*, *Porphyromonas*, *Prevotella*, *Prevotella dentalis* (formally *Hallela*), *Streptococcus*, *Shuttleworthia*, and *Veillonella* were all identified (Figure 3.1 and Table 3.2). These organisms have been routinely cultured and sequenced in oral samples (Downes *et al.*, 2002; Jousimies-Somer & Summanen, 2002; Lazarevic *et al.*, 2009; Williams & Collins, 1995). The oropharyngeal bacteria have also been sequenced in other CF studies (Bittar *et al.*, 2008; Rogers *et al.*, 2006).

It is difficult to define the role of the organisms in the CF lung without identifying the species. Each sequenced genus contains species that are part of the normal oral microbiota and organisms which cause infections in other areas of the body. For example *Neisseria* contains two pathogenic species which cause gonorrhea and meningitis (Liljemark & Gibbons, 1971). It is doubtful the organism is associated with lung disease

and the DNA from the oral strains were most likely present in the sputa. *Leptotrichia* is another genus that contains both nonpathogenic species such as *L. buccalis* which is part of the healthy microbiota yet other species have been associated with bacteremia, neutropenia, and endocarditis (Bhally *et al.*, 2005; Eribe & Olsen, 2008). Some oral genera such as, *Capnocytophaga*, *Porphyromonas*, and *Lactobacillus* have species that are known to be pathogenic in the human oral cavity but an association with the lung has not been established (Bunyaratavej, 2003; Cannon *et al.*, 2005; Geisler *et al.*, 2001).

Studies are beginning to make connections between oral flora and lung health. Certain oral bacteria are present more often in those suffering from asthma and chronic obstruction pulmonary disease (Arbes *et al.*, 2006; Deo *et al.*, 2009). It has been proposed that pneumonia or ventilator-associated pneumonia is partly attributed to oral bacteria and proper oral hygiene in nonambulatory patients will reduce the risk of infection (Bahrani-Mougeot *et al.*, 2007; Paju & Scannapieco, 2007; Scannapieco, 2006). In addition, *Capnocytophaga gingivalis*, an oral microbe, was isolated from bronchoalveolar lavage (BAL) fluid of a patient with pneumonia (Geisler *et al.*, 2001). Thus, it is possible for oral bacteria to reside in the lungs.

Recently, oral microbes belonging to *Streptococcus*, *Actinomyces*, *Veillonella*, and *Prevotella* have been shown to colonize the oral cavity and the lungs (Hilty *et al.*, 2010; Lazarevic *et al.*, 2009; Mabeza & Macfarlane, 2003; Shah *et al.*, 2008; Sibley *et al.*, 2008). *Prevotella* and *Veillonella* have recently been isolated from the lower lungs of healthy individuals disproving the opinion that the lower lung is sterile (Hilty *et al.*, 2010). Oral microbiota has also been attributed to lung infections. *Veillonella* has been

associated with chronic anaerobic pneumonitis in a non-CF individual (Shah *et al.*, 2008). *Actinomyces* also causes pulmonary actinomycosis (Mabeza & Macfarlane, 2003).

It could be hypothesized that certain species within a genus are a part of the normal lung microbiota while others are pathogens. In addition, it is also likely an organism can become pathogenic if certain conditions are present. This has recently shown true for *Streptococcus milleri*. The oral microbe has been cultured from the CF lung and shown to be pathogenic when present in high amounts (Sibley *et al.*, 2008). It is also possible the organism contributes to the disease through interactions with known pathogens. For example, a strain of oral *Streptococcus* can induce a positive regulation of virulence factors in *P. aeruginosa* (Duan *et al.*, 2003). Virulence factors are known to help *P. aeruginosa* establish an acute infection in the CF lung (Goodman *et al.*, 2004). Understanding the role of oropharyngeal bacteria in lung progression may lead to novel treatment methods.

**3.5.4 Rarely cultured bacteria are readily identified using molecular methods.** There were a number of sequence reads corresponding to organisms that are rarely cultured but identified in molecular studies. In this study, two samples contained *Campylobacter* sequences at low abundances (1.03 and 1.22 %) (Figure 3.1 and Table 3.2). Though this genus is not associated with lung infections, it has been previously identified in a CF sample based on a band in a T-RFLP profile (Rogers *et al.*, 2003). However, the organism was never sequenced from the sample (Rogers *et al.*, 2003). It has been sequenced from healthy oral samples (Keijser *et al.*, 2008).

*Eikenella*, identified in this study in one sample, is a commensal bacterium in the mouth and upper respiratory system (Figure 3.1 and Table 3.2). It can cause lung infections; although, to date it has not been associated with CF patients (Hoyler & Antony, 2000).

*Gemella*, another anaerobic bacterium, was present in eight samples (Figure 3.1 and Table 3.2). Certain species are opportunistic pathogens and have been associated with endocarditis and isolated from lung abscesses (da Costa *et al.*, 1996; La Scola & Raoult, 1998; Valipour *et al.*, 2005). This organism has been sequenced from CF samples and cultured from BAL samples (Bittar *et al.*, 2008; Harris *et al.*, 2007; Tunney *et al.*, 2008). High concentration of *Gemella* in the CF lung can lead to exacerbations (Michael Surette; personal communication)

*Eikenella* and *Gemella* may be more prevalent in oral and lung communities than previously thought. Routine culturing methods for CF samples do not usually include the anaerobic growth conditions that are required to culture these organisms. A recent study has shown that with proper culturing techniques numerous anaerobic bacteria in high numbers can be isolated from the CF lung (Tunney *et al.*, 2008). The significance of many anaerobes in the lung is unknown. However, as that *Gemella* can lead to illness, other anaerobes may also contribute to lung disease.

**3.5.5 Unusual organisms present in CF sputa.** The most abundant genus in 14 samples was *Chryseomonas* which is closely related to *Pseudomonas* (Figure 3.1 and Table 3.2). Some species such as *C. luteola* have been reclassified as belonging to the *Pseudomonas* genus due to the high degree of similarity between their 16S rRNA gene

sequence (Anzai *et al.*, 1997). In addition, Ribosomal Database Project II (RDP; Release 10) no longer includes the genus *Chryseomonas* and species were placed under *Pseudomonas*. There is 100 % homology of the full length 16S rRNA gene between *P. aeruginosa*, *P. flavescens*, *P. mendocina* and *Chryseomonas/Pseudomonas luteola*. The MTPS reads were analyzed with RDP II (Release 8) which may have had *Chryseomonas* separated from *Pseudomonas*. Reanalysis with the newer release version is necessary to determine if the sequences belong to *Pseudomonas* or a closely related group. If the sequences are indeed *P. luteola* then it may be the first report of this bacterium in the CF lung. The organism is known to cause infections in the eye, brain, and around implanted devices such as prosthetic valves leading to bacteremia (Casalta *et al.*, 2005; Gaschet *et al.*, 2009; Uy *et al.*, 2007). Like *P. aeruginosa*, *P. luteola* expresses a beta-lactamase making it more resistant to antibiotics (Doublet *et al.*, 2010). Its ability to colonize humans and its resistant to antibiotics increases the need to confirm the presence of *P. luteola*.

*Flavimonas* sequences were present in eight samples with a high abundance of 23.1 % in CF2 (Figure 3.1 and Table 3.2). However, the other seven samples were between one and two percent abundance. This genus was reclassified as *Pseudomonas* in RDP II (Release 10) (Anzai *et al.*, 1997). Formerly *Flavimonas orzihabitans*, *P. orzihabitans* is a rare pathogen but has been implicated in catheter associated infections (Marin *et al.*, 2000). There is a high homology between the 16S rRNA sequence from strains of *P. orzihabitans* and various strains of *P. aeruginosa*, yet the V1\_V2 region of the gene is not the same between all species which indicates a *Flavimonas* species was present. Species-specific PCR could be used to confirm the presence of *P. orzihabitans*.

Some sequences were present in very low abundance. *Oribacterium* was present in two samples with less than 2 % abundance (Figure 3.1 and Table 3.2). This genus, discovered in 2004 from sinus pus obtained from a young boy, has not been cultured from any other source (Carlier *et al.*, 2004). However, it was sequenced from healthy oral samples (Keijser *et al.*, 2008)

*Moryella*, present in one sample (2.44 %), contains one species *indoligenes* which has only been isolated from three human abscesses (Figure 3.1 and Table 3.2) (Carlier *et al.*, 2007). This organism has not been well characterized and its relationship to human lungs and oral cavity is unknown.

*Megasphaera* sequences were present in two samples in low abundances (Figure 3.1 and Table 3.2). The bacterium is normally associated with the rumens of animals. Phylogentic analysis of 16S rRNA sequences from uncultured bacteria have shown *Megasphaera*-like organisms present in the vaginal and oral cavity (subgingival plaque and endodontic infections) samples (Marchandin *et al.*, 2003; Spear *et al.*, 2008; Zozaya-Hinchliffe *et al.*, 2008). In this study, the bacteria are most likely associated with the oral cavity.

Two other taxa, *Schlegellella* and *Shigella*, were identified. *Schlegellella* spp. were present in one sample at 7.37 %. This genus contains aquatic bacteria some of which are associated with hot springs (Chou *et al.*, 2006; Elbanna *et al.*, 2003). Two samples also had low abundances for *Shigella* which causes dysentery (Yang *et al.*, 2005).

*Oribacterium*, *Moryella*, *Megasphaera*, *Schlegellella* and *Shigella* sequences were present in few samples and at low abundances. Their role in the CF lung may be quite

minor. It is possible that *Oribacterium* and *Megasphaera* are from the oral cavity as they have been sequenced from oral samples or are potentially colonizing the upper respiratory system. Sequences from *Moryella* may be a rare member of the eubacteriome and attempts should be made to isolate the organisms from sputa. *Schlegellella* and *Shigella* sequences are perplexing to find in the CF sputa and could possibly be attributed to sequencing error. *Chryseomonas* and *Flavimonas* are potentially novel bacteria in the CF lung. These organisms were present in most samples at large abundances indicating that the organisms were metabolically active. Due to the high 16S rRNA homology, it should be confirmed that these organisms are not in fact *P. aeruginosa*.

**3.5.6 Identifications above genus level are difficult to interpret.** Not all sequences obtained could be identified to the genus level. Five taxa, Actinomycineae, Carnobacteriaceae, Coriobacterineae, Micrococcineae, and Propionibacterineae were identified at high taxonomic levels that contain many environmentally divergent organisms. It is difficult to conjecture the significance of the taxon in the CF samples.

Actinomycineae sequences may potentially be from common oral bacteria. Carnobacteriaceae sequences may indicate the presence of *Carnobacterium*, *Dolosigranulum*, and *Granulicatella*, organisms sequenced in other CF samples (Bittar *et al.*, 2008; Harris *et al.*, 2007; Rogers *et al.*, 2009). The suborder Coriobacterineae contains many organisms that have been isolated from the human oral cavity (Lazarevic *et al.*, 2009). Micrococcineae is a suborder of Gram-positive organisms mostly found in soil (Hennessee & Li, 2010). Propionibacterineae is a suborder that contains a diverse group of bacteria living in dairy products to intestinal tracts of animals (Goldschmidt *et*



*al.*, 2009) (Figure 3.1 and Table 3.2). Sequencing the samples again with a different region of the 16S rRNA gene may provide better resolution of the community for these genera.

**3.5.7 Gender and ethnicity of the host may not affect the bacterial community.** Previous studies and this work has shown that the bacterial communities are not the same in CF individuals but rather unique (Bittar *et al.*, 2008; Harris *et al.*, 2007; Rogers *et al.*, 2003; Rogers *et al.*, 2009). It is not known why some CF patients have diverse communities while others may be dominated by one organism, such as *P. aeruginosa*. This study attempted to determine if there was a correlation between the gender and ethnicity of the host and the bacterial community. Using PCO and the Bray-Curtis similarity matrix, it was determined that there was no correlation between the bacterial community (as a whole) and the host's gender and ethnicity (separately or combined) (Figure 3.2). Though this study did not find a link between these factors and the community, it does not necessarily indicate that gender and ethnicity do not play a role in the community. A larger sample size is needed to increase the power of statistics. In addition, other races such as African-American should be added to the analysis.

**3.5.8 Healthy and CF bacterial communities are statistically different despite shared bacteria.** The Bray- Curtis similarity matrix calculated from the abundance of the identified taxa in each CF sputum sample and the healthy oral rinse were used to determine the similarity between the CF and healthy samples. The PCO analysis of the CF sputa communities and the healthy oral rinse communities clearly indicates that the

overall eubacteriomes between the two datasets were different (Figure 3.3). The CF samples were not as tightly clustered indicating the dissimilarity between CF samples. The healthy samples were clustered together and away from the CF samples. The separation of the two datasets show that the overall bacterial community in CF sputa is different from the community present in healthy oral rinses, yet there are organisms common to both groups. Within the communities, 27 taxa were only present in healthy samples, 12 groups were associated with CF, and 22 taxa were present in both groups (Table 3.4).

**3.5.8.1 CF pathogens mostly absent in healthy eubacteriomes.** Twelve taxa were unique to the CF dataset. It was not surprising to find *Proteus*, *Pseudomonas*, *Staphylococcus*, and *Stenotrophomonas* sequences in the CF dataset (Table 3.4). Each genera contains pathogenic species that have been cultured from the lung (Patient-registry, 2008; Whittier, 2001). *Sphingobacterium*, an uncommon CF pathogen, was also unique to the CF samples (Reina *et al.*, 1992). *Chryseomonas* and *Flavimonas*, both of which have been reclassified as *Pseudomonas*, were not found in the healthy samples (Anzai *et al.*, 1997). This further supports the hypothesis that these two taxa are potentially pathogens.

*Shigella* which is known to cause dysentery was present in two CF samples (Yang *et al.*, 2005). It is likely the bacteria were transient members of the oral community or identified incorrectly due to sequencing error. *Lactobacillus*, *Shuttleworthia*, and *Eikenella* were not found in any healthy samples which is surprising as they are known to be associated with the oral cavity (van houte *et al.*, 1972). These groups may have not

competed well with the rest of the communities in the healthy samples. *Moryella* was not present in the healthy oral rinses. This anaerobe may potentially be a rare bacterium in the lung.

**3.5.8.2 CF and healthy samples share bacteria.** Twenty-two taxa were present in both datasets (Table 3.4). *Achromobacter*, *Burkholderia*, and *Haemophilus* are pathogens routinely cultured from CF patients. *Achromobacter* sequences were present in one healthy individual (H22, 21.0 %) (Table 3.3). This organism can cause infections in non-CF patients and it is possible the healthy individual was actually suffering from an infection. *Burkholderia* in a healthy individual (H22, 1.18 %) was unexpected. The *Haemophilus* sequences, present in all samples but H3, were most likely from species that live in the oral cavity (Keijser *et al.*, 2008; Kraut *et al.*, 1972; Sixou *et al.*, 1991). It is possible the six CF samples with *Haemophilus* DNA were not from *H. influenzae* but rather from the oral species.

The largest subset of shared taxa which included *Actinomycineae*, *Campylobacter*, *Capnocytophaga*, *Gemella*, *Hallella*, *Fusobacterium*, *Lepotrichia*, *Megasphaera*, *Neisseria*, *Porphyomonas*, *Prevotella*, *Streptococcus*, and *Veillonella* are known colonizers of the mouth (Table 3.4). Species identification is needed to establish if the bacteria is unique to the normal oral sample or shared between the datasets.

*Oribacterium* was present in both datasets (Table 3.4). The organism is not well studied but it has been previously cultured from sinus pus and sequenced from healthy saliva (Carlier *et al.*, 2004; Keijser *et al.*, 2008). Due to its low abundance in samples, it is doubtful the organism is a major player in the mouth or lung communities.

*Schlegelella* sequences were present and this group of bacteria is most often associated with aquatic environments (Chou *et al.*, 2006). It is likely the organism is not a resident of the mouth. The taxonomic identification of Carnobacteriaceae 2, Coriobacterineae, Micrococcineae, and Propionibacterineae sequences are too broad to conjecture a hypothesis as to their role in the mouth.

**3.5.8.3 Bacterial diversity decreased in CF samples.** Twenty-seven taxa were identified in the healthy samples and not in the CF eubacteriomes (Table 3.4). The diversity of the healthy eubacteriome was much greater than in the sputa. It is not known if these taxa are a part of the core oral eubacteriome. The significance of the organisms may be revealed when the oral core biome is determined. Perhaps in the future, the composition of the oral community could be used as an indicator of health.

**3.5.9 Fungi are prevalent in CF sputa.** Through the use of culturing and species specific PCR, multiple species of *Aspergillus* and *Candida* are known to culture the CF lung (Bakare *et al.*, 2003). There have also been reports of *Scedosporium apiospermum*, *Exophiala* sp., and *Wangiella dermatitidis* infections (Defontaine *et al.*, 2002; Dixon & Polak-Wyss, 1991; Horr  *et al.*, 2004). Yet, the mycobiome in the CF lung has not been characterized. This is the first study identifying fungi in CF sputa.

From the 18 samples, 33 genera were identified (Figure 3.4 and Table 3.5). Sputum samples contained between 13 and 30 different taxa with an average of  $26 \pm 3$  (Table 3.5). All but one sample (CF7) had more than 24 genera present. These results highlight a complex mycobiome despite clinical results which report no fungal infections

in all patients. The majority of the taxa were present in low abundances which may indicate low levels of growth at the time of sampling. Shifts in the microbiome may promote growth of fungi which can lead to CF exacerbations. Therefore, it is important to identify and understand the role of the mycobiome in the CF lung.

**3.5.10 Pathogenic fungi undetected by clinicians.** Though the average diversity of mycobiome appears as large as the eubacteriome, there were few fungi previously associated with the CF lung present. *Candida albicans* sequences were present in eight samples (1.39 – 5.42 % abundance) (Supplemental Table 3.8). This species is often ruled as a commensal organism in the mouth. However, antibiotic treatment can shift the microbiome and cause rampant growth of *Candida* which leads to candidiasis. If this was the case in the CF patients, the *Candida* sequences would most likely be more abundant.

In addition, *Aspergillus terreus* (18/18), *Aspergillus ustus* (18/18), *Cryptococcus neoformans* (17/18), *Exophiala* (17/18) and *Wangiella dermatitidis* (1/18) sequences were present in the samples (Supplemental Table 3.8). Though the abundances were low (< 2%), their presence is still cause for concern. These organisms are pathogenic and known to colonize CF patients.

*Pichia* (*P. anomala*, *P. membranaefaciens*, and *Pichia spp*) sequences were the most abundant in all samples except for CF7 (15 - 22 %). Many strains of *Pichia* are environmental and clinician would consider it a contaminate if cultured. This reaction is beginning to change as *P. anomala*, *P. membranaefaciens* and *P. ohmeria* have been shown to be causative agents of opportunistic mycoses. Though other pathogenic species

were sequenced, the higher abundance of *Pichia* sequences indicated that this is the dominant colonizer in the mycobiome.

**3.5.11 Environmental fungi abundant in CF sputa.** Of the 33 genera present, the species from 24 genera are most likely environmental contaminants and unassociated with human mycobiomes (Supplemental Table 3.8). Most of the species sequenced from the following genera are not associated with humans or known to be human pathogens: *Amanita*, *Blastocladiella*, *Bulleromyces*, *Capronia*, *Cladosporium*, *Debaryomyces*, *Eutypella*, *Glomus*, *Geotrichum*, *Hortaea*, *Hypocrea*, *Metschnikowia*, *Mucor*, *Myrothecium*, *Olpidium*, *Penicillium*, *Phialophora*, *Rhinocladiella*, *Rhizoctonia*, *Rhodosporidium*, *Rhodotorula*, *Saccharomyces*, *Schizosaccharomyces*, and *Tricholoma*. These fungi are either mushrooms, household mold, or associated with water, soil, and plants.

Environmental fungi can cause chromoblastomycosis, which is a fungal infection of the skin and subcutaneous tissue. The mycobiome contained four organisms, *Cladophialophora arxii*, *Cladophialophora devriesii*, *Phialophora verrucosa*, and *Fonsecae pedrosoi* which can occasionally cause chromoblastomycosis.

In total, 27 genera identified from the sputa are most likely derived from consumed food and air. The sequences were present in low abundances which supports the conjecture that these fungi are transient members of the CF mycobiome.

**3.5.12 Gender and ethnicity may not effect fungal community composition.** This study attempted to determine if there was a correlation between the gender and ethnicity of the host and the fungal community (Figure 3.5). Using PCO, it was

determined that there was no correlation between the fungal community (as a whole) and the host's gender and ethnicity (separately or combined). As discussed with the bacterial community, these results do not necessarily indicate that gender and ethnicity do not play a role in the community. A larger sample size and other races should be added to the analysis.

**3.5.13 Healthy oral cavity contains a diverse mycobiome.** Using MTPS, Ghannoum et al. identified 74 culturable and 11 non-culturable fungal genera representing a total of 101 species for all 20 participants (Table 3.6 and Supplemental Table 3.9) (Ghannoum *et al.*, 2010). Over 60 genera are commonly found in the environment and not associated with human infections. The reoccurrence of the genera in multiple patients led the authors to state that the organisms were most likely in the oral cavity as a result from eating and breathing.

The largest percentage of fungi identified in the mouth rinse samples was identified as non-culturable fungi. Four pathogenic fungi were present in at least 20 % of the subjects. Of these pathogens, *Candida* was present in 75 % of the samples. The prevalence was much higher than other studies and the authors attributed it to better technology. *Aspergillus*, *Fusarium*, and *Cryptococcus* sequences were surprising as that they have never been shown to colonize the mouth. The authors conjectured that the organisms' pathogenic behaviors were controlled by other fungi in the mycobiome. The study determined the mycobiome in the healthy mouth is diverse but the clinical relevance of the fungi are unknown (Ghannoum *et al.*, 2010).

**3.5.14 CF sputa and oral rinses contain unique mycobiomes.** Caution must be taken when comparing the CF and healthy mycobiome data sets. The sequences from the two datasets were not analyzed together as one dataset. The separate analyses may have led to different identifications for the sequences. In addition, the CF analysis did not include uncultured fungi, rather if a sequenced was identified as uncultured the next closest aligned sequence from a cultured organism was used for the analysis. The absence of this category in the CF samples skews any abundance-based statistical analysis. Therefore, the two communities were compared with Sørensen similarity index based on presence/absence of the genera (Sorensen, 1948). The results were shown on a dendograph (Figure 3.6). Healthy mycobiomes were more similar to one another than with the CF communities. This clustering pattern of the samples was attributed to 63 genera only present in healthy samples and 22 genera unique to the CF dataset (Table 3.7). At the species level, *C. albicans*, an opportunistic pathogen, was the only shared organism between the two datasets.

**3.5.14.1 Pathogenic fungi were present in healthy and CF subjects.**

*Aspergillus*, *Candida*, *Cryptococcus*, and *Pichia* sequences were associated with healthy and CF individuals. Of the four species of *Candida*, *C. albicans* sequences were identified in the sputa samples. In addition, *Candida* was much more prevalent in the healthy samples. The oral samples contained six *Aspergillus* species, two *Cryptococcus* species and one *Pichia* species, none of which were present in the CF individuals (Supplemental Table 3.8 and 3.9). It was unexpected to identify pathogenic fungi in the healthy mouth. As mentioned by Ghannoum et al., these fungi may be controlled by



other members of the community preventing them from behaving as pathogens in the healthy individual.

**3.5.14.2 Environmental fungi abundant in both CF sputa and healthy oral samples but no species are in common.** Twenty genera of environmentally associated fungi were present in the CF samples. In addition, most if not all of the fungi unique to the mouth mycobiomes were from the environment. No species were present in both datasets. It is not surprising to find such a small overlap in genera between the two groups. The participants were from south Florida or Cleveland, Ohio; two entirely different geographically locations. The individuals were exposed to different fungi and that is reflected in their mycobiomes. Moreover, subjects from the same environment do not share all the same fungi. These differences could be attributed to where a person works or what they eat.

**3.5.15 Conclusion.** This study addressed the eubacteriome and mycobiome of the CF lung. Through 454 sequencing, the polymicrobial nature of the CF lung was explored. Pathogenic bacteria and microbes associated with the oral cavity were abundant. Previous studies have shown that sputum is not contaminated by oral bacteria (Rogers *et al.*, 2006). Therefore, oral bacteria may play a role in the progression of lung disease. Although the lung and mouth are different environments, the comparison is still noteworthy, demonstrating the overlap of oral bacteria present in the mouth and the sputa. The results illustrate the need to understand the role of oral bacteria in the airways.

South Florida patients also have an extensive fungal community in their sputa. Although most of the fungi present were attributed to environmental sources, pathogenic species were present in patients negative for fungal infections.

This sequencing study clearly demonstrates the complexity of the CF microbiome. Future work should examine the significance of the members in the community.

### **3.6 Acknowledgements**

I would like to thank MA Ghannoum from Case Western Reserve University for access to the healthy microbiome data and PM Gillevet for his analysis of the healthy data. Lisa Schneper was instrumental in this project aiding in statistical analyses and general discussions regarding the results. CF sputa were obtained with the help of Michael Light and the University of Miami CF clinic. Research was funded by the laboratory of Kalai Mathee under the FIU College of Medicine. MSD was funded by the Cystic Fibrosis Foundation Training Grant (DoudOH08) and MBRS RISE (NIH\NIGMS R25 GM061347).

**Supplemental Table 3.8:** Frequency distribution (%) of fungal species present with at least 1 % abundance in the CF sputa mycobiome

Samples	CF1	CF2	CF3	CF4	CF5	CF6	CF7	CF8	CF9	CF11	CF12	CF13	CF14	CF15	CF16	CF17	CF18	CF19	#
<i>Amanita phalloides</i>															1.1				1
<i>Aspergillus terreus</i>	1.0	1.4	1.3	1.5	1.3	1.4		1.1	1.4	1.3	1.1	1.3	1.5	1.2	1.1	1.3	1.1		16
<i>Aspergillus ustus</i>	1.1	1.5	1.3	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Blastocladiella emersonii</i>	1.0	1.4	1.3	1.6	1.3	1.4		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Bulleromyces albus</i>	1.1	1.5	1.3	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Candida albicans</i>		5.4	5.0		3.3	4.8			4.5		1.8		5.2			1.4			8
<i>Candida dubliniensis</i>																			0
<i>Capronia acutiseta</i>	1.1	1.5	1.4	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Capronia coronata</i>	1.4																		1
<i>Capronia epimyces</i>	1.1	1.5	1.4	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Capronia pilosella</i>	1.1	1.5	1.4	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Capronia pulcherrima</i>	1.1	1.5	1.4	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Cladophialophora arxii</i>	1.1	1.5	1.4	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Cladophialophora devriesii</i>	1.1	1.5	1.4	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Cladophialophora sp.</i>	1.4																		1
<i>Cladosporium cladosporioides</i>	1.3																		1
<i>Cladosporium macrocarpum</i>	1.3																		1
<i>Cladosporium sp.</i>	2.8								1.0										2
<i>Cladosporium tenuissimum</i>	1.4																		1
<i>Cryptococcus neoformans</i>	1.1	1.5	1.4	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Debaryomyces hansenii</i>	1.1	1.5	1.4	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Dipodascus albidus</i>	1.1	1.5	1.4	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Eutypella vitis</i>	1.0																		1
<i>Exophiala sp.</i>	1.1	1.5	1.4	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Fonsecaea pedrosoi</i>	1.1	1.5	1.4	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Geotrichum candidum</i>	1.1	1.5	1.4	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Glomus mosseae</i>	1.1	1.5	1.4	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Hortaea werneckii</i>	2.8														1.3				2
<i>Hypocrea koningii</i>	1.1						2.2	1.1		1.1				1.2	1.0				6
<i>Hypocrea muroiana</i>																			0
<i>Hypocrea rufa</i>	1.1						2.2	1.1		1.1				1.2	1.0				6
<i>Metschnikowia bicuspidata</i>	1.0	1.4	1.3	1.5	1.3	1.4		1.1	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Mucor racemosus</i>	1.0	1.4	1.3	1.6	1.3	1.4		1.1	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Myrothecium atroviride</i>	1.1						2.2	1.1		1.1	0.8			1.2	1.0		1.1		8
<i>Myrothecium cinctum</i>	2.3				1.5		4.6	2.3		2.3	1.7	2.0	0.3	2.3	2.0	1.8	2.1		12
<i>Myrothecium inundatum</i>	1.1						2.2	1.1		1.1				1.1	1.0		1.1		7
<i>Myrothecium leucotrichum</i>	1.1						2.3	1.2		1.1				1.2	1.0		1.1		7
<i>Myrothecium roridum</i>	4.3		1.7		2.8		9.0	4.5		4.4	3.2	3.9		4.6	4.1	3.6	3.9	1.1	13
<i>Myrothecium verrucaria</i>	2.2				1.4		4.5	2.3		2.2	1.6	2.0		2.3	2.1	1.8	2.2		11
<i>Olpidium brassicae</i>	1.1	1.5	1.4	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Paracentrotus lividus</i>	1.1	1.5	1.3	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Penicillium argillaceum</i>	1.1	1.5	1.4	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Penicillium freii</i>	1.1	1.5	1.3	1.6	1.3	1.5		1.3	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.3	1.8	17
<i>Phialophora sp.</i>	3.2	1.7	2.2	1.7	2.7	1.9	5.4	3.5	1.9	3.5	2.9	3.3	1.8	3.5	3.2	3.1	3.2	1.9	18
<i>Phialophora verrucosa</i>	1.1	1.5	1.4	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Pichia anomala</i>	1.1	1.5	1.4	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Pichia membranaefaciens</i>	6.2	8.6	7.9	9.3	7.7	8.6	5.5	6.8	8.2	7.7	7.9	7.7	8.7	7.5	6.4	7.6	7.0	7.8	18
<i>Pichia sp.</i>	8.3	11.4	10.5	12.3	10.3	11.4	7.3	9.1	10.9	10.3	10.5	10.2	11.7	10.0	8.5	10.2	9.4	10.3	18
<i>Rhinocladiella atrovirens</i>	1.1	1.5	1.4	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Rhinocladiella phaeophora</i>	1.1	1.5	1.4	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Rhizoctonia sp.</i>	14.8	19.5	17.8	21.0	17.6	19.3	13.2	16.3	18.4	18.5	18.1	17.9	19.5	18.0	14.7	18.1	16.5	17.3	18
<i>Rhodospiridium toruloides</i>	1.1	1.5	1.4	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Rhodotorula glutinis</i>	1.1	1.5	1.4	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Saccharomyces cerevisiae</i>															8.2				1
<i>Schizosaccharomyces fimicola</i>	1.1						2.2	1.1		1.1				1.1	1.0		1.1	0.3	8
<i>Schizosaccharomyces pombe</i>	2.2	3.0	2.7	3.2	2.6	2.9	1.9	2.5	2.8	2.6	2.7	2.6	3.0	2.6	2.3	2.6	2.5	2.7	18
<i>Schizosaccharomyces sclerotiorum</i>	1.1	1.5	1.4	1.6	1.3	1.5		1.2	1.4	1.3	1.3	1.3	1.5	1.3	1.1	1.3	1.2	1.3	17
<i>Tricholoma matsutake</i>	2.2	3.0	2.7	3.2	2.6	2.9	1.9	2.5	2.8	2.6	2.7	2.6	3.0	2.6	2.3	2.6	2.5	2.7	18
<i>Wangiella dermatitidis</i>	1.38																		1
Count*	54	37	38	36	40	37	15	45	38	45	41	39	38	45	48	40	43	37	

# is the total samples that contained that specific taxon

\* is the total number of taxa present per sample

**Supplemental Table 3.9:** Frequency distribution (%) of fungal species present with at least 1 % abundance in the healthy oral rinses mycobiome

Organism	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	H16	H17	H18	H19	H20	#
<i>Alternaria</i> sp.	2.0																				1
<i>Alternaria tenuissima</i>									1.1								1.0				2
<i>Alternaria triticina</i>																1.9					1
<i>Amanita bisporigera</i>															5.0						1
Ascomycete sp.										3.0		2.3									2
<i>Aspergillus amstelodami</i>	2.2																	2.4			2
<i>Aspergillus caesiellus</i>																	9.5				1
<i>Aspergillus flavus</i>								3.0													1
<i>Aspergillus oryzae</i>							1.0														1
<i>Aspergillus penicillioides</i>							1.0														1
<i>Aspergillus ruber</i>	1.1						2.3														2
<i>Aspergillus</i> sp.	3.3	2.6					3.0	29.0									2.0		4.0		6
<i>Atractylodes macrocephala</i>	6.0						7.2														2
<i>Aureobasidium pullulans</i>			6.3		4.6			2.3		1.9		10.9			11.8		2.8	6.0	4.5	6.4	10
<i>Avena fatua</i>															5.0						1
<i>Bagnisiella examinans</i>								8.0													1
<i>Campylobacter concisus</i>											2.5										1
<i>Candida albicans</i>			9.0	10.4		7.8	10.7			14.2				14.1			2.0		2.4		8
<i>Candida khmerensis</i>							1.3														1
<i>Candida metapsilosis</i>																				2.7	1
<i>Candida parapsilosis</i>			2.4			3.2	1.0														3
<i>Candida</i> sp.																	1.0				1
<i>Candida tropicalis</i>	4.7															2.5					2
<i>Chaetomium globosum</i>												1.2									1
<i>Cladosporium cladosporioides</i>	1.2						1.6			1.3	46.8	2.8	26.1		32.8	2.3	11.5	11.8		1.1	10
<i>Cladosporium herbarum</i>							1.8														2
<i>Cladosporium</i> sp.							4.5	1.3		1.1	8.3		5.7		8.4	1.3	2.4	3.6			9
<i>Cladosporium sphaerospermum</i>											1.1										1
<i>Cladosporium tenuissimum</i>							1.2														1
<i>Claviceps purpurea</i>										1.1											1
<i>Clavispora lusitaniae</i>								2.5													1
<i>Conoplea fusca</i>								3.4													1
<i>Corynespora cassicola</i>	2.2																				1
<i>Cryptococcus cellulolyticus</i>																3.8					1
<i>Cryptococcus diffluens</i>					3.3																1
<i>Cryptococcus</i> sp.																1.8					1
<i>Cystofilobasidium lari-marini</i>																3.6					1
<i>Davidiella tassiana</i>										1.1											1
<i>Dothideomycete</i> sp.	10.7								1.5	4.1											3
<i>Dothioraceae</i> sp.			3.0		1.7						1.9		5.3		4.2		1.6	2.8	1.3	2.3	9
<i>Emericella rugulosa</i>					1.2												1.0				2
<i>Eurotium amstelodami</i>	1.1																		1.4		2
<i>Eurotium</i> sp.	9.6	1.3					4.0			1.1								15.5			5
<i>Filobasidiales</i> sp.																					0
<i>Flammulina velutipes</i>											8.8										1
<i>Fusarium culmorum</i>									1.9							4.3					2
<i>Fusarium oxysporum</i>												3.1									1
<i>Fusarium poae</i>										1.1											1
<i>Fusarium</i> sp.	4.0						2.1			1.1											3
<i>Gibberella</i> sp.				1.7		1.6			3.8							16.9					4
<i>Gigaspora margarita</i>					1.2																1
<i>Glomus fulvum</i>									1.8												1
<i>Glomus mosseae</i>	5.4				13.1		2.0														3
<i>Glomus</i> sp.																7.1					1
<i>Guignardia mangiferae</i>															3.4						1
<i>Guignardia</i> sp.															1.7						1
<i>Hormonema</i> sp.													1.0						1.2		2
<i>Hypocreales</i> sp.		1.0																		5.2	2
<i>Issatchenkia orientalis</i>							1.6														1

Organism	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12	H13	H14	H15	H16	H17	H18	H19	H20	#		
<i>Lasiodiplodia theobromae</i>	2.8																					1	
<i>Lecania gerlachei</i>																					4.4	1	
<i>Lewia infectoria</i>				1.1					2.2													2	
<i>Macrophomina phaseolina</i>		3.5																				1	
<i>Nectria cinnabarina</i>													1.0									1	
<i>Nigrospora oryzae</i>										1.2		48.4										2	
<i>Ochroconis gamsii</i>	3.2																7.6					2	
<i>Ophiotoma floccosum</i>																					40.0	1	
<i>Ophiotoma pulvinisporum</i>																					1.3	1	
<i>Orpinomyces</i> sp.								2.1														1	
<i>Paecilomyces lilacinus</i>																				1.6		1	
<i>Penicillium brevicompactum</i>						1.8																1	
<i>Penicillium glabrum</i>			11.1																			1	
<i>Penicillium</i> sp.	1.8																					1	
<i>Penicillium spinulosum</i>			1.6																			1	
<i>Peziza violacea</i>											3.7											1	
<i>Phaeosphaeria nodorum</i>						1.5																1	
<i>Phoma foveata</i>					2.2																	1	
<i>Phoma plurivora</i>										1.7												1	
<i>Pichia guilliermondii</i>																				1.1		1	
<i>Pleosporaceae</i> sp.	8.4									1.1												2	
<i>Pongo abelii</i>																		4.0				1	
<i>Pueraria montana</i>										1.4												1	
<i>Pyrenophora tritici-repentis</i>										2.1												1	
<i>Rhizopogon</i> sp.			6.1																			1	
<i>Saccharomyces bayanus</i>										3.5												2.3	2
<i>Saccharomyces cerevisiae</i>				8.9					43.2			2.1						1.1	2.4	25.8		6	
<i>Saccharomyces ellipsoideus</i>									2.4													1.3	2
<i>Saccharomycetales</i> sp.			5.6	13.7	1.1	12.4	6.0			8.0			1.8	18.1	0.0	0.0	1.4		7.3	0.0		13	
<i>Schizophyllum commune</i>			14.7													2.6						2.2	3
<i>Schizosaccharomyces japonicus</i>	20.0																					1	
<i>Schizosaccharomyces pombe</i>						1.1																1	
<i>Sebacina incrustans</i>		1.8																				1	
<i>Sordariomycete</i> sp.		1.7														1.3						2	
<i>Stachybotrys chartarum</i>						6.2		1.3		1.2												3	
<i>Stemphylium solani</i>																				6.3		1	
<i>Subulispora britannica</i>												2.5										1	
<i>Teratosphaeria</i> sp.							1.2	1.9			17.1	1.1	17.7	16.8	1.1	10.9	5.3					9	
<i>Torulaspora delbrueckii</i>													2.1									1	
<i>Tremellales</i> sp.								1.9														1	
<i>Trichaptum abietinum</i>						2.2																1	
<i>Trichocomaceae</i> sp.						2.6	2.8												2.3			3	
<i>Trichosporon loubieri</i>																4.1						1	
<i>Non-Cultivable ascomycete</i>	4.1	2.1	25.5	37.2		25.0	1.2		3.5	28.5			5.5	48.2		1.9	5.6		21.3			13	
<i>Non-Cultivable basidiomycete</i>					37.1	2.2			2.1									3.1				4	
<i>Non-Cultivable ectomycorrhizal</i>														1.4	2.5							2	
<i>Non-Cultivable endophytic</i>	1.1	2.5		1.1				1.2	2.9							1.2		1.3	1.1			8	
<i>Non-Cultivable eukaryote</i>	2.4						3.1															2	
<i>Non-Cultivable fungus</i>		5.1	3.1	2.3	1.6	3.3	3.2		3.2	1.2			11.7	1.9	3.4	6.1	29.0	17.6	9.7	3.2		16	
<i>Non-Cultivable glomeromycete</i>			4.3							7.1				8.2								3	
<i>Non-Cultivable Glomus</i>				6.2		5.4	2.2						1.4				3.0	8.0	1.8			7	
<i>Non-Cultivable Leptosphaeriaceae</i>		9.6																				1	
<i>Non-Cultivable soil</i>	1.4	11.7			9.4	3.8	3.8	2.4	1.6	0.0	0.0	2.8							1.9	1.1		12	
<i>Non-Cultivable Unknown</i>	4.4	1.5		3.1	6.3	11.3	14.3	18.2	5.4	4.0	0.0	4.7				20.8		2.2	16.3			14	
<i>Verticillium</i> sp.																1.4						1	
<i>Wallemia sebi</i>							7.3															1	
<i>Xylariales</i> sp.		7.7											31.3									2	
<i>Zygosaccharomyces pseudorouxii</i>									1.9													1	
<i>Zygosaccharomyces rouxii</i>								2.1														1	
Count*	19	21	17	17	14	27	22	19	21	25	10	15	15	12	10	22	13	20	23	19			

# is the total samples that contained that specific taxon

\* is the total number of taxa present per sample

## **Chapter 4**

### **“SPR Detection of Bacteria Associated with the Cystic Fibrosis Lung”**

#### 4.1 Abstract

Traditional bacterial identification systems use phenotypic and biochemical analysis of the organism and/or antibody-based detection systems like ELISA. Antibody-antigen interactions can be observed on a nanoscale level using more sensitive techniques such as surface plasmon resonance (SPR). This technique detects bacteria by indirectly monitoring the oscillations of electrons that exist at the surface of an antibody-coated gold sensor chip. As a result of antibodies capturing bacteria, electron oscillations will change which is detectable by measuring the angle of light that is absorbed by the electrons. This study proposed to detect the interactions of *Pseudomonas aeruginosa* and *Staphylococcus aureus* with their cognate antibodies using SPR. The immunosensor detected *P. aeruginosa* and *S. aureus* at concentrations of  $10^7$  to  $10^8$  colony forming units per milliliter (cfu/mL). The sensitivity of the *P. aeruginosa* immunosensor was increased by a factor of ten when bovine serum albumin (BSA) was not added to the sensor. The immunosensor was specific showing only a minimum reaction to *Escherichia coli* at  $10^8$  cfu/mL. Optimization of the immunosensor assembly will increase the sensitivity and accuracy of the biosensors.

**Keywords:** Surface plasmon resonance, *Pseudomonas aeruginosa*, *Staphylococcus aureus*

**Abbreviations:** SPR, Surface Plasmon Resonance; SAM, Self-Assembled Monolayer, BSA, Bovine Serum Albumin; cfu, colony forming units; cystic fibrosis (CF); phosphate buffer saline (PBS); overnight (ON)

## 4.2 Introduction

Per every million people, 2.4 people die in the United States as a result of a bacterial infection (World Health Organization, 2008). Despite being an industrialized nation, the US is ranked 4<sup>th</sup> (for bacterial associated death) out of the 47 countries that give health statistics to the World Health Organization. Many bacterial infections are treatable. If not treated, an infection can lead to sepsis and ultimately death. An accurate and timely diagnosis is key in eliminating the bacterial infection and preventing the spread of the organisms to uninfected people.

Identification of the infecting organism is critical for successful treatment. Culturing for phenotypical and biochemical assays is the mainstay in most clinical laboratories. Using various media, the sample from the infected person is inoculated onto plates or broth and then the organism is allowed to grow for a period of time (Murray & Baron, 2007). Based on the phenotype of the organism and the media it is grown on, an identification of the bacteria is made. Besides visual phenotypes, biochemical tests are performed. Currently, multiple biochemical tests can be performed simultaneously using diagnostic kits such as the BioMérieux API system or the VITEK 2 by BioMérieux (Bosshard *et al.*, 2006). These products identify bacteria within two to 24 hours. Both of these assays are not error-proof leading to misidentifications (Bosshard *et al.*, 2006). The accuracy of the VITEK2 ID-GNB card (BioMérieux, Marcy l'Etoile, France) has been examined (Bosshard *et al.*, 2006; Brisse *et al.*, 2002; O'Hara & Miller, 2003). This system accurately identified 93.0 % of 414 clinically relevant enteric bacteria strains and 95.1 % of the 103 glucose-fermenting and nonfermenting nonenteric strains (O'Hara & Miller, 2003). Another study compared the



VITEK2 ID-GNB card (BioMérieux) and API 20 NE (BioMérieux) with sequencing data from the same 107 isolates (Bosshard *et al.*, 2006). The API system was able to identify 54 % of the isolates to species level, 7 % to the genus level and failed to identify 39 % of the isolates. Of the later, 26 % belonged to species that were not in the API database (Bosshard *et al.*, 2006). The VITEK 2 system had similar results whereas 53 % of the isolates were identified to species level, one percent to genus level, and 46 % (49 isolates) were unidentifiable. Of the 49 isolates, 39 belonged to a species not represented in the VITEK 2 database (Bosshard *et al.*, 2006). The VITEK has also been compared with the BD Phoenix (BD Diagnostic System, Sparks, MD) to determine the accuracy in identifying isolates of the *B. cenocepacia* complex, a known CF pathogen (Brisse *et al.*, 2002). The study determined an overall identification rate of 50 % for both systems (Brisse *et al.*, 2002). The accuracy of commercial systems varies greatly but no system has a 100 % identification rate. Misidentification can lead to improper treatment which can result in poor patient outcome and lead to a rise in antibiotic resistance (Levy, 2000).

Pathogens can also be identified based on antibody – antigen interactions using an enzyme-linked immunosorbent assay (ELISA). There are multiple ELISA protocols but the original method begins by attaching capture antibodies to a polystyrene plate. Then, the sample is placed into the system and if present the bacteria will bind to the antibody. The antibody-antigen interaction is detected using either chromogenic or fluorogenic methods (Engvall & Perlmann, 1971). The ELISA is not portable but still useful when analyzing clinical samples.

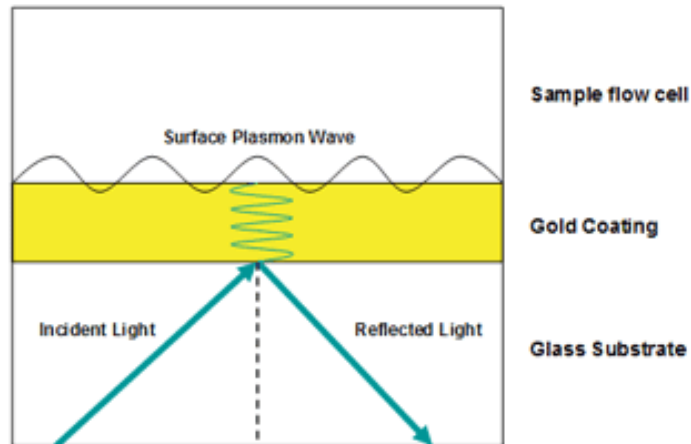
Rapid and portable diagnostic systems have been developed based on the concept used in ELISA. One category of antibody-based detection systems are referred to as immunochromatographic tests. One of the most well known examples of this type is the at home pregnancy test (Vaiukaitis *et al.*, 1972). The dipstick is coated with mouse monoclonal antibodies that recognize the human chorionic gonadotrophin molecules that are present in urine when pregnant (Vaiukaitis *et al.*, 1972). The molecule binds to the antibody and the complex moves down the strip where it is ensnared by polyclonal antibodies which bind to the human chorionic gonadotropin. At the location of the second reaction are dye molecules which undergo an enzymatic reaction to produce the colored line (Vaiukaitis *et al.*, 1972). Similar detection systems have been developed for diphtheria (Engler *et al.*, 2002), HIV (Esfandiari *et al.*, 2004), *Mycobacterium tuberculosis* (Esfandiari *et al.*, 2004), vaginal *Candida* species (Matsui *et al.*, 2009), and *Escherichia coli* O157:H7 (Park *et al.*, 2008). These systems can use various bodily fluids such as urine and sera as well as cultured bacteria in liquid solutions.

The newest generation of immunochromatographic tests use nanotechnology. The secondary capture antibody in these tests have been conjugated with colloidal gold nanoparticles (Biagini *et al.*, 2006; Chiao *et al.*, 2008; Rong-Hwa *et al.*, 2010; Shyu *et al.*, 2002). Using lateral flow, the analyte binds to the primary antibody and then moves towards the labeled secondary antibody forming another complex. The sandwich complex flows to the capture pad on the test strip which will turn red when sufficient nanoparticles accumulate. These tests have been developed to detect bacteria such as *Staphylococcus aureus* and numerous toxins such as ricin, botulism, and staphylococcal enterotoxin B (Biagini *et al.*, 2006; Chiao *et al.*, 2008; Rong-Hwa *et al.*, 2010; Shyu *et*

*al.*, 2002). These tests may be the future of diagnostics as that they are rapid, portable, user-friendly, cost-efficient, and have a long shelf-life (Rong-Hwa *et al.*, 2010).

Antibody-antigen interactions can also be observed with instruments that use surface plasmon resonance (SPR). The SPR effect is a phenomenon where electrons that exist at the interface between any two materials oscillate and produce waves (Wood, 1902, 1912). This phenomenon can be used to determine if an antigen binds to an antibody (Kretschmann & Raether, 1968; Otto, 1968). Using a SPR instrument, light is passed through a glass chip thinly coated with gold (Figure 4.1). A specific frequency of light passes into the gold creating an evanescent wave. The wave excites the plasmons causing them to resonate and form a plasmon wave which travels along the gold film interacting with the media in contact with the gold. The remaining frequencies of light are reflected at a specific angle. The angle of the light reflection, which can be measured, shifts as molecules adsorb onto the surface of the gold (Kretschmann & Raether, 1968; Schasfoort & McWhirter, 2008).

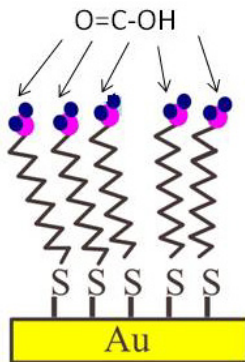
One specific type of SPR biosensors is an immunosensor which detects interactions between antibodies and other molecules such as bacteria (Geddes *et al.*, 1994; Mayo & Hallock, 1989). Essentially, SPR immunosensors detect antigens that bind to antibodies on the gold sensor by measuring the angle of light reflection (Kretschmann & Raether, 1968). SPR immunosensors have been used to identify *Escherichia coli* O157:H7 (Fratamico *et al.*, 1998; Meeusen *et al.*, 2005; Oh *et al.*, 2003; Su & Li, 2005; Subramanian *et al.*, 2006a; Taylor *et al.*, 2005), *Vibrio cholera*



**Figure 4.1:** Surface plasmon resonance phenomenon. Incident light penetrates through a gold coated glass chip and a specific frequency of light will create an evanescent wave. This wave excites the plasmons and cause a plasmon wave to propagate along the surface of the gold. The remaining light is reflected. Adapted from [www.icxt.com](http://www.icxt.com).

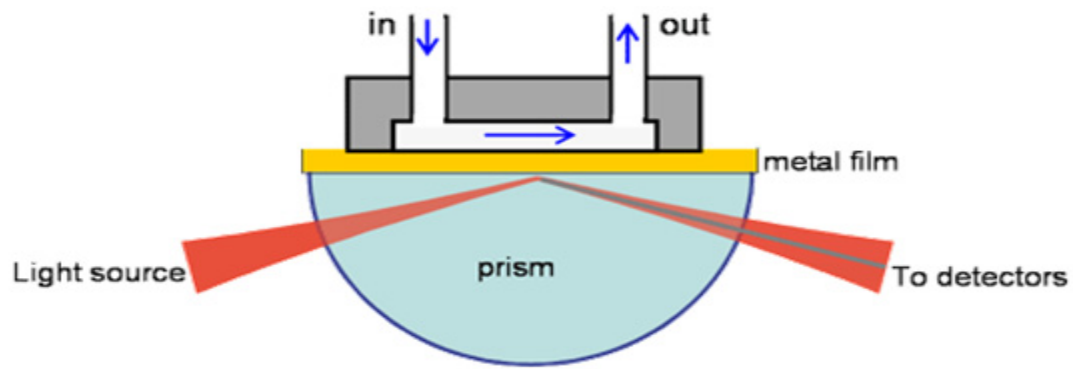
(Jyoung *et al.*, 2006), *Salmonella enteritis* (Bokken *et al.*, 2003), *Salmonella typhimurium* (Oh *et al.*, 2004a), *Salmonella paratyphi* (Oh *et al.*, 2004b), *Listeria monocytogenes* (Hearty *et al.*, 2006; Koubova *et al.*, 2001; Leonard *et al.*, 2005), *Yersinia enterocolitica* (Oh *et al.*, 2005), and *Campylobacter jejuni* (Taylor *et al.*, 2006).

Immunosensors are designed using a layer-by-layer self-assembly method (Cui *et al.*, 2003). As that the design protocol can vary, the method used in this study is explained in depth. The immunosensor self-assembly method starts with a glass chip that is thinly coated with gold on one side. The chip is prepared by soaking it in 11-mercaptoundecanoic acid ( $\text{HSC}_{10}\text{H}_{20}\text{COOH}$ ) or similar alkanethiol solutions to allow a self-assembled monolayer (SAM) formation. In this case, the gold molecule binds to the thiol group in 11-mercaptoundecanoic acid. The carbon chain extends from the thiol and is terminated with a carboxylic acid group (Figure 4.2) (Stettner *et al.*, 2009).



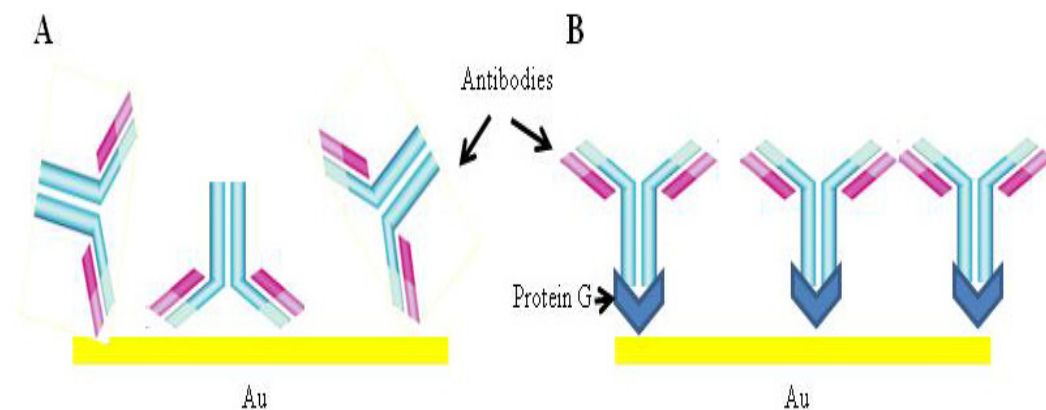
**Figure 4.2:** Self-assembled monolayer (SAM) formation on a gold coated chip. A gold sensor chip is immersed in 11-mercaptoundecanoic acid. The thiols bind to the gold leaving a chain of carbons ending with a carboxycyclic acid group (Chang & Liao, 2008).

After the formation of SAM, chip is placed onto the SPR instrument for layer-by-layer self assembly. The gold sensor is set onto the window of the instrument using a matching liquid (refractive index = 1.515) to ensure that the refractive index of the sensor is the same as the glass window (Biosensing Instrument, 2006). The instrument uses microfluidics which allows a carrier solution to constantly run over the sensor. The carrier is a reference buffer which establishes the baseline angle of light reflection. Molecules such as antibodies are injected into a port and carried to the sensor by the carrier buffer. After the injection is complete, the port closes and the carrier solution washes away unbound injected molecules. Light is constantly beamed onto the sensor and reflected as the carrier solution and injected chemicals are delivered to the sensor (Figure 4.3) (Biosensing Instrument, 2006). The angle of reflection will change as



**Figure 4.3.** Schematic diagram of SPR instrument. The carrier buffer and/or injected chemicals flow into one port, across the gold sensor chip, and then the buffer with any unbound chemicals exits via the out port. As the buffer flows, a light source is continuously beamed through a prism and into the gold chip. The light is reflected and the angle of reflection is measured by a detector. The angle of reflection is dependent on the molecules binding to the gold chip (Biosensing Instrument, 2006).

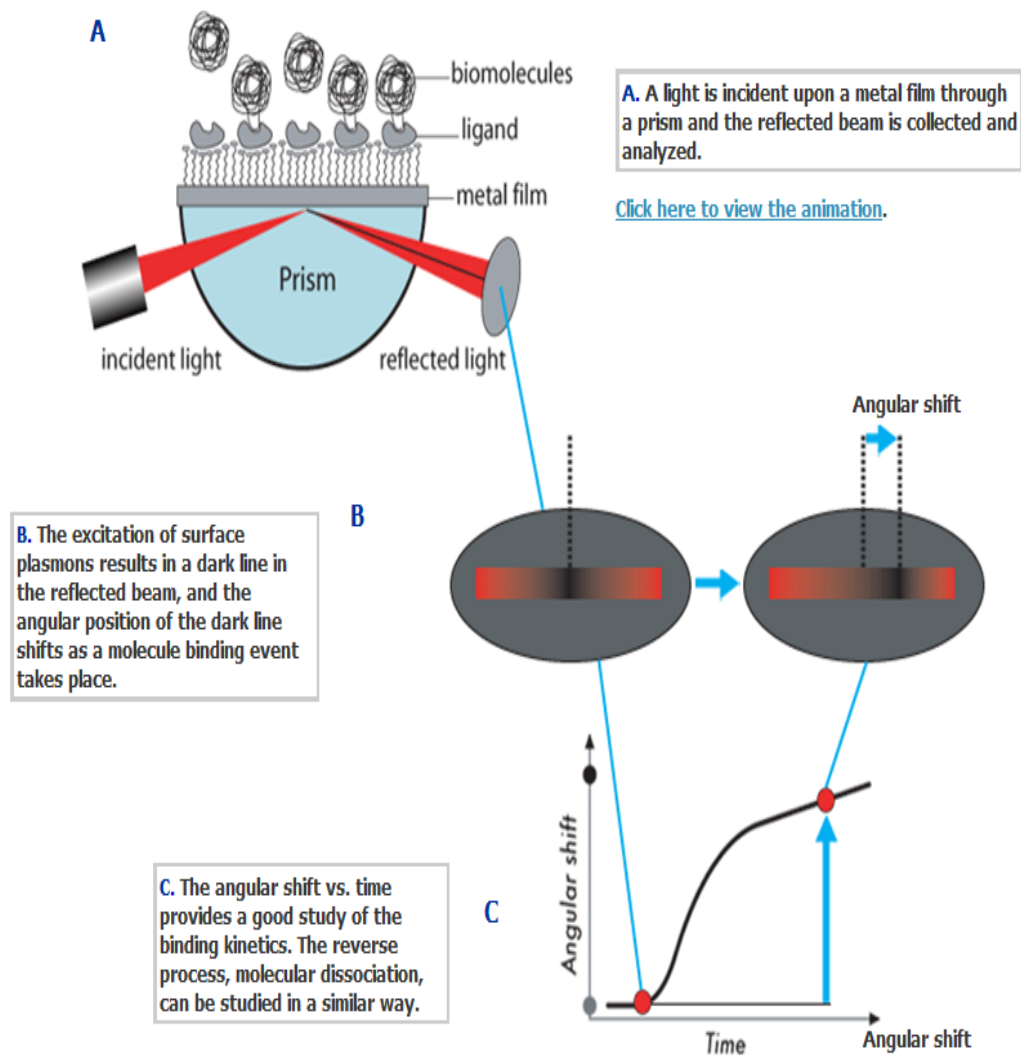
molecules adsorb onto the gold sensor. In between injections, the carrier/reference buffer flows over the sensor and the angle will stay constant forming a baseline measurement. bonds between carboxylic acids and amines. EDC converts the carboxylic acid into a reactive intermediate (an ester) which can be attacked by amines. The NHS is used to stabilize the intermediate for a longer period allowing an amide to be formed (Staros *et al.*, 1986). A solution of Protein G is then injected, and the nitrogen groups on the protein attack the ester groups forming an amide bond between the carboxylic acid group on the SAM and the amines on the Protein. Protein G molecules increase the layer of thickness on the gold sensor which affects the surface plasmons, thereby changing the angle of reflection. Protein G is added to ensure the antibodies bind in the proper orientation (Figure 4.4) (Lee *et al.*, 2007; Schmid *et al.*, 2006).



**Figure 4.4.** Antibody orientation. Panel A demonstrates antibodies binding to gold in multiple orientations. Many antibodies bind in an orientation that compromises the ability of the antibodies to bind to antigens. Panel B shows antibodies in the proper orientation as a result of the Protein G molecules attached to the gold layer.

Any remaining unbound activated ester groups are deactivated by an ethanolamine solution. The antibodies are injected and bind to the Protein G molecules in the proper orientation leaving the antigen binding sites on the antibodies exposed (Figure 4.3). Antibodies may not have bound to all Protein G molecules leaving Protein G free to bind to non-selective antigens. To prevent a false-positive reaction, the Protein G molecules are blocked by bovine serum albumin (BSA) which finishes the layer-by-layer self assembly of the immunosensor. Each addition increases the thickness of the layer adsorbed to the gold which directly affects the reflection of light. The shifts in angle of reflection can be measured in real-time (Figure 4.5).

To date, there are no SPR immunosensors for whole cell *Staphylococcus aureus* and *Pseudomonas aeruginosa*, two common opportunistic pathogens that are feared in



**Figure 4.5:** The SPR response. Panel A depicts a layer-by-layer self assembled gold sensor chip on the SPR window. Incident light enters through a prism hitting the gold sensor. The light is reflected. Panel B is a close up view of the reflected beam of light. The incident light excited surface plasmons that were in contact with the layers above the gold film. As a result of the electrons absorbing energy at a particular frequency of light, the reflected beam has a dark line within it. The angle of the dark line within the beam is the SPR response and it shifts depending on surface layers above the gold sensor chip. Panel C is a sensograph. The time and angular shift are measured throughout a SPR experiment.



hospitals. *S. aureus* is a Gram-positive cocci that can be found in the skin flora of healthy humans (Hartmann, 1978). The bacteria can also cause numerous infections due to its ability produce virulence factors and toxins (Freer & Arbuthnott, 1982; Jakubicz, 1972; Kapral *et al.*, 1965; Parker, 1924). *S. aureus* has been associated with a range of infections from mild acne to life-threatening diseases such as pneumonia and septicemia (Corner, 1940; Noble, 1998; Olansky & McCormick, 1958; Schneck, 1957). The bacteria can also cause respiratory infections in cystic fibrosis (CF) children and recent research has shown that the bacteria are still present in the lungs of adult CF patients (Beringer & Appleman, 2000; Kahl *et al.*, 1998; Patient-registry, 2008). In addition, *S. aureus* is one of the leading causes of nosocomial infections (Noskin *et al.*, 2005; Wertheim *et al.*, 2005). These infectious could once be treated with penicillin, but now the bacteria are resistant to it and other beta-lactam antibiotics (Ashley & Brindle, 1960; Zetola *et al.*, 2005). The emerging resistant strains of *S. aureus* make it more of a challenge to treat a patient thus driving up the cost of treatment (Kim *et al.*, 2001).

*P. aeruginosa* is a Gram-negative rod-shaped bacterium that is known for its environmental versatility and ubiquity (lives in water, soil, and in animal tissues) (Ringgen & Drake, 1952; Winslow *et al.*, 1920). The organism is capable of surviving in numerous environments due to its ability to metabolize a vast array of carbon sources (Conrad *et al.*, 1979; Haynes, 1951; Niel, 1943; Ringgen & Drake, 1952). *P. aeruginosa* produces numerous virulence factors and toxins which enable it to infect plants, animals, and humans that are either immunocompromised, have burn wounds or have CF (Caldwell *et al.*, 2009; Doggett *et al.*, 1964; Elrod & Braun, 1942; Finck-Barbançon *et al.*, 1997; Griffith *et al.*, 1989; Lowbury & Fox, 1954; Niilo, 1959; Wretling &

Pavlovskis, 1981; Young, 1980). Up to 80 % of all CF adults have a *P. aeruginosa* lung infection (Beringer & Appleman, 2000; Patient-registry, 2008). It is also one of the leading pathogens in nosocomial infections (Aloush *et al.*, 2006). *P. aeruginosa* is intrinsically resistant to many antibiotics and is capable of acquire resistance to other drugs over the course of the infection (Fullbrook *et al.*, 1970; Li *et al.*, 1994; Li *et al.*, 1998).

Both *S. aureus* and *P. aeruginosa* cause infections which can be difficult and expensive to treat. Identifying the agents of infection is the first step in treatment. In this study, SPR biosensors will be developed using *Staphylococcus*-specific and *P. aeruginosa*-specific antibodies. The corresponding bacteria will be detected based on the change of light refraction that is measured when the bacteria binds to the antibody that is fixed on the gold sensor.

### **4.3 Materials and methods**

**4.3.1 Bacterial isolates.** *Escherichia coli* DH5 $\alpha$  (New England Biolabs, Ipswich, MA), *Pseudomonas aeruginosa* PAO1 (Holloway & Morgan, 1986), and *Staphylococcus aureus* ATCC 17672 were grown overnight (ON) at 37 °C in Luria Bertani (LB) broth (10 g/L tryptone, 5g/L yeast extract, 5 g/L NaCL). One milliliter of the ON was placed into 5 mL of fresh LB and grown until an optical density at 600 nm (OD<sub>600</sub>) was between 0.4 and 0.5. The cultures were pelleted by centrifugation at 16.6 X g for 10 minutes at 4 °C. The pellets were resuspended in 1X phosphate buffer saline (PBS) pH 7.0. The samples were pelleted and washed two more times with 1X PBS. The pellets were

resuspended in the PBS to achieve a final concentration of  $10^8$  colony forming units per milliliter (cfu/mL).

**4.3.2 Determination of bacterial concentration.** For each isolate, the washed pellet was resuspended in 10 mL of phosphate buffer saline (PBS). Serial dilutions were made up till  $10^6$  dilution factor. One hundred microliters for the  $10^4$ ,  $10^5$ , and  $10^6$  dilutions were plated on LB plates. The plates were grown ON at 37 °C. Colonies were counted the next day to determine the colony forming units per milliliter (cfu/mL) for each isolate at its particular OD<sub>600</sub>. Based on the cfu/mL determination, a specific amount of 1X PBS was added to each pellet to make a  $10^8$  cfu/mL solution. Bacteria for each experiment was then grown to the OD<sub>600</sub> reading used in the cfu/mL calculation so that fresh bacteria could be used in the SPR experiments.

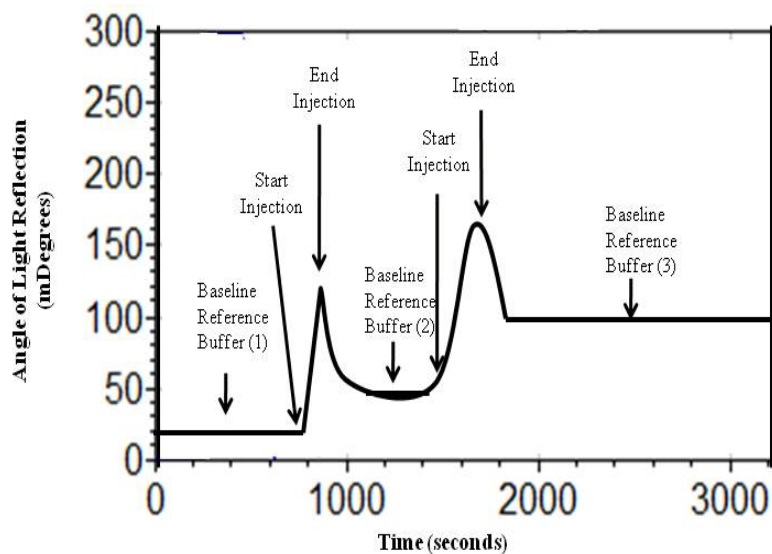
**4.3.3 Reagents and instrumentation.** The Biosensing Instrument 3000 and the associated Biosensing software and gold sensor chips (Biosensing Instrument, Tempe, AZ) were used for all experiments. All chemicals were purchased from Sigma-Aldrich (St. Louis, MO) unless otherwise noted. The antibody that was reactive with *P. aeruginosa* was rabbit polyclonal IgG anti-*P. aeruginosa* whole cell (ABCAM ab68538; Abcam, Cambridge, MA). Anti-*Staphylococcus* enterotoxin B, polyclonal was used to detect *S. aureus* (Thermo Scientific PA17248; Thermo Fischer Scientific, Waltham, MA).

**4.3.4 Preparation of gold sensor chip.** Gold sensor chips were immersed in 50 mM 11-mercaptoundecanoic acid overnight at 4 °C to allow self-assembled monolayer (SAM) formation. The chips were washed with deionized water and dried.

**4.3.5 Immobilization of antibodies.** The SAM modified gold sensor was placed onto the window of the instrument using a matching liquid (refractive index = 1.515). To allow for baseline stabilization, the reference solution which was 1X PBS (pH 7.0) was allowed to flow through the system at a flow rate of 0.150 mL/min. The flow rate was decreased to 0.050 mL/min once a baseline was established. The flow injection method was used throughout all experiments. The SAM was activated by injecting 125 cc of a 1:1100 mM N-hydroxysuccinimide (NHS) and 400 mM N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (EDC). The reference solution was allowed to run for 60 sec between injections to flush out the system. To increase the directional binding of the antibody, Protein G was coupled to the activated SAM by injecting 125 cc of 10 µg/mL Protein G in 1X PBS (pH 7.0) into the instrument. Any remaining unbound activated ester groups were deactivated by 125 cc 1 M aqueous ethanolamine solution (pH 8.0). Sixty seconds later, 125 cc of 1 mg/ml antibody was added. The Protein G molecules that have no attached antibodies were blocked with 125 cc of 1 % BSA (depending on the experiment). Excess BSA was removed from the sensor by running the reference buffer for 60 seconds. All injection times were 10 µL/min except for the antibodies and bacteria which were set at 5 µL/min.

The assembly process was monitored continuously by measuring the angle of light reflection (SPR response) in mDegrees. The data was plotted on to a sensograph

with time in seconds on the x-axis versus resonance in mDegrees on the y-axis. Peaks in the sensograph indicated changes in SPR response caused by the injected chemicals coming into contact with the functionalized gold surface. Unbound molecules were washed away by carrier buffer causing the peak to decrease and a new baseline to form (Figure 4.6). The difference between post-injection and pre-injection baselines (difference in angles) indicated the change in plasmon resonance caused by molecules adsorbing onto the gold layer (Figure 4.6). The angle of light reflection shifted as each compound was added to the sensor.



**Figure 4.6:** A representative sensograph. A theoretical SPR sensograph is depicted with time in seconds on the x-axis and angle reflection on the y-axis. Baseline reflections are stable in between injections. Injections cause a change in reflection. The reference buffer washes away the excess injected chemical. The adsorption of the chemical on to the gold layer is determined by subtracting the pre-injection baseline reference angle from the post-injection baseline reference angle i.e. Baseline Reference Buffer 2 – Baseline Reference Buffer 1.

Throughout the experiment, the carrier buffer of PBS continuously flowed past the sensor. The SPR response to the reference buffer forms a baseline measurement which is recorded in a separate sensograph. The sensograph produced from the injected compounds is subtracting from the reference sensograph, thereby removing background noise. The resulting normalized data is plotted on another sensograph and analyzed using the Biosensing Real Time Data Analysis software (Biosensing Instrument, Tempe, AZ).

**4.3.6 Detection of *S. aureus* and *P. aeruginosa*.** The prepared immunosensors were tested using 125 cc of bacteria at difference concentrations. Bacterial concentrations ranged from  $10^6$  to  $10^8$  cfu/mL for *S. aureus* and  $10^5$  to  $10^8$  cfu/mL for *P. aeruginosa*. SPR was used to detect the interaction between *S. aureus* and anti-*Staphylococcus* antibodies, and *P. aeruginosa* and anti-*P. aeruginosa* antibodies.

**4.3.7 Determination of the sensitivity for the *P. aeruginosa* biosensor.** A second immunosensor was developed for *P. aeruginosa*. In the previous experiments, BSA was used to decrease non-selective binding. However, BSA is a large molecule that increases the thickness of the layers on top of the gold surface. SPR is less sensitive when the layers above the gold surface becomes too thick. In addition, the instrument cannot detect SPR when the distance between the bacteria and the gold becomes too great. Thus to increase the sensitivity of the sensor, BSA was not added during the assembly process. The immunosensor was tested using *P. aeruginosa* concentrations from  $10^5$  to  $10^8$  cfu/mL.

**4.3.8 Determination of the specificity for the *P. aeruginosa* sensor.** To determine if BSA was necessary to reduce non-selective binding, another *P. aeruginosa* biosensor was developed without BSA. The sensor was first tested with  $10^8$  cfu/mL of *E. coli*. After the reference buffer washed away any unbound bacteria, the sensor was tested using  $10^7$  and  $10^8$  cfu/mL of *P. aeruginosa*.

#### **4.4 Results**

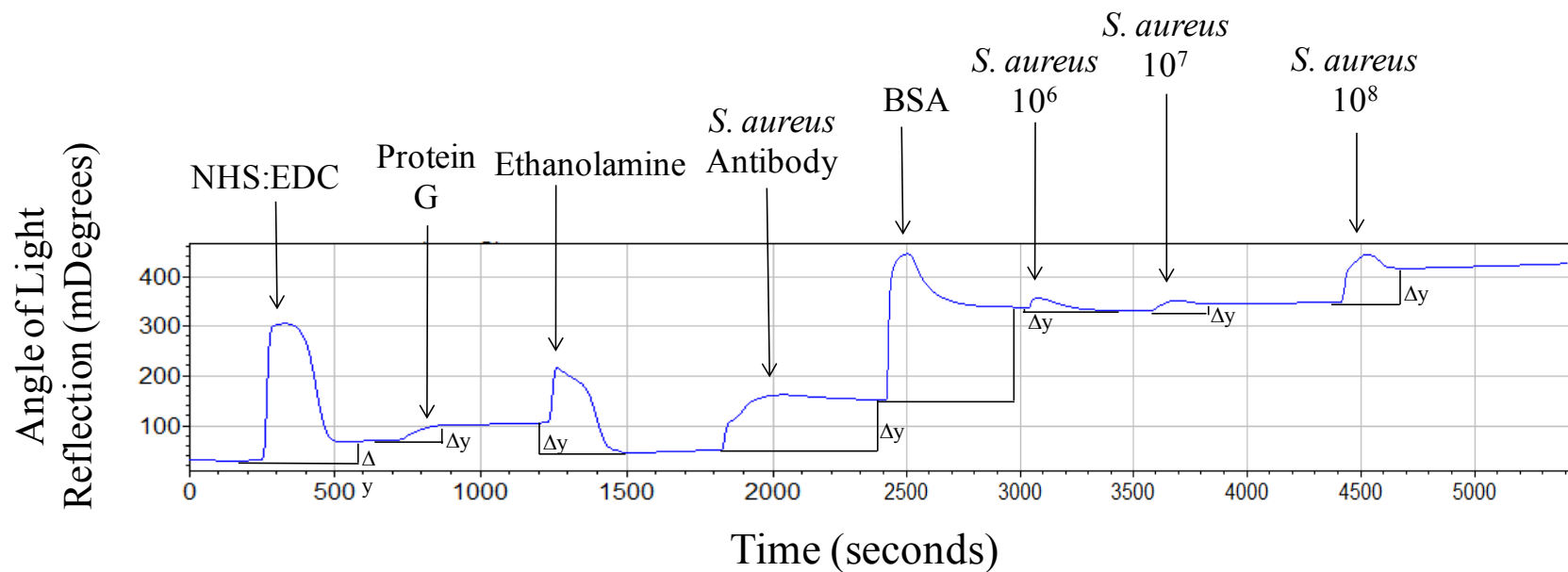
In order to detect the presence of bacteria in a solution, immunosensors were manufactured using a layer-by-layer self assembly protocol. The immersion of the chip in 50 mM 11-mercaptoundecanoic acid allowed self-assembled monolayer (SAM) formation of thiols which were activated using a 1:1 ratio of NHS:EDC. The generated reactive ester groups formed amide bonds with Protein G molecules. The antibodies were then anchored to the chip via Protein G. Free Protein G molecules were blocked with BSA to finish the assembly of the sensor. The SPR instrument monitored each molecule absorbing on to the gold layer in real time. The molecules absorbed onto the gold increased the thickness of the layer on top of the gold surface changing the plasmon resonance which affected the angle of light reflection. After each chemical was injected into the carrier solution, the change in SPR angle (mDegrees) was determined. The baseline between peaks on the sensograph indicated the angle of incident for the reference or carrier buffer flowing pass the sensor which is dependent on the thickness of the layer above the gold chip. The peaks represented the chemical injected into the carrier solution flowing over the sensor. The angles of reflection within the peak area do not indicate molecules adsorbing onto the gold. Rather, adsorption of molecules onto the

gold sensor is determined by the difference in baseline measurements before and after injection (Figure 4.6).

**4.4.1 Self assembly of *S. aureus* immunosensor.** In the *S. aureus* immunosensor, the reference buffer flowing past the sensor gave a baseline angle of light incident at 37.02 mDeg (Figure 4.7). This angle shifts within a degree or two as the buffer flows above the sensor due to the limitation of the SPR instrument. The injection of NHS:EDC caused the reflection to change 41.9 mDeg in comparison to the initial baseline (Figure 4.7). The reference buffer washed away unbound molecules. A solution of Protein G was injected and the reflection changed 47.523 mDeg (Figure 4.7). The reference buffer continued to run as the solution of the anti-*Staphylococcus* antibodies was injected. The SPR response was an angle 100.63 mDeg of light reflection (Figure 4.7). Unbound 1 % BSA was injected to block free Protein G molecules. The angle of incidence was changed to 187.29 mDeg (Figure 4.7).

**4.4.2 Detection of *S. aureus*.** The layer-by-layer self assembled immunosensor was used to detect various concentration of *S. aureus*. The sensor was first tested with  $10^6$  cfu/mL of *S. aureus* in PBS (Figure 4.7). The peak in the sensogram after the injection showed a 22.36 mDeg angle of reflection. Baseline measurements before and after injection were not different; there was no change in light reflection (Figure 4.7). There was a slight change in the angle of light reflection for  $10^7$  cfu/mL at 16.73 mDeg. However, a more pronounced reflection was observed for  $10^8$  cfu/mL at 78.27 mDeg.

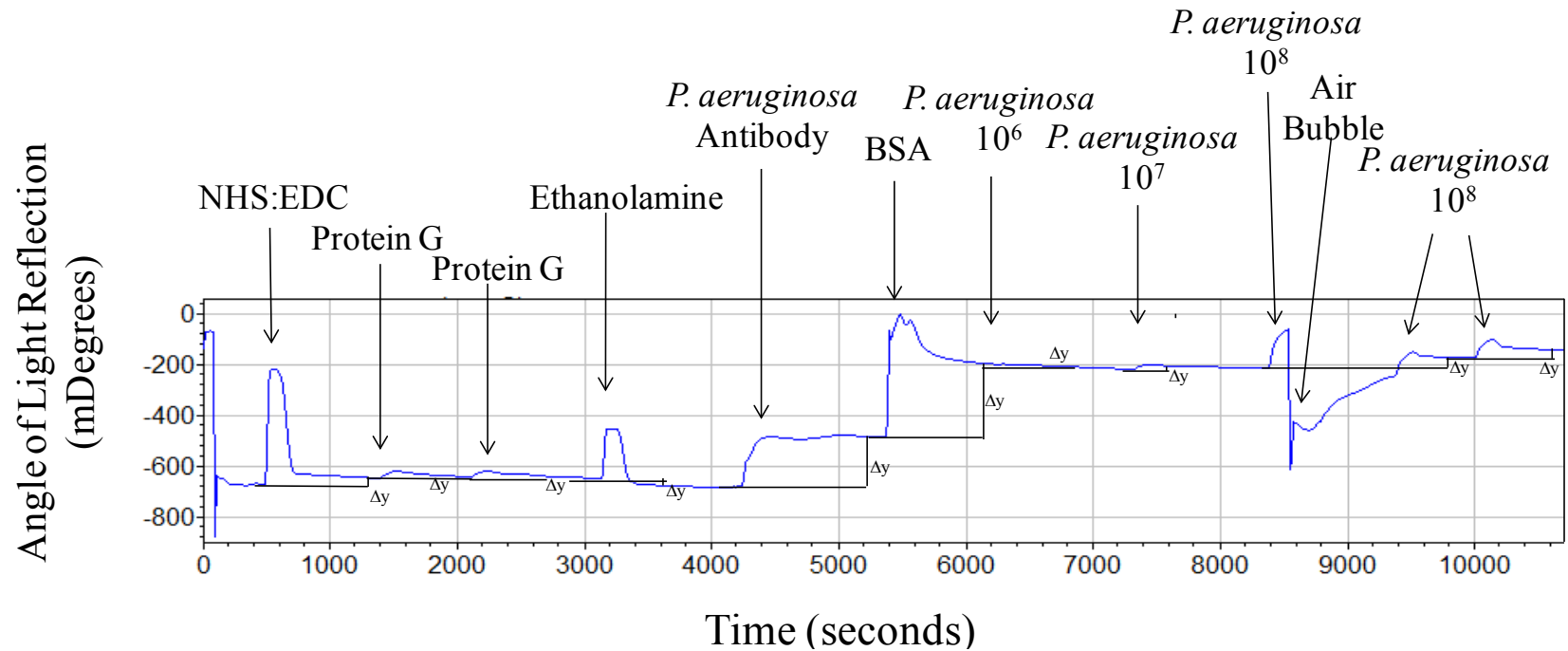




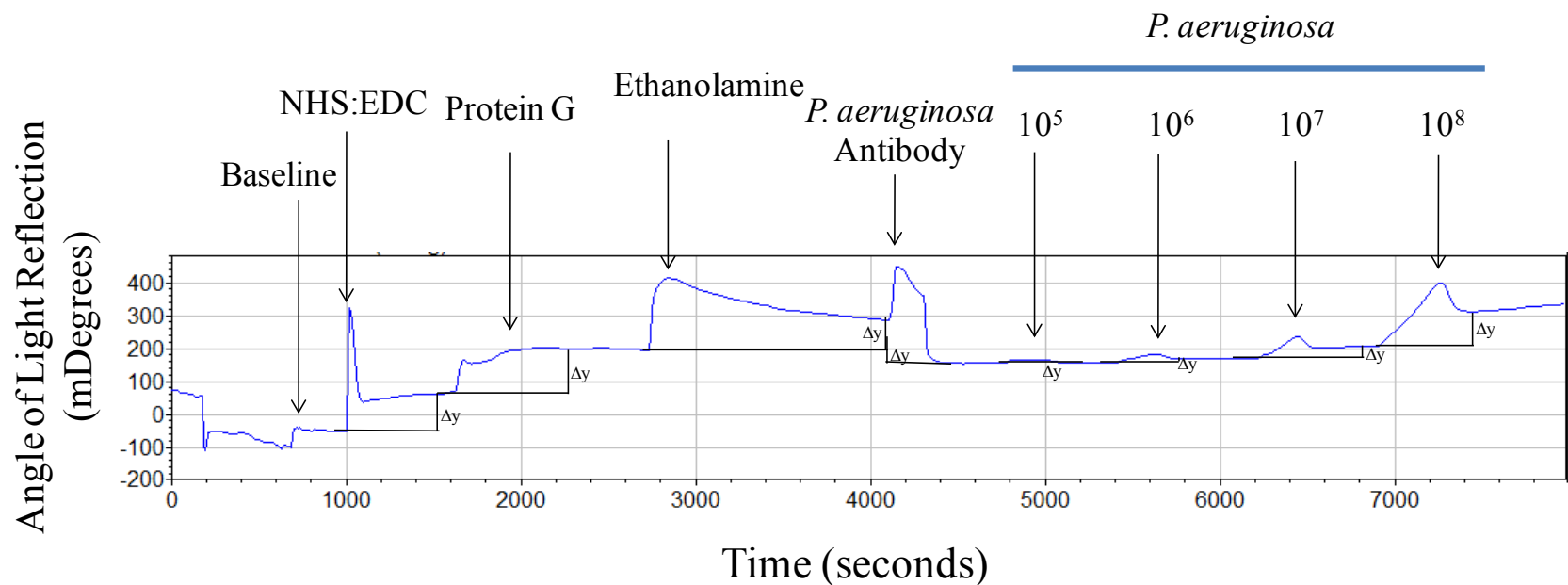
**Figure 4.7:** Detection of *S. aureus* using an immunosensor.  $\Delta y$  is the difference in angle of light reflection between post-injection and pre-injection. The sensor was self assembled by injecting NHS:EDC which allowed the Protein G to bind to the SAM layer on the gold chip. Ethanolamine deactivated any unbound ester groups. The anti-*Staphylococcus* antibodies were injected at 1800 seconds and a change in angle reflection was seen. Injected BSA was used to block free Protein G. Solutions of *S. aureus* at  $10^6$ ,  $10^7$ , and  $10^8$  cfu/mL were injected into the SPR instrument. The angles of light reflection changed at  $10^7$  and  $10^8$  cfu/mL injections.

**4.4.3 Detection of *P. aeruginosa*.** The  $10^7$  cfu/mL solution of *P. aeruginosa* increased the angle of reflection 16.15 mDeg from the pre-injection (Figure 4.8). The first injection of the  $10^8$  cfu/mL solution contained an air bubble which flowed over the sensor changing the refractive index of the sensor interfering with the plasmon waves (Figure 4.8). The next injection increased the angle by 36.34 mDeg as compared to the pre-injection measurement. The same concentration was injected one more time showing a reflection difference of 80.77 mDeg (Figure 4.8).

**4.4.4 *P. aeruginosa* immunosensor used for sensitivity tests.** An immunosensor was developed to determine if sensitivity of the SPR response could be increased in the absence of BSA. The baseline of the thiol gold coated sensor was established. The injection of NHS:EDC changed the angle of light reflection 150 mDeg (Figure 4.9). Protein G injection caused a 131.70 mDeg change in angle (Figure 4.9). The ethanolamine injection decreased the incident angle below the pre-injection baseline by 131.46 mDeg (Figure 4.9). The injection of anti-*P. aeruginosa* antibodies changed the light reflection 165.48 mDeg (Figure 4.9). The  $10^5$  cfu/mL of *P. aeruginosa* in PBS was washed over the immunosensor causing a peak in the sensograph. There was no difference in angle between the pre and post-injection. Angles of reflection for  $10^6$ ,  $10^7$ , and  $10^8$  cfu/mL were 13.14, 30.64, and 127.62 mDeg, respectively (Figure 4.9).

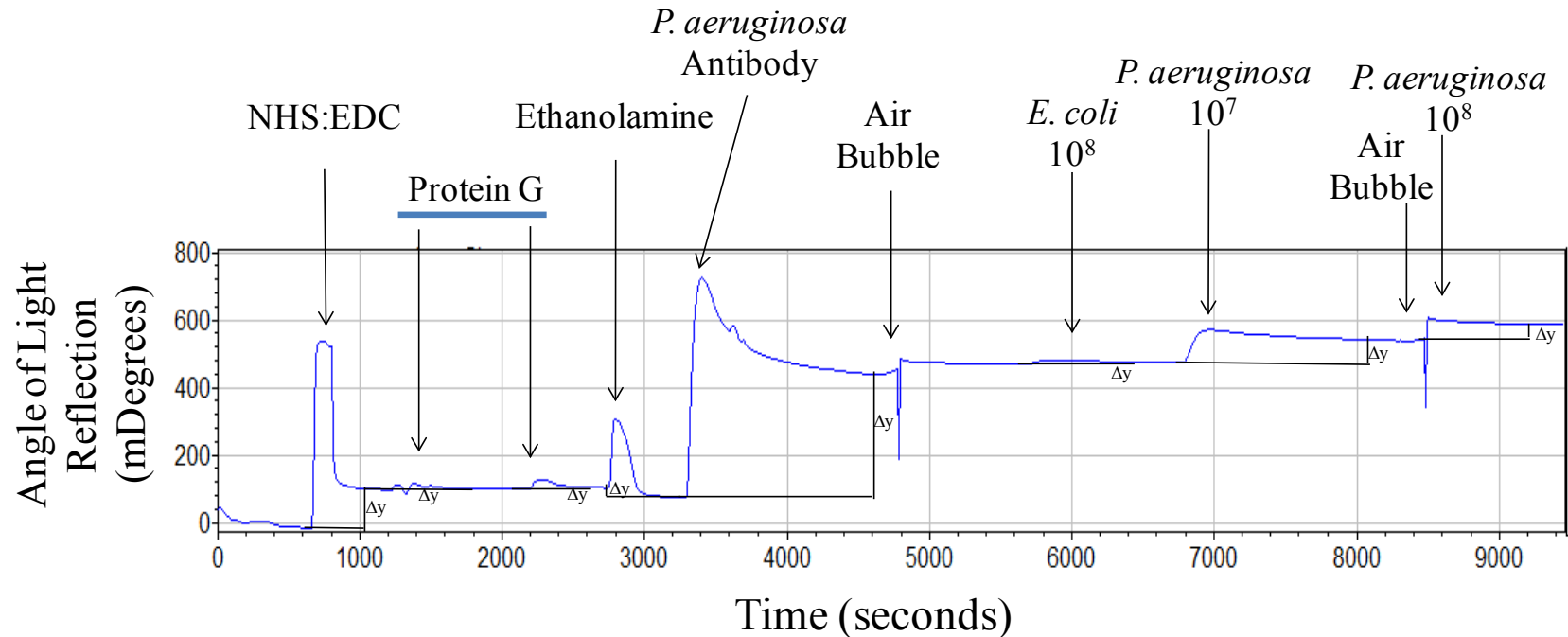


**Figure 4.8:** *P. aeruginosa* detection using an immunosensor.  $\Delta y$  is the difference in angle of light reflection between post-injection and pre-injection. The sensor was self assembled by injecting NHS:EDC which allowed the Protein G to bind to the SAM layer. Ethanolamine deactivated any unbound ester groups. The anti-*P. aeruginosa* antibodies were injected at 4300 seconds which shifted the angle of reflection. Injected BSA was used to block free Protein G. Solutions of *P. aeruginosa* at  $10^6$ ,  $10^7$ , and  $10^8$  cfu/mL were injected into the SPR instrument. The first  $10^8$  cfu/mL injection contained an air bubble which caused error in the SPR response. The solution was injected two more times. The  $10^7$  and both  $10^8$  cfu/mL solutions were detected.



**Figure 4.9:** Sensitivity of the *P. aeruginosa* immunosensor.  $\Delta y$  is the difference in angle of light reflection between post-injection and pre-injection. The sensor was assembled by injecting NHS:EDC which allowed the Protein G to bind to the SAM layer on the gold chip. Ethanolamine deactivated any unbound ester groups. The anti-*P. aeruginosa* antibodies were injected at 3300 seconds and a change in angle reflection was seen. The carrier solution contained air and the bubble passed over the sensor at approximately 4100 seconds. The  $10^5$  cfu/mL *P. aeruginosa* solution did not change the angle of light reflection. Injections of *P. aeruginosa* solutions  $10^6$ ,  $10^7$ , and  $10^8$  cfu/mL were followed by changes in angles of reflection.

**4.4.5 The effect of BSA on non-selective binding using the *P. aeruginosa* immunosensor.** BSA is added to immunosensors to block free Protein G molecules to reduce non-selective binding (Dong *et al.*, 2008; Jung *et al.*, 2006; Jung *et al.*, 2007; Lu *et al.*, 2001). The addition of BSA increases the thickness of the layers which can decrease the sensitivity of the SPR response. In addition, not all published immunosensors use BSA (Fratamico *et al.*, 1998; Waswa *et al.*, 2006). Therefore, an experiment was performed to determine if the *P. aeruginosa* immunosensor would give false positive results if BSA was not excluded (Figure 4.10). An immunosensor was designed as described earlier to include the anti-*P. aeruginosa* antibody but exclude BSA. Briefly, the baseline of the thiol gold coated sensor was established at -10 mDeg. The injection of NHS:EDC changed the angle of light reflection 139.01 mDeg (Figure 4.10). Protein G injection did not affect the light reflection. The solution was injected again and the reflection changed 18.58 mDeg (Figure 4.10). The ethanolamine injection decreased the incident angle below the pre-injection baseline by 41.81 mDeg (Figure 4.10). The anti-*P. aeruginosa* antibodies injection changed the light reflection 357.73 mDeg (Figure 4.10). The reference buffer contained an air bubble which caused a spike in the reflection. To test for non-selective binding, a solution of *E. coli*, which should not bind to anti-*P. aeruginosa* antibodies, was used. A  $10^8$  cfu/mL solution of *E. coli* was injected and changed the light reflection by 7.32 mDeg. The  $10^7$  and  $10^8$  cfu/mL *P. aeruginosa* solutions produced an angle of reflection of 103.96 and 74.26 mDeg, respectively (Figure 4.10).



**Figure 4.10:** Specificity of *P. aeruginosa* immunosensor.  $\Delta y$  is the difference in angle of light reflection between post-injection and pre-injection. The sensor was assembled by injecting NHS:EDC which allowed the Protein G to bind to the SAM layer on the gold chip. Ethanolamine deactivated any unbound ester groups. The anti-*P. aeruginosa* antibodies changed the angle of reflection. The carrier solution contained air and the bubble passed over the sensor at approximately 4800 seconds. Once the baseline was stable, a 10<sup>8</sup> cfu/mL *E. coli* solution was injected at 5700 seconds and no change in angle was seen. Angular shifts occurred for 10<sup>7</sup>, and 10<sup>8</sup> cfu/mL *P. aeruginosa* solutions.

## 4.5 Discussion

Bacteria can be identified using antibody-antigen interactions. SPR immunosensors have been shown to detect numerous bacteria based on antibody-antigen interaction. This study attempted to develop two immunosensors to detect bacteria, specifically *S. aureus* and *P. aeruginosa*. These bacteria cause infections in humans, are responsible for the bulk of nosocomial infections, and can be difficult to treat due to antibiotic resistance (Aloush *et al.*, 2006; Ashley & Brindle, 1960; Fullbrook *et al.*, 1970; Li *et al.*, 1994; Li *et al.*, 1998; Noskin *et al.*, 2005; Wertheim *et al.*, 2005; Zetola *et al.*, 2005). The spread and the disease process of these organisms can be reduced by rapid detection in patients.

**4.5.1 SPR immunosensor detected *S. aureus*.** Using layer-by-layer self assembled, a *Staphylococcus* immunosensor was made. At the beginning of the assembly, the sensor had layer of gold and a SAM of 11-mercaptoundecanoic acid molecules (Figure 4.7). NHS:EDC chemically reacted with the 11-mercaptoundecanoic allowing Protein G to bind thereby shifting the angle of reflection (Figure 4.7). The anti-*Staphylococcus* antibodies attached to the Protein G molecules fixed on the sensor causing the 100.63 mDeg change in angle (Figure 4.7). Unbound 1 % BSA was injected to block free Protein G molecules and the angle of reflection changed 187.29 mDeg (Figure 4.7).

Upon the completion of the sensor, solutions of *S. aureus* were used to determine if the sensor could detect bacteria. A change in light reflection was observed while the  $10^6$  cfu/mL solution flowed past the sensor as determined by the peak in the sensograph

(Figure 4.7). Upon completion of the injection, the reference buffer carried away any unbound molecules. The angle of reflection shifted and stabilized at the same angle of reflection that was observed before the injection. Thus, no bacteria in the  $10^6$  cfu/mL solutions were captured by the antibodies on the sensor. Increasing the concentration of bacteria ( $10^7$  and  $10^8$  cfu/mL) increased the angle of reflection proportionally (Figure 4.7). As that angular shift of light reflection indicates a change in the layer adsorbed on the gold sensor, it can be stated that the immunosensor assembled detected interactions between anti-*Staphylococcus* antibodies and *S. aureus*.

The sensor was unable to detect cells less than  $10^7$  cfu/mL. Sensitivity is a crucial characteristic of diagnostic assays. Immunosensors for other bacteria and toxins have been developed with detection limits of  $10^2$  cfu/mL (Oh *et al.*, 2004a; Oh *et al.*, 2004b; Oh *et al.*, 2005). These sensors used SAM with various chemicals, Protein G/A, and no BSA (Oh *et al.*, 2004a; Oh *et al.*, 2004b; Oh *et al.*, 2005). Optimization of the self assembly protocol may increase the sensitivity of the sensor. In addition, the effect of BSA on the sensitivity and specificity of the sensor should be examined.

**4.5.2 *P. aeruginosa* immunosensor detected solutions containing *P. aeruginosa*.** A *P. aeruginosa* immunosensor was assembled in the same manner as the *Staphylococcus* sensor (Figure 4.8). The NHS:EDC chemically changed the SAM layer causing the angle of reflection to change by 41.07 mDeg (Figure 4.8). There was no change in angle after the Protein G injection which indicated that no molecules were adsorbed on to the SAM layer. To ensure that the sensor was properly designed, the



Protein G solution was injected again and the reflection changed 17.60 mDeg suggesting that molecules did bind to the ester group on the SAM layer (Figure 4.8).

The ethanolamine injection deactivated any remaining ester groups. The angle of light reflection shifted resulting in an angle that was less than the pre-injection angle (Figure 4.8). The angular shift would indicate the molecules that were adsorbed onto the sensor before the ethanolamine injection were removed by the carrier solution containing the ethanolamine. Removal of molecules decreased the thickness of the layer adsorbed onto the gold sensor and decreasing the angle of reflection. This may be result of flow rate. The carrier buffer ran for 60 seconds between each injection to ensure that any unbound molecules were carried away from the sensor. This may have been too short of a time to ensure that unbound Protein G was removed from the surface of the sensor. Thus, the carrier buffer with ethanolamine continued to remove unbound Protein G decreasing the thickness of the layer on the gold and shifting the angle of reflection. The decreased angle of reflection may have also resulted from the activity of ethanolamine. If a portion of the SAM layer was free of Protein G and still reactive then the ethanolamine would deactivate the SAM layer, changing its shape and thus decreasing the thickness of the SAM layer. The angle of reflection would decrease as more of the SAM layer was deactivated. Both Protein G injections showed minimal changes in light reflection which suggests very few molecules attached and much of the SAM layer was left reactive which supports the ethanolamine explanation.

The angle of reflection after the ethanolamine injection is still greater than initial baseline angle. Continuing the assembly of the sensor, the anti-*P. aeruginosa* antibodies were injected. The angle of light reflection shifted by 205.38 mDeg (Figure 4.8). This

significant shift indicated that many antibodies attached to the sensor by coupling to Protein G molecules. The angular shift of 275.80 mDeg was observed upon the addition of BSA indicating free Protein G was blocked by BSA (Figure 4.8).

The self assembled immunosensor was tested using  $10^6$ ,  $10^7$ , and  $10^8$  cfu/mL of *P. aeruginosa* (Figure 4.8). There was no angular shift with the first injection which showed the sensor did not detect the specific bacteria in the  $10^6$  cfu/mL solution. There was an angular shift for the other concentrations (Figure 4.8). The change in angle was proportional to the increase in cell concentrations. The second  $10^8$  injection caused the greatest change in angle indicating that the sensor had not reached saturation with the previous injections.

Despite the unexpected angular shifts observed during the assembly process, the sensor clearly detected the presence of *P. aeruginosa* as shown by the angular shifts observed with the  $10^7$  and  $10^8$  cfu/mL injections. As with the *Staphylococcus* immunosensor, the sensor should ideally be able to detect less concentrated solutions.

**4.5.3 BSA decreases the sensitivity of the *P. aeruginosa* sensor.** An immunosensor was developed to determine if sensitivity of the SPR response could be increased if BSA was excluded. The angular shifts for the NHS:EDC, Protein G, ethanolamine, and anti-*P. aeruginosa* antibody injections were similar to the previous *P. aeruginosa* sensor (Figure 4.9). The assembled sensor was tested with various concentrations of *P. aeruginosa* solutions. The lowest limit of detection for an immunosensor was  $10^6$  cfu/mL for the *P. aeruginosa* sensor assembled without BSA. The presence of BSA lowered the threshold of detection to  $10^7$  cfu/mL for the other two

chips. This suggests that BSA does influence the sensitivity of the *P. aeruginosa* biosensor. Even without the BSA, the sensitivity of the sensor is much lower than published chips (Oh *et al.*, 2004a; Oh *et al.*, 2004b; Oh *et al.*, 2005).

**4.5.4 Non-BSA chips shows selective binding of bacterial cells with cognate antibodies.** Accuracy and sensitivity are both important characteristics of a diagnostic tool. Therefore, an experiment was performed to determine the selectivity of a sensor (Figure 4.10). To test for non-selective binding, a solution of  $10^8$  cfu/mL *E. coli* was injected into the instrument. The angle of light reflection shifted by 7.32 mDeg. The shift indicated that high concentrations of *E. coli* non-specifically adsorbed to the sensor. The  $10^7$  and  $10^8$  cfu/mL *P. aeruginosa* solutions produced an angle of reflection of 103.96 and 74.26 mDeg, respectively (Figure 4.10). These results proved that the sensor was assembled correctly, functional, and specific. The SPR response for the *E. coli* was insignificant in comparison to the response for the *P. aeruginosa* solution at the same concentration. A similar insignificant angular shift was observed with *E. coli* on a *Salmonella* specific sensor (Waswa *et al.*, 2006). Future experiments need to be performed to determine if BSA is needed to prevent non-selective binding of other bacteria.

**4.5.5 Optimization of immunosensors.** In order to use SPR immunosensors as a diagnostic tool, the accuracy, specificity, and sensitivity of the sensors need to be improved. SPR response can be affected by the instrumentation, environment and by the choice in chemicals used in the layer-by-layer self assembly process.

The instrument is very sensitive to the environment and should be placed in a temperature controlled area (Dudak & Boyaci, 2009). Lack of temperature control causes an increase in standard deviation between readings (Dutra *et al.*, 2007; He *et al.*, 2007; Matsumoto *et al.*, 2005; Naimushin *et al.*, 2003). Additionally, the instrument uses microfluidics to transfer chemicals to the gold sensor chip (Biosensing Instrument, 2006). The flow rate of the reference buffer and the injections can be modified to increase the adsorption of molecules to the sensor (Dudak & Boyaci, 2009). Ideally, the flow rate will enable bacteria to bind to the antibodies before the reference buffer washes them away. Though if the rate is too slow, the antigens will form aggregates and not be able to bind to the antibodies. On the other hand if the rate is too fast, then the mass transfer limitations will be minimized and antigens will flow past the antibodies unable to bind (Dudak & Boyaci, 2009; Skottrup *et al.*, 2008). The flow rate used in this study was an intermediate rate (10  $\mu\text{L}/\text{min}$  for reagents and 5  $\mu\text{L}/\text{min}$  for antibodies and bacteria); other published work used anywhere from 3 to 100  $\mu\text{L}/\text{min}$  (Fratamico *et al.*, 1998; Oh *et al.*, 2004a; Oh *et al.*, 2004b; Oh *et al.*, 2005; Subramanian & Irudayaraj, 2006; Waswa *et al.*, 2006). In addition, the reference buffer must run in between injections. Enough time needs to be allowed to ensure that excess or unbound molecules are washed away from the sensor. Sixty seconds was used in this study which may have not been enough time to clean the sensor. This factor is rarely detailed in published material. Some of the published studies do not use microfluidics for assembly; rather the chip is immersed in Protein G or the antibody for hours to days and then rinsed off with a buffer (Oh *et al.*, 2004a; Oh *et al.*, 2004b; Oh *et al.*, 2005). Longer incubation times with more effective rinses may increase the efficiency of the molecules binding to the sensor.

Solutions used in the SPR sensor assembly also affect the SPR response. First, the ionic solution and pH of the buffers may need to be optimized. SPR sensitivity is affected by ionic strength as it directly affects the supramolecular interactions between antibody and the antigen (Srilatha & Murthy, 2006). Besides PBS, HEPES buffer saline has been used (Fratamico *et al.*, 1998; Waswa *et al.*, 2006). The composition of the PBS may have not provided the optimum ionic strength that would lead to the appropriate binding affinity constants. The pH of the PBS solution also needs to be optimized for molecule interaction, especially the binding of antibodies to Protein G (Keeler *et al.*, 2007; Srilatha & Murthy, 2006; Waswa *et al.*, 2006). One study has shown that more antibodies bind to Protein G if a pH 5 is used while pH 7 is better for a sensor without Protein G (Fratamico *et al.*, 1998). In this work, pH 7 was used and that may have affected the angular shift of light reflection in response to the antibody (Fratamico *et al.*, 1998).

The properties of the sensor can also be improved by using other methods to fix the antibodies to the sensor, such as coating the sensor with streptavidin and labeling the antibody with biotin (Dudak & Boyaci, 2007). This method has been used to detect *E. coli* with a sensitivity of  $10^2$  cfu/mL (Dudak & Boyaci, 2007). This conjugation system may increase the detection limit for the *P. aeruginosa* and *Staphylococcus* immunosensors.

The assembly process can also be optimized by changing the chemical used to form SAM. There are some studies that conclude mixed SAMs (using two different alkanethiols) can increase the SPR response (Taylor *et al.*, 2006; Tsai & Li, 2009).

Once all these optimizations are determined, a sensor can be developed that is accurate, reproducible, and sensitive. At this time, the SPR immunosensors developed in this study lack the sensitivity to be used to detect *S. aureus* and *P. aeruginosa* in a sample.

#### **4.6 Acknowledgements**

I would like to thank Dr. Chenzhong Li from FIU Department of Biomedical Engineering for allowing me to work on this project in his laboratory. I am grateful to Ali Mirwah for teaching me how to use the instrument and Chang Liu for supplying me with reagents and access to the laboratory.

## **Chapter 5**

### **Summary and General Discussion**

## 5.1 Overview

Cystic fibrosis is a genetic disorder that affects approximately 30,000 children and adults in the United States (Patient-registry, 2008). In addition, one out of 31 people are have a mutation in one copy of the *cftr* gene encoding the cystic fibrosis transmembrane conductance regulator protein (Patient-registry, 2008). If a person inherits two copies of the gene each carrying a mutation then the individual can aspect to have difficulties with their digestive, reproduction, and respiratory systems (Patient-registry, 2008). During the 1950's few of these patients would live to attend elementary school. Due to advancements in science and medicine, the median age of survival has risen to 37.4 years in 2008 (Patient-registry, 2008). The leading cause of morbidity and mortality in CF patients is lung disease which stems from chronic microbial infection.

The CF lung contains a dehydrated layer of mucus which reduces the ability of the cilia to beat the microorganisms out of the lungs (Andersen, 1938). The limited mucosal clearance, the inability of the host immune system to eradicate the infection, and development of antibiotic resistance leads to chronic infections that destroy the lungs leaving the patient in need of a lung transplant (Patient-registry, 2008). In 2008, 158 people received a transplant to extend their lives (Patient-registry, 2008). Lung transplants are risky and the supply of lungs is limited (Adler *et al.*, 2009). The most likely way to increase CF life expectancy is to decrease bacterial and fungal lung load in the lungs. In order to treat and reduce infections, one must identify the causative agent and understand the role the organism plays in the lung.

This dissertation served to further understand the CF lung infection. First, the composition of the bacterial community in the CF lung and its dynamics over time were



determined using LH-PCR (Chapter 2). Second, the members of the community both bacteria and fungi were identified using 454 sequencing (Chapter 3). In addition, correlations between the diseased and healthy bacterial communities were identified (Chapter 3). Thirdly, a bacterial diagnostic tool targeted towards CF pathogens was developed using SPR (Chapter 4).

## **5.2 Bacterial community profiling of the CF lung**

Based on culturing methods, it was thought only a few bacteria including *Staphylococcus aureus*, *Haemophilus influenzae* and *Pseudomonas aeruginosa* infected the CF lung (Whittier, 2001). *S. aureus*, the predominant pathogen in children, is usually succeeded by *H. influenzae* later in childhood and *P. aeruginosa* becomes the predominant pathogen during adolescence, reaching a prevalence rate of 80% in adults (Beringer & Appleman, 2000; Patient-registry, 2008). With the advent of molecular tools, it became known that the CF lung community is larger and more diverse than culturing results had show (Armougom *et al.*, 2009; Bittar *et al.*, 2008; Harris *et al.*, 2007; Rogers *et al.*, 2003). One of the earliest studies analyzed CF patient sputa and bronchoscopy samples with T-RFLP (Rogers *et al.*, 2003). The common pathogens, *P. aeruginosa*, *Burkholderia cenocepacia*, *Staphylococcus aureus*, and *H. influenzae* were identified in the CF samples (Rogers *et al.*, 2003). In addition, *Achromobacter (Alcaligenes) xylooxidans*, *Rhodotorula mucilaginosa*, *Abiotrophia* spp, *Bacteroides gracilis*, *Eubacterium brachy*, *Mycobacterium mucilaginosus*, *Mycoplasma salivarium*, *Porphyromonas salivae*, *Ralstonia* spp, *Staphylococcus hominis*, *Streptococcus anginosus*, *Treponema vincentii*, *Veillonella* spp, *Burkholderia gladioli*,

*Stenotrophomonas maltophilia* and *Pandoraea* atypical were also shown to be members of this polymicrobial infection (Rogers *et al.*, 2003). Large scale Sanger sequencing projects continued to resolve the bacterial community and revealed the presence of organisms that were not previously associated with the CF lung. *Moxarella cattharalis*, *Granulicatella elegans*, *Inquilinus limosus*, *Granulicatella adiacens*, *Rothia mucilaginosa*, *Dialister pneumosintes*, *Dolosigranulum pigrum*, *Lysobacter*, *Coxiellaceae* were identified from either bronchoalveolar or sputum samples from CF patients (Bittar *et al.*, 2008; Harris *et al.*, 2007; Rogers *et al.*, 2003). These molecular studies indicate that the total ecology of the CF lung niche remains an open question. This research continued to identify the bacterial community in the CF lung by examining snap shots of the overall bacterial community.

Though the composition of the lung niche is being teased apart, a larger question still needs to be addressed. The interactions between these organisms, the host, and the disease process are not well understood. There are some correlations with patient age and infection. As noted earlier, *S. aureus* is most often identified in children and *P. aeruginosa* is associated with adulthood (Beringer & Appleman, 2000; Patient-registry, 2008). Yet, the molecular studies have identified *S. aureus* in adults (Bittar *et al.*, 2008). Culturing results also vary throughout a patient's life indicating that some bacteria are transient while others are constantly present in the community regardless of antibiotic therapy (Beringer & Appleman, 2000; Patient-registry, 2008; Rogers *et al.*, 2010; Smith *et al.*, 2006). Therefore it can be hypothesized that the bacterial community in the CF lung is dynamic. This study used LH-PCR to obtain a snap shot image of the bacterial community in a CF sample at the time of collection. The presumptive identities of the

organisms present in the sample were determined. The bacterial profiles were also used in an attempt to understand how the community as a whole can be affected by the host and other community members over a period of time.

In this study, 19 CF sputa samples from adult patients in Miami, FL were analyzed for the presence of bacteria using LH-PCR (Chapter 2; (Doud *et al.*, 2009)). The profiles generated demonstrate the complexity and uniqueness of the bacterial CF lung community (Chapter 2, Supplemental Figure 2.6 and 2.7). Each peak in a profile represents at least one bacterium. The V1\_V2 profiles for 12 of the CF samples contained multiple peaks and were different from one another. The most complex sample contained eight amplicons (Chapter 2, Figure 2.1A). Seven of the samples had identical V1\_V2 profiles containing one peak at 342 bp (Chapter 2, Figure 2.1B). In contrast, all V1 profiles contained multiple amplicons (10-30) and each profile was different demonstrating that lung communities are unique to a patient (Chapter 2, Figure 2.7 and Table 2.1). The profiles from the V1\_V2 region were more similar than the V1 LH profiles. Thus, the V1 region had a higher power of discrimination than the V1\_V2 region. The LH profiles provided snap shot pictures of the bacterial communities further highlighting the diversity and complexity of the CF lung infections.

The drawback to LH-PCR is the lack of species identification. It was first thought that organisms within in one genus would have the same length for a particular variable region. A bioinformatic analysis using the newly designed AmpliQué highlighted the presence of extreme length variability in the V1\_V2 region of the 16S rRNA gene (Gonzalez *et al.*, 2009). Not only do strains within a species vary in amplicon length, multiple genera can produce the same amplicon length. Identification of bacteria in an

LH profile based on one amplicon length is nearly impossible. Identification using LH-PCR may be attainable if multiple regions of the 16S rRNA gene are used to profile the same sample. Cloning and sequencing of the amplicons is the only way to identify an organism.

Identification using other community profiling techniques can also be challenging. T-RFLP has been used to analyze CF samples (Rogers *et al.*, 2003; Rogers *et al.*, 2004; Rogers *et al.*, 2006; Rogers *et al.*, 2009; Rogers *et al.*, 2010). This technique produces the same variable length fragments as LH-PCR and then the amplicons are cut with a restriction enzyme. This additional step increases the discriminatory power of the technique (Rogers *et al.*, 2003). Bacteria identification is still presumptive as that the length of the amplicon generated from the restriction digestion can still be the same for multiple organisms. Identification of bacteria is confirmed by cloning and sequencing (Rogers *et al.*, 2003; Rogers *et al.*, 2004). Identity of bacteria in a community is not easily determined with T-RFLP or LH-PCR but the techniques can easily provide an estimate of bacterial diversity associated with the sample (Doud *et al.*, 2009; Rogers *et al.*, 2003; Rogers *et al.*, 2004). However, community profiling can provide information on community dynamics.

Microbial community dynamics have traditionally been studied in the field of ecology. Many environmental factors such as nutrients, temperature, fertilization practices, and metal pollutants have been found to have an effect on the bacterial community in soil (Entry *et al.*, 2008; Gunapala & Scow, 1998; Mills *et al.*, 2003; Moreno *et al.*, 2006; Vinas *et al.*, 2005) and in marine and fresh water (Boer *et al.*, 2009; Carrero-Colon *et al.*, 2006; Crump *et al.*, 2003; Simek *et al.*, 2001).

LH-PCR and T-RFLP are quite commonly used in ecology to observe and understand the relationship between members of the community and external factors. The microbial ecology methodologies can be applied to microbial communities associated with infections.

This study which used LH-PCR is the first to observe bacterial community dynamics in the CF lung. One patient produced expectorated sputa multiple times in one day and then again one year and two years later. Minor variations in the LH profile were seen in the samples taken from one day at all time points (Chapter 2, Figure 2.5). In contrast, the bacterial profiles generated from samples at different time points (M1, M3, M16, and M17) were quite different from one another and from those samples given on the same day in M22. Hence, there is a temporal variability in the CF lung bacterial community indicating the community is not static but rather dynamic (Chapter 2, Figure 2.5).

Other human associated bacterial communities have been found to be constant rather than dynamic. Pyrosequencing was used to determine the composition of bacteria communities associated with various spots on the healthy human body (Costello *et al.*, 2009). Samples were collected on two consecutive days and then a month later for two additional consecutive days. The researchers concluded that each habitat has its own characteristic community and the communities were relatively stable over time (Costello *et al.*, 2009). The healthy human may contain a more stabilized microbiota because the human is not influenced by pathogenic microbes and therapeutic drugs (Hilty *et al.*, 2010)

The instability of the CF lung eubacteriome could be attributed to antimicrobial therapy. To address this hypothesis, LH profiles were generated over a six month period from two individuals undergoing antibiotic treatment (Chapter, Table 2.2). The variation in profiles from the medicated patients was much more dramatic than the changes in the communities from the stable patient. In addition, the most dramatic changes (drastic changes in abundance and disappearance of peaks) correlated with changes in the patient's antibiotic treatment (Chapter 2; Table 2.2). Therefore, antibiotics are a potential driver of the CF bacterial community. CF individuals from infancy are under periods of antibiotic treatment. The continuous use of antibiotics has most likely disrupted the lung flora. Antibiotic treatment has already been established as a driver of bacterial communities present in the mouse gastrointestinal track, the trachea of intubated people, and chronic wounds (Antonopoulos *et al.*, 2009; Flanagan *et al.*, 2007; Price *et al.*, 2009).

This study is the first to track the bacterial community in the CF lung over time. Although the study used a small sample size over a relatively short period of time, changes in the presence and abundance of amplicons in the LH profiles could be seen. To accurately identify the drivers of the community, a more comprehensive long term study needs to be performed. Armed with the patient's antibiotic information and long term sampling, it should be possible to determine the effectiveness of a drug against pathogens in the CF lung. This could eventually lead to more effective treatments. In addition, other drivers of the bacterial community need to be defined. The composition of the bacterial community may be influenced by the patient's age, ethnicity, *cftr* mutation or daily environment.

**Future studies.** To understand the community dynamics of the CF lung, the core biome or baseline community of the healthy human mouth and nasopharyngeal regions must be determined. Several recent studies have determined the microbiome in the oral cavity and lungs from healthy individuals (Ghannoum *et al.*, 2010; Hilty *et al.*, 2010; Keijsers *et al.*, 2008; Lazarevic *et al.*, 2009; Zaura *et al.*, 2009). This information should be compared to CF data to discern the presence of unique organisms and their role in CF pulmonary failure. Healthy samples, in addition to CF samples, can also be collected and analyzed as controls during a long term study. It would be advantageous to use age, gender, and race matched samples. Samples from healthy individuals that live with the CF patient could also be used to look at the role the environment has on the community.

A much large sample size is needed to corroborate the preliminary results of this study. Samples from CF patients from both genders, multiple races/ethnicities, different ages, and *cftr* mutations are needed. In addition, the health of the patient (weight, lung function, disease progression) should be recorded. Samples should be collected quite often, at least monthly though more often would be preferred for a long period of time (a year or more).

Ideally, the analysis of sputa or BAL samples using a community profiling technique such as LH-PCR should be performed in conjunction with a sequencing technique, such as 454 pyrosequencing. LH-PCR would provide bacterial profiles of the samples. Sample profiles that show change from one time point to the next can then be further analyzed using sequencing. The identity of the organisms and their relative abundance from each community would be compared with other time periods. Using

multivariate statistics, changes in community can be identified and correlated with patient characteristics to identify the drivers of the CF community.

### **5.3 High throughput 454 sequencing identifies bacteria and fungi associated with the CF lung**

Community profiling and Sanger sequencing projects have been used to study the polymicrobial nature of the CF lung (Armougom *et al.*, 2009; Bittar *et al.*, 2008; Harris *et al.*, 2007). These studies have shown that the bacterial communities are diverse and unique to each patient. The list of bacteria associated with the CF lung as noted in the previous section is getting longer with every study.

The CF lung can also be inhabited by fungi. Species-specific PCR and culturing have identified a few fungi, such as *Aspergillus* and *Candida* (Bakare *et al.*, 2003; Cimon *et al.*, 2001; Mearns *et al.*, 1967). There have been reports of *Scedosporium apiospermum*, *Exophiala sp.*, and *Wangiella dermatitidis* being cultured from CF sputa (Defontaine *et al.*, 2002; Dixon & Polak-Wyss, 1991; Hennequin *et al.*, 1997; Horr  *et al.*, 2004; Horr  *et al.*, 2009; Rogers *et al.*, 2003; Williamson *et al.*, 2001). Recent statistics have shown that there is an increase in the prevalence of fungi being detected in the CF patient which may be a consequence of aggressive antibacterial therapy (Nagano *et al.*, 2007). Yet, the diversity of the fungal community has not been previously analyzed with community profiling techniques or sequencing. A more comprehensive study is needed to determine if the fungal community is as diverse and unique as the bacterial community.



To resolve the microbiome, high throughput 454 sequencing was used to analyze the sputa from 19 south Floridian CF patients. This study is the first to use this technology to analyze the bacterial and fungal communities simultaneously in the same samples. Multivariate statistics was used to identify similarities between the communities and the relationship of the host's gender and ethnicity with the microbial community. The CF microbial communities were also compared to the communities sequenced from healthy oral rinses.

The unique polymicrobial nature of the CF lung was shown by the data. Sequencing results identified numerous bacteria in the CF lung, some of which were not previously identified in other metagenomic studies (Armougom *et al.*, 2009; Bittar *et al.*, 2008; Harris *et al.*, 2007; Rogers *et al.*, 2003; Tunney *et al.*, 2008). Thirty-four taxa were identified in the 18 samples. All samples contained numerous organisms with the average of  $9.6 \pm 3.3$  taxa present. *Pseudomonas* or *Streptococcus* were present in 16/18 samples. *Chryseomonas* was present in 15 samples and was the most abundant organism in 14 samples. Fourteen samples contained sequences from *Micrococcineae* and *Prevotella*. *Chryseomonas*, *Flavimonas*, and *Moryella* may potentially be novel bacteria in CF sputa.

Many genera that contain common CF pathogens were sequenced including *Pseudomonas*, *Haemophilus*, *Burkholderia*, *Staphylococcus*, *Achromobacter*, and *Stenotrophomonas* (Chapter 3, Table 3.2). In addition, many sequences corresponded to organisms that are known to be present in the oral cavity. Some of these organisms are commensal while others are known to cause infections. These oral sequences in the sputa may be contributed by oral bacteria that come in contact with the sputa during sampling

process or from bacteria that colonize the throat and lungs. The role of oral associated organisms in the CF lung is unknown but it can be hypothesized that they play a role in lung disease progression (Bittar *et al.*, 2008; Rogers *et al.*, 2006).

Some sequences were identified as organisms not cultured from CF samples but identified through sequencing (Bittar *et al.*, 2008; Harris *et al.*, 2007). This analysis identified sequences belonging to three genera that have never been cultured or sequenced from CF sputa or BAL (Chapter 3, Table 3.2). Sequence hits for *Chryseomonas* and *Flavimonas* were obtained for some MTPS reads. Both of these genera have been reclassified as *Pseudomonas*. The species associated with those genera can be pathogenic and have not been associated with the CF lung. Obtaining a full length 16S rRNA sequence is needed to verify the identity of these sequences. Sequences identified as *Moryella* were present. Little is known about the anaerobe and this is the first time it has been sequenced from CF samples. Overall, the 454 sequencing of the south Florida CF patients showed a more diverse eubacteriome in the CF sputa than other studies (Armougom *et al.*, 2009; Bittar *et al.*, 2008; Harris *et al.*, 2007; Rogers *et al.*, 2003; Rogers *et al.*, 2004; Rogers *et al.*, 2006; Rogers *et al.*, 2009; Rogers *et al.*, 2010).

Many samples shared multiple taxa but statistically the communities were unique to each patient (Chapter 3, Figure 3.2). Similarities between communities could not be attributed to patient's gender or ethnicity (Chapter 3, Figure 3.2). A much larger sample size is needed to determine if gender and ethnicity affect the composition of the bacterial community.

The bacterial communities in the CF sputa were compared to communities from healthy oral rinses. PCO clustered healthy bacterial communities closely together and

separated from CF bacterial communities (Chapter 3, Figure 3.3). Statistically, the overall bacterial community in the CF lung was quite different than the healthy oral samples. Twelve taxa were present in both datasets most of which were associated with the oral cavity (Chapter 3, Table 3.4). *Achromobacter* was present in one healthy sample which may indicate the person was not actually healthy. *Burkholderia* was present in both groups; without the species, it cannot be determined if the pathogenic organism was present in the healthy sample. Many of the common CF pathogens including *Pseudomonas*, *Staphylococcus*, and *Stenotrophomonas* were only sequenced in the CF samples. The comparison of the communities indicated that some microbes sequenced in the CF samples may originate from the mouth.

This study also used sequencing to identify the fungal taxa associated with the CF lung. Previous to this research, the fungal community was thought to consist of a handful of fungi, mainly species of *Aspergillus* and *Candida* (Bakare *et al.*, 2003; Cimon *et al.*, 2001; Mearns *et al.*, 1967). In contrast, 454 sequencing identified 33 taxa from 18 CF samples (Chapter 3, Figure 3.4 and Table 3.5 and Supplemental Table 3.8). The samples had an average of  $26 \pm 3$  taxa. Ten genera including *Aspergillus* were present in all samples. In regards to abundance, the most sequenced genus in 17 samples was *Pichia* with sequences from *P. anomala*, *P. membranaefaciens*, and *Pichia* sp. The second most abundant genera in all 17 samples was *Rhizoctonia* (Chapter 3, Table 3.5).

The opportunistic pathogen *Candida albicans* was present in eight samples. In addition, the pathogens *Aspergillus terreus*, *Aspergillus ustus*, *Cryptococcus neoformans*, *Exophiala*, and *Wangiella dermatitidis* were present in many samples (Chapter 3, Table 3.5 and Supplemental Table 3.8). None of these patients were being treated for a fungal

infection yet the most abundant sequence in 17 samples was *Pichia* (*P. sp.*, *P. membranefaciens*, and *P. anomala*). Attempts should be made to corroborate the presence of the pathogenic fungi.

Of the 33 genera present, 26 genera were environmentally associated fungi. Their associated low abundances indicated the fungi were not colonizing the lung but rather transient members of the community originating from consumed food and inhaled air.

Comparisons of the mycobiome established that each patient had a unique fungal community (Chapter 3, Table 3.5). PCO statistically showed that gender and/or ethnicity did not contribute to the likelihood of organisms being shared between patients (Chapter 3, Figure 3.5). A correlation may be present when a larger sample size is analyzed.

The mycobiomes in the CF sputa were compared to the fungal communities identified in healthy oral rinses (Chapter 3, Figure 3.6). Based on the presence of genera, CF mycobiomes were clustered together and apart from healthy mycobiomes. The two datasets shared 11 genera but at the species level only five organisms were in common. Four organisms were associated with the environment. The remaining shared fungus was *Candida albicans*, an opportunistic pathogen that is commonly found in the mouth. Both groups of samples contained numerous fungi present in the environment. These fungi varied between the datasets which may be a result of the individuals living in two very different areas within the US.

The CF sputa eubacteriome was established to be diverse containing pathogens, oral microbes, and environmental fungi. The role of the organism in the progression of lung disease is unknown for the majority of the identified organisms. The core biome of the oral cavity, nasal regions, throat, and bronchial tree should be determined to establish

which microbes are transient, beneficial to the host, or truly associated with the CF lung infection. Deep sequencing may potentially determine the core biome but the significance of the other organisms not associated with the healthy community cannot be determined by sequencing alone.

**Future studies.** Patients who share the same *cfr* mutation can have different predominant pathogens in their microbiomes. In addition, patients with the same pathogen can have completely different disease outcome. These observations led researchers to study the CF microbiome. In this study, the members of the microbiome were identified using deep sequencing. These results have left more questions than answers. First, is the assigned identity correct? Secondly, are the sequences from metabolically active organisms? Thirdly, what role does the community member play in the community? All future CF community studies should address these three questions.

To verify the identity of the organisms, the same CF samples can be sequenced using multiple genetic markers. Using a bioinformatic program, the sequences can be joined to make longer contigs which are then used for identification. Longer contigs will be more specific and thus organisms are more accurately identified. If technology is improved, longer read lengths of one marker will be possible and multiple markers will not be necessary.

Culturing the sequenced organism from the sample will also confirm its presence. For this method, it must be assumed that the identification obtained from a short sequencing read is correct. Using selective media and growth conditions, the sequenced organism may be cultured. The identities of the pure isolates can be verified by

sequencing the entire 16S rRNA gene for bacteria, 18S rRNA gene or the ITS regions located between the rRNA genes for fungi. If the organism is cultured then it is obviously metabolically active in the lung. Using T-RFLP, and sequencing, *Streptococcus milleri* group was detected and subsequently cultured from CF samples (Sibley *et al.*, 2008; Sibley *et al.*, 2010). This species is not readily identified using routine clinical culturing protocols (Sibley *et al.*, 2010). Creating growth media to isolate organisms identified by sequencing is laborious and will not always work for all microbes.

A non-culturing molecular approach can also be used to determine if the sequence bacteria is metabolically active. The active bacteria can be identified by analyzing the meta-transcriptome (Leininger *et al.*, 2006; Urich *et al.*, 2008). In this method, the metagenomic RNA is isolated from the CF samples. The total RNA is reversely transcribed into cDNA and then directly pyrosequenced. The resulting output is large numbers of cDNA that are rRNA-tagged and mRNA-tagged sequences which provide the sequence-based transcriptome of the community. Using bioinformatic programs like MEGAN, the sequences can be binned, aligned and identified (Leininger *et al.*, 2006; Urich *et al.*, 2008). By analyzing the meta-transcriptome, the metabolically active bacteria in the CF lung can be identified.

Determining the role of the organism in the community will be the most challenging issue to address. Traditionally, Koch's postulates must be fulfilled to state that an organism causes a disease (Koch, 1884). Two requirements are that organism causing disease be isolated and grown in culture and that the pure isolated organism causes disease when introduced to a healthy person. First, it will be difficult to

impossible to culture all microbes in the CF lung (Staley & Konopka, 1985). In addition, it is more likely that a combination of microbes increases lung disease in CF rather than one specific organism. Furthermore, CF lung disease progression is also affected by the host's immune system (Makam *et al.*, 2009). The complexity of these relationships makes future studies on dynamics challenging.

Bacterial interactions can be studied by growing combinations of organisms together in media or in artificial sputum (Sriramulu *et al.*, 2005). Competition studies can determine if bacteria can live in the same niche or if one bacteria flourishes in the community. In addition, the growth media can be analyzed for bacterial secreted compounds that may be promote or decrease growth of another organism. These experiments should start with only two organisms and then other CF organisms should be added in succession. However, the bacterial interactions monitored in these conditions may not replicate the dynamics that occur in the complex CF lung.

Recreating the microbial community in a model system will provide information on how the microbial community interacts with the host. A porcine and mouse lung models have been developed that can be used to study lung infections (Hoffmann *et al.*, 2005; Rogers *et al.*, 2008). These animal models can be infected with various combinations of microbes and the affect on the lung can be studied. In addition, BAL samples from the lungs can be analyzed to determine the persistency of a microbe that was colonized in the lung at the beginning of the infection. These studies will provide a glimpse into the human CF lung.

#### **5.4 SPR as a bacterial detection tool**

During a routine or a sick visit, a clinician will take a sputum sample from a CF patient and send it to the lab for analysis. The sample is cultured on various media and based on a phenotypic and biochemical assays, identification is made (Murray & Baron, 2007). Empirical treatment is initiated before microbial identification, based on the site of infection and if the isolate is Gram positive or Gram negative. Within a day or two, the clinician is notified of the laboratory result and adjusts treatment. For limited microbes, identifications can also be obtained using methods based on antibody-antigen interactions, such as the ELISA (Engvall & Perlmann, 1971; Pedersen *et al.*, 1987; Voller *et al.*, 1978). Molecular tools are favored for rapid detection of organisms. Many of these methods were discussed in Chapter 1. Diagnostic tools for microbial identification are constantly being upgraded or developed anew to achieve faster, cheaper, and more accurate identifications. This study looked at the feasibility of using surface plasmon resonance (SPR) as a bacterial diagnostic tool.

SPR is a phenomenon in which molecules between the interface of two media vibrate (Wood, 1902, 1912). This phenomenon can be used to determine if an antigen binds to an antibody (Kretschmann & Raether, 1968; Otto, 1968). Using a SPR instrument, light is passed through a glass chip thinly coated with gold (Chapter 4, Figure 4.1). A specific frequency of light passes into the gold creating an evanescent wave. The wave excites the plasmons causing them to resonate and form a plasmon wave which travels along the gold film interacting with the media in contact with the gold. The remaining frequencies of light are reflected at a specific angle. The angle of the light



reflection, which can be measured, shifts as molecules adsorb onto the surface of the gold (Kretschmann & Raether, 1968; Schasfoort & McWhirter, 2008).

SPR immunosensor detects interactions between antibodies and other molecules such as bacteria (Geddes *et al.*, 1994; Mayo & Hallock, 1989). Essentially, SPR immunosensors detect antigens that bind to antibodies on the gold sensor by measuring the angle of light reflection (Kretschmann & Raether, 1968). SPR immunosensors have been used to identify *Escherichia coli* O157:H7 (Fratamico *et al.*, 1998; Meeusen *et al.*, 2005; Oh *et al.*, 2003; Su & Li, 2005; Subramanian *et al.*, 2006a; Taylor *et al.*, 2005), *Vibrio cholera* (Jyoung *et al.*, 2006), *Salmonella enteritis* (Bokken *et al.*, 2003), *Salmonella typhimurium* (Oh *et al.*, 2004a), *Salmonella paratyphi* (Oh *et al.*, 2004b), *Listeria monocytogenes* (Hearty *et al.*, 2006; Koubova *et al.*, 2001; Leonard *et al.*, 2005), *Yersinia enterocolitica* (Oh *et al.*, 2005), and *Campylobacter jejuni* (Taylor *et al.*, 2006).

This research study developed SPR immunosensors to detect *Staphylococcus* and *Pseudomonas aeruginosa* using the self assembled layer-by-layer approach. The gold sensor chip was functionalized with SAM and then activated to allow binding of Protein G molecules which enables either *Pseudomonas aeruginosa* or *Staphylococcus* antibodies to bind to the sensor in the proper orientation (Chapter, 4, Figure 4.7). The sensitivity and specificity of the sensors were determined using various bacterial solutions.

The *Staphylococcus* sensor was able to detect  $10^7$  and  $10^8$  cfu/mL of *Staphylococcus aureus* (Chapter 4, Figure 4.7). A *P. aeruginosa* immunosensor detected  $10^7$  and  $10^8$  cfu/mL of *P. aeruginosa* (Chapter 4, Figure 4.8).

SPR immunosensors have been capable of detecting bacteria down to the  $10^2$  cfu/mL (Oh *et al.*, 2004a; Oh *et al.*, 2004b; Oh *et al.*, 2005). Attempts were made to

improve the sensitivity of the *P. aeruginosa* sensor by excluding BSA in the assembly process (Chapter 4, Figure 4.9). This modification enabled the sensor to detect  $10^6$  cfu/mL. Therefore, the sensor without BSA was more sensitive. BSA has been shown to decrease the sensitivity of other immunosensors (Dong *et al.*, 2008; Jung *et al.*, 2006; Jung *et al.*, 2007; Lu *et al.*, 2001).

The absence of BSA may increase false positives for the sensor. Therefore, an experiment was performed to determine the selectivity of a *P. aeruginosa* sensor made without BSA (Chapter 4, Figure 4.10). The sensor detected a minimal angular shift for a  $10^8$  cfu/mL *E. coli* solution demonstrating the specificity of the sensor.

In summary, four immunosensors were successfully assembled. These sensors could detect either *Staphylococcus* or *P. aeruginosa* cells in a solution. The sensitivity and specificity of the sensors were not ideal and not as good as other published sensors. Optimization of the self-assembling process may lead to sensors that could be used for bacterial detection in a clinical setting.

**Future studies.** The qualifications for any detection system are reproducibility and accuracy. In addition, for marketing reasons, systems should be sensitive, cost efficient, easy to use, portable, and provide diagnosis in a short period of time. At this time, SPR immunosensors lack the vast majority of these characteristics.

The immunosensors developed in this work do not meet many of the criteria needed for clinical use. The assembly and detection protocols need to be optimized. A detailed description of potential optimization steps was discussed in Chapter 4 (Chapter 4.5.5). Briefly, all future work should be performed in a temperature control setting

(Dudak & Boyaci, 2009; Dutra *et al.*, 2007; He *et al.*, 2007; Matsumoto *et al.*, 2005; Naimushin *et al.*, 2003). Flow rates of both the reference buffer and the injections should be determined so that enough time is allowed for molecules to interact on the gold surface without causing unbound molecule aggregates to form on the surface (Dudak & Boyaci, 2009; Skottrup *et al.*, 2008). Different methods can be used to ensure proper orientation of the antibodies onto the sensor without decreasing the sensitivity of the sensor (Dudak & Boyaci, 2007). In addition, the ionic composition and pH of buffers should be adjusted to maximize the detection limit (Srilatha & Murthy, 2006). Choice of antibodies also plays a role in adjusting the sensitivity and accuracy of the detection system.

In general, SPR immunosensors are more cost efficient than using sequencing for bacterial detection but more expensive than the commonly used phenotypic detection systems (Dalvie *et al.*, 2005; Rantakokko-Jalava *et al.*, 2006). The largest part of the SPR cost is the gold chip that costs around \$30. It is possible to reuse the chip after the first detection experiment by injecting a harsh chemical that removes all material adsorbed onto the gold (Waswa *et al.*, 2006). This method does affect the gold on the glass chip which will decrease the sensitivity of the next detection experiment. Validation studies would have to be performed to determine how significant the decrease is. The loss of sensitivity may outweigh the need to reduce cost. Low labor costs also makes the system more cost efficient. No additional training would be required for the technician to use SPR. The instrument is relatively easy to use with the only labor being manual injections into the instrument.

The time required for results is on par with ELISA and commercial identification systems (Funke & Funke-Kissling, 2005; Janyapoon *et al.*, 2000; Rantakokko-Jalava *et al.*, 2006). The lengthiest part of the analysis for all these methods is isolating and culturing a bacterium from a sample. Once the sensor has been assembled and the bacteria have grown, the antibody-antigen interaction can be observed within three to ten minutes depending on the injection flow rate. Changes in SPR response are observed in real time and no additional wait is needed for the results. With optimization of buffers, chips, antibodies, and assembly protocols it may be possible to create an immunosensor that can use human fluids and is faster and cheaper than current identification tools.

Currently developed SPR immunosensors could only be marketed to test for one type of bacteria in a pure culture or with samples containing two or three isolates that do not cross react with the antibodies on the sensor chip. As discussed earlier, a CF sputum sample contains a multitude of bacteria. A multiplex SPR immunosensor could be developed to analyze complex samples like CF sputa. This sensor must be made with multiple species-specific polyclonal antibodies affixed to the chip. In addition, the SPR response for each type of antibody on the chip would have to be distinguishable from each other. Advances in nanotechnology may allow SPR instruments to detect multiple antibody – antigen detections at faster rates and better detection limits.

## **5.5 Conclusions**

This work further proved the CF lung contains numerous bacteria and fungi and the lung community is dynamic affected by multiple drivers. By understanding the members in the community, one can begin to learn the role of the organisms within the

community and how they affect the host. These relationships are the key to decreasing lung damage and increasing the median age of survival for the patient, which is 37.7 years of age (Patient-registry, 2008). In the future, antibiotics may be prescribed to shift the community dynamic to a healthier state instead of killing one bacterium which leaves the door open for other pathogens.

The research projects in this study used two groups of CF patients, with 18-19 participants. These patients were all adults and mostly Caucasians. Future studies should use a larger population size that includes various races, ethnicity, ages, and *cfr* mutations. These factors as well as diet, health, and antibiotic use should be used to statistically determine if there is a correlation between the factors and the microbiome. In addition, healthy samples need to be analyzed to determine the organisms associated with the normal oral/lung core biome. Through the use of 454 sequencing and LH-PCR, the organisms in the lung can be identified, tracked and then related to host factors. Then, it can be determined if the presence of some organisms are necessary for good lung function. Contrary to common belief, the bronchial tree of the healthy person is not sterile and these microbes may play a role in maintaining lung health (Hilty *et al.*, 2010). Certain personal factors may also increase the likelihood of a particular microbe colonizing the lung. Determining how a microbial community causes pathogenicity is the next step to understanding the CF lung biome.

The studies addressed in this work can also be implemented with infants and children. By using LH profiling and 454 sequencing, the transition from a sterile newborn lower lung to the complex adult lung infection can be observed. Determining the organism that first cultures the lung and how the infection progresses throughout

childhood may change how the United States treats CF children. In Europe, CF patients are treated early and aggressively with antibiotics with the goal being eradication (Hansen *et al.*, 2008). In the US, treatment is only post infection (Flume *et al.*, 2009). Understanding the community in the neonate may lend one course of treatment better than the other and perhaps extend the life span considerably.

CF lung infections will need to be diagnosed, understood, and treated until *cfr* mutations are no longer being passed on to new generations. Continuing this research using 454 sequencing and community profiling with a large population of CF and healthy individuals may provide the knowledge needed to stabilize the lung community and reduce the pathogens that cause lung deterioration which ultimately kills the CF individual.

The advances of sequencing technology has made it cheaper, faster, and easier to identify microbial DNA in a CF sample. The name of the organism is but the first step in understanding the microscopic community living in the CF lung. Understanding how bacteria interact with each other and with the host is the next step in the puzzle.

## BIBLIOGRAPHY

- Aas, J., Paster, B., Stokes, L., Olsen, I. & Dewhirst, F. (2005). Defining the normal bacterial flora of the oral cavity. *J Clin Microbiol* 43, 5721 - 5732.
- Abel, K., de Schmertzing, H. & Peterson, J. I. (1963). Classification of microorganisms by analysis of chemical composition. I. Feasibility of utilizing gas chromatography. *J Bacteriol* 86, 1039-1044.
- Adam, D. (2001). Individual genomes targeted in sequencing revolution. *Nature* 411, 402-402.
- Adler, F. R., Aurora, P., Barker, D. H., Barr, M. L., Blackwell, L. S., Bosma, O. H., Brown, S., Cox, D. R., Jensen, J. L., Kurland, G., Nossent, G. D., Quittner, A. L., Robinson, W. M., Romero, S. L., Spencer, H., Sweet, S. C., van der Bij, W., Vermeulen, J., Verschuuren, E. A. M., Vrijlandt, E. J. L. E., Walsh, W., Woo, M. S. & Liou, T. G. (2009). Lung transplantation for cystic fibrosis. *Proc Am Thorac Soc* 6, 619-633.
- Ahmadian, A., Ehn, M. & Hober, S. (2006). Pyrosequencing: History, biochemistry and future. *Clin Chim Acta* 363, 83-94.
- Alonso, J. L., Amoros, I., Chong, S. & Garelick, H. (1996). Quantitative determination of *Escherichia coli* in water using CHROMagar *E. coli*. *J of Microbiol Meth* 25, 309-315.
- Aloush, V., Navon-Venezia, S., Seigman-Igra, Y., Cabili, S. & Carmeli, Y. (2006). Multidrug-resistant *Pseudomonas aeruginosa*: Risk factors and clinical impact. *Antimicrob Agents Chemother* 50, 43-48.
- Andersen, D. H. (1938). Cystic fibrosis of the pancreas and its relation to celiac disease: a clinical and pathological study. *Am J Dis Child* 56, 344-399.
- Anderson, M., Rich, D., Gregory, R., Smith, A. & Welsh, M. (1991). Generation of cAMP-activated chloride currents by expression of CFTR. *Science* 251, 679-682.
- Antonopoulos, D. A., Huse, S. M., Morrison, H. G., Schmidt, T. M., Sogin, M. L. & Young, V. B. (2009). Reproducible community dynamics of the gastrointestinal microbiota following antibiotic perturbation. *Infect Immun* 77, 2367-2375.
- Anzai, Y., Kudo, Y. & Oyaizu, H. (1997). The phylogeny of the genera *Chryseomonas*, *Flavimonas*, and *Pseudomonas* supports synonymy of these three genera. *Int J Syst Bacteriol* 47, 249-251.
- Arbes, J. S. J., Cohen, E. A., Sever, M. L. & Zeldin, D. C. (2006). Oral pathogens and asthma: A potential connection. *J Allergy Clin Immunol* 117, S114-S114.

- Armougom, F., Bittar, F., Stremmer, N., Rolain, J. M., Robert, C., Dubus, J. C., Sarles, J., Raoult, D. & La Scola, B. (2009). Microbial diversity in the sputum of a cystic fibrosis patient studied with 16S rDNA pyrosequencing. *Eur J Clin Microbiol Infect Dis* 28, 1151-1154.
- Ashley, D. J. & Brindle, M. J. (1960). Penicillin resistance in staphylococci isolated in a casualty department. *J Clin Pathol* 13, 336-338.
- Aston, F. W. (1919). A positive ray spectrograph. *Phil Mag (Series 6)* 37, 707-714.
- Atkins, S. D. & Clark, I. M. (2004). Fungal molecular diagnostics: A mini review. *J Appl Genet* 45, 3-15.
- Bahrani-Mougeot, F. K., Paster, B. J., Coleman, S., Barbuto, S., Brennan, M. T., Noll, J., Kennedy, T., Fox, P. C. & Lockhart, P. B. (2007). Molecular analysis of oral and respiratory bacterial species associated with ventilator-associated pneumonia. *J Clin Microbiol* 45, 1588-1593.
- Bakare, N., Rickerts, V., Bargon, J. & Just-Nubling, G. (2003). Prevalence of *Aspergillus fumigatus* and other fungal species in the sputum of adult patients with cystic fibrosis. *Mycoses* 46, 19-23.
- Balajee, S. A., Nickle, D., Varga, J. & Marr, K. A. (2006). Molecular studies reveal frequent misidentification of *Aspergillus fumigatus* by morphotyping. *Eukaryot cell* 5, 1705-1712.
- Balasubramanian, S., Klenerman, D. & Barnes, C. (2002). Arrayed polynucleotides and their use in genome analysis. Solexa, Ltd. US 2003/0022207 A1
- Baumgartner, C., Freydiere, A. M. & Gille, Y. (1996). Direct identification and recognition of yeast species from clinical material by using *albicans* ID and CHROMagar *Candida* plates. *J Clin Microbiol* 34, 454-456.
- Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J. & Wheeler, D. L. (2008). GenBank. *Nucleic Acids Res* 36, 25-30.
- Beran, R. (1977). Minimum Hellinger distance estimates for parametric models. *Ann Stat* 5, 445-463.
- Berggard, I. & Edelman, G. M. (1963). Normal counterparts to Bence-Jones proteins: Free L polypeptide chains of human gamma-globulin. *Proc Natl Acad Sci U S A* 49, 330-337.
- Beringer, P. (2001). The clinical use of colistin in patients with cystic fibrosis. *Curr Opin Pulm Med* 7, 434-440.



- Beringer, P. M. & Appleman, M. D. (2000). Unusual respiratory bacterial flora in cystic fibrosis: Microbiologic and clinical features. *Curr Opin Pulm Med* 6, 545-550.
- Bernhard, A. E., Colbert, D., McManus, J. & Field, K. G. (2005). Microbial community dynamics based on 16S rRNA gene profiles in a Pacific Northwest estuary and its tributaries. *FEMS Microbiol Ecol* 52, 115-128.
- Bhally, H. S., Lema, C., Romagnoli, M., Borek, A., Wakefield, T. & Carroll, K. C. (2005). *Leptotrichia buccalis* bacteremia in two patients with acute myelogenous leukemia. *Anaerobe* 11, 350-353.
- Biagini, R. E., Sammons, D. L., Smith, J. P., MacKenzie, B. A., Striley, C. A. F., Snawder, J. E., Robertson, S. A. & Quinn, C. P. (2006). Rapid, sensitive, and specific lateral-flow immunochromatographic device to measure anti-anthrax protective antigen immunoglobulin G in serum and whole blood. *Clin Vaccine Immunol* 13, 541-546.
- Biosensing Instrument (2006). SPR technology. [www.biosensingusa.com](http://www.biosensingusa.com)
- Bittar, F., Richet, H., Dubus, J. C., Reynaud-Gaubert, M., Stremler, N., Sarles, J., Raoult, D. & Rolain, J.-M. (2008). Molecular detection of multiple emerging pathogens in sputa from cystic fibrosis patients. *PLoS ONE* 3, e2908.
- Bjerketorp, J., Ng Tze Chiang, A., Hjort, K., Rosenquist, M., Liu, W. T. & Jansson, J. K. (2008). Rapid lab-on-a-chip profiling of human gut bacteria. *J Microbiol Meth* 72, 82-90.
- Blackwood, C. B., Marsh, T., Kim, S.-H. & Paul, E. A. (2003). Terminal restriction fragment length polymorphism data analysis for quantitative comparison of microbial communities. *Appl Environ Microbiol* 69, 926-932.
- Boer, S. I., Hedtkamp, S. I. C., van Beusekom, J. E. E., Fuhrman, J. A., Boetius, A. & Ramette, A. (2009). Time and sediment depth related variations in bacterial diversity and community structure in subtidal sands. *ISME J* 3, 780-791.
- Bokken, G., Corbee, R., van Knapen, F. & Bergwerff, A. (2003). Immunochemical detection of *Salmonella* group B, D, and E using an optical surface plasmon resonance biosensor. *FEMS Microbiol Lett* 222, 75-82.
- Boni, M. F. & Feldman, M. W. (2005). Evolution of antibiotic resistance by human and bacterial niche construction. *Evolution* 59, 477-491.
- Bosshard, P. P., Zbinden, R., Abels, S., Boddington, B., Altwegg, M. & Bottger, E. C. (2006). 16S rRNA gene sequencing versus the API 20 NE system and the VITEK 2 ID-GNB card for identification of nonfermenting Gram-negative bacteria in the clinical laboratory. *J Clin Microbiol* 44, 1359-1366.

- Bouchara, J.-P., Hsieh, H. Y., Croquefer, S., Barton, R., Marchais, V., Pihet, M. & Chang, T. C. (2009). Development of an oligonucleotide array for direct detection of fungi in sputum samples from patients with cystic fibrosis. *J Clin Microbiol* 47, 142-152.
- Boyanton, B. L., Jr., Luna, R. A., Fasciano, L. R., Menne, K. G. & Versalovic, J. (2008). DNA pyrosequencing-based identification of pathogenic *Candida* species by using the internal transcribed spacer 2 region. *Arch Pathol Lab Med* 132, 667-674.
- Bravo, L. T. C., Tuohy, M. J., Ang, C., Destura, R. V., Mendoza, M., Procop, G. W., Gordon, S. M., Hall, G. S. & Shrestha, N. K. (2009). Pyrosequencing for rapid detection of *Mycobacterium tuberculosis* resistance to rifampin, isoniazid, and fluoroquinolones. *J Clin Microbiol* 47, 3985-3990.
- Bray, J. R. & Curtis, J. T. (1957). An ordination of the upland forest communities in southern Wisconsin. *Ecol Monograph* 27, 3226-3349.
- Brett, M. M., Ghoneim, A. T. & Littlewood, J. M. (1988). An ELISA to detect antipseudomonal IgA antibodies in sera of patients with cystic fibrosis. *J Clin Pathol* 41, 1130-1134.
- Brisse, S., Stefani, S., Verhoef, J., van Belkum, A., Vandamme, P. & Goessens, W. (2002). Comparative evaluation of the BD Phoenix and VITEK 2 automated instruments for identification of isolates of the *Burkholderia cepacia* complex. *J Clin Microbiol* 40, 1743-1748.
- Bunyaratavej, P. (2003). The high frequency of *Porphyromonas gingivalis*, *Prevotella intermedia*, and *Prevotella nigrescens* in subgingival plaque may be associated with periodontal diseases in subjects aged 30 to 49 years. *J Evid Based Dent Pract* 3, 103-104.
- Burns, J. L., Emerson, J., Stapp, J. R., Yim, D. L., Krzewinski, J., Loudon, L., Ramsey, B. W. & Clausen, C. R. (1998). Microbiology of sputum from patients at cystic fibrosis centers in the United States. *Clin Infect Dis* 27, 158-163.
- Buzzetti, R., Salvatore, D., Baldo, E., Forneris, M. P., Lucidi, V., Manunza, D., Marinelli, I., Messori, B., Neri, A. S., Raia, V., Furnari, M. L. & Mastella, G. (2009). An overview of international literature from cystic fibrosis registries: 1. Mortality and survival studies in cystic fibrosis. *J Cyst Fibros* 8, 229-237.
- Cain, T., Lubman, D. & Webber, W., Jr. (1994). Differentiation of bacteria using protein profiles from MALDI-TOF/MS. *Rapid Commun Mass Spectrom* 8, 1026-1030.
- Caldwell, C. C., Chen, Y., Goetzmann, H. S., Hao, Y., Borchers, M. T., Hassett, D. J., Young, L. R., Mavrodi, D., Thomashow, L. & Lau, G. W. (2009). *Pseudomonas*

*aeruginosa* exotoxin pyocyanin causes cystic fibrosis airway pathogenesis. *Am J Pathol* 175, 2473-2488.

- Cannon, J., Lee, T., Bolanos, J. & Danziger, L. (2005). Pathogenic relevance of *Lactobacillus*: A retrospective review of over 200 cases. *Eur J Clin Microbiol Infect Dis* 24, 31-40.
- Canton, R., Cobos, N., de Gracia, J., Baquero, F., Honorato, J., Gartner, S., Alvarez, A., Salcedo, A., Oliver, A. & Garcia-Quetglas, E. (2005). Antimicrobial therapy for pulmonary pathogenic colonisation and infection by *Pseudomonas aeruginosa* in cystic fibrosis patients. *Clin Microbiol Infect* 11, 690-703.
- Carbonnelle, E., Beretti, J.-L., Cottyn, S., Quesne, G., Berche, P., Nassif, X. & Ferroni, A. (2007). Rapid identification of Staphylococci isolated in clinical microbiology laboratories by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *J Clin Microbiol* 45, 2156-2161.
- Carrier, J.-P., K'ouas, G., Bonne, I., Lozniewski, A. & Mory, F. (2004). *Oribacterium sinus* gen. nov., sp. nov., within the family 'Lachnospiraceae' (phylum Firmicutes). *Int J Syst Evol Microbiol* 54, 1611-1615.
- Carrier, J.-P., K'ouas, G. & Han, X. Y. (2007). *Moryella indoligenes* gen. nov., sp. nov., an anaerobic bacterium isolated from clinical specimens. *Int J Syst Evol Microbiol* 57, 725-729.
- Carrero-Colon, M., Nakatsu, C. H. & Konopka, A. (2006). Effect of nutrient periodicity on microbial community dynamics. *Appl Environ Microbiol* 72, 3175-3183.
- Casalta, J.-P., Fournier, P.-E., Habib, G., Riberi, A. & Raoult, D. (2005). Prosthetic valve endocarditis caused by *Pseudomonas luteola*. *BMC Infectious Diseases* 5, 82.
- Case, R. J., Boucher, Y., Dahllorf, I., Holmstrom, C., Doolittle, W. F. & Kjelleberg, S. (2007). Use of 16S rRNA and rpoB genes as molecular markers for microbial ecology studies. *Applied and Environmental Microbiology* 73, 278-288.
- Ceppellini, R., Dray, S. & Edelman, G. M. (1964). Nomenclature for human immunoglobulins. *WHO Bulletins* 30, 447-450.
- Chang, C. W. & Liao, J. (2008). Nano-indentation at the surface contact level: Applying a harmonic frequency for measuring contact stiffness of self-assembled monolayers adsorbed on Au. *Nanotechnology* 19, 315703 (online).
- Chen, Y. C., Eisner, J. D., Kattar, M. M., Rassoulian-Barrett, S. L., LaFe, K., Yarfitz, S. L., Limaye, A. P. & Cookson, B. T. (2000). Identification of medically important yeasts using PCR-based detection of DNA sequence polymorphisms in the internal transcribed spacer 2 region of the rRNA genes. *J Clin Microbiol* 38, 2302-2310.

- Chen, Y. C., Eisner, J. D., Kattar, M. M., Rassouljian-Barrett, S. L., Lafe, K., Bui, U., Limaye, A. P. & Cookson, B. T. (2001). Polymorphic internal transcribed spacer region 1 DNA sequences identify medically important yeasts. *J Clin Microbiol* 39, 4042-4051.
- Chiao, D.-J., Wey, J.-J., Shyu, R.-H. & Tang, S.-S. (2008). Monoclonal antibody-based lateral flow assay for detection of botulinum neurotoxin type A. *Hybridoma* 27, 31-35.
- Chou, Y. J., Sheu, S. Y., Sheu, D. S., Wang, J. T. & Chen, W. M. (2006). *Schlegelella aquatica* sp. nov., a novel thermophilic bacterium isolated from a hot spring. *Int J Syst Evol Microbiol* 56, 2793-2797.
- Cimon, B., Carrere, J., Chazalette, J. P., Vinatier, J. F., Chabasse, D. & Bouchara, J. P. (1999). Chronic airway colonization by *Penicillium emersonii* in a patient with cystic fibrosis. *Med Mycol* 37, 291-293.
- Cimon, B., Symoens, F., Zouhair, R., Chabasse, D., Nolard, N., Defontaine, A. & Bouchara, J.-P. (2001). Molecular epidemiology of airway colonisation by *Aspergillus fumigatus* in cystic fibrosis patients. *J Med Microbiol* 50, 367-374.
- Cimon, B., Zouhair, R., Symoens, F., Carrere, J., Chabasse, D. & Bouchara, J. P. (2003). *Aspergillus terreus* in a cystic fibrosis clinic: Environmental distribution and patient colonization pattern. *J Hosp Infect* 53, 81-82.
- Cimon, B., Challier, S., Beguin, H., Carrere, J., Chabasse, D. & Bouchara, J. P. (2005). Airway colonization by *Acrophialophora fusispora* in patients with cystic fibrosis. *J Clin Microbiol* 43, 1484-1487.
- Clarke, K. R., Somerfield, P. J., Airoidi, L. & Warwick, R. M. (2006). Exploring interactions by second-stage community analyses. *J Exp Marine Biol* 338, 179-192.
- Cocolin, L., Manzano, M., Cantoni, C. & Comi, G. (2001). Denaturing gradient gel electrophoresis analysis of the 16S rRNA gene V1 region to monitor dynamic changes in the bacterial population during fermentation of Italian sausages. *Appl Environ Microbiol* 67, 5113-5121.
- Coenye, T., Liu, L., Vandamme, P. & LiPuma, J. J. (2001). Identification of *Pandoraea* species by 16S ribosomal DNA-based PCR assays. *J Clin Microbiol* 39, 4452-4455.
- Coenye, T., Goris, J., Spilker, T., Vandamme, P. & LiPuma, J. J. (2002). Characterization of unusual bacteria isolated from respiratory secretions of cystic fibrosis patients and description of *Inquilinus limosus* gen. nov., sp. nov. *J Clin Microbiol* 40, 2062-2069.

- Coenye, T., Goris, J., De Vos, P., Vandamme, P. & Lipuma, J. J. (2003). Classification of *Ralstonia pickettii*-like isolates from the environment and clinical samples as *Ralstonia insidiosa* sp. nov. *Int J Syst Evol Microbiol* 53, 1075-1080.
- Coenye, T. & Vandamme, P. (2003). Intragenomic heterogeneity between multiple 16S ribosomal RNA operons in sequenced bacterial genomes. *FEMS Microbiol Lett* 228, 45-49.
- Coenye, T., Spilker, T., Reik, R., Vandamme, P. & Lipuma, J. J. (2005). Use of PCR analyses to define the distribution of *Ralstonia* species recovered from patients with cystic fibrosis. *J Clin Microbiol* 43, 3463-3466.
- Coleman, D. C., Sullivan, D. J., Bennett, D. E., Moran, G. P., Barry, H. J. & Shanley, D. B. (1997). Candidiasis: The emergence of a novel species, *Candida dubliniensis*. *AIDS* 11, 557-567.
- Conrad, R. S., Wulf, R. G. & Clay, D. L. (1979). Effects of carbon sources on antibiotic resistance in *Pseudomonas aeruginosa*. *Antimicrob Agents Chemother* 15, 59-66.
- Cooke, V. M., Miles, R. J., Price, R. G., Midgley, G., Khamri, W. & Richardson, A. C. (2002). New chromogenic agar medium for the identification of *Candida* spp. *Appl Environ Microbiol* 68, 3622-3627.
- Corner, B. (1940). *Staphylococcus aureus* septicaemia with recovery, M & B 693. *Proc R Soc Med* 33, 627-628.
- Costello, E. K., Lauber, C. L., Hamady, M., Fierer, N., Gordon, J. I. & Knight, R. (2009). Bacterial community variation in human body habitats across space and time. *Science* 326, 1694-1697.
- Cox-Foster, D. L., Conlan, S., Holmes, E. C., Palacios, G., Evans, J. D., Moran, N. A., Quan, P.-L., Briese, T., Hornig, M., Geiser, D. M., Martinson, V., vanEngelsdorp, D., Kalkstein, A. L., Drysdale, A., Hui, J., Zhai, J., Cui, L., Hutchison, S. K., Simons, J. F., Egholm, M., Pettis, J. S. & Lipkin, W. I. (2007). A metagenomic survey of microbes in honey bee colony collapse disorder. *Science* 318, 283-287.
- Crump, B. C., Kling, G. W., Bahr, M. & Hobbie, J. E. (2003). Bacterioplankton community shifts in an arctic lake correlate with seasonal changes in organic matter source. *Appl Environ Microbiol* 69, 2253-2268.
- Cui, X., Pei, R., Wang, Z., Yang, F., Ma, Y., Dong, S. & Yang, X. (2003). Layer-by-layer assembly of multilayer films composed of avidin and biotin-labeled antibody for immunosensing. *Biosens Bioelectron* 18, 59-67.
- Curtis, T. P., Sloan, W. T. & Scannell, J. W. (2002). Estimating prokaryotic diversity and its limits. *Proc Natl Acad Sci U S A* 99, 10494-10499.

- da Costa, C. T., Porter, C., Parry, K., Morris, A. & Quoraishi, A. H. (1996). Empyema thoracis and lung abscess due to *Gemella morbillorum*. *Eur J Clin Microbiol Infect Dis* 15, 75-77.
- da Silva Filho, L. V., Tateno, A. F., Velloso Lde, F., Levi, J. E., Fernandes, S., Bento, C. N., Rodrigues, J. C. & Ramos, S. R. (2004). Identification of *Pseudomonas aeruginosa*, *Burkholderia cepacia* complex, and *Stenotrophomonas maltophilia* in respiratory samples from cystic fibrosis patients using multiplex PCR. *Pediatr Pulmonol* 37, 537-547.
- da Silva Filho, L. V. F., Levi, J. E., Bento, C. N. O., da Silva Ramos, S. R. T. & Rozov, T. (1999). PCR identification of *Pseudomonas aeruginosa* and direction detection in clinical samples from cystic fibrosis patients. *J Med Microbiol* 48, 357-361.
- Dalvie, M. A., Sinanovic, E., London, L., Cairncross, E., Solomon, A. & Adam, H. (2005). Cost analysis of ELISA, solid-phase extraction, and solid-phase microextraction for the monitoring of pesticides in water. *Environ Res* 98, 143-150.
- Davis, P. B. (1987). Pathophysiology of cystic fibrosis with emphasis on salivary gland involvement. *J Dent Res* 66, 667-671.
- De Baets, F., Schlestraete, P., van Sabine, D., Haerynck, F. & Vaneechoutte, M. (2007). *Achromobacter xylosoxidans* in cystic fibrosis: Prevalence and clinical relevance. *J Cyst Fibros* 6, 75-78.
- De Leuw, J. & Heiser, W., (editors) (1982). *Theory of multi-dimensional scaling*. Amsterdam: North-Holland.
- Defontaine, A., Zouhair, R., Cimon, B., Carrere, J., Bailly, E., Symoens, F., Diouri, M., Hallet, J. N. & Bouchara, J. P. (2002). Genotyping study of *Scedosporium apiospermum* isolates from patients with cystic fibrosis. *J Clin Microbiol* 40, 2108-2114.
- Degand, N., Carbonnelle, E., Dauphin, B., Beretti, J. L., Le Bourgeois, M., Sermet-Gaudelus, I., Segonds, C., Berche, P., Nassif, X. & Ferroni, A. (2008). Matrix-assisted laser desorption ionization-time of flight mass spectrometry for identification of nonfermenting gram-negative bacilli isolated from cystic fibrosis patients. *J Clin Microbiol* 46, 3361-3367.
- Demko, C. A., Stern, R. C. & Doershuk, C. F. (1998). *Stenotrophomonas maltophilia* in cystic fibrosis: Incidence and prevalence. *Pediatr Pathol* 25, 304-308.
- Dempster, A. J. (1918). A new method of positive ray analysis. *Phys Rev* 11, 316-325.

- Deo, V., Bhongade, M., Ansari, S. & Chaven, R. (2009). Periodontitis as a potential risk factor for chronic obstructive pulmonary disease: A retrospective study. *Indian J Dent Res* 20, 466-470.
- DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., Huber, T., Dalevi, D., Hu, P. & Andersen, G. L. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 72, 5069-5072.
- Dickie, I. A. & FitzJohn, R. G. (2007). Using terminal restriction fragment length polymorphism (T-RFLP) to identify mycorrhizal fungi: A methods review. *Mycorrhiza* 17, 259-270.
- Dixon, D. M. & Polak-Wyss, A. (1991). The medically important dematiaceous fungi and their identification. *Mycoses* 34, 1-18.
- Doggett, R. G., Harrison, G. M. & Wallis, E. S. (1964). Comparison of some properties of *Pseudomonas aeruginosa* isolated from infections in persons with and without cystic fibrosis. *J Bacteriol* 87, 427-431.
- Dole, M., Mack, L. L., Hines, R. L., Mobley, R. C., Ferguson, L. D. & Alice, M. B. (1968). Molecular beams of macroions. *J Chem Phys* 49, 2240-2249.
- Dong, Y., Wilkop, T., Xu, D., Wang, Z. & Cheng, Q. (2008). Microchannel chips for the multiplexed analysis of human immunoglobulin G-antibody interactions by surface plasmon resonance imaging. *Anal Biochem* 390, 1575-1583.
- Doublet, B., Robin, F., Casin, I., Fabre, L., Le Fleche, A., Bonnet, R. & Weill, F.-X. (2010). Molecular and biochemical characterization of the natural chromosome-encoded class A {beta}-lactamase from *Pseudomonas luteola*. *Antimicrob Agents Chemother* 54, 45-51.
- Doud, M., Zeng, E., Schneper, L., Narasimhan, G. & Mathee, K. (2009). Approaches to analyse dynamic microbial communities such as those seen in cystic fibrosis lung. *Hum Genomics* 3, 246-256.
- Downes, J., Munson, M. A., Radford, D. R., Spratt, D. A. & Wade, W. G. (2002). *Shuttleworthia satelles* gen. nov., sp. nov., isolated from the human oral cavity. *Int J Syst Evol Microbiol* 52, 1469-1475.
- Duan, K., Dammel, C., Stein, J., Rabin, H. & Surette, M. G. (2003). Modulation of *Pseudomonas aeruginosa* gene expression by host microflora through interspecies communication. *Mol Microbiol* 50, 1477-1491.
- Dudak, F. C. & Boyaci, I. H. (2007). Development of an immunosensor based on surface plasmon resonance for enumeration of *Escherichia coli* in water samples. *Food Res Int* 40, 803-807.

- Dudak, F. C. & Boyaci, I. H. (2009). Rapid and label-free bacteria detection by surface plasmon resonance (SPR) biosensors. *Biotechnol J* 4, 1003-1011.
- Dunbar, J., Ticknor, L. O. & Kuske, C. R. (2000). Assessment of microbial diversity in four southwestern United States soils by 16S rRNA gene terminal restriction fragment analysis. *Appl Environ Microbiol* 66, 2943-2950.
- Dunbar, J., Ticknor, L. O. & Kuske, C. R. (2001). Phylogenetic specificity and reproducibility and new method for analysis of terminal restriction fragment profiles of 16S rRNA genes from bacterial communities. *Appl Environ Microbiol* 67, 190-197.
- Dutra, R. F., Mendes, R. K., Lins da Silva, V. & Kubota, L. T. (2007). Surface plasmon resonance immunosensor for human cardiac troponin T based on self-assembled monolayer. *J Pharmaceut Biomed* 43, 1744-1750.
- Edelman, G. M., Benacerraf, B., Ovary, Z. & Poulik, M. D. (1961). Structural differences among antibodies of different specificities. *Proc Natl Acad Sci U S A* 47, 1751-1758.
- Edelman, G. M. & Poulik, M. D. (1961). Studies on structural units of the gamma-globulins. *J Exp Med* 113, 861-884.
- Edelman, G. M. & Benacerraf, B. (1962). On structural and functional relations between antibodies and proteins of the gamma-system. *Proc Natl Acad Sci U S A* 48, 1035-1042.
- Elbanna, K., Lutke-Eversloh, T., van Trappen, S., Mergaert, J., Swings, J. & Steinbuchel, A. (2003). *Schlegelella thermodepolymerans* gen. nov., sp. nov., a novel thermophilic bacterium that degrades poly(3-hydroxybutyrate-co-3-mercaptopropionate). *Int J Syst Evol Microbiol* 53, 1165-1168.
- Elie, C. M., Lott, T. J., Reiss, E. & Morrison, C. J. (1998). Rapid identification of *Candida* species with species-specific DNA probes. *J Clin Microbiol* 36, 3260-3265.
- Elrod, R. P. & Braun, A. C. (1942). *Pseudomonas aeruginosa*: Its role as a plant pathogen. *J Bacteriol* 44, 633-645.
- Emmerich, P., Thome-Bolduan, C., Drosten, C., Gunther, S., Ban, E., Sawinsky, I. & Schmitz, H. (2006). Reverse ELISA for IgG and IgM antibodies to detect Lassa virus infections in Africa. *J Clin Virol* 37, 277-281.
- Engler, K. H., Efstratiou, A., Norn, D., Kozlov, R. S., Selga, I., Glushkevich, T. G., Tam, M., Melnikov, V. G., Mazurova, I. K., Kim, V. E., Tseneva, G. Y., Titov, L. P. & George, R. C. (2002). Immunochromatographic strip test for rapid detection of



- diphtheria toxin: Description and multicenter evaluation in areas of low and high prevalence of diphtheria. *J Clin Microbiol* 40, 80-83.
- Engvall, E. & Perlmann, P. (1971). Enzyme-linked immunosorbent assay (ELISA) quantitative assay of immunoglobulin G. *Immunochemistry* 8, 871-874.
- Engvall, E. & Perlmann, P. (1972). Enzyme-linked immunosorbent assay, Elisa. 3. Quantitation of specific antibodies by enzyme-labeled anti-immunoglobulin in antigen-coated tubes. *J Immunol* 109, 129-135.
- Entry, J. A., Mills, D., Mathee, K., Jayachandran, K., Sojka, R. E. & Narasimhan, G. (2008). Influence of irrigated agriculture on soil microbial diversity. *Appl Soil Ecol* 40, 146-154.
- Eribe, E. R. & Olsen, I. (2008). *Leptotrichia* species in human infections. *Anaerobe* 14, 131-137.
- Erlich, H. A., D. Gelfand, and J. Sninsky (1991). Recent advances in the polymerase chain reaction. *Science* 252, 1643-1651.
- Esfandiari, J., Lyashchenko, K., Greenwald, R. & Greenwald, D. (2004). Rapid lateral-flow test for detection of TB/HIV co-infection. In *International Conference on AIDS*. Bangkok, Thailand.
- Fahy, J. V., Keoghan, M. T., Crummy, E. J. & FitzGerald, M. X. (1991). Bacteraemia and fungaemia in adults with cystic fibrosis. *J Infect* 22, 241-245.
- Falla, T. J., Crook, D. W., Brophy, L. N., Maskell, D., Kroll, J. S. & Moxon, E. R. (1994). PCR for capsular typing of *Haemophilus influenzae*. *J Clin Microbiol* 32, 2382-2386.
- Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. & Whitehouse, C. M. (1989). Electrospray ionization for mass spectrometry of large biomolecules. *Science* 246, 64-71.
- Fierer, N. & Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. *Proc Natl Acad Sci U S A* 103, 626-631.
- Finck-Barbançon, V., Goranson, J., Zhu, L., Sawa, T., Wiener-Kronish, J. P., Fleiszig, S. M. J., Wu, C., Mende-Mueller, L. & Frank, D. W. (1997). ExoU expression by *Pseudomonas aeruginosa* correlates with acute cytotoxicity and epithelial injury. *Mol Microbiol* 25, 547-557.
- Flanagan, J. L., Brodie, E. L., Weng, L., Lynch, S. V., Garcia, O., Brown, R., Hugenholtz, P., DeSantis, T. Z., Andersen, G. L., Wiener-Kronish, J. P. & Bristow, J. (2007). Loss of bacterial diversity during antibiotic treatment of

- intubated patients colonized with *Pseudomonas aeruginosa*. *J Clin Microbiol* 45, 1954-1962.
- Flume, P. A., Mogayzel, P. J., Jr., Robinson, K. A., Goss, C. H., Rosenblatt, R. L., Kuhn, R. J., Marshall, B. C. & the Clinical Practice Guidelines for Pulmonary Therapies Committee (2009). Cystic fibrosis pulmonary guidelines: Treatment of pulmonary exacerbations. *Am J Respir Crit Care Med* 180, 802-808.
- Fox, A. (2006). Mass Spectrometry for Species or Strain Identification after Culture or without Culture: Past, Present, and Future, pp. 2677-2680.
- Fratamico, P., Strobaugh, T., Medina, M. & Gehring, A. (1998). Detection of *Escherichia coli* P157:H7 using a surface plasmon resonance biosensor. *Biotechnol Tech* 12, 571-576.
- Freer, J. H. & Arbuthnott, J. P. (1982). Toxins of *Staphylococcus aureus*. *Pharmacol Ther* 19, 55-106.
- Fullbrook, S., Elson, S. & Slocombe, B. (1970). R-factor mediated beta-lactamase in *Pseudomonas aeruginosa*. *Nature* 226, 1054-1056.
- Funke, G. & Funke-Kissling, P. (2005). Performance of the new VITEK 2 GP card for identification of medically relevant gram-positive cocci in a routine clinical laboratory. *J Clin Microbiol* 43, 84-88.
- Gaillot, O., Di Camillo, P., Berche, P., Courcol, R. & Savage, C. (1999). Comparison of CHROMagar *Salmonella* medium and Hektoen Enteric agar for isolation of *Salmonellae* from stool samples. *J Clin Microbiol* 37, 762-765.
- Garrett, R. & Grisham, C. M. (2010). *Biochemistry*, 4th edn. Belmont, CA: Brooks/Cole, Cengage Learning.
- Gaschet, A., Piau, C., Violette, J., Donnio, P. Y., Jolivet-Gougeon, A., Engrand, C., Betremieux, P., Pladys, P. & Tattevin, P. (2009). Multiple brain abscesses caused by *Pseudomonas luteola*. *Pediatr Infect Dis J* 28, 1144-1146.
- Geddes, N. J., Martin, A. S., Caruso, F., Urquhart, R. S., Furlong, D. N., Sambles, J. R., Than, K. A. & Edgar, J. A. (1994). Immobilisation of IgG onto gold surfaces and its interaction with anti-IgG studied by surface plasmon resonance. *J Immunol Methods* 175, 149-160.
- Geisler, W. M., Malhotra, U. & Stamm, W. E. (2001). Pneumonia and sepsis due to fluoroquinolone-resistant *Capnocytophaga gingivalis* after autologous stem cell transplantation. *Bone Marrow Transplant* 28, 1171-1173.

- Ghannoum, M. A., Jurevic, R. J., Mukherjee, P. K., Cui, F., Sikaroodi, M., Naqvi, A. & Gillevet, P. M. (2010). Characterization of the oral fungal microbiome (Mycobiome) in healthy individuals. *PLoS Pathog* 6, e1000713.
- Gharizadeh, B., Kalantari, M., Garcia, C. A., Johansson, B. & Nyren, P. (2001). Typing of human Papillomavirus by pyrosequencing. *Lab Invest* 81, 673-679.
- Gillevet, P. (2006). Multitag Sequencing and Ecogenomic Analysis. BioSpherex LLC.
- Goldschmidt, P., Costa Ferreira, C., Degorge, S., Benallaous, D., Boutboul, S., Laroche, L., Batellier, L. & Chaumeil, C. (2009). Rapid detection and quantification of *Propionibacteriaceae*. *Br J Ophthalmol* 93, 258-262.
- Gonzalez, G., Doud, M., Mathee, K. & Narasimhan, G. (2009). Computer-assisted bacterial identification using 16S rRNA sequence data. In *25th Southern Biomedical Engineering Conference: IFMBE Proceedings*, pp. 239-240. Miami, Florida: Springer Berlin Heidelberg.
- Gonzalez, J. M., Simo, R., Massana, R., Covert, J. S., Casamayor, E. O., Pedros-Alio, C. & Moran, M. A. (2000). Bacterial community structure associated with a dimethylsulfoniopropionate-producing North Atlantic algal bloom. *Appl Environ Microbiol* 66, 4237-4246.
- Goodacre, R., Heald, J. K. & Kell, D. B. (1999). Characterisation of intact microorganisms using electrospray ionisation mass spectrometry. *FEMS Microbiol Lett* 176, 17-24.
- Goodman, A. L., Kulasekara, B., Rietsch, A., Boyd, D., Smith, R. S. & Lory, S. (2004). A signaling network reciprocally regulates genes associated with acute infection and chronic persistence in *Pseudomonas aeruginosa*. *Dev Cell* 7, 745-754.
- Goss, C. H., Mayer-Hamblett, N., Aitken, M. L., Rubenfeld, G. D. & Ramsey, B. W. (2004). Association between *Stenotrophomonas maltophilia* and lung function in cystic fibrosis. *Thorax* 59, 955-959.
- Govan, J. R. & Nelson, J. W. (1993). Microbiology of cystic fibrosis lung infections: Themes and issues. *J R Soc Med* 86 Suppl 20, 11-18.
- Gowans, J. L., Humphrey, J. H. and Mitchison, N.A., (1971). A discussion on cooperation between lymphocytes in the immune response. *Proc Roy Soc London B* 176 369-481.
- Graf, B., Adam, T., Zill, E. & Gobel, U. B. (2000). Evaluation of the VITEK 2 system for rapid identification of yeasts and yeast-like organisms. *J Clin Microbiol* 38, 1782-1785.

- Gram, H. (1884). Über die isolierte Färbung der Schizomyceten in Schnitt- und Trockenpräparaten. *Fortschritte der Medizin* 2, 185-189.
- Grant, A. & Ogilvie, L. A. (2003). Terminal restriction fragment length polymorphism data analysis. *Appl Environ Microbiol* 69, 6342; author reply 6342-6343.
- Griffith, S., Nathan, C., Selander, R., Chamberlin, W., Gordon, S., Kabins, S. & Weinstein, R. (1989). The epidemiology of *Pseudomonas aeruginosa* in oncology patients in a general hospital. *J Infect Dis* 160, 1030-1036.
- Griffiths, R. I., Whiteley, A. S., O'Donnell, A. G. & Bailey, M. J. (2000). Rapid method for coextraction of DNA and RNA from natural environments for analysis of ribosomal DNA- and rRNA-based microbial community composition. *Appl Environ Microbiol* 66, 5488-5491.
- Griffiths, W. J., Jonsson, A. P., Liu, S., Rai, D. K. & Wang, Y. (2001). Electrospray and tandem mass spectrometry in biochemistry. *Biochem J* 355, 545-561.
- Gross, J. H. (2004). *Mass Spectrometry: A Textbook*. Heidelberg: Springer.
- Gunapala, N. & Scow, K. M. (1998). Dynamics of soil microbial biomass and activity in conventional and organic farming systems. *Soil Biol Biochem* 30, 805-816.
- Gundes, S. G., Gulenc, S. & Bingol, R. (2001). Comparative performance of Fungichrom I, Candifast and API 20C Aux systems in the identification of clinically significant yeasts. *J Med Microbiol* 50, 1105-1110.
- Hansen, C. R., Pressler, T. & Høiby, N. (2008). Early aggressive eradication therapy for intermittent *Pseudomonas aeruginosa* airway colonization in cystic fibrosis patients: 15 years experience. *J Cyst Fibros* 7, 523-530.
- Harris, J. K., De Groote, M. A., Sagel, S. D., Zemanick, E. T., Kapsner, R., Penvari, C., Kaess, H., Deterding, R. R., Accurso, F. J. & Pace, N. R. (2007). Molecular identification of bacteria in bronchoalveolar lavage fluid from children with cystic fibrosis. *Proc Natl Acad Sci U S A* 104, 20529-20533.
- Harrison, F. (2007). Microbial ecology of the cystic fibrosis lung. *Microbiology* 153, 917-923.
- Hartmann, A. A. (1978). Staphylococci of the normal human skin flora. *Arch Dermatol Res* 261, 295-302.
- Hartmann, M., Frey, B., Kolliker, R. & Widmer, F. (2005). Semi-automated genetic analyses of soil microbial communities: Comparison of T-RFLP and RISA based on descriptive and discriminative statistical approaches. *J Microbiol Methods* 61, 349-360.

- Haynes, W. C. (1951). *Pseudomonas aeruginosa*-its characterization and identification. *J Gen Microbiol* 5, 939-950.
- He, Y. B., Luo, H. Q. & Li, N. B. (2007). Thermodynamic and kinetic analysis of the interaction between hepatitis B surface antibody and antigen on a gold electrode modified with cysteamine and colloidal gold via electrochemistry. *Biosens Bioelectron* 22, 2952-2957.
- Hearty, S., Leonard, P., Quinn, J. & O'Kennedy, R. (2006). Production, characterisation and potential application of a novel monoclonal antibody for rapid identification of virulent *Listeria monocytogenes*. *J Microbiol Meth* 66, 294-312.
- Hegde, V., Leon-Velarde, C. G., Stam, C. M., Jaykus, L.-A. & Odumeru, J. (2007). Evaluation of BBL CHROMagar *Listeria* agar for the isolation and identification of *Listeria monocytogenes* from food and environmental samples. *J Microbiol Meth* 68, 82-87.
- Heller, L. C., Jones, M. & Widen, R. H. (2008). Comparison of DNA pyrosequencing with alternative methods for identification of *Mycobacteria*. *J Clin Microbiol* 46, 2092-2094.
- Henegariu, O., Heerema, N. A., Dlouhy, S. R., Vance, G. H. & Vogt, P. H. (1997). Multiplex PCR: Critical parameters and step-by-step protocol. *Biotechniques* 23, 504-511.
- Hennequin, C., Benailly, N., Silly, C., Sorin, M., Scheinmann, P., Lenoir, G., Gaillard, J. L. & Berche, P. (1997). In vitro susceptibilities to amphotericin B, itraconazole, and miconazole of filamentous fungi isolated from patients with cystic fibrosis. *Antimicrob Agents Chemother* 41, 2064-2066.
- Hennessee, C. & Li, Q. (2010). Micrococcineae: *Arthrobacter* and relatives. In *Handbook of hydrocarbon and lipid microbiology*, vol. 19, pp. 1853-1864. Edited by K. Timmis. Berlin Springer.
- Henry, T., Iwen, P. C. & Hinrichs, S. H. (2000). Identification of *Aspergillus* species using internal transcribed spacer regions 1 and 2. *J Clin Microbiol* 38, 1510-1515.
- Heremans, J. F., Heremans, M. T. & Schultze, H. E. (1959). Isolation and description of a few properties of the beta 2A-globulin of human serum. *Clin Chim Acta* 4, 96-102.
- Hill, T. C. J., Walsh, K. A., Harris, J. A. & Moffett, B. F. (2003). Using ecological diversity measures with bacterial communities. *FEMS Microbiol Ecol* 43, 1-11.
- Hillis, D. M. & Dixon, M. T. (1991). Ribosomal DNA: Molecular evolution and phylogenetic inference. *Q Rev Biol* 66, 411-453.

- Hilty, M., Burke, C., Pedro, H., Cardenas, P., Bush, A., Bossley, C., Davies, J., Ervine, A., Poulter, L., Pachter, L., Moffatt, M. F. & Cookson, W. O. (2010). Disordered microbial communities in asthmatic airways. *PLoS ONE* 5, e8578.
- Hinrikson, H. P., Hurst, S. F., Lott, T. J., Warnock, D. W. & Morrison, C. J. (2005). Assessment of ribosomal large-subunit D1-D2, internal transcribed spacer 1, and internal transcribed spacer 2 regions as targets for molecular identification of medically important *Aspergillus* species. *J Clin Microbiol* 43, 2092-2103.
- Hiraishi, A., Iwasaki, M. & Shinjo, H. (2000). Terminal restriction pattern analysis of 16S rRNA genes for the characterization of bacterial communities of activated sludge. *J Biosci Bioeng* 90, 148-156.
- Hodges, E., Xuan, Z., Balija, V., Kramer, M., Molla, M. N., Smith, S. W., Middle, C. M., Rodesch, M. J., Albert, T. J., Hannon, G. J. & McCombie, W. R. (2007). Genome-wide in situ exon capture for selective resequencing. *Nat Genet* 39, 1522-1527.
- Hoffmann, N., Rasmussen, T. B., Jensen, P., Stub, C., Hentzer, M., Molin, S., Ciofu, O., Givskov, M., Johansen, H. K. & Hoiby, N. (2005). Novel mouse model of chronic *Pseudomonas aeruginosa* lung infection mimicking cystic fibrosis. *Infect Immun* 73, 2504-2514.
- Holland, R. D., Wilkes, J. G., Rafii, F., Sutherland, J. B., Persons, C. C., Voorhees, K. J. & J. O. Lay, J. (1996). Rapid identification of intact whole bacteria based on spectral patterns using matrix-assisted laser desorption/ionization with time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom* 10, 1227-1232.
- Holloway, B. W. & Morgan, A. F. (1986). Genome organization in *Pseudomonas*. *Annu Rev Microbiol* 40, 79-105.
- Horré, R., Schaal, K. P., Siekmeier, R., Sterzik, B., de Hoog, G. S. & Schnitzler, N. (2004). Isolation of fungi, especially *Exophiala dermatitidis*, in patients suffering from cystic fibrosis. *Respiration* 71, 360-366.
- Horré, R., Marklein, G., Siekmeier, R., Nidermajer, S. & Reiffert, S. M. (2009). Selective isolation of *Pseudallescheria* and *Scedosporium* species from respiratory tract specimens of cystic fibrosis patients. *Respiration* 77, 320-324.
- Hoyler, S. L. & Antony, S. (2000). *Eikenella corrodens* causing severe parapneumonic infection with empyema in an immunocompetent patient: Case report and review of the literature. *Antimicrobics and Infect Dis Newsletter* 18, 29-31.
- Hsieh, S.-Y., Tseng, C.-L., Lee, Y.-S., Kuo, A.-J., Sun, C.-F., Lin, Y.-H. & Chen, J.-K. (2008). Highly efficient classification and identification of human pathogenic bacteria by MALDI-TOF MS. *Mol Cell Proteomics* 7, 448-456.

- Hughes, D. T. & Russell, N. J. (1982). The use of trimethoprim-sulfamethoxazole in the treatment of chest infections. *Rev Infect Dis* 4, 528-532.
- Hull, J. (2003). Basic science of cystic fibrosis. *Current Paediatrics* 13, 253-258.
- Huppert, M., Oliver, D. J. & Sun, S. H. (1978). Combined methenamine-silver nitrate and hematoxylin & eosin stain for fungi in tissues. *J Clin Microbiol* 8, 598-603.
- Huson, D. H., Auch, A. F., Qi, J. & Schuster, S. C. (2007). MEGAN analysis of metagenomic data. *Genome Res* 17, 377-386.
- Jaccard, P. (1901). Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bulletin del la Société Vaudoise des Sciences Naturelles* 37, 547-579.
- Jackson, O. L., Jr. (2001). MALDI-TOF mass spectrometry of bacteria. *Mass Spectrom Rev* 20, 172-194.
- Jakubicz, P. (1972). Studies on the isolation and biological properties of enterotoxin Staphylococci A and B. *Rocz Akad Med Im Juliana Marchiewskiego Bialymst* 24, 105-122.
- Janyapoon, K., Korbsrisate, S., Thamapa, H., Thongmin, S., Kanjanahareutai, S., Wongpredee, N. & Sarasombath, S. (2000). Rapid detection of *Salmonella enterica* serovar Choleraesuis in blood cultures by a dot blot enzyme-linked immunosorbent assay. *Clin Diagn Lab Immunol* 7, 977-979.
- Jefferis, R., Reimer, C. B., Skvaril, F., de Lange, G., Ling, N. R., Lowe, J., Walker, M. R., Phillips, D. J., Aloisio, C. H., Wells, T. W. & et al. (1985). Evaluation of monoclonal antibodies having specificity for human IgG sub-classes: Results of an IUIS/WHO collaborative study. *Immunol Lett* 10, 223-252.
- Johansen, P., Anderson, C. & Hadorn, B. (1968). Cystic fibrosis of the pancreas: A generalised disturbance of water and electrolyte movement in exocrine tissues. *The Lancet* 291, 455-460.
- Jonasson, J., Olofsson, M. & Monstein, H.-J. (2002). Classification, identification and subtyping of bacteria based on pyrosequencing and signature matching of 16S rDNA fragments. *Acta Pathol Microbiol Immunol Scand* 110, 263-272.
- Jones, A. M., Dodd, M. E., Govan, J. R., Barcus, V., Doherty, C. J., Morris, J. & Webb, A. K. (2004). *Burkholderia cenocepacia* and *Burkholderia multivorans*: Influence on survival in cystic fibrosis. *Thorax* 59, 948-951.
- Jordan, W. (2005). Competitive enzyme-linked immunosorbent assay. In *Methods in Molecular Biology*, vol. 295, pp. 215-225. Edited by R. Burns. New York City: Humana Press.

- Jousimies-Somer, H. & Summanen, P. (2002). Recent taxonomic changes and terminology update of clinically significant anaerobic Gram-negative bacteria (excluding spirochetes). *Clin Infect Dis* 35, S17-S21.
- Jung, J., Jung, S., Kim, H., Yuk, J., Park, J., Kim, Y., Han, J., Kim, P. & Ha, K. (2006). High-throughput analysis of GST-fusion protein expression and activity dependant protein interaction on GST-fusion protein arrays with a spectral surface plasmon resonance biosense. *Proteomics* 6, 1110-1120.
- Jung, Y., Lee, J. M., Jung, H. & Chung, B. H. (2007). Self-directed and self-oriented immobilization of antibody by Protein GDNA conjugat e. *Anal Chem* 79, 6534-6541.
- Jyoung, J., Hong, S., Lee, W. & Choi, J. (2006). Immunosensor for the detection of *Vibrio cholerae* O1 using surface plasmon resonance. *Biosens Bioelectron* 21, 2315-2319.
- Kae, H. (2009). Genome projects: Uncovering the blueprints of biology. In *The Science Creative Quarterly*, August 2003 edn: University of British Columbia.
- Kahl, B., Everding, Anne S., Mathias Herrmann, Koch, Hans G., Becker, K., Harms, E., Proctor, Richard A. & Peters, G. (1998). Persistent infection with small colony variant strains of *Staphylococcus aureus* in patients with cystic fibrosis. *J Infect Dis* 177, 1023-1029.
- Kaplan, E., Shwachman, H., Perlmutter, A. D., Rule, A., Khaw, K. T. & Holsclaw, D. S. (1968). Reproductive failure in males with cystic fibrosis. *N Engl J Med* 279, 65-69.
- Kapral, F., Keogh, A. & Taubler, J. (1965). The nature of alpha toxin production by *Staphylococcus aureus* grown in vivo. *Proc Soc Exp Biol Med* 119, 74-77.
- Karas, M. & Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10,000 Daltons. *Anal Chem* 60, 2299-2301.
- Keeler, C., Jablonski, E. M., Albert, Y. B., Taylor, B. D., Myszka, D. G., Clevenger, C. V. & Hodsdon, M. E. (2007). The kinetics of binding human prolactin, but not growth hormone, to the prolactin receptor vary over a physiologic pH range. *Biochemistry* 46, 2398-2410.
- Keijser, B. J., Zaura, E., Huse, S. M., van der Vossen, J. M., Schuren, F. H., Montijn, R. C., ten Cate, J. M. & Crielaard, W. (2008). Pyrosequencing analysis of the oral microflora of healthy adults. *J Dent Res* 87, 1016-1020.
- Kerem, B.-S., Rommens, J. M., Buchanan, J. A., Markiewicz, D., Cox, T. K., Chakravarti, A., Buchwald, M. & Tsui, L.-C. (1989). Identification of the cystic fibrosis gene: Genetic analysis. *Science* 245, 1073-1078.



- Keys, C. J., Dare, D. J., Sutton, H., Wells, G., Lunt, M., McKenna, T., McDowall, M. & Shah, H. N. (2004). Compilation of a MALDI-TOF mass spectral database for the rapid screening and characterisation of bacteria implicated in human infectious diseases. *Infect Genet Evol* 4, 221-242.
- Kim, J. O., Garofalo, L., Blecker-Shelly, D. & McGowan, K. L. (2003). *Candida dubliniensis* infections in a pediatric population: Retrospective identification from clinical laboratory isolates of *Candida albicans*. *J Clin Microbiol* 41, 3354-3357.
- Kim, T., Oh, P. I. & Simor, A. E. (2001). The economic impact of methicillin-resistant *Staphylococcus aureus* in Canadian hospitals. *Infect Control Hosp Epidemiol* 22, 99-104.
- Klappenbach, J. A., Saxman, P. R., Cole, J. R. & Schmidt, T. M. (2001). rrndb: the ribosomal RNA operon copy number database. *Nucleic Acids Research* 29, 181-184.
- Klein, M., Neuport-Sautes, C., Ellerson, J. R. & Fridman, W. H. (1977). Binding site of human IgG subclasses and their domains for Fc receptors of activated murine T cells. *J Immunol* 119, 1077-1083.
- Klockgether, J., Wurdemann, D., Reva, O., Wiehlmann, L. & Tummler, B. (2007). Diversity of the abundant pKLC102/PAGI-2 family of genomic islands in *Pseudomonas aeruginosa*. *J Bacteriol* 189, 2443-2459.
- Knutsen, A. P., Mueller, K. R., Hutcheson, P. S. & Slavin, R. G. (1994). Serum anti-*Aspergillus fumigatus* antibodies by immunoblot and ELISA in cystic fibrosis with allergic bronchopulmonary aspergillosis. *J Allergy Clin Immunol* 93, 926-931.
- Koch, R. (1884). 2 die aetiologie der tuberkulose. *Mitt Kaiser Gesundh*, 1-88.
- Kolak, M., Karpati, F., Monstein, H.-J. & Jonasson, J. (2003). Molecular typing of the bacterial flora in sputum of cystic fibrosis patients. *IJMM* 293, 309-317.
- Kopito, L. E., Kosasky, H. J. & Shwachman, H. (1973). Water and electrolytes in cervical mucus from patients with cystic fibrosis. *Fertil Steril* 24, 512-516.
- Kotloff, R., Fitzsimmons, S. & Fiel, S. (1992). Fertility and pregnancy in patients with cystic fibrosis. *Clin Chest Med* 13 623-635.
- Koubova, V., Brynda, E., Karasova, L., Skvor, J., Homola, J., Dostalek, J., Tobiska, P. & Rosicky, J. (2001). Detection of foodborne pathogens using surface plasmon resonance biosensors. *Sensors Actuator* 74, 100-105.

- Kraut, M. S., Attebery, H. R., Finegold, S. M. & Sutter, V. L. (1972). Detection of *Haemophilus aphrophilus* in the human oral flora with a selective medium *J Infect Dis* 126, 189-192.
- Kretschmann, E. & Raether, H. (1968). Radiative decay of nonradiative surface plasmons excited by light *Z NATURFORSCH PT A* 23, 2135-2136.
- Kubak, B. M. (2002). Fungal infection in lung transplantation. *Tranpl Infect Dis* 4, 24-31.
- Kumeda, Y. & Asao, T. (1996). Single-strand conformation polymorphism analysis of PCR-amplified ribosomal DNA internal transcribed spacers to differentiate species of *Aspergillus* section Flavi. *Appl Environ Microbiol* 62, 2947-2952.
- Kunzelmann, K. (1999). The cystic fibrosis transmembrane conductance regulator and its function in epithelial transport. *Rev Physiol Biochip P* 137, 1-70.
- La Scola, B. & Raoult, D. (1998). Molecular identification of *Gemella* species from three patients with endocarditis. *J Clin Microbiol* 36, 866-871.
- Lambiase, A., Rossano, F., Del Pezzo, M., Raia, V., Sepe, A., de Gregorio, F. & Catania, M. (2009). *Sphingobacterium* respiratory tract infection in patients with cystic fibrosis. *BMC Res Notes* 2, 262.
- Landing, B. H., Wells, T. R. & Wang, C. I. (1969). Abnormality of the epididymis and vas deferens in cystic fibrosis. *Arch Pathol* 88, 569-580.
- Lauber, C. L., Hamady, M., Knight, R. & Fierer, N. (2009). Soil pH as a predictor of soil bacterial community structure at the continental scale: A pyrosequencing-based assessment. *Appl Environ Microbiol* 75, 5111-5120.
- Lazarevic, V., Whiteson, K., Huse, S., Hernandez, D., Farinelli, L., Østerås, M., Schrenzel, J. & François, P. (2009). Metagenomic study of the oral microbiota by Illumina high-throughput sequencing. *J Microbiol Meth* 79, 266-271.
- Leaw, S. N., Chang, H. C., Sun, H. F., Barton, R., Bouchara, J. P. & Chang, T. C. (2006). Identification of medically important yeast species by sequence analysis of the internal transcribed spacer regions. *J Clin Microbiol* 44, 693-699.
- Lee, J., Park, H., Yongwon, J. & Jung, S. (2007). Direct immobilization of protein G variants with various numbers of cysteine residues on a gold surface. *Anal Chem* 79, 2680-2687.
- Leininger, S., Urich, T., Schloter, M., Schwark, L., Qi, J., Nicol, G. W., Prosser, J. I., Schuster, S. C. & Schleper, C. (2006). Archaea predominate among ammonia-oxidizing prokaryotes in soils. *Nature* 442, 806-809.

- Leonard, P., Hearty, S., Wyatt, G., Quinn, J. & O'Kennedy, R. (2005). Development of a surface plasmon resonance-based immunoassay for *Listeria monocytogenes*. *J Food Prot* 68, 728-735.
- Levy, S. B. (2000). Antibiotic and antiseptic resistance: Impact on public health. *Pediatr Infect Dis J* 19, S120-S122.
- Lewin, B. (2002). Protein Synthesis. In *Genes VII*, pp. 100, 139-140, 159-164. Oxford: University Press.
- Lewin, L. O., Byard, P. J. & Davis, P. B. (1990). Effect of *Pseudomonas cepacia* colonization on survival and pulmonary function of cystic fibrosis patients. *J Clin Epidemiol* 43, 125-131.
- Li, X.-Z., Zhang, L., Srikumar, R. & Poole, K. (1998). beta-lactamase inhibitors are substrates for the multidrug efflux pumps of *Pseudomonas aeruginosa*. *Antimicrob Agents Chemother* 42, 399-403.
- Li, X. Z., Livermore, D. M. & Nikaido, H. (1994). Role of efflux pump(s) in intrinsic resistance of *Pseudomonas aeruginosa*: Resistance to tetracycline, chloramphenicol, and norfloxacin. *Antimicrob Agents Chemother* 38, 1732-1741.
- Liesack, W. & Stackebrandt, E. (1992). Unculturable microbes detected by molecular sequence and probes. *Biodiversity and Conserv* 1, 250-262.
- Liljemark, W. F. & Gibbons, R. J. (1971). Ability of *Veillonella* and *Neisseria* species to attach to oral surfaces and their proportions present indigenously. *Infect Immun* 4, 264-268.
- Logotheti, M., Kotsovili-Tseleni, A., Arsenis, G. & Legakis, N. I. (2009). Multiplex PCR for the discrimination of *A. fumigatus*, *A. flavus*, *A. niger* and *A. terreus*. *J Microbiol Meth* 76, 209-211.
- Lowbury, E. J. & Fox, J. (1954). The epidemiology of infection with *Pseudomonas pyocyanea* in a burns unit. *J Hyg (Lond)* 52, 403-416.
- Lozupone, C. & Knight, R. (2005). UniFrac: A new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* 71, 8228 - 8235.
- Lozupone, C., Hamady, M. & Knight, R. (2006). UniFrac - an online tool for comparing microbial community diversity in a phylogenetic context. *BMC Bioinformatics* 7, 371-383.
- Lozupone, C. A., Hamady, M., Kelley, S. T. & Knight, R. (2007). Quantitative and qualitative  $\beta$  diversity measures lead to different insights into factors that structure microbial communities. *Appl Environ Microbiol* 73, 1576-1585.

- Lu, H., Homola, J., Campbell, C., Nenninger, G., Yee, S. & Ratner, B. (2001). Protein contact printing for a surface plasmon resonance biosensor with on-chip referencing. *Sensors Actuat B-Chem* 74, 91-99.
- Luna, R. A., Fasciano, L. R., Jones, S. C., Boyanton, B. L., Jr., Ton, T. T. & Versalovic, J. (2007). DNA pyrosequencing-based bacterial pathogen identification in a pediatric hospital setting. *J Clin Microbiol* 45, 2985-2992.
- Lyczak, J. B., Cannon, C. L. & Pier, G. B. (2002). Lung infections associated with cystic fibrosis. *Clin Microbiol Rev* 15, 194-222.
- Mabeza, G. F. & Macfarlane, J. (2003). Pulmonary actinomycosis. *Eur Respir J* 21, 545-551.
- Magurran, A. E. & Henderson, P. A. (2003). Explaining the excess of rare species in natural species abundance distributions. *Nature* 422, 714-716.
- Maidak, B. L., Cole, J. R., Parker, J., C.T., Garrity, G. M., Larsen, N., Li, B., Lilburn, T. G., McCaughey, M. J., Olsen, G. L., Overbeek, R., Pramanik, S., Schmidt, T. M., Tiedje, J. M. & Woese, C. R. (1999). A new version of the RDP (Ribosomal Database Project). *Nucleic Acids Res* 27, 171-173.
- Makam, M., Diaz, D., Laval, J., Gernez, Y., Conrad, C. K., Dunn, C. E., Davies, Z. A., Moss, R. B., Herzenberg, L. A., Herzenberg, L. A. & Tirouvanziam, R. (2009). Activation of critical, host-induced, metabolic and stress pathways marks neutrophil entry into cystic fibrosis lungs. *Proc Natl Acad Sci U S A* 106, 5779-5783.
- Manter, D. K. & Vivanco, J. M. (2007). Use of the ITS primers, ITS1F and ITS4, to characterize fungal abundance and diversity in mixed-template samples by qPCR and length heterogeneity analysis. *J Microbiol Meth* 71, 7-14.
- Marchandin, H., Jumas-Bilak, E., Gay, B., Teyssier, C., Jean-Pierre, H., Simeon de Buochberg, M., Carriere, C. & Carlier, J.-P. (2003). Phylogenetic analysis of some Sporomusa sub-branch members isolated from human clinical specimens: description of *Megasphaera micronuciformis* sp. nov. *Int J Syst Evol Microbiol* 53, 547-553.
- Mardis, E. R. (2008). Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet* 9, 387-402.
- Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bembien, L. A., Berka, J., Braverman, M. S., Chen, Y. J., Chen, Z., Dewell, S. B., Du, L., Fierro, J. M., Gomes, X. V., Godwin, B. C., He, W., Helgesen, S., Ho, C. H., Irzyk, G. P., Jando, S. C., Alenquer, M. L., Jarvie, T. P., Jirage, K. B., Kim, J. B., Knight, J. R., Lanza, J. R., Leamon, J. H., Lefkowitz, S. M., Lei, M., Li, J., Lohman, K. L.,

- Lu, H., Makhijani, V. B., McDade, K. E., McKenna, M. P., Myers, E. W., Nickerson, E., Nobile, J. R., Plant, R., Puc, B. P., Ronan, M. T., Roth, G. T., Sarkis, G. J., Simons, J. F., Simpson, J. W., Srinivasan, M., Tartaro, K. R., Tomasz, A., Vogt, K. A., Volkmer, G. A., Wang, S. H., Wang, Y., Weiner, M. P., Yu, P., Begley, R. F. & Rothberg, J. M. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437, 376-380.
- Marin, M., Garcia de Viedma, D., Martin-Rabadan, P., Rodriguez-Creixems, M. & Bouza, E. (2000). Infection of hickman catheter by *Pseudomonas* (formerly *Flavimonas*) *oryzihabitans* traced to a synthetic bath sponge. *J Clin Microbiol* 38, 4577-4579.
- Marklein, G., Josten, M., Klanke, U., Muller, E., Horre, R., Maier, T., Wenzel, T., Kostrzewa, M., Bierbaum, G., Hoerauf, A. & Sahl, H. G. (2009). MALDI-TOF mass-spectrometry for fast and reliable identification of clinical yeast isolates. *J Clin Microbiol* 47, 2912-2917.
- Martin, K. & Rygiewicz, P. (2005). Fungal-specific PCR primers developed for analysis of the ITS region of environmental DNA extracts. *BMC Microbiology* 5, 28.
- Mastella, G., Rainisio, M., Harms, H. K., Hodson, M. E., Koch, C., Navarro, J., Strandvik, B. & G., M. S. (2000). Allergic bronchopulmonary aspergillosis in cystic fibrosis. A European epidemiological study. *Eur Respir J* 16, 464-471.
- Mathee, K., Ciofu, O., Sternberg, C., Lindum, P. W., Campbell, J. I., Jensen, P., Johnsen, A. H., Givskov, M., Ohman, D. E., Molin, S., Hoiby, N. & Kharazmi, A. (1999). Mucoid conversion of *Pseudomonas aeruginosa* by hydrogen peroxide: A mechanism for virulence activation in the cystic fibrosis lung. *Microbiology* 145 (Pt 6), 1349-1357.
- Matsui, H., Hanaki, H., Takahashi, K., Yokoyama, A., Nakae, T., Sunakawa, K. & Omura, S. (2009). Rapid detection of vaginal *Candida* species by newly developed immunochromatography. *Clin Vaccine Immunol* 16, 1366-1368.
- Matsumoto, K., Torimaru, A., Ishitobi, S., Sakai, T., Ishikawa, H., Toko, K., Miura, N. & Imato, T. (2005). Preparation and characterization of a polyclonal antibody from rabbit for detection of trinitrotoluene by a surface plasmon resonance biosensor. *Talanta* 68, 305-311.
- Mayo, C. S. & Hallock, R. B. (1989). Immunoassay based on surface plasmon oscillations. *J Immunol Methods* 120, 105-114.
- McArthur, R. H. (1957). On the relative abundance of bird species. *Proc Natl Acad Sci U S A* 43, 293-295.

- McDowell, A., Mahenthiralingam, E., Moore, J. E., Dunbar, K. E. A., Webb, A. K., Dodd, M. E., Martin, S. L., Millar, B. C., Scott, C. J., Crowe, M. & Elborn, J. S. (2001). PCR-based detection and identification of *Burkholderia cepacia* complex pathogens in sputum from cystic fibrosis patients. *J Clin Microbiol* 39, 4247-4255.
- McGowan Jr, J. E. (2006). Resistance in nonfermenting gram-negative bacteria: Multidrug resistance to the maximum. *Am J Infect Control* 34, S29-S37.
- McKernan, K. J., Peckham, H. E., Costa, G. L., McLaughlin, S. F., Fu, Y., Tsung, E. F., Clouser, C. R., Duncan, C., Ichikawa, J. K., Lee, C. C., Zhang, Z., Ranade, S. S., Dimalanta, E. T., Hyland, F. C., Sokolsky, T. D., Zhang, L., Sheridan, A., Fu, H., Hendrickson, C. L., Li, B., Kotler, L., Stuart, J. R., Malek, J. A., Manning, J. M., Antipova, A. A., Perez, D. S., Moore, M. P., Hayashibara, K. C., Lyons, M. R., Beaudoin, R. E., Coleman, B. E., Laptewicz, M. W., Sannicandro, A. E., Rhodes, M. D., Gottimukkala, R. K., Yang, S., Bafna, V., Bashir, A., MacBride, A., Alkan, C., Kidd, J. M., Eichler, E. E., Reese, M. G., De La Vega, F. M. & Blanchard, A. P. (2009). Sequence and structural variation in a human genome uncovered by short-read, massively parallel ligation sequencing using two-base encoding. *Genome Res* 19, 1527-1541.
- McMenamin, J. D., Zacccone, T. M., Coenye, T., Vandamme, P. & LiPuma, J. J. (2000). Misidentification of *Burkholderia cepacia* in US cystic fibrosis treatment centers: an analysis of 1,051 recent sputum isolates. *Chest* 117, 1661-1665.
- Mearns, M., Longbottom, J. & Batten, J. (1967). Precipitating antibodies to *Aspergillus fumigatus* in cystic fibrosis. *Lancet* 1, 538-539.
- Meeusen, C., Alocilja, E. & Osburn, W. (2005). Detection of *E. coli* O157:H7 using a miniaturized surface plasmon resonance biosensor. *Trans Asae* 48, 2409-2416.
- Melchers, W. J., Verweij, P. E., van den Hurk, P., van Belkum, A., De Pauw, B. E., Hoogkamp-Korstanje, J. A. & Meis, J. F. (1994). General primer-mediated PCR for detection of *Aspergillus* species. *J Clin Microbiol* 32, 1710-1717.
- Mentasti, M., Morelli, P., Melioli, G. & Manno, G. (2008). Molecular epidemiology of *Stenotrophomonas maltophilia* (SM) in the cystic fibrosis (CF) centre of Genova. *J Cyst Fibros* 7, S13.
- Mercier, B., Verlingue, C., Lissens, W., Silber, S. J., Novelli, G., Bonduelle, M., Audrezet, M. P. & Ferec, C. (1995). Is congenital bilateral absence of vas deferens a primary form of cystic fibrosis? Analyses of the CFTR gene in 67 patients. *Am J Hum Genet* 56, 272-277.
- Merlino, J., Siarakas, S., Robertson, G. J., Funnell, G. R., Gottlieb, T. & Bradbury, R. (1996). Evaluation of CHROMagar Orientation for differentiation and

- presumptive identification of gram-negative bacilli and *Enterococcus* species. *J Clin Microbiol* 34, 1788-1793.
- Miles, L. E. & Hales, C. N. (1968). Labelled antibodies and immunological assay systems. *Nature* 219, 186-189.
- Miller, S. Z. (2003). *Molecular identification of microbes in the sputum of cystic fibrosis patients*. Masters in Forensic Science thesis, Florida International University, Miami.
- Mills, D. K., Fitzgerald, K., Litchfield, C. D. & Gillevet, P. M. (2003). A comparison of DNA profiling techniques for monitoring nutrient impact on microbial community composition during bioremediation of petroleum-contaminated soils. *J Microbiol Meth* 54, 57-74.
- Mills, D. K., Entry, J. A., Voss, J. D., Gillevet, P. M. & Mathee, K. (2006). An assessment of the hypervariable domains of the 16S rRNA genes for their value in determining microbial community diversity: The paradox of traditional ecological indices. *FEMS Microbiol Ecol* 57, 496-503.
- Mohammed, A. M., Yong-Seung, S., Shaofeng, Z., Richard, M. & Franco, B. (2007). Desorption electrospray ionization mass spectrometry of intact bacteria. *J Mass Spectrom* 42, 1186-1193.
- Monstein, H.-J., Nikpour-Badr, S. & Jonasson, J. (2001). Rapid molecular identification and subtyping of *Helicobacter pylori* by pyrosequencing of the 16S rDNA variable V1 and V3 regions. *FEMS Microbiol Lett* 199, 103-107.
- Moreno, L. I., Mills, D. K., Entry, J., Sautter, R. T. & Mathee, K. (2006). Microbial metagenome profiling using amplicon length heterogeneity-polymerase chain reaction proves more effective than elemental analysis in discriminating soil specimens. *Journal of Forensic Sci* 51, 1315-1322.
- Muder, R. R., Harris, A. P., Muller, S., Edmond, M., Chow, J. W., Papadakis, K., Wagener, M. W., Bodey, G. P. & Steckelberg, J. M. (1996). Bacteremia due to *Stenotrophomonas (Xanthomonas) maltophilia*: A prospective, multicenter study of 91 episodes. *Clin Infect Dis* 22, 508-512.
- Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. & Erlich, H. (1986). Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction. *Cold Spring Harb Symp Quant Biol* 51, 263-273.
- Murray, P. R. & Baron, E. J. (2007). *Manual of Clinical Microbiology*, 9th edn. Washington, D.C.: ASM Press.

- Nagano, Y., Millar, B. C., Johnson, E., Goldsmith, C., Elborn, J. S., Rendall, J. & Moore, J. E. (2007). Fungal infections in patients with cystic fibrosis. *Rev Med Microbiol* 18, 11-17.
- Naimushin, A., Soelberg, S., Bartholomew, D., Elkind, J. & Furlong, C. (2003). A portable surface plasmon resonance (SPR) sensor system with temperature regulation. *Sensors Actuat B-Chem* 96, 253-260.
- Nassif, X. (2009). Editorial commentary: A revolution in the identification of pathogens in clinical laboratories. *Clin Infect Dis* 49, 552-553.
- Niel, C. B. V. (1943). Biochemistry of microorganisms. *Annu Rev Biochem* 12, 551-586.
- Niilo, L. (1959). Some observations on *Pseudomonas* infections in poultry. *Can J Comp Med Vet Sci* 23, 329-330, 335-337.
- Nisonoff, A., Wissler, F. C., Lipman, L. N. & Woernley, D. L. (1960). Separation of univalent fragments from the bivalent rabbit antibody molecule by reduction of disulfide bonds. *Archives of Biochem Biophys* 89, 230-244.
- Noble, W. C. (1998). Skin bacteriology and the role of *Staphylococcus aureus* in infection. *Br J Dermatol* 139 Suppl 53, 9-12.
- Nomura, M. & Erdmann, V. A. (1970). Reconstitution of 50S ribosomal subunits from dissociated molecular components. *Nature* 228, 744-748.
- Nomura, M., Morgan, E. A. & Jaskunas, S. R. (1977). Genetics of bacterial ribosomes. *Annu Rev Genet* 11, 297-347.
- Noskin, G. A., Rubin, R. J., Schentag, J. J., Kluytmans, J., Hedblom, E. C., Smulders, M., Lapetina, E. & Gemmen, E. (2005). The burden of *Staphylococcus aureus* infections on hospitals in the United States: An analysis of the 2000 and 2001 nationwide inpatient sample database. *Arch Intern Med* 165, 1756-1761.
- Nossal, G. J. V. & Ada, G. L. (1971). *Antigens, lymphoid cells, and the immune response*. New York: Academic Press.
- Novotný, J., Bruccoleri, R., Newell, J., Murphy, D., Haber, E. & Karplus, M. (1983). Molecular anatomy of the antibody binding site. *J Biol Chem* 258, 14433-14437.
- O'Hara, C. M. & Miller, J. M. (2003). Evaluation of the Vitek 2 ID-GNB assay for identification of members of the family Enterobacteriaceae and other nonenteric gram-negative bacilli and comparison with the Vitek GNI+ card. *J Clin Microbiol* 41, 2096-2101.



- Odds, F. C. & Bernaerts, R. (1994). CHROMagar *Candida*, a new differential isolation medium for presumptive identification of clinically important *Candida* species. *J Clin Microbiol* 32, 1923-1929.
- Oh, B., Lee, W., Lee, W. & Choi, J. (2003). Nano-scale probe fabrication using self-assembly technique and application to detection of *Escherichia coli* O157:H7. *Biotechnol Bioprocess Eng* 8, 227-232.
- Oh, B., Kim, Y., Park, K., Lee, W. & Choi, J. (2004a). Surface plasmon resonance immunosensor for the detection of *Salmonella typhimurium*. *Biosens Bioelectron* 19, 1497-1504.
- Oh, B., Lee, W., Kim, Y., Lee, W. & Choi, J. (2004b). Surface plasmon resonance immunosensor using self-assembled Protein G for the detection of *Salmonella paratyphi*. *J Biotech* 111, 1-8.
- Oh, B., Lee, W., Chun, B., Bae, Y., Lee, W. & Choi, J. (2005). Surface plasmon resonance immunosensor for the detection of *Yersinia enterocolitica*. *Colloids Surf A* 257-258.
- Ojeda-Vargas, M., Pacheco, A., Elia, M., Villaverde, R. & Baquero, F. (1990). *Proteus mirabilis* as a cause of recurrent lung infection in a cystic fibrosis patient. *Eur J Clin Microbiol Infect Dis* 9, 234-235.
- Olansky, S. & McCormick, G. J. (1958). Triacetyloleandomycin: Its use in the treatment of acne and pyodermas caused by resistant *Staphylococcus aureus*. *Antibiot Annu* 6, 265-267.
- Olsen, G. J. & Woese, C. R. (1993). Ribosomal RNA: A key to phylogeny. *FASEB Journal* 7, 113-123.
- Oppenheimer, E. A., Case, A. L., Esterly, J. R. & Rothberg, R. M. (1970). Cervical mucus in cystic fibrosis: A possible cause of infertility. *Am J Obstet Gynecol* 108, 673-674.
- Oppenheimer, E. H. & Esterly, J. R. (1969). Observations on cystic fibrosis of the pancreas. V. Developmental changes in the male genital system. *J Pediatr* 75, 806-811.
- Oppenheimer, E. H. & Esterly, J. R. (1970). Observations on cystic fibrosis of the pancreas. VI. The uterine cervix. *J Pediatr* 77, 991-995.
- Otto, A. (1968). Excitation of nonradiative surface plasma waves in silver by the method of frustrated total reflection. *Z PHYS A-HADRON NUCL* 216, 398-410.
- Paju, S. & Scannapieco, F. (2007). Oral biofilms, periodontitis, and pulmonary infections. *Oral Dis* 13, 508-512.

- Pankiewicz, R., Remlein-Starosta, D., Schroeder, G. & Brzezinski, B. (2006). Biological activity and ESI MS study of oxaalkyl and hydroksyoaalkyl lasalocid esters. *J Mol Struct* 783, 136-144.
- Pape, J., Wadlin, J. & Nachamkin, I. (2006). Use of BBL CHROMagar MRSA medium for identification of methicillin-resistant *Staphylococcus aureus* directly from blood cultures. *J Clin Microbiol* 44, 2575-2576.
- Paradowski, L. J. (1997). Saprophytic fungal infections and lung transplantation--revisited. *J Heart Lung Transplant* 16, 524-531.
- Park, S., Kim, H., Paek, S., Hong, J. W. & Kim, Y. K. (2008). Enzyme-linked immunostrip biosensor to detect *Escherichia coli* O157:H7. *Ultramicroscopy* 108, 1348-1351.
- Parker, J. (1924). The production of an exotoxin by certain strains of *Staphylococcus aureus*. *J Exp Med* 40, 761-772.
- Patel, J. B. (2001). 16S rRNA gene sequencing for bacterial pathogen identification in the clinical laboratory. *Mol Diagn* 6, 313-321.
- Patient-registry (2008). Cystic fibrosis foundation patient registry annual data report: Cystic Fibrosis Foundation.
- Patrick, K. (2007). 454 life sciences: Illuminating the future of genome sequencing and personalized medicine. *Yale J Biol Med* 80, 191-194.
- Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Phil Mag (Series 5)* 50, 157-175.
- Pedersen, S. S., Espersen, F. & Hoiby, N. (1987). Diagnosis of chronic *Pseudomonas aeruginosa* infection in cystic fibrosis by enzyme-linked immunosorbent assay. *J Clin Microbiol* 25, 1830-1836.
- Pernis, B., Chiappino, G. & Rowe, D. S. (1966). Cells producing IgD immunoglobulins in human spleen. *Nature* 211, 424-425.
- Phillips, I., King, A. & Shannon, K. (2000). Comparative in-vitro properties of the quinolones. In *The quinolones*, pp. 99-137. Edited by V. Andriole. San Diego: Academic Press.
- Pierre-Audigier, C., Ferroni, A., Sermet-Gaudelus, I., Le Bourgeois, M., Offredo, C., Vu-Thien, H., Fauroux, B., Mariani, P., Munck, A., Bingen, E., Guillemot, D., Quesne, G., Vincent, V., Berche, P. & Gaillard, J. L. (2005). Age-related prevalence and distribution of nontuberculous mycobacterial species among patients with cystic fibrosis. *J Clin Microbiol* 43, 3467-3470.

- Porter, R. R. (1959). The hydrolysis of rabbit  $\gamma$ -globulin and antibodies with crystalline papain. *Biochem J* 73, 119-126.
- Price, L. B., Liu, C. M., Melendez, J. H., Frankel, Y. M., Engelthaler, D., Aziz, M., Bowers, J., Rattray, R., Ravel, J., Kingsley, C., Keim, P. S., Lazarus, G. S. & Zenilman, J. M. (2009). Community analysis of chronic wound bacteria using 16S rRNA gene-based pyrosequencing: Impact of diabetes and antibiotics on chronic wound microbiota. *PLoS ONE* 4, e6462.
- Przyklenk, B., Bauernfeind, A., Horl, G. & Emminger, G. (1987). Serologic response to *Candida albicans* and *Aspergillus fumigatus* in cystic fibrosis. *Infection* 15, 308-310.
- Quince, C., Lanzen, A., Curtis, T. P., Davenport, R. J., Hall, N., Head, I. M., Read, L. F. & Sloan, W. T. (2009). Accurate determination of microbial diversity from 454 pyrosequencing data. *Nat Methods* 6, 639-641.
- Ramani, R., Gromadzki, S., Pincus, D. H., Salkin, I. F. & Chaturvedi, V. (1998). Efficacy of API 20C and ID 32C systems for identification of common and rare clinical yeast isolates. *J Clin Microbiol* 36, 3396-3398.
- Ramsey, B. W., Pepe, M. S., Quan, J. M., Otto, K. L., Montgomery, A. B., Williams-Warren, J., Vasiljev, K. M., Borowitz, D., Bowman, C. M., Marshall, B. C., Marshall, S. & Smith, A. L. (1999). Intermittent administration of inhaled tobramycin in patients with cystic fibrosis. Cystic fibrosis inhaled tobramycin study group. *N Engl J Med* 340, 23-30.
- Rantakokko-Jalava, K., Elo-Lehtonen, E. & Meurman, O. (2006). Comparison of workflow and accuracy of identification and antimicrobial susceptibility testing of clinical isolates of *Enterobacteriaceae*, *Pseudomonas aeruginosa* and enterococci by Vitek 2 and routine methods. *APMIS* 114, 43-49.
- Reddy, L. V., Kumar, A. & Kurup, V. P. (1993). Specific amplification of *Aspergillus fumigatus* DNA by polymerase chain reaction. *Mol Cell Probe* 7, 121-126.
- Rees, G. N., Baldwin, D. S., Watson, G. O., Perryman, S. & Nielsen, D. L. (2004). Ordination and significance testing of microbial community composition derived from terminal restriction fragment length polymorphisms: Application of multivariate statistics. *Antonie Leeuwenhoek* 86, 339-347.
- Reina, J., Borrell, N. & Figuerola, J. (1992). *Sphingobacterium multivorum* isolated from a patient with cystic fibrosis. *Eur J Clin Microbiol Infect Dis* 11, 81-82.
- Reischl, U., Pulz, M., Ehret, W. & Wolf, H. (1994). PCR based detection of *Mycobacterium* in sputum samples using a simple and reliable DNA extraction protocol. *Biotechniques* 17, 844-845.

- Retsema, J., Girard, A., Schelkly, W., Manousos, M., Anderson, M., Bright, G., Borovoy, R., Brennan, L. & Mason, R. (1987). Spectrum and mode of action of azithromycin (CP-62,993), a new 15-membered-ring macrolide with improved potency against gram-negative organisms. *Antimicrob Agents Chemother* 31, 1939-1947.
- Ringen, L. M. & Drake, C. H. (1952). A study of the incidence of *Pseudomonas aeruginosa* from various natural sources. *J Bacteriol* 64, 841-845.
- Riordan, J. R., Rommens, J. M., Kerem, B.-S., Alon, N., Rozmahel, R., Grzelczak, Z., Zielenski, J., Lok, S., Plavsic, N., Chou, J.-L., Drumm, M. L., Iannuzzi, M. C., Collins, F. & Tsui, L.-C. (1989). Identification of the cystic fibrosis gene: Cloning and characterization of complementary DNA. *Science* 245, 1066-1068.
- Ritchie, N. J., Schutter, M. E., Dick, R. P. & Myrold, D. D. (2000). Use of length heterogeneity PCR and fatty acid methyl ester profiles to characterize microbial communities in soil. *Appl Environ Microbiol* 66, 1668-1675.
- Rogers, C. S., Abraham, W. M., Brogden, K. A., Engelhardt, J. F., Fisher, J. T., McCray, P. B., Jr., McLennan, G., Meyerholz, D. K., Namati, E., Ostedgaard, L. S., Prather, R. S., Sabater, J. R., Stoltz, D. A., Zabner, J. & Welsh, M. J. (2008). The porcine lung as a potential model for cystic fibrosis. *Am J Physiol Lung Cell Mol Physiol* 295, L240-263.
- Rogers, G., Daniels, T., Tuck, A., Carroll, M., Connett, G., David, G. & Bruce, K. (2009). Studying bacteria in respiratory specimens by using conventional and molecular microbiological approaches. *BMC Pulm Med* 9, 14.
- Rogers, G. B., Hart, C. A., Mason, J. R., Hughes, M., Walshaw, M. J. & Bruce, K. D. (2003). Bacterial diversity in cases of lung infection in cystic fibrosis patients: 16S ribosomal DNA (rDNA) length heterogeneity PCR and 16S rDNA terminal restriction fragment length polymorphism profiling. *J Clin Microbiol* 41, 3548-3558.
- Rogers, G. B., Carroll, M. P., Serisier, D. J., Hockey, P. M., Jones, G. & Bruce, K. D. (2004). Characterization of bacterial community diversity in cystic fibrosis lung infections by use of 16S ribosomal DNA terminal restriction fragment length polymorphism profiling. *J Clin Microbiol* 42, 5176-5183.
- Rogers, G. B., Carroll, M. P., Serisier, D. J., Hockey, P. M., Jones, G., Kehagia, V., Connett, G. J. & Bruce, K. D. (2006). Use of 16S rRNA gene profiling by terminal restriction fragment length polymorphism analysis to compare bacterial communities in sputum and mouthwash samples from patients with cystic fibrosis. *J Clin Microbiol* 44, 2601-2604.

- Rogers, G. B., Skelton, S., Serisier, D. J., van der Gast, C. J. & Bruce, K. D. (2010). Determining cystic fibrosis-affected lung microbiology: Comparison of spontaneous and serially induced sputum samples by use of terminal restriction fragment length polymorphism profiling. *J Clin Microbiol* 48, 78-86.
- Rommens, J. M., Iannuzzi, M. C., Kerem, B., Drumm, M. L., Melmer, G., Dean, M., Rozmahel, R., Cole, J. L., Kennedy, D., Hidaka, N. & et al. (1989). Identification of the cystic fibrosis gene: chromosome walking and jumping. *Science* 245, 1059-1065.
- Ronaghi, M., Uhlen, M. & Nyren, P. (1998). A sequencing method based on real-time pyrophosphate. *Science* 281, 363-365.
- Rong-Hwa, S., Shiao-Shek, T., Der-Jiang, C. & Yao-Wen, H. (2010). Gold nanoparticle-based lateral flow assay for detection of staphylococcal enterotoxin B. *Food Chem* 118, 462-466.
- Rozzak, D. B. & Colwell, R. R. (1987). Survival strategies of bacteria in the natural environment. *Microbiol Res* 51, 365-379.
- Rothberg, J. M. & Leamon, J. H. (2008). The development and impact of 454 sequencing. *Nat Biotech* 26, 1117-1124.
- Rowen, J. L., Tate, J. M., Nordoff, N., Passarell, L. & McGinnis, M. R. (1999). *Candida* isolates from neonates: Frequency of misidentification and reduced fluconazole susceptibility. *J Clin Microbiol* 37, 3735-3737.
- Rudi, K., Zimonja, M., Trosvik, P. & Næs, T. (2007). Use of multivariate statistics for 16S rRNA gene analysis of microbial communities. *Int J Food Prop* 120, 95-99.
- Saeed, E. N. & Hay, R. J. (1981). Immunoperoxidase staining in the recognition of *Aspergillus* infections. *Histopathology* 5, 437-444.
- Sakaguchi, M., Inouye, S., Miyazawa, H. & Tamura, S. (1989). Measurement of antigen-specific mouse IgE by a fluorometric reverse (IgE-capture) ELISA. *J Immunol Methods* 116, 181-187.
- Sanders, S., Kubagawa, H., Suzuki, T., Butler, J. & Cooper, M. (1987). IgM binding protein expressed by activated B cells. *J Immunol* 139, 188-193.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74, 5463-5467.
- Scannapieco, F. A. (2006). Pneumonia in nonambulatory patients: The role of oral bacteria and oral hygiene. *J Am Dent Assoc* 137, 21S-25.

- Schasfoort, R. & McWhirter, A. (2008). SPR instrumentation In *Handbook of surface plasmon resonance*, pp. 35-80. Edited by R. Schasfoort & A. Tudos. Cambridge: The Royal Society of Chemistry.
- Scheld, W. M. (2003). Maintaining fluoroquinolone class efficacy: Review of influencing factors. *Emerg Infect Dis* 9, 1-9.
- Schmid, A., Stanca, S., Thakur, M., Ravindranathan Thampi, K. & Raman Suri, C. (2006). Site-directed antibody immobilization on gold substrate for surface plasmon resonance sensors. *Sensors Actuat B-Chem* 113, 297-303.
- Schmidt, D. & Rath, P.-M. (2003). Faster genetic identification of medically important aspergilli by using gellan gum as gelling agent in mycological media. *J Med Microbiol* 52, 653-655.
- Schneck, H. (1957). Recurrent *Staphylococcus aureus* pneumonia and empyema treated with novobiocin; a review of the literature, with report of a case. *J Pediatr* 50, 570-578.
- Schuster, S. C. (2008). Next-generation sequencing transforms today's biology. *Nature Methods* 5, 16-18.
- Schwarz, J. (1982). The diagnosis of deep mycoses by morphologic methods. *Hum Pathol* 13, 519-533.
- Seng, P., Drancourt, M., Gouriet, F., La Scola, B., Fournier, P. E., Rolain, Jean M. & Raoult, D. (2009). Ongoing revolution in bacteriology: Routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Clin Infect Dis* 49, 543-551.
- Shade, A., Jones, S. E. & McMahon, K. D. (2008). The influence of habitat heterogeneity on freshwater bacterial community composition and dynamics. *Environ Microbiol* 10, 1057-1067.
- Shah, A., Panjabi, C., Nair, V., Chaudhry, R. & Thukral, S. S. (2008). *Veillonella* as a cause of chronic anaerobic pneumonitis. *Int J Infect Dis* 12, e115-e117.
- Shelly, D. B., Spilker, T., Gracely, E. J., Coenye, T., Vandamme, P. & LiPuma, J. J. (2000). Utility of commercial systems for identification of *Burkholderia cepacia* complex from cystic fibrosis sputum culture. *J Clin Microbiol* 38, 3112-3115.
- Shyu, R.-H., Shyu, H.-F., Liu, H.-W. & Tang, S.-S. (2002). Colloidal gold-based immunochromatographic assay for detection of ricin. *Toxicon* 40, 255-258.
- Sibley, C. D., Parkins, M. D., Rabin, H. R., Duan, K., Norgaard, J. C. & Surette, M. G. (2008). A polymicrobial perspective of pulmonary infections exposes an

- enigmatic pathogen in cystic fibrosis patients. *Proc Natl Acad Sci U S A* 105, 15070-15075.
- Sibley, C. D., Grinwis, M., Field, T. R., Parkins, M. D., Noorgard, J. C., Gregson, D. B., Rabin, H. R. & Surette, M. (2010). McKay agar enables routine quantification of the *Streptococcus milleri* group in cystic fibrosis patients. *J Med Microbiol* 48, 395-401.
- Simek, K., Pernthaler, J., Weinbauer, M. G., Hornak, K., Dolan, J. R., Nedoma, J., Masin, M. & Amann, R. (2001). Changes in bacterial community composition and dynamics and viral mortality rates associated with enhanced flagellate grazing in a mesoeutrophic reservoir. *Appl Environ Microbiol* 67, 2723-2733.
- Sixou, M., Duffaut-Lagarrigue, D. & Lodter, J. P. (1991). The distribution and prevalence of *Haemophilus actinomycetemcomitans* in the oral cavity. *J Biol Buccale* 19, 221-228.
- Skottrup, P. D., Nicolaisen, M. & Justesen, A. F. (2008). Towards on-site pathogen detection using antibody-based sensors. *Biosens Bioelectron* 24, 339-348.
- Smith, E. E., Buckley, D. G., Wu, Z., Saenphimmachak, C., Hoffman, L. R., D'Argenio, D. A., Miller, S. I., Ramsey, B. W., Speert, D. P., Moskowitz, S. M., Burns, J. L., Kaul, R. & Olson, M. V. (2006). Genetic adaptation by *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients. *Proc Natl Acad Sci U S A* 103, 8487-8492.
- Sogin, M. L., Morrison, H. G., Huber, J. A., Mark Welch, D., Huse, S. M., Neal, P. R., Arrieta, J. M. & Herndl, G. J. (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc Natl Acad Sci U S A* 103, 12115-12120.
- Song, Y., Talaty, N., Tao, W. A., Pan, Z. & Cooks, R. G. (2007). Rapid ambient mass spectrometric profiling of intact, untreated bacteria using desorption electrospray ionization. *Chem Commun*, 61-63.
- Sorensen, T. (1948). A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. *Kongelige Danske Videnskabernes Selskab* 5, 1-34.
- Spear, G. T., Sikaroodi, M., Zariffard, M. R., Landay, A. L., French, A. L. & Gillevet, P. M. (2008). Comparison of the diversity of the vaginal microbiota in HIV-infected and HIV-uninfected women with or without bacterial vaginosis. *J Infect Dis* 198, 1131-1140.

- Spicuzza, L., Sciuto, C., Vitaliti, G., Di Dio, G., Leonardi, S. & La Rosa, M. (2009). Emerging pathogens in cystic fibrosis: Ten years of follow-up in a cohort of patients. *Eur J Clin Microbiol Infect Dis* 28, 191-195.
- Spilker, T., Coenye, T., Vandamme, P. & LiPuma, J. J. (2004). PCR-based assay for differentiation of *Pseudomonas aeruginosa* from other *Pseudomonas* species recovered from cystic fibrosis patients. *Journal of Clinical Microbiology* 42, 2074-2079.
- Srilatha, N. & Murthy, G. (2006). Electrostatic interaction in BIAcore binding studies: A cause for anomaly. *Curr Sci* 90, 677-682.
- Sriramulu, D. D., Lunsdorf, H., Lam, J. S. & Romling, U. (2005). Microcolony formation: A novel biofilm model of *Pseudomonas aeruginosa* for the cystic fibrosis lung. *J Med Microbiol* 54, 667-676.
- Staley, J. T. & Konopka, A. (1985). Measurement of in situ activities of nonphotosynthetic microorganisms in aquatic and terrestrial habitats. *Annu Rev Microbiol* 39, 321-346.
- Stanworth, D. R. (1969). Immunochemistry of IgE. *Proc R Soc Med* 62, 971-974.
- Staros, J., Wright, R. & Swingle, D. (1986). Enhancement by N-hydroxysulfosuccinimide. *Anal Biochem* 156, 220-222.
- Stepán, J., Pantucek, R., Ruzicková, V., Rosypal, S., Hájek, V. & Doskar, J. (2001). Identification of *Staphylococcus aureus* based on PCR amplification of species specific genomic 826 bp sequence derived from a common 44-kb Sma I restriction fragment. *Mol Cell Probe* 15, 249-257.
- Stettner, J., Frank, P., Griesser, T., Trimmel, G., Schennach, R., Resel, R. & Winkler, A. (2009). Characterization of 11-MUA SAM formation on gold surfaces. In *Interface Controlled Organic Thin Films*, vol. 129, pp. 101-105. Berlin Heidelberg: Springer.
- Stevens, D., Moss, R., Kurup, V. P., Knutsen, A. P., Greenberger, P., Judson, M., Denning, D., Cramer, R., Brody, A., Light, M., Skov, M., Maish, W., Mastella, G. & the cystic fibrosis foundation consensus conference (2003). Allergic bronchopulmonary aspergillosis in cystic fibrosis-state of the art: Cystic fibrosis foundation consensus conference. *Clin Infect Dis* 37, Suppl 3, S225-264.
- Stralin, K., Backman, A., Holmberg, H., Fredlund, H. & Olcen, P. (2005). Design of a multiplex PCR for *Streptococcus pneumoniae*, *Haemophilus influenzae*, *Mycoplasma pneumoniae* and *Chlamydia pneumoniae* to be used on sputum samples. *APMIS* 113, 99-111.



- Su, X. & Li, Y. (2005). Surface plasmon resonance and quartz crystal microbalance immunosensors for detection of *Escherichia coli* O157:H7. *Trans Asae* 48, 405-413.
- Su, Y., Duan, J. & Wu, W. (2005). Selectivity and specificity of a chromogenic medium for detecting *Vibrio parahaemolyticus*. *J Food Prot* 68, 1454-1456.
- Subhash, K. M. (2004). Beyond bacteria: Interpreting fungal elements in the Gram stain. *Clin Microbiol Newsletter* 26, 108-112.
- Subramanian, A., Irudayaraj, I. & Ryan, T. (2006a). A mixed self-assembled monolayer-based surface plasmon immunosensor for detection for *E. coli* O157:H7 in apple juice. *Trans Asae* 49, 1257-1262.
- Subramanian, A. & Irudayaraj, J. (2006). Surface plasmon resonance based immunosensing of *E. coli* O157:H7 in apple juice. *Trans Asabe* 49, 1257-1262.
- Subramanian, A., Irudayaraj, J. & Ryan, T. (2006b). Mono and dithiol surfaces on surface plasmon resonance biosensors for detection of *Staphylococcus aureus*. *Sensors Actuat B-Chem* 114, 192-198.
- Suzuki, M., Rappe, M. S. & Giovannoni, S. J. (1998). Kinetic bias in estimates of coastal picoplankton community structure obtained by measurements of small-subunit rRNA gene PCR amplicon length heterogeneity. *Appl Environ Microbiol* 64, 4522-4529.
- Suzuki, M. T., Rappe, M. S., Haimberger, Z. W., Winfield, H., Adair, N., Strobel, J. & Giovannoni, S. J. (1997). Bacterial diversity among small-subunit rRNA gene clones and cellular isolates from the same seawater sample. *Appl Environ Microbiol* 63, 983-989.
- Takahashi, N., Tetaert, D., Debuire, B., Lin, L. C. & Putnam, F. W. (1982). Complete amino acid sequence of the delta heavy chain of human immunoglobulin D. *Proc Natl Acad Sci U S A* 79, 2850-2854.
- Tang, Y.-W. & Stratton, C. W. (2006). *Advanced techniques in diagnostic microbiology*. New York, N.Y.: Springer.
- Taylor, A., Yu, Q., Chen, S., Homola, J. & Jiang, S. (2005). Comparison of *E. coli* O157:H7 preparation methods used for detection with surface plasmon resonance sensor. *Sensors Actuat B-Chem* 107, 202-208.
- Taylor, A., Ladd, J., Yu, Q., Chen, S., Homola, J. & Jiang, S. (2006). Quantitative and simultaneous detection of four foodborne bacterial pathogens with a multi-channel SPR sensor. *Biosens Bioelectron* 22, 752-758.

- Tokeshi, M. (1993). Species abundance patterns and community structure. In *Advances in Ecological Research*, vol. 24, pp. 111-179. Edited by M. Begon & A. H. Fitter. San Diego: Academic Press Inc.
- Tomasi, T. B. (1973). Production of a noncovalently bonded pentamer of immunoglobulin M: Relationship to J Chain. *Proc Natl Acad Sci U S A* 70, 3410-3414.
- Tsai, W.-C. & Li, I.-C. (2009). SPR-based immunosensor for determining staphylococcal enterotoxin A. *Sensors Actuat B-Chem* 136, 8-12.
- Tunney, M. M., Field, T. R., Moriarty, T. F., Patrick, S., Doering, G., Muhlebach, M. S., Wolfgang, M. C., Boucher, R., Gilpin, D. F., McDowell, A. & Elborn, J. S. (2008). Detection of anaerobic bacteria in high numbers in sputum from patients with cystic fibrosis. *Am J Respir Crit Care Med* 177, 995-1001.
- Tuohy, M. J., Hall, G. S., Sholtis, M. & Procop, G. W. (2005). Pyrosequencing<sup>(TM)</sup> as a tool for the identification of common isolates of *Mycobacterium* sp. *Diagn Microb Infect Dis* 51, 245-250.
- Turnbaugh, P., Hamady, M., Yatsunenko, T., Cantarel, B., Duncan, A., Ley, R., Sogin, M., Jones, W., Roe, B., Affourtit, J., Egholm, M., Henrissat, B., Heath, A., Knight, R. & Gordon, J. (2009). A core gut microbiome in obese and lean twins. *Nature* 457, 480 - 484.
- Urich, T., Lanzen, A., Qi, J., Huson, D. H., Schleper, C. & Schuster, S. C. (2008). Simultaneous assessment of soil microbial community structure and function through analysis of the meta-transcriptome. *PLoS ONE* 3, e2527.
- Uy, H. S., Leuenberger, E. U., de Guzman, B. B. & Natividad, F. F. (2007). Chronic, postoperative *Pseudomonas luteola* endophthalmitis. *Ocul Immunol Inflamm* 15, 359-361.
- Vaidyanathan, S., Rowland, J. J., Kell, D. B. & Goodacre, R. (2001). Discrimination of aerobic endospore-forming bacteria via electrospray-ionization mass spectrometry of whole cell suspensions. *Anal Chem* 73, 4134-4144.
- Vaidyanathan, S., Kell, D. B. & Goodacre, R. (2002). Flow-injection electrospray ionization mass spectrometry of crude cell extracts for high-throughput bacterial identification. *J Am Soc Mass Spectr* 13, 118-128.
- Vaiukaitis, J., Braunstein, G. & Ross, G. (1972). A radioimmunoassay which specifically measures human chorionic gonadotropin in the presence of human luteinizing hormone. *Am J Obstet Gynecol* 113, 751-758.

- Valipour, A., Koller, H., Setinek, U. & Burghuber, O. C. (2005). Pleural empyema associated with *Gemella morbillorum*: Report of a case and review of the literature. *Scand J Infect Dis* 37, 378-381.
- Valouev, A., Ichikawa, J., Tonthat, T., Stuart, J., Ranade, S., Peckham, H., Zeng, K., Malek, J. A., Costa, G., McKernan, K., Sidow, A., Fire, A. & Johnson, S. M. (2008). A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. *Genome Res* 18, 1051-1063.
- van Belkum, A., Renders, N. H., Smith, S., Overbeek, S. E. & Verbrugh, H. A. (2000). Comparison of conventional and molecular methods for the detection of bacterial pathogens in sputum samples from cystic fibrosis patients. *FEMS Immunol Med Microbiol* 27, 51-57.
- van de Peer, Y., Chapelle, S. & De Wachter, P. (1996). A quantitative map of nucleotide substitution rates in bacterial rRNA. *Nucleic Acids Res* 24, 3381-3391.
- van houte, J., Gibbons, R. J. & Pulkkinen, A. J. (1972). Ecology of human oral lactobacilli. *Infect Immun* 6, 723-729.
- van Pelt, C., Verduin, C. M., Goessens, W. H. F., Vos, M. C., Tummler, B., Segonds, C., Reubsaet, F., Verbrugh, H. & van Belkum, A. (1999). Identification of *Burkholderia* spp. in the Clinical Microbiology Laboratory: Comparison of Conventional and Molecular Methods. *J Clin Microbiol* 37, 2158-2164.
- van Weemen, B. K. & Schuurs, A. H. W. M. (1971). Immunoassay using antigen--enzyme conjugates. *FEBS Letters* 15, 232-236.
- Vinas, M., Sabate, J., Espuny, M. J. & Solanas, A. M. (2005). Bacterial community dynamics and polycyclic aromatic hydrocarbon degradation during bioremediation of heavily creosote-contaminated soil. *Appl Environ Microbiol* 71, 7008-7018.
- Vliegen, I., Jacobs, J. A., Beuken, E., Bruggeman, C. A. & Vink, C. (2006). Rapid identification of bacteria by real-time amplification and sequencing of the 16S rRNA gene. *J Microbiol Meth* 66, 156-164.
- Voller, A., Bartlett, A. & Bidwell, D. E. (1978). Enzyme immunoassays with special reference to ELISA techniques. *J Clin Pathol* 31, 507-520.
- Wainwright, B. J., Scambler, P. J., Schmidtke, J., Watson, E. A., Law, H. Y., Farrall, M., Cooke, H. J., Eiberg, H. & Williamson, R. (1985). Localization of cystic fibrosis locus to human chromosome 7cen-q22. *Nature* 318, 384-385.
- Wang, Y. & Qian, P. Y. (2009). Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. *PLoS One* 4, e7401.

- Ward, D., Weller, R. & Bateson, M. (1990a). 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community. *Nature* 345, 63-65.
- Ward, D. M., Weller, R. & Bateson, M. M. (1990b). 16S rRNA sequences reveal uncultured inhabitants of a well-studied thermal community. *FEMS Microbiology Review* 6, 105-115.
- Waswa, J. W., Debroy, C. & Irudayaraj, J. (2006). Rapid detection of *Salmonella enteritidis* and *Escherichia coli* using sensor surface plasmon resonance biosensor. *J Food Process Eng* 29, 373-385.
- Wertheim, H. F., Melles, D. C., Vos, M. C., van Leeuwen, W., van Belkum, A., Verbrugh, H. A. & Nouwen, J. L. (2005). The role of nasal carriage in *Staphylococcus aureus* infections. *Lancet Infect Dis* 5, 751-762.
- White, T., Bruns, T., Lee, S. & Taylor, J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In *PCR protocols: A guide to methods and applications*, pp. 315-322. Edited by M. Innis, D. Gelfand, J. Sninsky & T. White. New York: Academic Press, Inc.
- Whittier, S. (2001). Update on the microbiology of cystic fibrosis: Traditional and emerging pathogens. *Clin Microbiol Newsletter* 23, 67-71.
- Wigley, P. & Burton, N. F. (1999). Genotypic and phenotypic relationships in *Burkholderia cepacia* isolated from cystic fibrosis patients and the environment. *J Appl Microbiol* 86, 460-468.
- Williamson, E. C. M., Speers, D., Arthur, I. H., Harnett, G., Ryan, G. & Inglis, T. J. J. (2001). Molecular epidemiology of *Scedosporium apiospermum* infection determined by PCR amplification of ribosomal intergenic spacer sequences in patients with chronic lung disease. *J Clin Microbiol* 39, 47-50.
- Williems, A. & Collins, M. D. (1995). NOTES: 16S rRNA gene similarities indicate that *Hallella seregens* (Moore and Moore) and *Mitsuokella dentalis* (Haapasalo et al.) are genealogically highly related and are members of the genus *Prevotella*: emended description of the genus *Prevotella* (Shah and Collins) and description of *Prevotella dentalis* comb. nov. *Int J Syst Bacteriol* 45, 832-836.
- Winslow, C., Broadhurst, J., Buchanan, R., Krumwiede, C. J., Rogers, L. & Smith, G. (1920). The families and genera of the bacteria: Final report of the committee of the society of American bacteriologists on characterization and classification of bacterial types. *J Bacteriol* 5, 191-229.
- Woese, C. R. (1987). Bacterial evolution. *Microbiol Rev* 51, 221-271.
- Wood, C. & Russel-Bell, B. (1983). Characterization of pigmented fungi by melanin staining. *Am J Dermatopath* 5, 77-82.

- Wood, R. W. (1902). On a remarkable case of uneven distribution of light in a diffraction grating spectrum. *Phil Mag Series 6* 4, 396 - 402.
- Wood, R. W. (1912). Diffraction gratings with controlled groove form and abnormal distribution of intensity. *Phil Mag Series 6* 23, 310 - 317.
- Wretling, B. & Pavlovskis, O. R. (1981). The role of proteases and exotoxin A in the pathogenicity of *Pseudomonas aeruginosa* infections. *Scand J Infect Dis Suppl* 29, 9-13.
- Yang, C., Mills, D., Mathee, K., Wang, Y., Jayachandran, K., Sikaroodi, M., Gillevet, P., Entry, J. & Narasimhan, G. (2006). An eco-informatics tool for microbial community studies: Supervised classification of Amplicon Length Heterogeneity (ALH) profiles of 16S rRNA. *J Microbiol Meth* 65, 49-62.
- Yang, F., Yang, J., Zhang, X., Chen, L., Jiang, Y., Yan, Y., Tang, X., Wang, J., Xiong, Z., Dong, J., Xue, Y., Zhu, Y., Xu, X., Sun, L., Chen, S., Nie, H., Peng, J., Xu, J., Wang, Y., Yuan, Z., Wen, Y., Yao, Z., Shen, Y., Qiang, B., Hou, Y., Yu, J. & Jin, Q. (2005). Genome dynamics and diversity of *Shigella* species, the etiologic agents of bacillary dysentery. *Nucleic Acids Res* 33, 6445-6458.
- Yarris, L. (2003). X-ray crystal images shed more light on protein synthesis. <http://www.lbl.gov/Science-Articles/Archive/ALS-ribosome-structure.html>
- Young, L. (1980). The role of exotoxins in the pathogenesis of *Pseudomonas aeruginosa* infections. *J Infect Dis* 142, 626-630.
- Zaura, E., Keijsers, B., Huse, S. & Crielaard, W. (2009). Defining the healthy "core microbiome" of oral microbial communities. *BMC Microbiology* 9, 259.
- Zeng, E., Mathee, K. & Narasimhan, G. (2007a). IEM: An algorithm for iterative enhancement of motifs using comparative genomics data. *Comput Syst Bioinformatics Conf* 6, 227-235.
- Zeng, J. S., Sutton, D. A., Fothergill, A. W., Rinaldi, M. G., Harrak, M. J. & de Hoog, G. S. (2007b). Spectrum of clinically relevant *Exophiala* species in the United States. *J Clin Microbiol* 45, 3713-3720.
- Zetola, N., Francis, J. S., Nuermberger, E. L. & Bishai, W. R. (2005). Community-acquired methicillin-resistant *Staphylococcus aureus*: An emerging threat. *Lancet Infect Dis* 5, 275-286.
- Zozaya-Hinchliffe, M., Martin, D. H. & Ferris, M. J. (2008). Prevalence and abundance of uncultivated *Megasphaera*-like bacteria in the human vaginal environment. *Appl Environ Microbiol* 74, 1656-1659.

## VITA

MELISSA S. DOUD

January 11, 1981	Born, Elmira, New York
1999-2003	Elmira College Elmira, New York B.A. Biology/Chemistry June 2003 <i>Magna cum laude</i>
2003-2006	Florida International University Miami, Florida Master of Science in Forensic Science May 2006
2006-2010	Florida International University Miami, Florida Doctoral Candidate in Biology

## PUBLICATIONS AND PRESENTATIONS

M. Doud, M. Light, G. Gonzalez, G. Narasimhan, and K. Mathee. 2010. Combination of 16S rRNA variable regions provides a detailed analysis of bacterial community dynamics in the lungs of cystic fibrosis patients. *Human Genomics* 4(3):147-169.

M. Doud, E. Zeng, L. Schneper, G. Narasimhan, K. Mathee. 2009. Approaches to analyzing dynamic microbial communities such as those seen in cystic fibrosis lung. *Human Genomics* 3(3): 246-256.

G. Gonzalez, M. Doud, K. Mathee, and G. Narasimhan. 2009. Computer-assisted bacterial identification using 16S rRNA sequence data. *IFMBE Proceedings* 24: 239-240.

M. Doud, R. Grimes-Zeppego, E. Molina, N. Miller, D. Balachandar, L. Schneper, R. Poppiti, K. Mathee. 2009. A *k<sub>2</sub>A*-positive *Klebsiella pneumoniae* causes liver and brain abscess in a Saint Kitt's man. *International Journal of Medical Sciences* 6: 301-304.

H. Vedala, D.H. Kim, M. Doud, K. Mathee, and W. Choi. 2008. Effect of environmental factors on electrical conductivity of single oligo-DNA molecule measured using single-walled carbon nanotube nanoelectrodes. *Nanotechnology* 19: 1-6 (265704).

K. Mathee, G. Narasimhan, C. Valdes, X. Qiu, J. M. Matewish, M. Koehrsen, A. Rokas, C. N. Yandava, R. Engels, E. Zeng, R. Olavarietta, M. Doud, R. Smith, P. Montgomery, J. White, P. A. Godfrey, C. Kodira, B. Birren, J. Galagan and S. Lory. 2008. Dynamics of *Pseudomonas aeruginosa* genome evolution. *Proceedings of National Academy of Sciences USA* 105(8): 3100-3105.

S. Roy S, H. Vedala, A.D. Roy, D.H. Kim, M. Doud, K. Mathee, H.K. Shin, N. Shimamoto, V. Prasad and W. Choi. 2008. Direct electrical measurements on single-molecule genomic DNA using single-walled carbon nanotubes. *Nano Letters* 8: 26-30.

#### AWARDS

- MBRS RISE Fellow – Awarded August 2006 till May 2010
- Student Training Grant – Cystic Fibrosis Foundation – Awarded 3 months: 2008-2009
- NIH-BRI/MBRS Summer Grant – Awarded Summer 2008
- NIH-BRI/MBRS Summer Grant – Awarded Summer 2005