Fall 12-16-2016

The Effect of Agricultural Exposures and Genetics on Respiratory Phenotypes

Lisa W. Boyd
University of Nebraska Medical Center

THE EFFECT OF AGRICULTURAL EXPOSURES AND GENETICS ON RESPIRATORY PHENOTYPES

by

Lisa Weissenburger-Moser Boyd

A DISSERTATION

Presented to the Faculty of

the University of Nebraska Graduate College

in Partial Fulfillment of the Requirements

for the Degree of Doctor Philosophy

Epidemiology Graduate Program

Under the Supervision of Professor Tricia D. LeVan

University of Nebraska Medical Center
Omaha, Nebraska

November, 2016

Supervisory Committee:

Debra Romberger, MD

Stephen Rennard, MD

Ted Mikuls, MD MSPH

Fang Yu, PhD

ACKNOWLEDGEMENTS

I would like to acknowledge the many people who have made this dissertation possible. First, I would like to thank my committee for all their effort and constant support on these projects. A very special thanks goes to Dr. Tricia LeVan, without whom my work on these projects would not be possible.

I would like to thank the patients at the VA Nebraska-Western Iowa Health Care System for continuing to be a part of these studies and their willingness to work with me.

To Dr. Jane Meza and Dr. Lynette Smith for all they do.

Thanks also to Dr. James Merchant and Dr. Kevin Kelly from the University of Iowa and their help in the Keokuk Country Rural Health Study database.

Finally, I would like to thank my family and friends for their continuous understanding, sacrifice, and encouragement; specifically, my parents, brothers, and husband, and also my friend, who helped edit this dissertation, Rachel Bowles.

Table of Contents

LIST OF TABLES

Chapter II

Chapter III

Chapter IV

LIST OF FIGURES

Chapter III

Chapter IV

LIST OF ABBREVIATIONS

COPD	Chronic Obstructive Pulmonary Disease
TLR4	Toll-like receptor 4
LPS	Lipopolysaccharide
SNP	Single Nucleotide Polymorphism
BMI	Body Mass Index
LD	Linkage Disequilibrium
CI	Confidence Interval
PFA	Principal Factor Analysis
FEV1	Forced Expiratory Volume in 1 second
FVC	Force Vital Capacity

ABSTRACT

Lisa Weissenburger-Moser, PhD

University of Nebraska Medical Center, December 2016

Supervisor: Tricia D LeVan, PhD

Agricultural workers are at risk for respiratory diseases, such as chronic obstructive pulmonary disease (COPD) and chronic bronchitis, due to exposures and inflammatory agents found in their work environment. While previous studies have been conducted to examine agricultural effects of chronic respiratory diseases more closely, there are still several significant gaps in the existing literature.

This dissertation explores the relationship between agricultural exposures and respiratory diseases. Utilizing two populations, the Keokuk Country Rural Health Study and the AgLung Study, this research: 1) used principal factor analysis (PFA) to distill exposure data into essential variables characterizing long-term agricultural exposures in order to examine the relationship between agricultural exposures and chronic respiratory diseases, 2) examined the relationship between *TLR4* 299/399 polymorphisms with chronic bronchitis and pro-inflammatory cytokines in two agriculturally-exposed populations, and 3) examined the lung microbiome longitudinally in agriculturally-exposed individuals with stable COPD. The results of the studies included in this dissertation may assist in continuing to fill the gaps in knowledge in this area of study in order to improve respiratory health in agricultural workers.

CHAPTER I.

INTRODUCTION

Respiratory Disease in Agricultural workers. Respiratory disease associated with agriculture work was one of the first recognized occupational hazards (1). As far back as the 16th century, the potential complications of inhaling dry grain dust and being exposed to hays and animals were described in farmers (1). In the United States, there are 3.2 million farmers operating 2.1 million farms (2). Chronic upper and lower airway diseases found in farmers include rhinitis, mucus membrane inflammation syndrome, sinusitis, asthma, asthma-like syndrome, chronic bronchitis, and chronic obstructive pulmonary disease (COPD) among others (3-9).

Agricultural respiratory symptoms have an overlapping spectrum (10), in that certain symptoms can be found in several diseases. The definition of COPD and its subtypes (chronic bronchitis, emphysema, and asthma) and the interrelationships between the closely related disorders that cause airflow limitation, provide a foundation for understanding the spectrum of patient presentations. In this dissertation, COPD and chronic bronchitis were examined, but it is important to define other subtypes of know COPD (i.e. emphysema) (11). Because patients present with a spectrum of manifestations of COPD, it is helpful to understand these subtypes.

COPD

COPD disease burden. COPD is a major global health problem. This debilitating condition is currently the third leading cause of death in the world and the United States (12,13), and is projected to be the fifth most common cause of disability in the world by 2020. More than 15.7 million Americans have been diagnosed with the disease, making

it a disease of public health importance (14). The NHLBI/WHO Global Initiative for Chronic Obstructive Lung Disease (GOLD) has defined COPD as: “a common preventable and treatable disease that is characterized by persistent airflow limitation that is usually progressive and associated with an enhanced chronic inflammatory response in the airways and the lung to noxious particles or gases.” (11). Exposure to tobacco smoke is recognized as the number one risk factor for COPD. Another important risk factor for COPD is exposure to agricultural dust (15).

There were several population studies and occupational cohort studies published during the 1970s and 1980s that determined occupational exposure to dust was associated with COPD (16-18). A review by the American Thoracic Society found that 15% of COPD is due to occupational exposures, independent of smoking (15). Although, it has been difficult in determine the effects of agricultural dusts and COPD independent of smoking, there are some studies that have disentangled these two risk factors and have shown that occupational dusts and fumes have an effect on lung function, making those with exposure more likely to develop COPD. Among farmers, the prevalence of COPD was found to be 30.2% in 2007 (19). In the same study, the attributable risk for COPD due to farming exposures was 7.7%. In other words, removing farming exposures could prevent 1 in 13 COPD cases (19). Not only is there a need to evaluate types of agricultural exposures and how they contribute to the disease, but there is also a need to better understand the biological processes that lead to inflammation and lung function decline in agricultural workers.

Chronic bronchitis

Chronic bronchitis disease burden. Chronic bronchitis, an inflammatory condition that affects the central bronchi, is one of two lung disease by which patients with COPD are

characterized. Excessive mucus secretion differentiates it from the emphysema (20). Other symptoms of chronic bronchitis related to lung inflammation and heavy mucus production include cough and production of sputum (20). In 2008, more than 9.8 million Americans reported having a chronic bronchitis diagnosis (21). Similar to COPD, tobacco smoke is recognized as the number one risk factor for chronic bronchitis; however, agricultural dust has proven to be a significant cause for the disease (22).

Farmers perform a variety of activities that potentially put them at risk for chronic bronchitis (23). Studies of the respiratory health of farmers, that included specialized production (i.e. raising food crops) (25) and a few larger population studies of farmers (22), have indicated that organic dust exposure in agriculture may initiate chronic bronchitis and related symptoms (24). Approximately 25% of agricultural workers commonly complain of cough and sputum production, both characteristics of the disease (23,25).

Emphysema and Asthma

Emphysema is a pathological term that describes some of the structural changes (i.e. abnormal and permanent enlargement of the airspaces distal to the terminal bronchioles that is accompanied by destruction of the airspace walls, without obvious fibrosis) sometimes associated with COPD (26). Asthma, although not a subtype of COPD, is defined as "a chronic inflammatory disorder of the airways in which many cells and cellular elements play a role. This disease is associated with airway responsiveness that leads to recurrent episodes of a variety of symptoms including wheezing, breathlessness, chest tightness, and coughing, particularly, both at night or in the early morning. These symptoms have been found to be associated with widespread, but variable, airflow obstruction within the lung that is often reversible either spontaneously or with treatment (27). Although not a subtype of COPD, asthma has an overlap of

syndromes (found in those with COPD) of people with asthma that develop non-reversible obstruction compared to those with asthma that is reversible.

Exposures. Working in agriculture is considered one of the most hazardous occupations. The industry is very diverse and includes multiple occupational and environmental exposures (10). Farmers, their families, and other agricultural workers can be exposed to a number of things: organic dusts (from moldy hay, grain, animals, and silage) that contain microorganisms and bacterial endotoxins, chemical toxicants, animal wastes, and other particular matter, all contributing to airway inflammation. These exposures and inflammatory agents are believed to contribute to respiratory diseases such as COPD.

Organic dust. Organic dust is mixture of vegetable products, insect fragments, animal dander, bird and rodent feces, pesticides, microorganisms, endotoxins, and pollens. Animal feces, endotoxins, and pollens are the primary sources of toxic and allergenic contributors in animal confinement facilities. Respirable dusts penetrate to the level of the terminal bronchioles and alveoli, where the gas exchange occurs (28). Production, harvest, transfer, storage, and processing procedures produce grain dust and the highest levels are associated with grain cleaning (29). Silos, chopping straw, unloading grain silos, shoveling feed, opening bales of hay for feed, and cleaning old animal housing structures also generate grain dust (10).

Organic dust can also contain biologically active proteins that may be allergenic and pro-inflammatory. It is the biologically active compounds, as well as toxicant gases (i.e. CO_2, NO_2), that contribute to respiratory conditions (23). As mentioned, agricultural

workers are exposed to high concentrations of airborne inhalable organic dust. And in this dust, there is high amounts of gram-positive and gram-negative bacteria components and fungi (30,31). Lipopolysaccharide (LPS) is a component of gram-negative bacterial cell wall (31,32), and exposure to LPS among workers has been associated with increased respiratory disorders, including chronic bronchitis (33-35). How these exposures influence long-term respiratory health is an area of active research.

Conducting Exposure Analysis. There is substantial published data on respiratory diseases, such as COPD, in the farming population. However, understanding COPD is challenging as causal inferences from epidemiological studies of chronic disease in agricultural populations is often limited due to a lack of long-term exposure measurements, limited sample sizes, phenotype outcomes, confounding, and heterogeneity (36).The types of methods used to assess agricultural exposures include direct measurement of personal exposure (37), biomarkers of exposure within a short period of time (38), or self-report questionnaire surveys (37). Although this direct exposure method is often an accurate approach, it may not be relevant for studies of disease with long latency periods such as cancer and COPD due to cost (39). Accurate estimation of long-term agricultural exposures based on questionnaire data has been used to improve the validity of epidemiologic investigations and subsequent evaluation of the association between agricultural exposures and chronic diseases (40). Questionnaires are generally designed to ask a large set of questions on the subject's attributes with the purpose of obtaining enough information for chronic exposure assessment. However, the designed questions in the questionnaire may not be direct indicators of the true exposure. Sorting out useful information from the large amount of

questionnaire data is challenging, and is essential in obtaining objective, unbiased, and interpretable exposure assessments in an epidemiological study.

There are many diverse agricultural tasks that characterize a person's exposures; therefore, it is difficult to have only one or two variables to explain the broad range of agricultural exposures. To date, several studies on chronic diseases such as cancer and COPD have utilized surveys and these surveys have assessed up to 50 exposures (37,40-42). It is uncertain if numerous questions are necessary to obtain accurate exposure information. Detailed questionnaires have been used in several agricultural studies (37). For this thesis, I used principal factor analyses (PFA), to summarize a large amount of important agricultural exposure variables from questionnaires.

I examined questionnaire data used to measure agricultural exposures in epidemiological studies where we identified the most important and biologically relevant exposures and distilled the number of variables that need to be collected moving forward. As a first step in understanding whether these agricultural exposures could be influenced by genetic risk factors, we examined the association of tagging TLR4 polymorphisms in the second aim of my dissertation, that have been previously implicated in respiratory diseases in other populations. I wanted to see if these risk factors are also associated with chronic respiratory disease such as chronic bronchitis among our agriculturally exposed populations.

Mechanisms of LPS induced inflammation: Studies have shown that there are high levels of inhalant LPS in grain threshing and sieving (43), flax threshing (44), herb processing (43,45), wood processing (46), waste collection and sorting (47), handling dry sludge (48), on pig farms (49,50), in cucumber and tomato nurseries (51), in plants

using straw and wood chips as biofuel (52), and in many other work environments (53). The biologically active lipid (Lipid A) moiety of LPS has been shown to be responsible for its adverse health effects (54,55).

Among different species, the LPS activity of various gram-negative bacteria varies. The genera of gram-negative bacteria, *Pseudomonas*, *Bacillus*, *Cornyebacterium, Pasteurella, Vibrio*, and *Enterobacter,* have been shown to shed LPS in agricultural settings (10). Activities of animals and humans produce organic dust which contains endotoxin and is found in animal confinement structures (swine and poultry), livestock farming, grain elevators, cotton industry, potato processing, flax industry, and animal feed industry (10).

TLR4. Once LPS is in the lungs, it induces an inflammatory process mediated through the CD14 receptor and Toll Like Receptor-4 (TLR-4). This leads to the impairment of lung function and respiratory diseases such as bronchitis, asthma, and toxic pneumonitis (organic dust toxic syndrome) (53,56).

Toll-like receptor 4 (TLR4) has been studied among the TLRs that recognize gram-positive (57) and gram-negative bacteria (58). The *TLR4* gene codes for the TLR4 protein and these receptors recognize LPS, in combination with CD14 (59) and myeloid differentiation protein 2 (MD-2) (60), to induce a cascade leading to the activation of NF-κB, AP-1, the innate immune system (61), and the production of pro-inflammatory cytokines (62,63). Both IL-6 and TNF-α, pro-inflammatory cytokines, are found in high levels in the lung among those with respiratory symptoms and disease (64-66).

Functional polymorphisms in *TLR4* have been studied and their role investigated in a number of inflammatory diseases (67). A large number of studies have focused on

the missense Asp299Gly polymorphism. Gly299 has been shown to cause reduced levels of pro-inflammatory cytokines and inflammatory hypo-responsiveness to inhaled LPS (63,68,69). Another study found that the frequency of this polymorphism was protective in patients with COPD (70), while other studies suggested that the presence of the *TLR4* polymorphism did not have any significant impact on lung function (67,71). Budulac et al. also showed that none of the *TLR4* single nucleotide polymorphisms (SNPs) were associated with FEV_1 in COPD patients, but found that some SNPs were significantly associated with accelerated or reduced decline of FEV_1 (72). However, the association of the Asp299Gly polymorphism and chronic bronchitis has not been studied in individuals who are chronically exposed to agricultural exosures.

COPD and the Lung Microbiome

Understanding the pathogenesis of COPD development and progression is challenging as it is a heterogeneous disease caused by a number of factors (i.e., environmental exposures and genetics). To date, there has been controversy regarding the role of lower respiratory tract bacteria in its pathogenesis (73). It was believed that the lungs of healthy individuals were sterile, while the lungs of COPD individuals were colonized. Recently, with the help of new culture-independent methods, it has been discovered that healthy lungs are not sterile and that the lung microbiome is modified in those with lung diseases. The lung microbiome of healthy individuals has been shown to consist primarily of bacteria from the Bacteriodetes, Firmicutes, Proteobacteria, and Actinobacteria phyla (74). Pathogens from the phylum Proteobacteria are also commonly colonized in COPD individuals, including genera *Haemophilus* and *Moraxella*. (74,75). From bronchoalveolar lavage samples, a study showed that *Haemophilus* species were more frequent in COPD patients, while *Bacteroidetes* were more prevalent

in control patients. Studies have been conducted to understand the microbiome of the lungs among “healthy” (no signs for disease or decreased lung function) smokers, never-smokers, and patients with COPD (73,75,76). These studies have consistently shown that smoking status and COPD severity play a role in bacterial colonization.

Several studies have shown that the lung microbiome becomes less diverse with increasing COPD severity (73,75); however Sze et al. failed to show significant differences in bacterial diversity with worsening COPD severity in lung tissue (76). Smoking in and of itself has shown not to alter the lung microbiome (77), and many studies of the lung microbiome of COPD patients have included smokers (75).

Few longitudinal studies examining the change of the lung microbiome over time have been conducted (78,79). Sethi et al. examined the lung microbiome during acute exacerbations and found lower abundances of genera *Moraxella* and *Streptococcus* in sputum samples compared with samples taken at time of “clinical stability” (78). Millares et al., on the other hand, identified increases in *Haemophilus, Pseudomonas*, and *Moraxella* during exacerbations compared to paired sampling from periods of clinical stability in COPD patients (79). Respiratory viral infections, especially rhinoviruses, are a major cause of COPD exacerbations (80). Molyneaux et al. investigated the effect of rhinovirus infection on the airway bacterial microbiome and discovered that rhinovirus infection in COPD patients alters the respiratory microbiome (81) with a significant increase of *Haemophilus influenzae* in patients with COPD. This was not observed in healthy individuals. These studies have shown that exacerbations and respiratory infections can alter the lung microbiome in patients with COPD (82-84). To date, there are no longitudinal studies comparing baseline sputum samples with samples taken several months later in COPD patients without any sort of intervention (i.e. antibiotic, transplantation, etc.). Understanding the stability of the lung microbiome in COPD is

critical for implementation of antimicrobial interventions that may reduce inflammation and slow disease progression.

Specific Aims

COPD and chronic bronchitis are both debilitating respiratory conditions caused by smoking and exposures to agricultural dusts. While there is a significant amount of literature examining the relationship between agricultural diseases and respiratory health, this existing body of literature has not yet fully examined the assessment of agricultural exposures, such as animals, crops, and farm tasks based on factor analysis, the association of *TLR4* polymorphisms with chronic bronchitis, and the stability of the lung microbiome in COPD patients.

The purpose of this dissertation is to examine the effects of complex environmental exposures as they relate to chronic respiratory diseases in order to fill these gaps in the literature. The first study of this dissertation (Chapter II) aims to determine the necessary and useful information to describe long-term agricultural exposures by summarizing a large amount of agricultural exposure variables using principal factor analyses. The aims of the second study of this dissertation (Chapter III) are to: examine the association of *TLR4* single nucleotide polymorphisms (SNPs) with chronic bronchitis in two novel agriculturally-exposed populations, and to determine the association of *TLR4* Asp299Gly and pro-inflammatory cytokines (IL-6 and TNF-α). The third study (Chapter IV) aims to determine the short term stability of the lung microbiome in COPD patients with chronic exposure to agricultural dusts using Illumina platform of the 16S rRNA gene.

Together, these studies will help to lead to advances in our overall understanding of COPD and other respiratory phenotypes in an agricultural exposed population. Although the proposed studies are primarily correlative in nature, these efforts are the pivotal steps in assessing 1) the relationship among a variety of agricultural exposures, 2) *TLR4* polymorphisms as it relates to chronic bronchitis, and 3) the stability of the

indigenous lung microbiome in COPD patients. We anticipate that this project, altogether, will impact agriculturally- exposed individuals with COPD and chronic bronchitis by contributing to the knowledge of the association between these exposures and respiratory diseases.

CHAPTER II. A PRINCIPAL FACTOR ANALYSIS TO CHARACTERIZE AGRICULTURAL EXPOSURES AMONG NEBRASKA VETERANS

Abstract

Agricultural workers are at increased risk of developing chronic respiratory disorders. Accurate estimation of long-term agricultural exposures based on questionnaires has been used to improve the validity of epidemiologic investigations and subsequent evaluation of the association between agricultural exposures and chronic diseases. Our aim was to use principal factor analysis (PFA) to distill exposure data into essential variables characterizing long-term agricultural exposures. This is a cross-sectional study of veterans between the ages of 40-80 years and worked on a farm for ≥2 years. Participant characteristics: 98.1% were white males with a mean age 65 ± 8 (SD) years and 39.8% had chronic obstructive pulmonary disease. The final model included four factors and explained 16.6% of the variance in the exposure data. Factor 1 was a heterogeneous factor; however, Factor 2 was exclusively composed of exposure to livestock such as hogs, dairy and poultry. Factor 3 included exposures from jobs on or off the farm such as wood dust, mineral dust, asbestos and spray paint. Crop exposure loaded exclusively in Factor 4 and included lifetime hours of exposure and maximum number of acres farmed in the participants' lifetime. The factors in the final model were interpretable and consistent with farming practices.

Introduction

Agricultural workers are at increased risk of developing chronic respiratory disorders including chronic bronchitis, occupational asthma and obstructive lung disease, and these diseases are likely caused by multiple agricultural exposures (85). Epidemiological studies can provide evidence of an exposure-response relationship, an important factor for the suggestion of a causal association (37). However, causal inference from epidemiological studies of chronic disease in agricultural populations is often limited due to a lack of long-term exposure measurements (36). The types of methods used to assess agricultural exposures have included direct measurement of personal exposure (86,87), biomarkers of exposure (88,89), and self-report questionnaires (90,91). Although the direct measurement method is often a precise approach, it may not be relevant for studies of disease with long latency periods such as obstructive lung diseases. Accurate estimation of long-term agricultural exposures based on questionnaire data has been used to improve the validity of epidemiologic investigations and subsequent evaluation of the association between agricultural exposures and chronic lung diseases (90). The questionnaire is usually designed to ask a large set of questions about agricultural tasks and exposures with the purpose of obtaining enough information for chronic exposure assessment. However, oftentimes the designed questions are not direct indicators of the true exposure. Sorting out useful information from the large amount of questionnaire data is challenging, yet essential in obtaining objective, unbiased, and interpretable exposure assessments in an epidemiological study.

Here we use a statistical method, principal factor analysis (PFA), to summarize a large amount of important agricultural exposure variables from questionnaires designed to assess the relationship between agricultural exposures and respiratory disease. PFA is a statistical method that has been proposed to characterize heterogeneous exposures

when exposure monitoring is unavailable and short-term exposure measurements are inadequate (92). To our knowledge there has been no assessment of agricultural exposures, such as animals, crops, and farm tasks based on factor analysis.

Our overall objective was to identify a set of essential agriculturally-related exposures that should be considered when assessing respiratory outcomes. Using data from a cross sectional study of veterans that worked on a farm or in production agriculture as an adult for ≥ 2 years, we applied the method of factor analysis to two questionnaires. Questionnaire 1 assessed agricultural exposures in two-hundred and sixty-three individuals, while Questionnaire 2 (extended version) evaluated exposure in another four-hundred and eighteen individuals. We first compared the pattern of clustered agricultural exposures of Questionnaire 1 to Questionnaire 2. Second, we ascertained whether utilization of dichotomous (yes/no) versus intensity exposure variables (years) yielded similar factor loading models. Finally, we evaluated whether there was greater variation explained using agricultural intensity exposure variables coded as total lifetime hours compared to exposure intensity variables coded as total lifetime years.

Methods

Study Population

We used agricultural exposure data from a cross-sectional study designed to assess the relationship between agricultural exposures and chronic respiratory disease in veterans utilizing the VA Nebraska Western Iowa Health Care System. Potential study participants were approached in the primary care outpatient clinics if they had worked on a farm as an adult for ≥ two years. Eligibility criteria for the study included individuals between the ages of 40 and 80 years. Individuals who had been diagnosed by a physician with asthma, lung cancer or interstitial lung disease such as pulmonary fibrosis, sarcoidosis and hypersensitivity pneumonitis were excluded from the study. Recruitment into the study began March 2008 and continued through December 2013 with a total of 681 participants. Demographic information, smoking status and agricultural-related exposures were obtained at the time of enrollment. COPD was defined as post-bronchodilator $FEV_1/FVC < 0.70$ by the Global Initiative for Chronic Obstructive Lung Disease (GOLD) classification criteria (93). The study was approved by the VA Nebraska Western Iowa Healthcare Systems Institutional Review Board and all participants signed a written informed consent document.

Exposure Questionnaires

Agricultural exposures were assessed using Questionnaire 1 (Q1) from March 2008 to July 2010. Questionnaire 2 (Q2) was developed to obtain more detailed agricultural exposure data and was utilized from August 2010 to December 2013. All participants answered either Q1 or Q2.

Questionnaire 1

Q1 was a telephone questionnaire conducted by the Nebraska Department of Health and Human Services. Participants were contacted at their preferred phone number within 30 days of enrollment. For Q1, participants were asked to provide total years of working or living on a farm, as well as intensity of farm work (weeks per year, hours per week) during their 20's, 40's and 60's. Twenty-four percent of the population was under the age of 60 and did not have agricultural exposure data during this time period (60's), thus the PFA for Q1 only examined intensity of farm work in participant's 20's and 40's. Information (yes/no) on their farm and off the farm exposures (farm tasks, livestock, crops, and 'other exposures', i.e. wood dust, grain dust, silica/mineral dust, asbestos, smoke other than cigarette, chemical solvents, spray paint, welding fumes) and whether they worked on a farm (yes/no) during their 20's and 40's were obtained. Farm tasks were assessed by asking "What were the tasks you performed on the farm?" and included spread manure, grind animal feed, handle silage, grind hay, till soil, drive combines, drive diesel tractors, and repair engines. Total years worked or lived on the farm were calculated by taking the age last lived or worked on a farm minus the age first lived or worked on a farm and subtracting any time between these two points when the participant did not live or work on a farm. The variables for weeks per year (≤ 4, 5-20, 21-40, > 41) and hours per week (< 20, 20-40, 41-60, >60) working on the farm during the participants' 20's and 40's were collected as categorical variables.

Questionnaire 2

Q2 was administered in-person by the study coordinator at the time of enrollment. In contrast to Q1, Q2 assessed lifetime exposures (birth to 80 years) and more detailed information about intensity of farm work (hours per week, weeks per year, total years), farm tasks (ever/never), livestock (total years, maximum # livestock), crops (total years, maximum # acres), and 'other exposures' on and off the farm (hours per week, weeks

per year, total years). A composite intensity exposure variable, total lifetime hours, was calculated as [total years* total hours/week * total weeks/year]. Additional exposure variables were collected in Q2 such as worked with diesel powered farm equipment (maximum days per year, total years) and worked with gas powered farm equipment (maximum days per year, total years). In order to compare Q1 and Q2, we recoded Q2 to represent exposures during the participant's 20's and 40's similar to Q1 (yes/no), except Q2 data for farm tasks were utilized as ever/never. The intensity of farm work variables were collected as continuous variables in Q2 then coded as categorical variables (≤ 4, 5-20, 21-40, > 41 weeks per year) and (< 20, 20-40, 41-60, >60 hours per week) during the participants' 20's and 40's for Q1 and Q2 comparisons.

Principal factor analysis

Analyses were conducted using SAS/STAT® software for Windows version 9.2 (SAS Institute, Cary, NC, USA). We first standardized all time-related exposure variables that were continuous and described as total years or total hours to zero mean and unit variance, so that all these variables would enter PFA under similar scales. PFA was conducted using SAS PROC FACTOR using a polychoric correlation covariance matrix; a method for estimating correlations among theorized normally distributed continuous latent variables from observed ordinal variables (94,95). With this method, factors, that are independent of each other, were extracted in descending order of importance with respect to the proportion of the variance accounted for by each factor (96). For example, the first factor was derived from a weighted linear combination of agricultural variables that accounted for the largest total variation in the data. The second factor derived contained another linear combination of agricultural variables and accounted for variance not accounted for by the first factor.

The number of factors in the model was determined based on the following criteria: at least two variables with a loading score of ≥ 0.5 in a factor; factors must have an eigenvalue > 1.0; and each factor must account for at least 1 percent of the total variance. For every variable in each factor, a factor loading score was calculated, which represents the correlations between each of the variables included in each factor, similar to Pearson correlation coefficients (97,98). Generally, a factor loading score of 0.30 to 0.40 are considered meaningful (96-99); however we used a factor loading score of ≥ 0.5 to identify the most highly correlated variables in each factor. Also, the eigenvalue for each factor was calculated and an eigenvalue > 1 indicated that the factor explained more of the variance than could be accounted for by any one variable (96,99). We used a promax (oblique) rotated factor pattern because we assumed that the factors were correlated (92). We determined the number of factors using the scree test plot. The scree test plots the factors on the X-axis and the corresponding eigenvalues on the Y-axis (100). The test drops factors after the break of inflexion. This test is reliable when the sample size is at least 200 (96). The scree test plot was first viewed to determine the number of factors to include and PROC FACTOR was again conducted where the number of factors were specified.

In total, four models were run. Model 1 used data from Q1 and was compared to Q2. Models 1 and 2 differed only by the way farm task questions were asked, i.e. For Q1 farm tasks were asked as “yes/no” during the participants 20’s and 40’s and for Q2 farm tasks were asked as “ever/never” during their lifetime. Model 2 was then compared to Model 3 to ascertain whether utilization of dichotomous versus intensity exposure variables (years), respectively, yielded similar factor loading models. And finally, Model 3 was compared to Model 4, to determine whether agricultural exposure variables coded as total lifetime years compared to total lifetime hours, respectively, generated a greater percentage of variation explained.

Results

A total of 263 eligible subjects were enrolled using Q1 and 418 participants enrolled using Q2, all with the exposure questionnaire completed. The two populations were primarily white males with ~ 55% of the participants having greater than a high school education (Table 1). The prevalence of COPD in this population was 39.8%. Of note, participants enrolled using the Q2 were older ($p = 0.007$, Q1 = 63.5 years ± 8.1 SD vs. Q2 = 65.3 years ± 8.7 SD), worked on a farm for longer ($p = 0.001$, Q1 = 24.6 years ± 19.6 SD vs. Q2 = 29.6 years ± 18.5 SD) and were more likely to be exposed to hogs in open pen, beef cattle, dairy cattle, poultry and crops than those enrolled with Q1.

Questionnaire 1

For development of Model 1, agricultural exposure data were obtained from Q1. Q1 collected mostly dichotomous exposure data (yes/no) during the participants' 20's and 40's, except duration (years lived/worked) and intensity of farm work (weeks per year and hours per week) were obtained as continuous variables. The factors for Model 1 yielded eigenvalues greater than 1 and explained 24.4% of the variance in the exposure data (Table 2). Factor 1 explained 7.3% of the variance in the observed data, Factor 2 explained 7.0%, Factor 3 and Factor 4 explained 7.0% and 3.1%, respectively. The proportion of variance explained by each of the remaining factors was 6.2% and these factors were not included in the final model due to our *a priori* inclusion criteria. Variables loading high on Factor 1 (i.e. factor loading scores equal to or greater than 0.50) were exclusively 'other exposures' from a job on or off the farm during the participants 20's or 40's, including wood dust, grain dust, rock dust, asbestos, smoke other than cigarette, chemical solvents, spray paint, and welding fumes. Loading high on Factor 2 were live/work on farm (weeks per year, hours per week) during their 20's, farm tasks such as spread manure, handle silage, and grind hay during their 20's or

40's, and exposure to many types of livestock. Variables substantial to Factor 3 were total lifetime years lived or worked on the farm as well as worked on the farm during their 40's (weeks/year and hours/week). Farm tasks performed during their 20's or 40's, such as grinding animal feed, driving combines, driving diesel tractors, along with exposure to pesticides, loaded high in Factor 3. Factor 4 included two variables, exposure to hogs in closed lots and crops.

Questionnaire 2

Because there were two questionnaires, two phases of population recruitment, and more detailed exposure information collected in Q2 compared to Q1, we wanted to determine if the factor models obtained by each questionnaire were qualitatively comparable when using similar exposure variables. Data for Q2 were recoded to represent exposures (lived/worked on a farm and variables for exposure to livestock, crops and 'other exposures') during the participants' 20's and 40's. Data for farm tasks were utilized as lifetime exposure (ever, never). In Model 2, four factors were retained in the model and explained 14.5% of the total variance in the observed data (Table 3). The remaining factors accounted for 5.3% of the variance. Variables loading high on Factor 1 were heterogeneous and included worked on a farm during the participants' 20's (weeks per year and hours per week) and exposure to hogs in open lots, beef cattle, dairy cattle, poultry, crops, and grain dust in their 20's or 40's. Factor 1 explained 4.3% of the variance in the observed data. Factor 2 explained 3.9% of the variance in the observed data and was a homogeneous factor comprised of many farming tasks performed in their lifetime such as spread manure, grind animal feed, handle silage, grind hay, till soil, and drive combines and diesel tractors. Variables included in Factor 3 were years lived and worked on the farm and worked on the farm (weeks per year and hours per week) in the

participants' 40's. Factor 3 explained 3.6% of the variance in the observed data. Factor 4 explained 2.7% of the variance and included exposure to wood dust, rock dust, asbestos, chemical solvents, and spray paint during their 20's or 40's with asbestos, smoke, chemical solvents and welding fumes near the cutoff loading score of 0.5.

Q2 collected detailed exposure data over the participant's lifetime. We wanted to ascertain whether utilization of these intensity exposure variables (years) yielded more homogeneous factors compared to using dichotomous (yes/no) exposure variables. In Model 3, we incorporated lifetime agricultural exposures (continuous variables) and compared the factors and factor loading scores to Model 2, where dichotomous exposure variables (20's and 40's) were utilized. For Model 3, three factors explained 10.5% of the total variance in the observed data (Table 4). The proportion of variance explained by the remaining factors was 5.6%. Factor 1 was a heterogeneous factor explaining 4.7% of the variance and included years lived and worked on the farm, years worked with beef cattle, crops, grain dust, and pesticide. Factor 2 in Model 3 loaded similar variables as Factor 2 in Model 2 and explained 3.5% of the variance, i.e. farming tasks such as spread manure, grind animal feed, handle silage, grind hay, till soil, and drive combines. Factor 3 explained 2.3% of the variance and included the lifetime exposure (years) to wood dust, rock dust, asbestos, chemical solvents, and spray paint.

We developed Model 4 to assess if more detailed lifetime intensity variables resulted in unique principal factors and exposure patterns that captured a greater variation than Model 3. Model 4 employed total lifetime hours for worked on farm, worked with livestock, exposure to crops and 'other exposures' (Table 5). Additional variables utilized in Model 4 were the summation of maximum number of livestock, maximum number of acres of crops and diesel/gas exposure. Model 4 included four factors and explained 16.6% of the variance. The remaining factors accounted for 11.5% of the variance. Factor 1 explained 7.8% of the total variance and included years lived

on the farm, total hours worked on the farm, total years worked with diesel power, total days/year worked with gas powered equipment, and farm tasks performed over a lifetime, such as till soil, drive combines, and drive diesel tractors. Total years worked with beef cattle, total years worked with crops, total number of acres of crops, and total hours exposed to grain dust, pesticides, and diesel fuel were also included in Factor 1. Factor 2 included total years exposed to hogs in open lots, total years of exposure and number of dairy cattle and poultry. Factor 2 explained 3.6% and Factor 3 explained 2.7% of the total variance. Factor 3 included lifetime total hours exposed to rock dust and spray paint. Factor 4 included total years and acres of other crops and explained 2.5% of the total variance.

In order to reduce bias, a sensitivity analysis was performed for Model 4 by stratifying by COPD status (Table E1). Similar clustering patterns were found for Factors 1 and 2 in the total population and those with COPD and those without COPD, and were identical when the factor loading score was relaxed to 0.4. Factor 3 in Model 4 for the total population loaded similar variables to those with COPD, while Factor 4 contained variables from both COPD and no COPD. In addition, age and smoking status were tested in all models; however, these variables had a loading score <0.5 and thus, were not included in the final models.

Discussion

The ultimate goal of the veteran cohort is to describe long-term agricultural exposures and their relation to respiratory outcomes. Existing studies have shown the harmful effects of the farming environment on COPD, asthma, and other airway diseases (85). Specifically, exposures such as animals, hay, and grains are known to have an adverse effect on respiratory health (101), as well as agricultural pesticides (102). Long-term work in large animal-feeding operations, particularly swine confinement facilities and cattle feedlots (103), also contribute to chronic respiratory disease with dairy farming associated specifically with COPD (104).

In this exploratory statistical analysis, we utilized principal factor analysis to examine the correlation among a large number of exposure variables as well as to reduce the number of variables into domains of agricultural exposure patterns without loss of a significant amount information. Model 1 utilized Q1 that collected dichotomous (yes/no) exposure data during the participants' 20's and 40's. Models 2, 3, and 4 utilized variables collected from Q2 which quantitated lifetime agricultural exposures as total years, weeks per year and hours per week. Overall, we found that duration and intensity of farm work, farm tasks, livestock exposure, crop exposure and 'other exposures' were independent entities and their clustering within a model was modified by the intensity units of exposure (dichotomous vs. continuous).

There were four principal factors derived for Model 1 using Q1. Factor 1 had a homogeneous cluster composed of variables in the 'other exposures' category and represented job exposures on or off the farm such as wood dust, grain dust, rock dust, asbestos, smoke other than cigarette, chemical solvents, spray paint, and welding fumes. These exposures are often categorized as vapor, dust and smoke, and have been associated with occupational respiratory disease such as asthma and COPD (105). Factor 2 was heterogeneous yet interpretable and included variables such as duration of

farm work during the participants' 20's, select farm tasks and livestock exposures. Of note, the farm tasks in this factor were related to animal husbandry such as spread manure and exposure to dairy cattle. Individuals who farmed during their 20's were more likely to have exposure to animals than those that farmed during their 40's. In contrast, individuals that farmed during their 40's were more likely to perform less strenuous tasks such as drive combines and diesel tractors and this pattern was observed in Factor 3. There are many reasons why younger farmers have different exposures than older farmers. Open cabbed tractors, while rare today, were the norm for older farmers and therefore, they were more exposed to pesticides and dust (106). We see this in Model 1 where working on the farm in the participant's 40's clustered with driving of combines and diesel tractors as well as pesticides. There was a clear separation between all factors such that each variable loaded significantly on only one factor. The variables with loading scores of ± 0.50 or higher within a factor were correlated most likely due to the fact that many of the variables within a cluster, such as farm tasks, are done collectively when working in agriculture.

Model 2 was derived using Q2 variables that were recoded to replicate exposure variables similar to Q1. The variables were dichotomous for exposure during the participants' 20's and 40's. As in Model 1, there were four factors and each variable loaded significantly on only one factor. The first factor included working on a farm during their 20's and this was correlated to animal exposures such as beef and dairy cattle and hogs in open lots (marginal correlation). This pattern was also observed in Model 1. In addition to animal exposures during the participants 20's, crops and grain dust were included in Factor 1 and are consistent with livestock production practices. Factor 2 aligned with many of the farming tasks, while Factor 3 consisted of lifetime years lived and worked on a farm along with intensity of farm work during the participants 40's. This clustering of lifetime years and intensity of farm work was similar to that observed in

Model 1. Factor 4 included variables from 'other exposures' and this same pattern was seen for Model 1. Overall the factors in Model 1 and 2 were similar with clustering of lifetime years worked/lived on a farm, intensity of farm work, livestock exposure, and 'other exposures'. The major difference between the two models was that farm tasks loaded heavily in Model 2 compared to Model 1. In Q2, these farm task questions were asked as "ever/never" during their lifetime, whereas in Q1 these questions were asked with "yes/no" answers for their 20's and 40's. These observations suggest that collecting information on farm tasks is important in accounting for the variability in agricultural exposures due to their heavy loading in the model and that the "ever/never" during a person's lifetime would be more all-inclusive. Furthermore, the dissimilarities of factors in Models 1 and 2 may be due to the different age structure of these two populations. The population from Q2 had a greater proportion of people older than 70 than the population from Q1. Of note, the percentage of variation explained is a measurement of fit. Q2 had a lower percentage of variation explained compared to Q1, which may be due to greater variability as it was used on a larger population with more workers (86% vs 59%) working on a farm for more than 10 years.

For Model 3, we used lifetime exposures with intensity units as total years, except for farm tasks as ever/never. The principal factors for Model 3 had three distinct patterns. The first factor contained heterogeneous exposure variables including live/work on the farm, livestock, crops, and 'other exposures'. Farming tasks clustered and loaded heavily in Factor 2 as with 'other exposure' variables in Factor 3. Also, these domains were predominant in Model 2. Even though the percentage of variance explained in Model 3 was less than that in Model 2, there was utilization of more complete exposure variables (lifetime) in Model 3 compared to Model 2 (20's and 40's).

As a final Model, we included all of the collected exposure and intensity variables as total lifetime hours, maximum lifetime number of livestock or acres of crops or

ever/never farm tasks. We observed four distinct factors in Model 4. Factor 1 was a heterogeneous factor that included exposures related to crops and livestock, whereas the main Factor 2 domain was livestock. Factor 3 included variables from 'other exposures' and Factor 4 was solely 'other crops'. Model 4 captured a higher percentage of variance, which suggests that detailed intensity variables for agricultural exposure are advantageous in capturing a greater percentage of variance than dichotomous (yes/no) or even the variables coded as total years. We observed that diesel/gas exposure variables were important to include in Model 4 as it loaded high in Factor 1. Model 4 included additional crop variables that were asked in Q2 and resulted in a distinct factor pattern of crops (Factor 4). This was not found in previous models.

Many studies have found the utility of factor analysis. The Agricultural Health Study utilized factor analysis to identify clusters of pesticide exposures that relate to prostate cancer (107). Another study clustered respiratory phenotypes of COPD to explain the heterogeneity of COPD (108). PFA is not only used to assess the effect of occupational exposures on respiratory diseases, but is also used to evaluate the reproducibility and validity of questionnaires as Hammond et al. tested the validity and reliability of the English Evaluation of Daily Activity Questionnaire (109).

In this study, factor analysis was used to extract the useful information from a complex dataset to interpret the agricultural exposure data. Studies have found the importance in including the use of solvents, paint, exposure to welding fumes (110), and pesticide use (98) when investigating exposure-respiratory disease associations. We found these exposure variables to also be important in our analysis in describing long-term agricultural exposures.

This study has some important strengths. First, the exposure data were comprehensive, including hours per week, weeks per years, and total years, and were

collected by trained study personnel. Secondly, the agricultural population is large and all have worked in Nebraska or Iowa, thus have similar exposures. Finally, the statistical methods used allow unbiased analyses that are not based on any a priori assumptions. This study does have limitations. Recall bias is probable since participants were asked to retrospectively recall their lifetime farming exposures. This could have resulted in overestimation or underestimation of the exposure, which could ultimately impact factor weighting and subsequent regression analysis. There is a potential for interviewer bias as there were two methods to obtain exposure information, telephone interviews for Q1 and in-person interviews for Q2. It would be difficult to determine if this would be an over- or under-reporting of exposures. Additionally, there is the issue of generalizability of these results. The population included veterans with agricultural exposure utilizing the VA Nebraska Western Iowa Health Care System. They were primarily white males with a mean age of 64 years; therefore, their agricultural exposures may be different from younger workers due to technological advances in farming. In addition, direct measurement of agricultural exposures was not performed.

In summary, we found that PFA was an effective statistical method for characterizing exposure patterns in our population of agricultural workers. We have identified clusters in a large dataset that describes the heterogeneity of exposures including duration and intensity of farm work, farm tasks, livestock exposure, crop exposure and ‘other exposures’. We examined four models and found that Model 4, with the most detailed exposure information, captured the highest percentage of variance compared to the other models. The resulting factor patterns were clearly interpretable and logical in terms of farming practice. From this study, we also determined that the most important exposure variables to be asked in questionnaires when evaluating agricultural exposures and respiratory diseases are years worked on a farm, farm tasks, and exposure to livestock, crops and ‘other exposures’ as these consistently loaded high

across the four models. The next step is to further explore these patterns in Model 4 to examine the relationship between agricultural exposures and respiratory diseases such as COPD in this population.

Table 1. Study Population Characteristics

Characteristic	Questionnaire 1 n=263	Questionnaire 2 n=418
Sex*		
Male	261 (99)	406 (97)
Female	2 (1)	12 (3)
Age (yrs)*		
≤49	14 (5)	20 (5)
50-59	59 (22)	71 (17)
60-69	126 (48)	191 (46)
70-80	64 (24)	136 (32)
Race*^		
White	259 (98)	391 (95)
Other	4 (2)	20 (5)
Education**		
≤ High school	99 (39)	180 (44)
> High school	131 (51)	230 (56)
Refused	27 (10)	0 (0)
Smoking Status^		
Current	53 (21)	86 (21)
Former	147 (59)	240 (58)
Never	49 (20)	87 (21)
Refused	1 (<1)	0 (0)
Worked on Farm (yrs)#**	23.9 ± 19.7	29.4 ± 18.6
Agricultural Exposures (Yes/No)	58 (26)	84 (20)
Hogs, confinement	115 (51)	296 (71)
Hogs, open pen**	136 (60)	309 (74)
Beef cattle**	67 (30)	243 (58)
Dairy cattle**	85 (38)	286 (69)
Poultry**	155 (68)	399 (96)
Crops**		
COPD^		
Yes	102 (42)	151 (37)
No	141 (58)	257 (63)

Data are presented as n (%)
**P* value < 0.05
***P* value < 0.001
^Numbers don't add up due to missing values
#Data presented as mean ± SD

Table 2. Principal factor analysis results for Questionnaire 1 (Model 1), (n= 263)*

Exposures		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4
Live/Work on Farm					
(20's, 40's)	Lived on farm, yrs (lifetime)	-28	1	**82**	6
	Worked on farm, yrs (lifetime)	3	9	**61**	8
	Worked on farm, wks/yr (20's)	-1	**63**	27	7
	Worked on farm, wks/yr (40's)	10	-8	**109**	-32
	Worked on farm, hrs/wk (20's)	5	**55**	19	12
	Worked on farm, hrs/wk (40's)	1	-14	**110**	-16
Farm Tasks					
yes/no, (20's, 40's)^	Spread manure	17	**59**	36	-7
	Grind animal feed	7	39	**52**	17
	Handle silage	21	**50**	19	12
	Grind hay	10	**67**	30	7
	Till soil	20	34	48	19
	Drive combines	9	33	**67**	0
	Drive diesel tractors	34	7	**50**	23
	Repair engines	38	28	25	-4
Livestock					
yes/no, (20's, 40's)^	Hogs in closed lots	14	-8	10	**85**
	Hogs in open lots	-12	**62**	24	29
	Beef cattle	20	**55**	27	12
	Dairy cattle	-18	**120**	-22	-9
	Poultry	8	**69**	1	5
	Other livestock	31	**83**	-16	33
Crops					
yes/no, (20's, 40's)^	Corn, soybeans, hay, grain sorghum, wheat, oats	-5	0	-25	**115**
Other Exposures	Wood dust	**89**	7	-4	-25
yes/no, (20's, 40's)^	Grain dust	**53**	32	22	6
	Silica/sand/rock/mineral dust	**84**	22	-30	3
	Asbestos	**76**	-20	10	9
	Smoke other than cigarette	**90**	13	-18	3
	Chemical solvents	**82**	2	10	17
	Spray paint	**80**	-12	14	7
	Welding fumes	**74**	2	23	2
	Pesticide	28	11	**56**	4
Eigenvalue		19.0	2.1	1.8	1.4

*For ease of presentation, all values were multiplied by 100 and rounded to the nearest integer.
^If they answered yes in either their 20's and/or 40's
Bolded values represent factor loading score of +0.50 or higher

Table 3. Principal factor analysis results for Questionnaire 2 (Model 2), n = 418*

Exposures		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4
Live/Work on Farm (20's, 40's)	Lived on farm, yrs (lifetime)	8	28	**52**	-23
	Worked on farm, yr (lifetime)	9	18	**65**	-17
	Worked on farm wks/yr (20's)	**66**	0	-13	-13
	Worked on farm wks/yr (40's)	-11	-16	**105**	8
	Worked on farm hrs/wk(20's)	**50**	-7	-13	-8
	Worked on farm hrs/wk (40's)	9	-10	**92**	4
Farm Tasks ever/never (lifetime)	Spread manure	1	**62**	-1	7
	Grind animal feed	3	**77**	6	-10
	Handle silage	0	**73**	-12	25
	Grind hay	-3	**52**	-13	7
	Till soil	3	**70**	17	-13
	Drive combines	23	**75**	5	0
	Drive diesel tractors	16	**51**	29	34
	Repair engines	-9	20	17	35
Livestock yes/no (20's, 40's)^	Hogs, closed lots	24	7	5	4
	Hogs, open lots	**58**	27	8	-18
	Beef cattle	**56**	7	18	2
	Dairy cattle	**65**	17	-15	2
	Poultry	**79**	2	-3	10
	Other livestock	12	5	7	33
Crops yes, no (20's, 40's)^	Corn, soybeans, hay, grain sorghum, wheat, or oats	**76**	-5	30	4
Other Exposures yes, no (20's, 40's)^	Wood dust	10	-1	-6	**63**
	Grain dust	**74**	-8	21	17
	Silica/sand/rock/mineral dust	16	-9	-15	**62**
	Asbestos	-20	14	0	**51**
	Smoke, not cigarette	-12	-6	17	45
	Chemical solvents	-12	-2	6	**52**
	Spray paint	-9	20	-17	**59**
	Welding fumes	20	-3	-1	44
	Pesticides	35	8	36	7
Eigenvalue		7.3	2.9	2.3	2.0

*For ease of presentation, all values were multiplied by 100 and rounded to the nearest integer.

^If they answered yes in either their 20's and/or 40's

Bolded values represent factor loading score of +0.50 or higher

Table 4. Principal factor analysis results using Questionnaire 2 (Model 3), n = 418*

Exposures		FACTOR 1	FACTOR 2	FACTOR 3
Live/Work on Farm (lifetime)	Lived on farm, yrs	**68**	11	10
	Worked on farm, yrs	**94**	-9	0
Farm Tasks ever/never (lifetime)	Spread manure	-3	**65**	5
	Grind animal feed	15	**72**	-9
	Handle silage	-16	**77**	14
	Grind hay	-15	**58**	1
	Till soil	17	**70**	-15
	Drive combines	18	**72**	-7
	Drive diesel tractors	23	46	24
	Repair engines	20	13	40
Livestock total yrs (lifetime)	Hogs in closed lots	5	24	2
	Hogs in open lots	49	19	-13
	Beef cattle	**61**	7	1
	Dairy cattle	23	19	2
	Poultry	17	0	6
	Other livestock	10	5	17
Crops total yrs (lifetime)	Corn, soybeans, hay, grain sorghum, wheat, oats	**96**	-5	0
Other Exposures total yrs (lifetime)	Wood dust	-4	-5	**56**
	Grain dust	**84**	-9	11
	Silica/sand/rock/mineral dust	-15	0	**62**
	Asbestos	6	-1	**57**
	Smoke other than cigarette	8	-3	44
	Chemical solvents	8	11	**50**
	Spray paint	-16	15	**53**
	Welding fumes	22	8	44
	Pesticide	**68**	8	7
Eigenvalue		6.8	2.7	1.4

*For ease of presentation, all values were multiplied by 100 and rounded to the nearest integer.
Bolded values represent factor loading score of +0.50 or higher

Table 5. Principal factor analysis results using Questionnaire 2 (Model 4), n = 418*

Exposures		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4
Live/Work on Farm (lifetime)	Lived on farm, yrs	**64**	15	-25	3
	Worked on farm, total hrs^	**88**	3	-21	10
Farm Tasks never, ever (lifetime)	Spread manure	20	48	25	-20
	Grind animal feed	36	36	8	-42
	Handle silage	15	33	42	-29
	Grind hay	7	27	20	-26
	Till soil	**53**	22	10	-7
	Drive combines	**50**	14	11	-39
	Drive diesel tractors	**64**	-21	37	-10
	Repair engines	27	-12	38	-3
Livestock total hrs^ & max no.+ (lifetime)	Hogs in closed lots, total yrs	12	17	5	-8
	Hogs in open lots, total yrs	36	**51**	-31	-15
	Max # of hogs in open lots	30	38	-19	-34
	Beef cattle, total yrs	**58**	23	-13	7
	Max # beef cattle	43	7	6	5
	Dairy cattle, total yrs	4	**66**	-2	6
	Max # dairy cattle	-4	**58**	6	0
	Poultry, total yrs	-8	**78**	-5	20
	Max # poultry	-26	**72**	-6	14
	Other livestock, total yrs	17	30	31	40
	Max # other livestock	10	29	36	35
Crops total hrs^ & max no.+ (lifetime)	Corn, soybeans, hay, grain sorghum, wheat, oats, total yrs	**85**	2	-22	6
	Max # of acres, corn, soybeans, hay, grain sorghum, wheat, oats	**71**	-25	6	-4
	Other crops, yrs	11	17	6	**80**
	# acres, other crops	15	12	9	**79**
Other Exposures total hrs^ (lifetime)	Wood dust	-14	12	49	8
	Grain dust	**79**	-1	-4	12
	Silica/sand/rock/mineral dust	-24	14	**54**	-7
	Asbestos	-3	-8	49	0
	Smoke other than cigarette	8	11	37	31
	Chemical solvents	-2	-4	39	9
	Spray paint	-6	-10	**57**	1
	Welding fumes	14	0	40	-6
	Pesticide	**70**	3	1	12
Diesel/Gas (lifetime)	Worked diesel power, yrs	**93**	-23	10	13
	Worked gas powered equipment, days/yr	**72**	0	-3	-4
	Diesel fuel/fumes/exhaust	**75**	-7	16	6
Eigenvalue		9.3	2.9	2.4	2.2

*For ease of presentation, all values were multiplied by 100 and rounded to the nearest integer. Bolded values represent factor loading score of +0.50 or higher

^Total hours in lifetime = (total yrs)*(total hrs/wk)*(total wks/yr)

+Maximum # = average number of livestock or average number acres of crops in lifetime

Supplement

A Principal Factor Analysis to Characterize Agricultural Exposures Among Nebraska Veterans

Table 6. Principal factor analysis results using Questionnaire 2 (Model 4, stratified by COPD status), n =418*

Exposures		FACTOR 1		FACTOR 2		FACTOR 3		FACTOR 4	
		COPD	No COPD	COPD	No COPD	COPD	No COPD	COPD	No COPD
Live/Work on Farm (lifetime)	Lived on farm, yrs	46	**73**	38	7	5	-21	-20	8
	Worked on farm, total hrs^	**83**	**90**	19	-7	8	-15	-20	14
Farm Tasks never, ever (lifetime)	Spread manure	14	18	15	**62**	**55**	12	41	8
	Grind animal feed	35	32	24	**54**	49	7	3	-21
	Handle silage	29	3	9	**51**	31	31	49	-19
	Grind hay	38	-8	-12	45	**66**	19	0	-11
	Till soil	49	**51**	-3	37	24	-1	15	7
	Drive combines	49	43	22	25	39	11	9	-33
	Drive diesel tractors	**55**	**60**	-13	-11	7	41	31	-14
	Repair engines	10	35	0	-8	-9	47	29	2
Livestock total hrs^ & max no.+ (lifetime)	Hogs in closed lots, total yrs	12	13	12	14	16	3	-1	7
	Hogs in open lots, total yrs	34	35	**63**	**50**	9	-33	-16	-12
	Max # of hogs in open lots	21	27	**50**	48	22	-29	5	-35
	Beef cattle, total yrs	46	**62**	40	16	-6	-11	1	6
	Max # beef cattle	34	42	13	8	-2	0	20	5
	Dairy cattle, total yrs	23	-4	**65**	**69**	-12	1	-1	15
	Max # dairy cattle	9	-10	**52**	**64**	-8	4	8	13
	Poultry, total yrs	-13	4	**74**	**60**	-2	1	7	38
	Max # poultry	-36	-14	**80**	**60**	-22	3	5	24
	Other livestock, total yrs	11	25	6	16	3	21	**50**	**61**
	Max # other livestock	4	16	1	18	8	23	**59**	**56**
Crops total hrs^ & max no.+ (lifetime)	Corn, soybeans, hay, grain sorghum, wheat, oats, total yrs	**81**	**86**	17	-4	2	-16	-25	4
	Max # of acres, corn, soybeans, hay, grain sorghum, wheat, oats	**69**	**68**	-22	-20	11	6	0	-6
	Other crops, yrs	37	-6	8	5	**-75**	-17	23	**84**
	# acres, other crops	42	-2	6	2	**-81**	-10	21	**81**
Other Exposures total hrs^ (lifetime)	Wood dust	-15	-16	4	17	-7	48	**53**	10
	Grain dust	**78**	**81**	-5	-2	-1	-2	-4	13
	Silica/sand/rock/mineral dust	-49	-12	12	12	6	**54**	**71**	1
	Asbestos	-9	1	-10	-3	0	**60**	33	9
	Smoke other than cigarette	16	6	-8	12	-22	38	39	37
	Chemical solvents	1	0	-13	-2	-13	47	27	10
	Spray paint	24	-20	**-50**	13	-4	**64**	32	-9
	Welding fumes	14	15	1	3	-8	43	40	-11
	Pesticide	**69**	**67**	10	10	-28	9	8	-2
Diesel/Gas (lifetime)	Worked diesel power, yrs	**85**	**94**	-14	-20	-14	14	6	6
	Worked gas powered equipment, days/yr	**72**	**69**	-3	7	10	-2	-6	-3
	Diesel fuel/fumes/exhaust	**69**	**78**	-6	-2	-16	18	14	-1
Eigenvalue		9.6	9.3	3.1	3.3	2.8	2.9	2.2	2.1

*For ease of presentation, all values were multiplied by 100 and rounded to the nearest integer.
Bolded values represent factor loading score of +0.50 or higher
^Total hours in lifetime = (total yrs)*(total hrs/wk)*(total wks/yr)
+Maximum # = average number of livestock or average number acres of crops in lifetime

CHAPTER III. TLR4 ASP299GLY IS ASSOCIATED WITH CHRONIC BRONCHITIS AND DECREASED LPS-STIMULATED TNF- α PRODUCTION IN AGRICULTURAL WORKERS

Abstract

Introduction: Exposure to lipopolysaccharide (LPS) in the agricultural environment has been associated with increased respiratory symptoms among workers. The toll-like receptor 4 (*TLR4*) gene encodes the receptor for LPS. A mutation at amino acid position 299 (Asp299Gly; rs4986790) of the *TLR4* gene has been associated with *acute* airway and inflammatory hypo-responsiveness to inhaled LPS. We hypothesized that agricultural workers *chronically* exposed to LPS and carry the Asp299Gly polymorphism would be protected from respiratory symptoms.

Methods: We examined the association of *TLR4* tagging single nucleotide polymorphisms (SNPs) with chronic bronchitis in the Keokuk County Rural Health Cohort and the veterans AgLung population. Participants were between the ages of 40-90 years and worked on a farm for at least two years. Chronic bronchitis was defined as having chronic cough and chronic phlegm for three consecutive months for at least two years. Whole blood was stimulated with LPS and assayed for TNF-α by ELISA.

Results: The Keokuk population was primarily non-smoking white males (60.6%) with a mean age of 68.8 ± 11.0 (SD) years. Participants reported working on a farm for 33.3 ± 20.7 (mean ± SD) years. Of the 279 individuals, 14.7% had chronic bronchitis. Farmers with the *TLR4* rs4986790 AG or GG genotypes had a significantly greater odds of chronic bronchitis compared with farmers with the AA genotype (OR_{adj} = 3.27; 95% CI:

1.33, 8.06), even after adjustment for age, body mass index, education, sex, years worked on farm and smoking status. These results were confirmed in the AgLung population (n = 588) (OR_{adj} = 1.88; 95% CI: 0.99, 3.55). The AgLung population consisted of 94.2% white males with a mean age of 64.6 ± 8.5 (SD). Individuals in the AgLung cohort reported working on a farm for 27.7 ± 19.0 (mean ± SD) years. The prevalence of chronic bronchitis in the AgLung population was 32.1%. Individuals with the minor allele at rs4986790 had decreased responsiveness to LPS as measured by TNF- α (p_{adj}=0.029) compared to those with the AA genotype.

Conclusion: We provide evidence that a missense polymorphism in the *TLR4* gene (Asp299Gly) is associated with increased odds of chronic bronchitis in two agriculturally-exposed populations. We also show that the minor allele at Asp299Gly was associated with decreased LPS-stimulated production of TNF- α in a whole blood assay.

Introduction

Chronic bronchitis is a common complaint in farmers, with a prevalence of approximately 25% (23,25). The disease is characterized by airway inflammation and mucus hypersecretion (111,112), and is defined as having chronic cough and chronic phlegm for greater than two years (111,113). Though tobacco smoke is recognized as the number one risk factor for chronic bronchitis, agricultural dust has proven to be a significant cause for the disease (22). Agricultural workers are exposed to high concentrations of airborne inhalable organic dust (114,115), which contains high amounts of lipopolysaccharide (LPS), a component of the gram-negative bacterial cell wall (31,32). LPS exposure has been linked to inflammatory outcomes and several studies have shown a dose-response relationship between LPS and respiratory disease, including chronic bronchitis (33-35,116).

The innate immune response in the airways involves recognizing conserved pathogen-associated molecular patterns (PAMPs) by Toll-like receptors (TLRs) (72,117), ten of which have been identified to date. The TLR4 receptor recognizes LPS, in combination with CD14 (59) and myeloid differentiation protein 2 (MD-2) (60), and activates NF-κB and AP-1 (61) resulting in production of pro-inflammatory cytokines (62,63) such as IL-6 and TNF-α. These inflammatory cytokines are found in high levels among those with respiratory symptoms (64-66); thus, TLR4 may have a role in the development and/or progression of chronic respiratory diseases, such as chronic bronchitis (72).

Functional polymorphisms in *TLR4* have been studied and their role investigated in a number of inflammatory diseases (67), such as disseminated candidiasis (118), sepsis (119), and inflammatory bowel disease (120). These studies have focused on the missense Asp299Gly polymorphism, with the rare allele Gly299 associated with reduced levels of pro-inflammatory cytokines. In the study by Arbour et al, Gly299 was associated

with hypo-responsiveness to inhaled LPS as measured by FEV_1 and displayed increased inflammation in response to LPS inhalation (63,68,69). In another study, investigators found that the Gly299 allele was protective in patients with chronic obstructive pulmonary disease (COPD) (70), while other studies suggested that the presence of the *TLR4* polymorphism, including Gly299 did not have any significant impact on lung function (67,71). These contrasting results were also found by Budulac et al. that showed that none of the *TLR4* single nucleotide polymorphisms (SNPs) were associated with FEV_1 in COPD patients, but found that some were significantly associated with accelerated or reduced decline of FEV_1 (72). However, none of these studies were performed using populations exposed to high levels of LPS.

As part of the present study, we utilized cross-sectional data from two well-characterized populations of agricultural workers to determine if *TLR4* SNPs were associated with decreased production of inflammatory cytokines and a lower risk of developing chronic bronchitis. We hypothesized that agricultural workers chronically exposed to LPS and carrying the minor allele at Asp299Gly would have a decreased prevalence of chronic bronchitis and decreased LPS-stimulated production of IL-6 and TNF-α.

Methods

Study Population and Clinical Assessments. The Keokuk County Rural Health Study (KCRHS) is a population-based, prospective study of health status and environmental exposures of a large, stratified, random sample of residents in one rural Iowa county (121). A cohort of 1004 households (461 lived in town, 341 on farms, 202 in non-farm households) was recruited in Round 1 (1994 to 1998), and of these households, 1002 were followed up in Round 2 (1999-2004), and 662 in Round 3 (2006-2011). In-person interviews were performed at each round to collect information on demographics, smoking habits, general health, and respiratory symptoms. For this study, individuals that had blood collected for genotyping assays (Round 3), had worked on a farm for ≥ 2 years, and were ≥ 40 years of age at baseline were included in the analysis (n=279). All participants provided written consent and the research protocol was approved by the IRB at the University of Iowa and the University of Nebraska Medical Center. The University of Iowa approved transfer of de-identified DNA samples and the Keokuk database to the University of Nebraska Medical Center for genotyping and data analysis, respectively.

The AgLung population is a cross-sectional study designed to assess the relationship between agricultural exposures and chronic respiratory diseases in veterans utilizing the VA Nebraska Western Iowa Health Care System (122). Eligibility criteria for the study included individuals between the ages of 40 and 80 years that had worked on a farm as an adult for ≥ 2 years. Individuals who had been diagnosed by a physician with asthma, lung cancer or interstitial lung disease such as pulmonary fibrosis, sarcoidosis or hypersensitivity pneumonitis were excluded from the study. Recruitment into the study began March 2008 and continued through December 2013 with a total of 681 participants. Demographic information, smoking habits and respiratory symptoms were obtained at the time of enrollment by in-person and telephone interview. The study was

approved by the VA Nebraska Western Iowa Healthcare Systems Institutional Review Board and all participants signed a written informed consent document.

For both populations, chronic bronchitis was defined by the American Thoracic Society guidelines as having chronic cough and chronic phlegm for three consecutive months for at least two years (123). In the Keokuk population, if participants answered yes to both chronic cough and chronic phlegm in either of the three rounds, they were considered to have chronic bronchitis.

In the Keokuk population, airway obstruction was defined as having a pre-bronchodilator $FEV_1/FVC < 0.7$. Whereas, in the AgLung population, airway obstruction was diagnosed based on post-bronchodilator spirometry as $FEV_1/FVC < 0.7$.

Whole Blood Assay. The whole blood assay was performed for the AgLung study and not the Keokuk study. Heparinized blood was diluted in a ratio of 1:1 with antibiotic-free L-glutamine-RPMI 1640 (Life Technologies, Grand Island, NY). Cells (1 mL/tube) were stimulated with either phosphate buffered saline (unstimulated control) or LPS (1 ng/ml). After 24 hours of incubation at 37 °C with 5% CO_2, blood cultures were centrifuged (500 x g, 5 min) and cell-free supernates were collected and stored at −80°C for later cytokine analysis. All samples for the whole blood assay were processed within 2 hours of collection as this has been shown to be important for optimal cytokine measurement (124,125).

TNF-a and IL-6 ELISAs. A lab-developed sandwich ELISA was utilized (126). Flat-bottomed polystyrene microtiter plates were coated with 200 µl/well of purified (goat) anti-human IL-6 or (mouse) anti-human TNF-α antibody (2 µg/mL) (both from R & D Systems) in carbonate buffer (pH 9.6) overnight at 4°C. After three washings in phosphate buffered saline/Tween 20 (PBS-T), the supernates from cell-free whole blood

assay were dispensed in duplicate wells and incubated at room temperature for 2 hours. Plates were again washed three times with PBS-T and incubated with (rabbit) anti-human IL-6 antibody (Sigma-Aldrich, St. Louis, MO) diluted 1:1000 or biotinylated (goat) anti-human TNF-α (1:250) (R & D Systems) in PBS-T/BLOTTO (0.2% instant nonfat milk, PBS-T/B) for 1 hour. After three more washes, human serum-absorbed peroxidase conjugated (goat) anti-rabbit IgG (Rockland Immunochemicals, Limerick, PA) was added at 1:2000 (IL-6) or streptavidin-HRP (1:200, for TNF-α) (R & D Systems) in PBS-T/B for 1 hour. After the plates were washed three times, 200 µl/well of peroxidase substrate (10 ng/ml orthophenylenediamine) containing 0.003% H_2O_2 (Sigma-Aldrich) was added to IL-6 plates, and 100 µL/well TMB substrate (R & D Systems) for the TNF-α plates. The reaction ended with 27.5 µl/well of 8M sulfuric acid, and plates were read using the VERSAmax microplate reader at 490nm or 450nm. An integrated 8-point curve using purified recombinant human proteins was used to determine cytokine concentrations. The limits of detectability for human cytokine assays were: IL-6, 60 pg/mL and TNF-α, 15 pg/mL.

Genetic analysis. Genomic DNA was isolated from whole blood. The complete coding region of *TLR4*, intronic sequence, 5 kb in the promoter region and 2 kb on the 3' prime end was analyzed for adequate coverage. Tagging single nucleotide polymorphisms (SNPs) were chosen based on a minor allele frequency > 10% and linkage disequilibrium (LD) < 0.7 (127). Additional SNPs were included based on their functional significance and relevant citations in the literature. The following SNPs were analyzed for this study: rs11536878, rs11536898, rs4986790 (Asp299Gly), rs4986791 (Thr399Ile), rs5030717, rs5030728, and rs1927911.

Samples were genotyped using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (Sequenom Inc., San Diego, CA). Using the SpectroDESIGNER software (Sequenom Inc.), multiplex PCR assays and associated extension reactions were designed. Primer extension products were loaded onto a 384-element chip with a nanoliter pipetting system (Sequenom Inc.) and analyzed by a MassARRAY mass spectrometer (Bruker Daltonik GmbH, Bremen, Germany). SpectroTYPER RT 4.0 software (Sequenom Inc.) was used to analyze the resulting mass spectra for peak identification. For genotyping quality control, Hardy-Weinberg calculations were performed to ensure that each marker was within the expected allelic population equilibrium.

Statistical Analyses. Chronic bronchitis was compared by patient characteristics using chi-square tests. Because there were a small number of individuals homogenous for the minor allele in many SNPs, we combined with those carrying the heterozygous genotype. A combination variable (smoke intensity) was created and contained 5 categories: 1) current and heavy smokers, 2) current and light smokers, 3) former and heavy smokers, 4) former and light smokers, and 5) never smokers. Heavy and light smokers were defined as having more than (>) and less than or equal to (≤) the median pack-years, respectively.

Multivariable logistic regression models were examined looking at the association between chronic bronchitis and *TLR4* polymorphisms, while adjusting for age, body mass index (BMI), smoke intensity, education, sex, race, and years worked on a farm. We controlled for BMI because it has been shown that *TLR4* Asp299Gly is a potential risk factor for reduced FEV_1/FVC among obese patients (BMI ≥ 30 kg/m^2) (71). Quantile regression, to specify changes in the median response, was used to determine the association between lymphocyte-adjusted LPS-stimulated IL-6 and TNF-α levels with

TLR4 polymorphisms. The Benjamini Hochberg false discovery rate (FDR) method was used to adjust for multiple comparisons due to examining associations with multiple SNPs.

Results

Subject characteristics.

Study population characteristics stratified by chronic bronchitis status for both populations are summarized in Table 7. There were 279 participants from the Keokuk population and 588 participants from the AgLung population, all with complete data on *TLR4* SNPs and covariates included in the analysis. The majority of the participants from the Keokuk population were male (60%) with a mean age of 68.8 ± 11.0 (SD) (range from 40-90). The prevalence of chronic bronchitis was 15%. In this population, participants with chronic bronchitis were more likely to be older (≥ 81), have less education, a greater history of smoking, and worked on the farm for a greater number of years compared to individuals without chronic bronchitis. The overall prevalence of airway obstruction in this population was very low (2.5%). Reflecting demographic trends in the VA population nationally (128), the AgLung participants were predominately white males (98%) with a mean age of 64.6 ± 8.5 (SD) years (range 40-80). This population included 143 (24%) patients with chronic bronchitis. Participants who had chronic bronchitis were more likely to be current smokers and have airway obstruction. The characteristics of both study populations are compared and presented in Table 11.

Association of TLR4 polymorphisms with chronic bronchitis in the Keokuk and AgLung Population.

The Keokuk and AgLung populations were genotyped for seven polymorphisms at locations rs11536878, rs11536898, rs4986790, rs4986791, rs5030717, rs5030728, and rs1927911 (Table 12). The observed minor allele frequencies are presented in Table E2 and showed no deviation from a population in Hardy-Weinberg equilibrium ($p > 0.05$). The allele frequencies were consistent with those reported from the HapMap Project (http:/www.ncbi.nlm.nih.gov/SNP/index.html) (129) in individuals with Western European

ancestry (Table 12). Linkage disequilibrium among these polymorphisms was modest, except for rs4986790 and rs4986791 that had $r^2 > 0.80$ in both populations (Figure 1).

Because exposure to LPS from the agricultural environment has been associated with increased respiratory symptoms among workers, and that TLR4 is the receptor for LPS, we investigated the association between *TLR4* polymorphisms and chronic bronchitis. Results from univariate and multivariable analyses for the Keokuk population are summarized in Table 8 and Figure 2. In the univariate analysis, we found individuals carrying the minor allele in two polymorphisms (rs4986790 and rs4986791) had a greater prevalence of chronic bronchitis than those homozygous for the major allele (Table 8). Farmers with the *TLR4* rs4986790 genotypes AG or GG had a significantly greater odds of chronic bronchitis compared with farmers with the AA genotype (OR = 3.20; 95% CI: 1.40, 7.30, p-value=0.002). After adjustment for age, body mass index, education, sex, years worked on farm and smoke intensity, the association remained significant (OR_{adj} = 3.27; 95% CI: 1.33, 8.06, p_{adj}-value=0.010) (Figure 2). As expected, *TLR4* rs4986791 was in strong LD with rs4986790 and thus followed a similar trend for association (p=0.001, p_{adj}=0.06; OR_{adj} = 4.53; 95% CI: 1.674, 12.234). In contrast, those with the AG or GG genotype for rs5030728 polymorphism were less likely to have chronic bronchitis (p = 0.001) in the univariate analysis and had a decreased odds of having chronic bronchitis than those with the AA genotype (OR_{adj} = 0.30; 95% CI: 0.14, 0.63, p_{adj} =0.002). After correcting for multiple comparisons using the FDR of 5%, we found rs4986790, rs4986791, rs5030728 were still significantly associated with chronic bronchitis.

The association of rs4986790 and rs4986791 with chronic bronchitis was confirmed in the AgLung population (OR_{adj} = 1.88; 95% CI: 0.99, 3.55 and OR_{adj} = 1.83; 95% CI: 1.00, 3.35, respectively) (Table 9), though did not pass a 5% FDR. To determine whether our results were being driven by the high prevalence of airway

obstruction in the AgLung population, we performed a sensitivity analysis by stratifying by airflow obstruction (yes/no; FEV1/FVC < 0.7). We found that the association between rs4986790, rs4986791 and chronic bronchitis was primarily among those without airway obstruction, yet our statistical power was limited in this analysis (p-value=0.026 for rs4986790) (data not shown). Rs5030728 was not associated with a lower risk of chronic bronchitis in the AgLung population (Table 9).

Association of IL-6 and TNF-α *levels with TLR4 polymorphisms in the AgLung population.*

We examined the association between inflammatory cytokine levels (IL-6 and TNF- α) and *TLR4* polymorphisms. In the unadjusted analysis, rs4986790 AG/GG genotype had decreased levels of stimulated IL-6 and TNF-α production (p = 0.027 and 0.012, respectively) compared to those with the AA genotype (Table 10). The associations did not remain significant in the adjusted analysis (p_{adj}=0.065 and 0.005 respectively). The association did remain significant in the adjusted analysis for rs4986791 with IL-6 and TNF-α production (p = 0.036 and <0.001, respectively) (Table 10). However, when stratifying by airway obstruction, we found significant associations of rs4986790 and rs4986791 with IL-6 among those without obstruction (padj = 0.032 and padj = 0.029) (Tables 13 and 14).

Discussion

This study explored whether *TLR4* polymorphisms are associated with chronic bronchitis among agricultural workers. We hypothesized that workers who are exposed to agricultural dust containing LPS and who are carriers of the minor allele at *TLR4* Asp299Gly would be protected against chronic bronchitis. Our study, however, found the converse and presents new evidence for the positive relationship between Asp299Gly polymorphism and chronic bronchitis. Agricultural workers with the minor allele at *TLR4* Asp299Gly or Thr399Ile, had a 2-4 greater odds of chronic bronchitis than those homozygous for the major allele. Our results also indicate that individuals with the *TLR4* Asp299Gly or Thr399Ile polymorphism had lower levels of LPS-stimulated IL-6 and TNF-α from a whole blood assay.

In the Keokuk population, the prevalence of the Gly299 allele was 4.2% in those without chronic bronchitis and 12.2% in participants with chronic bronchitis. In the AgLung population, the prevalence was 4.3% in those without chronic bronchitis and 7.0% in participants with chronic bronchitis. The overall prevalence of the minor homozygous variants and heterozygous variants was low in both populations, therefore we used a dominant model. We looked at the association of *TLR4* polymorphisms and chronic bronchitis in two agriculturally-exposed populations. In the Keokuk population, we found that three out of the seven SNPs in the *TLR4* were of significance: two with a greater odds of chronic bronchitis (rs4986790 and rs4986791) and one with a lower odds of chronic bronchitis (rs5030728). In the AgLung population, *TLR4* 299/399 had borderline significance with an increased odds of chronic bronchitis most likely due to the population's level of smoking compared to the Keokuk population. These SNPs were also associated with lower levels of LPS-stimulated IL-6 and TNF- α.

TLRs are a part of the innate immune system which is the first line of defense against infectious and non-infectious microorganisms. They have been largely studied in

microbial and viral infections, inflammation and immune cells (130), but not extensively in respiratory diseases such as COPD and chronic bronchitis. Many studies found no significant associations between lung function and *TLR4* SNPs (71,131). Rohdea et al., however, found that the frequency of the Asp299Gly polymorphism was decreased in patients with COPD (70).

It is difficult to reconcile the inconsistencies of previous studies and this present study since chronic bronchitis is a form of COPD. In this study, we examined the relationship between *TLR4* 299/399 and chronic bronchitis in one fairly healthy population (Keokuk) and one unhealthy population (AgLung). The Keokuk population comprised former or never smokers, with only a few demonstrating evidence of airway obstruction. The AgLung population better represented previous studies, including smokers and patients with airway obstruction. The borderline relationship found in the AgLung population could be masked by smoking. The association between *TLR4* 299/399 and increased odds of chronic bronchitis could also be due to other mechanisms being driven by 1) chronic exposure to agricultural dust since both populations have a long history of working on a farm, 2) polymorphisms affecting mucus production, and/or 3) different TLR pathways.

Agricultural dust may be contributing to an increased risk of chronic bronchitis. Workers, especially those in swine confinements, have been found to be at a greater risk of chronic respiratory symptoms (32,132). In addition, high LPS levels in this environment were associated with the presence of respiratory complaints in workers. In our study, we only adjusted for years worked on the farm; we did not adjust for different exposures nor for current farming practices. In order to unravel the contribution of the agricultural environment to chronic bronchitis in workers, future studies should look at other specific agricultural exposures and include a control population that is not exposed.

It is known that airway mucus hypersecretion is a common pathological feature of chronic obstructive airway diseases including chronic bronchitis (133). The major protein components of mucus, mucins, are directly induced by TLR signaling (134). *TLR2* and co-receptor asialoganglioside gangliotetraosylceramide (asialoGM1) have been found to stimulate IL-8 production and MUC2 expression through NF-κB activation (135-137). This shows the possibility that other genes may be contributing to the increased odds of chronic bronchitis in these two populations. We did not assay polymorphisms in the *TLR2*, *MUC2* or *MUC5AC* gene, work that may be conducted in the future.

Also of note, not all bacterial products signal through the *TLR4* pathway. LPS from gram-negative bacteria signal through other TLRs as well (138). Furthermore, other ligands in organic dust, such as gram-positive bacteria, could be triggering through TLR pathways in these two highly exposed populations.

Many studies have found a relationship between *TLR4* SNPs and reduced levels of pro-inflammatory cytokines and hypo-responsiveness to inhaled LPS (63,68,69). Similar to these studies, our study showed a significant association between *TLR4* 299/399 polymorphisms and decreased levels of LPS-stimulated production of IL-6 and TNF-α (63,68,69) in individuals carrying the risk allele at *TLR4* 299/399. However, in a study conducted by Long et al., it was Asp299Gly, not the Thr399Ile, that was found to be the functional polymorphism as it blunted *TLR4* function, through decreased NF-kβ activity and consequent decreased IL-8 production in response to LPS (139). Therefore, it's the Asp299Gly variant that is likely responsible for impaired responsiveness of TLR4 to LPS and corresponding activation of NF-κB. It seems that those with the *TLR4* polymorphism are more likely to have less airway or alveolar space forms of endotoxin-induced inflammation; however, these individuals may be more susceptible to a systemic inflammatory response initiated or exacerbated by endotoxin (63,140) which could lead to respiratory disease.

Smoking is the primary environmental risk factor for chronic bronchitis, but not all smokers develop the disease. We found the association between *TLR4* 299/399 and chronic bronchitis in two populations with different smoking patterns. The AgLung population consisted of heavy smokers and of those who were current smokers, approximately 13.0% had the rs4986790 polymorphism. The more significant finding came among never smokers in the Keokuk population. A previous study did not find 299Gly in COPD patients who had never smoked (70). However, we found the presence of *TLR4* 299/399 polymorphisms in never smokers. Among those with chronic bronchitis, 22.2% had the Asp299Gly polymorphisms. This suggests that agricultural exposures, independent of smoking, and the Asp299Gly variant may contribute to the development of chronic bronchitis.

One strength of this study is that it is a population-based study in the relatively healthy Keokuk population, compared to previous studies that included patients with COPD and smokers. Another strength is that both populations have worked in Nebraska or Iowa, thus presumably have similar farming practices. A limitation of this study is that a “healthy worker effect” might have occurred, since a large number of people leave the industry because they have developed respiratory issues, and therefore, are unable to remain employed in this industry. This may have caused an underestimation of the results as individuals with respiratory problems might not work in environments with high exposures. We were unable to investigate this further due to the cross-sectional nature of this study. Whole blood assays were used to assess responsiveness to LPS and stimulated cytokine levels were measured in whole blood assays, only in the AgLung population; however, that may not reflect the airway. Additionally, generalizability of these results is limited as both populations included mainly Caucasian farmers from Iowa and Nebraska.

In summary, our study is the first to investigate the association between *TLR4* tagging SNPs and chronic bronchitis in two agriculturally exposed populations. We found Asp299Gly and Thr399Ile are risk factors for chronic bronchitis in agriculturally-exposed populations. These associations provide insights for future investigations to look at the associations between specific agricultural exposures with *TLR4* SNPs and chronic bronchitis in different populations. In addition, future studies should look at the association between TLR4 299/399 and chronic bronchitis in people without agricultural exposures.

Figure Legends

Figure 1. Linkage disequilibrium (LD) between 7 tagging SNPs in the *TLR4* gene. Top: Keokuk population; Bottom: AgLung population. LD values presented as r^2*100.

Figure 2. Adjusted-Association of *TLR4* Polymorphisms with Chronic Bronchitis in the Keokuk Population.
Forest plot of odds ratios (95% CI) of Chronic Bronchitis vs. No Chronic Bronchitis after adjustment for age, BMI, education, sex, smoke intensity, and years worked on a farm. A dominant model was assumed. The following SNPs: rs4986790, rs4986791, and rs5030728 passed a false discovery rate adjustment at the 5% level.

Table 7. Characteristics of the Keokuk and AgLung Study Populations Stratified by Chronic Bronchitis*

Characteristic	Keokuk		AgLung	
	Chronic Bronchitis n=41	No Chronic Bronchitis n=238	Chronic Bronchitis n=143	No Chronic Bronchitis n=445
Sex				
Male	28 (68.3)	141 (59.2)	140 (97.9)	435 (97.8)
Female	13 (31.7)	97 (40.8)	3 (2.1)	10 (2.2)
Age (yrs)				
≤50	3 (7.3)	19 (8.0)	8 (5.6)	31 (7.0)
51-60	7 (17.1)	31 (13.0)	36 (25.2)	86 (19.3)
61-70	6 (14.6)	80 (33.6)	63 (44.0)	214 (48.1)
71-80	13 (31.7)	83 (34.9)	36 (25.2)	114 (35.6)
≥81	12 (29.3)	25 (10.5)	0 (0.0)	0 (0.0)
Race				
White	37 (100.0)§	222 (100.00)§	137 (95.8)	428 (96.2)
Other	0 (0.0)	0 (0.0)	6 (4.2)	17 (3.8)
BMI				
<25	5 (12.2)	33 (13.9)	19 (13.3)§	62 (13.9)
25-29.9	14 (34.1)	92 (38.7)	47 (32.9)	130 (29.2)
≥ 30	22 (53.7)	113 (47.5)	77 (53.8)	253 (56.9)
Education				
≤ High school	26 (63.4)	131 (55.0)	61 (43.9)§	187 (43.3) §
> High school	15 (36.6)	107 (45.0)	78 (56.1)	245 (56.7)
Smoking Status				
Current	0 (0.0)	0 (0.0)	55 (38.5)	69 (15.5)
Former	14 (34.2)	68 (28.6)	77 (53.9)	266 (59.8)
Never	27 (65.8)	170 (71.4)	11 (7.7)	110 (24.7)
Pack-years (median, range)	23.0 (9.0, 43.5)	12.5 (2.7, 30.0)	34.3 (11.0, 54.23)	14.3 (0, 43.0)
Airflow obstruction				
Yes	3 (7.3)	4 (1.7)	77 (54.2) §	150 (34.1) §
No	38 (92.7)	234 (98.3)	65 (45.8)	290 (65.9)
Worked on Farm (yrs)	40.8 (21.6)	32.1 (20.3)	25.1 (18.6)	28.5 (19.1)

Abbreviations and Definitions: Airflow obstruction, $FEV_1/FVC < 0.70$; BMI, body mass index (kg/m^2); chronic bronchitis, chronic cough + chronic phlegm for 3 consecutive months for ≥ 2 years.

*Data are presented as n (%) except for: pack-years, median (interquartile range); years worked on a farm, mean ± SD.

§Category totals do not add up to column totals due to missing values.

Table 8. Association of *TLR4* Polymorphisms with Chronic Bronchitis in the Keokuk Population*

Polymorphisms	Chronic Bronchitis MAF (n, %)	No Chronic Bronchitis MAF (n, %)	p-value
rs11536878	28 (9.8)	34 (12.2)	0.673
rs11536898	36 (13.9)	32 (12.2)	0.811
rs4986790	34 (12.2)	12 (4.2)	0.002^
rs4986791	32 (12.5)	9 (3.4)	0.001^
rs5030717	42 (15)	26 (9.1)	0.105
rs5030728	62 (19.5)	98 (30.7)	0.001^
rs1927911	87 (31.1)	69 (24.6)	0.179

*Assuming a dominant model.
^Passed a false discovery rate adjustment at the 1% level.

Table 9. Association of *TLR4* Polymorphisms with Chronic Bronchitis in the AgLung Population^

Polymorphisms	Chronic Bronchitis MAF (n, %)	No Chronic Bronchitis MAF (n, %)	p-value	p_{adj}-value*	Odds Ratio (95% CI)*
rs11536878	72 (12.2)	71 (12.0)	0.702	0.403	1.226 (0.761, 1.9750
rs11536898	84 (14.6)	78 (13.7)	0.613	0.589	1.133 (0.720, 1.785)
rs4986790	41 (7.0)	25 (4.3)	0.037	0.053	1.876 (0.991, 3.552)
rs4986791	43 (7.5)	28 (4.9)	0.083	0.050	1.829 (0.999, 3.348)
rs5030717	60 (10.1)	69 (11.8)	0.733	0.895	0.967 (0.590, 1.585)
rs5030728	189 (32.4)	194 (33.1)	0.765	0.262	0.790 (0.523, 1.193)
rs1927911	156 (26.6)	142 (24.2)	0.984	0.779	1.060 (0.704, 1.597)

Abbreviations and Definitions: MAF, Minor Allele Frequency; OR, Odds Ratio; CI, Confidence Interval.

^Assuming a dominant model.

*Multivariable results are adjusted for age, BMI, education, sex, smoke intensity, race, and years worked on a farm.

Table 10. Association of rs4986790, rs4986791, and rs5030728 Polymorphisms with IL-6 and TNF-α levels in the AgLung Population^

Polymorphism	LPS-Stimulated Cytokine§	β	SE	95% CI	p-value	β_{adj}*	SE_{adj}*	95% CI_{adj}*	p-$value_{adj}$*
rs11536878									
	IL-6	-0.039	0.064	-0.164, 0.086	0.542	-0.063	0.057	-0.176, 0.049	0.271
	TNF-α	-0.150	0.111	-0.369, 0.068	0.178	-0.310	0.113	-0.532, -0.088	0.006
rs11536898									
	IL-6	-0.020	0.071	-0.159, 0.120	0.784	-0.035	0.060	-0.153, 0.083	0.560
	TNF-α	-0.060	0.103	-0.262, 0.142	0.558	-0.119	0.098	-0.310, 0.073	0.225
rs4986790									
	IL-6	-0.225	0.096	-0.414, -0.036	0.027	-0.231	0.125	-0.476, 0.015	0.065
	TNF-α	-0.414	0.159	-0.726, -0.101	0.012	-0.359	0.129	-0.613, 0.106	0.005
rs4986791									
	IL-6	-0.231	0.087	-0.403, -0.060	0.012	-0.198	0.094	-0.383, -0.013	0.036
	TNF-α	-0.421	0.155	0.119, 0.722	0.008	-0.459	0.114	-0.683, -0.234	<0.001
rs5030717									
	IL-6	-0.050	0.059	-0.121, 0.111	0.930	-0.035	0.053	-0.138, 0.068	0.505
	TNF-α	0.005	0.128	-0.246, 0.256	0.967	-0.012	0.107	-0.222, 0.198	0.909
rs5030728									
	IL-6	0.034	0.053	-0.069, 0.138	0.516	0.057	0.054	-0.049, 0.163	0.293
	TNF-α	0.143	0.096	-0.046, 0.332	0.138	0.148	0.103	-0.055, 0.351	0.152
rs1927911									
	IL-6	-0.009	0.053	-0.114, 0.096	0.868	-0.032	0.055	-0.141, 0.077	0.564
	TNF-α	-0.106	0.104	-0.311, 0.100	0.310	-0.082	0.095	-0.268, 0.105	0.392

Abbreviations and Definitions: SE, Standard Error; CI, Confidence Interval; BMI, body mass index (kg/m^2).
^Assuming a dominant model.
*Multivariable results are adjusted for age, BMI, education, sex, smoke intensity, race, and years worked on a farm.
§Whole blood stimulated ± lipopolysaccharide (1 ng/ml) for 24hrs. Cytokines measured by ELISA.

Figure 1.

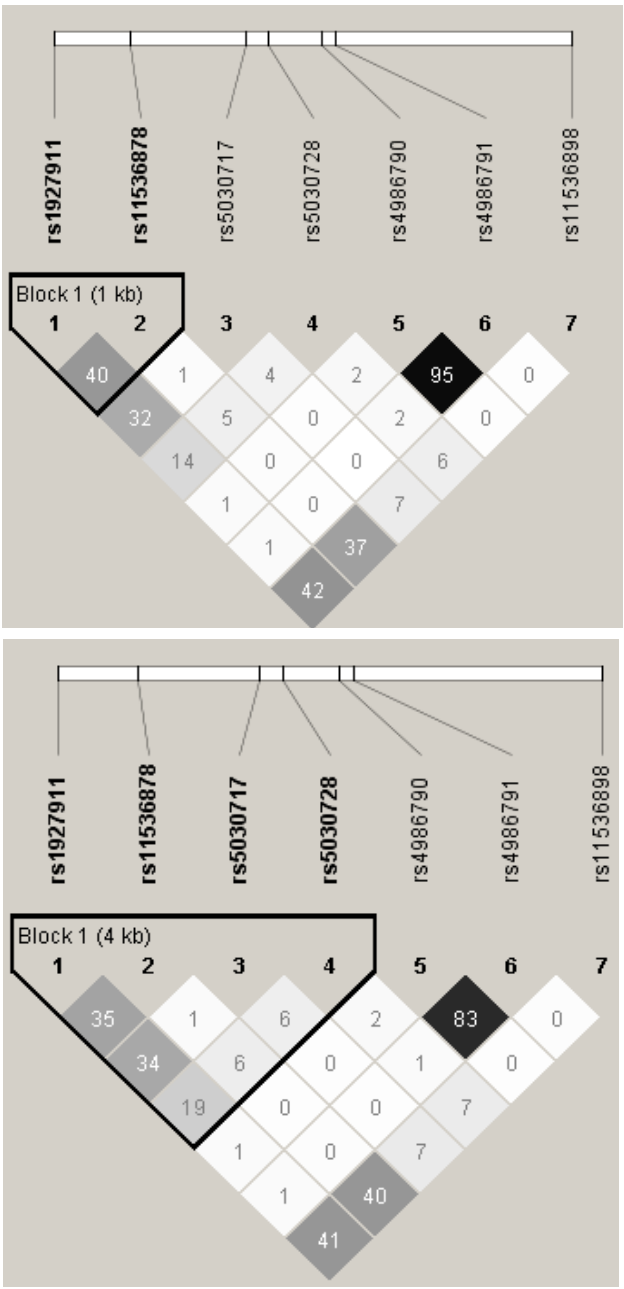
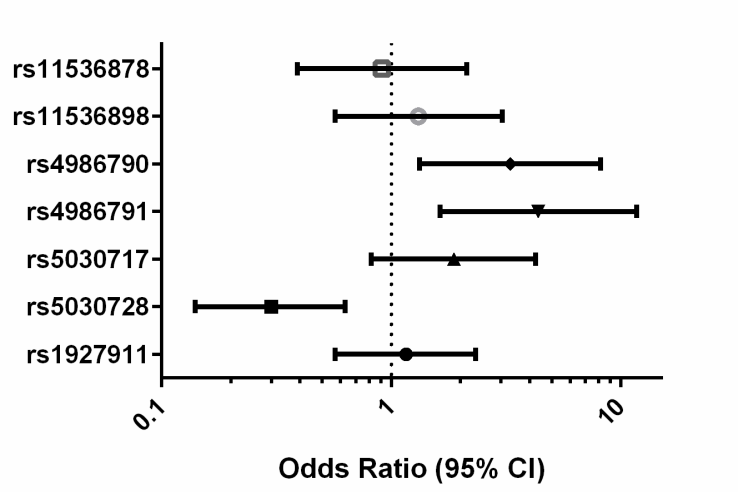

rs1927911
rs11536878
rs5030717
rs5030728
rs4986790
rs4986791
rs11536898
Block 1 (1 kb)
1 2 3 4 5 6 7
40 1 4 2 95 0
32 5 0 2 0
14 0 0 6
1 0 7
1 37
42
rs1927911
rs11536878
rs5030717
rs5030728
rs4986790
rs4986791
rs11536898
Block 1 (4 kb)
1 2 3 4 5 6 7
35 1 6 2 83 0
34 6 0 1 0
19 0 0 7
1 0 7
1 40
41

Figure 2


rs11536878
rs11536898
rs4986790
rs4986791
rs5030717
rs5030728
rs1927911
0.1
1
10
Odds Ratio (95% CI)

Supplementary Data

TLR4 Asp2999Gly is Associated with Chronic Bronchitis and Decreased LPS-Stimulated TNF- α Production in Agricultural Workers

Table 11. Characteristics of Study Populations*

Characteristic	Keokuk n=279	AgLung n=588	p-value§
Sex			<0.001
Male	169 (60.6)	575 (97.8)	
Female	110 (39.4)	13 (2.2)	
Age (yrs)			<0.001
≤50	22 (7.9)	39 (6.6)	
51-60	38 (13.6)	122 (20.8)	
61-70	86 (30.8)	277 (47.1)	
71-80	96 (34.4)	150 (25.5)	
≥81	37 (13.3)	0 (0.0)	
Race			<0.001
White	259 (100.0)^	565 (96.1)	
Other	0 (0.0)	23 (3.9)	
BMI			0.075
<25	38 (13.6)	81 (13.8)	
25-29.9	106 (38.0)	177 (30.1)	
≥ 30	135 (48.4)	330 (56.1)	
Education			<0.001
≤ High school	157 (56.3)	248 (43.4)^	
> High school	122 (43.7)	323 (56.6)	
Smoking Status			<0.001
Current	0 (0.0)	124 (21.1)	
Former	82 (29.4)	343 (58.3)	
Never	197 (70.6)	121 (20.6)	
Pack-years (median, range)	15.0 (3.2, 31.0)	20.0 (2.2, 46.5)	0.361
Airflow obstruction			<0.001
Yes	7 (2.5)	227 (39.0)^	
No	272 (97.5)	355 (61.0)	
Worked on Farm (yrs), mean ± SD	33.3 (20.7)	27.7 (19.0)	<0.001

Abbreviations and Definitions: Airflow obstruction, $FEV_1/FVC < 0.70$; BMI, body mass index (kg/m^2).
*Data are presented as n (%) except for: pack-years, median (interquartile range); years worked on a farm, mean ± SD.
^Category totals do not add up to column totals due to missing values.
§Comparison between the Keokuk and AgLung population.

Table 12. Minor Allele Frequencies of *TLR4* Polymorphisms in the Keokuk and AgLung Population

SNP	Location	Minor/Major Allele	Minor Allele Frequency		
			Keokuk n(%)	AgLung n(%)	HapMap-CEU n(%)
rs11536878	Intron	A/C	33 (11.8)	65 (12.2)	27 (11.9)
rs11536898	3' UTR	A/C	32 (12.2)	72.5 (13.9)	13 (11.4)
rs4986790	Asp299Gly	G/A	15 (5.4)	27 (5.1)	24 (3.5)
rs4986791	Thr399Ile	T/C	12.5 (4.8)	29 (5.6)	10 (4.5)
rs5030717	Intron	G/A	27.5 (9.8)	61.5 (11.5)	25 (11.1)
rs5030728	Intron	A/G	87 (30.9)	176.5 (33.2)	69 (30.5)
rs1927911	Intron	T/C	70.5 (25.2)	146 (27.4)	60 (26.5)

Abbreviations and Definitions: MAF, Minor Allele Frequency; UTR, untranslated region; CEU, Utah Residents with Northern and Western European Ancestry.

Table 13. Associations of rs4986790 and rs4986791 Polymorphisms with IL-6 and TNF-α levels in the AgLung Population Among those with Airflow Obstruction^

Cytokine§	Polymorphism	β	SE	95% CI	p-value	β_{adj}*	SE_{adj}*	95% CI_{adj}*	p-$value_{adj}$*
	rs4986790								
IL-6, 1 ng/mL		-0.0073	0.2144	-0.4299, 0.4154	0.9730	-0.0542	0.1667	-0.3830, 0.3200	0.7456
TNF-α, 1 ng/mL		-0.0337	0.2432	-0.5131, 0.4458	0.8899	-0.4016	0.2090	-0.8139, 0.0107	0.0562
	rs4986791								
IL-6, 1 ng/mL		-0.0590	0.1524	-0.3595, 0.2416	0.6992	-0.0390	0.1406	-0.3164, 0.2385	0.7821
TNF-α, 1 ng/mL		-0.0695	0.2243	-0.5118, 0.3728	0.7570	-0.3524	0.1928	-0.7327, 0.0278	0.0691

Abbreviations and Definitions: SNP, Single Nucleotide Polymorphism; SE, Standard Error; CI, Confidence Interval.

*Multivariable results are adjusted for age, BMI, education, sex, smoke intensity, race, and years worked on a farm.

^Assuming a dominant model.

§Whole blood stimulated ± lipopolysaccharide (1 ng/ml) for 24hrs. Cytokines measured by ELISA.

Table 14. Associations of rs4986790 and rs4986791 Polymorphisms with IL-6 and TNF-α levels in the AgLung Population Among those with no Airflow Obstruction^

Cytokine	Polymorphism	β	SE	95% CI	p-value	β_{adj}*	SE_{adj}*	95% CI_{adj}*	p-$value_{adj}$*
	rs4986790								
IL-6, 1 ng/mL		-0.3893	0.1615	-0.7070, -0.0715	0.0165	-0.2918	0.1351	-0.5576, -0.0260	0.0315
TNF-α, 1 ng/mL		-0.5756	0.1532	-0.8770, -0.2743	0.0002	-0.3835	0.2090	-0.7948, 0.0278	0.0675
	rs4986791								
IL-6, 1 ng/mL		-0.3794	0.1365	-0.6479, -0.1108	0.0058	-0.2586	0.1179	-0.4907, -0.0266	0.0290
TNF-α, 1 ng/mL		-0.5531	0.1725	-0.8925, -0.2137	0.0015	-0.3983	0.1897	-0.7716, 0.0250	0.0366

Abbreviations and Definitions: SNP, Single Nucleotide Polymorphism; SE, Standard Error; CI, Confidence Interval.

*Multivariable results are adjusted for age, BMI, education, sex, smoke intensity, race, and years worked on a farm.

^Assuming a dominant model.

§Whole blood stimulated ± lipopolysaccharide (1 ng/ml) for 24hrs. Cytokines measured by ELISA.

CHAPTER IV. STABILITY OF THE LUNG MICROBIOME OF COPD PATIENTS

Abstract

Chronic obstructive pulmonary disease (COPD) is an inflammatory disorder characterized by incompletely reversible airflow obstruction. The complexity of the lung microbial community in COPD patients has been highlighted in recent years. Evidence suggests that transplantation, medications, age, and disease severity influence microbial community membership. However, the dynamics of the lung microbiome in stable COPD remain poorly understood.

In this study, we completed a longitudinal 16S ribosomal RNA survey of the lung microbiome on sputum samples collected from 7 COPD subjects at baseline and follow-up (≤9 months). Our analysis revealed no significant shifts in the abundance (>2-fold) of taxa between the two time points. Bacterial composition and the number of operational taxonomic units (OTUs) remained the same over time (p-value = 0.75). Also, we collected replicate samples from four patients over a two-day period and found strong clustering of samples for each individual. Results from this study suggest that the lung microbiome is relatively resilient compared to other human microbiomes which have been shown to change over a similar period of time. This study furthers our understanding of the dynamics of the lung microbiome in COPD patients.

Introduction

Chronic obstructive pulmonary disease (COPD), a disease characterized by persistent airflow limitation and chronic inflammation (12), is the third-leading cause of death in the United States. Although tobacco smoke is the leading risk factor for COPD, it is estimated that 15% of COPD cases are caused by occupational exposures, independent of smoking (15). Understanding COPD can be very challenging as it is a heterogeneous disease caused by a number of factors (i.e., environmental exposures and genetics).

To date, there has been controversy regarding the role of lower respiratory tract bacteria in COPD pathogenesis (73). It was previously believed that the lungs of healthy individuals were sterile, while the lungs of COPD individuals were colonized. Recently, with the help of new culture-independent methods, it has been discovered that healthy lungs are not sterile but are colonized by commensal bacteria. This microbial community is modified in individuals with lung disease. Pathogenic bacteria of the phylum Proteobacteria are more commonly colonized in COPD individuals including *Haemophilus* spp. and *Moraxella* spp. (74,75). From bronchoalveolar lavage samples, a study showed that *Haemophilus* species were more frequent in COPD patients, while *Bacteroidetes* were more prevalent in control patients (75).

Few longitudinal studies examining the change of the lung microbiome over time have been conducted (78,79). One study examined the lung microbiome during acute exacerbations and found lower abundances of genera *Moraxella* and *Streptococcus* in sputum samples compared with samples taken from the same COPD patients characterized as being clinically stabile (78). Another study identified increases in *Haemophilus, Pseudomonas, and Moraxella* during exacerbations compared to paired sampling from periods of clinical stability in the same severe COPD patients (79). Respiratory viral infections, especially rhinoviruses, are a major cause of COPD exacerbations (80). Molyneaux et al. investigated the effect of rhinovirus infection on the

airway bacterial microbiome and discovered that rhinovirus infection in COPD alters the respiratory microbiome (81). An increase in Proteobacteria, most notably in *Haemophilis*, was observed in patients with the infection compared to healthy individuals. All of these studies have shown that changes in the lung microbiome in patients with COPD occur when there are exacerbations and respiratory infections (82-84). As yet, there are no longitudinal studies comparing baseline sputum samples with samples collected over time in stable COPD patients without any sort of intervention.

The aim of this study was to examine the lung microbiome longitudinally in individuals with stable COPD using 16s rRNA sequencing. We hypothesized that the lung microbiome of COPD patients is relatively stable within a 9 month time period. To assess the inter-variability of the lung microbiome, we collected induced sputum at baseline and at follow-up (≤ 9 months) from 7 patients with stable COPD (stage 2) and determined whether the microbial community changed over this time period. To assess intra-variability of the lung microbial community, we collected replicate samples from four patients over a two-day period and examined the community variability within each patient.

Methods

Patient recruitment

The AgLung cohort is a cross-sectional study of agricultural exposures and respiratory disease among veterans visiting the General Medicine clinics at the VA Nebraska-Western Iowa Health Care System in Omaha, NE (122). Other than working on a farm for more than two years, eligibility criteria included being between 40 and 80 years of age and no history of lung cancer, metastatic cancer to the lungs or interstitial lung disease such pulmonary fibrosis, asthma, sarcoidosis, hypersensitivity pneumonitis. Seven COPD patients from the AgLung cohort were included in the present analysis. Patients were eligible for this study if they were former smokers and their post-bronchodilator forced expiratory volume in one second (FEV_1)/forced vital capacity (FVC) ratio was ≤ 0.7 and their FEV_1 was 50-79% predicted (moderate COPD, stage 2) (11). COPD was defined as FEV_1/FVC ≤ 0.7 or a diagnosis of COPD from a pulmonologist. Eligible individuals had stable respiratory symptoms, were afebrile, and had not taken any antibiotics or corticosteroids for two months prior to sputum induction.

Sputum collection

For the longitudinal phase of this study (inter-subject variability), two induced sputum samples were obtained, one at baseline and another within nine months of follow-up, using an established induced sputum protocol developed by the NIH-sponsored SPIROMICS study for COPD (141). Briefly, three 7-minute inhalations of nebulized hypertonic saline (3%) was followed by expectoration of the sputum. To minimize oral contamination, all subjects performed an oral rinse with mouthwash (Cepacol, Reckitt Benckiser, Parsippany, NJ, USA), sterile water, and molecular grade water in triplicate, respectively, prior to obtaining an induced sputum sample.

For the intra-subject variability experiments, replicate samples from four patients over a two-day period were collected. Participants went through the same three 7-minute inhalation protocol as described above. A baseline sample was collected in the morning on day one, another 4-5 hours later, and then again the next morning. One patient gave a fourth sample in the afternoon on the second day.

Sputum processing, DNA extraction, PCR amplification, and pyrosequencing

Sputum was processed in a modification of the method developed by Neil Alexis (142). Briefly, freshly-collected sputum was weighed and 0.9 mL of the sample including all mucous plugs was solubilized in 0.1% dithiothreitol, diluted 4-fold with EDTA, and filtered through a 0.48 mm mesh strainer. Cells isolated from this fraction were assessed for squamous cell contamination (>40% squamous contamination excluded the sample). DNA was isolated from the remaining solubilized sputum using a bead beating, solvent extraction method (PowerSoil DNA isolation Kit, Mo Bio) according to the manufacturer's instructions. An approximately 460 bp sized fragment of the V3 / V4 region of the 16S rRNA gene was amplified (25 cycles of PCR) for each of the DNA samples beginning with 12.5 ng of DNA per Illumina's recommended protocol outlined in the 16s Metagenomic Sequencing Library Preparation protocol (Illumina, San Diego, CA). 16S Amplicon PCR Forward Primer = 5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG 16S Amplicon PCR Reverse Primer = 5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC. Following generation of the amplicons, dual indices and Illumina sequencing adapters were added (8 cycles of PCR) using the Nextera XT Index kit (Illumina catalog # FC-131-1001). Resultant libraries were multiplexed and 300 bp paired end sequencing was performed on an Illumina MiSeq instrument using V3 chemistry per Illumina's

recommendations. An ample yield of approximately 10,000 reads/sample was used to characterize the bacterial community from each sample.

Data Processing and Bioinformatics Analysis

Samples with paired-end Illumina reads were filtered to trim the low quality bases from 3' as well as 5' ends using Cutadapt package (DOI: http://dx.doi.org/10.14806/ej.17.1.200). Given a significantly lower quality of the bases at 3' end of the reads, the quality-value (Phred Score) cutoffs for 3' and 5' bases were 35 and 30, respectively. Quality-trimmed read pairs with the minimum length of 150bp were retained, and used in the subsequent analysis performed using Illumina Base-Space cloud platform for processing the 16s rRNA data. The total read number in each of our samples ranged between 10,056—40,133.

In the downstream analysis, Qiime Preprocessing and Visualization apps were used (143). Post-QC reads were aligned against the Greengene database (144). Following the taxonomy assignment, the number of sequences assigned to a particular phylotype and the percentage of these sequences in the microbial community were calculated for each sample. Samples with less than 5,000 total reads were excluded. Reads with more than 97% identity were tallied to make the counts and percentages tables, with each row representing a different phylotype. Alpha (Shannon Index) and Beta (Unifrac weighted) diversity scores were calculated. Unifrac distance metric was used for calculating the Principal Component Analysis (PCoA) plots. Community membership and structure were examined using the PCoA plots to determine relatedness among samples. Heat maps showing similarity or dissimilarity among samples were generated.

Statistical Analysis

For inter-subject variability experiment, samples from patients were grouped according to the time when the sputum sample was taken, baseline (T1) versus months later (T2). The taxon-based method was used to analyze the 16S rRNA sequence of each sample. The diversity, richness, and composition of the microbial communities were compared among samples, as was the association of the relative abundance of bacteria, using paired t-tests. For Figure 3B showing the differences in prevalence (%) of taxonomy, T2 minus T1 was calculated.

A Wald test was used to compare the prevalence of the taxa between the two sample groups (T1 vs. T2) as follows. At different taxonomic levels, the relative abundance was compared by the DESeq2 method to identify the taxa among the groups, while accounting for the paired nature of the data (143). The Benjamini and Hochberg false discovery rate (FDR) approach was used to adjust the raw p-values to account for multiple comparisons (145). Taxa yielding an FDR adjusted p-value < 0.05 were considered statistically significant. The diversity and richness indices and their 95% CIs were calculated using the Shannon Index, which is a measure of both richness and evenness.

Results

Lung microbiome profiles

Sputum samples from seven subjects were obtained at baseline and at follow-up (≤ 10 months). Patient demographics are summarized in Table 15, sorted by sample collection date within patient. The majority of the participants (87.5%) was over the age of 60 and worked on the farm for greater than 20 years (62.5%). The mean FEV_1/FVC ratio and % predicted FEV_1 was 63.77 and 67, respectively. DNA sequencing reads (337,474) were selected after demultiplexing and quality control filtering, with each sample averaging 24,105 sequences. Unique OTUs (1,632) were identified across 14 samples and the number of OTUs observed at 97% identity ranged from 10,056-40,133 (Table 15). There were no significant differences in the numbers of sequences obtained from baseline and at the later time point (p-value = 0.86) (Table 15). Additionally, no differences were found in the richness between the two groups [Time 1 (T1) vs Time 2 (T2)] (p-value = 0.75) (Figure 8). A heat map of OTU abundances, created using R, visually confirmed this result (Figure 6A and 6B). Rarefraction curves were calculated for all samples and maximum OTUs were detected at 1500 sequences and leveled off, showing that additional sampling would not have detected additional OTUs (Figure 8). There was no difference in the Shannon diversity (alpha-diversity) between the two time points (p-value = 0.32) (Figure 8).

Bacterial community composition between Time =1 and Time =2 (inter-subject variability).

Based on overall phyla composition, the samples were composed of four major groups: Proteobacteria, Firmicutes, Bacteroidetes, and Actinobacteria (Figure 3A). Approximately 90.5% of sequences belonged to one of four phyla: Firmicutes (44.1%), Bacteroidetes (23.3%), Proteobacteria (12.3%), or Actinobacteria (10.8%). Of the 183

genera identified, the most abundant were *Veillonella* (17.2%), *Prevotella* (11.8%), *Streptococcus* (11.1%), *Rothia* (7.6%), and *Haemophilus* (3.9%), all of which are typical members of the lung microbiota (146) (Figure 4A). These top ten genera were 100% prevalent across all samples from both visits (T1 and T2), suggesting that for our cohort the lung microbiome was relatively stable over this time period (147).

The phylum level abundance variations between the two time points are presented in Figure 3B. The largest change was found in patient ID 2150. Actinobacteria changed from 8.2% to 28.9%, Firmicutes increased from 38.2% to 64.8%, and Proteobacteria dropped from 37.9% to 1.5%. The sample taken six months later (Shannon Index of 2.52) was less diverse than this patient's baseline sample (Shannon Index of 2.91). For example, Fusobacteria decreased to 0% from 6.6%.

The genus level abundance variations between two time points are presented in Figure 4B. Again, patient ID 2150 had the largest variations. *Rothia* increased to 25.8% from 8.0%, *Streptococcus* increased to 28.7% from 3.7%, and *Haemophilis* decreased to 0.6% from 13.4%. Of note, ID 2150 at time = 1 had the largest percentage of *Haemophilus* (13.4%) compared to the rest of the samples.

In order to evaluate the inter-subject variability over time, principal coordinate analysis (PCoA) was performed using weighted UniFrac (Figure 5). This analysis did not reveal any major clustering between the two groups (T1 and T2) suggesting small variations between the two groups. All but two samples clustered around PC1. Patient ID 2150 did not cluster with the other samples or with his other time point. This patient had complex medical issues and was the only patient who worked on the farm for fewer than 20 years.

To further determine if there were any differences between T1 and T2 taxa, we used the Wald test to detect whether there was evidence of differential relative abundance in taxa between samples. However, zero OTUs were significantly different

among the two groups (T1 and T2). To examine the clinically stability at two time points, we used a volcano plot (Figure 7) and, again, determined there were no taxa that exhibited differential relative abundance, based both on visual inspection and statistical thresholds. However, there were many OTUs that had log2-fold change greater than 1 (indicated in gold).

Bacterial community composition within patients (intra-subject variability).

We also examined the intra-subject variability of the lung microbiome of four patients over a two-day period. All patients were seen at least three times: Day 1, morning (D1_1); Day 1, afternoon (D1_2); Day 2, morning (D2_1); and one patient was seen Day 2, afternoon (D2_2). Sample D2_1 for patient ID 1204 had very low reads (<5,000), and therefore was excluded from the analysis. Because our sample size was very small, we could not effectively use the Wald test to detect differentially abundance features between samples. Instead, we evaluated the Shannon Index for all samples (Figure 11). Similar to the inter-subject variability, the rarefraction curves showed that additional sampling would not have provided detection of additional OTUs. In addition, the Shannon Index was visibly similar among samples (Figure 11). PCoA plots showed samples taken from the same patient clustered with each other (Figure 10). Similar OTU abundance banding at each time point for each patient were observed (Figure 13), suggesting there is little intra-subject variability of the lung microbiome in these patients.

Similar to the inter-subject variability part of this study, 88.9% of sequences belonged to one of four phyla: Firmicutes (38.9%), Bacteroidetes (25.1%), Proteobacteria (17.9%), and Actinobacteria (7.0%) (Figure 12A). Again, the genus *Veillonella* (14.8%) was the most abundant in these 12 samples, followed by *Prevotella* (8.8%), *Streptococcus* (8.8%), and *Rothia* (5.7%). *Haemophilus* was the 10th top genera among these samples (4.3%) (Figure 12B).

Discussion

This study uses induced sputum samples to explore the short-term stability of the lung microbiome in clinically stable COPD patients. We hypothesized that the lung microbiome of individuals with COPD would have similar bacterial compositions and diversities during a short time period (≤ 2 days) or longitudinally (≤ 9 months). Our data shows some variation in the microbial profile of some COPD patients over 2-9 months. Furthermore, our results suggest a high degree of similarity of the lung microbiome when sampled over 2 consecutive days.

In this study, we investigated the sputum bacterial microbiome of seven COPD patients using 16S rRNA gene sequencing. The identified dominant phyla in these 14 samples represent common bacterial compositions found in the lung (74,148). The most abundant genera were *Veillonella*, *Prevotella*, *Streptococcus*, *Rothia*, and *Haemophilus*. Another genus that was represented in these samples was [*Prevotella*] (7.1%). The square brackets indicate taxonomic changes that were recommended, mainly based on genome trees. This means that this is a suggested, but not a verified taxonomy. We identified similar diversities and OTU abundances when comparing T1 and T2 for each patient. We did not observe any statistically significant differences in the diversity, nor did we observe differences in OTU abundances among the two groups (T1 vs T2). These results suggest a stability of the lung microbiome within a 9 month period.

There were no differences in the Shannon Diversity Index between T1 and T2 samples for each patient. Figure S-2 confirmed this using other alpha-diversities (i.e. Chaos Index, Simpson Index). Most T1 and T2 samples clustered around PC1 in the PCoA plot, except for patient ID 2150.

The intra-subject variability of 2-4 sputum samples from 4 patients was examined over a two-day period. Similar bacterial compositions were identified in these observations. The dominant phyla included Proteobacteria, Firmicutes, Bacterioidetes,

and Actinobacteria. The abundant genera included *Veillonella*, *Prevotella*, *Streptococcus*, and *Rothia*, all of which are typical members of the lung microbiota (healthy and COPD) (146). *Haemophilus* represented the 10th top genus in these samples.

The lung microbiome of healthy individuals has been shown to consist of bacteria from the Bacteriodetes, Firmicutes, Proteobacteria, and Actinobaterica phyla (74,148), which are all phyla found in our samples. In this study, Firmicutes was the most prevalent phylum, a phylum primarily composed of gram-positive microorganisms, followed by Bacteriodetes. A study conducted by Hilty et al. showed Bacteroidetes, particularly *Prevotella* spp., was more prevalent in controls than in COPD patients (75), while Proteobacteria, a gram-negative phylum, was more prevalent in COPD patients compared to controls (75).

We demonstrated similar genera in COPD patients compared to prior studies including *Streptococcus, Prevotella, Moraxella,* and *Haemophilus* (73,149). Different techniques such as expectorated sputum, bronchial aspirate, bronchoalveolar lavage (BAL), and bronchial mucosal biopsies, have all found these genera in COPD patients (149). In this study, only *Streptococcus* and *Haemophilus* (not *Pseudomonas* or *Moraxella*) were represented in the top ten genera of the 14 samples. Of note, *Haemophilus* was most prevalent in 2150-T1, possibly due to complex medical issues of this patient as described below.

The lung microbiome can be difficult to study, particularly using longitudinal and cross-sectional study designs. In previous studies, factors including interventions (transplantation and medications), age, exacerbations, and disease severity have been shown to influence microbial community membership (75,78,84,147,150). We attempted to control for these covariates during the patient selection phase as described below. Patient ID 2150 showed the greatest differences in the bacterial composition between

the two time points. This patient's medical chart indicated multiple health issues including heart failure, chest pain, and arrhythmia. Although patient ID 2150 was not taking antibiotics two months prior to collection of induced sputum samples, he had taken clindamycin after the first sample and prior to the second sample. Antibiotics have shown to increase microbial diversity, though this did not occur in our observations of this patient (83,147). Our observations did show a large decrease in Proteobacteria and an increase in Firmicutes in this patient, which is in agreement with previous studies regarding antibiotic use among COPD patients (83,147). Age has been shown to be associated with microbial diversity (150). Pragman et al. showed younger age is associated with less microbial diversity in COPD patients (150). Although patient ID 2150 was the youngest compared to others in this study, only his second sample showed the lowest diversity among all samples.

COPD is characterized by natural histories that are punctuated by periods of acute exacerbations. A study conducted by Sethi et al. identified lower densities of *Moraxella catarrhalis* and *Streptococcus pneumonia* in sputum samples collected during acute exacerbations compared with samples during clinical stability (78). Another study examined sputum from COPD patients and identified increases in *Haemophilus, Pseudomonas,* and *Moraxella* during exacerbations compared to paired sampling from periods of clinical stability (79). Both, Huang et al. and Molyneaux et al., found an increase in the phylum Proteobacteria during COPD exacerbation (84,147). What's interesting to note is that *Moraxella* was found in only one sample of this study, ID 2396 (time = 2) at 0.7%.

The relationship between bacterial diversity and COPD severity remains disputed. Studies have reported that bacterial diversity decreases with increased COPD severity (73,75); however, other studies using lung tissue samples have failed to show significant differences in bacterial diversity with increasing COPD severity (76). All of the

patients in our study had stage 2 COPD at the time of enrollment into the AgLung study; therefore, we were unable to investigate the relationship between COPD stage and bacterial diversity. Although patient ID 2326 had a borderline diagnosis of COPD, the bacterial composition and diversity of his samples were similar to those with COPD stage two in this study. Notably, both of patient ID 2326's samples clustered with all other samples (except ID 2150) on the PCoA plot.

This study only included former smokers as current smoking has been shown to have an impact on the lung microbiome in COPD patients (151). Smoking in and of itself does not alter the lung microbial community (77); however, current smokers were found to have lower 16S rRNA copy numbers. And while 16S rRNA copy number cannot be directly compared with measures of bacterial community diversity, as some bacteria can possess multiple copies of the 16S rRNA gene, some studies have shown smokers exhibit greater variation in the relative abundance and composition of bacteria inhabiting the nasal or oropharynx (152). Hilty et al. examined the lung microbiome using airway brushings and included COPD patients that were current smokers (80%) (75). They found Proteobacteria to be the most abundant phylum in COPD patients. Although Proteobacteria was one of the top four phyla in our study, Firmicutes was the most abundant. A handful of studies have evaluated the lung microbiome of COPD patients who smoke; however, additional studies with larger sample sizes are required to definitively determine if the lung microbiome is altered by smoking in those with COPD.

The assessment of the lung microbiome in agricultural workers has not been studied**.** However, shotgun pyrosequencing metagenomic analyses of DNA from dusts from swine confinement facilities and grain elevators have been performed (153). Boissy et al. identified Firmicutes as the predominate phylum in dust from a swine confinement facility. In contrast, Proteobacteria was the most abundant phylum in grain elevator dust (153). Approximately 71.5 % of individuals in this study worked with hogs, either in

confinement or open lots, in their lifetime, while 100% were exposed to grain dust. All patients were former farmers in this study; and therefore, most likely the influence of the agricultural environment did not play a role in influencing the lung microbiome unless there were long term effects. Further studies are needed to evaluate the effect of agricultural exposure on the lung microbiome in those with and without COPD.

To our knowledge, this study is the first to present a longitudinal analysis of the lung microbiome in patients with COPD during clinical stability. One of the admitted limitations of this study is the sample size. In fact, relatively small sample sizes have plagued many lung microbiome studies and most likely limited our ability to observe significant differences between our groups. Inducing sputum from patients is difficult and time-consuming and for safety reasons, we only induced sputum in those with COPD stage two. Additionally, sputum samples have been shown to have significantly lower diversity than other sample types such as BAL, cell-free BAL supernatants, and biopsies (149).

The time between samples taken varied among our patients. Patient ID 1053 had the largest elapsed time period between induced sputum samples (9 months). In this patient, the phylum Firmicutes in T2 decreased from 54.50% to 27.5% and the genus *Prevotella* increased by approximately 27% from the baseline sample (T1=10.1%). Patient ID 2150 had approximately six months between the two time points and we saw a decrease in his microbial diversity. However, in this patient we saw an increase in Firmicutes and a decrease in *Prevotella*. Additional longitudinal analyses with greater patient numbers are needed to help better understand the dynamics of the lung microbiome in relation to COPD. We were unable to control for the time intervals in this study. Coordinating with these patients was difficult as patients were busy, some became sick, and one patient had to cancel due to a sick family member.

In summary, we have presented the induced sputum bacterial profiles of agriculturally-exposed, male patients with moderate COPD based on 16S rRNA gene sequencing. We showed some variation in the lung microbiome of COPD patients over a two to nine month time period, but the lung microbiome remains relatively stable over 2 consecutive days. This study adds further insights into the lung microbiome of COPD patients, with the inclusion of repeated and longitudinal sampling.

Figure Legends

Figure 3. A. Taxonomic Identification at the Phylum Level (Inter-Variability Samples). Taxonomic results at the phylum level are displayed for each sample at two different time points (T1 and T2). B. Differences in Percent of Taxonomic Identification at the Phylum Level (Inter-Variability Samples) (T2-T1) For example, patient 2150 decreased in Proteobacteria by 36.4%. All sequences were submitted to Qiime for taxonomic identification. Top 8 taxonomic results at the phylum level are displayed. Figures were created in Excel.

Figure 4. A. Taxonomic Identification at the Genus Level. B. Differences in percent of taxonomic Identification at the Genus Level (T2-T1). All sequences were submitted to Qiime for taxonomic identification. Top 10 taxonomic results at the genus level are displayed. Figures were created in Excel.

Figure 5. Principal Coordinate Analysis Demonstrates Clustering of Baseline Samples and Samples taken at a Later Time. Principal coordinate analysis was performed using Qiime and Weighted UniFrac, and the results for principal coordinates 1 and 2 and 3 are shown. T1 samples (red) do not cluster separately from T2 samples (blue). However, there are 2 samples that did not cluster with the other samples. Three PCs explained approximately 84% of the total variance in the data.

Figure 6. A. Heat map of OTU Abundances between Baseline and Samples taken at a Later Time. B. Heat map of differences of OTU abundances between T1 and T2 (T2-T1). Lighter blue shows greater abundance compared to dark blue/black. Figures were created in R (package phyloseq)(154).

Figure 7. Volcano plot indicating that no taxa are significantly increased or decreased in the pairwise comparisons indicated, using t-tests (R package calibrate). Results shown are from all 7 participants. Color differences show relative abundance of at least 2-fold, or $\log_2$ equal to or less than 1 (black) versus greater than 1 (gold). Figure was created in R (package phyloseq).

Table 15. Patient Characteristics and Sequencing Results.

Sample ID	Sputum Induction	Age	FEV_1	FEV_1/FVC	BMI	Worked on farm (yrs)	Trimmed Sequences	OTUs^	Shannon Index
2195-T1	Baseline	62	69	60	≥30	21	41796	40133	3.13
2195-T2	2 months						35728	32448	2.98
2150-T1	Baseline	53	59	67	≥30	14	27460	25185	2.91
2150-T2	6 months						10527	10056	2.52
2326-T1	Baseline	70	69	70	≥30	43	16822	16114	3.00
2326-T2	2 months						27872	26533	3.29
1053-T1	Baseline	59	52.5	71	≥30	49	24588	22705	3.23
1053-T2	9 months						38683	35736	3.41
1204-T1	Baseline	67	63.9	62	≥30	48	25139	24421	2.76
1204-T2	3 months						11574	11117	2.72
2397-T1	Baseline	61	74	68	<25	20	16313	15470	3.40
2397-T2	4 months						16548	15674	3.20
2285-T1	Baseline	66	59	71	≥30	48	19517	18763	3.17
2285-T2	4 months						24907	21459	2.65

Abbreviations and Definitions: FEV_1, % predicted forced expiration volume in 1 second; FEV_1/FVC, ratio of forced expiratory volume in 1 second/forced vital capacity (L); BMI, body mass index (kg/m^2), OTU, Operational Taxonomic Units.
^Number of sequences that classified to OTUs.

Figure 3A.

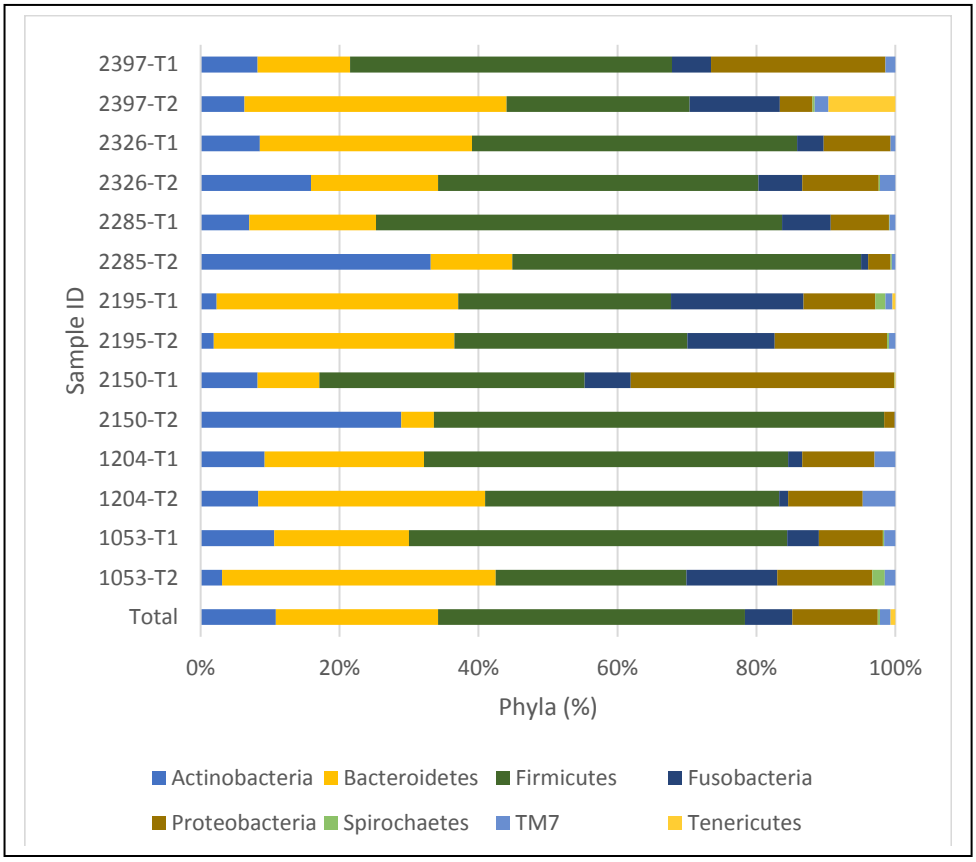

2397-T1
2397-T2
2326-T1
2326-T2
2285-T1
2285-T2
2195-T1
2195-T2
2150-T1
2150-T2
1204-T1
1204-T2
1053-T1
1053-T2
Total
Sample ID
0%
20%
40%
60%
80%
100%
Phyla (%)
Actinobacteria
Bacteroidetes
Firmicutes
Fusobacteria
Proteobacteria
Spirochaetes
TM7
Tenericutes

Figure 3B.

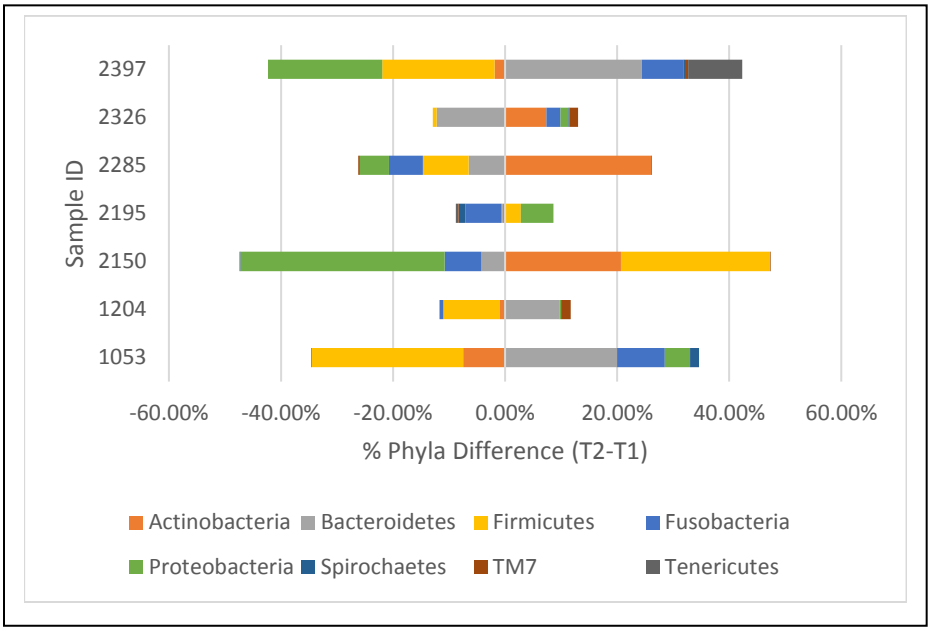

Sample ID
2397
2326
2285
2195
2150
1204
1053
-60.00%
-40.00%
-20.00%
0.00%
20.00%
40.00%
60.00%
% Phyla Difference (T2-T1)
Actinobacteria
Bacteroidetes
Firmicutes
Fusobacteria
Proteobacteria
Spirochaetes
TM7
Tenericutes

Figure 4A.

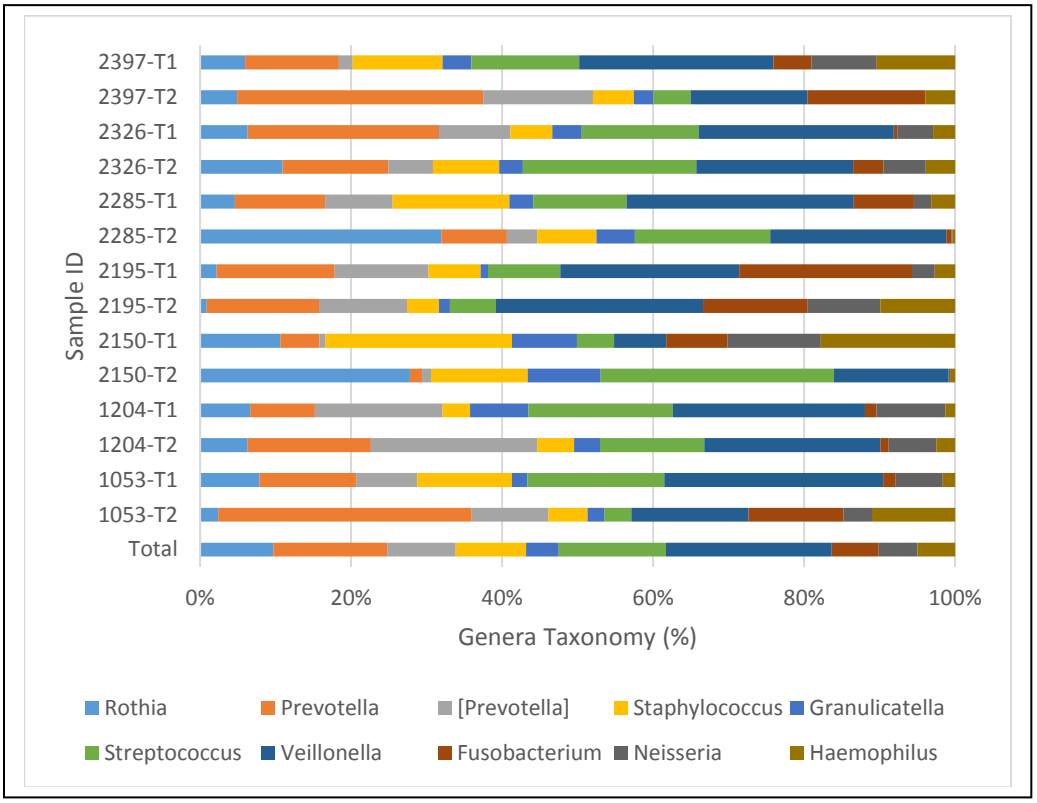

2397-T1
2397-T2
2326-T1
2326-T2
2285-T1
2285-T2
2195-T1
2195-T2
2150-T1
2150-T2
1204-T1
1204-T2
1053-T1
1053-T2
Total
Sample ID
0%
20%
40%
60%
80%
100%
Genera Taxonomy (%)
Rothia
Prevotella
[Prevotella]
Staphylococcus
Granulicatella
Streptococcus
Veillonella
Fusobacterium
Neisseria
Haemophilus

Figure 4B.

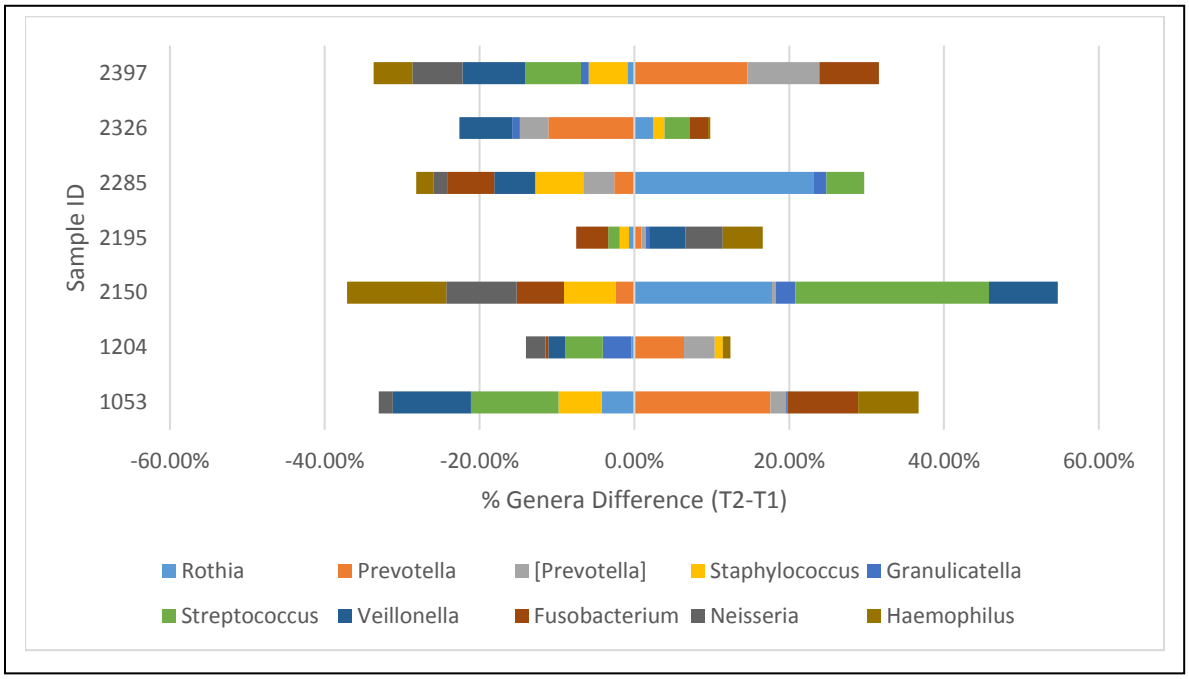

Sample ID
2397
2326
2285
2195
2150
1204
1053
-60.00%
-40.00%
-20.00%
0.00%
20.00%
40.00%
60.00%
% Genera Difference (T2-T1)
Rothia
Prevotella
[Prevotella]
Staphylococcus
Granulicatella
Streptococcus
Veillonella
Fusobacterium
Neisseria
Haemophilus

Figure 5.

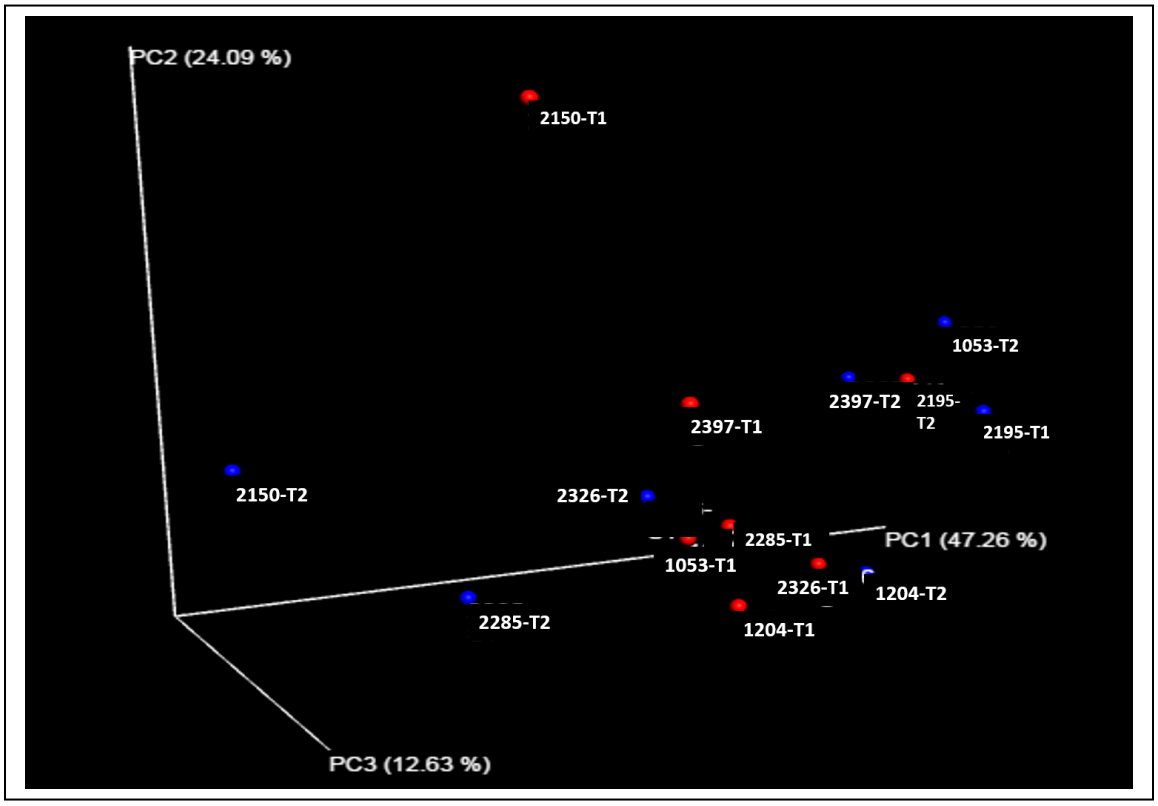

PC2 (24.09 %)
2150-T1
1053-T2
2397-T2
2195-T2
2397-T1
2195-T1
2150-T2
2326-T2
2285-T1
PC1 (47.26 %)
1053-T1
2326-T1
1204-T2
2285-T2
1204-T1
PC3 (12.63 %)

Figure 6A.

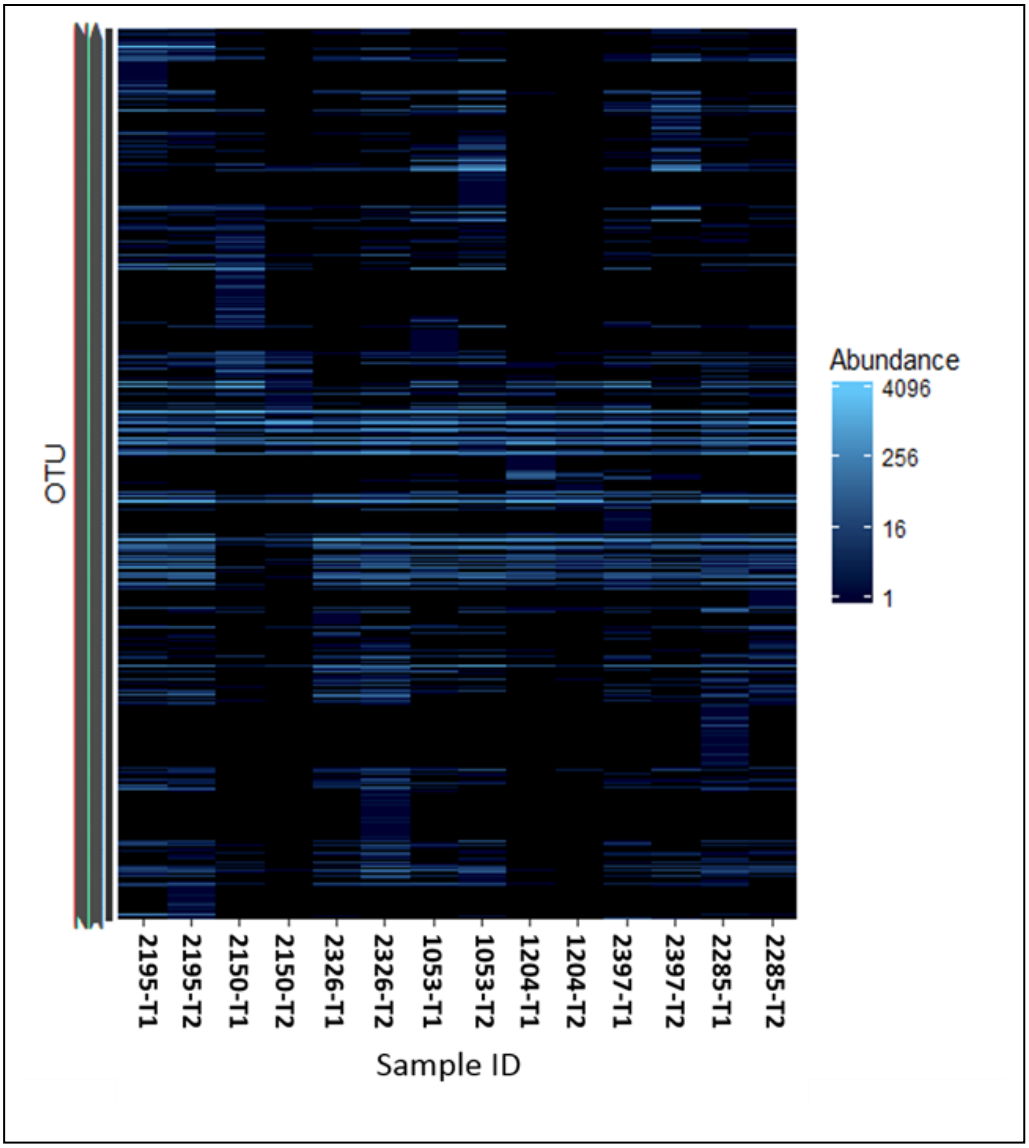

OTU
Abundance
4096
256
16
1
2195-T1
2195-T2
2150-T1
2150-T2
2326-T1
2326-T2
1053-T1
1053-T2
1204-T1
1204-T2
2397-T1
2397-T2
2285-T1
2285-T2
Sample ID

Figure 6B.

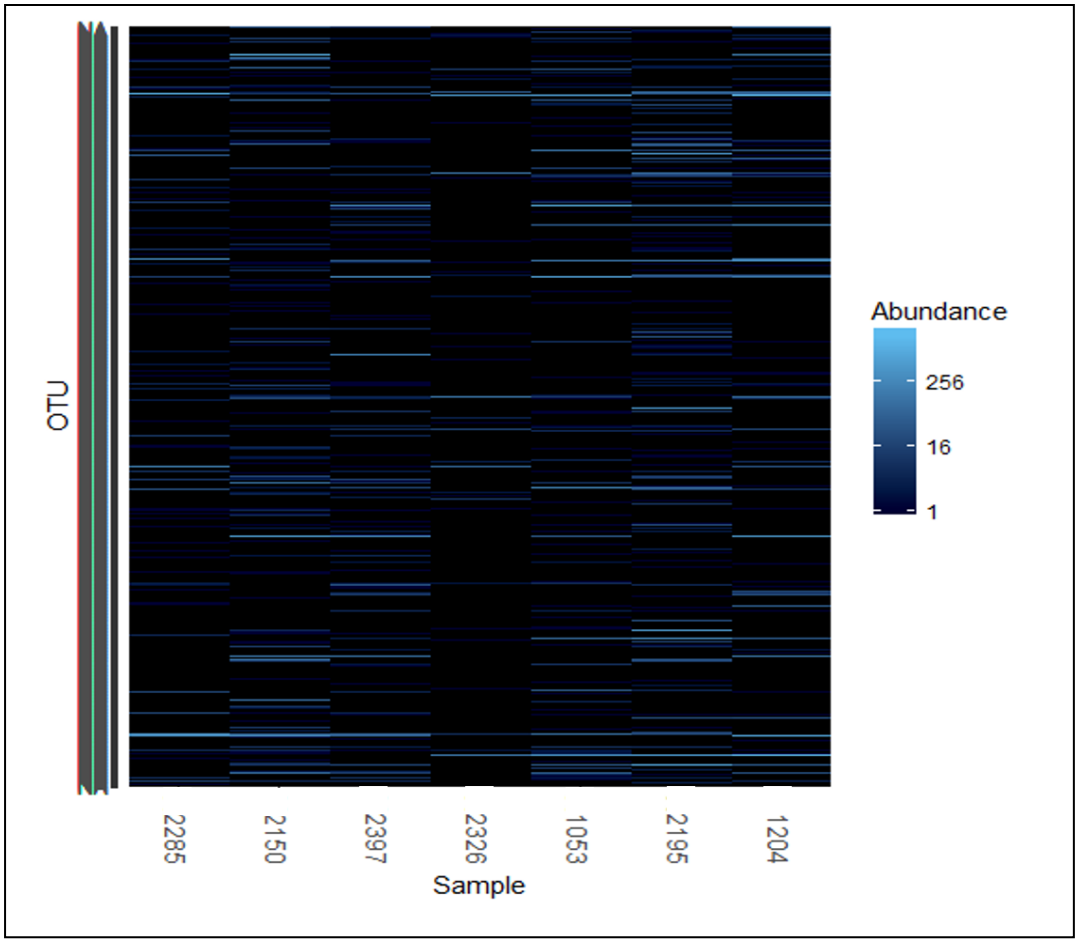
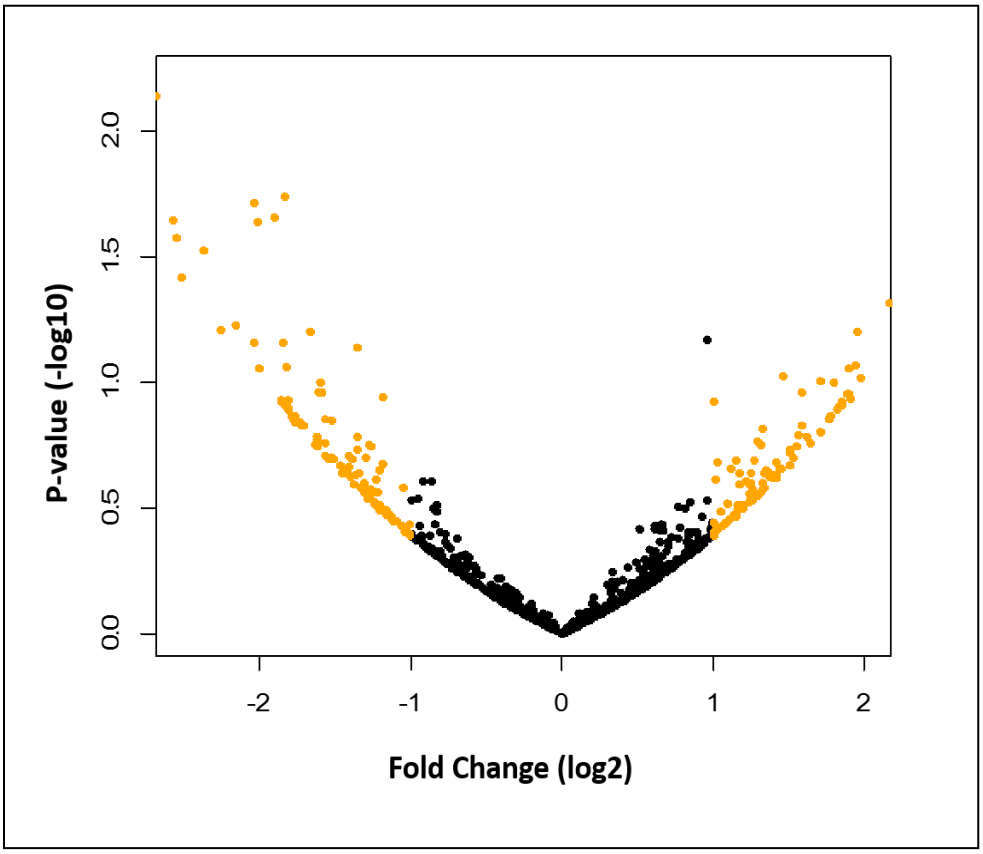

OTU
Abundance
256
16
1
2285
2150
2397
2326
1053
2195
1204
Sample

Figure 7.



Stability of the Lung Microbiome in COPD Patients

Supplementary Data

Figure 8. Shannon Diversity Index for Inter-Subject Variability Samples. X-axis is Samples per Sequence. Y-axis is Shannon Rarefraction Measures. At 1500 sequences/sample, all unique OTUs had been identified. Legend shows what each color represents. Figure was created in Qiime.

Figure 9. Diversity Indexes for Inter-Subject Variability Samples. Boxplots for different alpha diversities, Observed, Chao1, Ace, Shannon, Simpson, InvSimpson (1/Simpson), and Fisher. Boxplots show no significant differences in alpha diversities between T1 and T2. Figure was created in R (package ggplot2).

Figure 10. Principal Coordinate Analysis Demonstrates Clustering of Intra-Subject Variability Samples. Principal coordinate analysis was performed using Qiime and Weighted UniFrac, and the results for principal coordinates 1 and 2 and 3 are shown. For every patient, samples clustered together.

Figure 11. Shannon Diversity Index for Intra-Subject Variability Samples. Samples. X-axis is Samples per Sequence. Y-axis is Shannon Rarefraction Measures. At 1500 sequences/sample, all unique OTUs had been identified. Legend shows what each color represents. Figure was created in Qiime.

Figure 12. A. Taxonomic Identification at the Phylum Level for Intra-Subject Variability Samples. Taxonomic results at the phylum level are displayed for each sample. B. Taxonomic Identification for the top 8 Genus Level for intra-subject variability samples. Taxonomic results at the genus level are displayed for each sample.

Figure 13. Heat Map of OTU Abundances for Intra-Subject Variability Samples. Figure was created in R (package phyloseq). Lighter blue shows greater abundance compared to dark blue/black.

Figure 8.

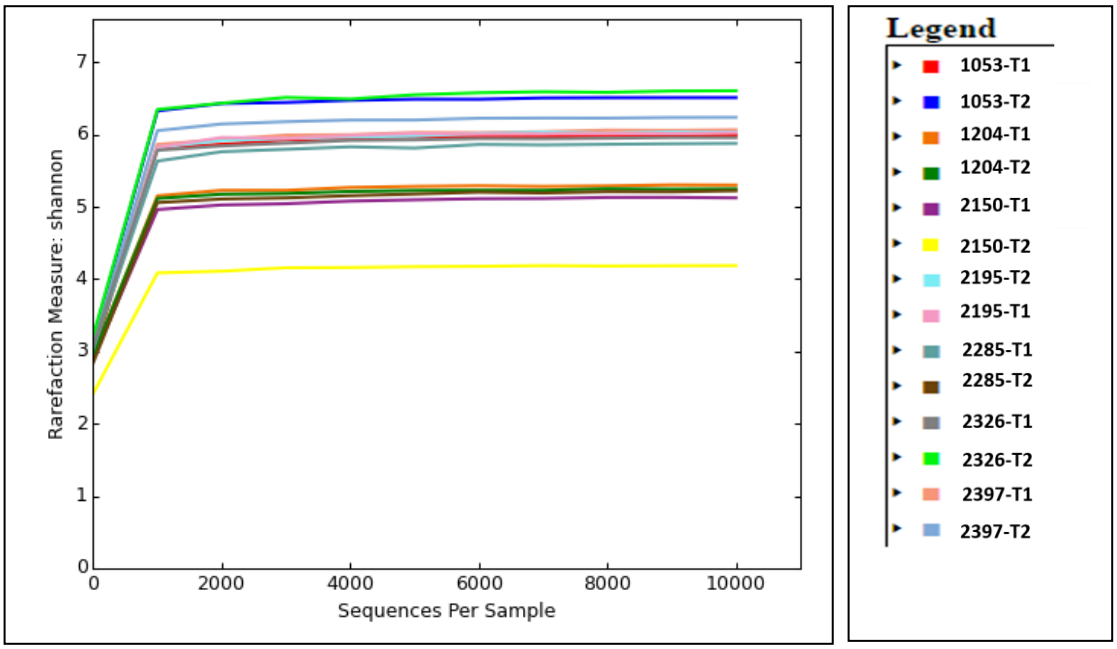

Rarefaction Measure: shannon
Sequences Per Sample
0
1
2
3
4
5
6
7
0
2000
4000
6000
8000
10000
Legend
1053-T1
1053-T2
1204-T1
1204-T2
2150-T1
2150-T2
2195-T2
2195-T1
2285-T1
2285-T2
2326-T1
2326-T2
2397-T1
2397-T2

Figure 9.

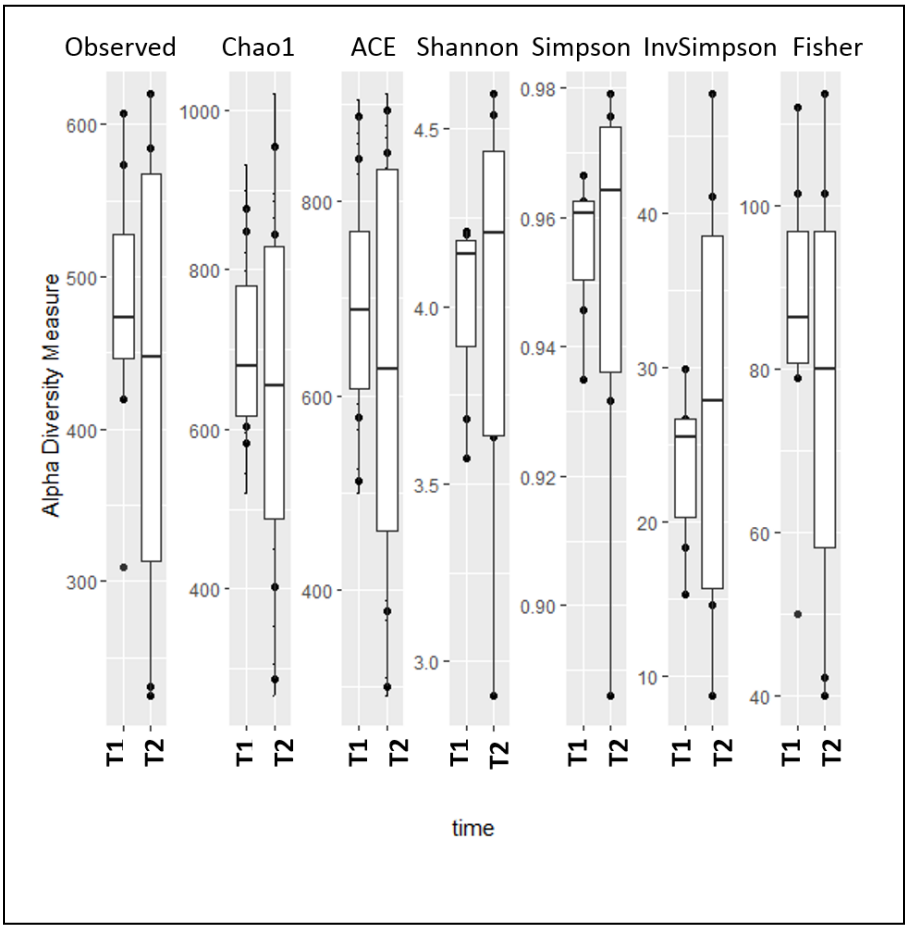

Observed
Chao1
ACE
Shannon
Simpson
InvSimpson
Fisher
Alpha Diversity Measure
600
500
400
300
1000
800
600
400
800
600
400
4.5
4.0
3.5
3.0
0.98
0.96
0.94
0.92
0.90
40
30
20
10
100
80
60
40
T1
T2
time

Figure 10.

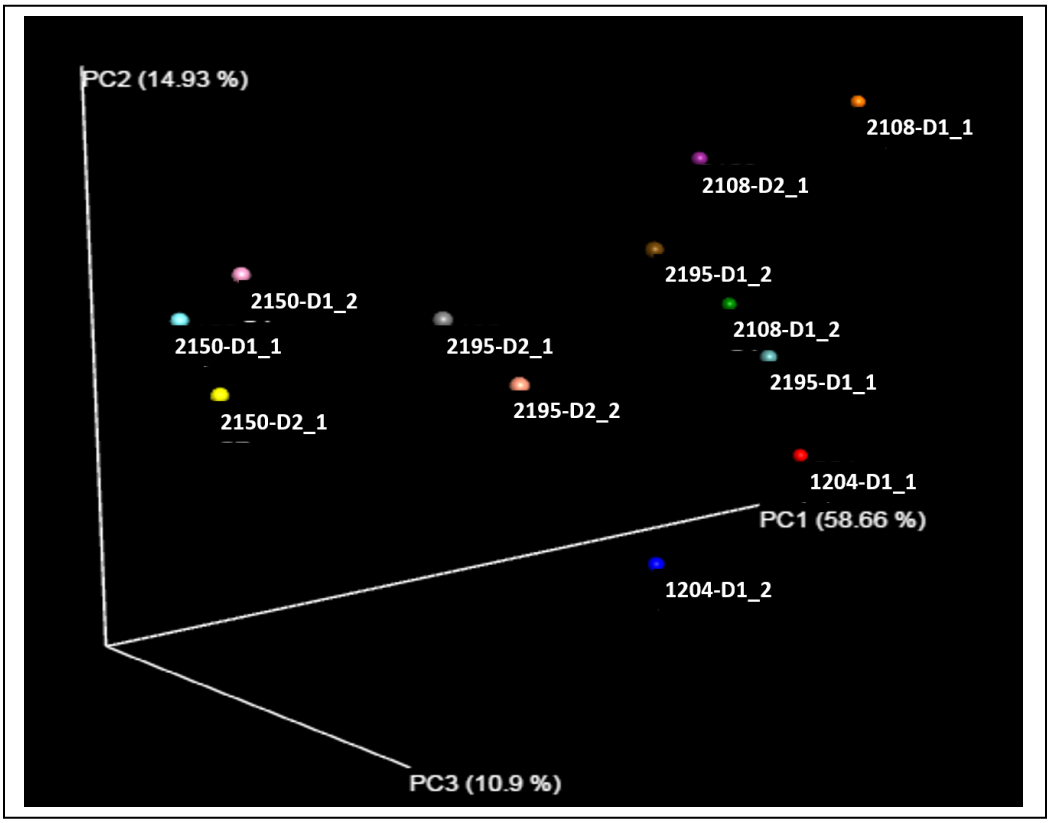

PC2 (14.93 %)
2108-D1_1
2108-D2_1
2195-D1_2
2150-D1_2
2150-D1_1
2195-D2_1
2108-D1_2
2195-D1_1
2150-D2_1
2195-D2_2
1204-D1_1
PC1 (58.66 %)
1204-D1_2
PC3 (10.9 %)

Figure 11.

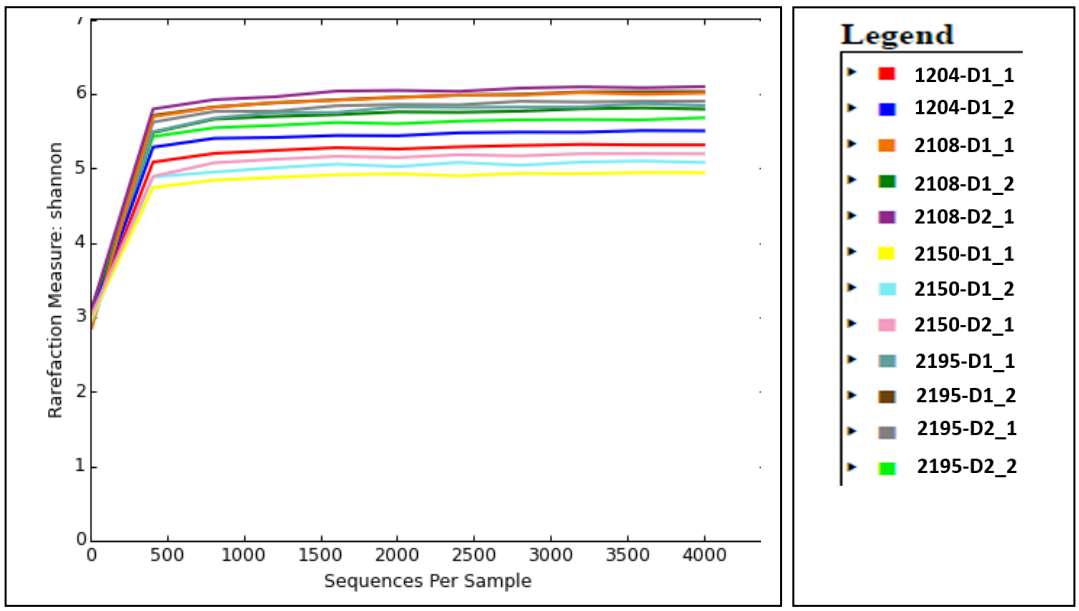

Rarefaction Measure: shannon
Sequences Per Sample
0
1
2
3
4
5
6
0
500
1000
1500
2000
2500
3000
3500
4000
Legend
1204-D1_1
1204-D1_2
2108-D1_1
2108-D1_2
2108-D2_1
2150-D1_1
2150-D1_2
2150-D2_1
2195-D1_1
2195-D1_2
2195-D2_1
2195-D2_2

Figure 12A.

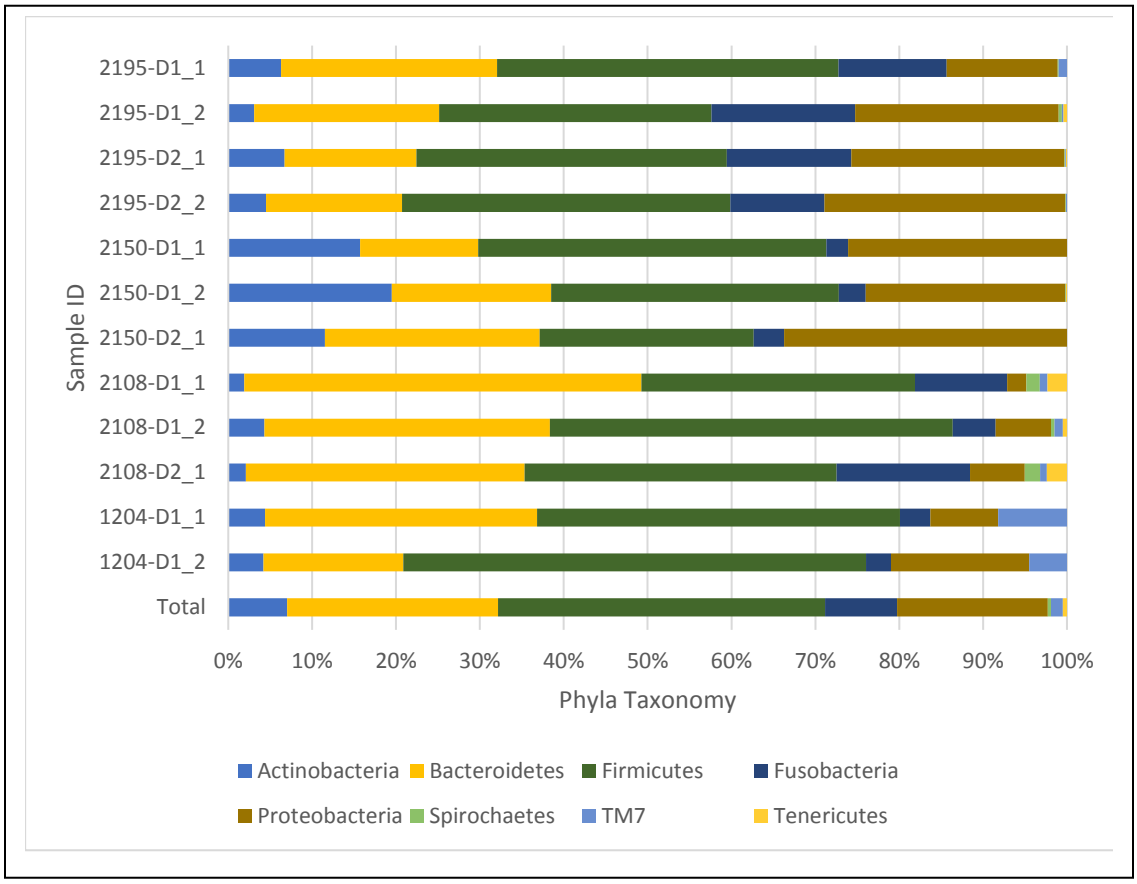

2195-D1_1
2195-D1_2
2195-D2_1
2195-D2_2
2150-D1_1
2150-D1_2
2150-D2_1
2108-D1_1
2108-D1_2
2108-D2_1
1204-D1_1
1204-D1_2
Total
Sample ID
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
Phyla Taxonomy
Actinobacteria
Bacteroidetes
Firmicutes
Fusobacteria
Proteobacteria
Spirochaetes
TM7
Tenericutes

Figure 12B.

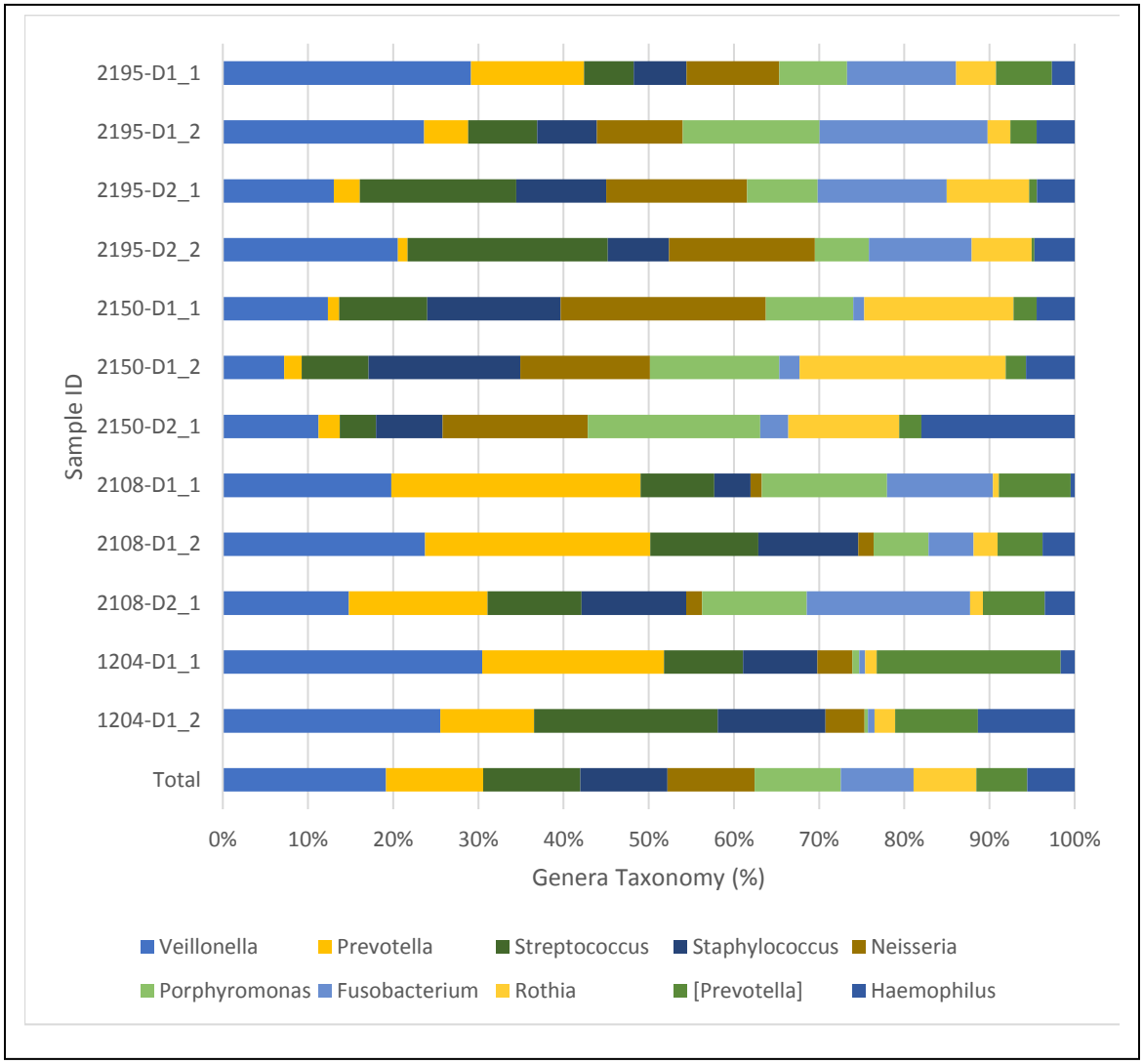
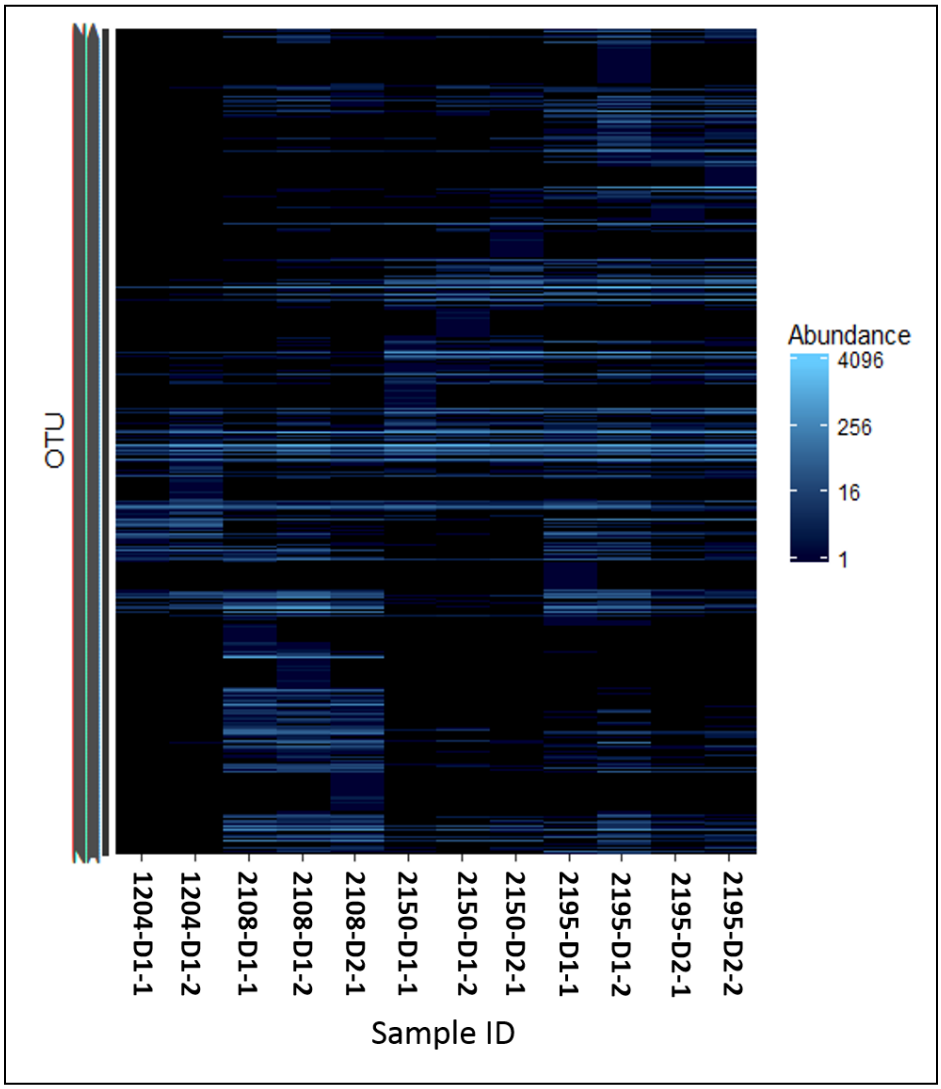

2195-D1_1
2195-D1_2
2195-D2_1
2195-D2_2
2150-D1_1
2150-D1_2
2150-D2_1
2108-D1_1
2108-D1_2
2108-D2_1
1204-D1_1
1204-D1_2
Total
Sample ID
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
Genera Taxonomy (%)
Veillonella
Prevotella
Streptococcus
Staphylococcus
Neisseria
Porphyromonas
Fusobacterium
Rothia
[Prevotella]
Haemophilus

Figure 13.


OTU
Abundance
4096
256
16
1
1204-D1-1
1204-D1-2
2108-D1-1
2108-D1-2
2108-D2-1
2150-D1-1
2150-D1-2
2150-D2-1
2195-D1-1
2195-D1-2
2195-D2-1
2195-D2-2
Sample ID

CHAPTER V. DISCUSSION

Summary of Current Research

Agricultural workers are at risk for respiratory diseases, such as chronic obstructive pulmonary disease (COPD) and chronic bronchitis, due to exposures and inflammatory agents found in their work environment. These exposures include: organic dusts, microorganisms, bacterial endotoxins, chemical toxicants, and bacterial degradation of grain and animal wastes. In addition, farmers can be exposed to many potential respiratory irritants including diesel exhaust, pesticides, hays, and grains.

While previous studies have been conducted to examine agricultural effects of chronic respiratory diseases more closely, there are still significant gaps in the literature regarding the understanding of how agricultural workers are afflicted with chronic respiratory diseases. This dissertation sought to help narrow some of the existing gaps in the literature and study the relationship of agricultural exposures and respiratory diseases.

The AgLung database and the Keokuk database have both provided an opportunity to explore the relationship between agricultural exposures and chronic respiratory diseases. The AgLung database allowed us to utilize multiple questionnaires to identify what questions were necessary to gain an accurate exposure assessment. Further, we recruited COPD patients from this database so that we could examine the stability of lung microbiome longitudinally over a limited period of time. Because blood was taken for genotyping from both the AgLung and Keokuk populations, we were also able to examine the relationship between *TLR4* polymorphisms with chronic respiratory disease outcomes and the expression of pro-inflammatory cytokines.

Chapter 2 of this dissertation used principal factor analysis (PFA) to summarize a large amount of important agricultural exposure variables collected from questionnaires. The two questionnaires from the AgLung cohort were designed to assess the relationship between agricultural exposures and respiratory diseases. Previous studies on chronic diseases have utilized questionnaires that have assessed up to 50 exposures. However, it was uncertain if numerous questions were necessary to obtain accurate exposure information. The results of Chapter 2 give a sufficient model that is interpretable and consistent with farming practice. Four models were examined to characterize agricultural exposures. Model 1 used data from a different questionnaire than the other three models, asking “yes/no” for farm tasks during the participants’ 20s and 40s. The other questionnaires, used in Models 2, 3, and 4, asked about farm tasks as “ever/never” during their lifetime and included lifetime total years and total hours for many of the exposures. We determined the pattern of clustered agricultural exposures differed between Questionnaire 1 (Model 1) and Questionnaire 2 (Model 2). We also showed that the utilization of intensity exposure variables (total years found in Model 3) yielded different factors than dichotomous (yes/no found in Model 2) variables. In addition, we showed greater variation explained using total lifetime hours, found in Model 4, compared to total lifetime years (Model 3).

Utilizing both the AgLung and Keokuk populations, Chapter 3 focused on *TLR4* polymorphisms. Functional polymorphisms in *TLR4* have been studied and their role investigated in a number of inflammatory diseases. Most studies, including ours, have focused on the missense Asp299Gly polymorphism, as the rare allele Gly299 has been shown to cause reduced levels of pro-inflammatory cytokines and inflammatory hypo-responsiveness to inhaled LPS. The association of the Asp299Gly polymorphism and chronic bronchitis had not been studied in individuals who are chronically exposed to

high levels of LPS. The results of Chapter 3 showed that the Asp299Gly is associated with chronic bronchitis in two agriculturally-exposed populations. We also showed that this polymorphism was associated with decreased LPS-stimulated production of IL-6 and TNF-α in a whole blood assay.

The final study of this dissertation (Chapter 4) examined the stability of the lung microbiome. Changes in the lung microbiome in patients with COPD have been found to occur when there are respiratory infections, airway inflammation as a result of an exacerbation, and the use of antibiotics. There were no longitudinal studies comparing baseline sputum samples with samples taken several months later in COPD patients during clinical stability. The results of Chapter 4 showed little variability of the lung microbiome in COPD patients over a period of time (≤ 10 months).

Together, the studies from these chapters offer new insight and address some of the gaps in knowledge regarding the assessment of agricultural exposures based on factor analysis, the association of TLR4 polymorphisms with chronic bronchitis, and the stability of the lung microbiome in COPD patients.

Implications of Current Research

While cigarette smoking is the leading contributor to COPD and chronic bronchitis, a summary by the American Thoracic Society found that 15% of the occurrence of COPD is due to occupational exposures, independent of smoking. Farming is among the top occupations associated with COPD. Many studies have shown that there is an important role of agricultural exposures in the pathophysiology of COPD. In this dissertation, we studied agricultural exposures somewhat independent of smoking.

The first study of this dissertation examined the relationship between a variety of agricultural exposures. This study compared four different PFA models to identify the essential questions to ask in order to characterize long-term agricultural exposures. The questions that were identified as pertinent included duration and intensity (i.e. total lifetime hours) of farm work, farm tasks, livestock exposure, crop exposure and "other exposures".

Chapter 3 of this dissertation found that the Asp299Gly variant is associated with chronic bronchitis in two agricultural-exposed populations. This association was observed in two somewhat different populations. The Keokuk population was a fairly healthy population (without airflow obstruction and smoking exposure). The AgLung population, on the other hand, had a large percentage of individuals with airflow obstruction and who were current smokers. The association was strongest in the Keokuk population possibly suggesting that agricultural exposures, independent of smoking, contributes to the development of chronic bronchitis among those with the *TLR4* Asp299Gly polymorphism. Similar to previous studies, this study also found the association between Gly299 and reduced levels of LPS-stimulated TNF-α.

Chapter 4 presented the bacterial profiles of agriculturally-exposed, male patients with moderate COPD based on 16S rRNA gene sequencing. All patients were former smokers who worked on the farm for a large number of years. This chapter has the most clinical implications as it has given us a better understanding of the stability of the lung microbiome of COPD patients over a short period of time. Our results indicated that the lung microbiome remains relatively stable over a short period of time.

Suggested Future Research

The work in this dissertation adds to the existing literature on the relationship between agricultural exposures and chronic respiratory diseases. However, there are still many gaps in the literature in which future research in this area should focus. Based on the results of this dissertation, several future directions are suggested.

The second chapter of this dissertation focused on the development of a model for understanding the relationships between a variety of agricultural exposures. The results from this study could help us refine future measurement efforts by developing a questionnaire that encompasses the most important exposure variables found here in this study which could also greatly reduce the burden both to the researcher and the participant in that we would measure a smaller number of variables rather than 36 plus variables that were found in our questionnaires. It also may help us hone in on the most biologically relevant exposures. Next steps include conducting field tests and formal tests of validation. Then we could explore who will this questionnaire predict respiratory outcomes in an agriculturally-exposed population, particularly compared to more standard questionnaires.

This model may be helpful in guiding future research, but first should be used in its current population (AgLung) and then, in populations similar to this population (i.e. Keokuk, Midwest, Great Plains). In addition, the model should include genetic variants to see how agricultural exposures and genetic susceptibility factors underlie the pathogenesis of chronic respiratory diseases. I encourage those working on databases similar to ours to use this model to examine their respiratory disease outcomes.

The results from the second study found the opposite to what was hypothesized. Because the *TLR4* polymorphism (Asp299Gly) has been associated with acute airway and inflammatory hypo-responsiveness to inhaled LPS, we hypothesized that agricultural workers chronically exposed to LPS and carrying the Asp299Gly

polymorphism would be protected from respiratory symptoms, i.e. chronic bronchitis. We observed the opposite and were left with one main question: why are *TLR4* 299/399 polymorphisms associated with an increased odds of chronic bronchitis in these agricultural-exposed populations? To further examine these results, other areas need to be studied, in addition to the TLR4 pathway. These areas include polymorphisms affecting mucus production and different pathways including the TLR2 pathway.

Airway mucus hypersecretion is a common pathological feature of chronic obstructive airway diseases including chronic bronchitis. Mucins, the major protein components of mucus, are directly induced by TLR signaling. Other genes (i.e. MUC5AC, MUC2) may be contributing to the increased odds of chronic bronchitis in these two agriculturally-exposed populations.

In addition, not all bacterial products signal through the TLR4 pathway. LPS also signals through TLR2. Unlike TLR4, TLR2 polymorphisms responded the same as the wild-type to LPS. Therefore, the production of pro-inflammatory cytokines that contribute to chronic bronchitis might be due to LPS being signaled through the TLR2 pathway. Furthermore, other ligands in organic dust, such as gram-positive bacteria, could be triggering signaling through the TLR2 pathway. Although endotoxin is found in organic dust, many studies have failed to link endotoxin exposure to the development of disease (155-160). Studies have found that gram-positive bacteria are more prevalent in organic dust than gram-negative bacteria (159-161). Peptidoglycan (PGN), a non-endotoxin component, has been found to initiate the innate immune inflammatory response to swine facility animal farming dusts in vitro (159,162,163). TLR2 is an important receptor of gram-positive bacteria as it recognizes PGNs that are part of the cell wall of gram-positive bacteria (138,164). Therefore, in these two highly exposed populations, the presence of gram-positive bacteria in organic dust may be contributing to the increased

risk of chronic bronchitis through the TLR2 pathway (155). Future studies need to assay polymorphisms in the TLR2, MUC2, and/or MUC5AC gene in individuals with and without exposure to agriculture.

Additionally, in future analysis, the PFA model from Chapter 2 should be applied to the findings from Chapter 3. This model should model chronic bronchitis as the outcome and include the four factors (from Chapter 2) and the Asp299Gly variant. In addition, this model can include demographic variables such as smoking variables, age, and education.

The results of the fourth chapter of this dissertation showed that the lung microbiome of agriculturally-exposed COPD patients remains relatively stable over a two day time period and shows variations among some COPD patients. Future longitudinal studies should include a larger sample size. Additionally, different time periods between the two samples should be examined. Bronchoalveolar lavage (BAL), cell-free BAL supernatants, and biopsies, should be studied as well since these samples provide different insights into the processes of the lung microbiology. For further comparison, longitudinal studies should be conducted in agricultural-exposed patients (current farmers) without COPD to determine how these exposures affect the lung microbiome over time.

Limitations

There are limitations in each of the studies in this dissertation. In the first study, recall bias is probable as participants were asked to retrospectively recall their lifetime farming exposures. In addition, there is a potential for interviewer bias as there were two methods to obtain exposure information (telephone interviews vs in-person interviews). Furthermore, the mean age of this population was 64 years, therefore, their agricultural

exposures may be different from younger workers because of technological advances in farming. Direct measurement of agricultural exposures was not performed.

In the second study, a “healthy worker effect” might have occurred as a large number of people leave the industry after developing respiratory issues. They are, therefore, unable to remain employed in this industry. This factor may have caused an underestimation of the results as individuals with respiratory problems might not work in environments with high exposures. Whole blood assays were used to access responsiveness to LPS and stimulated cytokine levels were measured in a whole blood assays; however, that may not reflect the airway.

In the final study, the sample size was quite small making it difficult to observe more significant differences between the groups (T1 and T2). In addition, sputum samples have been shown to have significantly lower diversity than the other sample types such as total BAL, cell-free BAL supernatants, and biopsies. In all studies, there is the issue of generalizability of these results. Both populations included mainly Caucasian farmers from Iowa and Nebraska.

Conclusions

This dissertation contributed to the many gaps in knowledge in the relationship between agricultural exposures and respiratory diseases. This dissertation utilized PFA to determine interpretable and logical factors to examine the relationship between agricultural exposures and respiratory diseases. It showed *TLR4* 299/399 polymorphisms are associated with chronic bronchitis in two agriculturally-exposed populations, and suggested that the lung microbiome of agriculturally-exposed individuals is relatively stable. Continuing to fill the gaps in knowledge in this area of study is crucial to improving the respiratory health in agricultural workers (141).

Bibliography

(1) Respiratory health hazards in agriculture. Am J Respir Crit Care Med 1998 Nov;158(5 Pt 2):S1-S76.

(2) USDA Census of Agriculture. Farm Demographics-U.S. Farmers. 2012.

(3) Eduard W, Douwes J, Omenaas E, Heederik D. Do farming exposures cause or prevent asthma? Results from a study of adult Norwegian farmers. Thorax 2004 May;59(5):381-386.

(4) Eduard W, Pearce N, Douwes J. Chronic bronchitis, COPD, and lung function in farmers: the role of biological agents. Chest 2009 Sep;136(3):716-725.

(5) Girard M, Cormier Y. Hypersensitivity pneumonitis. Curr Opin Allergy Clin Immunol 2010 Apr;10(2):99-103.

(6) Kogevinas M, Anto JM, Sunyer J, Tobias A, Kromhout H, Burney P. Occupational asthma in Europe and other industrialised areas: a population-based study. European Community Respiratory Health Survey Study Group. Lancet 1999 May 22;353(9166):1750-1754.

(7) May S, Romberger DJ, Poole JA. Respiratory health effects of large animal farming environments. J Toxicol Environ Health B Crit Rev 2012;15(8):524-541.

(8) Monso E, Riu E, Radon K, Magarolas R, Danuser B, Iversen M, et al. Chronic obstructive pulmonary disease in never-smoking animal farmers working inside confinement buildings. Am J Ind Med 2004 Oct;46(4):357-362.

(9) Omland O. Exposure and respiratory health in farming in temperate zones--a review of the literature. Ann Agric Environ Med 2002;9(2):119-136.

(10) Kirkhorn SR, Garry VF. Agricultural lung diseases. Environ Health Perspect 2000 Aug;108 Suppl 4:705-712.

(11) Global Initiative for Chronic Obstructive Lung Diseas. Global Strategy for the Diagnosis, Management and Prevention of COPD, Global Initiative for Chronic Obstructive Lung Disease (GOLD). 2016; . Accessed October 12, 2016.

(12) World Health Organization. Chronic Obstructive Pulmonary Disease (COPD). 2014; Available at: http://www.who.int/respiratory/copd/en/. Accessed September, 5, 2013.

(13) National Center for Health Statistics (US). 2016 May.

(14) Wheaton AG, Cunningham TJ, Ford ES, Croft JB, Centers for Disease Control and Prevention (CDC). Employment and activity limitations among adults with chronic

obstructive pulmonary disease--United States, 2013. MMWR Morb Mortal Wkly Rep 2015 Mar 27;64(11):289-295.

(15) Balmes J, Becklake M, Blanc P, Henneberger P, Kreiss K, Mapp C, et al. American Thoracic Society Statement: Occupational contribution to the burden of airway disease. Am J Respir Crit Care Med 2003 Mar 1;167(5):787-797.

(16) Lebowitz MD. Occupational exposures in relation to symptomatology and lung function in a community population. Environ Res 1977 Aug;14(1):59-67.

(17) Becklake MR. Chronic airflow limitation: its relationship to work in dusty occupations. Chest 1985 Oct;88(4):608-617.

(18) Becklake MR. Occupational exposures: evidence for a causal association with chronic obstructive pulmonary disease. Am Rev Respir Dis 1989 Sep;140(3 Pt 2):S85-91.

(19) Lamprecht B, Schirnhofer L, Kaiser B, Studnicka M, Buist AS. Farming and the prevalence of non-reversible airways obstruction: results from a population-based study. Am J Ind Med 2007 Jun;50(6):421-426.

(20) Stockley RA, Rennard SI, Rabe K, Celli B. Struction-function Relationships. The Pathophysiology of Airflow Obstruction. In: Niewohner DE, editor. Chronic Obstructive Pulmonary Disease; 2008. p. Chapter 1.

(21) Pleis JR, Lethbridge-Cejku M. Summary health statistics for U.S. adults: National Health Interview Survey, 2006. Vital Health Stat 10 2007 Dec;(235)(235):1-153.

(22) Mullen JB, Wright JL, Wiggs BR, Pare PD, Hogg JC. Structure of central airways in current smokers and ex-smokers with and without mucus hypersecretion: relationship to lung function. Thorax 1987 Nov;42(11):843-848.

(23) Von Essen S, Donham K. Illness and injury in animal confinement workers. Occup Med 1999 Apr-Jun;14(2):337-350.

(24) Melbostad E, Eduard W, Magnus P. Chronic bronchitis in farmers. Scand J Work Environ Health 1997 Aug;23(4):271-280.

(25) Thorne P, Konham K, Dosman Jea. Report of the Occupational Health Group: Proceedings from "Understanding the Impacts of Large Scale Swine Production: an Interdisciplinary Scientific Workshop". 1995.

(26) Rennard SI. COPD: overview of definitions, epidemiology, and factors influencing its development. Chest 1998 Apr;113(4 Suppl):235S-241S.

(27) Global Initiative for Asthma (GINA). Global Strategy for Asthma Management and Prevention. 2015.

(28) ACGIH. Cincinnati, OH:American Council of Government Industrial Hygienists. 2000;2016.

(29) Simpson JC, Niven RM, Pickering CA, Oldham LA, Fletcher AM, Francis HC. Comparative personal exposures to organic dusts and endotoxin. Ann Occup Hyg 1999 Feb;43(2):107-115.

(30) Zhiping W, Malmberg P, Larsson BM, Larsson K, Larsson L, Saraf A. Exposure to bacteria in swine-house dust and acute inflammatory reactions in humans. Am J Respir Crit Care Med 1996 Nov;154(5):1261-1266.

(31) Boissy RJ, Romberger DJ, Roughead WA, Weissenburger-Moser L, Poole JA, LeVan TD. Shotgun pyrosequencing metagenomic analyses of dusts from swine confinement and grain facilities. PLoS One 2014 Apr 18;9(4):e95578.

(32) Donham KJ, Popendorf W, Palmgren U, Larsson L. Characterization of dusts collected from swine confinement buildings. Am J Ind Med 1986;10(3):294-297.

(33) Schwartz DA, Donham KJ, Olenchock SA, Popendorf WJ, Van Fossen DS, Burmeister LF, et al. Determinants of longitudinal changes in spirometric function among swine confinement operators and farmers. Am J Respir Crit Care Med 1995 Jan;151(1):47-53.

(34) Schwartz DA, Thorne PS, Yagla SJ, Burmeister LF, Olenchock SA, Watt JL, et al. The role of endotoxin in grain dust-induced lung disease. Am J Respir Crit Care Med 1995 Aug;152(2):603-608.

(35) Basinas I, Sigsgaard T, Kromhout H, Heederik D, Wouters IM, Schlunssen V. A comprehensive review of levels and determinants of personal exposure to dust and endotoxin in livestock farming. J Expo Sci Environ Epidemiol 2015 Mar-Apr;25(2):123-137.

(36) Pope CA,3rd, Burnett RT, Thurston GD, Thun MJ, Calle EE, Krewski D, et al. Cardiovascular mortality and long-term exposure to particulate air pollution: epidemiological evidence of general pathophysiological pathways of disease. Circulation 2004 Jan 6;109(1):71-77.

(37) Thomas KW, Dosemeci M, Coble JB, Hoppin JA, Sheldon LS, Chapa G, et al. Assessment of a pesticide exposure intensity algorithm in the agricultural health study. J Expo Sci Environ Epidemiol 2010 Sep;20(6):559-569.

(38) Tucker JD, Moore DH,2nd, Ramsey MJ, Kato P, Langlois RG, Burroughs B, et al. Multi-endpoint biological monitoring of phosphine workers. Mutat Res 2003 Apr 20;536(1-2):7-14.

(39) Kurmi OP, Semple S, Simkhada P, Smith WC, Ayres JG. COPD and chronic bronchitis risk of indoor air pollution from solid fuel: a systematic review and meta-analysis. Thorax 2010 Mar;65(3):221-228.

(40) Tsukuya G, Matsumoto K, Fukuyama S, Crawford B, Nakanishi Y, Ichinose M, et al. Validation of a COPD screening questionnaire and establishment of diagnostic cut-points in a Japanese general population: The Hisayama study. Allergol Int 2015 Jan;64(1):49-53.

(41) Darby AC, Waterhouse JC, Stevens V, Billings CG, Billings CG, Burton CM, et al. Chronic obstructive pulmonary disease among residents of an historically industrialised area. Thorax 2012 Oct;67(10):901-907.

(42) Golka K, Abreu-Villaca Y, Anbari Attar R, Angeli-Greaves M, Aslam M, Basaran N, et al. Bladder cancer documentation of causes: multilingual questionnaire, 'bladder cancer doc'. Front Biosci (Elite Ed) 2012 Jun 1;4:2809-2822.

(43) Golec M, Reichel C, Mackiewicz B, Skorska C, Curzytek K, Lemieszek M, et al. Cathelicidin LL-37, granzymes, TGF-beta1 and cytokines levels in induced sputum from farmers with and without COPD. Ann Agric Environ Med 2009;16(2):289-297.

(44) Gora A, Mackiewicz B, Krawczyk P, Golec M, Skorska C, Sitkowska J, et al. Occupational exposure to organic dust, microorganisms, endotoxin and peptidoglycan among plants processing workers in Poland. Ann Agric Environ Med 2009;16(1):143-150.

(45) Tsapko VG, Chudnovets AJ, Sterenbogen MJ, Papach VV, Dutkiewicz J, Skorska C, et al. Exposure to bioaerosols in the selected agricultural facilities of the Ukraine and Poland - a review. Ann Agric Environ Med 2011;18(1):19-27.

(46) Pipinic IS, Varnai VM, Lucic RB, Cavlovic A, Prester L, Orct T, et al. Endotoxin exposure assessment in wood-processing industry: airborne versus settled dust levels. Arh Hig Rada Toksikol 2010 Jun;61(2):161-166.

(47) Park DU, Ryu SH, Kim SB, Yoon CS. An assessment of dust, endotoxin, and microorganism exposure during waste collection and sorting. J Air Waste Manag Assoc 2011 Apr;61(4):461-468.

(48) Heldal KK, Madso L, Huser PO, Eduard W. Exposure, symptoms and airway inflammation among sewage workers. Ann Agric Environ Med 2010;17(2):263-268.

(49) Szadkowska-Stanczyk I, Brodka K, Buczynska A, Cyprowski M, Kozajda A, Sowiak M. Exposure to bioaerosols among CAFO workers (swine feeding). Med Pr 2010;61(3):257-269.

(50) Demanche A, Bonlokke J, Beaulieu MJ, Assayag E, Cormier Y. Swine confinement buildings: effects of airborne particles and settled dust on airway smooth muscles. Ann Agric Environ Med 2009;16(2):233-238.

(51) Madsen AM, Hansen VM, Nielsen SH, Olsen TT. Exposure to dust and endotoxin of employees in cucumber and tomato nurseries. Ann Occup Hyg 2009 Mar;53(2):129-138.

(52) Madsen AM. Exposure to airborne microbial components in autumn and spring during work at Danish biofuel plants. Ann Occup Hyg 2006 Nov;50(8):821-831.

(53) Lawniczek-Walczyk A, Gorny RL. Endotoxins and beta-glucans as markers of microbiological contamination--characteristics, detection, and environmental exposure. Ann Agric Environ Med 2010;17(2):193-208.

(54) Cohen BS. Industrial hygiene measurement and control. In: Lippincott-Raven, editor. Environmental and Occupational Medicine. 3rd ed. Philadelphia; 1998. p. 1741-1755.

(55) Shahan TA, Sorenson WG, Lewis DM. Superoxide anion production in response to bacterial lipopolysaccharide and fungal spores implicated in organic dust toxic syndrome. Environ Res 1994 Oct;67(1):98-107.

(56) Rylander R. Organic dust induced pulmonary disease - the role of mould derived beta-glucan. Ann Agric Environ Med 2010;17(1):9-13.

(57) Kurt-Jones EA, Mandell L, Whitney C, Padgett A, Gosselin K, Newburger PE, et al. Role of toll-like receptor 2 (TLR2) in neutrophil activation: GM-CSF enhances TLR2 expression and TLR2-mediated interleukin 8 responses in neutrophils. Blood 2002 Sep 1;100(5):1860-1868.

(58) Akira S. Mammalian Toll-like receptors. Curr Opin Immunol 2003 Feb;15(1):5-11.

(59) Chow JC, Young DW, Golenbock DT, Christ WJ, Gusovsky F. Toll-like receptor-4 mediates lipopolysaccharide-induced signal transduction. J Biol Chem 1999 Apr 16;274(16):10689-10692.

(60) Shimazu R, Akashi S, Ogata H, Nagai Y, Fukudome K, Miyake K, et al. MD-2, a molecule that confers lipopolysaccharide responsiveness on Toll-like receptor 4. J Exp Med 1999 Jun 7;189(11):1777-1782.

(61) O'Neill LA, Bryant CE, Doyle SL. Therapeutic targeting of Toll-like receptors for infectious and inflammatory diseases and cancer. Pharmacol Rev 2009 Jun;61(2):177-197.

(62) De Schepper S, De Ketelaere A, Bannerman DD, Paape MJ, Peelman L, Burvenich C. The toll-like receptor-4 (TLR-4) pathway and its possible role in the pathogenesis of Escherichia coli mastitis in dairy cattle. Vet Res 2008 Jan-Feb;39(1):5.

(63) Arbour NC, Lorenz E, Schutte BC, Zabner J, Kline JN, Jones M, et al. TLR4 mutations are associated with endotoxin hyporesponsiveness in humans. Nat Genet 2000 Jun;25(2):187-191.

(64) de Moraes MR, da Costa AC, Correa Kde S, Junqueira-Kipnis AP, Rabahi MF. Interleukin-6 and interleukin-8 blood levels' poor association with the severity and clinical profile of ex-smokers with COPD. Int J Chron Obstruct Pulmon Dis 2014 Jul 29;9:735-743.

(65) Bitko V, Garmon NE, Cao T, Estrada B, Oakes JE, Lausch RN, et al. Activation of cytokines and NF-kappa B in corneal epithelial cells infected by respiratory syncytial virus: potential relevance in ocular inflammation and respiratory infection. BMC Microbiol 2004 Jul 15;4:28.

(66) Noah TL, Henderson FW, Wortman IA, Devlin RB, Handy J, Koren HS, et al. Nasal cytokine production in viral acute upper respiratory infection of childhood. J Infect Dis 1995 Mar;171(3):584-592.

(67) Sabroe I, Whyte MK, Wilson AG, Dower SK, Hubbard R, Hall I. Toll-like receptor (TLR) 4 polymorphisms and COPD. Thorax 2004 Jan;59(1):81.

(68) Kiechl S, Wiedermann CJ, Willeit J. Toll-like receptor 4 and atherogenesis. Ann Med 2003;35(3):164-171.

(69) Michel O, LeVan TD, Stern D, Dentener M, Thorn J, Gnat D, et al. Systemic responsiveness to lipopolysaccharide and polymorphisms in the toll-like receptor 4 gene in human beings. J Allergy Clin Immunol 2003 Nov;112(5):923-929.

(70) Rohde G, Klein W, Arinir U, Hagedorn M, Duerig N, T Bauer T, et al. Association of the ASP299GLY TLR4 polymorphism with COPD. Respir Med 2006 May;100(5):892-896.

(71) Pahwa P, Karunanayake CP, Rennie DC, Chen Y, Schwartz DA, Dosman JA. Association of the TLR4 Asp299Gly polymorphism with lung function in relation to body mass index. BMC Pulm Med 2009 Sep 21;9:46-2466-9-46.

(72) Budulac SE, Boezen HM, Hiemstra PS, Lapperre TS, Vonk JM, Timens W, et al. Toll-like receptor (TLR2 and TLR4) polymorphisms and chronic obstructive pulmonary disease. PLoS One 2012;7(8):e43124.

(73) Erb-Downward JR, Thompson DL, Han MK, Freeman CM, McCloskey L, Schmidt LA, et al. Analysis of the lung microbiome in the "healthy" smoker and in COPD. PLoS One 2011 Feb 22;6(2):e16384.

(74) Charlson ES, Bittinger K, Haas AR, Fitzgerald AS, Frank I, Yadav A, et al. Topographical continuity of bacterial populations in the healthy human respiratory tract. Am J Respir Crit Care Med 2011 Oct 15;184(8):957-963.

(75) Hilty M, Burke C, Pedro H, Cardenas P, Bush A, Bossley C, et al. Disordered microbial communities in asthmatic airways. PLoS One 2010 Jan 5;5(1):e8578.

(76) Sze MA, Dimitriu PA, Hayashi S, Elliott WM, McDonough JE, Gosselink JV, et al. The lung tissue microbiome in chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2012 May 15;185(10):1073-1080.

(77) Morris A, Beck JM, Schloss PD, Campbell TB, Crothers K, Curtis JL, et al. Comparison of the respiratory microbiome in healthy nonsmokers and smokers. Am J Respir Crit Care Med 2013 May 15;187(10):1067-1075.

(78) Sethi S, Sethi R, Eschberger K, Lobbins P, Cai X, Grant BJ, et al. Airway bacterial concentrations and exacerbations of chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2007 Aug 15;176(4):356-361.

(79) Millares L, Ferrari R, Gallego M, Garcia-Nunez M, Perez-Brocal V, Espasa M, et al. Bronchial microbiome of severe COPD patients colonised by Pseudomonas aeruginosa. Eur J Clin Microbiol Infect Dis 2014 Jul;33(7):1101-1111.

(80) Hewitt R, Farne H, Ritchie A, Luke E, Johnston SL, Mallia P. The role of viral infections in exacerbations of chronic obstructive pulmonary disease and asthma. Ther Adv Respir Dis 2016 Apr;10(2):158-174.

(81) Molyneaux PL, Mallia P, Cox MJ, Footitt J, Willis-Owen SA, Homola D, et al. Outgrowth of the bacterial airway microbiome after rhinovirus exacerbation of chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2013 Nov 15;188(10):1224-1231.

(82) Miravitlles M, Anzueto A. Role of infection in exacerbations of chronic obstructive pulmonary disease. Curr Opin Pulm Med 2015 May;21(3):278-283.

(83) Wang Z, Bafadhel M, Haldar K, Spivak A, Mayhew D, Miller BE, et al. Lung microbiome dynamics in COPD exacerbations. Eur Respir J 2016 Apr;47(4):1082-1092.

(84) Molyneaux PL, Mallia P, Cox MJ, Footitt J, Willis-Owen SA, Homola D, et al. Outgrowth of the bacterial airway microbiome after rhinovirus exacerbation of chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2013 Nov 15;188(10):1224-1231.

(85) Poole JA, Romberger DJ. Immunological and inflammatory responses to organic dust in agriculture. Curr Opin Allergy Clin Immunol 2012 Apr;12(2):126-132.

(86) Basinas I, Sigsgaard T, Erlandsen M, Andersen NT, Takai H, Heederik D, et al. Exposure-affecting factors of dairy farmers' exposure to inhalable dust and endotoxin. Ann Occup Hyg 2014 Jul;58(6):707-723.

(87) Moran RE, Bennett DH, Garcia J, Schenker MB. Occupational exposure to particulate matter from three agricultural crops in California. Int J Hyg Environ Health 2014 Mar;217(2-3):226-230.

(88) Ljubicic A, Varnai VM, Vucemilo M, Matkovic K, Milic D, Macan J. Exhaled breath condensate pH and FeNO as biomarkers of acute and chronic exposure to hazards at swine farms. J Occup Environ Med 2014 Sep;56(9):946-952.

(89) Sutoluk Z, Kekec Z, Daglioglu N, Hant I. Association of chronic pesticide exposure with serum cholinesterase levels and pulmonary functions. Arch Environ Occup Health 2011;66(2):95-99.

(90) Wurtz ET, Schlunssen V, Malling TH, Hansen JG, Omland O. Occupational Chronic Obstructive Pulmonary Disease in a Danish Population-Based Study. COPD 2014 Nov 21.

(91) Valcin M, Henneberger PK, Kullman GJ, Umbach DM, London SJ, Alavanja MC, et al. Chronic bronchitis among nonsmoking farm women in the agricultural health study. J Occup Environ Med 2007 May;49(5):574-583.

(92) Gorsuch RL. Factor Analysis. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates.; 1983.

(93) Global Strategy for Diagnosis, Management and Prevention of COPD. 2014; . Accessed November 20, 2015.

(94) Olsson U. Maximum likelihood estimation of the polychoric correlation coefficient. Psychometrika 1979;44(4):443-460.

(95) Baglin J. Improving Your Exploratory Factor Analysis for Ordinal Data: A Demonstration Using FACTOR. Practical Assessment, Research and Evaluation 2014;19(5).

(96) Yong AG, Pearce S. A Beginner's Guide to Factor Analysis: Focusing on Exploratory Factor Analysis. Tutorials in Quantitative Methods for Psychology 2013;9(2):79-94.

(97) FJ Floyd KW. Factor analysis in the development and refinementof clinical assessment instruments. Psychological assessment 1995;7(3):286-299.

(98) Samanic C, Hoppin JA, Lubin JH, Blair A, Alavanja MC. Factor analysis of pesticide use patterns among pesticide applicators in the Agricultural Health Study. J Expo Anal Environ Epidemiol 2005 May;15(3):225-233.

(99) Kaiser HF. The application of electronic computers to factor analysis. Educational and Psychological Measurement 1960;20:141-151.

(100) Colgan P. The scientific use of factor analysis in behavioral and life sciences Raymond B. Cattell. Plenum Press, New York and London, 1978. xxii + 618 pp., US $ 32.50, ISBN 0-306-30939-4. Behav Processes 1981 Dec;6(4):385-386.

(101) Respiratory health hazards in agriculture. Am J Respir Crit Care Med 1998 Nov;158(5 Pt 2):S1-S76.

(102) Hoppin JA. Pesticides and respiratory health: where do we go from here? Occup Environ Med 2014 Feb;71(2):80-2013-101876. Epub 2013 Nov 12.

(103) Wells AD, Poole JA, Romberger DJ. Influence of farming exposure on the development of asthma and asthma-like symptoms. Int Immunopharmacol 2014 Nov;23(1):356-363.

(104) Marescaux A, Degano B, Soumagne T, Thaon I, Laplante JJ, Dalphin JC. Impact of farm modernity on the prevalence of chronic obstructive pulmonary disease in dairy farmers. Occup Environ Med 2015 Dec 16.

(105) LeVan TD, Koh WP, Lee HP, Koh D, Yu MC, London SJ. Vapor, dust, and smoke exposure in relation to adult-onset asthma and chronic respiratory symptoms: the Singapore Chinese Health Study. Am J Epidemiol 2006 Jun 15;163(12):1118-1128.

(106) Voaklander D, Day L, Dosman J, Hagel L, Pickett W. Older farmers and machinery exposure-cause for concern? Am J Ind Med 2012 Nov;55(11):1044-1050.

(107) Alavanja MC, Samanic C, Dosemeci M, Lubin J, Tarone R, Lynch CF, et al. Use of agricultural pesticides and prostate cancer risk in the Agricultural Health Study cohort. Am J Epidemiol 2003 May 1;157(9):800-814.

(108) Postma DS, Anzueto AR, Jenkins C, Make BJ, Similowski T, Ostlund O, et al. Factor analysis in predominantly severe COPD: identification of disease heterogeneity by easily measurable characteristics. Respir Med 2013 Dec;107(12):1939-1947.

(109) Hammond A, Tennant A, Tyson SF, Nordenskiold U, Hawkins R, Prior Y. The reliability and validity of the English version of the Evaluation of Daily Activity Questionnaire for people with rheumatoid arthritis. Rheumatology (Oxford) 2015 Sep;54(9):1605-1615.

(110) Coble J, Hoppin JA, Engel L, Elci OC, Dosemeci M, Lynch CF, et al. Prevalence of exposure to solvents, metals, grain dust, and other hazards among farmers in the Agricultural Health Study. J Expo Anal Environ Epidemiol 2002 Nov;12(6):418-426.

(111) Pace E, Ferraro M, Giarratano A, Cipollina C, Gjomarkaj M. TLR4 up-regulation and reduced Foxp3 expression in mechanically ventilated smokers with obstructive chronic bronchitis. COPD 2013 Apr;10(2):147-155.

(112) REID L. Pathology of chronic bronchitis. Proc R Soc Med 1956 Oct;49(10):771-773.

(113) Vestbo J, Prescott E, Lange P. Association of chronic mucus hypersecretion with FEV1 decline and chronic obstructive pulmonary disease morbidity. Copenhagen City Heart Study Group. Am J Respir Crit Care Med 1996 May;153(5):1530-1535.

(114) Sahlander K, Larsson K, Palmberg L. Altered innate immune response in farmers and smokers. Innate Immun 2010 Feb;16(1):27-38.

(115) Poole JA, Romberger DJ. Immunological and inflammatory responses to organic dust in agriculture. Curr Opin Allergy Clin Immunol 2012 Apr;12(2):126-132.

(116) Reynolds SJ, Donham KJ, Whitten P, Merchant JA, Burmeister LF, Popendorf WJ. Longitudinal evaluation of dose-response relationships for environmental exposures and pulmonary function in swine production workers. Am J Ind Med 1996 Jan;29(1):33-40.

(117) Janeway CA,Jr, Medzhitov R. Innate immune recognition. Annu Rev Immunol 2002;20:197-216.

(118) van der Graaf CA, Netea MG, Drenth IP, te Morsche RH, van der Meer JW, Kullberg BJ. Candida-specific interferon-gamma deficiency and toll-like receptor polymorphisms in patients with chronic mucocutaneous candidiasis. Neth J Med 2003 Nov;61(11):365-369.

(119) Chantratita N, Tandhavanant S, Seal S, Wikraiphat C, Wongsuvan G, Ariyaprasert P., et al. TLR4 genetic variation is associated with inflammatory responses in Gram-positive sepsis. Clin Microbiol Infect 2016 Sep 8.

(120) Franchimont D, Vermeire S, El Housni H, Pierik M, Van Steen K, Gustot T, et al. Deficient host-bacteria interactions in inflammatory bowel disease? The toll-like receptor (TLR)-4 Asp299gly polymorphism is associated with Crohn's disease and ulcerative colitis. Gut 2004 Jul;53(7):987-992.

(121) Merchant JA, Stromquist AM, Kelly KM, Zwerling C, Reynolds SJ, Burmeister LF. Chronic disease and injury in an agricultural county: The Keokuk County Rural Health Cohort Study. J Rural Health 2002 Fall;18(4):521-535.

(122) Weissenburger-Moser L, Meza J, Yu F, Shiyanbola O, Romberger DJ, LeVan TD. A principal factor analysis to characterize agricultural exposures among Nebraska veterans. J Expo Sci Environ Epidemiol 2016 Apr 6.

(123) American Thoracic Society. Patient Information Series. 2013.

(124) Alavi-Naini R, Salimi S, Sharifi-Mood B, Davoodikia AA, Moody B, Naghavi A. Association between the CD14 gene C-159T polymorphism and serum soluble CD14 with pulmonary tuberculosis. Int J Tuberc Lung Dis 2012 Oct;16(10):1383-1387.

(125) van der Linden MW, Huizinga TW, Stoeken DJ, Sturk A, Westendorp RG. Determination of tumour necrosis factor-alpha and interleukin-10 production in a whole blood stimulation system: assessment of laboratory error and individual variation. J Immunol Methods 1998 Sep 1;218(1-2):63-71.

(126) Romberger DJ, Bodlak V, Von Essen SG, Mathisen T, Wyatt TA. Hog barn dust extract stimulates IL-8 and IL-6 release in human bronchial epithelial cells via PKC activation. J Appl Physiol (1985) 2002 Jul;93(1):289-296.

(127) Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, Nickerson DA. Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. Am J Hum Genet 2004 Jan;74(1):106-120.

(128) Affairs USDoV. *National survey of veterans*. 2010.

(129) Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 2001 Jan 1;29(1):308-311.

(130) Sarir H, Henricks PA, van Houwelingen AH, Nijkamp FP, Folkerts G. Cells, mediators and Toll-like receptors in COPD. Eur J Pharmacol 2008 May 13;585(2-3):346-353.

(131) LeVan TD, Von Essen S, Romberger DJ, Lambert GP, Martinez FD, Vasquez MM, et al. Polymorphisms in the CD14 gene associated with pulmonary function in farmers. Am J Respir Crit Care Med 2005 Apr 1;171(7):773-779.

(132) Donham KJ, Reynolds SJ, Whitten P, Merchant JA, Burmeister L, Popendorf WJ. Respiratory dysfunction in swine production facility workers: dose-response relationships of environmental exposures and pulmonary function. Am J Ind Med 1995 Mar;27(3):405-418.

(133) Chu HW, Jeyaseelan S, Rino JG, Voelker DR, Wexler RB, Campbell K, et al. TLR2 signaling is critical for Mycoplasma pneumoniae-induced airway mucin expression. J Immunol 2005 May 1;174(9):5713-5719.

(134) McClure R, Massari P. TLR-Dependent Human Mucosal Epithelial Cell Responses to Microbial Pathogens. Front Immunol 2014 Aug 12;5:386.

(135) McNamara N, Khong A, McKemy D, Caterina M, Boyer J, Julius D, et al. ATP transduces signals from ASGM1, a glycolipid that functions as a bacterial receptor. Proc Natl Acad Sci U S A 2001 Jul 31;98(16):9086-9091.

(136) DiMango E, Ratner AJ, Bryan R, Tabibi S, Prince A. Activation of NF-kappaB by adherent Pseudomonas aeruginosa in normal and cystic fibrosis respiratory epithelial cells. J Clin Invest 1998 Jun 1;101(11):2598-2605.

(137) Soong G, Reddy B, Sokol S, Adamo R, Prince A. TLR2 is mobilized into an apical lipid raft receptor complex to signal infection in airway epithelial cells. J Clin Invest 2004 May;113(10):1482-1489.

(138) Takeuchi O, Hoshino K, Kawai T, Sanjo H, Takada H, Ogawa T, et al. Differential roles of TLR2 and TLR4 in recognition of gram-negative and gram-positive bacterial cell wall components. Immunity 1999 Oct;11(4):443-451.

(139) Long H, O'Connor BP, Zemans RL, Zhou X, Yang IV, Schwartz DA. The Toll-like receptor 4 polymorphism Asp299Gly but not Thr399Ile influences TLR4 signaling and function. PLoS One 2014 Apr 2;9(4):e93550.

(140) O'Brien AD, Rosenstreich DL, Taylor BA. Control of natural resistance to Salmonella typhimurium and Leishmania donovani in mice by closely linked but distinct genetic loci. Nature 1980 Oct 2;287(5781):440-442.

(141) Couper D, LaVange LM, Han M, Barr RG, Bleecker E, Hoffman EA, et al. Design of the Subpopulations and Intermediate Outcomes in COPD Study (SPIROMICS). Thorax 2014 May;69(5):491-494.

(142) Alexis N, Soukup J, Ghio A, Becker S. Sputum Phagocytes from Healthy Individuals Are Functional and Activated: A Flow Cytometric Comparison with Cells in Bronchoalveolar Lavage and Peripheral Blood. Clinical Immunology 2000;97(1):21-32.

(143) Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods 2010 May;7(5):335-336.

(144) DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. Appl Environ Microbiol 2006 Jul;72(7):5069-5072.

(145) Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Royal Stat Soc Ser B 1995;57:289-300.

(146) Bathoorn E, Kerstjens H, Postma D, Timens W, MacNee W. Airways inflammation and treatment during acute exacerbations of COPD. Int J Chron Obstruct Pulmon Dis 2008;3(2):217-229.

(147) Huang YJ, Sethi S, Murphy T, Nariya S, Boushey HA, Lynch SV. Airway microbiome dynamics in exacerbations of chronic obstructive pulmonary disease. J Clin Microbiol 2014 Aug;52(8):2813-2823.

(148) Marsland BJ, Gollwitzer ES. Host-microorganism interactions in lung diseases. Nat Rev Immunol 2014 Dec;14(12):827-835.

(149) Cabrera-Rubio R, Garcia-Nunez M, Seto L, Anto JM, Moya A, Monso E, et al. Microbiome diversity in the bronchial tracts of patients with chronic obstructive pulmonary disease. J Clin Microbiol 2012 Nov;50(11):3562-3568.

(150) Pragman AA, Kim HB, Reilly CS, Wendt C, Isaacson RE. The lung microbiome in moderate and severe chronic obstructive pulmonary disease. PLoS One 2012;7(10):e47305.

(151) Huang YJ, Lynch SV. The emerging relationship between the airway microbiota and chronic respiratory disease: clinical implications. Expert Rev Respir Med 2011 Dec;5(6):809-821.

(152) Charlson ES, Chen J, Custers-Allen R, Bittinger K, Li H, Sinha R, et al. Disordered microbial communities in the upper respiratory tract of cigarette smokers. PLoS One 2010 Dec 20;5(12):e15216.

(153) Boissy RJ, Romberger DJ, Roughead WA, Weissenburger-Moser L, Poole JA, LeVan TD. Shotgun pyrosequencing metagenomic analyses of dusts from swine confinement and grain facilities. PLoS One 2014 Apr 18;9(4):e95578.

(154) McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 2013 Apr 22;8(4):e61217.

(155) Poole JA, Wyatt TA, Kielian T, Oldenburg P, Gleason AM, Bauer A, et al. Toll-like receptor 2 regulates organic dust-induced airway inflammation. Am J Respir Cell Mol Biol 2011 Oct;45(4):711-719.

(156) Rask-Andersen A, Malmberg P, Lundholm M. Endotoxin levels in farming: absence of symptoms despite high exposure levels. Br J Ind Med 1989 Jun;46(6):412-416.

(157) Cormier Y, Israel-Assayag E, Racine G, Duchaine C. Farming practices and the respiratory health risks of swine confinement buildings. Eur Respir J 2000 Mar;15(3):560-565.

(158) Vogelzang PF, van der Gulden JW, Folgering H, Heederik D, Tielen MJ, van Schayck CP. Longitudinal changes in bronchial responsiveness associated with swine confinement dust exposure. Chest 2000 May;117(5):1488-1495.

(159) Poole JA, Alexis NE, Parks C, MacInnes AK, Gentry-Nielsen MJ, Fey PD, et al. Repetitive organic dust exposure in vitro impairs macrophage differentiation and function. J Allergy Clin Immunol 2008 Aug;122(2):375-82, 382.e1-4.

(160) Romberger DJ, Bodlak V, Von Essen SG, Mathisen T, Wyatt TA. Hog barn dust extract stimulates IL-8 and IL-6 release in human bronchial epithelial cells via PKC activation. J Appl Physiol (1985) 2002 Jul;93(1):289-296.

(161) Nehme B, Letourneau V, Forster RJ, Veillette M, Duchaine C. Culture-independent approach of the bacterial bioaerosol diversity in the standard swine confinement buildings, and assessment of the seasonal effect. Environ Microbiol 2008 Mar;10(3):665-675.

(162) Poole JA, Dooley GP, Saito R, Burrell AM, Bailey KL, Romberger DJ, et al. Muramic acid, endotoxin, 3-hydroxy fatty acids, and ergosterol content explain monocyte and epithelial cell inflammatory responses to agricultural dusts. J Toxicol Environ Health A 2010;73(10):684-700.

(163) Poole JA, Thiele GM, Alexis NE, Burrell AM, Parks C, Romberger DJ. Organic dust exposure alters monocyte-derived dendritic cell differentiation and maturation. Am J Physiol Lung Cell Mol Physiol 2009 Oct;297(4):L767-76.

(164) Akira S, Uematsu S, Takeuchi O. Pathogen recognition and innate immunity. Cell 2006 Feb 24;124(4):783-801.