© Copyright 2016

Ian Byrell Stanaway

Longitudinal Analysis of Genetic and Environmental Data from an Agricultural Cohort Living in Yakima Valley Reveals an Agricultural Exposome

Ian Byrell Stanaway

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2016

Reading Committee:

Elaine Faustman, Chair

Julia Cui

Ali Shojaie

Program Authorized to Offer Degree:

Environmental and Occupational Health

University of Washington

Abstract

Longitudinal Analysis of Genetic and Environmental Data from an Agricultural Cohort Living in Yakima Valley Reveals an Agricultural Exposome

Ian Byrell Stanaway

Chair of the Supervisory Committee:
Elaine Faustman, Ph.D., Chair
Professor, Department of Environmental and Occupational Health

In an agricultural community cohort, we investigated agricultural pesticide exposure associated changes in the oral buccal microbiota. We found a seasonally persistent association between blood detection of the insecticide Azinphos-methyl and the taxonomic composition of the buccal swab oral microbiome. Blood and buccal samples were collected concurrently from individual subjects across two seasons, spring-summer 2005 and winter 2006. Mass spectrometry quantified blood concentrations of the organophosphate insecticide Azinphos-methyl. Buccal microbiome samples were 16S rRNA gene DNA sequenced and assigned to the bacterial taxonomy. Non-parametric analysis of the transformed microbiome data for individuals with and without Azinphos-methyl blood detection showed significant perturbations in seven common bacterial genera (>0.5% of sample mean read depth), including significant reductions in the common oral bacteria, *Streptococcus*. Diversity in composition between individuals' microbiomes was also investigated using principal components analysis to reveal two primary clusters of microbiome types. The spring-summer 'exposed' cluster with significantly less bacterial diversity was enriched for farmworkers and contained 27 of the 30 adult individuals who also had Azinphos-methyl

agricultural pesticide exposure detected in the blood. The children of each adult nearly always clustered with their respective household adult. Additionally, we investigated the additive effects of genetic minor allele variants in the Azinphos-methyl pesticide target blood cholinesterase enzymes. There is a significant additive relationship in the regression framework with environmental and genetic associations to these blood cholinesterase enzymes activities.

TABLE OF CONTENTS

TABLE OF CONTENTS 3

1. DISSERTATION INTRODUCTION - AGRICULTURAL EXPOSURE & PHENOTYPES 7
   - 1.1. The "Agricultural Exposome" - Environment, Geography and Farming Practices of Yakima Valley 7
   - 1.2. Agricultural Insecticide Exposure 8
   - 1.3. Biological Phenotypes 8
   - 1.4. Hypotheses 9
   - 1.5. The Yakima Valley Children's Health Cohort Sample Biospecimen Collection - The Composition of the Cohort 9
   - 1.6. Analytical Methods and Background Rationale for Application to this Cohort 10
     - 1.6.1. Erythrocyte Acetylcholinesterase (AChE) and Plasma Butyrylcholinesterase (BChE) Blood Cholinesterase Enzyme Activities 10
     - 1.6.2. Blood Pesticide Concentrations 10
     - 1.6.3. Primer Directed Polymerase Chain Reaction (PCR) Based 16S rRNA DNA Sequencing 11
     - 1.6.4. Genome Wide Genotyping 11

2. THE AGRICULTURAL ORAL MICROBIOME 13
   - 2.1. Abstract 13
   - 2.2. Introduction 14
   - 2.3. Materials and Methods 16
   - 2.4. Results 19
   - 2.5. Discussion 30
   - 2.6. Conclusions 32

3. ADULT-CHILD CO-CLUSTERING OF BUCCAL MICROBIOME TYPES 33
   - 3.1. Abstract 33
   - 3.2. Introduction 33
   - 3.3. Methods 35
     - 3.3.1. Participants and Agricultural Setting 35
     - 3.3.2. Study Design & Cohort General Characteristics 35
     - 3.3.3. Total Adult Blood Azinphos-methyl Pesticide 36
     - 3.3.4. 16S rRNA Buccal Swab Sequencing 36
     - 3.3.5. Statistical and Numerical Methods 38
   - 3.4. Results 39
     - 3.4.1. Blood Measures of Azinphos-methyl Pesticide Exposure 39
     - 3.4.2. Census of 16S rRNA sequence data from adults and children. 40
     - 3.4.3. Non-parametric Associations Between Exposure Groups and Genera Proportion using the Wilcoxon's Rank Sum Test 41
     - 3.4.4. Household Pome Farmworker Status 46
     - 3.4.5. Principal Component Analysis of Buccal Sample Common OTU Proportions 51
     - 3.4.6. Winter Microbiome PCA comparison to Azinphos-methyl Detection and Pome Farmworker Status 53
     - 3.4.7. Top OTU Loadings in the Principal Components Analysis 56
     - 3.4.8. Adult-Child PC1 Co-clustering 58
     - 3.4.9. Alpha Diversity Analysis of PC1 Clusters 59
   - 3.5. Discussion 59
   - 3.6. Chapter 3 Supplemental Figures 62

4. MINOR ALLELE CONTRIBUTIONS TO CHOLINESTERASE ENZYMATIC ACTIVITY VARIABILITY IN ORGANOPHOSPHATE PESTICIDE EXPOSED FARMWORKERS 66
4.1. Abstract 66
4.2. Introduction 66
4.3. Methods 68
4.3.1. Study Design, Participants and Agricultural Setting 68
4.3.2. Blood Azinphos-methyl Pesticide Concentration 68
4.3.3. Blood AChE and BChE Activities 68
4.3.4. Genotyping 69
4.3.5. Statistical and Numerical Methods 70
4.4. Results 72
4.4.1. Genotype Array Variation 72
4.4.2. Blood Phenotypes 73
4.4.3. AChE Enzyme Activity Linear Models 75
4.4.4. BChE Enzyme Activity Linear Models 77
4.4.5. AChE-BChE Enzyme-Enzyme Activity Inhibition Phenotype Relationship 79
4.4.6. Regression of Azinphos-methyl Exposed AChE and BChE Gene Minor Allele Carriers with AChE and BChE Enzyme Activity Inhibition 80
4.5. Discussion 84
4.6. Supplemental Figures 88
5. DISSERTATION CONCLUSIONS 103
5.1. Special Acknowledgements 105
6. REFERENCES 106
7. APPENDICES 117

Figures

Figure 2-1 Spring/Summary 2005 Common Buccal Sample Bacteria by Blood Pesticide Detection 22
Figure 2-2 Winter 2006 Common Buccal Sample Bacteria by Blood Pesticide Detection 23
Figure 2-3 PCA Exploration of the Oral Buccal Microbiota Relative Abundances 25
Figure 2-4 Abundances of Commonly Detected OTUs by PC Loadings and Clusters 28
Figure 2-5 Adult Alpha Diversity by Exposure Groups 29
Figure 3-1 Adult Blood Azinphos-methyl Detection 40
Figure 3-2 Spring-summer Adult and Child Buccal Bacteria by Azinphos-methyl Detection 43
Figure 3-3 Winter Adult and Child Buccal Bacteria by Azinphos-methyl Detection 45
Figure 3-4 Spring-summer Adult and Child Buccal Bacteria by Pome Farmworker Occupation 48
Figure 3-5 Winter Adult and Child Buccal Bacteria by Pome Farmworker Occupation 50
Figure 3-6 Spring-summer Adult and Child Microbiome PCA 52
Figure 3-7 Winter Adult and Child Microbiome PCA 54
Figure 3-8 PCA Loadings, Cluster Membership and Streptococcus Abundance 57
Figure 3-9 Adult and Child Alpha Diversity by PCA Cluster 59

Figure 3-10	Spring/Summary 2005 Combined Adult and Child Buccal Genera by Adult Blood AZM	62
Figure 3-11	Winter 2006 Combined Adult and Child Buccal Genera by Adult Blood AZM Detection	63
Figure 3-12	Spring-Summer 2005 Combined Adult and Child Buccal Genera by Household Occupation	64
Figure 3-13	Winter 2006 Combined Adult and Child Buccal Genera by Household Occupation	65
Figure 4-1	Blood Assay Results for Spring/Summer 2005 and Winter 2006	74
Figure 4-2	AChE Blood Phenotype Gene + Environment Generalized Linear Model	75
Figure 4-3	BChE Blood Phenotype Gene + Environment Generalized Linear Model	77
Figure 4-4	AChE Inhibition of Activity Plotted by BChE Enzyme Activity in Azinphos-methyl Exposed Individuals	79
Figure 4-5	AChE Inhibition by Azinphos-methyl	81
Figure 4-6	Spring-Summer AChE Enzyme Activity Model Residuals Diagnostic Plots	88
Figure 4-7	Spring-Summer AChE Enzyme Activity Residuals	89
Figure 4-8	Spring-Summer Partial Residuals Plot	89
Figure 4-9	Winter AChE Enzyme Activity Model Residuals Diagnostic Plots	90
Figure 4-10	Winter AChE Enzyme Activity Model Residuals Diagnostic Plots	91
Figure 4-11	Winter AChE Enzyme Activity Model Residuals Diagnostic Plots	91
Figure 4-12	Spring-Summer AChE Enzyme Activity Time-Series Model Residuals Diagnostic Plots	92
Figure 4-13	Spring-Summer AChE Enzyme Activity Time-Series Model Residuals Diagnostic Plots	92
Figure 4-14	Spring-Summer BChE Enzyme Activity Model Residuals Diagnostic Plots	93
Figure 4-15	Spring-Summer BChE Enzyme Activity Model Residuals Diagnostic Plots	94
Figure 4-16	Spring-Summer BChE Enzyme Activity Model Residuals Diagnostic Plots.	94
Figure 4-17	Winter BChE Enzyme Activity Model Residuals Diagnostic Plots	95
Figure 4-18	Winter BChE Enzyme Activity Model Residuals Diagnostic Plots	96
Figure 4-19	Winter BChE Enzyme Activity Model Residuals Diagnostic Plots	96
Figure 4-20	Spring-Summer AChE Enzyme Activity Time-Series Model Residuals Diagnostic Plots	97
Figure 4-21	Spring-Summer AChE Enzyme Activity Time-Series Model Residuals Diagnostic Plots	97
Figure 4-22	AChE Enzyme Inhibition of Activity Model Residuals Diagnostic Plots	98
Figure 4-23	AChE Enzyme Inhibition of Activity Model Residuals Diagnostic Plots	99
Figure 4-24	AChE Enzyme Inhibition of Activity Model Residuals Diagnostic Plots	99
Figure 4-25	BChE Enzyme Inhibition of Activity Model Residuals Diagnostic Plots	100
Figure 4-26	BChE Enzyme Inhibition of Activity Model Residuals Diagnostic Plots	101

Figure 4-27 BChE Enzyme Inhibition of Activity Model Residuals Diagnostic Plots 101
Figure 7-1 Email from Oksana Lockridge Concerning BCHE Variants and Risk 117

Tables

Table 2-1 Wilcoxon's Rank Sum Test of Azinphos-methyl Blood Detection Groups for Adult Buccal Microbiome Compositional Perturbation 20
Table 3-1 Wilcoxon's Rank Sum Tests of Spring-Summer 2005 Adult and Child Buccal Microbiota by Adult Azinphos-methyl Blood Detection 42
Table 3-2 Wilcoxon's Rank Sum Tests of Winter 2006 Adult and Child Buccal Microbiota by Adult Azinphos-methyl Blood Detection 44
Table 3-3 Wilcoxon's Rank Sum Tests of Spring-summer 2005 Adult and Child Buccal Microbiota by Household Pome Farmworker Status 47
Table 3-4 Wilcoxon's Rank Sum Tests of Winter 2006 Adult and Child Buccal Microbiota by Household Pome Farmworker Status 49
Table 3-5 Azinphos-methyl Detection Spring-Summer 2005 Fisher's Exact Tests of PCA Cluster Membership 53
Table 3-6 Pome Farmworkers Spring-summer 2005 Fisher's Exact Tests of PCA Cluster Membership 53
Table 3-7 Azinphos-methyl Detection Winter 2006 Fisher's Exact Tests of PCA Cluster Membership 55
Table 3-8 Pome Farmworkers Winter 2006 Fisher's Exact Tests of PCA Cluster Membership 55
Table 3-9 Household Adult-Child Pairs Principal Component Co-Clustering Match Counts 58
Table 4-1 AChE and BChE Protein Coding Minor Alleles Tested in the Gene + Environment Model 73
Table 4-2 Summary of AChE Activity Regression Results by Season 76
Table 4-3 Summary of AChE Activity Time Series Regression Results 76
Table 4-4 Summary of BChE Activity Regression Results by Season 78
Table 4-5 Summary of BChE Activity Time Series Regression Results 78
Table 4-6 Summary of AChE and BChE Activity Inhibition Regression Results 80

1. DISSERTATION INTRODUCTION - AGRICULTURAL EXPOSURE & PHENOTYPES

1.1. The "Agricultural Exposome" - Environment, Geography and Farming Practices of Yakima Valley

The Yakima River Valley in central Washington State, situated on the east side of the Cascade mountains, is an intensely cultivated agricultural district. The valley benefits from the Cascade mountains' rain shadow effect making it an arid dry high mountain desert climate with many days of sun throughout the year situated on fertile volcanic soils. Despite the dry sunny conditions, farming is readily performed here due to irrigation and public water projects that utilize the region's many rivers. The Yakima River is a large tributary to the Columbia River which bisects the state originating in Canada, flowing south to the Oregon-Washington border forming the state border, and continuing west to the Pacific Ocean. The Columbia River as the fourth largest river in the United States by volume of water is supported by a drainage basin roughly the size of France with many farming valleys like the Yakima that make the region a natural economic resource to Washington State's development in both historically and in modern times.

Farming activities in the Yakima River Valley include livestock, grain, hops, vineyards, cherries, apples, pears, cherries and peach production, giving a wide variety of crops and livestock that make the valley's economic contributions to Washington State very important. Among these diverse crops, Washington Apples and other pome fruit, such as pears, have gained important prominent positions in the economic visibility and product placement of Washington State. These farms' economic production is maintained by the essential efforts performed by farmworkers. Farming as an occupation has a unique hazard profile with exposure to many dangers ranging from chemical and biological hazards to obvious physical hazards like tractors and large moving equipment. Humans have endured exposure to natural hazards since the invention of agriculture ~12,000 years ago when man began cultivating and extracting resources from our environment. In the last century, newer hazards associated with the chemical and physical dangers of modern industrial agriculture have been introduced to this historic farming practice.

This dissertation, assess susceptibility to chemical hazards by analyzing genetic and exposure data derived of blood, urine and oral buccal swab samples collected from ~200 Yakima Valley households. This effort was funded by a Children's Health Cohort (CHC) grant from the National Institute of Environmental Health Sciences (NIEHS) and United States (US) Environmental Protection Agency (EPA) directed at finding and quantifying agricultural exposures that may impact child health and development in an underserved agricultural community. This agricultural cohort at the University of Washington has a sample collection housed by the Center for Child Environmental Health Risks Research. The environmental pesticides and aspects of exposure are well studied in this cohort [1-15].

1.2. Agricultural Insecticide Exposure

The top three pesticides applied to the Washington State apple fruit production area as reported by the USDA in 2005 were the insecticides Azinphos-methyl, Carbaryl and petroleum distillates (USDA 2005 Washington State Pesticide Usage Report http://www.nass.usda.gov/Surveys/Guide_to_NASS_Surveys/Chemical_Use/). Both Azinphos-methyl and Carbaryl insecticides have a pesticidal mode of action where they inhibit cholinesterase enzyme activities leading to muscle and neuronal paralysis that high doses can result in death. There is a relatively large differential window in dose toxicity between mammals and insects, and the relatively low environmental persistence [16, 17] is what made Azinphos-methyl and other related organophosphate pesticides (OPs) attractive options for controlling insect pests in fruit crops and resulted in their extensive use. The main insect pest that the early summer pome fruit pesticide applications are timed to control is the Codling Moth (*Cydia pomonella*). The Codling Moth's ancestral ecological range originated in Asia Minor much like the apple and now they are found in temperate apple growing areas worldwide and can have 1 to 5 generations in a summer growing season. The larvae burrow into the apples destroying their economic value. Despite the effectiveness of Azinphos-methyl at controlling the Codling Moth and the short persistence it exhibits in the environment, it has been shown that Azinphos-methyl is too toxic for acute exposures to maintain a rational margin of exposure in farmworker populations. The EPA as of September 30th, 2013, has phased out the use of Azinphos-methyl in the United States. This makes the exposure information of the Yakima River Valley Children's Health Cohort a unique dataset to formulate hypotheses to inform our understanding of the "Agricultural Exposome" in the context of pesticide exposure and human metabolism.

1.3. Biological Phenotypes

The perceived idea of phenotype alludes to measured differences in the biologic composition of organisms of the same species which have different traits. This paradigm focuses on the organisms studied to make detailed inventories of the biological parts that are different between groups of organisms and usually attributes the different phenotypic traits to specific genotypes. While this definition is adequate for many static phenotypes, it leaves out the influence of the environment and exposure on the evolutionary shaping and expression of the phenotype. Phenotype can also be defined as a gene operating in an environment. This brings the idea that gene traits can be dynamic in their response to their environment. If a gene is not stimulated in the particular environment studied, then there will be no measured dynamic differences in phenotypic trait between groups of organisms under this environmental condition. Hence this gives rise to the phenotype formulation:

$$\text{Phenotype}_{ij} = \text{Gene}_{ij}\ x(+)\ \text{Environment}_{ij} \quad \text{Eq. 1}$$

The corollary that closely follows this supposition is that some static phenotypes may be expressed in an environment which they are not needed in anymore. In these cases, the relaxation of environmental selection for the static trait also leads to mutation driven erosion of phenotypic expression over generations of the organism population and hence becomes vestigial. As a biologist with a history in genetics and an active interest in defining

phenotype, this concept that phenotype can be changed by both mutation driven genotypes and environmental stimuli has driven my interest in studying Environmental Toxicology in graduate school to complement my experience as a geneticist. Another antecedent to phenotype is that a phenotype may be the commensal organisms that manage to inhabit a host given their genotypes and environment. This offers a different trophic layer which can have variability associated with a host and between hosts and the microbiome.

1.4. Hypotheses

In this dissertation, the hypothesis that environmental exposure to agricultural conditions is associated with altered phenotypes in this cohort was tested. When possible the genetic component was incorporated into the analytical model to show that a particular phenotype can have both environmental and genetic components contributing to the phenotypic expression. The two phenotypic systems that are examined are the oral-buccal swab microbiome composition and blood cholinesterase activities of this agricultural cohort.

The two hypotheses generally stated are:

General Hypothesis 1: The composition of the oral buccal microbiome is associated with agricultural exposures.

General Hypothesis 2: The enzyme activities of genes involved in pesticide xenobiotic metabolism are associated with genetic minor alleles and environmental exposure.

1.5. The Yakima Valley Children's Health Cohort Sample Biospecimen Collection - The Composition of the Cohort

The Yakima Valley Children's Health Cohort sample collection of blood, urine and buccal swabs studied here are part of a larger project sample collection that has had three funding cycles of sample collections over ~15 years of time. This study's cohort design was organized to assess Yakima River Valley farm worker and non-farm worker community members and their children's seasonal exposure to organophosphate pesticides, including Azinphos-methyl. In Summer 2005 and Winter 2006 sample collections studied in this dissertation, matched longitudinal samples of house dust, vehicle dust, urine and blood were collected from adults, from approximately 200 households. Nearly all subjects were self-identified as Hispanic and approximately 50% of these households were employed as farmworkers in the pome fruit (apple or pear) production. Sample collections were designed around the seasonal nature of agricultural pesticide application practices with a sample collection in the early summer and another in the winter. This sampling design reflects an exposure difference of the agricultural cycle between the seasons. In early summer, pome fruit farm workers enter the orchards shortly after pesticide applications to "thin" small fruits and buds allowing the remaining fruit to become larger and of more value. In the winter non-spray season there is relatively little farming activity that would expose the farm worker to pesticides as the last pesticide applications were months before and environmental degradation reduces the frequency of acute exposures significantly.

Some of the founding hypotheses tested in the original study design were to see if community awareness and education on methods of changing clothes and boots after working in the orchards would alter the pesticide exposure of the subjects, especially the children. Additionally, they aimed to elucidate the primary pathways of exposure and quantify exposures in households and individuals. In this success was found in showing dust in the households and vehicles has association to urine pesticide metabolites and parent compound blood concentrations of both adults and children studied. The urine metabolite data also shows a 10-fold difference in exposure of this population in comparison to that of both the non-farmworker portion of the cohort and National Health and Nutrition Examination Survey (NHANES) data representative of the rest of the US pesticide exposure. This exposure difference underlines the care that must be taken in assessing pesticides for occupational use.

1.6. Analytical Methods and Background Rationale for Application to this Cohort

1.6.1. Erythrocyte Acetylcholinesterase (AChE) and Plasma Butyrylcholinesterase (BChE) Blood Cholinesterase Enzyme Activities

Erythrocyte acetylcholinesterase (AChE) and plasma butyrylcholinesterase (BChE) blood cholinesterase enzyme activities were measured using the a kit (Blood Cholinesterase Testing Kit Manual: http://www.eqmresearch.com/Manual-E.pdf) that utilizes the method of Ellman et al [18]. AChE and BChE enzyme activities are respectively expressed in the units of activity Units / gram of hemoglobin (U/g Hgb) or activity Units / mL of blood plasma (U/mL plasma). The blood cholinesterase activity is inactivated by the CYP450 metabolite oxon form of the parent compound organophosphate pesticides and blood activity buffers the inhibition of cholinesterase activity in the central nervous which results in toxicity. The measurement of blood cholinesterases allows the indirect assessment of organophosphate pesticides (OP) dose based on the percent of inhibition from the activity level when no OP dose is present. When the level of inhibition is more than 20% a farmworker or pesticide handler is removed from work exposures until restoration of normal activity.

1.6.2. Blood Pesticide Concentrations

Blood pesticide concentrations were determined by sending sera to the Pesticide Laboratory in the National Center for Environmental Health at the CDC for pesticide concentration measurements. Blood sera concentrations of Azinphos-methyl were measured using isotope dilution gas chromatography-linked high resolution mass spectrometry (GC-HR-MS) by the method of Barr et al [19]. The limit of detection for Azinphos-methyl was 0.04 ng/g of blood. These pesticide blood level values are a biomarker of exposure. Blood concentrations should not be interpreted as an absolute measure of dose as we do not know where in the time course of metabolism and clearance excretion of the area under the concentration curve (AUC) approximation of total dose these individuals are at.

1.6.3. Primer Directed Polymerase Chain Reaction (PCR) Based 16S rRNA DNA Sequencing

Primer directed polymerase chain reaction (PCR) based 16S rRNA DNA sequencing of the buccal swab samples were used to determine the proportional bacterial abundances of the sampled oral microbiomes in summer and winter seasons. Sequencing was performed on the Ion Torrent platform. To identify the organisms present in the oral buccal microbiomes, bacterial 16S rRNA DNA sequences from all individuals were clustered to Operational Taxonomic Units (OTUs) at 97% nucleotide sequence identity where a representative sequence of the OTU cluster is then classified to the genus level using the Ribosomal Database Project Bayesian Classifier [20]. This produces a *de-novo* internally consistent taxonomic reference sequence set of bacterial organisms detected collectively across the set of samples generated by the sequencing pipeline. Following this taxonomic reference sequence set generation, the individual sample's reads are mapped back onto the *de-novo* taxonomic OTU reference representative sequences at 97% sequence identity providing a count of reads for each OTU for each sample. In most analyses this is used to calculate the proportional abundance of each taxonomic type of bacteria at the OTU level or the collapsed across OTUs classified to the appropriate taxonomic rank identified in the specific analysis.

Proportional_Abundance_of_Classified_Taxon =

Count_Sample_Taxon_Sequences / Total_Count_Sample_Sequences

Eq. 2

The rationale for characterizing the buccal oral microbiome is the idea that environmental exposure to agricultural pesticides and/or organisms would be reflected in the composition of the microbiome.

1.6.4. Genome Wide Genotyping

Genome wide genotyping of individuals was performed mainly on blood (and in a few cases buccal swab) DNA extracts with the Illumina 5M genotyping array chip (www.illumina.com). The design of the genotype site selection for this array was based on deep exome and genome sequencing of the collaborative efforts of the 1000 Genomes Project and the ~12,000 individuals sequenced by the Exome Sequencing Project. These two projects provide coverage of both common and rare genotype variants. Prior to these genotype discovery efforts, most array genotyping platforms only provided a selection of common genetic variants that were used for genetic association to phenotypes. Previously this cohort had also been genotyped on the DMET Affymetrix array at ~2000 common variants in ~225 drug metabolism genes.

<u>Rare and common variant site selection</u> on the recent Illumina arrays is informed by a minor allele population genetics model where with the explosive increase in the human population in recent history, there has been a large increase in the number of rare variants in the human population [21]. This has increased the burden of variation that has not been under purifying selection due to the recent nature of the rare variations introduction to the population. These

rare variant alleles are more likely to have a change of phenotype. This is of interest in susceptible populations for disease risk, especially in the context of environmental exposure where genetic variants are not likely to be deleterious unless the carrier is exposed to a condition where they may not metabolize an exogenous (or endogenous) compound as efficiently and be more susceptible to the toxic effects of said compounds.

2. THE AGRICULTURAL ORAL MICROBIOME

Human Oral Buccal Microbiomes are Associated with Farmworker Status and Azinphos-methyl Agricultural Pesticide Exposure

Ian B. Stanaway[1,2], James C. Wallace[1,2], Ali Shojaie[3], William C. Griffith[1,2], Sungwoo Hong[1,2], Carly S. Wilder[1,2], Foad H. Green[1,2], Jesse Tsai[1,2], Misty Knight[1,2], Tomomi Workman[1,2], Eric M. Vigoren[1,2], Jeffrey S. McLean[4], Beti Thompson[5] and Elaine M. Faustman[1,2*]

[1] Department of Environmental and Occupational Health Sciences, University of Washington, Seattle, Washington, United States of America
[2] Institute for Risk Analysis and Risk Communication, University of Washington, Seattle, Washington, United States of America
[3] Department of Biostatistics, University of Washington, Seattle, Washington, United States of America
[4] School of Dentistry, Periodontics, University of Washington, Seattle, Washington, United States of America
[5] Fred Hutchinson Cancer Research Center, Seattle, Washington, United States of America

*Corresponding Author:
Email: faustman@u.washington.edu (EMF)

2.1. Abstract

In a longitudinal agricultural community cohort sampling of 65 adult farmworkers and 52 adult non-farmworkers, we investigated agricultural pesticide exposure associated changes in the oral buccal microbiota. We found a seasonally persistent association between the detected blood concentration of the insecticide Azinphos-methyl and the taxonomic composition of the buccal swab oral microbiome. Blood and buccal samples were collected concurrently from individual subjects across two seasons, spring-summer 2005 and winter 2006. Mass spectrometry quantified blood concentrations of the organophosphate insecticide Azinphos-methyl. Buccal microbiome samples were 16S rRNA gene DNA sequenced, assigned to the bacterial taxonomy and analyzed after 'centered-log-ratio' transformation to handle the compositional nature of the proportional abundances of bacteria per sample. Non-parametric analysis of the transformed microbiome data for individuals with and without Azinphos-methyl blood detection showed significant perturbations in seven common bacterial genera (>0.5% of sample mean read depth), including significant reductions in the common oral bacteria, *Streptococcus.* Diversity in 'centered-log-ratio' composition between individuals' microbiomes was also investigated using principal components analysis (PCA) to reveal two primary PCA clusters of microbiome types. The spring-summer 'exposed' cluster with significantly less bacterial diversity was enriched for farmworkers and contained 27 of the 30 individuals who also had Azinphos-methyl agricultural pesticide exposure detected in the blood.

Importance

In this study, we show in human subjects that organophosphate pesticide exposure is associated with large-scale significant alterations of the oral buccal microbiota composition with extinctions of whole genera suggested in some individuals. The persistence of this association from the spring-summer to the winter also suggests long lasting effects on the commensal microbiota have occurred. The important health related outcomes of these agricultural community individuals' pesticide associated microbiome perturbations are not understood at this time. Future investigations should index medical and dental records for common and chronic diseases that may be interactively caused from this pesticide exposure-microbiome alteration association.

Key Words: Farmworkers, Azinphos-methyl, Oral, Microbiome, Bacteria, Buccal Mucosa, 16S rRNA, Sequencing

2.2. Introduction

Oral microbiome taxonomic diversity and community composition is variable both within and between individuals, but as a whole microbiomes are relatively stable in an individual adult developed host [22]. Xenobiotic-microbiome interactions have been targeted as important to drug metabolism and biotransformation [23] which also makes it an important interaction term in the study of toxicology. Understanding of the microbiome, environmental exposure and potential impacts on human health is currently growing: It has been elucidated that the microbiome can affect cognition [24-26], metabolic syndrome and obesity [27-29], and social development [30, 31]. Many common environmental exposures [32-39] can change the microbiome composition; these can be as simple as dietary differences [34, 35] and be induced by probiotics [37] and antibiotics [32]. It has also been shown that community composition changes in the microbiome are persistent when they occur as the result of extinction events and are rescued only when both nutrients (diet as environment) and microbiome transplantation (reseeding or exposure) allow recovery [40]. Additionally, changes in the microbiome can have particularly compelling effects on tertiary phenotypes, which can affect the well-being of individuals. For example a Danish cohort study of prenatal exposure to antibacterial agents found an increased risk of being obese and overweight later in life in individuals with prenatal exposure to systemic antibacterials [41]. Another study in Canada showed that the alteration of microbial compositions affects the risk of childhood asthma [42]. These reports point at the microbiome as a common determinant of other phenotypes with pleiotropic effects on human health. Understanding the environmental effects of exposure on the microbiome is thus particularly important for public health.

Organophosphate pesticides (OPs), including Azinphos-methyl, are a class of toxicants that have had favor in the agricultural industry due to their relatively low persistence in the environment [16, 17]. Acute OP toxicity in higher organisms primarily operates via a common mechanism of covalently binding a serine residue in the active site of the acetylcholinesterase enzyme (AChE); inactivating this site results in a systemic inhibition of enzyme activity [43]. AChE cleaves the neurotransmitter acetylcholine, which is released in synapses when they fire. By inhibiting the enzyme cleavage of acetylcholine, synapses are

not allowed to reset, which may result in death at high dose in adults, and in neuro-developmental abnormality in children, at lower doses. Residential studies of exposure to the OP chlorpyrifos have shown reductions in working memory, intelligence and brain anomalies in the babies of exposed mothers [44-48], as well as laboratory animal experiments [49]. Studies of acute exposure to these compounds have resulted in use restrictions of some OPs in both occupational and residential settings. As a result, the EPA has phased out any use of Azinphos-methyl as of September 30, 2013 (http://www.epa.gov/opp00001/reregistration/Azinphos-methyl/phaseout_fs.htm).

The ~200 household Yakima Valley community agricultural cohort studied here provides seasonal longitudinal measures of exposure to OPs using blood samples collected in the spring-summer of 2005 and winter 2006 for adults and children [3]. Sequenced buccal samples from 117 adult individuals were used to determine proportional abundances of oral microbiome bacteria detected in buccal swabs. Blood concentrations of Azinphos-methyl have been determined by mass spectrometry in this family-based community pesticide exposure study [3]. Samples were collected in spring-summer 2005 to measure pesticide exposure during the spring to summer fruit-thinning season, when the workers had a high likelihood of OP exposure. The winter 2006 samples assessed exposure during the off-season, when the likelihood of exposure was low. Seventy individuals had buccal samples 16S rRNA sequenced from both longitudinal collection seasons. Bacterial detection was performed by DNA sequencing of the V5 and V6 regions of the 16S ribosomal RNA (rRNA) gene from 206 adult buccal swab samples. Counts of reads mapping to Operational Taxonomic Units (OTUs) [50] with >97% sequence identity were converted to sample proportional abundances. The microbiota OTUs were assigned bacterial taxonomies by the Ribosomal Database Project (RDP) Classifier [20]. Proportional taxonomic data falls into the category of compositional data analysis with simplex geometry, where the correlation structure of reductions in abundance of one bacterial taxa is correlated with compensatory proportional increase in another taxa due to the remaining read counts being occupied by other compositional members [51]. The 'centered-log-ratio' of each sample's taxonomic proportional abundances applied here has been proposed as an appropriate transformation of the data to correct for this compositional nature of the data [51-53]. To avoid complications due to limits of detection of rare member of the microbiota when read depths vary between samples, we have chosen to analyze associations in only common taxonomic genera members of the community composition (>0.5% mean proportion of sample read depth) as defined by the spring-summer 2005 samples. We adopted a slightly less stringent definition for common (>0.1% sample mean OTU proportion) when conducting exploratory principal component analysis (PCA) at the OTU taxonomy resolution, versus the genera-level taxonomy used in hypothesis testing of exposed groups. To model read depths and the probability that we would detect a taxonomic proportion, we used the negative binomial probability, $Pr[missed\ taxon] = e^{-pn}$, where p is the proportion of reads identified to that taxon and n is the total number of reads for that sample [54, 55]. This allowed us to establish criteria from the samples' taxa mean proportions (e.g., genera >0.005, OTUs >0.001) for which taxa to investigate at the read depth of sequencing sampled. Using only deeply sequenced samples (min threshold: 2,500 reads; median: 38,454; mean: 55,321) we fix a high probability at the median read depth, $Pr[missed\ 0.001] < 1.99\times10^{-17}$ that we will observe the identified common taxa in the Dirichlet process if the taxa exists in the sample above the

proportion threshold. This common taxon analysis with deeply sequenced samples mitigates the false zero-inflation of non-detects, which can occur when analyzing varying depths of reads sampled and rarely detected taxa. We have also adopted this common genus detection limit threshold to limit the number of hypotheses tested to only common microbiota; to further reduce the possibility of bias, non-parametric Wilcoxon's Rank-Sum tests were used to test for differences in microbial abundances. Using this approach, we identify robust false discovery adjusted associations between the identified microbiota abundance and Azinphos-methyl exposure in this Hispanic Yakima Valley Washington State community agricultural cohort. We additionally extend this result into an exploratory PCA of the 'centered-log-ratio' common OTU sample proportions present in buccal microbiomes of pesticide exposed individuals and other members of the Yakima Valley Hispanic community.

2.3. Materials and Methods

Agricultural Cohort Description: Individuals are ~50% self-identified Hispanic apple and pear orchard farmworkers and ~50% Hispanic non-farmworker community members. The cohort was established in 2005 and consisted of 200 adults and 200 children from the same households [3], 117 of these individuals are studied in this buccal microbiome analysis report. The Center for Community Health promotion of the Fred Hutchinson Cancer Research Center collected the human subjects' samples in the Yakima Valley agricultural region of Washington State. Samples are housed by the University of Washington's Center for Child Environmental Health Risks Research.

Seasonal Sampling Design: Human oral buccal swabs and intravenous blood samples were collected in the spring-summer 2005 and the winter 2006. The spring-summer 2005 sample collection was timed for spring through early summer (April – July, 2005) when farmworkers thin the fruit trees and were potentially exposed to pesticides recently applied to control insects. The winter 2006 sample collection was timed for the off-season (December, 2005 – March, 2006) when pesticides were not in use. Each participating adult provided written informed consent. The Fred Hutchinson Cancer Research Center's Institutional Review Board reviewed and approved collection procedures for this study (File IR 5946).

Measurements of Azinphos-methyl Blood Concentrations: Isotope dilution gas chromatography-linked high resolution mass spectrometry (GC-HR-MS) [19] was used to determine blood sera concentrations of Azinphos-methyl at the Pesticide Laboratory in the National Center for Environmental Health at the Center for Disease Control.

Sample Collection of Oral Buccal Mucosa: The Catch-all™ Sample Collection Swab (Epicentre, Madison, WI) was used for oral buccal mucosa sample collection. Participants were asked to not smoke, eat or drink (except water) for 30 minutes before sampling. The oral cavity was rinsed twice with water before sampling. The swab was moved circularly inside the mouth on each cheek 20 times, then stored in a 15 mL polypropylene tube with the tip down in 1 mL of RNAlater (Qiagen, Venlo, Netherlands). Initially samples were stored in the field office at -10°C and moved to the University of Washington laboratory -80 °C freezer.

DNA Preparation: Buccal swab samples were thawed and diluted in 9 mL of phosphate buffered saline (PBS). The swabs were agitated in a circular motion continuously while the technician viewed the PBS for oral-buccal debris turbidity and then discarded the swab. The PBS suspensions were then centrifuged at 400g at room temperature for 10 minutes and the supernatants aspirated and discarded. The manufacturer's protocol instructions for the Maxwell® 16 Buccal Swab LEV DNA Purification Kit (Promega, Fitchburg, Wisconsin) for sample DNA extraction and elution were performed. Quantification of the sample extract's double stranded DNA concentrations was performed by the manufacturer's protocol instructions with the Quant-iT™ PicoGreen® dsDNA Assay Kit (Thermo Fisher Scientific, Waltham, Massachusetts, USA).

16S rRNA Polymerase Chain Reaction (PCR) and Bead DNA Cleanup: Eurofins (Eurofins Scientific, Luxembourg) primers, 10ng of sample DNA extract template and EmeraldAmp GT PCR Master Mix (Clontech Laboratories, Inc., Mountain View, CA) were combined in PCR tubes with Nuclease free water to primer concentrations of 0.9 µM in a volume of 50 µL. Molecular grade water negative control blanks were also thermocycled with the same reagents. The 16S ribosomal subunit DNA gene primers were designed following the gene placement described as in Cai et al [56]. Primer sequences targeted the 16S rRNA gene variable region 5 (V5, ATTAGATACCCNGGTAG) and variable region 6 (V6, CGACAGCCATGCANCACCT). Primer design was bidirectional with four oligonucleotides in total, the two above sequences and their reverse complements. Primer synthesis included the Ion™ sequencing adaptors and DNA barcodes. PCRs were thermocycled using a MJ Research PTC 200 Peltier Thermal Cycler (MJ Research, St. Bruno, Quebec, Canada) with the temperature cycle protocol (Step 1: 95 C, 10 min, Step 2: 95 C, 30 sec, Step 3: 50 C, 30 sec, Step 4: 72 C, 1 min 15 sec, Step 5: Goto Step 2, 1x, Step 6: 95 C, 30 sec, Step 7: 68 C, 30 sec, Step 8: 72 C, 10 min, Step 9: Goto Step 6,34x, Step 10: 72 C, 10 min, Step 11: 4 C). PCR amplicons were purified with the Agencourt AMPure XP PCR purification system (Beckman Coulter, Brea, CA).

Buccal DNA PCR Amplicon Pooling: Sample PCR amplicon fragment length and concentration for sequencing was quantified with the Agilent High Sensitivity DNA Kit (Agilent Technologies Inc., Santa Clara, CA). Amplicons with peaks greater than the High Sensitivity DNA reference markers were diluted and re-analyzed. Amplicon templates were diluted to a concentration of ~26pM in Low TE buffer and pooled for sequencing.

Template Preparation, Enrichment, and Sequencing: DNA template preparation of the PCR amplicons was performed using the Ion PGM Template OT2 400 kit (Life Technologies, Foster City, CA) on the Ion OneTouch 2 System (Life Technologies) and sequenced on the Ion Torrent PGM System (Life Technologies) with the Ion PGM Sequencing 400 Kit (Life Technologies) and Ion 318 Chip v2 (Life Technologies). The default Ion Torrent Browser template program for 16S Metagenomics Target Sequencing was used.

Read Sequence Preprocessing: Each barcoded samples' sequences were provided in individual FASTQ files by the Ion Torrent native software suite. Quality control included removing reads without both 5' and 3' primer sequences, primer trimming and removing reads with greater than three expected errors based on the PHRED base quality score [50].

This reduced the error per read on an Ion Torrent system to be less than 0.012 (3/247 bp mean read length).

De-novo 16S rRNA Operational Taxonomic Unit (OTU) Reference Generation: A single FASTQ file was concatenated from all samples and quality controlled as stated above resulting in a FASTA file. Exact duplicate read dereplication was performed via the USEARCH software [50] using the -derep_fullength and -sizeout flags. The FASTA file was organized in descending order of number of observations with singleton sequences being removed using the USEARCH -minsize 2 –sortbysize functions. Sequence reads were then clustered to OTUs of 97% sequence identity and the selected cluster consensus sequence representative for each OTU cluster was output to an OTU FASTA file with the USEARCH -cluster_otus function. Chimera read removal using the USEARCH -uchime_ref -strand plus flags and the -db gold.fa reference file removed PCR artifacts from the OTU FASTA file. This file was labeled with the USEARCH fasta_number.py script. The alignment target file for later use in individual sample read sequence alignments to OTUs was then made with the USEARCH -makeudb function. OTU reference sequences assignments to bacterial taxonomies down to the genus level were done with the default settings of the RDP Classifier v2.6 java program [20]. At a read sequence alignment of 0.97 in OTU generation and with a mean read length of 247 bp there will be in excess of ~7.41 bases different between OTUs, with less than 3 of the base differences between OTUs being sequencing errors inflating the count of rare members of the OTU taxonomy. This margin between OTUs sequence identity definition and known sequencing error profile gives a balanced taxonomic resolution in the face of known errors that all sequencers have and can be modeled efficiently with a PHRED score [57, 58].

Sample 16S rRNA Read Sequence Alignments: Read sequences were quality controlled as stated above. Two-step Chimera detection with the USEARCH -uchime_denovo and the USEARCH -uchime_ref -strand plus flags with the -db gold.fa reference file removed PCR artifacts. Each sample's reads were aligned by USEARCH with the commandline flags -usearch_global -strand both -id 0.97 to the *de-novo* OTU references to result in a count of reads matching each OTU.

Data Analysis: Custom Perl scripts were used to organize taxonomy, sample OTU count and sample categorical descriptor tables from the 16S rDNA USEARCH alignment files and RDP Classifier into formats for association and categorical analysis. The R package *phyloseq* [59] was used with these tables for species richness Chao alpha diversity measures. Analyses in *phyloseq* were conducted at the OTU classification level. When conducing total diversity measures that depend on the numbers of singletons, doubletons and so on in the tail of the distribution, no proportional frequency cutoffs were employed. Whenever analyses required statistical comparisons between orthogonal detection groups (i.e., Azinphos-methyl blood detection), OTUs were pooled by genera identified with the RDP classifier; genera identified at >0.5% mean proportional abundance in the spring-summer 2005 sample were used in subsequent analyses. PCA was performed using the R base function *prcomp*() on 'centered-log-ratio' transformed proportions via the R package *compositions* [60] and on the OTUs detected in >0.1% mean proportion of the spring-summer 2005 samples' reads. This PC analysis allowed an exploratory examination of the beta diversity differences between

individuals within the cohort. Cluster analysis of the PC1 and PC2 scores was performed by the *mclust*() R package [61]. PC loadings from the spring/summer data were used to project the winter OTU centered-log-ratio transformed proportions into the spring/summer data space generating projected PC1 and PC2 scores and to assign the winter cluster based on the spring/summer fit for the data. The base R implementations of statistical tests (Fisher's Exact Test, Wilcoxon Rank Sum test, Welch's t-test) were used where noted in the text to identify enrichment of categorical variables between PC clusters.

2.4. Results

Azinphos-methyl Blood Exposure Measures: The blood concentrations of Azinphos-methyl were measured by the GC-HR-Mass Spectrometry [19]. Of the 117 individuals presented here, Azinphos-methyl was detected in 36 farmworkers and one non-farmworker where we also had buccal microbiome data passing quality control. The detected concentration range in individuals was 0.021-6.192 ng Azinphos-methyl/g blood serum. The limit of detection was 0.04 ng/gram of plasma. One sample had a value less than the limit of detection that was non-zero: 0.021, this sample was treated as a non-detect. The remainder of the assayed non-detects were zeros. We formatted the range of the numbers such that values from the limit of quantification, 0.04 to 1 ng/g (the median) are coded as 0.1 and values greater than 1 are coded as 1.1 ng/g plasma. This was to help ensure another level of privacy for the human subjects in the published form and does not affect the results as we mainly use the data categorically as detect or non-detect in this analysis. Only farmworkers had detected levels of Azinphos-methyl in the spring-summer 2005 blood sera samples. All sampled individuals from the spring-summer collection had an Azinphos-methyl blood assay result. Blood samples collected in the winter 2006 non-spray season had one non-farmworker sample with a detected concentration of Azinphos-methyl. This individual was a non-detect in the matched spring-summer 2005 blood sample. Thirteen of 95 individuals with winter microbiome data did not have blood sample Azinphos-methyl assay results from the winter blood samples. They did have an Azinphos-methyl blood result from the spring-summer collection, and this value was used to categorize these individual as to their blood detection category in the spring-summer season microbiome analysis. Since we did not have a winter blood Azinphos-methyl assay value for these individuals, they were removed from the winter microbiome analyses.

Buccal Bacterial Taxonomic Assignments: We observed DNA bands at ~300 base pairs (bp) in size by agarose gel electrophoresis of the PCR amplicon of the 16S rRNA V5 and V6 regions as expected. Sequencing of the PCR amplicons was performed with the Ion Torrent DNA sequencer. Reads were quality control discarded if the primers were not detected during trimming, if the read had more than three expected errors or chimeras were detected [50]. The remaining mean read length was 247 bp (221-279 bp) putting the effective error rate at less than 0.012 error per base after quality control. The median sequencing depth of the 196 adult samples with more than 2,500 reads passing quality control was 38,454 total reads (mean: 55,321) making these samples deeply sequenced. Operational Taxonomic Unit (OTU) clustering by USEARCH [50] with 97% read DNA sequence identity resulted in 2,520 OTUs with measured differences between taxonomy members in excess of 7.41 nucleotide bases ((1-0.97) x 247bp) with less than three of these being sequencing errors. These OTUs classified to 286 bacterial reference taxonomies with the Ribosomal Database

Project (RDP) Classifier java program [20]. Sample bacterial DNA sequences were mapped onto the 2,520 *de-novo* OTU reference taxonomies with USEARCH [50]. Reads that did not match an OTU were assigned to the category OTU_NA and treated as the other OTUs.

Non-parametric Buccal Bacterial Analysis by Azinphos-methyl Blood Detection Group: The microbiome data from buccal samples in spring-summer 2005 (n = 101) and winter 2006 (n = 82) were categorized to factor groups by blood detection of Azinphos-methyl. Twenty-nine spring-summer individuals and 26 winter individuals who were sequenced had concurrent winter or previous spring-summer blood Azinphos-methyl detection by mass-spectrometry as evidence of exposure. Each buccal sample's genus proportion was transformed by the centered-log-ratio, and the common genera (n=22 genera, by spring-summer samples mean genera proportion > 0.005) were tested by the Wilcoxon's Rank Sum test [51, 52] between the sample groups in which Azinphos-methyl (AZM) was detected and not detected (ND). Table 2-1 presents the false discovery rate [62] adjusted significance, the estimated non-parametric location difference, and the unadjusted 95% confidence interval of each genus's 'centered-log-ratio' abundance between the AZM and ND sample groups for both seasons.

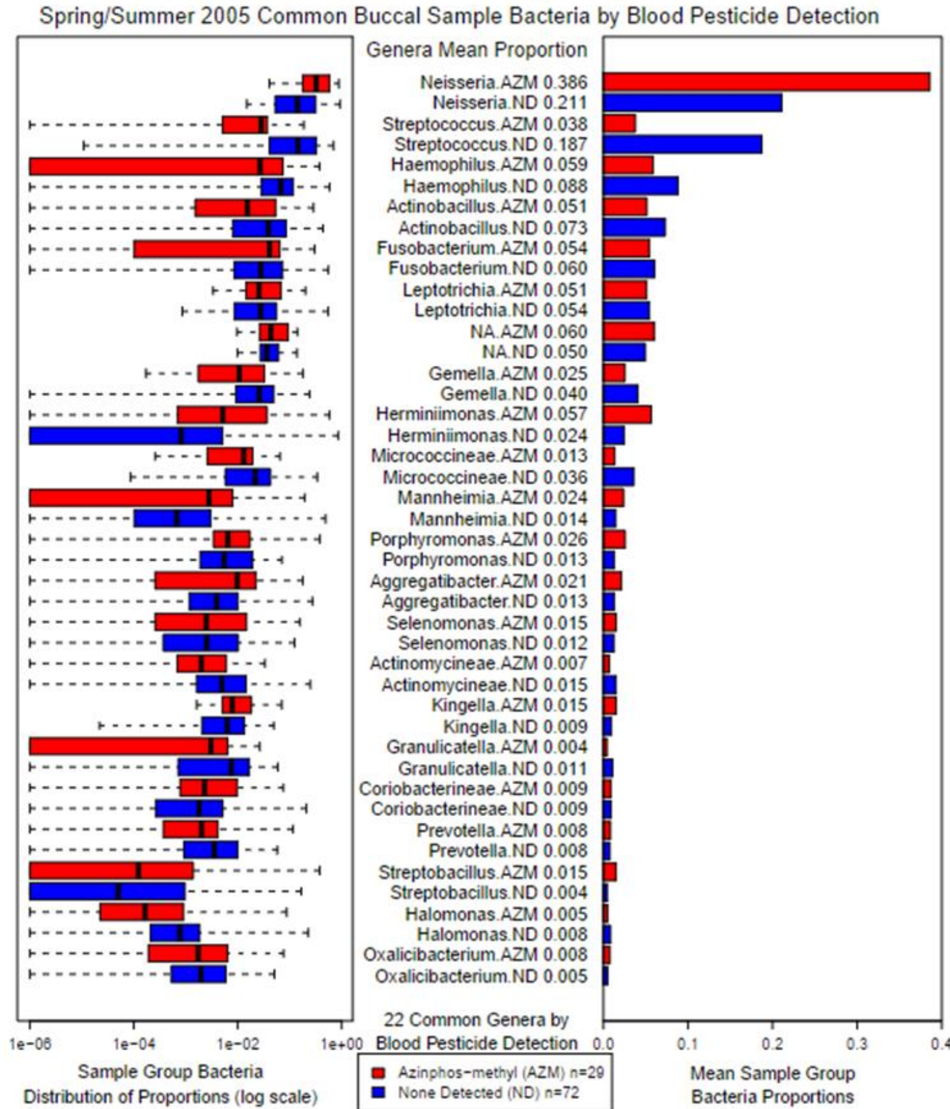
Table 2-1 Wilcoxon's Rank Sum Test of Azinphos-methyl Blood Detection Groups for Adult Buccal Microbiome Compositional Perturbation

Wilcoxon's Rank Sum Test of Azinphos-methyl Blood Detection Groups for Compositional Pertubation, Significance Results of Common (>0.5%) Genera

	Spring-Summer 2005					Winter 2006				
Genera	Location Difference	lower	upper	p-value	FDR	Location Difference	lower	upper	p-value	FDR
Streptococcus	-1.98	-2.74	-1.08	0.0001	0.001	-1.19	-2.11	-0.30	0.011	0.14
Micrococcineae	-1.06	-1.74	-0.43	0.0018	0.02	-0.78	-1.59	-0.09	0.030	0.22
Gemella	-0.99	-1.96	-0.25	0.009	0.05	-0.65	-1.58	0.33	0.171	0.4
Haemophilus	-1.05	-1.92	-0.24	0.008	0.05	-0.88	-1.97	0.07	0.09	0.3
Halomonas	-0.80	-1.49	-0.07	0.015	0.05	-1.04	-1.83	-0.19	0.013	0.14
Actinomycineae	-1.00	-1.72	-0.20	0.013	0.05	-0.69	-1.49	0.10	0.09	0.3
Granulicatella	-0.89	-1.64	0.00	0.026	0.08	-0.23	-0.86	0.38	0.5	0.8
Mannheimia	0.79	0.00	1.55	0.042	0.11	0.55	-0.25	1.31	0.3	0.5
Herminiimonas	1.01	-0.04	1.62	0.081	0.19	0.14	-0.60	0.87	0.6	0.8
Actinobacillus	-0.98	-2.09	0.13	0.095	0.19	-0.67	-1.70	0.24	0.1	0.4
Neisseria	0.45	-0.08	1.07	0.089	0.19	0.46	-0.17	1.11	0.1	0.4
Prevotella	-0.68	-1.47	0.13	0.11	0.19	-0.49	-1.25	0.32	0.2	0.4
Kingella	0.35	-0.24	0.98	0.2	0.3	0.60	-0.07	1.30	0.1	0.3
Fusobacterium	-0.27	-1.16	0.30	0.4	0.6	-0.17	-1.26	0.60	0.6	0.8
Oxalicibacterium	-0.12	-0.90	0.64	0.6	0.8	0.13	-0.83	1.37	0.7	0.8
Coriobacterineae	0.13	-0.62	0.94	0.6	0.8	0.12	-0.88	1.03	0.8	0.8
Selenomonas	-0.22	-1.11	0.81	0.6	0.8	-0.23	-1.27	1.02	0.7	0.8
Aggregatibacter	0.12	-0.60	1.13	0.6	0.8	-0.01	-1.06	0.73	0.8	0.8
Porphyromonas	-0.19	-0.90	0.62	0.5	0.8	-0.18	-0.97	0.60	0.7	0.8
otu_NA	-0.05	-0.68	0.43	0.8	0.9	0.18	-0.26	0.62	0.4	0.7
Streptobacillus	0.00	-0.25	0.56	1.0	1.0	0.00	0.00	0.47	0.7	0.8
Leptotrichia	0.00	-0.59	0.59	1.0	1.0	0.42	-0.16	1.06	0.2	0.4

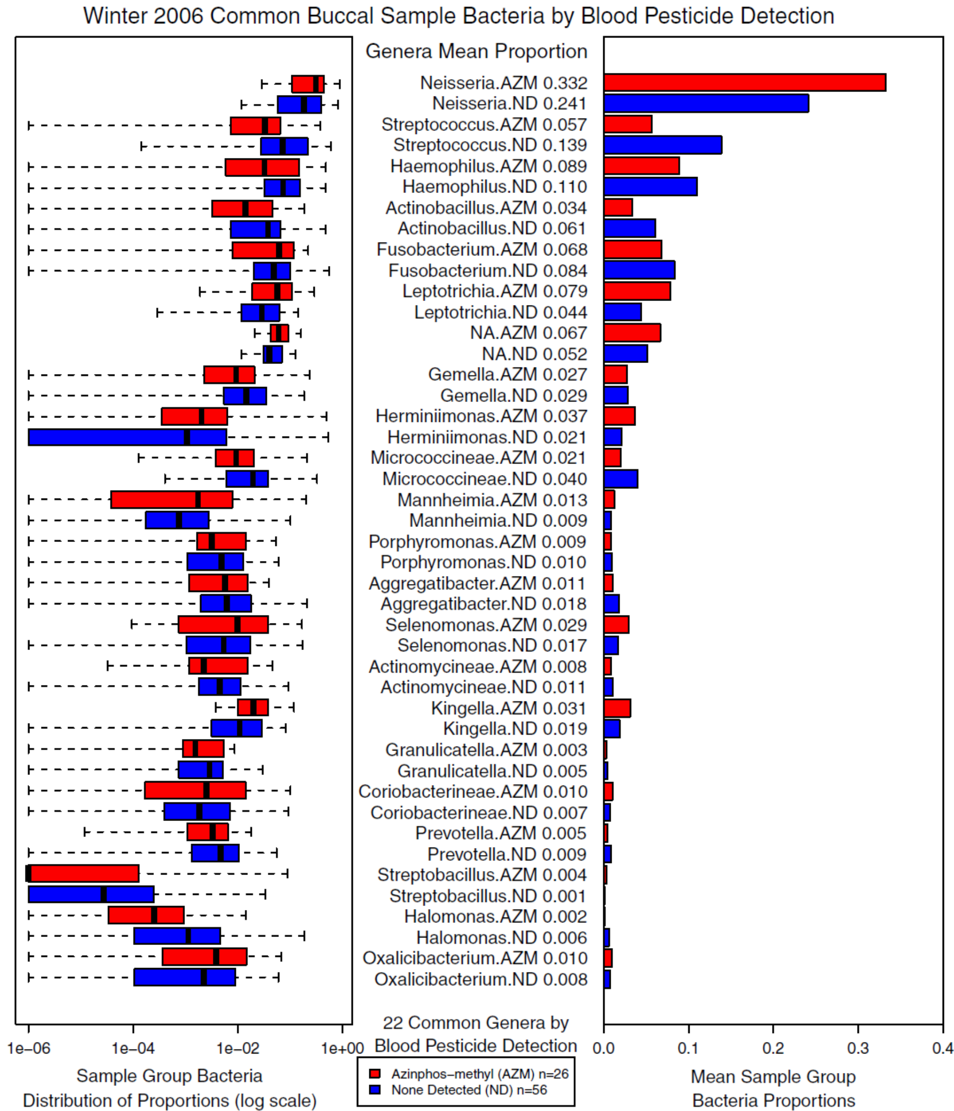
Seven taxa (*Streptococcus, Micrococcineae, Gemella, Haemophilus, Halomonas, Actinomycineae,* and *Granulicatella*) were significantly reduced in the Azinphos-methyl group at a false discovery rate (FDR) <0.1 in the spring-summer 2005 data (shaded dark grey in Table 2-1). Two of these (*Streptococcus* and *Halomonas* [shaded light grey in Table 2-1]) maintained a suggestive negative association significance (p-values ~0.01, FDR ~ 0.14) in the following winter. We observed winter values with the same direction but lower magnitude than the spring-summer collection for Wilcoxon's location difference between groups showing reduced abundance in those with detected Azinphos-methyl. This is not surprising as of the 26 individuals in the winter Azinphos-methyl detection group, only one (the non-farmworker) had concurrent winter exposure measured in blood. This suggests a relaxation but persistence of the exposure effect into the winter. *Mannheimia, Herminiimonas, Actinobacillus, Neisseria* and *Prevotella* appear to be suggestively perturbed in the Azinphos-methyl group. *Actinobacillus* and *Prevotella* were suggestively (0.1 < FDR < 0.2) perturbed taxa that decreased in the Azinphos-methyl detection group whereas the other three are increased in relative abundance. In Figure 2-1 (Spring-summer 2005) and Figure 2-2 (Winter 2006) each tested common taxon's mean proportional abundance in buccal samples by blood Azinphos-methyl detection group is presented.

Figure 2-1 **Spring/Summary 2005 Common Buccal Sample Bacteria by Blood Pesticide Detection**



Spring-summer 2005 common (>0.5% sample mean) bacterial genera from buccal swab sequencing organized by spring-summer Azinphos-methyl blood mass spectrometry detection into exposure groups greater than and less than the limit of detection (0.04 ng/gram of plasma). Taxa are ranked by the mean proportion of reads. Boxplots in the left panel are in log scale (with zero genera proportion samples set to 1e$^{-6)}$; black bars, boxes and whisker bars show the median, the central quartiles and the extreme values, respectively. The right panel and the central label text report the group mean proportion to show the relative abundances of various genera in relation to each other.

Figure 2-2 **Winter 2006 Common Buccal Sample Bacteria by Blood Pesticide Detection**



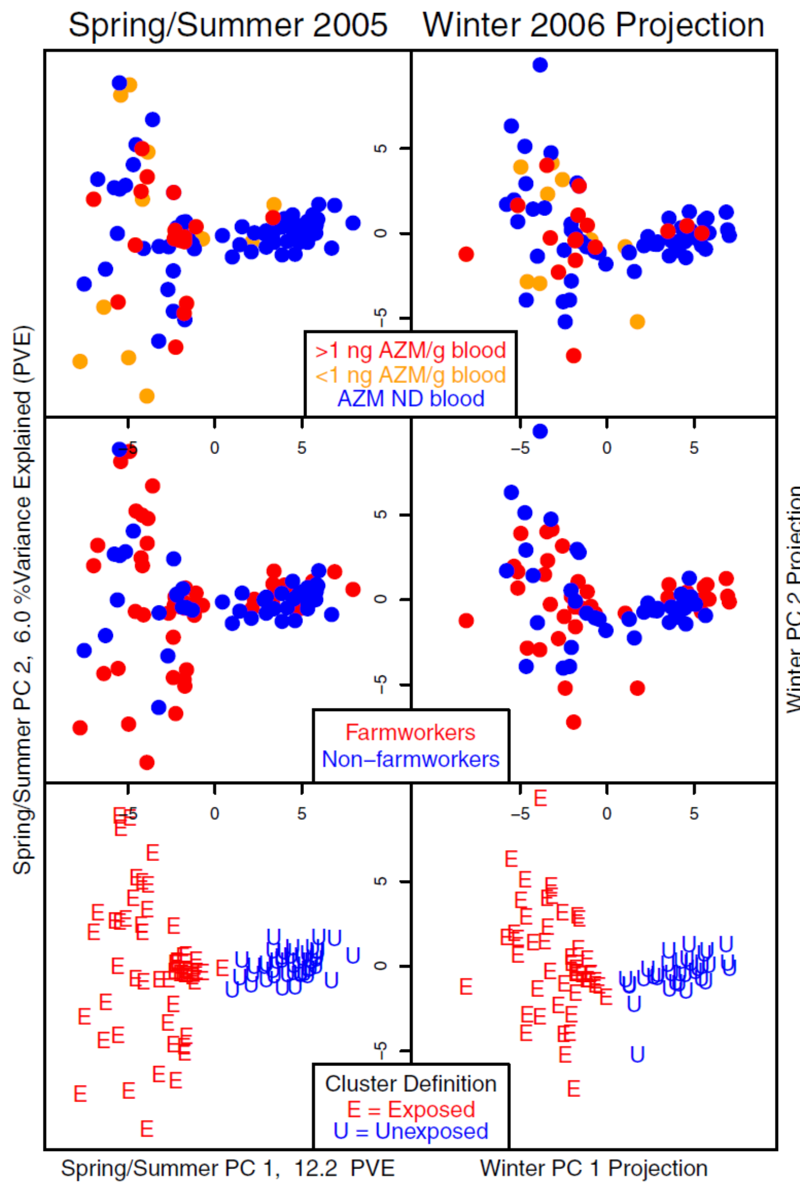
Winter 2006 Common (>0.5% sample mean) Bacterial Genera from Buccal swab sequencing organized by Azinphos-methyl blood mass spectrometry detection into exposure groups greater than and less than the limit of detection (0.04 ng/gram of plasma). Taxa are ranked by the mean proportion of reads from the spring-summer collection to facilitate graphical comparisons between seasons. Boxplots in the left panel are in log scale (with zero genera proportion samples set to 1e$^{-6)}$; black bars, boxes and whisker bars show the median, the central quartiles and the extreme values, respectively. The right panel and the central label text report the group mean proportion to show the relative abundances of various genera in relation to each other

Samples sequenced at very high depth have zero proportions of some genera. This can be seen in the whisker bars for many genera which extend to the 10^{-6} extreme value. The lowest common population genera proportions measured that were non-zero in individual samples were ~10^{-5}, two orders of magnitude below the mean proportion threshold and one order of magnitude above the zero-order of the log scale. Use of the deep depth of sequencing, common population sample mean proportion thresholds and the non-parametric Wilcoxon's

Rank Sum test of the proportions 'centered-log-ratios' mitigate the zero-inflation statistical effects that can occur in count based proportional data at varying depths sampled. As an example, *Streptococcus* is absent in three Azinphos-methyl detected spring-summer collection individuals (H161A, H153A, H125A) with sequenced samples at the depths (2,695; 3,884; 5,788 reads, *Pr[missed at* 0.005*]*: 1.4×10^{-6}, 3.7×10^{-9}, 2.7×10^{-13}, respectively), whereas *Streptococcus* is detected in all individuals from the group with Azinphos-methyl not detected. Two samples (H160A and H149A) from the winter collection with Azinphos-methyl detected also had no *Streptococcus* detected in their buccal microbiota data. The depth of sequencing for these samples (84,158 and 30,611 reads respectively) gave very low probabilities ($\sim 10^{-183}$ and 10^{-67}, respectively) at the mean proportion threshold (0.005) that we missed *Streptococcus* in these individuals, suggesting the true proportion is indeed very low in many of these exposed individuals. This analysis suggests that taxa proportions that are very low in some individuals and proportionally different between the groups are likely due to a true difference in the composition of the microbiota community structure. *Streptococcus* is a taxonomically large common oral microbiota member. Some members of *Streptococcus* have been shown to be beneficial community members to oral health (i.e., *S. salivarius, S. oligofermentans, and S. mitis*) while others have been implicated in oral disease (i.e., *S. mutan, S. pneumonia, S. sanguinis, S. sobrinus, S. gordonii, S. agalactiae, S. tigurinus, and S. anginosus*) [63-75]. This split in the beneficence profile of *Streptococcus* makes a diversity of disease and positive health potentials for the observed effect when the compositionally large member *Streptococcus* is perturbed. Similarly with *Haemophilus* [76-84] and other genera identified here, which are known to be common to the oral context [85], can have a diversity of disease and health profiles within any taxonomic level and can be contextually associated with different co-infections, plasmids, species and strains [86]. This makes drawing specific conclusions to the perturbed community member contributions to health difficult without specific information on health outcome in this agricultural context and additional investigations.

PCA Exploration of the Oral Buccal Microbiota Relative Abundances: The buccal samples in spring-summer 2005 and winter 2006 microbiome data revealed two primary clusters of oral microbiome types in each season. A two-dimensional PCA projection of sample OTU proportion 'centered-log-ratio' of OTUs detected >0.001 mean proportional abundance is plotted in Figure 2-3 for both seasons with point colors representing Azinphos-methyl detection in blood, farmworker occupational status and the Gaussian model based PC1 cluster calls.

Figure 2-3 PCA Exploration of the Oral Buccal Microbiota Relative Abundances



Oral buccal microbiome differences between individuals in bacteria projected by PC 1 and PC2 of the 'centered-log-ratio' transformed proportions of common OTUs (>0.001 mean proportion of spring-summer sample reads). The top panels show mass spectrometry results of Azinphos-methyl blood detection as colored points (red: >1 ng/g, orange: <1 ng/g, blue: none detected) overlaid at the PC coordinates of the buccal microbiome composition as the measure of exposure. The limit of detection is 0.04 ng/gram of plasma. The middle panels show the farmworker occupational status of individuals. The bottom panels depict the PC1 Gaussian model-based cluster calls of the data projected in the PC coordinates. Two groups are observed in both seasons, the left cluster group is enriched for farmworkers and Azinphos-methyl detection in blood samples. We interpret the left cluster as "Exposed" based on the co-clustering of individuals with blood detection of Azinphos-methyl and the enrichment of farmworkers. The OTU percent of variance explained (~18% PVE) by PC1 and PC2 is reported in the axis labels. The right winter panels are supplementary out of sample projections through the spring/summer PC1 and PC2 loadings.

We observed two clusters evident in both seasons with a tighter clustering in the spring-summer collection. The less aggregated left cluster is visibly enriched for those individuals with detected Azinphos-methyl in the blood (Figure 2-3, orange and red points, top panels) and farmworker status (middle panels). A similar but less aggregated clustering pattern is also seen in the following winter 2006 buccal microbiomes (right panels), suggesting a persistence to the microbiome state induced by previous exposure. Azinphos-methyl is only detected in one non-farmworker in the winter. If Azinphos-methyl was detected in either season the color in both seasons plot panels reflects this to track migration of pesticide exposed subjects' microbiome types while those same individuals are Azinphos-methyl non-detected in the winter. t tests of the PC1 scores for spring/summer and winter showed that the microbiome composition PC1 score was significant for Azinphos-methyl detection in the left direction of PC1. The spring-summer mean Azinphos-methyl group PC1 score was -2.8, versus 1.1 in the non-detect samples. In the winter projection, the Azinphos-methyl group PC1 mean score was -1.6, while the non-detected samples had a PC1 mean score of 0.73. The spring/summer P value of $4x10^{-7}$ and winter P-value of $1.7x10^{-4}$ show a persistent significant association between the buccal microbiome composition PC1 score and Azinphos-methyl detection. The data points for the samples with lower blood concentration (<1 ng Azinphos-methyl/g blood, the median) are colored orange to show that there is no simple evidence of a dose response within the exposed individuals. t test P-values of >0.49 result from the comparison of PC1 values for high (red) and low (orange) Azinphos-methyl concentrations in spring/summer and winter. The middle two panels of Figure 2-3 show the farmworker-versus-nonfarmworker distribution in the OTU microbiota PCA. In the spring-summer, the t test of the PC1 scores for farmworker occupational status also showed that farmworkers had significantly lower mean microbiome PC1 scores (-1.01) than nonfarmworkers (1.16) (Welch's t test, p-value ~0.01). This farmworker occupational PC1 microbiome exposure was not persistent in the winter (p-value ~0.45). The bottom two panels present the mClust cluster call assignments using the spring-summer 2005 PC1 as input to Gaussian model based cluster assignments [61]. To train the model, the spring-summer 2005 PC1 data was used to determine a decision value of 0.72 based on the midpoint between the maximum PC1 value in the spring-summer left cluster and the minimum PC1 value in the right cluster. This midpoint (0.72) between clusters was then used to independently classify the winter 2006 buccal samples to the respective clusters. This allowed us to control for overfitting when comparing PC1 cluster memberships between seasons. The left spring-summer and winter clusters are assumed as the "exposed" based on the co-clustering of Azinphos-methyl blood detected samples. Categorical analysis of occupation status and cluster membership shows the spring-summer 2005 buccal microbiomes' left exposed PC cluster contains 57 individuals (38 farmworkers and 19 non-farmworkers) and the right more tightly grouped cluster contained 44 individuals (16 farmworkers and 28 non-farmworkers). The difference of cluster membership was significant by the Fisher's Exact Test for occupational farmworkers status (p-value ~ 0.003, Odds Ratio ~ 3.5, 95% CI: 1.4 – 8.7). When the 29 farmworkers and one non-farmworker with both buccal microbiome data and detected blood concentrations of Azinphos-methyl were compared with spring-summer buccal microbiome cluster membership, the left cluster enriched with farmworkers contained nearly all (27/30) of the spring-summer microbiome data individuals with Azinphos-methyl detected (including the one non-farmworker with detection in the winter) and 30 individuals where none was detected. The right cluster contained almost exclusively individuals with non-detects of

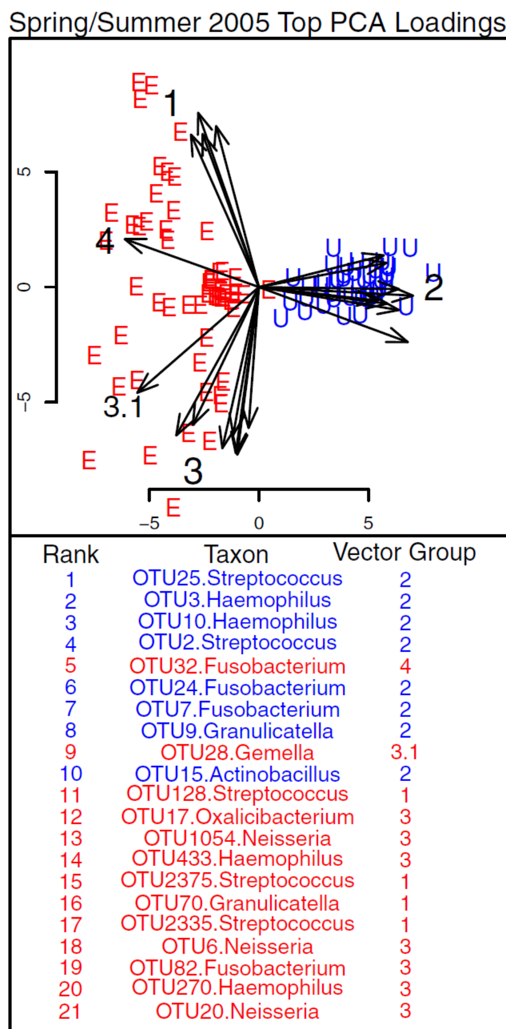
Azinphos-methyl (41/44) except three farmworkers with Azinphos-methyl detected in blood. Those with detected levels of Azinphos-methyl were significantly enriched in one of the two cluster from the spring-summer 2005 buccal collection (p-value ~ 6.5×10^{-6}, Odds Ratio [OR] ~ 12.0, 95% CI: 3.2 – 67). We will refer to the left cluster enriched with farmworkers and containing the majority of individuals with Azinphos-methyl detection as the "exposed" cluster and the right cluster enriched with non-farmworkers the "unexposed" cluster. In the winter 2006, 48 individuals (28 farmworkers and 20 non-farmworkers) cluster in the left exposed PC cluster and 34 individuals (14 farmworkers and 20 non-farmworkers) in the right unexposed cluster (Fisher's Exact Test, p-value ~ 0.18, OR ~ 1.98, 95% CI: 0.74 – 5.4). Comparison with Azinphos-methyl blood detection in any season to the winter 2006 microbiome PC clusters showed Azinphos-methyl detection in blood was still significantly enriched in the exposed cluster (Fisher's Exact Test, p-value ~ 0.008, OR ~4.4, 95% CI: 1.4 – 17.2), with 21 individuals having Azinphos-methyl detected and 27 with none detected as compared to the unexposed cluster containing five individuals with Azinphos-methyl detected and 29 with none detected.

Many individuals with Azinphos-methyl not detected in the blood (spring-summer: 30/57, winter: 27/48) cluster with individuals where it was detected. This would be expected as Azinphos-methyl is metabolized and excreted with a half-life clearance of ~30 hours [87], making the temporal window in which one must capture an exposure important. We analyzed the persistence of cluster membership using the 70 individuals with both spring-summer 2005 and winter 2006 microbiome sequence data. The majority (94%, 66/70) of individuals sampled remained in their original cluster the following winter. The one nonfarmworker with Azinphos-methyl detected in the winter of 2006 remained as an exposed cluster member in both seasons. This suggests a persistence to the exposed microbiome state as detected by model-based PC1 clustering state. The two individuals who migrated from the exposed to the unexposed microbiome cluster were a farmworker and a nonfarmworker. The farmworker who changed from the exposed to the unexposed cluster had Azinphos-methyl detected in blood in the spring/summer of 2005 but none detected in the winter. Among just the exposed spring/summer cluster individuals, the composition effect of the spring-summer Azinphos-methyl-exposed cluster as measured by PC cluster membership at the later winter collection date had 95% (39/41) remaining in the exposed cluster for both seasons. This cluster membership concordance in combination with the extremely low bacterial taxa abundances measured for some exposed individuals suggests that these microbiomes have undergone extinctions of the affected microbiota with minor regrowth. The consequence of bacterial extinction with little regrowth has been shown in mice when gut re-colonization is not facilitated by microbiome seeding [40]. Two nonfarmworkers with no Azinphos-methyl detected switched from unexposed to exposed microbiome PC clusters, suggesting exposure events since the last spring-summer's pesticide applications which we did not ascertain. We speculate more of the individuals from the unexposed cluster would migrate to the exposed cluster if sample collections had continued to include a following spring-summer 2006 agricultural cycle, when pesticide spraying begins again. The non-farmworkers' microbiomes that cluster with individuals with detected Azinphos-methyl blood concentrations suggests other unobserved exposures have occurred throughout the spring-summer, fall and into winter as observed in the one non-farmworker with detected Azinphos-methyl in the winter 2006 collection. We suspect stochastic differences in exposure in the

home and community environment, as well as differential geographic location, Azinphos-methyl dose and clearance duration may explain the heterogeneous clustering of non-farmworkers with farmworkers into the respective clusters.

Abundances of Commonly Detected OTUs by PC Loadings and Clusters: The top ranked 21 PC loadings of bacterial OTU for spring-summer are presented in Figure 2-4.

Figure 2-4 Abundances of Commonly Detected OTUs by PC Loadings and Clusters



Rank	Taxon	Vector Group
1	OTU25.Streptococcus	2
2	OTU3.Haemophilus	2
3	OTU10.Haemophilus	2
4	OTU2.Streptococcus	2
5	OTU32.Fusobacterium	4
6	OTU24.Fusobacterium	2
7	OTU7.Fusobacterium	2
8	OTU9.Granulicatella	2
9	OTU28.Gemella	3.1
10	OTU15.Actinobacillus	2
11	OTU128.Streptococcus	1
12	OTU17.Oxalicibacterium	3
13	OTU1054.Neisseria	3
14	OTU433.Haemophilus	3
15	OTU2375.Streptococcus	1
16	OTU70.Granulicatella	1
17	OTU2335.Streptococcus	1
18	OTU6.Neisseria	3
19	OTU82.Fusobacterium	3
20	OTU270.Haemophilus	3
21	OTU20.Neisseria	3

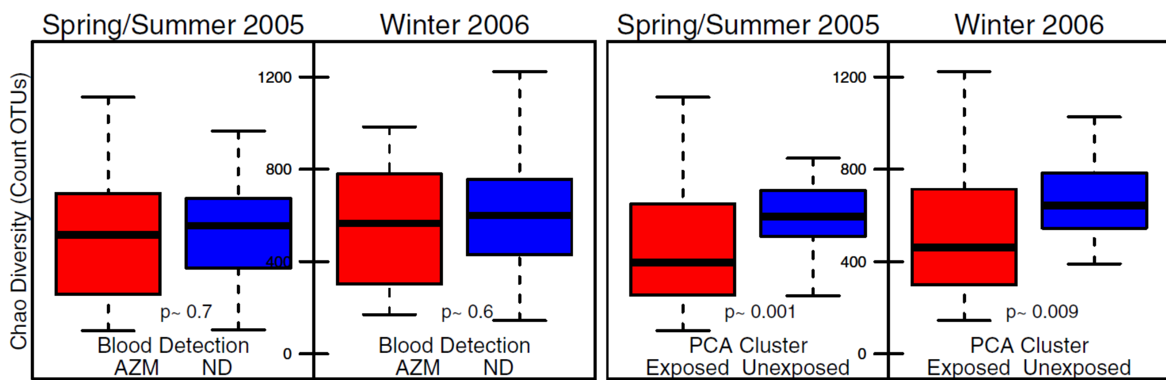
The ranked PC 1 and PC2 vector loadings of the top 21 OTUs in Spring-summer 2005 buccal samples are colored based on cluster (Red = Exposed, Blue = Unexposed) and with number labelled vector groups (1-8) based on which graph direction the OTUs are the highest in abundance.

In the spring-summer collection there are two primary vector groups labelled as 1 and 3 among exposed individuals; representative taxa present in these individuals are depicted in Figure 2-4. Vector group 2 includes taxa that are increased in individuals where we did not detect Azinphos-methyl in the blood. The high loading rank OTUs in the spring-summer 2005 data are representatives of several genera also detected in the Azinphos-methyl blood detection group non-parametric analysis of the common genus composition. These include the genera *Streptococcus, Gemella,* and *Haemophilus,* and suborder *Micrococcineae*. The *Streptococcus* OTUs were still present but at lower abundance in the exposed cluster appear more in the top portion of the plot vector groups 1 (spring-summer). *Streptococcus* is the most increased in samples in the direction of vector groups 2 (spring-summer) to the unexposed cluster. *Neisseria* is a common driver in the left exposed cluster direction of PC1. While not associated by the non-parametric analyses, it is possible that *Neisseria* is less susceptible to the exposure and as the most common microbiota detected, is among the majority of the reads left when other taxa are not present in the compositional relative abundance. This leaves *Neisseria* as even more common in the exposed individuals. *Neisseria* is a large genera with many diverse members, many of which do not normally cause disease [88]. The PC cluster discerning taxa is split as to which *Fusobacterium* OTUs are indicator taxa for the respective exposure clusters, reflective of the Wilcoxon's test where it is not significant (See Table 2-1,

Figures 2-1 & 2-2). OTUs in loading vector groups 1 and 3 are pointed in opposite directions of PC2 and not at the central mass in the left direction of PC1. These heterogeneous patterns between cluster directions in OTU genera may be explained in part by complex relationships in the composition. The perturbed taxa have lower abundance but are still present in some exposed individual subjects, as seen in *Streptococcus* in the nonparametric test. This PC2 heterogeneity may be interpreted as a less tractable community ecology and homeostatic dysbiotic state that would occur when a normal apex community member like *Streptococcus* is depleted. This community perturbation would have effects on many members of the community that are reliant on amino acids from syntrophic metabolic niches [89] generated by *Streptococcus*. The larger heterogeneity in the exposed PC2 scores may also be reflective of this general ecological dysbiosis of missing community members. The winter unexposed cluster also shows this less aggregated "dysbiotic" pattern, where individuals seem more heterogeneous. Dysbiosis has been observed in other microbiome organophosphate pesticide exposure studies of the gut using an *in-vitro* bioreactor as well as *in-vivo* in the rat and Japanese quail [90-94]. Differential susceptibility of OTUs representative of species-strains to this exposure is also suspected to contribute to heterogeneity between OTU and genera level comparisons, but would require additional investigation in another experimental context.

Alpha Diversity: Using both the Azinphos-methyl detection group subject identities and the cluster identity calls from PC analysis we investigated the alpha species richness diversity of the samples using the method of Chao [95]. The Chao diversity metric estimates the total number of organisms in a sample (count of OTUs) given an adjustment based on the number of rare members observed. Figure 2-5 presents boxplots of the seasons by Azinphos-methyl detection groups and exposure clusters definitions for Chao Alpha diversity values.

Figure 2-5 Adult Alpha Diversity by Exposure Groups



The Chao diversity metrics of are presented for exposure categories. The right panels depict each seasons' blood burden Azinphos-methyl (AZM, red) and none detection (ND, blue) groups' diversity. The left panels depict each seasons' PCA cluster group diversity labeled as "Exposed" (red) and "Unexposed" (blue). The black bar is the median with the boxes as the central quartiles. Whisker bars are the extreme values. The Chao Alpha diversity metric is an estimate of the total number of organisms given the number of low frequency OTUs detected. Significance was determined by Welch's Two Sample t-test.

There is no significant difference in the OTU richness or diversity in the Azinphos-methyl blood detected individuals in either season. However, within both the spring-summer and

winter sample collections the exposed PC cluster had significantly reduced counts of organisms as compared to the unexposed cluster as seen in the Chao diversity (Welch's t-test, p-values < 0.009).

2.5. Discussion

We identify both a direct association and a principle components clustering pattern of individuals indicating the compositional diversity profile of the oral buccal microbiome is associated with environmental exposure consistent with an agricultural signal. This agricultural signal was measured in members of the Yakima Valley Hispanic community in the Washington State by blood concentrations of the organophosphate insecticide Azinphos-methyl and self-reported apple and pear orchard farmworker status. These exposed individuals have reduced abundance of several common oral indicator genera in this molecular epidemiology analysis of the buccal microbiota profile of agricultural pome fruit orchard workers and community members. The spring-summer 2005 sample collection timing targeted occupational farmworker and community exposures at times when seasonal agricultural applications of Azinphos-methyl insecticide were known to occur. The microbiome PC cluster group membership of non-farmworkers with farmworkers exposed to Azinphos-methyl suggests that many individuals in the community may have unobserved exposures to pesticides which we did not capture in blood. We observe one such exposure in a non-farmworker from the winter 2006 collection season with a blood Azinphos-methyl detection when exposure is not likely based on the seasonal nature of pesticide applications. This underlines the need for datasets with fine temporal resolutions to discover modes of exposure for community pesticide exposures. Previously published household dust data from this same cohort suggests that Azinphos-methyl exposures may occur via take-home pathways in cars and houses, where dry stale air and low UV light conditions can allow environmental persistence [14].

Only three non-farmworker households have self-reported resident smokers, making smoking unable to explain the scale of exposure difference associated here with the microbiome composition. Soil bacteria have been known to degrade organophosphates in since the early 1970s [96-100], and have been well studied with the intent of developing a bioengineered solution to environmental remediation and industrial uses [101]. The earliest (2013) report that OPs like Azinphos-methyl may cause dysbiotic changes in the composition of a host microbiome were conducted with chlorpyrifos in an intestinal *in-vitro* simulation reactor model and *in-vivo* in the rat [91]. Azinphos-methyl, chlorpyrifos and other OPs are closely related compounds with the same covalent mechanism of primary toxicity, cholinesterase inactivation. The same group showed that chronic chlorpyriphos exposure changes the gut permeability and permits translocation of bacteria across the rat intestine [93] and quickly followed this *in-vivo* rat model with a perinatal exposure study where they showed dysbiotic microbiota and exposure affected the intestinal development of the rat pups [92]. Two other very recent articles show that other OPs cause changes in the gut microbiome. In one the OP, trichlorfon induced intestinal dysbiosis and reductions in many taxa at low dose in the Japanese quail (*Coturnix japonica*) [90]. In the other study, low level dosing of mice with the organophosphate diazinon induced gut microbiota and metabolomics changes in both sexes [94]. These orthogonal studies provide evidence to support our finding that human occupational and community exposure changes the composition of the human buccal

microbiome. These studies also found significant reductions in the abundances of *Lactobacillus*, the name defining genera member of the order *Lactobacillales*, of which *Streptococcus* is also a common phylogeny member. We also detect another common genera member of the order *Lactobacillales, Granulicatella* to be significantly decreased in abundance (Table 2-1, Spring-summer, -0.89, 95% CI difference: -1.64-0, p-value ~0.026, FDR ~0.08) and an indicator OTU in cluster analysis (See Figure 2-4). This suggests that the order *Lactobacillales* may in general be susceptible to organophosphate pesticide exposure. The *in-vivo* rat studies reported recently [91-93] also showed increased intestinal permeability, reductions in tight-junction gene expression, as well as increased translocation of bacteria to distal organs in chlorpyrifos dosed animals. The *in-vitro* digester microbiome gut model did not contain a membrane equivalent of the intestine or an immune system and showed differences in bacterial composition of human feces used to inoculate the culture system upon organophosphate exposure to chlorpyrifos [91]. These orthogonal findings suggest that at least two independent mechanisms of microbiome perturbation exist, host membrane permeability and direct microbial action. Additionally, the differences in gut metabolomics analysis shown by recent findings with the organophosphate diazinon dosing in rats indicate that there are changes in gut derived neurotransmitters and amino-acid metabolism products that would have the potential to affect both host behavior phenotypes (i.e., developmental autism in exposed fetuses) and the niches of other gut bacteria present [94]. Future functional validations of antibiotic and other etiological properties, as well as health outcomes by *in-vitro* and *in-vivo* methods could distinguish between the membrane permeability, immune, and inflammatory response and the potential for direct antimicrobial properties that could have produced these perturbations in the human buccal microbiome composition. The systemic bacterial ecology and interaction with the human immune system and tight junction tissue sites may contribute to the holistic compositional effect in the buccal context. In any of these cases, agricultural pesticide exposure is associated here with significant community composition changes in these Hispanic peoples' buccal microbiomes. The community ecology of organisms significantly dependent on *Streptococcus* will likely be affected as the second most common oral member detected with significant reductions in abundance.

We specifically timed the sampling of the population in the spring-summer and winter seasons to capture the Azinphos-methyl exposure well in the respective seasons to well-inform a molecular epidemiology exposure effect study such as this. The farmworkers were actively working in the recently Azinphos-methyl sprayed orchards thinning the fruit and pruning in the April–July, 2005 (spring-summer) sample collection time, as shown by direct blood detection as the measure of exposure. The health impacts of these agricultural exposure associated bacterial community structures and composition in the oral context and at other body sites (i.e., gut, nasal, etc.) have not been investigated in this cohort to date. Hispanics have a disproportionate disease burden for many common diseases related to vasculature (i.e., diabetes and cardiovascular disease) [102-110] that have also been shown to have microbiome related disease variables in many other experimental contexts [111-124]. Dental health, bacterial infection, and peripheral vascular disease have many independent associations [125] and some of the etiology could be affected by systemic loose tight junctions as induced by organophosphates [93]. In many other contexts, environmental pollutants have been shown to

affect the microbiome [126], here we show in humans association between the composition of the microbiome and agricultural pesticide exposure.

2.6. Conclusions

Given the large microbiome differences at the oral buccal site, the Hispanic cohort studied in this paper should continue to be followed for repeated measures of buccal samples with additional sampling, dental history and investigate the pleiotropy to other health outcomes that may have resulted from these agricultural exposures. These future investigations should also perform molecular functional studies on individual bacterial taxa and in community ecology systems biology *in*-vivo contexts for their susceptibility to organophosphates. Additionally, more evaluation of the oral microbiome composition as a simple to collect molecular biomarker of past exposures as a whole for use in epidemiological studies would be able to greatly inform public health and practice.

Authors' Contributions

IBS, JCW and EMF conceived of the project and experiments. IBS, FHG, CSW and SH designed experiments. IBS, CSW, FHG, JT, MK and SH performed experiments. EMV, WCG, SH, BT and EMF curated and collected samples. IBS wrote computer code and implemented analysis. IBS, JCW, TW, EMV, WCG, AS and EMF managed and analyzed data. IBS, EMF, JSM, CSW, AS and JCW wrote the paper. All authors reviewed and commented on the paper.

Acknowledgements

We would like to thank the Yakima Valley Community and Families who participated in this study and the staff members from the Fred Hutchinson Cancer Research Center who performed sample collections. The authors would also like to disclose that we have filed for a record of invention pertaining to the associated effects of Azinphos-methyl on microbiota.

3. ADULT-CHILD CO-CLUSTERING OF BUCCAL MICROBIOME TYPES

Children from Agricultural Households Have Significantly Altered Oral-Buccal Microbiomes

3.1. Abstract

The microbiome has been shown to be vulnerable to environmental exposure and be altered in association to many common diseases. This investigation shows that the composition of the oral buccal microbiome is altered in association with organophosphate pesticide blood detection and demographic measures of agricultural pesticide exposure in this retrospective cohort analysis of children's health. Azinphos-methyl exposure was quantified in the household adults by mass-spectrometry of the blood. Oral buccal swab 16S rRNA sequencing determined taxonomic microbiota proportional composition from concurrent seasonal samples. These samples included pome fruit farmworkers (n=65) and other local Hispanic community non-farmworkers (n=52) with 94 paired children from the same respective households in two sample collections, spring-summer 2005 and winter 2006 for a Yakima Valley Hispanic community agricultural cohort. Machine learning and false discovery adjusted non-parametric exposure group comparison methods were employed in hypothesis testing. We observed by mass spectrometry the organophosphate pesticide Azinphos-methyl in ~30 adult farmworkers in the spring-summer blood and one non-farmworker in the winter sampling. Principal component analysis of the microbiome data identified two primary clusters with association of PC1 to Azinphos-methyl blood detection and farmworker status. The children's buccal microbiota composition clustered with Azinphos-methyl andwith their household adult paired match in ~95% of the households. We show that agricultural pesticide exposure as measured by Azinphos-methyl pesticide detection in the adult blood and self-reported pome fruit farmworker household status is associated with significant alterations in the buccal oral microbiomes of Yakima Valley agricultural community members and their children.

3.2. Introduction

The microbiome is variable in association to many common phenotypes and diseases [24-28, 31, 34]. The microbiome has also been shown to be variable in response to many different environmental stimuli [32-39]. This puts the microbiome in the middle of a multivariate causal paradigm where microbiome variability may both reflect exposure and disease phenotypes while also contributing to our state of wellbeing.

When the microbiota diversity is disturbed, dysbiotic conditions may occur due to symbiotic relationships between key community members being broken. This dysbiosis can be more profound when common members of a community have large reductions in total numbers. The loss of one bacterial taxa can define dynamics leading to habitat niche loss for other taxa

[89] which compounds indirect population shifting effects from a single taxa community perturbation.

Interactions of the microbiome with both endogenous and exogenous chemical compounds have been recognized for many years. In the practice of pharmacology and toxicology, the metabolic potential of the gut microbiome is recognized as a hazard for specific community profiles with human variability [127]. Antibiotics [32], probiotics [37] and diet [34, 35] are all well understood mechanisms by which we can change our microbiomes. These kinds of alterations of microbiomes can result in child developmental outcomes of increased risk of later life obesity from prenatal exposure to systemic antibacterials [41].

Children have a different susceptibility profile than adults due to the developmental changes that occur during their development. The gut microbiome phenotype is stabilized in composition in most individuals by the age of two [128]. Immune development by commensal microbiota community-host signaling interactions has been shown to be necessary for normal gut and immune phenotypes [129]. Early life has many opportunities to affect the development of the microbiome [130]. Common diseases like asthma are associated with specific alterations in bacterial community structures that likely begin to develop when individuals are young [42]. The lifelong potential for childhood microbiome perturbation to affect the quality of life makes having a significant understanding of the environmental interactions with human microbiota important to public health.

In Washington State a Hispanic pome fruit farmworker and non-farmworker community cohort from Yakima Valley was formed in 2005 to study the potential for exposures to agricultural pesticides among farmworkers, local community members and their respective children. This cohort has an archived sample collection which includes blood and buccal swabs.

Azinphos-methyl (AZM) is an organophosphate pesticide that was used in the Yakima Valley to control the crop damaging coddling moth *(Cydia pomonella)* in apple and pear pome fruit orchards. It is usually applied in the spring-late summer seasons with timed re-entry intervals for orchard farmworkers designed to reduce exposure. Despite this, it is understood that significant agricultural community and farmworker exposures occur. These exposures oftentimes exceed the accepted risk standards of the EPA and thus this pesticide has been removed from registration with legal applications of Azinphos-methyl in the United States ceasing as of September 30, 2013 (http://www.epa.gov/opp00001/reregistration/Azinphos-methyl/phaseout_fs.htm).

Azinphos-methyl, and other similar organophosphates, has an acute mechanism of toxicity by a covalent bond to the serine-200 residue in the active site triad of cholinesterase [43]. This prevents cholinesterase that recycles acetylcholine in the neuronal synapse from clearing the neurotransmitter, resulting in synapse neurotransmitter overstimulation and loss of useful activity. Prenatal exposure to the organophosphate pesticide chlorpyrifos has been associated with a loss of working memory in New York City Dominicans and African Americans [47]. The babies of exposed mothers have been shown to have significantly different structural features measurable by MRI in association with working memory and

intelligence [44-48]. Controlled animal organophosphate pesticide dosing experiments have replicated similar neurological phenotypes related to these findings in model *in-vivo* systems [49]. The activity profile of organophosphates shows that other off target effects at low dose are evident which include oxidative stress [131-133]. It is likely that Azinphos-methyl can interact with other serine residues in similar serine hydrolase protein moieties creating other systemic losses in enzyme activity. This drives our scientific inquiry at assessing the holistic organism phenotype outcomes that may occur from off target toxicological effects on orthogonal body systems. Due to this, we sought to investigate the potential effects of this agricultural pesticide exposure on the simple to collect oral-buccal swab microbiome.

Of the 400-founding cohort individuals we sequenced the buccal swabs of 117 individual household's adults and 94 residence matched children. We used 16S rRNA primers targeting the V5V6 from adjacent conserved regions to identify a buccal microbiota taxonomy compositional phenotype for each buccal sample. Here using unsupervised machine learning techniques, we show that the children's buccal samples co-cluster with their household adult in this agricultural community cohort. This includes principal component co-clustering with those adults who appear to have significantly altered microbiota compositions in association with farmworker status and Azinphos-methyl exposure as detected in the adults' blood. The children's household resident adult occupational status and household adult blood Azinphos-methyl pesticide detection was used to dichotomize and non-parametrically test if the shared household environment of these adult-child pairs would have a similar effect on the children's microbiota. We show here in the reported data that the children of pome fruit farmworker households, and those adult individuals who have Azinphos-methyl detected in their blood, have children who have similar significantly altered oral buccal microbiome compositions with reduced diversity.

3.3. Methods

3.3.1. Participants and Agricultural Setting

In 2005, the participants of our study, the Children's Health Cohort 2 (CHC2), were recruited from 200 Yakima Valley agricultural community households. Recruitment included one child and one adult guardian from each household. The household adult was identified as the consenting guardian. Written and informed consent was obtained from each adult and consent of the guardian for the child. The Fred Hutchinson Cancer Research Center Institutional Review Board approved the study collection procedures. (File IR 5946)

3.3.2. Study Design & Cohort General Characteristics

The CHC2 is a cross-sectional longitudinal two-timepoint cohort study design. The initial well studied aim was to assess potential pesticide exposures of pome fruit orchard workers and their children by surveying 100 farmworker adult-child pairs in comparison to 100 matched control non-farmworker adult-child households [1, 2, 4-6, 9-15]. Timing of the first sample collection phase was synchronized with the local commercial agricultural applications of Azinphos-methyl pesticide to control Coddling Moths *(Cydia pomonella)*. At the same time is the fruit thinning season when pome farmworkers thin the excess fruit and trim branches giving them the potential for exposure to residues in the orchard. This begins

in the Spring and spans to Summer (April to July) 2005. The second sample collection for the same individuals occurred in the winter time spanning December 2005 to early March 2006 when pesticides are not normally used and environmental degradation would have reduced the likelihood of exposure. Whole blood from adult and buccal swabs from adults and children were collected in these two distinct agricultural seasons (1) Spring-summer 2005 and (2) Winter 2006. For the analyses in this manuscript we used the adult blood to assess organophosphate pesticide exposure and the buccal swabs collected at the same time to quantify the microbiota composition of the oral-buccal microbiome in both adults and children. The children had buccal swabs but no blood sample taken to minimize the invasiveness of the home visit. The adult blood Azinphos-methyl mass spectrometry detection results of pesticide exposure biomarkers were used to dichotomize participants into binomial exposure groups. These categorical exposed and unexposed groups were compared to the bacterial microbiota taxonomy composition. Sex, age, adult Pome Farmworker (FW) status, ancestry, and other general demographic variables were self-reported by the adult participants.

3.3.3. Total Adult Blood Azinphos-methyl Pesticide

The blood sample Azinphos-methyl concentrations were determined by mass spectrometry as described by Barr et al. [19],. The limit of detection is 0.04 ng/g plasma.

3.3.4. 16S rRNA Buccal Swab Sequencing

Oral buccal mucosa samples were collected using a Catch-all™ Sample Collection Swab (Epicentre, Madison, WI). Participants were instructed to abstain from eating, smoking or drinking any beverages except water for at least 30 minutes prior to providing a sample. Immediately preceding the sampling, the mouth was rinsed twice with water. Collection swabs were pressed against the cheek inside the oral cavity and moved in a circular motion 20 times on each cheek and placed in a 15 mL polypropylene tube with the tip submerged in 1 mL of RNAlater (Qiagen, Venlo, Netherlands). Samples were stored initially at -10°C in the field office and then transferred to the University of Washington laboratory and stored at -80 °C. For DNA preparation each sample was thawed then diluted with 9 mL of phosphate buffered saline (PBS). The sampling swab was stirred and removed from the tube. Sample suspensions were centrifuged at 400g for 10 minutes, then supernatant was aspirated out of each sample. Sample DNA extraction and elution was completed following the manufacturer's instructions for Maxwell® 16 Buccal Swab LEV DNA Purification Kit (Promega, Fitchburg, Wisconsin). The Quant-iT™ PicoGreen® dsDNA Assay Kit (Thermo Fisher Scientific, Waltham, Massachusetts, USA) was used to quantify the concentration of double stranded DNA in each sample DNA extract according to the manufacturer's specifications. Polymerase Chain Reaction (PCR) and bead DNA cleanup was performed on DNA extracts with a PicoGreen concentration greater than 30 ng/µL which were then diluted to a 2 ng/µL concentration for PCR DNA loading. Each PCR was loaded with 10 ng of sample DNA extract template. Eurofins (Eurofins Scientific, Luxembourg) primers were diluted to 18 µM concentration. Primers and the EmeraldAmp GT PCR Master Mix (Clontech Laboratories, Inc., Mountain View, CA) were added to each PCR tube along with samples for a final primer concentration of 0.9 µM with nuclease free water. Negative control molecular grade water blanks were thermocycled with samples to control for bacterial

contamination of reagents. Primer sequences were designed targeting the 16S rRNA gene variable region 5 (V5, ATTAGATACCCNGGTAG) and variable region 6 (V6, CGACAGCCATGCANCACCT) of the 16S ribosomal subunit DNA gene as described in Cai et al [56]. Primers were designed to be bidirectional with four oligonucleotides (including the reverse complements of the above sequences) that also contained both the Ion™ sequencing adaptors and 96 unique DNA barcodes. PCRs were performed with a MJ Research PTC 200 Peltier Thermal Cycler (MJ Research, St. Bruno, Quebec, Canada) using the temperature cycle protocol also presented in Chapter 2. The Agencourt AMPure XP PCR purification system (Beckman Coulter, Brea, CA) was used to purify the PCR amplicons. Sample PCR DNA amplicons were bioanalyzed for fragment length and quantified using the Agilent High Sensitivity DNA Kit (Agilent Technologies Inc., Santa Clara, CA) and diluted in Low TE buffer to yield the C2 of ~26pM. Sample PCR amplicons with unique barcodes were pooled in equimolar-volume ratios in groups of 16-30 for sequencing. The Ion Torrent manufacturer's protocols for PCR amplicon template preparation was performed with the Ion PGM Template OT2 400 kit (Life Technologies, Foster City, CA) with the Ion OneTouch 2 System (Life Technologies) with sequencing by the Ion Torrent PGM System (Life Technologies) utilizing the Ion PGM Sequencing 400 Kit (Life Technologies) and the Ion 318 Chip v2 (Life Technologies). The sequencing runs were set up following the default Ion Torrent Browser template for 16S Metagenomics Target Sequencing. The 16S rRNA sequence FASTQ files were processed by the Ion Torrent native software suite to split the DNA sequence reads into individual samples in each sequencing pool based on the 96 barcodes. DNA read sequence quality control discarded reads without the full primer sequences detected on both 5' and 3' ends. Primer sequences were trimmed. Reads with more than three expected errors based on the PHRED score [50, 57, 58] were discarded. To generate a *de-novo* 16S rRNA Operational Taxonomic Unit (OTU) reference to align the individual samples reads, all quality controlled reads from samples were concatenated into a single FASTQ file. Dereplication of exact duplicate reads was performed via the USEARCH/UPARSE software [50] with the -derep_fulllength and -sizeout flags, which adds to the sequence label the number of copies of that sequence in the original concatenated file. The USEARCH -minsize 2 –sortbysize functions organized the FASTA file from the most observations of a read sequence to the least number of observations and removed read sequences that were only seen once. The USEARCH -cluster_otus function then clustered the reads to OTUs at 97% sequence identity within a cluster and the cluster consensus sequence was selected as the representative OTU for the cluster. Chimera read detection was performed via the USEARCH -uchime_ref -strand plus flags with the -db gold.fa reference file provided with the USEARCH software. The resulting OTU FASTA file was labeled with the fasta_number.py script provided with the USEARCH software. The USEARCH -makeudb then made an alignment target file for later use in individual sample read sequence assignments to OTUs. Assignment of OTU reference sequences of bacterial taxa to the genus level was performed with the RDP Bayesian Classifier v2.6 java program with the default settings [20]. Alignment of each sample's quality controlled sample 16S rRNA read sequences to the OTU references was then performed by USEARCH with the commandline flags -usearch_global -strand both -id 0.97.

3.3.5. Statistical and Numerical Methods

Dichotomized hypothesis testing groups were determined by the adult Blood Azinphos-methyl detection (AZM >0.04 ng/g) or non-detection (ND) and household index adult pome fruit orchard farmworker status. Epidemiologic and exposure studies based on cohorts with low temporal resolution in the sampling design have unascertained individual toxico-pharmacokinetic differences in clearance, dose area under the curve (AUC), environmental exposure concentration, absorption and time since last exposure. This confounds inferring dose and the AUC from single measured concentration point estimates of exposure such as the blood detection data presented here. Due to this sparse sampling, binary dichotomized classification of exposed and unexposed states has been proposed as the most appropriate model by which to test cohorts with temporally sparse exposure sampling data structures over life course [134]. Dose-response models with sparse temporal biomarker data have dose group misclassification and loss of power due to the mixing between single point estimates of AUC groups. The sampling design performed here captures binary cross sectional estimates of exposure with two field collection operations occurring longitudinally from the two seasons.

We designed a balance between the OTU 97% sequence homology identity and the error rate at the 247 base pair mean PCR amplicon DNA fragment read length in the generation of the OTU molecular ontology. There is a greater than 7.41 nucleotide homology difference between OTUs (7.41=(1-0.97)x247bp) from the OTU *de-novo* cluster centroid generation. Quality control reduced the base identity error to <3 expected errors per read modeled on the PHRED score [50, 57, 58]. When the error rate is considered, the estimated difference between OTUs in base pair sequence homology provides a Hamming distance resolution greater than 4.4 (7.41 - 3) between OTUs that contribute to the meaningful molecular ontology of the DNA taxonomy.

The negative binomial identity, $Pr[missed\ taxon] = e^{-pn}$, with p proportion of taxon reads and n sample read depth of sequencing, quantifies the probability of not observing a taxonomy member of a composition if it truly exists there [54, 55]. Buccal samples were only considered for analysis if the depth of sequencing was >2,500 read threshold. This fixes the detection probability of members of the microbiota composition to be very high if the taxon is proportionally common. Analysis of only common taxa in deeply sequenced samples resolves the variable detection limits of rare members' false zero-inflation of non-detects. This occurs with randomly generated sample sequence depths as seen in pooled next generation sequencers. There are two taxonomies that may be queried in a 16S rRNA molecular sequence census. The molecular ontology taxonomy of *de-novo* OTU clustering is defined by the Hamming distances between categories of DNA read sequences with shared homology of >97% sequence identity. OTU clustering provides a physio-chemical DNA homology taxonomy structure independent of the assumed traditional nomenclature based taxonomies. We overlay on the molecular sequence OTU ontology the assumed Linnaean Kingdom to Genera level taxonomy by the RDP Bayesian Classifier [20]. The RDP taxonomic identification java tool uses known reference sequences from the Bergley's New Bacterial Taxonomy curation collection to assign OTUs to known bacterial isolates known to the genera level. OTU counts matching the same genera name were summed when making genera proportion calculations. Due to the differences of assumptions in these two (1)

molecular OTU and (2) Linnaean bacterial taxonomies, we used two different definitions of common for the interpretation of the data. Principal component machine learning methods adopted a less stringent common OTU definition to take exploratory advantage of the true molecular taxonomy while the more conservative the non-parametric methods adopted a more stringent common definition for hypothesis testing of differences in composition between exposure groups. Exploratory PC analysis were performed using the molecular taxonomy for common OTUs >0.1% sample proportion which at the minimum sample depth of 2,500 reads Pr[missed] ~0.08. To find group differences in common community composition at the Linnaean genera taxonomy level, common genera >0.5% proportion (Pr[missed genera at minimum 2,500 reads depth of sequencing] ~ 3.7×10^{-6}) were tested by Wilcoxon's Rank-Sum non-parametric association tests using the pesticide exposure categorical groups AZM and FW status. This common genera threshold also limits the number of genera-exposure group hypotheses tested. The index spring-summer 2005 buccal samples were used to select the definition for common and these same taxa were used for the winter inferences at the genera level. The proportional abundance correlation structure of compositional data simplex geometry posits that a compositional taxa's perturbation (i.e., reductions due to antibiotic exposure) has auto-correlated compensatory proportional increases in other taxa. This is because the remaining count of taxa reads are occupied by other remaining compositional members [51]. To adjust for bias that may occur when compensatory increases in other truly unassociated taxa are observed, each sample's taxonomic proportional abundances had the 'centered-log-ratio' transformation applied to correct for the auto-correlation nature of compositional data before subsequent statistical tests were applied [51-53].

The prcomp() base R function was used on the common OTU sample proportions once with the adult spring-summer buccal sequences. This adult PC analysis was then used to distinguish the beta diversity differences between samples in clusters generated via the model based cluster fitting algorithm mClust() available in R [61]. The adult spring-summer 2005 buccal sample collection was fitted to establish a decision value of from PC1 discriminating two clusters. To avoid overfitting both in the following winter and with the children when assessing cluster identity sharing between samples, the loadings from the adult's spring-summer PC1 and PC2 were used to project the winter buccal samples and the children's spring-summer OTU proportions. The adult spring-summer cluster decision value was then applied to the adult winter and children's buccal PC analysis in both seasons when assessing longitudinal cluster membership concordance between seasons for individual participants. The categorical variables adult AZM blood detection and farmworker status was then compared to PC1 scores by Welch's t-test and between the microbiome clusters by the Fisher's Exact Test to establish significance.

3.4. Results

3.4.1. Blood Measures of Azinphos-methyl Pesticide Exposure

Blood samples in both seasons collected from the adults were analyzed by mass spectrometry [19] for the presence of the compound organophosphate pesticide Azinphos-methyl. We did not have blood to analyze for the children for this study evaluation. After removal of samples with less than 2,500 reads, there are 72 adult individuals with microbiome data in

the spring-summer which had no detected Azinphos-methyl in the blood (Limit of Detection 0.04 ng/g plasma). Twenty-nine adult farmworkers had detection of blood AZM where we also had microbiome data in the spring-summer. A total of 101 adults from the summer had both microbiome data and blood measures of pesticide exposure. (See Figure 3-1 & 3-2) Of those samples with both blood pesticide exposure data from either season and winter microbiome sequencing data, 23 adult farmworkers had Azinphos-methyl detected in the spring-summer with the one non-farmworker which had Azinphos-methyl detected in the winter low exposure season. One sample has a non-zero result (0.021 ng/g) that was below the limit of detection (0.04 ng/g). This sample was treated as an Azinphos-methyl non-detect. All microbiome sequenced adults (n=101) from the summer collection had an AZM blood measure collected concurrently. There are 78 spring-summer children of adults with blood Azinphos-methyl concentration results and microbiome data. Thirteen adults with microbiome data from the winter collection had no winter measure of blood AZM but were also sampled and measured in the spring-summer for Azinphos-methyl. These samples and their children with no winter blood pesticide data were removed from the winter microbiome analysis leaving 82 adults and 62 children with both buccal microbiome and blood Azinphos-methyl in the winter.

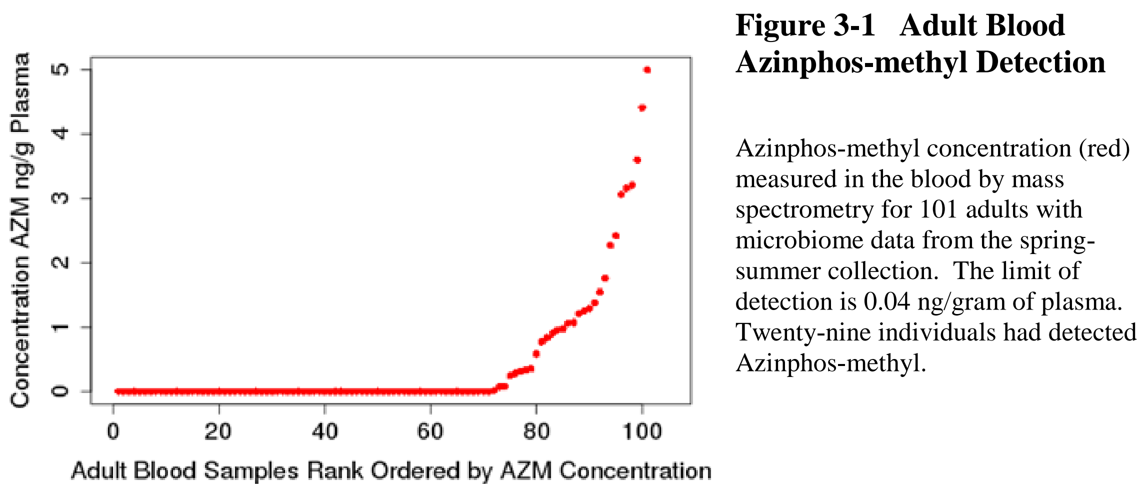


Figure 3-1 Adult Blood Azinphos-methyl Detection

Azinphos-methyl concentration (red) measured in the blood by mass spectrometry for 101 adults with microbiome data from the spring-summer collection. The limit of detection is 0.04 ng/gram of plasma. Twenty-nine individuals had detected Azinphos-methyl.

3.4.2. Census of 16S rRNA sequence data from adults and children.

De-novo operational taxonomic units (OTUs) were generated by USEARCH/UPARSE [50] from the sequenced reads with less than three expected errors and classified into the Bergley's New Bacterial Taxonomy down to the genera level using the RDP Classifier [20]. This resulted in 2520 OTUs with greater than 7.41 nucleotide differences between OTUs. These OTUs classified to 286 genera via the RDP classifier. Twenty-three of these genera were common at >0.5% mean proportion among children and adults. We used a minimum buccal sample sequencing depth threshold of 2,500 reads aligned to OTUs that passed quality control for inclusion of an individual's microbiome sample in further analysis. Reads from individual samples when aligned onto the OTUs resulted in a median buccal sample 16S rRNA read depth of 34,212 reads. These two depth of sequencing parameters, minimum and median depth of sequencing, gave us a minimum 0.92 and median >0.999 respective

probabilities of detecting all common genera (>0.5% of read) in individual samples. Among common OTUs (>0.1%) we had >0.999 probabilities of detection at minimum and median sampling depths.

3.4.3. Non-parametric Associations Between Exposure Groups and Genera Proportion using the Wilcoxon's Rank Sum Test

The Azinphos-methyl blood mass-spectrometry detection dichotomized the adults into AZM detect and non-detected groups. We used the household adult detection of Azinphos-methyl in the blood as a proxy exposure biomarker of the children due to the shared household environment of the sampled adult-child household pairs. This allowed for the group categorization of the children into Azinphos-methyl exposed and non-detected groups of their household adult. We ran the same non-parametric Wilcoxon's Rank Sum test on the adult and child samples independently and combined to test which bacteria may be perturbed in these high-risk exposure groups. Significant (FDR < 0.1) perturbations in twelve common (>0.5% proportion of reads) genera are associated in the spring-summer 2005 and/or winter 2006 collections with the adult blood Azinphos-methyl exposure detection groups among the sample groups. These genera included *Streptococcus, Herminiimonas, Haemophilus, Granulicatella, Micrococcineae, Gemella*, *Capnocytophaga, Actinobacillus, Halomonas, Actinomycineae, Prevotella,* and *Porphyromonas*. Each genera is summarized by blood Azinphos-methyl detection, season, adult and child subgroups in Figure 3-2 and Figure 3-3. Table 3-1 (Spring-Summer 2005) and Table 3-2 (Winter 2006) give a false discover adjusted report of the associated reductions in these bacterial organisms tested via the non-parametric Wilcoxon's Rank Sum test. Most all genera are reduced in abundance except *Herminiimonas* in the child subgroup and combined adult and child tests. The most consistent robust association is to the genera *Streptococcus* followed by *Haemophilus*.

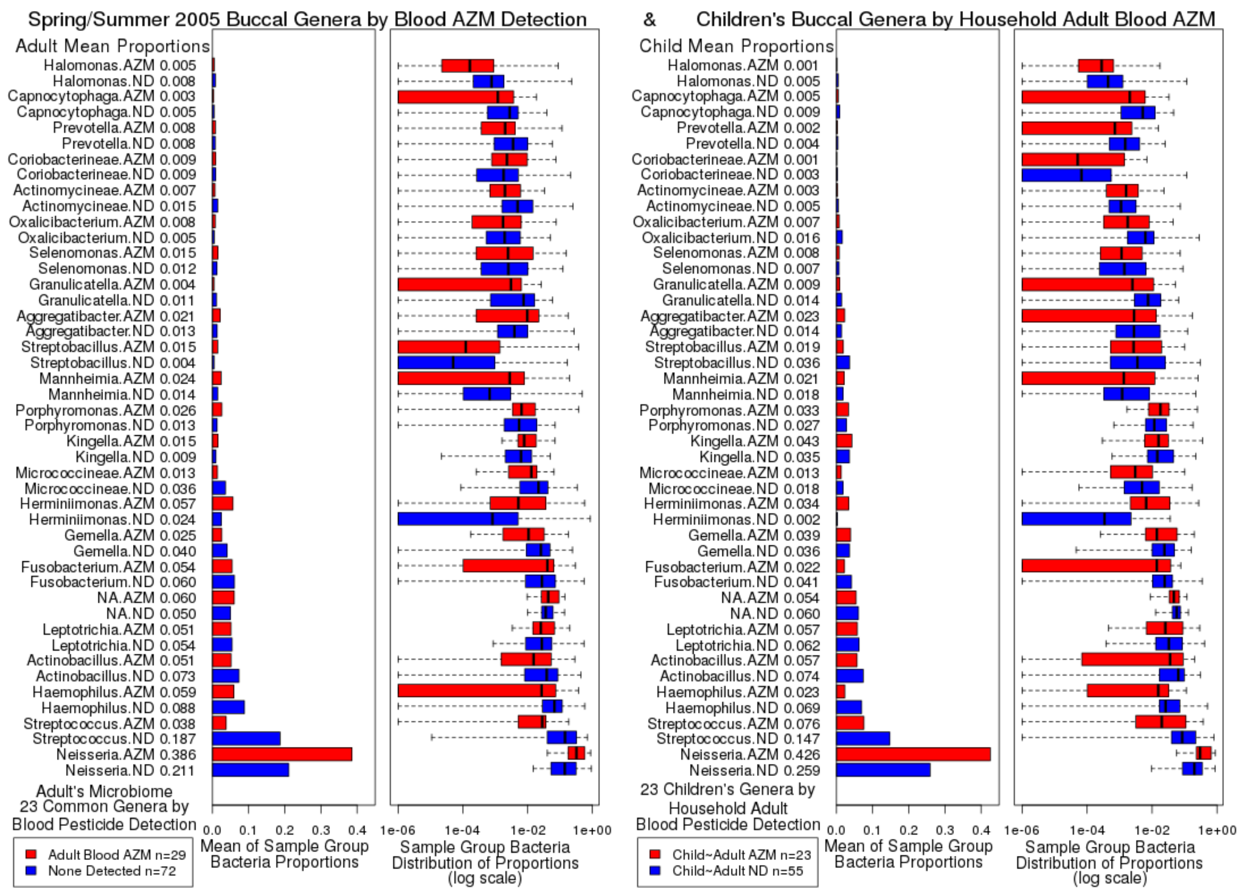
Table 3-1 **Wilcoxon's Rank Sum Tests of Spring-Summer 2005 Adult and Child Buccal Microbiota by Adult Azinphos-methyl Blood Detection**

Spring-Summer Microbiota Azinphos-methyl Wilcoxon's Rank Sum Test Results															
	Combined					Adults					Children				
Taxa	loc diffs	lower	upper	p-value	FDR	loc diffs	lower	upper	p-value	FDR	loc diffs	lower	upper	p-value	FDR
Streptococcus	**-1.70**	**-2.39**	**-1.05**	**0.000002**	**0.00004**	**-1.98**	**-2.74**	**-1.08**	**0.00005**	**0.001**	**-1.37**	**-2.59**	**-0.40**	**0.009**	**0.095**
Herminiimonas	**1.22**	**0.71**	**1.85**	**0.00005**	**0.0006**	1.01	-0.04	1.62	0.081	0.18	**1.85**	**0.97**	**2.95**	**0.00003**	**0.0008**
Haemophilus	**-1.13**	**-1.75**	**-0.49**	**0.0003**	**0.002**	**-1.05**	**-1.92**	**-0.24**	**0.008**	**0.05**	**-1.08**	**-2.15**	**-0.21**	**0.01**	**0.095**
Granulicatella	**-0.89**	**-1.50**	**-0.29**	**0.0013**	**0.007**	**-0.89**	**-1.64**	**0.00**	**0.03**	**0.09**	**-0.89**	**-1.83**	**-0.16**	**0.02**	**0.096**
Micrococcineae	**-0.88**	**-1.46**	**-0.31**	**0.004**	**0.02**	**-1.06**	**-1.74**	**-0.43**	**0.002**	**0.02**	-0.53	-1.59	0.57	0.34	0.56
Gemella	**-0.82**	**-1.44**	**-0.25**	**0.005**	**0.02**	**-0.99**	**-1.96**	**-0.25**	**0.009**	**0.05**	-0.53	-1.51	0.26	0.17	0.36
Capnocytophaga	**-0.57**	**-1.04**	**-0.04**	**0.02**	**0.05**	-0.57	-1.16	0.00	0.04	0.11	-0.61	-1.35	0.12	0.13	0.36
Actinobacillus	**-0.80**	**-1.55**	**-0.08**	**0.03**	**0.08**	-0.98	-2.09	0.13	0.09	0.18	-0.62	-1.55	0.24	0.16	0.36
Halomonas	**-0.54**	**-1.10**	**0.00**	**0.03**	**0.08**	**-0.80**	**-1.49**	**-0.07**	**0.01**	**0.06**	-0.12	-1.02	0.43	0.63	0.79
Mannheimia	0.49	0.00	1.16	0.05	0.12	0.79	0.00	1.55	0.04	0.11	0.13	-0.49	1.16	0.55	0.79
Neisseria	0.35	-0.02	0.72	0.06	0.12	0.45	-0.08	1.07	0.09	0.18	0.25	-0.25	0.73	0.33	0.56
Actinomycineae	-0.54	-1.13	0.02	0.06	0.12	**-1.00**	**-1.72**	**-0.20**	**0.01**	**0.06**	0.12	-0.56	0.71	0.69	0.79
Fusobacterium	-0.49	-1.16	0.00	0.08	0.14	-0.27	-1.16	0.30	0.37	0.57	-0.70	-1.68	0.10	0.10	0.34
Oxalicibacterium	-0.42	-1.13	0.05	0.09	0.15	-0.12	-0.90	0.64	0.62	0.75	-0.92	-2.06	0.03	0.06	0.26
Prevotella	-0.46	-0.98	0.08	0.10	0.16	-0.68	-1.47	0.13	0.11	0.19	-0.03	-0.87	0.65	0.81	0.85
otu_NA	-0.26	-0.64	0.10	0.16	0.2	-0.05	-0.68	0.43	0.84	0.92	-0.45	-1.00	0.07	0.08	0.30
Leptotrichia	-0.22	-0.70	0.27	0.4	0.5	0.00	-0.59	0.59	0.99	0.99	-0.51	-1.39	0.30	0.17	0.36
Selenomonas	-0.18	-0.80	0.38	0.4	0.6	-0.22	-1.11	0.81	0.65	0.75	-0.09	-1.06	0.60	0.61	0.79
Porphyromonas	-0.16	-0.64	0.37	0.5	0.6	-0.19	-0.90	0.62	0.54	0.75	-0.16	-0.87	0.57	0.72	0.79
Aggregatibacter	0.07	-0.40	0.75	0.6	0.7	0.12	-0.60	1.13	0.65	0.75	0.07	-0.72	1.05	0.72	0.79
Streptobacillus	0.00	-0.57	0.47	0.8	0.8	0.00	-0.25	0.56	0.98	0.99	-0.34	-1.56	0.78	0.62	0.79
Coriobacterineae	0.00	-0.41	0.69	0.8	0.8	0.13	-0.62	0.94	0.65	0.75	0.00	-0.47	0.76	0.93	0.93
Kingella	0.02	-0.46	0.51	0.9	0.9	0.35	-0.24	0.98	0.19	0.32	-0.44	-1.17	0.33	0.27	0.51

*upper and lower signify the Location Difference 95% confidence interval from the Wilcoxon's Rank Sum Test

The adult data show *Streptococcus, Haemophilus, Granulicatella, Micrococcineae, Gemella, Actinomycineae,* and *Halomonas* are significantly reduced in association with their blood Azinphos-methyl detection at an FDR <0.1 in the spring-summer season. The children show a reduction of *Streptococcus, Haemophilus, Granulicatella,* and increased *Herminiimonas.* The increased *Herminiimonas* among children is suggestively associated with a positive location difference (increased abundance) in the exposed adults. When the adults and children were combined *Streptococcus, Herminiimonas, Haemophilus, Granulicatella, Micrococcineae, Gemella*, *Capnocytophaga, Halomonas,* and *Actinobacillus* are significant at an FDR <0.1. All genera that were significant when combined were either suggestive or significant in the adult or child subgroups (Table 3-1).

Figure 3-2 Spring-summer Adult and Child Buccal Bacteria by Azinphos-methyl Detection



The spring-summer 2005 adults' and children's buccal microbiomes grouped by household adult Azinphos-methyl blood detection. Wilcoxon's Rank Sum test found significant (FDR < 0.1) associations for reductions in the genera *Streptococcus, Haemophilus, Granulicatella, Micrococcineae, Gemella, Actinomycineae* and *Halomonas* among the adults and reductions in the genera *Streptococcus, Haemophilus, Granulicatella* and increases in *Herminiimonas* among the children. Twenty-three common (>0.5% mean proportion) genera were tested for association. (See Table 3-1)

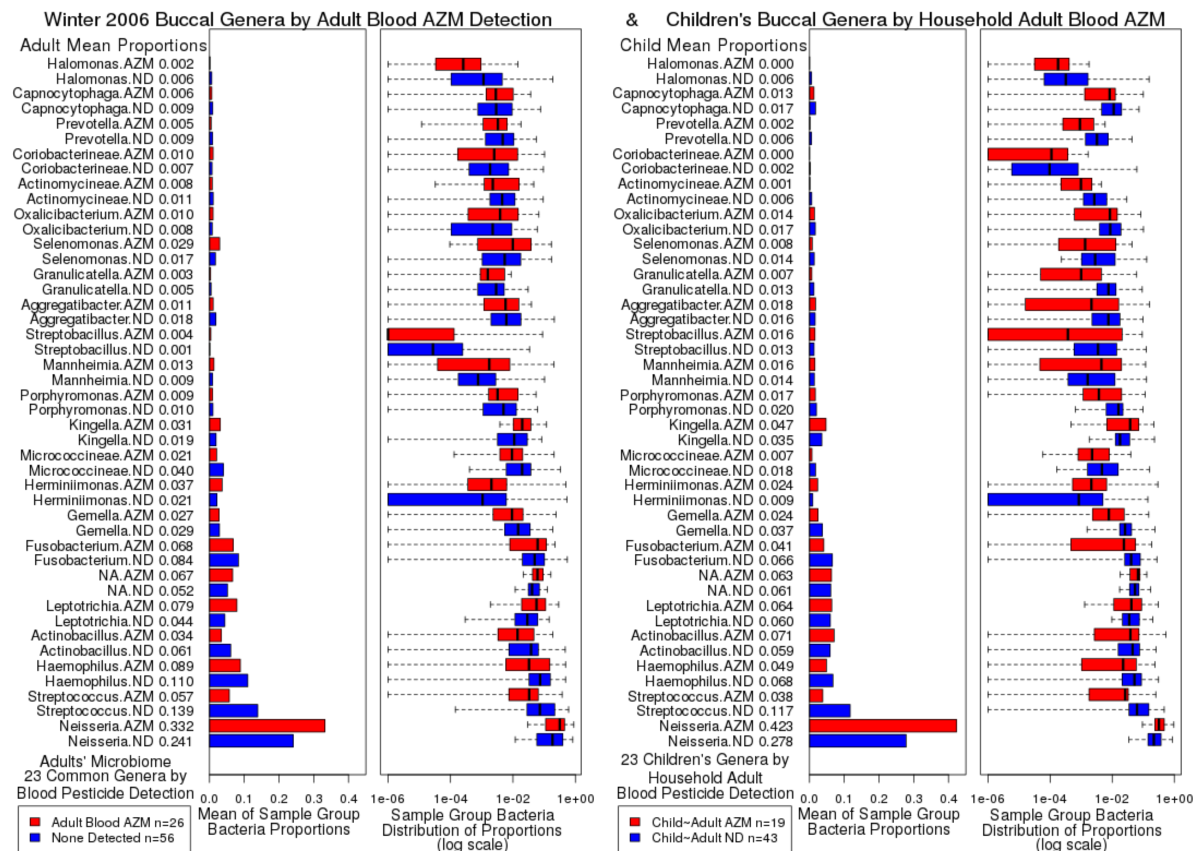
Table 3-2 **Wilcoxon's Rank Sum Tests of Winter 2006 Adult and Child Buccal Microbiota by Adult Azinphos-methyl Blood Detection**

Winter Microbiota Azinphos-methyl Wilcoxon's Rank Sum Test Results															
	Combined					Adults					Children				
Taxa	loc diffs	lower	upper	p-value	FDR	loc diffs	lower	upper	p-value	FDR	loc diffs	lower	upper	p-value	FDR
Streptococcus	**-1.33**	**-2.06**	**-0.66**	**0.0001**	**0.002**	**-1.19**	**-2.11**	**-0.30**	**0.01**	**0.14**	**-1.68**	**-3.05**	**-0.63**	**0.002**	**0.03**
Gemella	**-0.99**	**-1.59**	**-0.33**	**0.003**	**0.02**	-0.65	-1.58	0.33	0.17	0.39	**-1.37**	**-2.28**	**-0.45**	**0.004**	**0.03**
Actinomycineae	**-1.01**	**-1.65**	**-0.33**	**0.003**	**0.02**	-0.69	-1.49	0.10	0.09	0.34	**-1.43**	**-2.24**	**-0.30**	**0.006**	**0.03**
Halomonas	**-0.82**	**-1.43**	**-0.23**	**0.004**	**0.02**	**-1.04**	**-1.83**	**-0.19**	**0.01**	**0.14**	-0.41	-1.24	0.10	0.17	0.33
Granulicatella	**-0.72**	**-1.22**	**-0.14**	**0.008**	**0.04**	-0.23	-0.86	0.38	0.48	0.79	**-1.36**	**-2.37**	**-0.44**	**0.004**	**0.03**
Prevotella	**-0.86**	**-1.49**	**-0.20**	**0.009**	**0.04**	-0.49	-1.25	0.32	0.21	0.45	**-1.48**	**-2.49**	**-0.51**	**0.006**	**0.03**
Haemophilus	**-0.88**	**-1.68**	**-0.15**	**0.01**	**0.04**	-0.88	-1.97	0.07	0.09	0.34	-0.89	-2.20	0.08	0.07	0.20
Micrococcineae	**-0.82**	**-1.50**	**-0.16**	**0.02**	**0.04**	-0.78	-1.59	-0.09	0.03	0.23	-0.93	-1.89	0.19	0.10	0.24
Neisseria	0.45	0.01	0.89	0.04	0.103	0.46	-0.17	1.11	0.13	0.39	0.45	-0.26	1.06	0.19	0.34
Porphyromonas	-0.58	-1.14	-0.01	0.04	0.103	-0.18	-0.97	0.60	0.73	0.83	**-1.11**	**-1.88**	**-0.19**	**0.01**	**0.04**
Actinobacillus	-0.59	-1.32	0.05	0.09	0.19	-0.67	-1.70	0.24	0.15	0.39	-0.47	-1.70	0.48	0.41	0.50
Fusobacterium	-0.52	-1.26	0.07	0.10	0.20	-0.17	-1.26	0.60	0.59	0.83	-1.00	-2.08	-0.03	0.03	0.11
Mannheimia	0.48	-0.06	1.13	0.13	0.23	0.55	-0.25	1.31	0.28	0.54	0.47	-0.45	1.62	0.30	0.46
Kingella	0.35	-0.16	0.88	0.16	0.26	0.60	-0.07	1.30	0.08	0.34	0.14	-0.89	1.02	0.77	0.77
Capnocytophaga	-0.35	-1.01	0.19	0.19	0.30	-0.07	-0.93	0.63	0.79	0.83	-0.79	-1.58	0.06	0.08	0.22
Aggregatibacter	-0.35	-1.05	0.25	0.30	0.41	-0.01	-1.06	0.73	0.83	0.83	-0.83	-1.94	0.26	0.16	0.33
Selenomonas	-0.41	-1.19	0.38	0.29	0.41	-0.23	-1.27	1.02	0.73	0.83	-0.73	-1.89	0.45	0.22	0.36
Herminiimonas	0.17	-0.29	0.75	0.44	0.56	0.14	-0.60	0.87	0.64	0.83	0.17	-0.63	1.28	0.57	0.59
Leptotrichia	0.12	-0.35	0.57	0.65	0.79	0.42	-0.16	1.06	0.16	0.39	-0.36	-1.10	0.41	0.33	0.47
Streptobacillus	0.00	-0.54	0.24	0.78	0.84	0.00	0.00	0.47	0.66	0.83	-0.43	-1.68	0.63	0.39	0.50
Oxalicibacterium	-0.06	-0.82	0.61	0.80	0.84	0.13	-0.83	1.37	0.69	0.83	-0.37	-1.42	0.54	0.49	0.55
Coriobacterineae	-0.01	-0.74	0.61	0.76	0.84	0.12	-0.88	1.03	0.81	0.83	-0.03	-0.89	0.43	0.50	0.55
otu_NA	0.03	-0.29	0.37	0.89	0.89	0.18	-0.26	0.62	0.43	0.77	-0.24	-0.61	0.28	0.38	0.50

*upper and lower signify the Location Difference 95% confidence interval from the Wilcoxon's Rank Sum Test

The winter 2006 adults' Azinphos-methyl detection of common bacterial association results in Table 3-2 show a relaxation of the association in comparison to the spring-summer 2005 tests with no significant genera. The adult relaxation of association is in accordance with the seasonal variability in the pesticide exposure with suggestive FDRs (~ 0.14) in *Streptococcus* and *Halomonas*. The children appear to maintain more robust associations to their household adults Azinphos-methyl detection in the winter with the perturbed genera *Streptococcus, Gemella, Actinomycineae, Granulicatella, Prevotella,* and *Porphyromonas*. This may suggest the children's microbiomes do not recover as well as the adults. We only detect one non-farmworker in the winter with Azinphos-methyl in the blood. This individual also had a child sequenced. The tested exposure groups represent the aggregate exposure of the previous summer and the winter detection. The relaxation in association may signify that some microbiome recovery from the previous exposures in the spring-summer has occurred in the time elapsed. When the winter data for adults and children are combined *Streptococcus, Gemella, Actinomycineae, Halomonas, Granulicatella, Prevotella, Haemophilus,* and *Micrococcineae* are significant at an FDR < 0.1. See Table 3-2.

Figure 3-3 Winter Adult and Child Buccal Bacteria by Azinphos-methyl Detection



The winter 2006 adults' and children's buccal microbiomes grouped by household adult Azinphos-methyl blood detection versus non. Wilcoxon's Rank Sum test found suggestive evidence (FDRs ~0.14, p-values~0.01) of associations to reductions in the genera *Streptococcus* and *Halomonas* among the adults and found significant (FDR < 0.1) associations with reductions in the genera *Streptococcus, Gemella, Actinomycineae, Granulicatella, Prevotella* and *Porphyromonas* among the children. Twenty-three common (>0.5% mean proportion) genera were tested for association. (See Table 3-2)

3.4.4. Household Pome Farmworker Status

Significant perturbations in 10 common (>0.5% proportion of reads) genera are associated in either the spring-summer 2005 and/or winter 2006 collections with household pome fruit farmworker status. These taxa include *Streptococcus, Neisseria, Haemophilus, Gemella, Micrococcineae, Granulicatella, Actinobacillus, Leptotrichia, Kingella,* and *Halomonas*. Each taxon is summarized by household occupation, season, adult and child subgroups in Figures 3-4 (Spring-summer) and Figure 3-5 (Winter). Table 3-3 (Spring-summer) and Table 3-4 (Winter) give false discover adjusted reports of the associated reductions or increases (*Neisseria, Leptotrichia, Kingella*) in these bacterial organisms in association with household occupation tested via the non-parametric Wilcoxon's Rank Sum test for each season and adult-child subgroups.

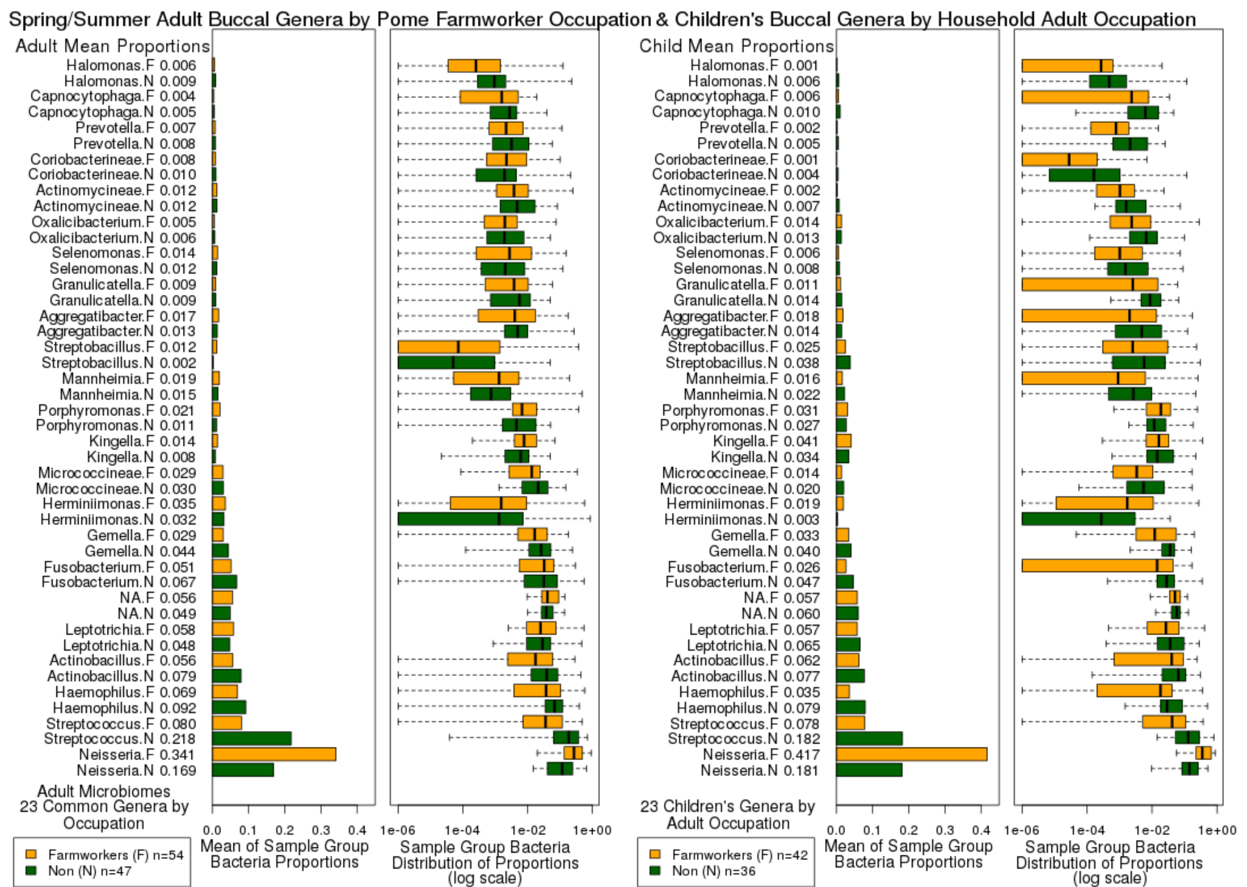
Table 3-3 Wilcoxon's Rank Sum Tests of Spring-summer 2005 Adult and Child Buccal Microbiota by Household Pome Farmworker Status

Spring-Summer Microbiota Farmworker Occupation Wilcoxon's Rank Sum Test Results															
	Combined					Adults					Children				
Taxa	loc diffs	lower	upper	p-value	FDR	loc diffs	lower	upper	p-value	FDR	loc diffs	lower	upper	p-value	FDR
Streptococcus	**-1.48**	**-2.11**	**-0.89**	**0.000001**	**0.000033**	**-1.48**	**-2.33**	**-0.61**	**0.00**	**0.01**	**-1.46**	**-2.54**	**-0.63**	**0.00**	**0.02**
Neisseria	**0.70**	**0.36**	**1.05**	**0.00004**	**0.00048**	**0.67**	**0.20**	**1.24**	**0.01**	**0.06**	**0.73**	**0.29**	**1.22**	**0.00**	**0.02**
Haemophilus	**-0.80**	**-1.36**	**-0.26**	**0.003**	**0.022**	-0.58	-1.31	0.00	0.06	0.32	**-0.96**	**-1.88**	**-0.15**	**0.02**	**0.08**
Gemella	**-0.73**	**-1.27**	**-0.22**	**0.004**	**0.03**	-0.58	-1.30	0.11	0.09	0.35	**-0.95**	**-1.82**	**-0.18**	**0.01**	**0.08**
Micrococcineae	**-0.68**	**-1.21**	**-0.11**	**0.018**	**0.08**	-0.59	-1.18	0.05	0.07	0.33	-0.66	-1.57	0.44	0.21	0.35
Actinobacillus	-0.67	-1.38	-0.05	0.03	0.13	-0.78	-1.84	0.21	0.13	0.43	-0.57	-1.43	0.18	0.14	0.29
Prevotella	-0.52	-1.01	0.00	0.04	0.14	-0.32	-1.04	0.35	0.39	0.78	-0.73	-1.39	0.03	0.08	0.21
Actinomycineae	-0.52	-1.05	0.00	0.05	0.15	-0.47	-1.19	0.22	0.21	0.54	-0.42	-1.10	0.23	0.19	0.34
Granulicatella	-0.41	-0.96	0.02	0.09	0.21	-0.01	-0.79	0.59	0.77	0.92	**-0.86**	**-1.71**	**-0.19**	**0.02**	**0.08**
Oxalicibacterium	-0.38	-0.90	0.04	0.09	0.21	-0.12	-0.73	0.65	0.72	0.92	-0.80	-1.72	0.02	0.06	0.17
Fusobacterium	-0.42	-1.00	0.05	0.10	0.21	-0.16	-0.88	0.41	0.60	0.92	-0.76	-1.67	0.00	0.05	0.15
Capnocytophaga	-0.32	-0.75	0.07	0.16	0.29	-0.16	-0.67	0.34	0.54	0.92	-0.60	-1.34	0.22	0.16	0.32
Halomonas	-0.27	-0.79	0.05	0.16	0.29	-0.61	-1.20	0.00	0.04	0.29	0.08	-0.62	0.59	0.71	0.75
Herminiimonas	0.19	0.00	0.76	0.21	0.35	0.00	-0.63	0.57	0.73	0.92	0.69	0.00	1.46	0.02	0.08
Leptotrichia	-0.18	-0.61	0.28	0.38	0.59	0.07	-0.45	0.67	0.80	0.92	-0.54	-1.25	0.18	0.10	0.24
Streptobacillus	0.00	-0.66	0.28	0.46	0.66	0.00	-0.38	0.08	0.59	0.92	-0.49	-1.53	0.63	0.55	0.72
Kingella	0.13	-0.31	0.56	0.50	0.68	0.34	-0.21	0.91	0.20	0.54	-0.17	-0.87	0.50	0.63	0.72
otu_NA	-0.08	-0.36	0.21	0.60	0.76	0.04	-0.41	0.40	0.88	0.96	-0.19	-0.64	0.27	0.43	0.65
Mannheimia	0.06	-0.35	0.64	0.63	0.76	0.38	-0.23	1.15	0.28	0.63	-0.19	-1.11	0.57	0.60	0.72
Selenomonas	-0.07	-0.66	0.45	0.69	0.76	0.00	-0.82	0.88	0.99	0.99	-0.11	-0.98	0.55	0.60	0.72
Coriobacterineae	0.00	-0.50	0.39	0.68	0.76	0.07	-0.57	0.80	0.77	0.92	0.00	-0.70	0.29	0.56	0.72
Porphyromonas	0.08	-0.35	0.57	0.73	0.76	0.26	-0.33	0.99	0.41	0.78	-0.15	-0.77	0.52	0.70	0.75
Aggregatibacter	0.00	-0.61	0.40	0.83	0.83	0.00	-0.76	0.65	0.92	0.96	0.02	-1.09	0.71	0.90	0.90

*upper and lower signify the Location Difference 95% confidence interval from the Wilcoxon's Rank Sum Test

The adult farmworkers showed significant (FDR <0.1) reductions in *Streptococcus* and increases in *Neisseria* in both seasons and additionally gained significance for *Haemophilus, Actinobacillus, Leptotrichia, Kingella,* and *Halomonas* in the winter. The children also showed significant (FDR <0.1) spring-summer reductions in *Streptococcus* and increases in *Neisseria* and as well as reductions in *Haemophilus, Gemella* and *Granulicatella*. When adult and child subgroups are combined *Streptococcus, Neisseria, Haemophilus, Gemella,* and *Micrococcineae* are all significant. The observed increases in *Neisseria* may be due to *Neisseria* being less susceptible to the exposure effects and be the most proportionally common remaining microbiota surveyed. *Streptococcus, Haemophilus, Gemella, Micrococcineae,* and *Halomonas* are all also associated with negative location differences from non-parametric testing in both pome fruit farmworker groups and blood Azinphos-methyl detection exposure groups.

Figure 3-4 **Spring-summer Adult and Child Buccal Bacteria by Pome Farmworker Occupation**



The spring-summer 2005 adult and child buccal microbiomes grouped by household occupational status for pome fruit farmworker. Wilcoxon's Rank Sum test found significant (FDR <0.1) associations for reductions in the genera *Streptococcus* and increases in *Neisseria* among the adults and children (See Table 3-3). The children also had significant reductions in *Haemophilus, Gemella,* and *Granulicatella*. Twenty-three common (>0.5% mean proportion) genera were tested for association.

Table 3-4 Wilcoxon's Rank Sum Tests of Winter 2006 Adult and Child Buccal Microbiota by Household Pome Farmworker Status

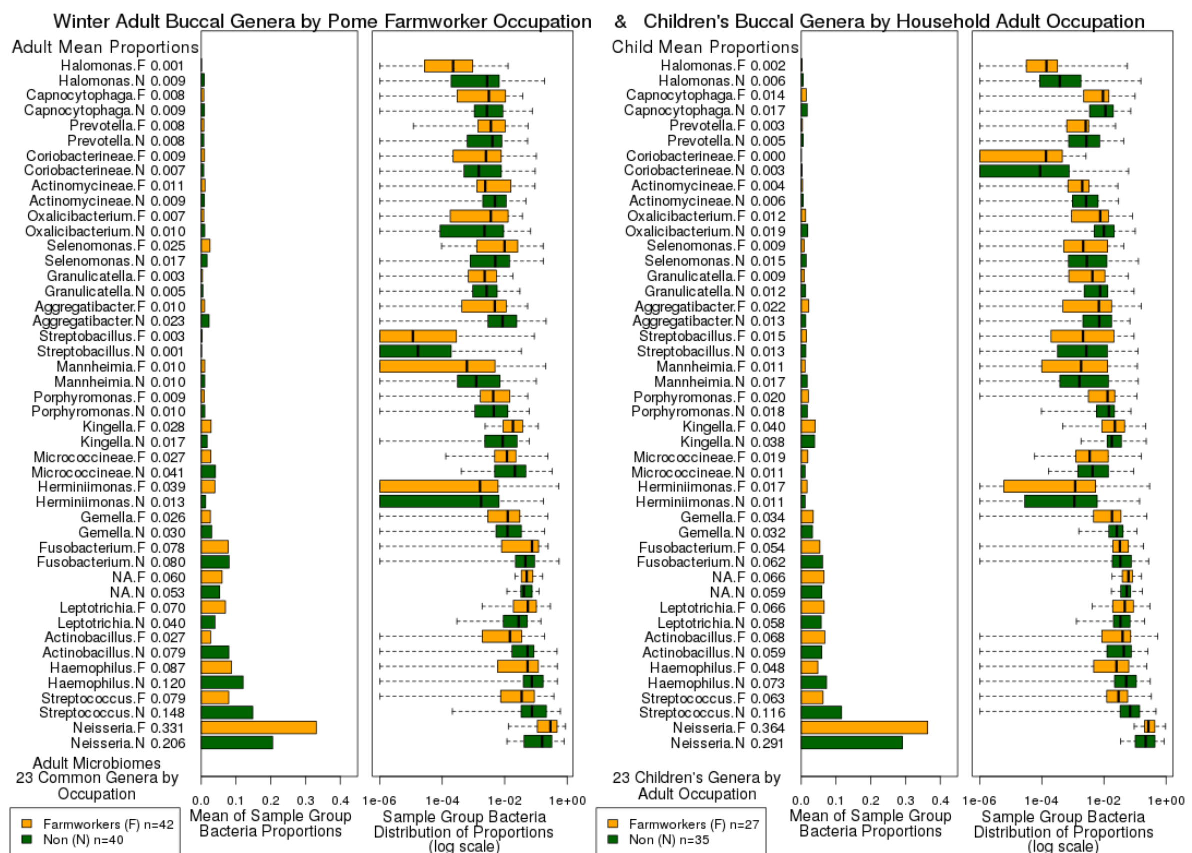
Winter Microbiota Farmworker Occupation Wilcoxon's Rank Sum Test Results

	Combined					Adults					Children				
Taxa	loc diffs	lower	upper	p-value	FDR	loc diffs	lower	upper	p-value	FDR	loc diffs	lower	upper	p-value	FDR
Halomonas	**-1.03**	**-1.59**	**-0.50**	**0.00004**	**0.001**	**-1.49**	**-2.29**	**-0.79**	**0.0001**	**0.002**	-0.52	-1.43	0.00	0.04	0.44
Streptococcus	**-0.94**	**-1.55**	**-0.34**	**0.002**	**0.03**	**-0.98**	**-1.84**	**-0.10**	**0.02**	**0.09**	-0.92	-1.93	-0.06	0.03	0.44
Neisseria	0.52	0.10	0.94	0.01	0.11	**0.74**	**0.13**	**1.30**	**0.02**	**0.07**	0.27	-0.34	0.86	0.39	0.81
Actinobacillus	-0.70	-1.32	-0.06	0.03	0.16	**-1.11**	**-2.04**	**-0.21**	**0.01**	**0.07**	-0.16	-1.07	0.61	0.69	0.89
Haemophilus	-0.67	-1.39	-0.01	0.04	0.16	**-0.60**	**-1.62**	**0.20**	**0.17**	**0.57**	-0.77	-1.84	0.09	0.08	0.58
Leptotrichia	0.41	0.01	0.81	0.05	0.17	**0.67**	**0.13**	**1.18**	**0.01**	**0.07**	0.07	-0.63	0.74	0.86	0.90
Kingella	0.45	-0.02	0.90	0.06	0.19	**0.86**	**0.25**	**1.49**	**0.01**	**0.07**	-0.02	-0.78	0.69	0.95	0.95
Gemella	-0.35	-0.95	0.15	0.18	0.50	-0.21	-1.07	0.64	0.63	0.83	-0.46	-1.24	0.17	0.16	0.76
Granulicatella	-0.23	-0.74	0.28	0.32	0.74	0.00	-0.65	0.63	0.87	0.90	-0.48	-1.43	0.34	0.22	0.76
Actinomycineae	-0.30	-0.89	0.27	0.31	0.74	-0.23	-0.94	0.48	0.63	0.83	-0.49	-1.53	0.35	0.29	0.81
Oxalicibacterium	-0.27	-0.93	0.31	0.36	0.76	0.01	-0.84	1.05	0.89	0.90	-0.54	-1.46	0.25	0.19	0.76
otu_NA	0.12	-0.20	0.40	0.52	0.86	0.21	-0.19	0.63	0.31	0.71	-0.11	-0.53	0.35	0.70	0.89
Micrococcineae	-0.17	-0.77	0.46	0.60	0.86	-0.37	-1.01	0.28	0.26	0.71	-0.23	-1.24	0.78	0.59	0.89
Aggregatibacter	-0.18	-0.89	0.34	0.52	0.86	-0.27	-1.19	0.46	0.49	0.82	-0.19	-1.10	0.71	0.71	0.89
Selenomonas	0.22	-0.48	0.97	0.52	0.86	0.50	-0.44	1.47	0.28	0.71	-0.32	-1.25	0.90	0.58	0.89
Coriobacterineae	0.12	-0.43	0.74	0.57	0.86	0.35	-0.50	1.19	0.39	0.74	0.00	-0.94	0.43	0.58	0.89
Porphyromonas	-0.10	-0.65	0.41	0.69	0.93	0.13	-0.64	0.88	0.68	0.83	-0.35	-1.06	0.37	0.36	0.81
Herminiimonas	0.00	-0.43	0.33	0.84	0.97	0.00	-0.70	0.55	0.83	0.90	0.00	-0.80	0.61	0.82	0.90
Streptobacillus	0.00	-0.45	0.33	0.80	0.97	0.00	0.00	0.55	0.50	0.82	-0.07	-1.28	0.95	0.74	0.89
Capnocytophaga	-0.02	-0.62	0.48	0.85	0.97	0.32	-0.50	1.08	0.38	0.74	-0.43	-1.26	0.34	0.23	0.76
Fusobacterium	0.00	-0.61	0.63	0.99	0.99	0.16	-0.77	1.04	0.66	0.83	-0.19	-1.06	0.61	0.53	0.89
Mannheimia	0.00	-0.54	0.58	0.96	0.99	0.00	-0.61	0.87	0.90	0.90	-0.06	-1.00	0.82	0.81	0.90
Prevotella	0.00	-0.63	0.55	0.99	0.99	0.21	-0.62	0.98	0.59	0.83	-0.39	-1.34	0.51	0.36	0.81

*upper and lower signify the Location Difference 95% confidence interval from the Wilcoxon's Rank Sum Test

The adult pome farmworker winter subgroup had reductions in the genera *Halomonas, Streptococcus, Actinobacillus, Haemophilus,* and increases in *Leptotrichia, Kingella,* and *Neisseria*. The children had no FDR adjusted significant genera associated with their winter pome farmworker status but do have *Halomonas* and *Streptococcus* with low p-values (<0.04). This relaxation in the child winter pome farmworker status association but not the adults is the opposite of the pattern observed with Azinphos-methyl exposure. This may suggest that there is some heterogeneity in what the pome farmworker status and Azinphos-methyl detection are measuring. When the adult and child subgroups are combined, pome farmworker associated genera include *Streptococcus* and *Halomonas*.

Figure 3-5 Winter Adult and Child Buccal Bacteria by Pome Farmworker Occupation



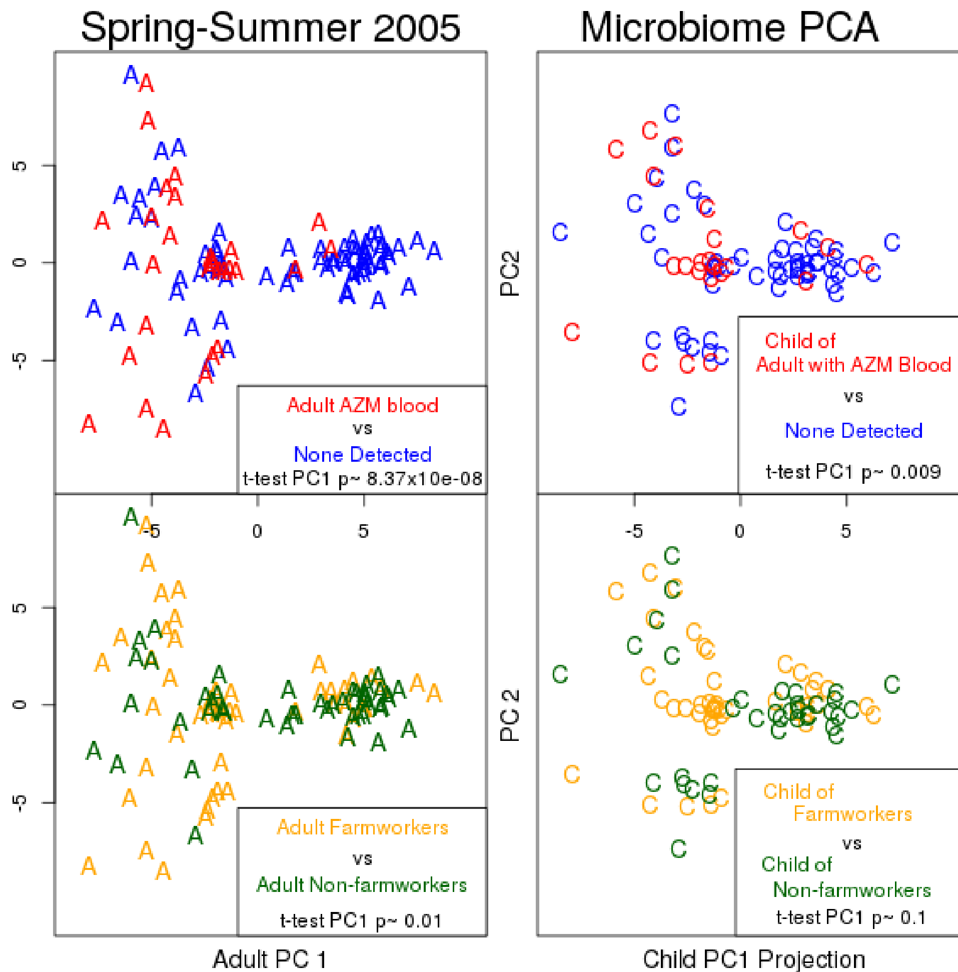
The winter 2006 adult and child buccal microbiomes grouped by household occupational status for pome fruit farmworker versus non-farmworker. Wilcoxon's Rank Sum test found reductions in the genera *Streptococcus, Actinobacillus, Haemophilus, Halomonas* and increases in *Leptotrichia, Kingella and Neisseria* (FDR < 0.1) in the adults. We found no false discovery adjusted associations to the genera detected in these agricultural community children. Twenty-three common (>0.5% mean proportion) genera were tested for association. (See Table 3-4)

3.4.5. Principal Component Analysis of Buccal Sample Common OTU Proportions

Summer Microbiome PCA comparison to Azinphos-methyl Detection and Pome Farmworker Status

The centered log ratio of each buccal sample's common OTUs (>0.1% mean, n = 153 spring-summer) were input to the PC analyses. The results from the summer buccal microbiome composition PCA are presented in Figure 3-6 by principal components 1 and 2 which explain 12.6% and 5.8% of the variance respectively in adults. The spring-summer loadings from the adults' PC1 and PC2 were used with the adults' winter OTUs and the children's OTUs in both seasons as out of sample projections. This allowed for more unbiased assessment of cluster generation and co-clustering relationships from spring-summer to winter and between adults and their matched household child. Visual inspection revealed two distinct primary clusters in the PC pattern. Clusters were called by the mClust() model based clustering algorithm [61] and generated a PC1 decision value of 0.189 from the spring-summer score distribution. Each of the categorical variables (AZM and household pome farmworker status) were compared to the microbiome beta diversity via Welch's t-test of the PC1 scores and by the Gaussian model based clusters generated from the PC1 scores. Significant associations of the categorical environmental exposure variables are seen between these two clusters defined by PC1.

Figure 3-6 **Spring-summer Adult and Child Microbiome PCA**



The spring-summer 2005 Buccal Swab Microbiome PCA by adult blood Azinphos-methyl detection and pome fruit farmworker household status for adults and children. The children's PC scores were generated by projecting the child OTUs through the adults' PC loadings. Note that each panel is color coded by the different categorical exposure variables with the left and right panels as adults and children respectively. The significance for each exposure measures association to microbiome composition defined by PC1 is reported in the lower right boxes in each panel. Agricultural exposure was defined by 1) adult blood AZM detection greater than the limit of detection (0.04 ng/g plasma) in the in the top left and right panels for adults and children respectively and 2) the combine adult and child cohort demographic farmworker household status in the lower left panel. The PCA of adult samples explained 12.6% of the variance with PC1 and 5.8% with PC2.

The two measures of agricultural exposure (adult blood AZM and farmworker status) were dichotomized into groups. Welch's t-tests between the respective groups were performed with PC1 and shown to be significant (See Figure 3-6: Summer 2005 Microbiome PCA; Adult AZM vs. ND: p-value ~ 8.37×10^{-8}, Child of Adult w/ AZM vs. ND: p-value ~ 0.009, Adult Farmworker vs. Non: p-value ~ 0.01, Child of Farmworker vs. Non: p-value ~ 0.1). The sample group classes are increased to the left in the direction of PC1. Using this PC1 directionality and a decision value of greater than and less than 0.189 determined by Gaussian model based cluster discriminant analysis [61] of the centered microbiome PC1, we classified the individuals to “Exposed” and “Unexposed” PC1 groups (See Figure 3-8). The right exposed cluster is more likely to have detection of blood AZM (Adults & Children

Combined Spring-summer Fisher's exact test p-value ~1.3×10^{-8}, Odds Ratio 9) and Farmworker household status (Adults & Children Combined Spring-summer Fisher's exact test p-value ~7.8×10^{-5}, Odds Ratio 3.47). Table 3-5 and Table 3-6 report the adult and child subgroup counts, odds, confidence intervals and summary statistics.

Table 3-5 Azinphos-methyl Detection Spring-Summer 2005 Fisher's Exact Tests of PCA Cluster Membership

Spring-Summer Microbiome PC 1 Clusters						
	*Combined		**Adults		***Children	
	Exposed	Unexposed	Exposed	Unexposed	Exposed	Unexposed
Azinphos-methyl	47	7	27	3	20	4
None-Detected	53	72	29	42	24	30

*p-value ~ 0.00000001; Odds Ratio: 9, 3.6 - 25 (95%ci)

**p-value ~ 0.003; Odds Ratio: 13, 3.4 - 72 (95%ci)

***p-value ~ 0.001; Odds Ratio: 6.1, 1.7 - 28 (95%ci)

Table 3-6 Pome Farmworkers Spring-summer 2005 Fisher's Exact Tests of PCA Cluster Membership

Spring-Summer Microbiome PC 1 Clusters						
	*Combined		**Adults		***Children	
	Exposed	Unexposed	Exposed	Unexposed	Exposed	Unexposed
Farmworkers	67	29	38	16	29	13
Non	35	50	18	29	15	21

*p-value ~ 0.000078; Odds Ratio: 3.47, 1.8 - 6.8 (95%ci)

**p-value ~ 0.001; Odds Ratio: 3.8, 1.5 – 9.5 (95%ci)

***p-value ~ 0.022; Odds Ratio: 3.1, 1.1 - 8.8 (95%ci)

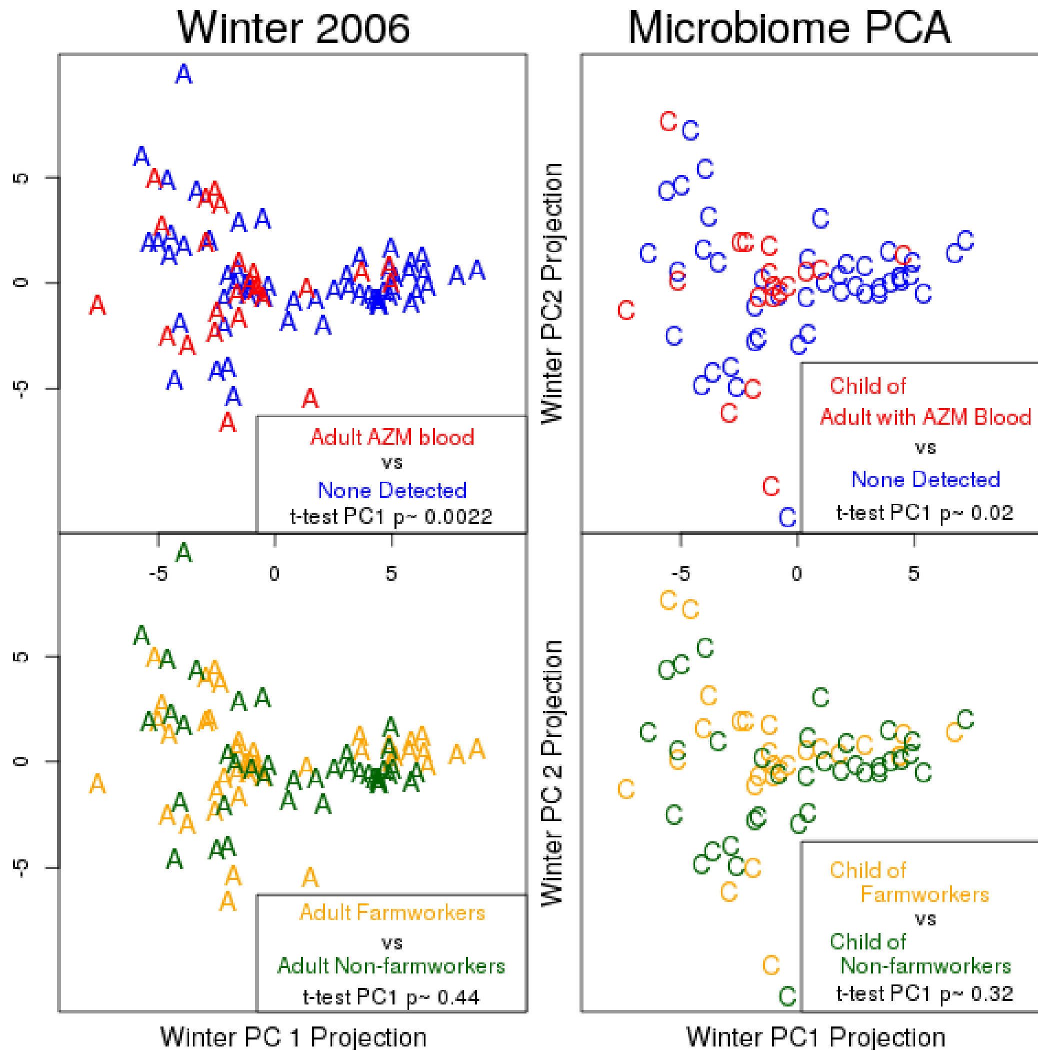
The samples that are in the left exposed cluster of PC1, generally have a more diverse score in PC2. This may suggest those individuals who have the altered microbiome profile are in a generally more chaotic dysbiotic state with more variation in community structure than the unexposed individuals as seen by the unexposed in the tighter grouping of the right cluster. This human cohort's high variability in dysbiotic compositional state is similar to results in quail and rats [90-94] where dysbiosis is also seen in the gut microbiomes of test animals exposed to organophosphate pesticides.

3.4.6. Winter Microbiome PCA comparison to Azinphos-methyl Detection and Pome Farmworker Status

In Figure 3-7 the winter 2006 buccal sample OTUs PC1 scores were compared to the exposure classes of adult AZM blood and farmworker status. These are again increased to the left direction of PC1, but are depressed in the winter association which is like the non-parametric tests. The blood AZM detection maintains robust significance (Welch's T-test

adults, p~0.002, children, p~0.02) with the winter microbiome PC1, but farmworker status becomes suggestive (p>0.32).

Figure 3-7 Winter Adult and Child Microbiome PCA



Winter 2006 Microbiome PCA projections of adults and children by categorical exposure variables. The winter PC scores were generated by projecting the winter OTUs through the adults' spring-summer PC loadings. Note that each panel is color coded by the different categorical exposure variables with the left and right panels as adults and children respectively. The significance for each exposure measures association to microbiome composition defined by PC1 is reported in the lower corner of the panels. Agricultural exposure was defined by 1) adult blood AZM greater than the limit of detection (0.04 ng/g plasma) in the in the top left and right panels for adults and children and 2) the adult and child cohort demographic farmworker household status in the lower panels.

This winter off-season depression in association is also evident by the lower separation between the modeled winter exposure group clusters in the top right panel of Figure 3-8 versus the summer modeled exposure clusters in the top left panel of Figure 3-8. The right winter exposed cluster is still more likely to have detection of the previous summer's adult blood AZM (Adults & Children Combined Spring-summer Fisher’s exact test p-value

~0.00005, Odds Ratio 5.3) and Farmworker household status (Adults & Children Combined Spring-summer Fisher's exact test p-value ~0.018, Odds Ratio 2.3). Table 3-7 and Table 3-8 provide Fisher's Exact Test combined and subgroup summary association statistics. Subgroup analysis of adults and children showed a strong association in the adults and the children both in the household farmworker status and household adult blood Azinphos-methyl detection categorical variables.

Table 3-7 Azinphos-methyl Detection Winter 2006 Fisher's Exact Tests of PCA Cluster Membership

Winter Microbiome PC 1 Clusters						
	*Combined		**Adults		***Children	
	Exposed	Unexposed	Exposed	Unexposed	Exposed	Unexposed
Azinphos-methyl	37	8	21	5	16	3
None-Detected	46	53	26	30	20	23

*p-value ~ 0.00005; Odds Ratio: 5.3, 2.1 - 14 (95%ci)

**p-value ~ 0.004; Odds Ratio: 4.8, 1.5 - 18 (95%ci)

***p-value ~ 0.006; Odds Ratio: 6, 1.4 - 36 (95%ci)

Table 3-8 Pome Farmworkers Winter 2006 Fisher's Exact Tests of PCA Cluster Membership

Microbiome PC 1 Clusters						
	*Combined		**Adults		***Children	
	Exposed	Unexposed	Exposed	Unexposed	Exposed	Unexposed
Farmworkers	47	22	28	14	19	8
Non	36	39	19	21	17	18

*p-value ~ 0.02; Odds Ratio: 2.3, 1.1 - 4.8 (95%ci)

**p-value ~ 0.12; Odds Ratio: 2.2, 0.82 - 5.9 (95%ci)

***p-value ~ 0.12; Odds Ratio: 2.5, 0.78 – 8.4 (95%ci)

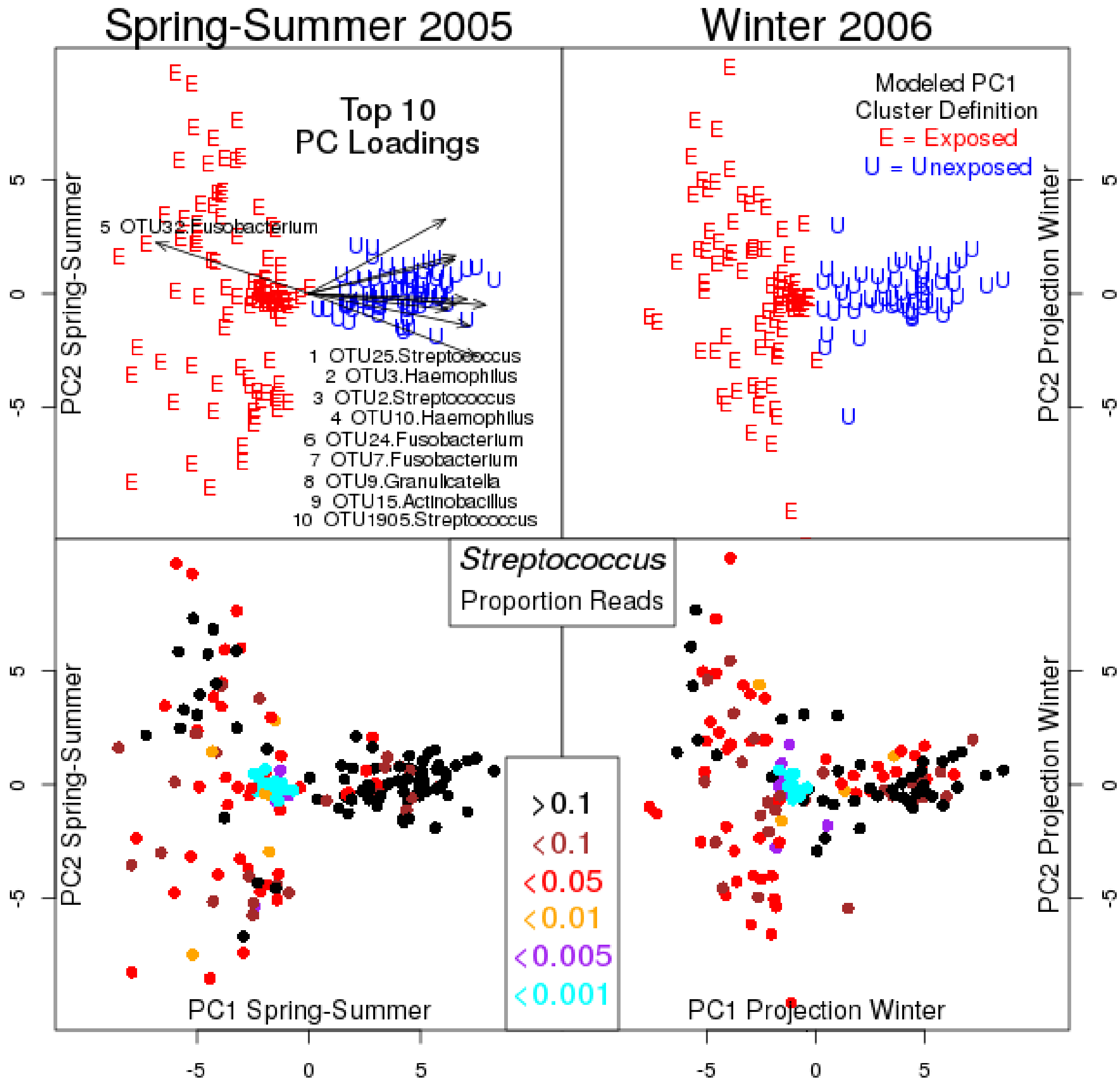
There is minor recovery of the host's microbiome compositional phenotypes in the winter when exposure is less frequently detected, despite the winter depression in association among subgroups. There is a low rate of return of the adult AZM blood detected exposed cluster individuals to the "normal" right hand unexposed clusters from summer to winter (Figure 3-8). Based on the fitted buccal microbiome PC1 modeled cluster phenotype, 4.3% (3/70) of exposed cluster people with data in both seasons migrate from the exposed to the unexposed PC1 cluster over the ~6-9 months between samplings. This low rate of return to the 'normal' unexposed microbiome cluster is an example of the persistence of the microbiome compositional state once it has been established. Other investigators have found that when microbiota extinctions occur in the rat gut microbiome, the altered microbiome state persists and remains extinct unless recolonization occurs via inoculation of the live organisms [40]. Thus, any individuals returning to the unexposed cluster could be an example of host re-

colonization from the human environment and from personal contact with community, kin and kindred. This relative stability is evident in the 96% (111/116) of individuals remaining on the same respective side of the centered PC1 fitted model decision value. This suggests that, similar to the rat gut microbiome study [40], once the buccal microbiome has been set into an altered state based on extinctions of specific genera, it persists in this reduced bacterial abundance state until bacterial exposure events recolonize the host.

3.4.7. Top OTU Loadings in the Principal Components Analysis

We assessed the top 10 OTU loadings in PC1 and PC2 which are driving the PC analysis of variance. In the Spring-summer sampling the top 10 PC drivers OTUs were mainly dominated in the right-hand direction of PC1 with nearly all of them (9 of 10 OTUs) increased in the direction of the unexposed cluster, see Figure 3-8 in the top two panels.

Figure 3-8 **PCA Loadings, Cluster Membership and *Streptococcus* Abundance**



PCA OTU loading vectors in the top left panel show the top 10 drivers of the PCAs for the spring-summer microbiome. The directionality of these vectors is almost exclusively contained by PC1 and points to the loading vectors being OTUs increased in proportion in the direction of the unexposed cluster. The top panels also show the mclust() modelled cluster calls. The bottom two panels show the aggregate of *Streptococcus* OTUs by the proportion of reads in each sample as colored in the scale legend. The left hand exposed cluster contains individuals with much lower proportions of *Streptococcus* with a tight grouping of individuals with very low (<0.001, cyan) proportions versus the left unexposed cluster which has almost exclusively individuals with high proportions (>0.1, black). The PCA of adult samples explained 12.6% of the variance with PC1 and 5.8% with PC2.

This increase of nearly all the top driver PC loadings in the direction of the unexposed cluster suggests that these individuals have more of these OTUs representative organisms. These classify to the genera *Streptococcus, Haemophilus, Granulicatella, Fusobacterium,* and *Actinobacillus*. *Fusobacterium* is split between directionality with the same two OTU7.*F*

and OTU32.*F*. The OTU25.*Streptococcus* taxon is the top ranked driver loading while the exposure is occurring in the spring-summer collection. *Streptococcus* is very common genera in the oral context. To illustrate the effect of the reduction of diversity on the genera *Streptococcus* we shaded the PC projection with the proportion of sample read alignments matching a *Streptococcus* RDP classified OTU. The color scale spans the range of < 0.001 to > 0.1 proportion as seen in the bottom two panels of Figure 3-9. Many individuals in the right exposed cluster direction of PC1 contain very low proportions (~<0.01) of this normally very common genera.

3.4.8. Adult-Child PC1 Co-clustering

Nearly all children's oral buccal microbiome compositions cluster with their household resident adult microbiome. In the summer collection 70 adult-child pairs were buccal 16S rRNA sequenced. Sixty-seven (96%) matched microbiome PCA clusters and three were unmatched. Of these, 37/38 (97.4%) adult-child pairs in the Azinphos-methyl exposed PCA cluster group co-clustered. One child of an adult (1/38) in the Azinphos-methyl exposed summer cluster had a child PC1 score in the unexposed group. In the summer unexposed cluster group, 30/32 (93.7%) adult-child pairs matched in the summer unexposed cluster definition. Two children (2/31) with a household adult in the unexposed summer cluster, clustered with those individuals in the Azinphos-methyl exposed summer cluster group. In the winter, 53 adult-child pairs were sequenced. Of these, 50/53 (94%) adults and their household children matched winter cluster definition within the household with three children not matching their household adult microbiome winter cluster. In the Azinphos-methyl exposed microbiome winter cluster, 31/33 (94%) children clustered with their household adult and three did not. Nineteen (95%; 19/20) children match their adult in the winter unexposed cluster with one child being unmatched into the winter exposed cluster. We assessed the significance by the Fisher's Exact Test of the non-random co-clustering of children with respective adults between the two PC clusters adult-child pair matching definitions as shown in Table 3-9.

Table 3-9 Household Adult-Child Pairs Principal Component Co-Clustering Match Counts

Table 3-9 Household Adult-Child Pairs Principal Component Co-Clustering Match Counts				
	*Spring/Summer PC Clusters		**Winter PC Clusters	
	Adult-Exposed	Adult-Unexposed	Adult-Exposed	Adult-Unexposed
Child-Exposed	37 match	2 non	31 match	1 non
Child-Unexposed	1 non	30 match	2 non	19 match

Significance by Fisher's Exact Test for Count Data

*p-value $< 2.6 \times 10^{-16}$; Odds Ratio: 407, 40 - 16384 (95%ci)

**p-value $< 3.3 \times 10^{-11}$; Odds Ratio: 217, 21 - 11202 (95%ci)

Forty-one adult-child pairs spanned both seasons' collections with all samples present. This allowed us to follow how individual household pairs co-clustering changed over time from summer to winter. Of these adult-child pairs 93% (38/41) co-clustered in their respective

clusters with the same household member for both seasons. Two adult-child pairs moved together from exposed to unexposed.

3.4.9. Alpha Diversity Analysis of PC1 Clusters

The Alpha diversity total count of the OTU richness was compared to the clusters generated from PC1. The diversity estimation method of Chao was used as it makes an adjustment to the observed OTU count based on the number of rare member counts [59, 95]. The exposed PC cluster had significant reductions in diversity with adult t-test p-values < 0.02 for spring-summer and winter using the Chao diversity estimate of the total number of taxa in both seasons. This indicates that those individuals in the exposed cluster have fewer taxa detected. Child subgroup analysis of PC1 clusters still showed significant (p-value~0.01) reduced diversity in the winter Azinphos-methyl exposed cluster but is suggestive in the spring-summer (p-value~0.09), see Figure 3-9. This Chao diversity analysis shows the unexposed cluster has more OTU diversity and supports the increase of the OTU loadings in the PC1 direction of the unexposed cluster also (See Figure 3-8).

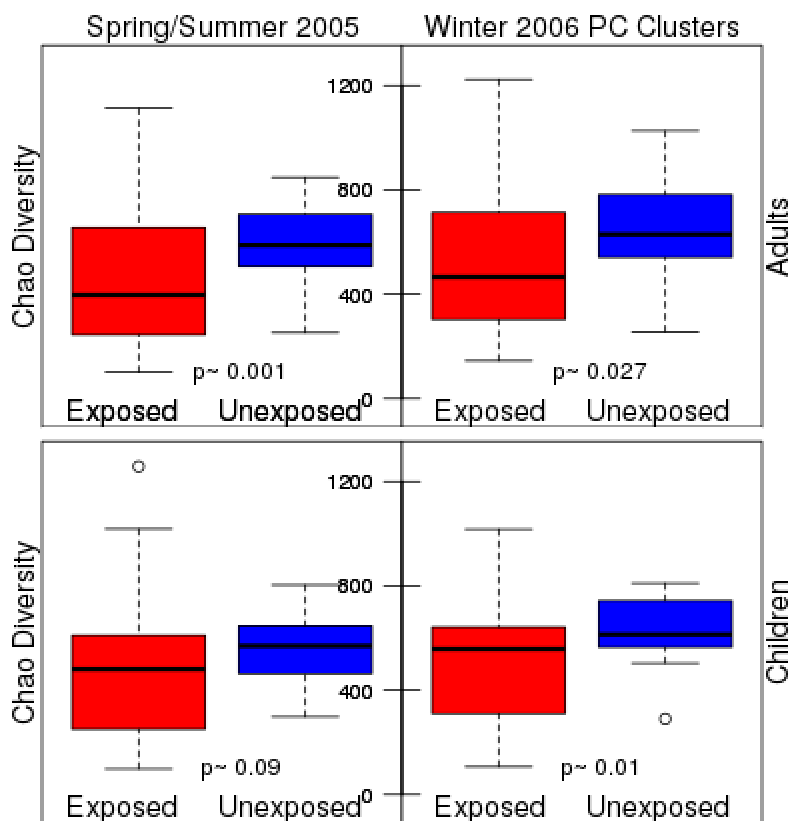


Figure 3-9 Adult and Child Alpha Diversity by PCA Cluster

Summer and winter exposure Gaussian modelled clusters by the Chao diversity of children (bottom panels) and adults (middle panels) and combined (top panels). PC clusters are for Spring-summer 2005 and Winter 2006 clusters as shown in Figure 3-8. Both seasons show a significant decrease in the diversity in the exposed PC cluster by Welch's T-test indicating fewer taxa detected.

3.5. Discussion

This analysis suggests that organophosphate pesticide exposure can be a large contributing determinant of the microbiome composition as seen in a human agricultural cohort.
Our results are consistent with a growing body of evidence currently published in five manuscripts that other related organophosphate pesticides, also induce perturbations in the commensal microbiota of experimental animals [90-94]. One of these replicated their observation of changes in the rat gut microbiome in an *in-vitro* bioreactor [91] suggesting a

host independent antibiotic mechanism of action for organophosphates. Four of the reports used rodent *in-vivo* gut microbiome model systems [91-94] and one used the Japanese quail (*Coturnix japonica*) [90]. Most interestingly these orthogonal studies also identified members of the order *Lactobacillales* as being reduced under organophosphate pesticide exposure conditions. *Streptococcus* and *Granulicatella* are common oral buccal members of the *Lactobacillales* and these taxa we consistently detected as reduced in our agricultural cohort.

There is a high percentage of co-clustering of children with their household adult in the two respective seasons. In the exposed cluster ~95% adult-child household pairs co-cluster and similarly in the unexposed ~95% of buccal microbiomes co-cluster. This suggests that shared environmental exposure risks are similar among cohabitating household adults and children. Those children that live with an adult farmworker may have enough take-home pesticide exposure to associate their microbiome composition with their respective household adult. The one child in the spring-summer and two children in the winter with altered microbiomes in the exposed clusters who have their paired household adult clustering with normal microbiomes may have received enough community encountered exposure to pesticides to modify their oral-buccal microbiome composition. We postulate that chronic low-dose home exposure to these adults was insufficient to modify their microbiome composition but their children somehow have encountered enough pesticide exposure to have the associated altering effect. Previously, pesticide exposure of these children has been documented by urinary and blood biomarkers with exposure likely occurring via dust transported home on clothing and dual usage of semi-closed environments shared by households. These semi-closed environments include family cars and trucks used for child transport and the work commute thus providing pathways of potential pesticide exposure to children [11, 14]. The association pattern observed, where the child microbiomes have a dysbiotic state with low PC1 scores and higher variability in PC2 like their household adult guardian, suggests that they have similarly modified microbiomes associated with exposure to agricultural pesticides. This association is stronger in the summer when the agricultural exposure to pesticides is the greatest. Our studies show the minimally invasive buccal swab appears to serve as a biomarker of Azinphos-methyl exposure. This susceptibility of the microbiome may be particularly useful to the larger epidemiological endeavor of understanding the disease burden associated with pesticide exposure. It is more difficult to obtain temporally rich blood biomarkers of exposure than a buccal swab.

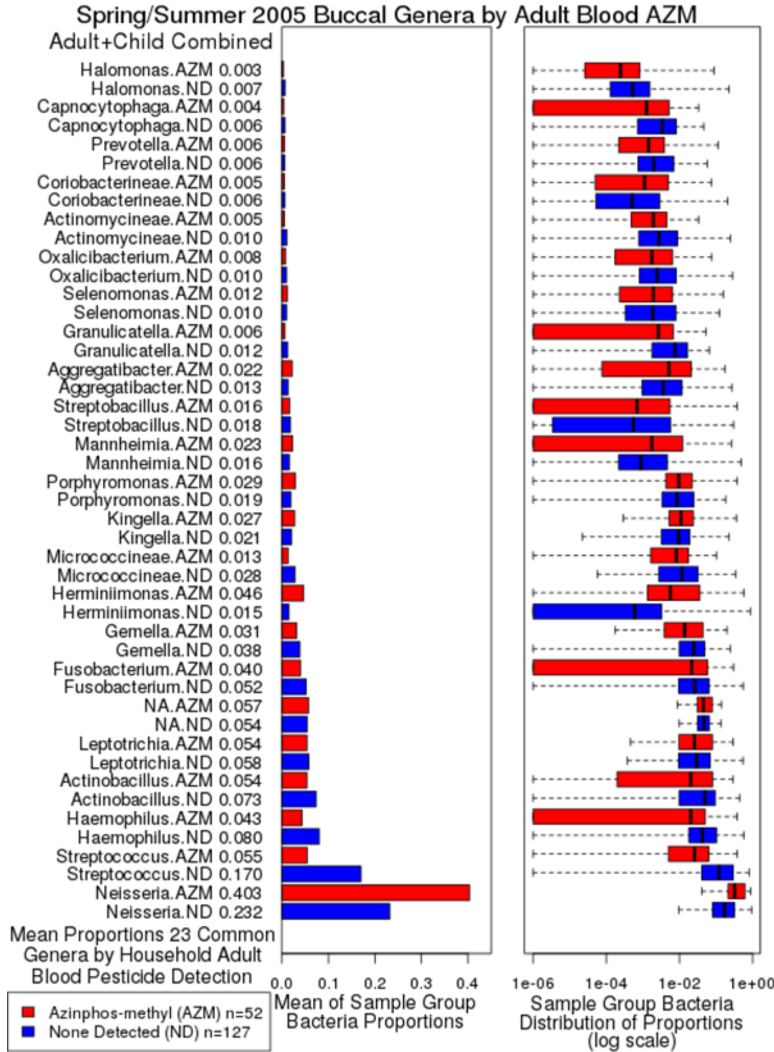
The two most prominent bacterial genera (*Streptococcus* and *Neisseria*) in our agricultural cohort microbiome study are significantly perturbed. There are many commensal organisms beneficial to host life in the order *Lactobacillaes*. As an example, strains of *Lactobacillus plantarum* help to maintain normal growth in chronically underfed infant mice [135]. Reductions in *Streptococcus* could also have beneficial outcomes due to the known pathogens in this genus. But these kinds of conclusions can be wrong due to the diversity of beneficial and benign organisms that are also among taxon. Interpreting the potential health impacts for the oral microbiome are complex. More studies with health outcome data are needed to support any conclusions of this nature. *Neisseria* is known to be a large taxa with many non-pathogenic members and therefore an increase may not be associated with or causal of adverse health outcomes [88]. These mixed pathogenicity profiles make drawing

specific conclusions very difficult at this taxonomic resolution. What is more compelling to use in accounting for potential disease impacts is the altered microbial interactions effect on many other metabolism and overall metabolic outcomes in the systems biology context [86]. Indeed there appears to be a replicable microbiome phenome that is associated with environmental exposures in many contexts [32-39, 136] and these finding should be followed up with more detailed health outcomes.

The associated altered microbiome effects on life course may be compounded in children due to the developmentally immature nature of their gastrointestinal tract, immune systems, and overall bodies. Dysfunctional outcomes of child development that can result from the immune system could be large and diverse. In humans the microbiome is established by the age of two [128], but child development is still occurring through juvenile stages into adulthood which gives many opportunities to affect the development of the microbiome [130]. The consequences of not having proper signaling with a normal commensal immune system have been documented in a few model systems. Activations of receptors by bacterial ligands have been shown to be important for normative immune cell maintenance and proliferation [137]. Microbiome associated states induced by continued antibacterial challenges [138] have the potential to induce disease-related dysfunction phenotypes [139]. Given this and the known microbiome connections to periodontal disease, obesity, cardiovascular disease, and diabetes [111-124], the long-term child developmental outcomes of these pesticide exposures on this community could compound risk factors with many common diseases. Hispanic adults also carry a disproportionate disease burden of diabetes and cardiovascular disease [102-110], both phenotypes that are connected to the metabolic dysfunction endophenotype which has been shown to be inducible by the microbiome. The collection of longer-term information on glucose-insulin regulation, triglyceride levels, and cardiovascular-related disease biomarkers in this cohort would be a logical next step to investigate the impacts of these observations. Additionally, dental records should be assessed and comparedd to incidence of dental decay and periodontal disease and if changes in dental health are evident. Dysbiosis with unstable community structures may allow altered microbial populations to persist in places like the oral, nasal, and gut contexts. Pesticide exposure and alterations of the microbiome could have multivariate causal pathways of influence on pleotropic common and rare diseases. These include systemically linked diseases (i.e., diabetes, cardiovascular disease, neurological outcomes, periodontal disease, etc.) that the microbiome appears associated with and can be a contributing factor [111-124]. We have shown that exposure to Azinphos-methyl is associated with alterations in microbiome composition. However longer term health outcome studies will be needed to ascertain the health impacts the pesticide associated microbial alterations will influence on the wellbeing of these agricultural community individuals and their children.

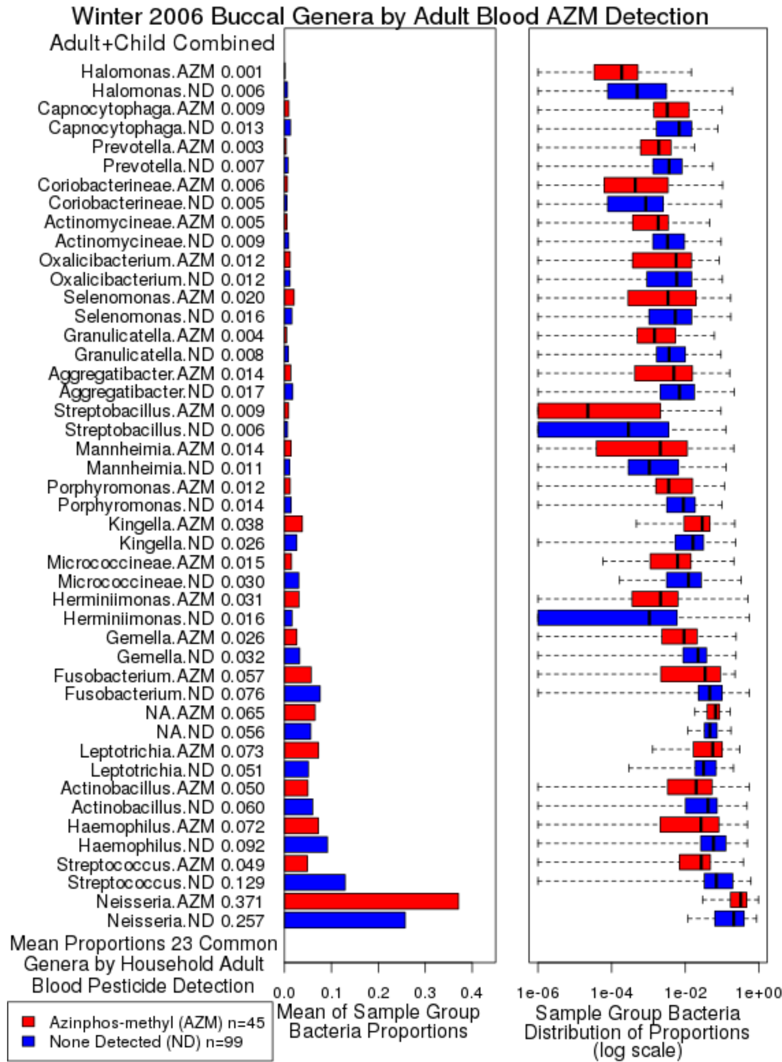
3.6. Chapter 3 Supplemental Figures

Figure 3-10 Spring/Summary 2005 Combined Adult and Child Buccal Genera by Adult Blood AZM



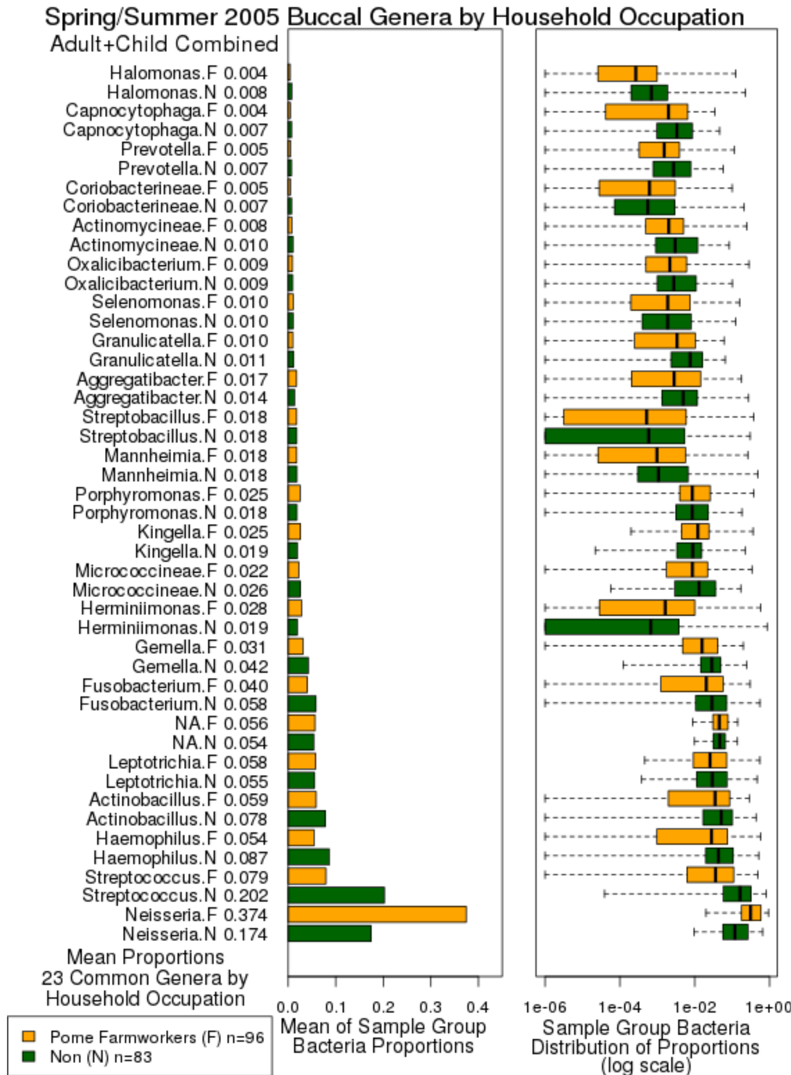
The combined adult and child spring-summer 2005 buccal microbiomes grouped by household adult Azinphos-methyl blood detection versus none detected. Wilcoxon's Rank Sum test found associations (FDR <0.1) to the genera *Streptococcus, , Herminiimonas Haemophilus, Granulicatella, Micrococcineae, Gemella, Capnocytophaga, Actinobacillus,* and *Halomonas* (See Table 3-1) among the 23 (>0.5% mean proportion) common genera tested.

Figure 3-11 **Winter 2006 Combined Adult and Child Buccal Genera by Adult Blood AZM Detection**



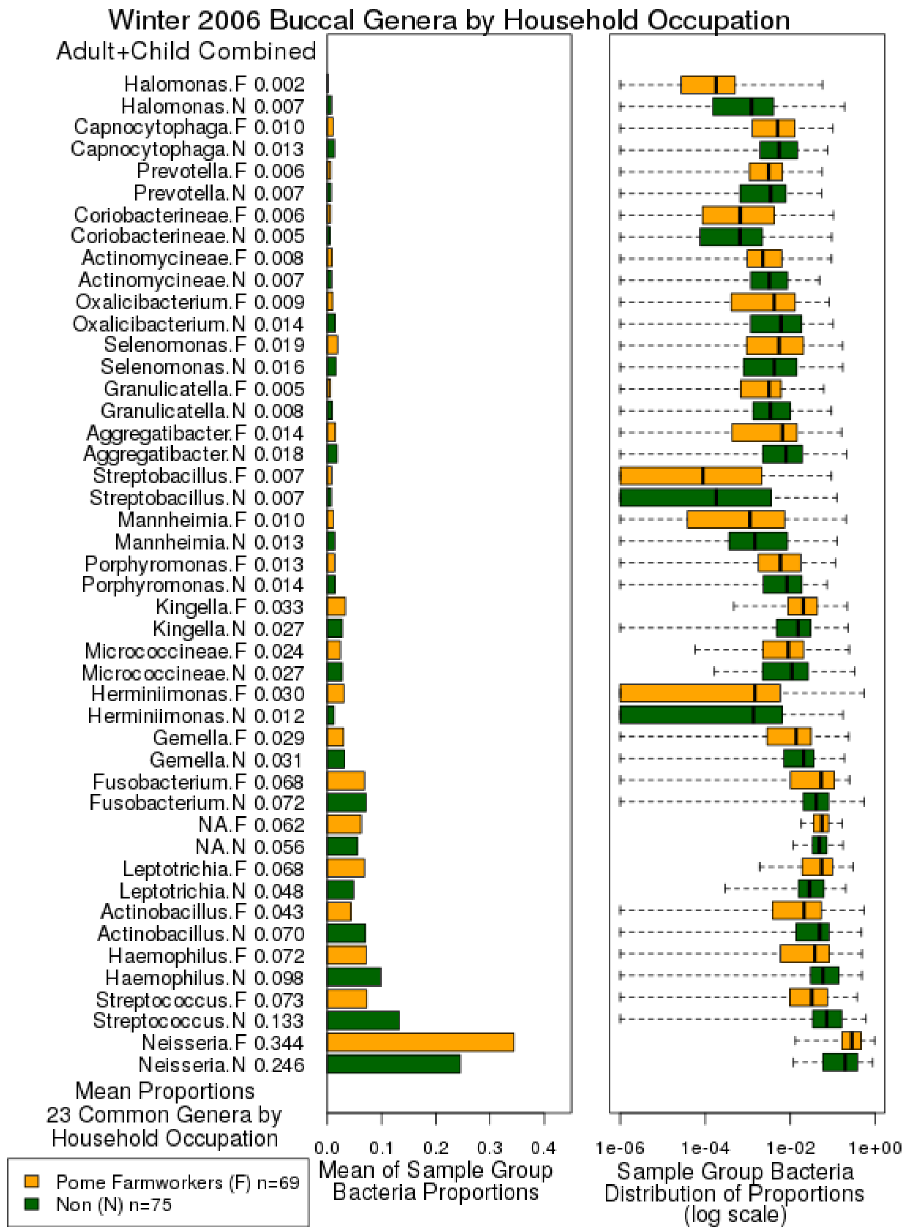
The combined adult and child winter 2006 buccal microbiomes grouped by household adult Azinphos-methyl blood detection versus non. Wilcoxon's Rank Sum test found associations (FDR<0.1) to the genera *Streptococcus, Gemella, Actinomycineae, Halomonas, Granulicatella, Prevotella, Haemophilus,* and *Micrococcineae* (See Table 3-2) among the 23 (>0.5% mean proportion) common genera tested.

Figure 3-12 **Spring-Summer 2005 Combined Adult and Child Buccal Genera by Household Occupation**



The combined adult and child spring-summer 2005 buccal microbiomes grouped by household occupational status for pome fruit farmworker versus non. Wilcoxon's Rank Sum test found associations (FDR <0.1) to the genera *Streptococcus, Neisseria, Haemophilus, Gemella* and *Micrococcineae* (See Table 3-3) among the 23 (>0.5% mean proportion) common genera tested.

Figure 3-13 **Winter 2006 Combined Adult and Child Buccal Genera by Household Occupation**



The combined adult and child winter 2006 buccal microbiomes grouped by household occupational status for pome fruit farmworker versus non. Wilcoxon's Rank Sum test found associations (FDR <0.1) to the genera *Streptococcus* and *Halomonas* (See Table 3-4) among the 23 (>0.5% mean proportion) common genera tested.

4. MINOR ALLELE CONTRIBUTIONS TO CHOLINESTERASE ENZYMATIC ACTIVITY VARIABILITY IN ORGANOPHOSPHATE PESTICIDE EXPOSED FARMWORKERS

Acetylcholinesterase and Butyrylcholinesterase Activity Minor Allele Carriers: Additive Gene and Environment Effects on Blood Cholinesterase in an Azinphos-methyl Exposed Agricultural Cohort.

Ian B. Stanaway, James C. Wallace, Ali Shojaie, William C. Griffith,
Sungwoo Hong, Tomomi Workman, Eric M. Vigoren, Beti Thompson, and Elaine M. Faustman

4.1. Abstract

Minor alleles in gene product enzymes are often associated with loss or reduction in the functionality of enzyme activity. We investigate the minor allele contribution to lower blood cholinesterase activities of the enzymes encoded by the genes acetylcholinesterase (AChE) and butyrylcholinesterase (BChE). The pesticide Azinphos-methyl and other similar organophosphates have an acute mode of action for toxicity via inhibition of AChE activity and altering neuronal synapse function. Both AChE and BChE gene products are expressed in the blood and are well studied in a toxico-pharmacological perspective of medicine and organophosphate pesticide exposure [140]. The hypothesis where these enzymes may have gene by environment interactions with certain organophosphate pesticides (and pharmaceuticals), which are dependent on minor allele contributions, has been posited for some time and shown in animal dosing models, but never convincingly shown in a pesticide exposed human population [140]. We relate the genetic minor allele significance to protection of susceptible populations by showing carriers of minor alleles in both AChE and BChE genes are also associated with lower blood cholinesterase enzyme activities and this genetic component is additive with the Azinphos-methyl exposure.

4.2. Introduction

Blood is an important diagnostic phenotype sampling matrix for biomarker based evaluation of human health related impacts. Mass spectrometry methods have catalogued thousands of protein moieties in blood plasma (http://www.plasmaproteomedatabase.org/). Given this, there are many complex phenotypes with biomarkers displayed in blood from which we may have diagnostic and hopefully predictive value for delivering higher standards of health care and preventative medicine. However, blood traits like any other phenotype are subject to both genotype and environmental exposure effects. In Washington State, the Department of Labor and Industries requires that blood cholinesterase activities of agricultural pesticide handlers be monitored for surveillance of worker-related exposure to Organophosphate Pesticides (OPs) (http://www.lni.wa.gov/Safety/Topics/AtoZ/Cholinesterase/). This has been instituted by state agricultural regulators as a diagnostic measure to prevent excessive occupational pesticide exposure. Two homologous genes confer the phenotype of blood

cholinesterase activity: the erythrocyte membrane bound acetylcholinesterase (AChE) and the free plasma enzyme butyrylcholinesterase (BChE) which is excreted by the liver into the blood. AChE is important for most higher organisms with enzyme activity in the neuronal synapse and it degrades the neurotransmitter acetylcholine (ACh) allowing a synapse to fire again with each successive wave of ACh. It is likely an evolved response to natural toxins which allowed the enzymes coded by BChE and AChE gene products to rise to fixation as expressed blood enzyme activity traits. These genes have many other exogenous substrates, some of which overlap in the activity profile between the two enzymes and some which do not. The substrates of BChE include muscle relaxants and several pharmaceuticals [140] which makes having an understanding of the AChE-BChE gene system important to the practice of medicine. Carriers of 'silent' alleles have been detected in the human population which make them susceptible to paralysis from some muscle relaxants (i.e. succinylcholine) but are quite rare in the population overall, and homozygous carriers of the silent alleles appear to be normal otherwise, while missing this activity [140]. Azinphos-methyl is an organophosphate pesticide that inactivates both AChE and BChE by covalently binding a serine residue in the active site. Inhibition of activity is defined as the fraction of the normal activity or abundance of cholinesterase that has been inactivated. Cholinesterase inhibition by pesticides would logically be larger in lower activity genetic variant carriers at the same dose as there is less active enzyme to buffer the central nervous system from the absorbed dose. This would effectively lower the acceptable safe dose for those individuals with lower activity. Rare genetic variants with increased likelihood of change of function in the gene enzyme products is likely to result in decreased functional activity [21]. The normal human whole blood activity of BChE is ~307x that of AChE activity by volume [140] giving heterozygous loss of BChE function the potential ability to limit activity down to ~153x of AChE. This would reduce the buffering ability of BChE at protecting the enzyme activity of AChE. Rescue of severely organophosphate poisoned individuals and test animals has been demonstrated by intravenous dosing of many different recombinant and purified preparations of BChE, effectively increasing their available buffering activity [140]. These direct clinical and experimental efficacy validations show that the available gene-activity dose of plasma BChE effects the toxicity experienced by the test subjects. One would expect to find cholinesterase alleles with differential susceptibility to toxicant induced AChE activity inhibition due to heterozygous reductions in activity. Pesticide-genetic interactions with cholinesterase activity have been documented. A study [141] of carriers of AChE regulatory transcription factor binding site variants show genetic minor allele carriers can have higher static cholinesterase activity, but one more acute sensitivity to dosage. The regulatory site genetic variation breaks the regulatory binding site enzyme activity induction capability when exposed to an anticholinesterase, resulting in an enzyme activity crash due to loss of transcription [141]. This illustrates the non-linear non-monotonic responses that can be mixed in a population of individuals' responses. Due to this mixed phenotype profile, we chose not to investigate regulatory non-coding alleles in our analysis, which also are likely to confer gain of function phenotypes by increasing the gene dosage. This bi-directional and non-monotonic possibility of mixed population phenotypes with regulatory alleles such as these would lead to a loss of power, especially at the samples sizes at which we are investigating. Thus, we have enriched the test population of variants for loss, reduction, and change of function protein coding alleles by only testing the variants in exons and splice site gene regions. This is a similar strategy as used in exome sequencing for monogenic

disorders [142]. In this report, we present an analysis of an agricultural cohort from the Yakima Valley of Washington State. Blood collections of the subjects were used to detect the organophosphate pesticide Azinphos-methyl and the phenotypes of interest: blood cholinesterase activities. We aimed to demonstrate an additive candidate genotype by environment association to blood cholinesterase enzyme activity showing both environmental and genetic minor allele components are significant.

4.3. Methods

4.3.1. Study Design, Participants and Agricultural Setting

In 2005, the participants of our study were recruited among 200 Yakima Valley agricultural community households. Recruitment included one child and one adult guardian from each household. Written and informed consent was obtained from each adult. The Fred Hutchinson Cancer Research Center Institutional Review Board approved the study collection procedures (File IR 5946). The study design is a cross-sectional longitudinal two timepoint cohort. The initial well studied aim was to assess potential pesticide exposures of pome fruit orchard workers and their children by surveying 100 farmworker adult-child pairs in comparison to 100 matched control non-farmworker adult-child households [1, 2, 4-6, 9-15]. Timing of the first sample collection phase was synchronized with the local commercial agricultural applications of Azinphos-methyl pesticide to control Coddling Moths *(Cydia pomonella)*. In the same pesticide spraying season is the fruit thinning season when pome farmworkers thin the excess fruit and trim branches giving them the potential for exposure to pesticide residues in the orchard. This began in the Spring (April) spanning to Summer (July) 2005. The second sample collection for the same individuals occurred in the Winter time spanning December 2005 to early March 2006 when pesticides are not normally used and environmental degradation would have reduced the likelihood of exposure. Whole blood from the adults was collected in these two distinct agricultural seasons (1) Spring/Summer 2005 and (2) Winter 2006. Demographic variables (gender, age, ancestry, occupation) were self-reported by the participant subjects. For the analyses in this manuscript adult blood was used to assess organophosphate pesticide exposure and the subjects' genotypes to investigate the association to blood cholinesterase activities of AChE and BChE with gender and age as covariates.

4.3.2. Blood Azinphos-methyl Pesticide Concentration

As described by Barr et al. [19], the blood sample Azinphos-methyl concentrations were determined by mass spectrometry (GC-HR-MS). The limit of detection is 0.04 ng/mL plasma.

4.3.3. Blood AChE and BChE Activities

Erythrocyte AChE and plasma BChE blood cholinesterase enzyme activities were measured using the Testmate Kit from EQM research (Blood Cholinesterase Testing Kit Manual http://www.eqmresearch.com/Manual-E.pdf) which utilizes the method of Ellman et al [18]. AChE and BChE enzyme activities are respectively expressed in the units of activity Units / gram of hemoglobin (U/g Hgb) or activity Units / mL of blood plasma (U/mL plasma) [143].

4.3.4. Genotyping

Genotyping of individuals was performed mainly on blood (and in a few cases, buccal swabs) DNA extracts with the Illumina 5M genotyping array chip (www.illumina.com) at the North-West Genome Center situated in the University of Washington's Department of Genomes Sciences. The design of the genotype site selection for this array was based on variant discovery from deep exome and genome sequencing of the collaborative efforts of the 1000 Genomes Project [144-147] and the ~12,000 individuals sequenced by the Exome Sequencing Project [148-154]. These two projects provide population based coverage of both common and rare genotype variants. Prior to these genotype discovery efforts, most array genotyping platforms only provided a selection of common genetic variants that were used for genetic association to phenotypes. Rare and common variant site selection on recent Illumina arrays is informed by a minor allele population genetics model. The explosive increase in the human population in recent history has increased the number of rare variants in the human population [21]. This has increased the burden of rare variation that has not been under purifying selection due to the recent *de-novo* mutations introduction to the population. These rare variant alleles are more likely to have a change of phenotype associated with them. This is of interest in susceptible populations for disease risk. In the context of environmental exposure, genetic variants are not likely to be deleterious unless the carrier is exposed to a condition where they may not metabolize an exogenous (or endogenous) compound as efficiently. This may make the carriers of minor alleles more susceptible to the toxic effects of said compounds. Due to this, we selected non-monomorphic exonic variants with minor alleles seen in our cohort in the two candidate genes (AChE and BChE) which interact with this pesticide from among the Illumina 5M genotyping array variants. We then tested for additive gene + environment associations to the enzyme activity phenotypes of their own coding products. All exonic minor alleles included are disease associated, non-synonymous or at splice sites in the mRNA gene model. Monomorphic sites with only one allele detected were not considered due to lack of variation detection. Minor alleles are defined by the population frequency of the less common nucleotide DNA base at a variant locus. In diploid organism with two of each chromosome like humans the minor allele frequency takes a range of 0-0.5 and is calculated as presented in Equation 1.

Given:

Count Allele A >= Count Allele B

Minor Allele Frequency = Count Allele B / (Count Allele A + Count Allele B) Eq. 1

The presence of minor alleles was used to categorize the individuals to chromosomally phased homozygous minor allele carriers (2 chromosomes with minor alleles in the gene), heterozygous major-minor allele carriers (1 chromosome with minor alleles in the gene) or major allele homozygotes (0 chromosomes with minor alleles in the gene) for each gene. This allowed testing of the joint effect of multiple low frequency genotypes in carriers (1 or

2) versus major alleles homozygotes (0). If variants had been evaluated individually it is likely that tests would be underpowered at this sample size (~100+ subjects) if the rare variants were considered individually. This variant burden approach is a strategy that has been utilized in the testing of rare variants to phenotypes in genome wide scans [155-159] which we are adapting to this candidate gene + environment association to their own gene products enzyme activity phenotype with alleles already associated with loss of function.

4.3.5. Statistical and Numerical Methods

BChE and AChE enzyme activities, age and blood Azinphos-methyl were treated as continuous variables. AChE and BChE activities and blood Azinphos-methyl concentration were log transformed. A value of one was added to each Azinphos-methyl concentration before log transformation to include the non-detects which would be undefined with zero log values. These transformations were applied to normalize the linear model residuals. The a genome wide selection of genotypes of the assayed individuals were input to the R package SNPRelate [160] and PRIMUS [161] to identify related individuals and any non-Hispanic population stratification which may skew allele frequencies. The sites on the genotyping array which were 10 kilobases flanking to the ChE genes were input to PHASE [162] to haplotype the variants and determine the carrier burden of minor alleles. Singleton variants with uncertain phases were assigned to the haplotype which PHASE selected in the output. From the maximal unrelated set of Hispanic individuals, the minor allele frequencies were calculated and the carriers of minor allele coding variants were individually coded as "0" “1” or "2" based on the chromosome haplotype phased alleles in the gene coding region. Gender was coded as "Male" or "Female". Analyses were conducted by gene (AChE and BChE) and season (Spring/Summer and Winter) for four independent generalized linear models. Within the gene and season an individual subject was removed from analysis if they were missing one of any of the variables. The model relationship used in analysis via the statistical R language call lm() [163] is presented in equation 2.

log(Gene Enzyme Activity) ~ Age + Gender + log(1+Azinphos-methyl) + Gene Minor Allele Carrier Count (0, 1, 2)

Eq2.

To control for outliers and non-normal residuals we used the ‘sandwich’ estimator vcovHC() with the HC4 estimator function methods on the fitted models with the coeftest() R function to correct the standard errors for robust p-value estimation [164, 165]. We also used the gls() function of the nlme R package [166] to fit a time series model with the AR(1) covariance structure by subject across seasons. The initial gls() regression model included the variables in Eq. 3 before variable selection.

log(Gene Enzyme Activity) ~ Age + Gender + log(1+Azinphos-methyl):Season + Gene Minor Allele Carrier Count (0, 1, 2)

Eq3.

Among the subgroup of Azinphos-methyl exposed individuals with detected blood concentrations, we assessed the BChE to AChE gene-gene blood cholinesterase activity interaction by measuring the difference in inhibition of AChE conferred by the available

activity of the generally more abundant BChE (~307x) by volume in the blood. Since both enzymes are inactivated by Azinphos-methyl and BChE is ~307x AChE activity in whole blood, it is expected that BChE enzyme activity will buffer AChE enzyme activity. An Azinphos-methyl exposed individual with less basal BChE activity, may be at risk of more AChE inhibition (less activity) compared to their normal basal AChE activity. We interpret the blood AChE biomarker as a risk marker of the inhibition the central nervous system would exhibit. We used the ratio of the exposed spring-summer samples' AChE activity values divided by the same individual's winter unexposed value to determine a fractional ratio (inhibition) of their basal AChE activities. The one-sided Pearson's Correlation test was used with the ratio of the Azinphos-methyl exposed to unexposed AChE activities in comparison to three measures of the BChE activity: (1) the total BChE activity while exposed in the spring-summer sample, (2) the total BChE activity while unexposed in the winter and (3) the exposed/unexposed BChE activities ratio. We chose these three measures of the BChE activity phenotype to balance the opportunity for the AChE inhibition phenotype tests to differently weight the somewhat inseparable effects of gene and environment. The exposed/unexposed BChE activities ratio is more driven by exposure due to the inhibition, while the total Azinphos-methyl exposed BChE activity will have more weight on the genetics. The comparison of AChE inhibition to only the basal winter BChE activity may provide the best control of the Azinphos-methyl exposed AChE inhibition interaction with the much more available (~150-300x) BChE buffering activity. The expectation is that lower BChE enzyme activities and minor alleles in the BChE gene will be associated with more inhibition of AChE activity. This expectation is supported by animal model tests where active exogenous BChE enzyme administered intravenously rescues severely poisoned animals [140] and where transgenic mice with higher BChE enzyme activity are resistant to organophosphate poisoning [167]. We test this in a linear regression model presented in Eq. 4 investigating the difference of AChE and BChE activity inhibition between allele carriers of 1 and 2 phased gene copies in AChE and BChE and those homozygous (0) with the common major allele haplotype with co-variates age and gender.

log(Ratio Summer ChE Activity / Winter Activity) ~ Age + Gender + log(1+Azinphos-methyl) + AChE Gene Minor Allele Carrier Count (0, 1, 2) + BChE Gene Minor Allele Carrier Count (0, 1, 2)

Eq4.

We checked final models against the null (Y~1) to show they were significant and report unadjusted slopes and un/adjusted errors, p-values and r^2 values of these models. The base R function anova() was used with initial unadjusted and final adjusted models to calculate the variance attributed to each variable based on the residuals of the sum of squares. We are interested in the genetic effects of these two enzymes variants on their own activities which are known to be a target of the Azinphos-methyl pesticide. Therefore, the sandwich adjusted variable selections for the final models included the genetic component regardless of its initial significance in addition to any other independent variables that showed significance in the first regression. The site selection of many of these alleles have been shown in other contexts to have change of phenotype in activity [168-170] and the remainder have associated disease [171, 172], genome placement [173, 174] and protein changing biochemical considerations [175] which make them worth collective investigation of blood

phenotype. This prior knowledge on their activity affects the interpretation of significance and the false discovery nature of this study. Since we are investigating alleles regression result at two loci with an expected environmental component pesticide known to target these genes, we elected to implement corrections for multiple comparisons for two genes using the Benjamini-Hochberg false-discovery-rate (FDR) methods [62]. In a corresponding gene-model regression result we submitted the sandwich corrected p-values of the two genes together to the R function p.adjust(). In example, FDR correction is applied to the spring-summer AChE minor allele independent variable result with the spring-summer BChE minor allele independent variable result. In the ChE enzyme activity inhibition models where we are interested in genetic effects of BChE minor allele carriers on inhibition of AChE and BChE enzyme activity, but not the effect of AChE minor allele carriers on BChE enzyme inhibition due to the 307x whole blood stoichiometric differences, we corrected for three tests: AChE inhibition ~ AChE alleles (#1) + BChE alleles (#2) and BChE inhibition ~ BChE alleles (#3).

4.4. Results

4.4.1. Genotype Array Variation

Using the genotypes of the 194 individuals assayed on the Illumina genotyping array, the R package SNPRelate [160] and PRIMUS [161] we identified a maximum set of 164 unrelated individuals that cluster with 1000 Genomes Hispanic reference genotypes [146]. We observed the AChE gene exonic and splice-site regions on the array contained eight variants. Of these five were monomorphic and three AChE variants we observed minor alleles. The BChE gene has 10 exonic or splice site variants on the array, five of these were monomorphic and five BChE coding variants have minor alleles. Table 4-1 provides a summary of the minor allele genetic variant set and cohort minor allele frequencies used in this study of 156 unrelated individuals of the with passing genotype data and references to the associated know loss of function reports in the literature.

Table 4-1 AChE and BChE Protein Coding Minor Alleles Tested in the Gene + Environment Model

Gene Variant ID	Chr: Position Hg19	Major(n)	minor(n)	Cohort-Frequency (%)?	Function	Notes	Refs.
BChE rs1803274	3:165491280	C(270)	T(42)	0.135	Non-synon	70% normal activity	[168]
BChE rs114664527	3:165503930	C(311)	T(1)	0.003	Splice	Within the Spliceosome	[174]
BChE rs28933390	3:165547569	C(311)	A(1)	0.003	Non-synon	Flouride-resistance	[170]
BChE rs1799807	3:165548529	A(305)	G(7)	0.022	Non-synon	50% normal activity	[169]
BChE rs1126680	3:165555125	G(303)	A(9)	0.029	Non-synon	Missense Early Start Codon	[174]
AChE rs1799806	7:100488658	G(242)	C(70)	0.22	Non-synon	Lower Activity & Cardio Traits	[172]
AChE rs7636	7:100490077	G(297)	A(15)	0.048	Synonymous	Diabetes Risk Allele	[171]
AChE rs1799805	7:100490797	C(303)	A(9)	0.029	Non-synon	Creates Alternate Glycosylation	[175]

The BChE gene variant rs1803274 and is a known non-synonymous minor allele with 70% of the normal activity [168]. The rs114664527 BChE gene variant is in the footprint of the spliceosome [174] so may produce alternately spliced exons with altered activity. The BChE gene variant rs28933390 is the Fluoride-2 non-synonymous variant associated with fluoride resistance [170]. The BChE gene rs1799807 non-synonymous variant is the 'atypical' variant with 50% of normal activity [169]. The BChE gene rs1126680 non-synonymous variant minor allele makes an early missense start codon in the 3' untranslated exon [174]. The AChE gene rs1799806 non-synonymous minor allele is known to have lower activity as found in a study of cardiovascular traits [172]. The AChE gene rs7636 is synonymous for the same amino acid but is in the coding region and has been associated with diabetes risk [171]. The AChE gene rs1799805 non-synonymous minor allele is responsible for the creation of an alternate blood group antigen on the surface of the enzyme [175]. Phasing of chromosomes revealed n=83 major allele homozygotes, n=61 major-minor allele heterozygotes and n=12 minor allele homozygotes of the AChE gene. Phasing of chromosomes revealed n=114 major allele homozygotes, n=39 major-minor and n=3 minor allele homozygotes of the BChE gene. These counts show these rare variations are collectively common in the population.

4.4.2. Blood Phenotypes

The blood erythrocyte AChE activity phenotype had a range of ~20-45 Activity Units/gram of hemoglobin. Histograms of spring through summer and winter assay results are in Figure 4-1 with the other study gene enzyme activities, gene allele counts and Azinphos-methyl variables of the unrelated subjects.

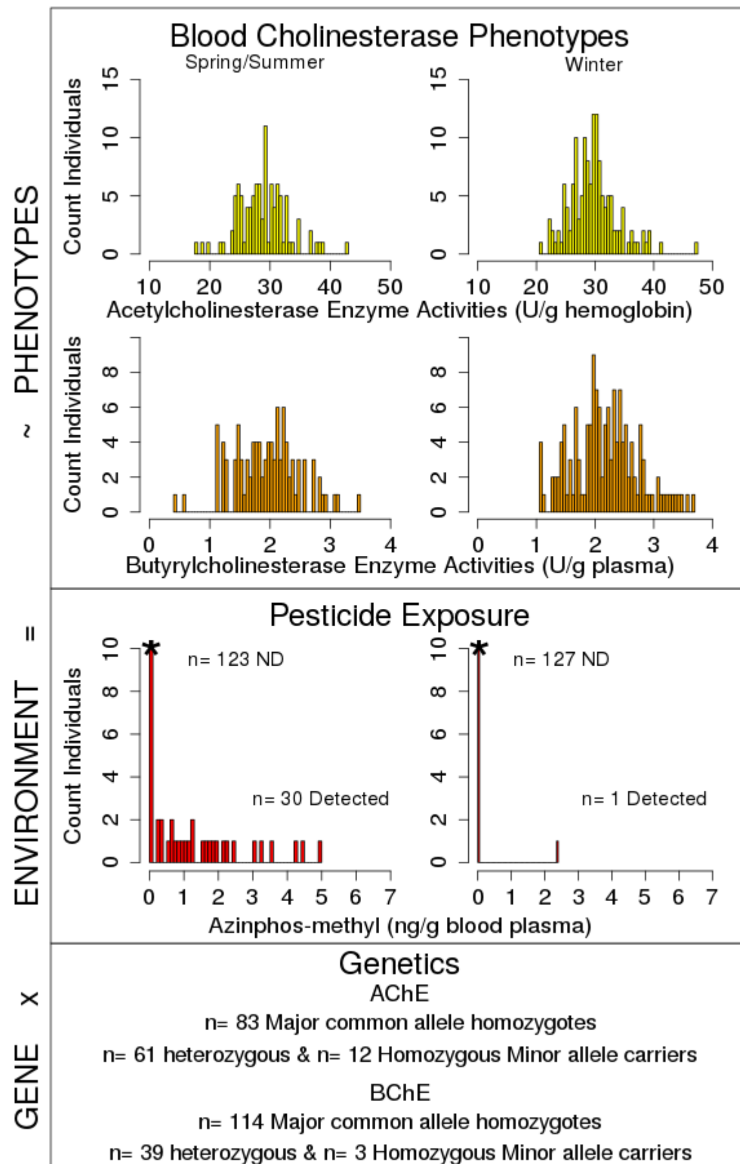


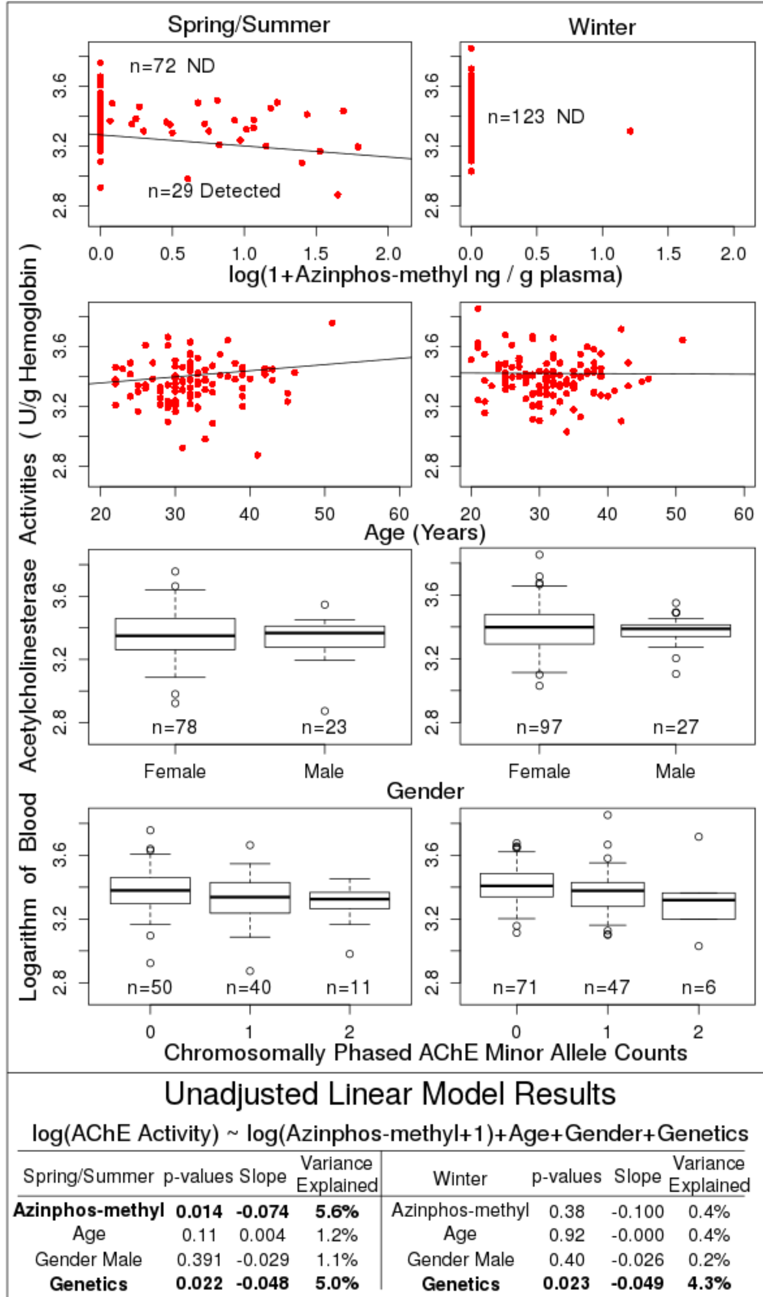
Figure 4-1 Blood Assay Results for Spring/Summer 2005 and Winter 2006

The top four panels are the blood cholinesterase enzyme activity blood phenotypes conferred by the AChE and BChE genes. The middle panel is the exposure to the organophosphate pesticide Azinphos-methyl measured by the blood concentration. The bottom panel reports the counts of carriers of genetic minor alleles in these two genes that may lead to a reduction in the activity of their enzyme products shown in the top panels. All panels only show the results of unrelated individuals.

The blood plasma BChE enzyme activity phenotype had a range of ~0.5-3.8 Activity Units / gram of plasma. We do not detect any apparent 'silent' BChE enzymes as all individuals have activity. Azinphos-methyl was detected in the blood of 30 individuals above the limit of detection (0.04 ng/g) in the spring-summer and one individual in the winter as the measure of environmental exposure to this pesticide. There were 123 non-detects in the spring-summer and 127 assayed non-detects in the winter blood collections. There were 83 major allele homozygotes, 61 major-minor allele heterozygotes and 12 homozygous carriers of the chromosomally phased minor alleles in the AChE gene coding region. The BChE gene had 114 major allele homozygotes, 39 major-minor allele heterozygotes and three minor allele homozygotes in the coding region. (See Figure 4-1, bottom panel)

4.4.3. AChE Enzyme Activity Linear Models

Sample attrition due to removal of related individuals and missing variables allowed the inclusion of 101 subjects in the spring-summer AChE model and 124 subjects in the winter model. Azinphos-methyl exposure is associated with lower activity in AChE and explained ~5.6% of the variance in blood AChE activity in the spring-summer when pesticide was detected in the blood. It was not associated in the winter sampling when there is little (n=1) exposure. See Figure 4-2 for model activity factored plots and unadjusted coefficient results.



Spring/Summer	p-values	Slope	Variance Explained	Winter	p-values	Slope	Variance Explained
Azinphos-methyl	**0.014**	**-0.074**	**5.6%**	Azinphos-methyl	0.38	-0.100	0.4%
Age	0.11	0.004	1.2%	Age	0.92	-0.000	0.4%
Gender Male	0.391	-0.029	1.1%	Gender Male	0.40	-0.026	0.2%
Genetics	**0.022**	**-0.048**	**5.0%**	**Genetics**	**0.023**	**-0.049**	**4.3%**

Figure 4-2 AChE Blood Phenotype Gene + Environment Generalized Linear Model

The Gene + Environment Acetylcholinesterase Blood Phenotype Generalized Linear Models. The top panels show regression of the AChE enzyme activity by Azinphos-methyl blood pesticide concentration. The next panels depict AChE enzyme activity by age, followed by gender and chromosomally phased genetic minor allele carrier counts with the unadjusted p-values and variance explained in the bottom of the figure. The middle lines in the boxplots are the median. The boxes are the middle 50% of the data with the edges as the 25th and 75th percentiles. The whiskers are the interquartile range past the percentiles and the circles are the extreme values.

Unadjusted and sandwich standard error interval estimator adjusted and FDR-adjusted results showed carriers of minor alleles in the AChE gene are significantly (FDR < 0.1) associated with lower enzyme activity in the spring/summer and winter blood collections (Figure 4-2 & Table 4-2).

Table 4-2 Summary of AChE Activity Regression Results by Season

Table 4-2	Summer AChE Activity										
	Unadjusted Linear Model					Sandwich Adjusted Model With Selected Variables					
Variables	Estimate	Std. Error	t-value	p-value	%Variance	Estimate	Std. Error	t-value	p-value	%Variance	
Age	0.004	0.003	1.62	0.11	1.2						Adj. R-sqr
Gender Male	-0.029	0.034	-0.86	0.39	1.1						0.083
log(1+Azm)	**-0.074**	**0.030**	**-2.50**	**0.014**	**5.55**	**-0.077**	**0.040**	**-1.92**	**0.058**	**6.1**	FDR
AChE Genetics	**-0.048**	**0.021**	**-2.34**	**0.022**	**4.96**	**-0.044**	**0.022**	**-2.03**	**0.045**	**4.1**	**0.09**
	Winter AChE Activity										
	Unadjusted Linear Model					Sandwich Adjusted Model With Selected Variables					
Variables	Estimate	Std. Error	t-value	p-value	%Variance	Estimate	Std. Error	t-value	p-value	%Variance	
Age	0.000	0.002	-0.10	0.92	0.38						Adj. R-sqr
Gender Male	-0.026	0.030	-0.85	0.40	0.19						0.032
log(1+Azm)	-0.100	0.113	-0.89	0.38	0.4						FDR
AChE Genetics	**-0.049**	**0.021**	**-2.31**	**0.023**	**4.3**	**-0.046**	**0.026**	**-1.77**	**0.079**	**4**	**0.079**

Age has a slight trending positive slope in the spring-summer but slight negative slope in the winter. The male gender has a non-significant similar negative slope in both seasons. Age and gender both only explain < 1.2% of the variance whereas Azinphos-methyl and AChE minor alleles are significant with ~4 to 6% of variance explained by each.
The genetic effects and pesticide effects are similarly sized in the effect on variance at this exposure level. The slopes are very similar before and after variable selection for the final model with sandwich error estimators. Inspection of the residuals (See Supplement Figures 4-6, 4-7 & 4-8) suggests that all linear model normality assumptions have been met for the spring-summer AChE activity models. For the winter, AChE enzyme activity residuals, some departure from normality is detected due to tailing and a low Shapiro-Wilk p-value (See Supplemental Figure 4-8) making the sandwich error estimator more important for the winter AChE activity model. The standard errors for the AChE genetic component were similar in the initial and selected models. A generalized least squares analysis of the AChE data in a two-point subject variance time series analysis supported the conclusions of the independent season regression results. Sample attrition across seasons allowed for the inclusion of 74 individuals for this evaluation. The results of the seasonal generalized least squares analysis with false discovery adjusted significance (FDR < 0.1) association found that the AChE minor allele carriers additive effect (See Table 4-3).

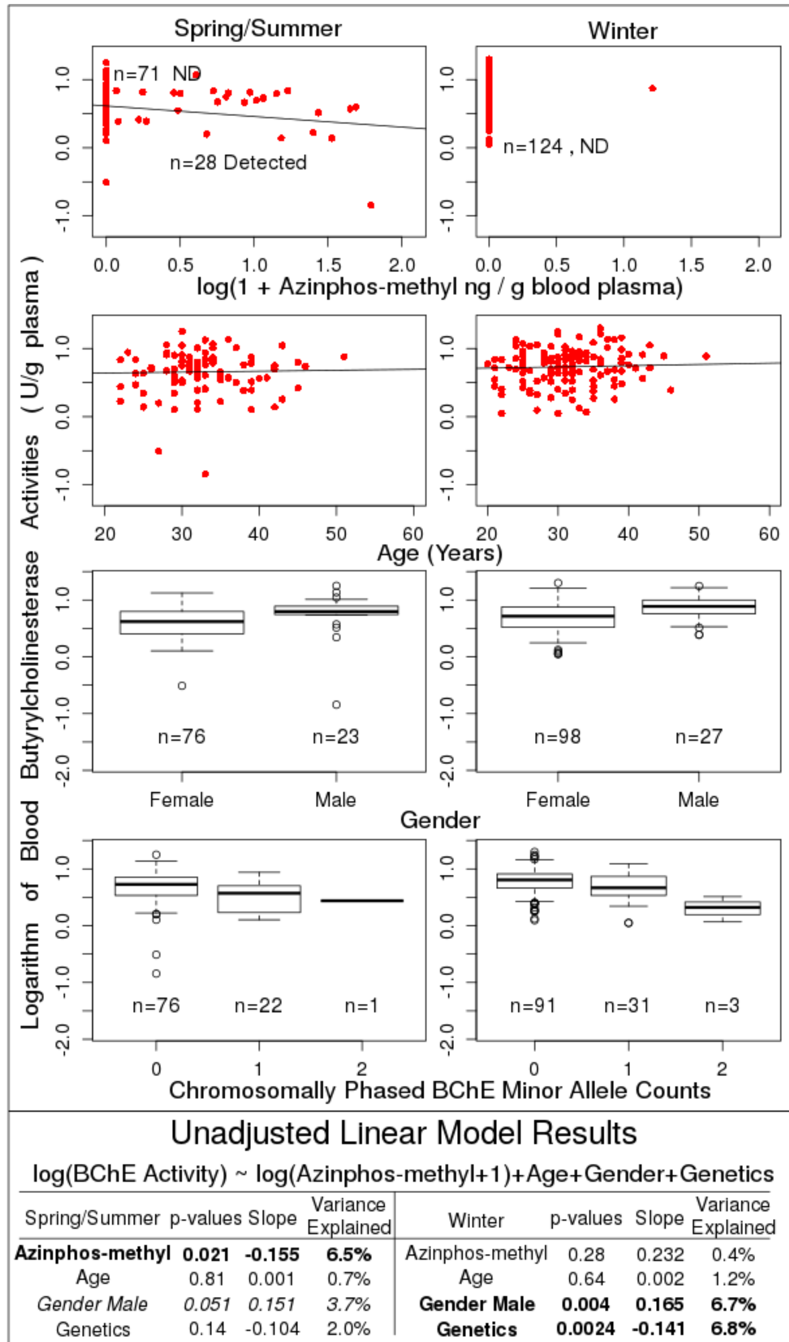
Table 4-3 Summary of AChE Activity Time Series Regression Results

Table 4-3	AChE Activity Time Series Generalized Least Squares Analysis								
	Full Model				Selected Variable Model				
Variables	Estimate	Std. Error	t-value	p-value	Estimate	Std. Error	t-value	p-value	
Age	**0.005**	**0.003**	**2.09**	**0.04**	**0.005**	**0.0025**	**1.85**	**0.066**	
Gender Male	-0.041	0.034	-1.20	0.23					
log(1+Azm):season	**-0.088**	**0.022**	**-4.03**	**0.0001**	**-0.089**	**0.022**	**-4.11**	**0.0001**	FDR
AChE Genetics	**-0.069**	**0.023**	**-2.95**	**0.004**	**-0.067**	**0.024**	**-2.84**	**0.0051**	**0.01**

In this time-series subject variance analysis Age gains suggestive significance as a co-variate with Azinphos-methyl and the AChE minor allele genetics in the final model. In the seasonal independent models age was only suggestive in the spring-summer blood collection. This may be due to differential exposure of the younger individuals. The AChE minor allele count genetics main effect remains significantly associated in the time-series model with lower activity and a slightly larger negative slope (~-0.068).

4.4.4. BChE Enzyme Activity Linear Models

Sample attrition due to missing variables allowed the inclusion of 99 subjects in the spring-summer BChE generalized linear model and 125 subjects in the winter model. Azinphos-methyl exposure is associated with lower BChE enzyme activity in the spring-summer blood collection (See Figure 4-3 & Table 4-4).



Spring/Summer	p-values	Slope	Variance Explained	Winter	p-values	Slope	Variance Explained
Azinphos-methyl	**0.021**	**-0.155**	**6.5%**	Azinphos-methyl	0.28	0.232	0.4%
Age	0.81	0.001	0.7%	Age	0.64	0.002	1.2%
Gender Male	*0.051*	*0.151*	*3.7%*	**Gender Male**	**0.004**	**0.165**	**6.7%**
Genetics	0.14	-0.104	2.0%	**Genetics**	**0.0024**	**-0.141**	**6.8%**

Figure 4-3 BChE Blood Phenotype Gene + Environment Generalized Linear Model

The top panels show regression of the BChE enzyme activity by Azinphos-methyl concentration. The next panels depict BChE activity by age, followed by gender and genetic minor alleles with the p-values and variance explained reported in the bottom of the figure. The middle lines in the boxplots is the median. The boxes are the middle 50% of the data with the edges as the 25th and 75th percentiles. The whiskers are the interquartile range past the percentiles and the circles are the extreme values.

Minor alleles in the BChE gene are also associated with lower enzyme activity, but appear to exhibit seasonal variability in association with the minor alleles. The reduced BChE enzyme activity signal is strongest in the winter when Azinphos-methyl exposure is detected in only one individual. The slope is more negative in the winter (-0.14) than the spring-summer (-0.10) even when adjusted by the sandwich estimator and model selection. The male gender is a significant associated co-variate in the winter with higher BChE enzyme

activities with similar slopes and variance explained in the spring-summer collection. Age is non-significant with BChE enzyme activity in both seasons. The plasma BChE enzyme activity phenotype genetic association appears to be dominated by exposure to the Azinphos-methyl pesticide in the spring-summer, while the genetic minor allele activity determinants are robustly associated in the winter when little exposure is measured. There is also an outlier, sample 140 identified in the Supplemental Figure 4-14 leverage plot with Cook's Distance. Removing this individual from the analysis and rerunning the analysis had similar effects as the sandwich estimator adjustment on the association statistics presented in Table 4-4. Inspection of the residual plots (Supplemental Figures 4-15 and 4-18) showed departure for normality in both spring-summer and winter models making the use of the sandwich estimator necessary for the BChE activity models. Slopes, standards errors, and significance were very similar before and after model selection and variable error sandwich adjustment.

Table 4-4 Summary of BChE Activity Regression Results by Season

Table 4-4	Summer BChE Activity										
	Unadjusted Linear Model					Sandwich Adjusted Model With Selected Variables					
Variables	Estimate	Std. Error	t-value	p-value	%Variance	Estimate	Std. Error	t-value	p-value	%Variance	
Age	0.001	0.006	0.25	0.8	0.67						Adj. R-sqr
Gender Male	**0.151**	**0.076**	**1.98**	**0.051**	**3.7**	**0.154**	**0.079**	**1.96**	**0.054**	**4.2**	0.101
log(1+Azm)	**-0.155**	**0.066**	**-2.36**	**0.021**	**6.5**	-0.156	0.133	-1.17	0.25	6.5	FDR
BChE Genetics	**-0.104**	**0.070**	**-1.48**	**0.14**	**2**	**-0.105**	**0.063**	**-1.66**	**0.10**	**2.1**	**0.10**
	Winter BChE Activity										
	Unadjusted Linear Model					Sandwich Adjusted Model With Selected Variables					
Variables	Estimate	Std. Error	t-value	p-value	%Variance	Estimate	Std. Error	t-value	p-value	%Variance	
Age	0.002	0.004	0.47	0.64	1.2						Adj. R-sqr
Gender Male	**0.165**	**0.057**	**2.92**	**0.0042**	**6.7**	**0.168**	**0.053**	**3.19**	**0.0018**	**7.6**	0.127
log(1+Azm)	0.232	0.212	1.10	0.28	0.37						FDR
BChE Genetics	**-0.141**	**0.045**	**-3.10**	**0.0024**	**6.8**	**-0.135**	**0.047**	**-2.86**	**0.005**	**6.4**	**0.01**

We also applied a generalized least squares analysis to the BChE activity data in a two-point subject time series analysis. Sample attrition across seasons and variables allowed for the inclusion of 74 individuals in the BChE enzyme activity time-series analysis. This analysis supports the results of the independent seasonal regression models. The false discovery adjusted non-significance (FDR~ 0.31) association observed in the BChE gene minor allele carriers (See Table 4-5) have a negative slope. We still observe the significant co-variates male gender and Azinphos-methyl pesticide blood concentration with a significant seasonal interaction effect on activity in the subject based time series generalized least squares analysis.

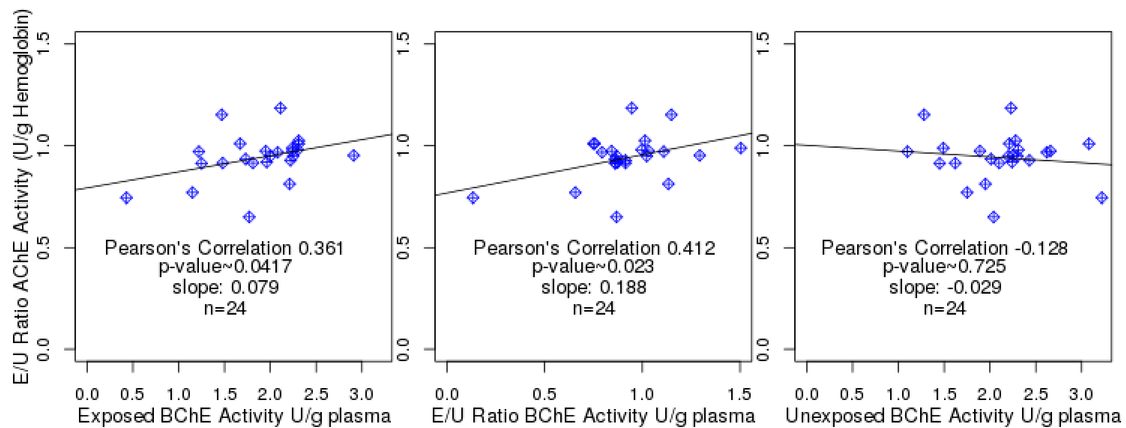
Table 4-5 Summary of BChE Activity Time Series Regression Results

Table 4-5	BChE Activity Time Series Generalized Least Squares Analysis								
	Full Model				Selected Variable Model				
Variables	Estimate	Std. Error	t-value	p-value	Estimate	Std. Error	t-value	p-value	
Age	0.002	0.005	0.370	0.71					
Gender Male	**0.151**	**0.064**	**2.370**	**0.02**	**0.156**	**0.062**	**2.500**	**0.014**	
log(1+Azm):season	**-0.209**	**0.052**	**-3.999**	**0.0001**	**-0.208**	**0.052**	**-4.007**	**0.0001**	FDR
BChE Genetics	-0.060	0.061	-0.988	0.32	-0.062	0.060	-1.026	0.31	0.31

4.4.5. AChE-BChE Enzyme-Enzyme Activity Inhibition Phenotype Relationship

Next, we evaluated the inhibition of the AChE activity phenotype and determined its relationship to activity of the BChE enzyme in blood of pesticides exposed individuals. Using three related metrics with the measured activity of the BChE enzyme blood phenotype: 1) the total BChE activity in the Azinphos-methyl spring-summer pesticide exposed individuals and 2) the exposed/unexposed BChE activities ratio and 3) the winter basal unexposed activity we investigated the relationship between AChE enzyme activity inhibition and available BChE enzyme activity. Since there is little exposure in the winter BChE activity in the winter should not be inhibited from the normal activity, which may give an unbiased estimate of the basal activity that the individual would be able to buffer. The total spring-summer Azinphos-methyl exposed individuals BChE activity (left panel Figure 4-4) and exposed individuals BChE inhibition ratio (middle panel Figure 4-4) show a significant relationship between the enzyme activity and inhibition of BChE and the inhibition of AChE enzyme activity respectively. This was expected due to the susceptibility of both these enzymes to inactivation by Azinphos-methyl. This demonstrates that Azinphos-methyl acts on both enzymes and is similar to pesticide exposed human subjects' results obtained by other investigators [176]. The winter pesticide unexposed basal BChE enzyme activity (right panel Figure 4-4) of the individuals that were exposed in the spring-summer does not appear to be significantly correlated with the AChE exposed/unexposed activity ratio.

Figure 4-4 AChE Inhibition of Activity Plotted by BChE Enzyme Activity in Azinphos-methyl Exposed Individuals



The Y-axis is the ratio of AChE activities in the Azinphos-methyl exposed spring-summer thinning season divided by the same individuals winter AChE activity. The X-axis of the left panel is the spring-summer BChE enzyme activity when Azinphos-methyl exposure occur. The X-axis of the middle panel is the ratio of BChE activity in spring-summer thinning season (E) divided by the same individual's Azinphos-methyl winter BChE enzyme activity (U). The X-axis of the right panel is the BChE enzyme activity in winter when individuals are not exposed to Azinphos-methyl. The Azinphos-methyl exposed BChE enzyme activity and the BChE enzyme activity ratio were significantly correlated with the activity ratio of the AChE enzyme.

4.4.6. Regression of Azinphos-methyl Exposed AChE and BChE Gene Minor Allele Carriers with AChE and BChE Enzyme Activity Inhibition

Since a normal individual has ~307x more BChE activity than AChE activity in the blood [140] we sought to see if carriers of BChE gene minor alleles would have effects on the ratio of exposed to unexposed enzyme activities of AChE and BChE. We did test whether AChE and BChE gene minor alleles would have an effect on the AChE enzyme inhibition while exposed. It is not expected that AChE gene minor alleles would have a measurable buffering effect on BChE activity due to the 1 AChE unit to 307 BChE units activity profile of whole blood [140] so we did not test the AChE minor alleles in the BChE enzyme activity model. We excluded individuals who did not have Azinphos-methyl detected in the blood because the variability of the marginally larger unexposed individuals confounds the specific effects in individuals which we know are concurrently pesticide exposed with their cholinesterase enzyme activity. Twenty-four unrelated genotyped subjects had Azinphos-methyl blood concentrations measured as exposure in the summer with both spring-summer and winter AChE and BChE enzyme activities to calculate the inhibitions.

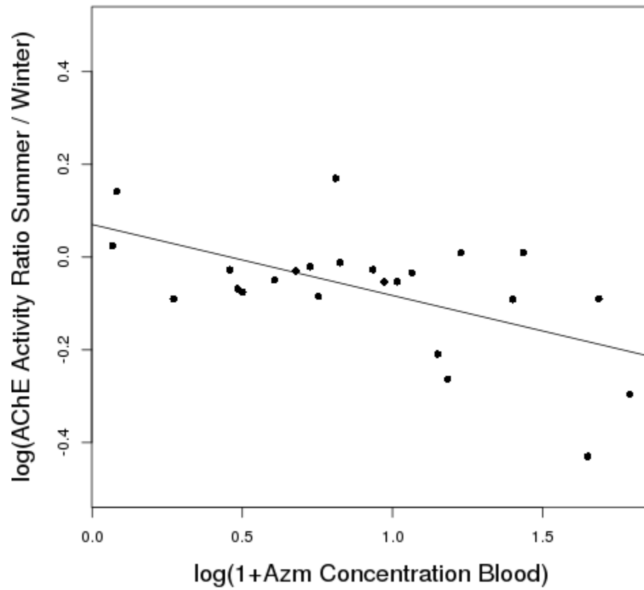
Table 4-6 Summary of AChE and BChE Activity Inhibition Regression Results

Table 4-6	AChE Activity Inhibition										
	Unadjusted Linear Model					Sandwich Adjusted Model With Selected Variables					
Variables	Estimate	Std. Error	t-value	p-value	%Variance	Estimate	Std. Error	t-value	p-value	%Variance	Adj. R-sqr
Age	-0.008	0.005	-1.46	0.16	**23.4**	**-0.009**	**0.005**	**-1.87**	**0.078**	**23.4**	0.446
Gender Male	-0.023	0.062	-0.37	0.71	1.6						
log(1+Azm)	**-0.144**	**0.045**	**-3.23**	**0.005**	**25**	**-0.148**	**0.052**	**-2.86**	**0.010**	**26.1**	FDR
AChE Genetics	-0.037	0.036	-1.03	0.3	3.3	-0.031	0.033	-0.96	0.35	1.9	0.47
BChE Genetics	0.039	0.055	0.71	0.5	1.3	0.050	0.058	0.85	0.41	2.8	0.47

	BChE Activity Inhibition										
	Unadjusted Linear Model					Sandwich Adjusted Model With Selected Variables					
Variables	Estimate	Std. Error	t-value	p-value	%Variance	Estimate	Std. Error	t-value	p-value	%Variance	Adj. R-sqr
Age	0.009	0.019	0.46	0.65	0.78						0.218
Gender Male	-0.276	0.227	-1.22	0.238	11.6						
log(1+Azm)	**-0.442**	**0.179**	**-2.47**	**0.023**	**21.6**	-0.496	0.360	-1.38	0.18	25.6	FDR
BChE Genetics	0.049	0.218	0.22	0.83	0.17	0.167	0.224	0.74	0.47	3	0.47

Age and Azinphos-methyl blood concentration were included as co-variates in the AChE activity ratio inhibition adjusted model. (See Table 4-6 and Figures 4-5 and 4-6) Slopes and significance are similar between initial and final selected models. The age associated AChE inhibition negative slope, sandwich adjusted p-value and explained variance suggests that older individuals are more likely to have more inhibition of their AChE enzyme while exposed.

Figure 4-5 **AChE Inhibition by Azinphos-methyl**



The Y-axis displays the ratio of the Azinphos-methyl exposed individuals spring-summer AChE enzyme activity divided by their corresponding unexposed winter blood sample activity. The X-axis displays the logarithm of the Azinphos-methyl concentration plus one. The linear regression results displayed in Table 4-6 where Azinphos-methyl is significant with both adjusted and unadjusted p-values < 0.01 and explains ~25% of the variance by ANOVA of the residuals (See Table 4-6).

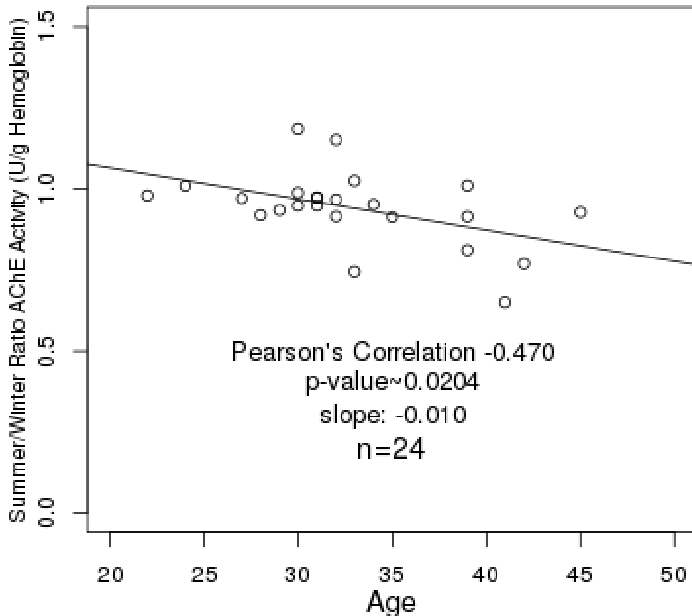
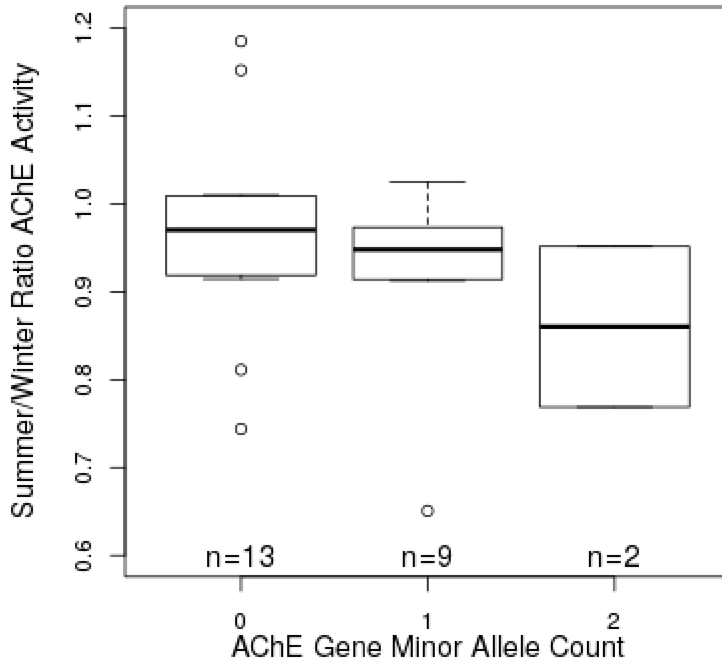


Figure 4-6 AChE Inhibition by Age in Azinphos-methyl Exposed Individuals

The Y-axis displays the ratio of the Azinphos-methyl exposed individuals spring-summer AChE enzyme activity divided by their corresponding unexposed winter blood sample activity. The X-axis displays the age of the corresponding individuals. The p-value shown is based on the Pearson's Correlation and is supported by linear regression results with co-variates where age appears to explain ~23% of the variance in AChE enzyme activity inhibition (See Table 4-6).

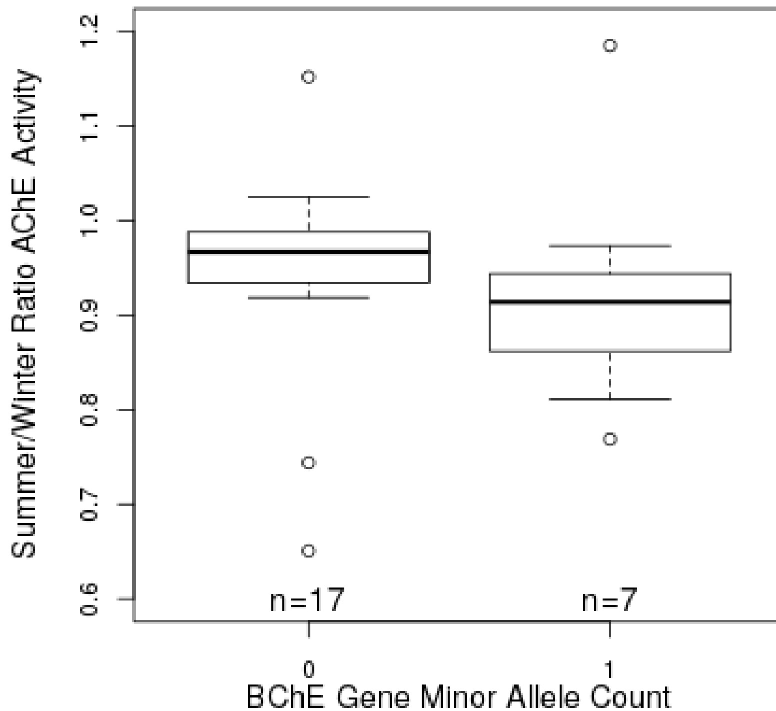
We did not find a significant genetic AChE or BChE minor allele count association to AChE enzyme activity inhibition by linear regression methods (See Table 4-6). Visual inspection of the boxplot of the inhibition data (See Figures 4-7 & 4-8) shows a negative trend in the spring-summer by winter AChE ratio with increasing minor allele counts of both genes having lower activity. Age-related inhibition variation and the co-variate adjustment in association obscures the directionality for the BChE minor allele carrier categories by giving it a positive slope. The sample size (n=24) is not likely enough to detect this potential effect of minor alleles on AChE enzyme activity inhibition.

Figure 4-7 AChE Inhibition by AChE Minor Allele Counts in Azinphos-methyl Exposed Individuals



The Y-axis displays the ratio of the Azinphos-methyl exposed individuals spring-summer AChE enzyme activity divided by their corresponding unexposed winter blood sample activity. The X-axis displays the chromosomally phased allelic minor allele counts of the AChE minor allele carriers with zero (0) as major allele homozygotes, one (1) as major-minor allele heterozygotes and two (2) as individuals that carry two chromosomal copies (homozygotes) of the minor allele forms of the AChE gene. There is a non-significant visible trend in the data. See Table 4-6 for association statistics. The middle lines on the boxes are the medians. The boxes are the middle 50% of the data with the edges as the 25th and 75th percentiles. The whiskers are the interquartile range past the percentiles and the circles are the extreme values.

Figure 4-8 AChE Inhibition by BChE Minor Allele Counts in Azinphos-methyl Exposed Individuals



The Y-axis displays the ratio of the Azinphos-methyl exposed individuals spring-summer AChE enzyme activity divided by their corresponding unexposed winter blood sample activity. The X-axis displays the chromosomally phased allelic minor allele counts of the AChE minor allele carriers with zero (0) as major allele homozygotes and one (1) as major-minor allele heterozygotes. See Table 4-6 for association statistics. We had no Azinphos-methyl exposed individuals that carried two chromosomal copies (homozygotes) of the minor allele forms of the BChE gene. The middle lines on the boxes are the medians. The boxes are the middle 50% of the data with the edges as the 25th and 75th percentiles. The whiskers are the interquartile range past the percentiles and the circles are the extreme values.

The BChE activity ratio enzyme inhibition model also has Azinphos-methyl as a significant co-variate in the unadjusted p-values so was included in the final sandwich adjusted model where it is suggestive while explaining ~25% of the variance in BChE enzyme inhibition. (See Table 4-6 & Figure 4-9)

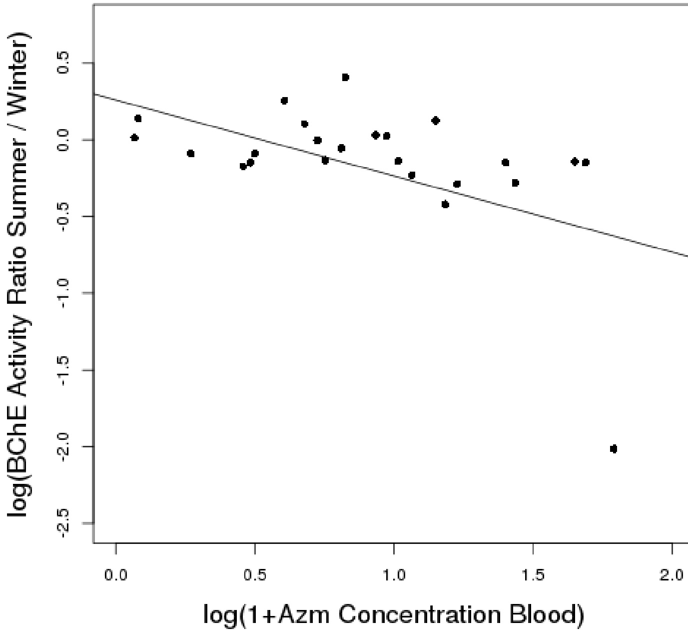
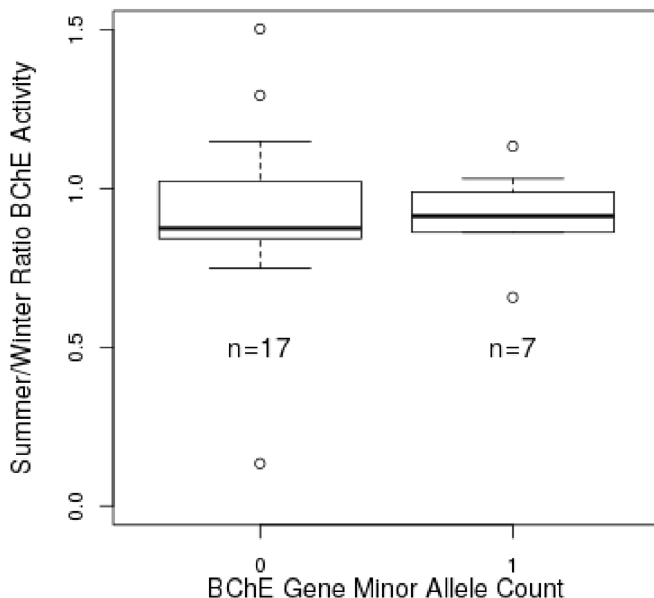


Figure 4-9 BChE Inhibition by Azinphos-methyl in Azinphos-methyl Exposed Individuals

The Y-axis displays the ratio of the Azinphos-methyl exposed individuals spring-summer BChE enzyme activity divided by their corresponding unexposed winter blood sample enzyme activity. The X-axis displays the logarithm of the Azinphos-methyl concentration plus one. The slope from the linear regression results displayed in Table 4-6 where Azinphos-methyl is not significant after sandwich interval error estimation and explains up to ~23% of the variance by ANOVA of the residuals.

Minor alleles in BChE showed no visual trend (See Figure 4-10) and were not associated with greater inhibition of BChE activity in Azinphos-methyl exposed individuals (See Table 4-6)

Figure 4-10 BChE Inhibition by BChE Minor Allele Counts in Azinphos-methyl Exposed Individuals



The Y-axis displays the ratio of the Azinphos-methyl exposed individuals spring-summer BChE enzyme activity divided by their corresponding unexposed winter blood sample activity. The X-axis displays the chromosomally phased allelic minor allele counts of the AChE minor allele carriers. Zero (0) depicts major allele homozygotes and one (1) depicts major-minor allele heterozygotes. See Table 4-6 for association statistics. We had no Azinphos-methyl exposed individuals that carried two chromosomal copies (homozygotes) of the minor allele forms of the BChE gene. The boxes are the middle 50% of the data with the edges as the 25th and 75th percentiles. The whiskers are the interquartile range past the percentiles and the circles are the extreme values.

4.5. Discussion

In this paper, we saw that lower blood enzyme activity in both cholinesterase genes is additively associated with minor alleles and blood pesticide measures of exposure in this agricultural community cohort candidate gene by environment analysis. The genetic variants, enzymes and pesticide exposure selected for this study are well known to be biochemically functional with these cholinesterases. Age and gender are also known to affect the systemic biology of most all systems making their inclusion in initial model tests important. Similar effects of minor allele have been demonstrated with pharmacogenetic variants in the BChE gene with specific drugs (i.e., succinylcholine muscle relaxant susceptibility) [140]. Using only known enzyme activity reducing, disease associated, protein changing exonic and splice site minor alleles of the AChE and BChE genes, these genetic influences on the measured enzymes activities isolate a signal of reductions in activity due to genetic effects. Genetic minor allele variants have the potential to be bi-directional and non-monotonic in the larger context of their effects on phenotype [141]. This genetic variant site selection emphasizes a mono-directional enzyme activity effect of these single nucleotide variant sites and is likely what has allowed us to have the power to see these additive gene plus environment associations with enzyme activity. This variant site selection also replicates the associated genetic loss of function reports for many of the variants included in this analysis, see Table 4-1 for a variant by variant list of references.

We have a crude measure of exposure using point blood concentration estimates of Azinphos-methyl dose as seen from mass-spectrometry detection. A time series analysis of pesticide blood concentrations with more blood samples collected in closer temporal proximity would give a better sense of the dose experienced by the exposed individuals. This more frequent sampling could be targeted to the spring-summer when exposures are more common. We show that the Azinphos-methyl pesticide exposure and the genetics of the pesticide target AChE and BChE enzymes can jointly have an additive negative effect on the enzyme activity phenotypes with both genetics and environment explaining ~ 5% of the variance in enzyme activity even after model selection and error estimation sandwich adjustments. Male gender also appears to be a significant co-variate with increased BChE enzyme activity in both seasons (Figure 4-3, Table 4-4). There appear to be different underlying model results in the AChE and BChE enzyme systems by which the additive genetic and environmental pesticide exposure are associated in the models. We show a more complex gene-environment relationship with BChE enzyme activity regression where BChE gene minor alleles are also associated with lower activity, but this signal is partially occluded by the spring-summer Azinphos-methyl exposure which dominates the genetic (and gender) effects and depresses the minor allele association in the pesticide exposed spring-summer unadjusted results. The BChE minor allele genetic association is not observed in the time series analysis (See Table 4-5). The BChE activity models are particularly driven by an outlier, sample 140 (See Supplemental Figures 4-14, 4-15, 4-25 & 4-26 Residuals plots) with sample 117 also being an outlier with similar direction but less extreme influence. Removal of sample 140 and rerunning of the BChE enzyme activity model appears to adjust the model association statistics (data not shown) with similar results to the sandwich error estimator adjustment where the Azinphos-methyl co-variate becomes non-significant upon sandwich adjustment and we regain the genetic variable association (See Table 4-4). This sample 140

also has the highest detected blood concentration of Azinphos-methyl in the model and has the most inhibition of BChE enzyme activity detected. This outlier is an individual that would have been pulled from the field based on the Washington State Labor and Industries Standards (http://www.lni.wa.gov/Safety/Topics/AtoZ/Cholinesterase/) and thus has a biologically plausible reason to be interpreted carefully. The organophosphate pesticide exposure induced inhibition regulatory exposure limit set by the Washington State Department of Labor and Industries for pesticide handlers' states workers are to cease work at 20% inhibition of blood cholinesterase. If more comprehensive time series blood samples were taken to estimate the covariance of the pesticide blood concentration with the BChE enzyme activity, it may be possible to observe both the genetic and environmental effects on activity simultaneously in this exposed individual. The outlier sample 140 is driving the environmental Azinphos-methyl effect on BChE enzyme activity. It is possible that the leverage effects of sample 140 in the Supplemental Figure 4-14 leverage plot as a BChE activity outlier is also due to the acuteness of this individual's exposure. The duty duration of the plasma BChE enzyme activity pesticide exposure effect is likely different than AChE enzymes on the surface of blood cells. Red blood cells circulate for ~100-120 days in an adult [177]. The blood plasma BChE enzyme is produced by the liver. The response and durational effects of these two enzymes expressed by different tissues into the blood are likely to differ in their toxicodynamic characteristics based on their source tissue production. The ~307x blood volume abundance of BChE in comparison to AChE [140] could reduce the dominance of the environment on AChE enzyme activity to a proportional marginal effect by buffering the stoichiometric portion of Azinphos-methyl that AChE enzyme encounters in the blood. The time window in which they show a depression of activity as the exposed individual returns to homeostasis is likely faster with the liver derived BChE enzyme rather than the blood cell born AChE enzyme. The exposed enzyme-enzyme activity when organophosphate pesticide is in the blood is correlated. The relatively low AChE inhibition slopes with respect to BChE activity metrics (Figure 4-4) indicate that AChE activity reduces at a slower rate than BChE while exposed. The regression analysis of AChE enzyme activity inhibition in Azinphos-methyl exposed individuals with minor alleles in the AChE and BChE genes did not find a genetic association, but plotting of allele carrier groups shows a promising visual trend and we have a small sample size with 24 individuals. We observed a age related negative slope trend in the AChE enzyme inhibition that appears significant investigated solely by Pearson's Correlation. A similar effect has been observed in rats where older organisms do not recover as quickly to their basal level of AChE enzyme activity and show inhibition for longer periods of time [178, 179]. This increased AChE enzyme inhibition observed in older subjects is supported by these rat models [178, 179]. Other studies have shown variable cholinesterase responses of both enzymes to organophosphate exposure between individuals [180] which would support further investigation into this phenomenon.

There is a wider window of natural variation in both blood cholinesterase activities shown in Figure 4-1 than the genetic variances can explain. This suggests we have not observed all the variation determinants including allelic content of these genes and the other possible pesticide exposure effects on the enzyme activities. It is likely that other unobserved alleles and exposures are also responsible for a portion of the low and high activity phenotypes seen in the histograms of Figure 4-1. It is likely that regulatory gene variants contribute to some

of this variability. We purposely only selected exonic variation to limit the heterogeneity in direction of effect to increase our power at this sample size. This excluded regulatory variation which should also be investigated in more depth by different extensions of similar methods to those presented here. There is also likely more rare variation in this cohort which was not on the genotyping array. Additionally we did not biochemically detect any of the silent variant homozygotes for BChE which are known to exists in the human population [181]. We did not sequence these individuals to find all the rare variants contained in this population. This would be an obvious next step to improve the model. The variant site selection of the Illumina 5M genotyping platform is derived from the multi-ethnic 1000 Genomes project populations [144-147] and ~12,000 African-American and Caucasian subjects of the Exome Sequencing Project [148-154] and does not capture all our cohort's private rare variants, despite the low frequency of many of the variants we do see. There is a potential loss of variance specificity and power in the estimated genetic effects due to these other unobserved rare alleles. To improve the utility of this analysis in risk assessment and occupational medicine contexts as well as clinical and personalized medicine, one would need to sequence these individuals to ascertain their personal rare variants affecting cholinesterase activity. More powered studies constructed like this one, but with sequencing based variant detection may begin to be able to define the unknown genetically determined basal cholinesterase activities. This would be useful with acute organophosphate exposure patients in hospital settings, thus informing the clinician of the expected basal enzyme activity and allowing better approximations of acute severity of poisoning. Electronic medical records may contain an individual's historical cholinesterase activity values and clinically actionable genotypes in estimating these basal and critical activities for individual patients. In clinical decision-analysis this would allow the delivery of a higher standard of prognostic care. There is early literature, circa 1950s to the present, regarding BChE enzyme activity variants with respect to succinylcholine muscle relaxant susceptibility [140]. Current hospital reports suggest that even though a functional understanding of BChE's importance to anesthetic metabolism susceptibility has been understood for more than a half century, precision medicine is still working to bridge the information communication and clinical practice gap in risk management [182, 183]. Both BChE and AChE genes will likely be selected as clinically informative diagnostic variables for a number on conditions due to the pleiotropic nature of their systemic distributions effects on many phenotypes and disease state [140]. This makes understanding the normal range of activities and the genetic basis of differences in activity important for the construction of prognostic models.

The ethics of protection of susceptible populations and occupational risk analysis settings poses several different dilemmas from the clinical medical establishment. Carriers of genetic minor alleles may be at greater risk in comparison to the major allele carriers. Knowledge of our own environmental exposure susceptible risk genotypes allows an individual to enable their personal prevention and reduction of assumed risks in occupational and exposure settings. Strict non-discrimination laws and genetic privacy codes are already in place to enforce public equity of individuals with genetic variants that may affect their susceptibility. Farmworker-pesticide exposure-effect studies like the cholinesterase example presented here are aimed to inform a more genetically equitable population risk assessment for the potentially susceptible individuals. Genetically susceptible individuals such as those minor allele carriers investigated here need to be included into the regulatory framework for

estimating acceptable population and occupational exposures a priori. This study's relatively small sample size (~100 samples after attrition, 24 pesticide exposed) shows that both the environment and genetics have additive terms in the model and these both explain similar amounts of the variance on these pesticide target enzymes. We were not able to show that the genetics had a significant effect on the cholinesterase's enzyme inhibition. We suspect the trend we see in AChE enzyme activity may become significant with more individuals sampled and more blood collections in time series. With the coming advent of electronic medical record integration with statistically informed clinical decision analysis, the integration of known genotypes into predictive disease models that are informed by hundreds to thousands of systematically ascertained clinical symptom and lab reports will allow for more robustly powered research and higher standards of care in diseases and exposures that have cholinergic etiologies and may further prevent avoidable exposures.

4.6. Supplemental Figures

Supplemental Regression Diagnostics Plots and Normal Assumptions Tests

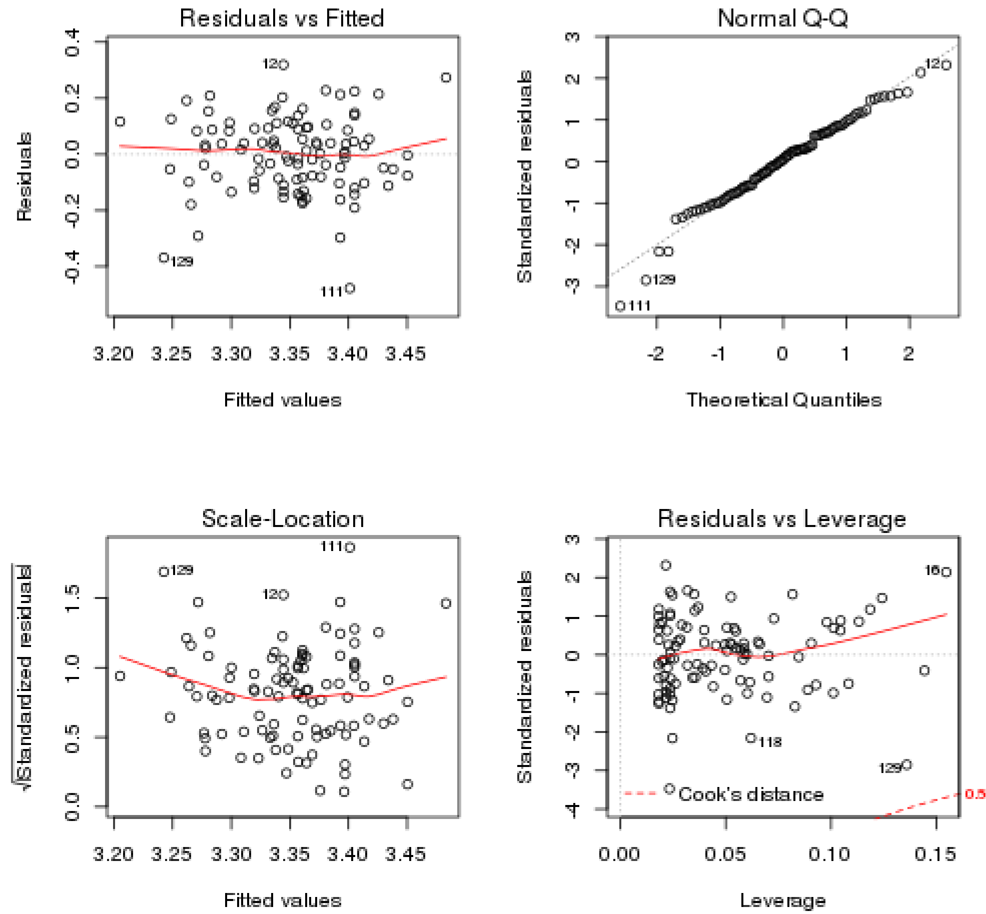


Figure 4-6 Spring-Summer AChE Enzyme Activity Model Residuals Diagnostic Plots

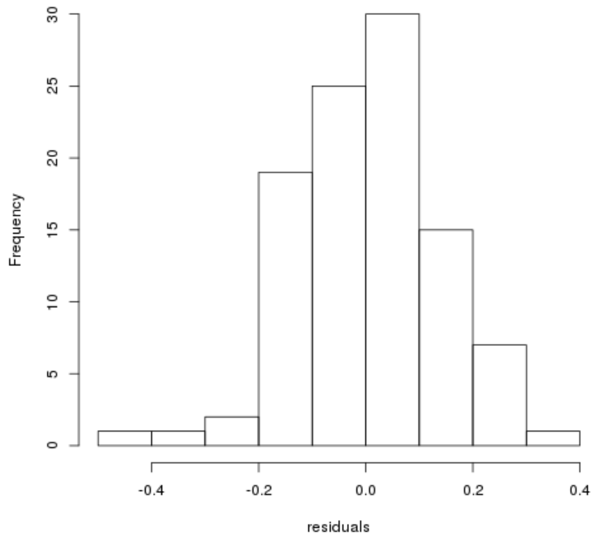


Figure 4-7 **Spring-Summer AChE Enzyme Activity Residuals**

Shapiro-Wilk normality test: W = 0.98108, p-value = 0.1567
(W = 0.98756, p-value = 0.469 final model)

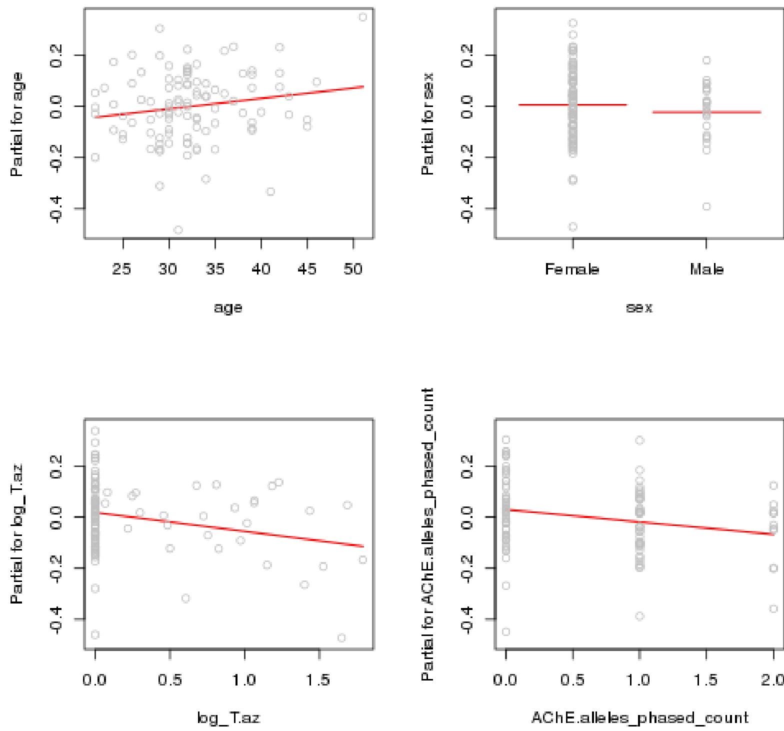


Figure 4-8 **Spring-Summer Partial Residuals Plot**

Shows the correction that the other co-variates have without the effect of the variable.
Durbin-Watson test of Collinearity: DW = 2.2316, p-value = 0.8685

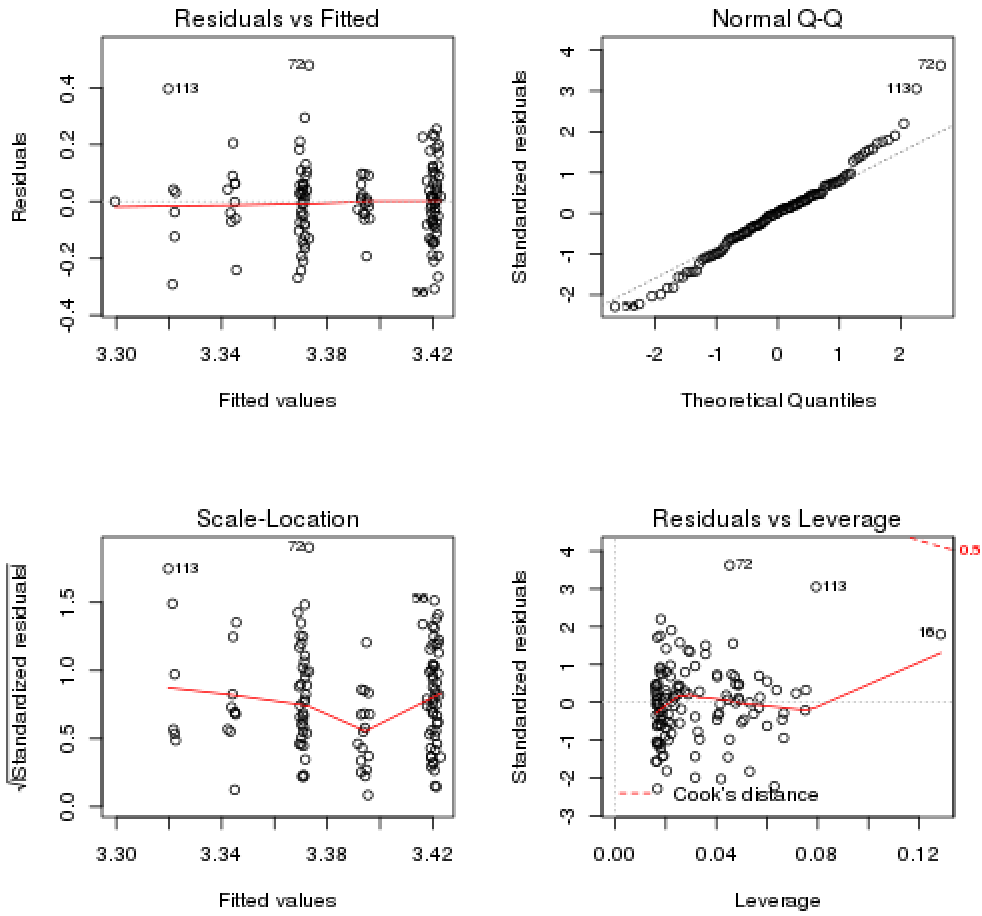


Figure 4-9 **Winter AChE Enzyme Activity Model Residuals Diagnostic Plots**

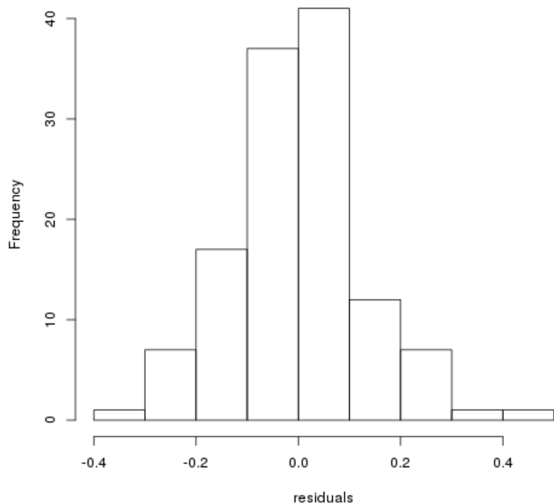


Figure 4-10 Winter AChE Enzyme Activity Model Residuals Diagnostic Plots

Shapiro-Wilk normality test : W = 0.97877, p-value = 0.0479
(W = 0.97809, p-value = 0.04129 final model)

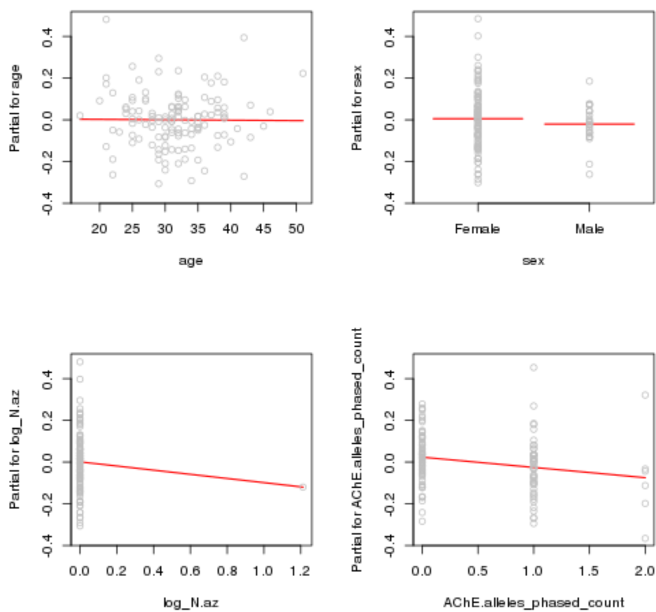


Figure 4-11 Winter AChE Enzyme Activity Model Residuals Diagnostic Plots

Shows the correction that the other co-variates have without the effect of the variable.
Durbin-Watson test of Collinearity: DW = 1.9674, p-value = 0.4193

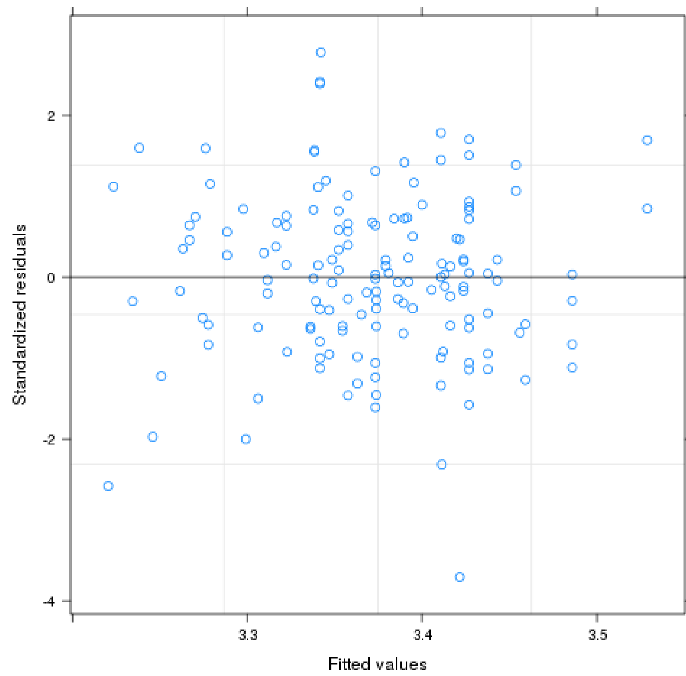


Figure 4-12 **Spring-Summer AChE Enzyme Activity Time-Series Model Residuals Diagnostic Plots**

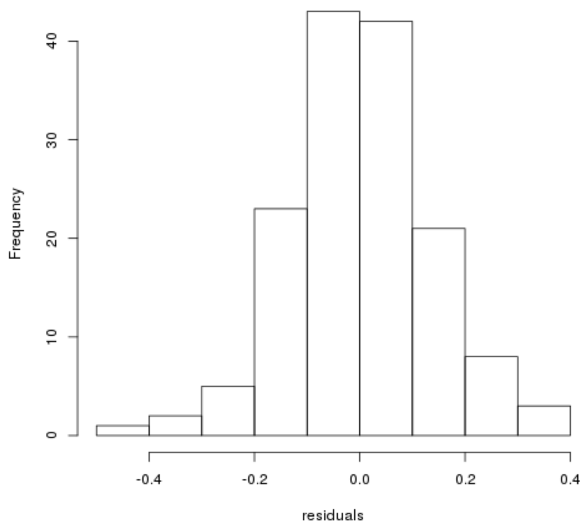


Figure 4-13 **Spring-Summer AChE Enzyme Activity Time-Series Model Residuals Diagnostic Plots**

Shapiro-Wilk normality test: W = 0.99053, p-value = 0.4233
(W = 0.99016, p-value = 0.3897 final model)
Durbin-Watson test of Collinearity: DW = 2.0759, p-value = 0.6661

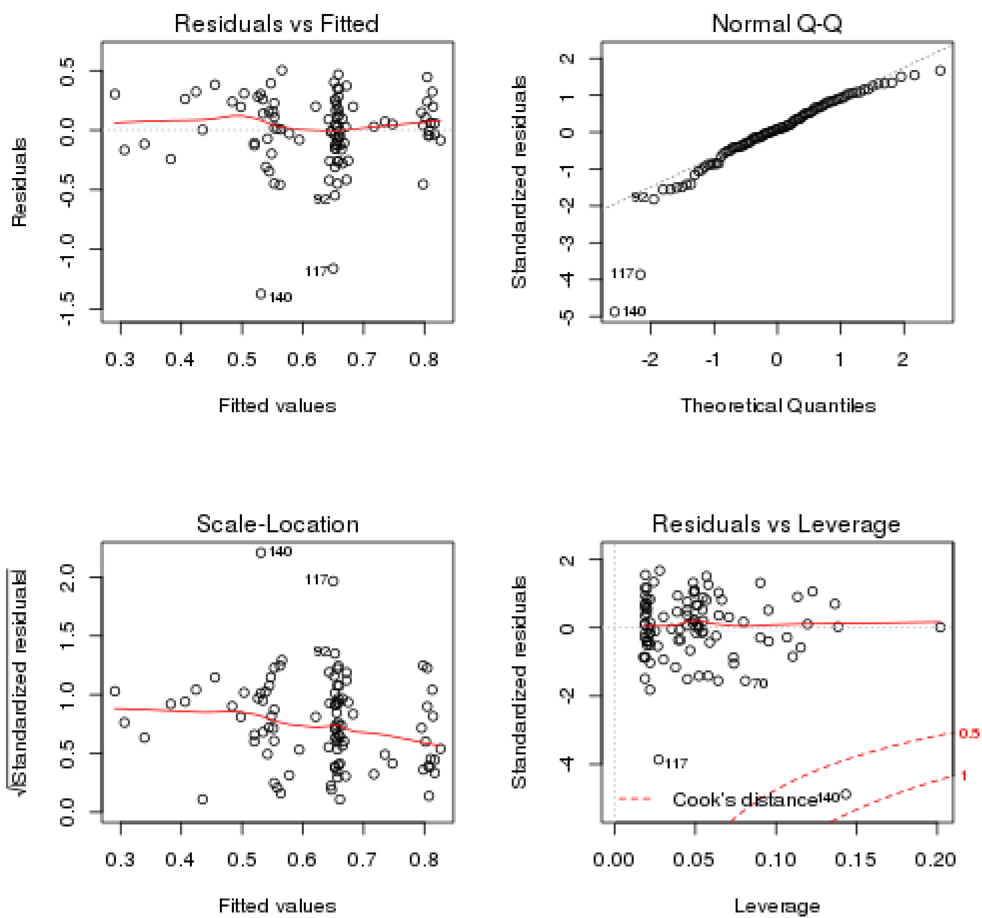


Figure 4-14 **Spring-Summer BChE Enzyme Activity Model Residuals Diagnostic Plots**

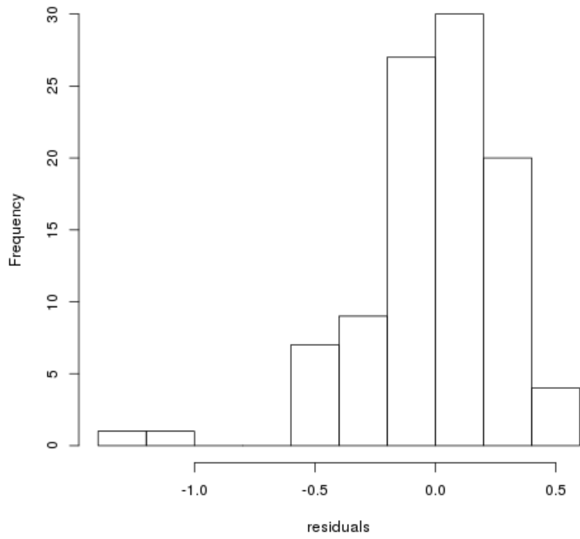


Figure 4-15 **Spring-Summer BChE Enzyme Activity Model Residuals Diagnostic Plots**

Shapiro-Wilk normality test: W = 0.89642, p-value = 1.07e-06
(W = 0.89498, p-value = 9.191e-07 final model)

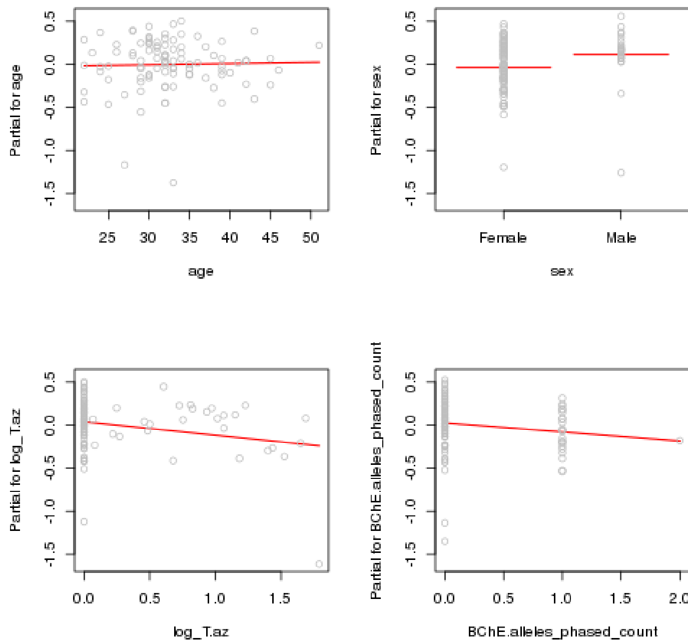


Figure 4-16 **Spring-Summer BChE Enzyme Activity Model Residuals Diagnostic Plots.**

Shows the correction that the other co-variates have without the effect of the variable.
Durbin-Watson test of Collinearity: DW = 2.2312, p-value = 0.8647

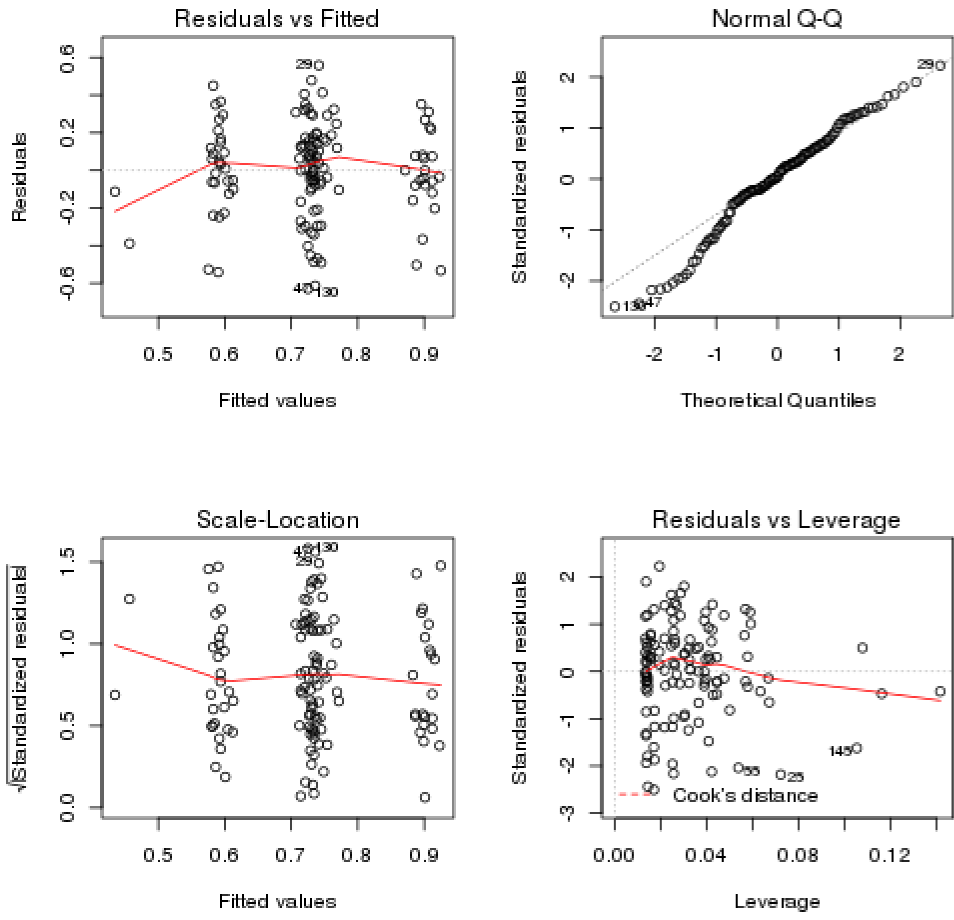


Figure 4-17 Winter BChE Enzyme Activity Model Residuals Diagnostic Plots

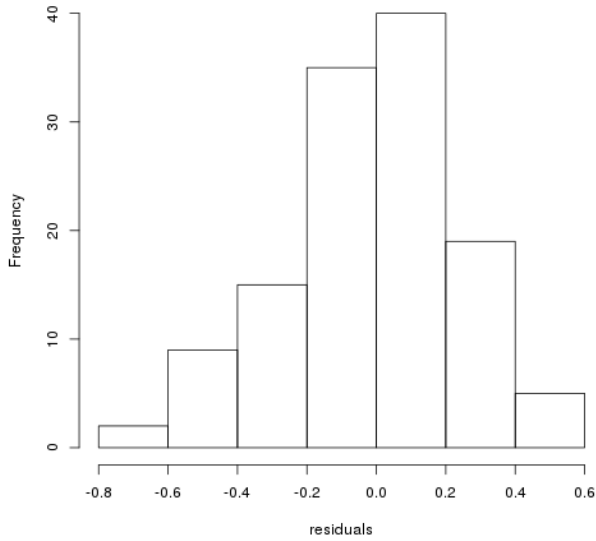


Figure 4-18 Winter BChE Enzyme Activity Model Residuals Diagnostic Plots

Shapiro-Wilk normality test: W = 0.9758, p-value = 0.02413
(W = 0.97573, p-value = 0.02376 final model)

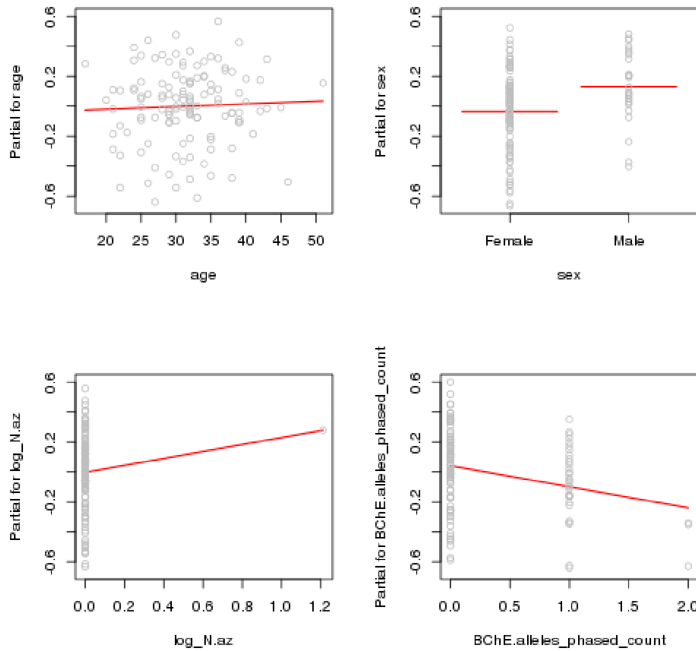


Figure 4-19 Winter BChE Enzyme Activity Model Residuals Diagnostic Plots

Shows the correction that the other co-variates have without the effect of the variable.
Durbin-Watson test of Collinearity: DW = 1.8553, p-value = 0.2027

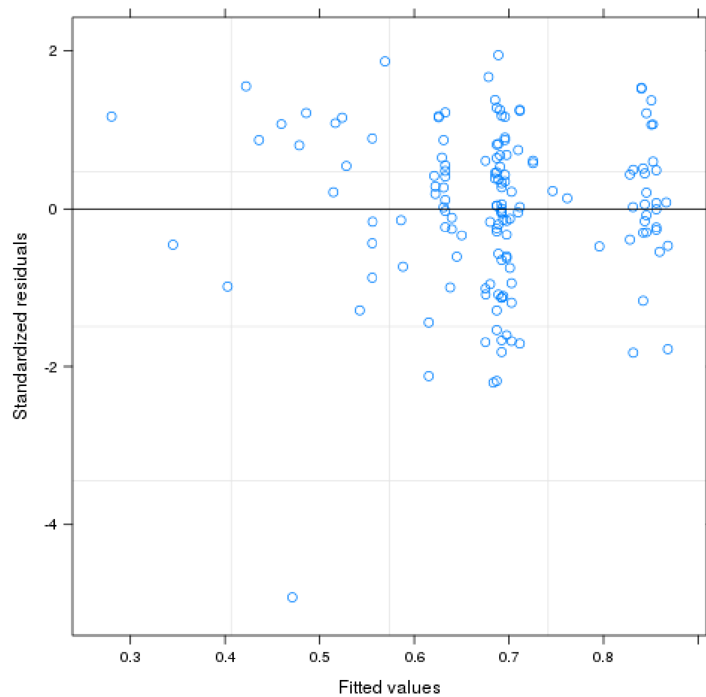


Figure 4-20 **Spring-Summer AChE Enzyme Activity Time-Series Model Residuals Diagnostic Plots**

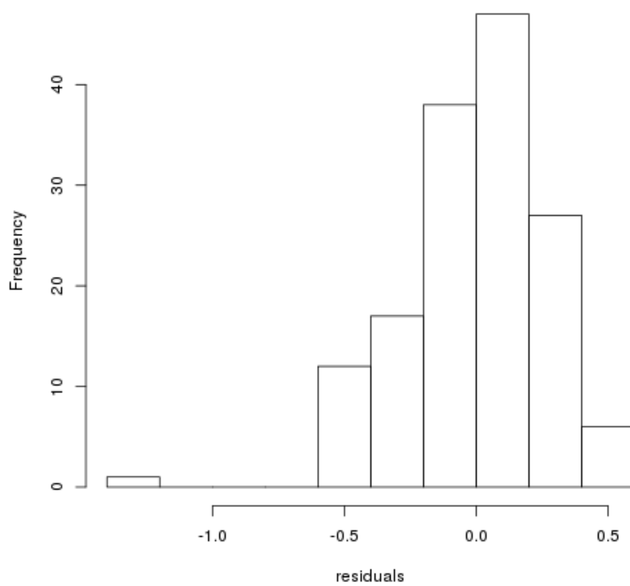


Figure 4-21 **Spring-Summer AChE Enzyme Activity Time-Series Model Residuals Diagnostic Plots**

Shapiro-Wilk normality test: W = 0.94875, p-value = 2.922e-05
(W = 0.94704, p-value = 2.123e-05 final model)
Durbin-Watson test of Collinearity: DW = 1.8853, p-value = 0.2293

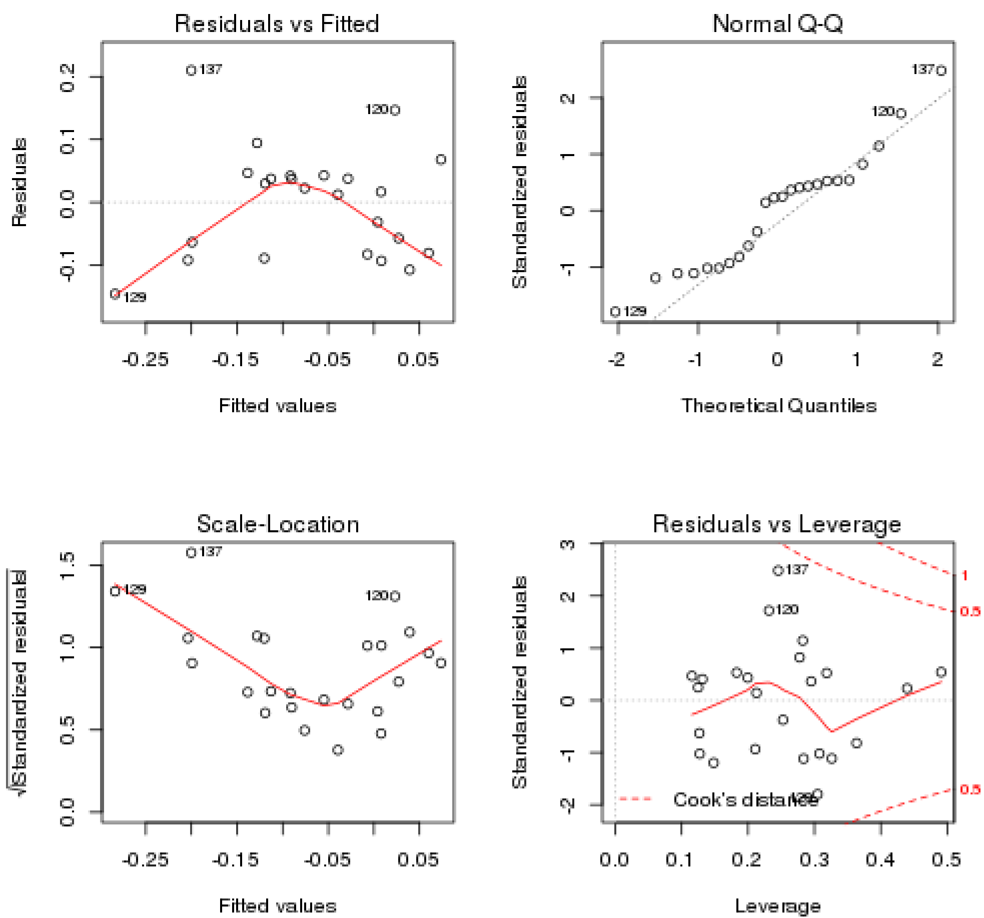


Figure 4-22 **AChE Enzyme Inhibition of Activity Model Residuals Diagnostic Plots**

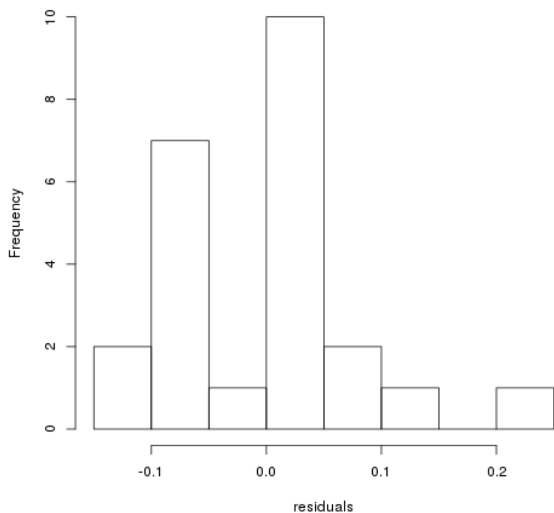


Figure 4-23 **AChE Enzyme Inhibition of Activity Model Residuals Diagnostic Plots**

Shapiro-Wilk normality test: W = 0.94473, p-value = 0.2077
(W = 0.9462, p-value = 0.2238 final model)

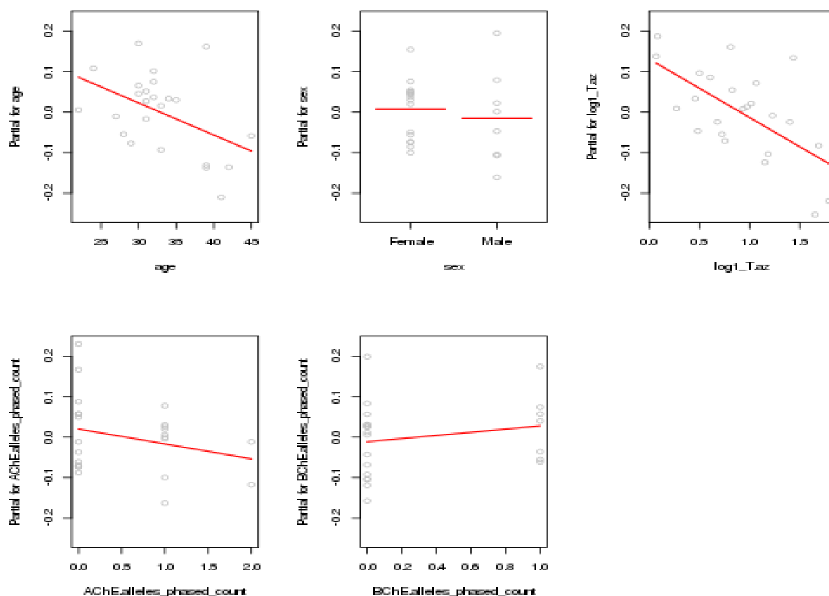


Figure 4-24 **AChE Enzyme Inhibition of Activity Model Residuals Diagnostic Plots**

Shows the correction that the other co-variates have without the effect of the variable.
Durbin-Watson test of Collinearity: DW = 1.7791, p-value = 0.2901

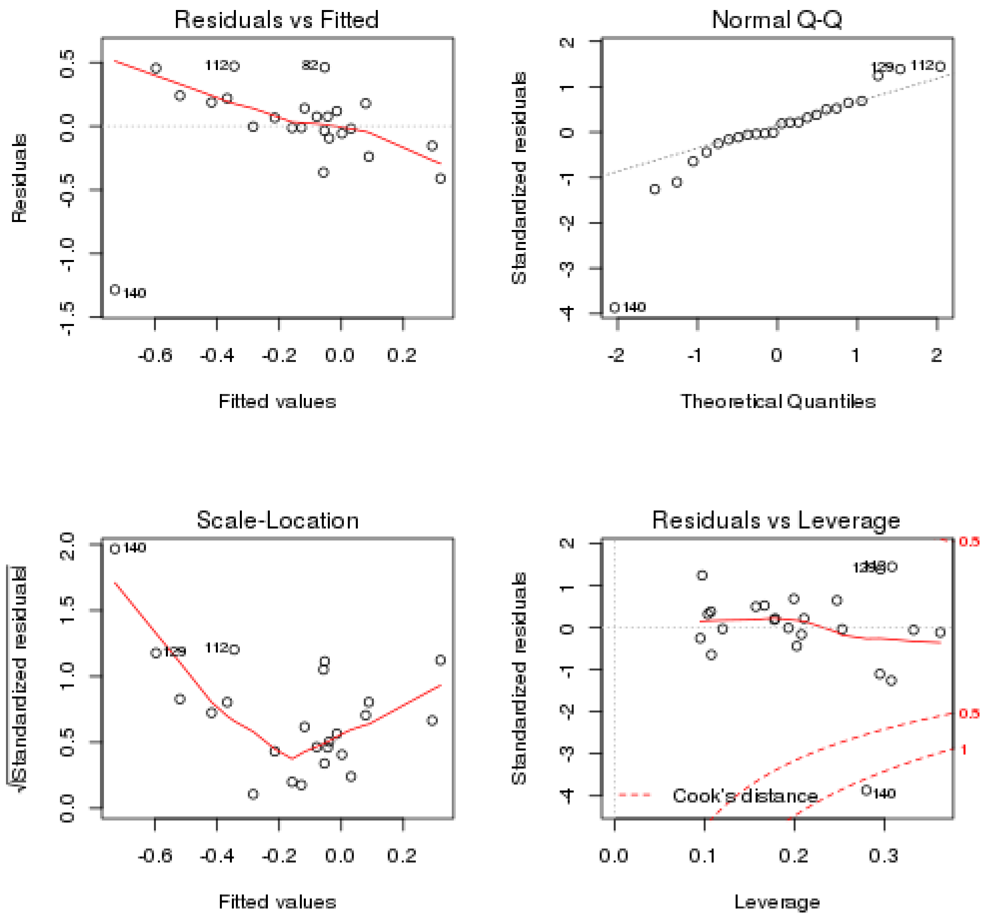


Figure 4-25 **BChE Enzyme Inhibition of Activity Model Residuals Diagnostic Plots**

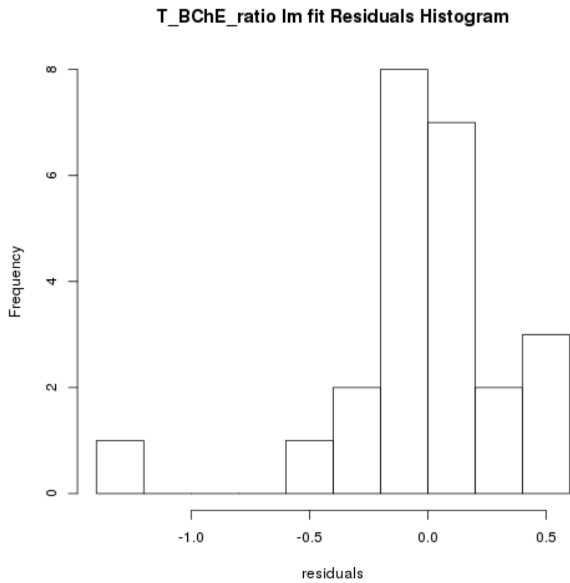


Figure 4-26 **BChE Enzyme Inhibition of Activity Model Residuals Diagnostic Plots**

Shapiro-Wilk normality test: W = 0.82091, p-value = 0.0006605
(W = 0.80807, p-value = 0.0004004 final model)

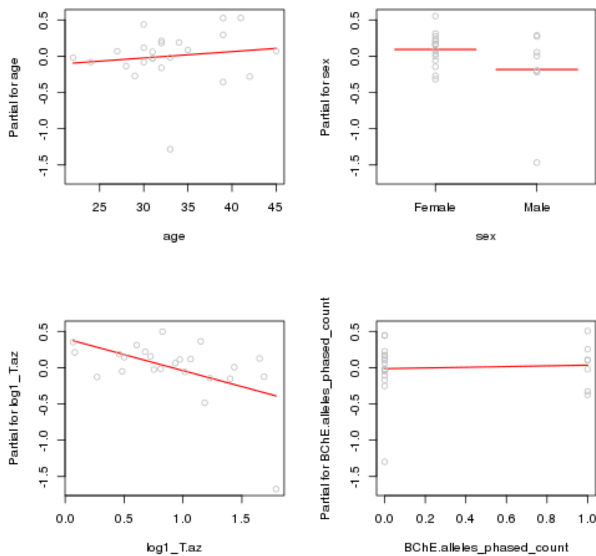


Figure 4-27 **BChE Enzyme Inhibition of Activity Model Residuals Diagnostic Plots**

Shows the correction that the other co-variates have without the effect of the variable.
Durbin-Watson test of Collinearity: DW = 1.9856, p-value = 0.4733

5. DISSERTATION CONCLUSIONS

In this dissertation I use molecular epidemiology with computational biology and machine learning techniques to show agricultural pesticide exposure is associated with alterations in three phenotypes: the blood biomarkers (1) Acetylcholinesterase and (2) Butyrylcholinesterase activities and (3) the novel findings that Azinphos-methyl exposed individuals are associated with significantly different oral-buccal microbiome compositions than those individuals where there is little evident of exposure.

Both phenotype systems (cholinesterase and microbiomes) have orthogonal literature supporting the direction and pattern of these observed effects. We see in this dissertation's study cohort the second most common oral-buccal genera detected, *Streptococcus* in the order *Lactobacillales* is significantly reduced in Azinphos-methyl exposed individuals. These microbiome findings were not in the literature when I began this program in the Fall of 2012. The first orthogonal article supporting this altered microbiome pattern that we detect in humans, appeared in 2013, with Joly et. al. [91] where the investigators show in an *in-vitro* bioreactor and an *in-vivo* rat model where significant perturbations of the resident gut bacteria were observed in animals dosed with the related organophosphate pesticide Chlorpyrifos. The perturbed microbiota shows reductions in members of the order *Lactobacillales*. The same authors followed this with two *in-vivo* rat gut microbiome papers with similar results and added neonatal and changes in gut epithelial tight junction proteins in exposed animals which also allows the bacteria to translocate [92, 93]. It seems that both direct antibacterial action of organophosphate pesticides on bacterial taxa and host pesticide exposure mediated effects can affect the quantity and distribution of bacteria in the oral cavity and body system. Two other recent groups also have *in-vivo* model organophosphate pesticide dosing studies that show perturbation of the microbiome. One in the rat gut microbiome with the OP Diazinon [94], and another in the Japanese Quail gut microbiome with the OP Trichlorfon [90]. This brings the literature reports to four other independent groups that have observed a pattern of association in model systems spanning host independent *in-vitro* methods to *in-vivo* methods in three organisms. Our study of this cohort is believed to be the first in humans and was started before the other evidence also emerged. This puts this research at the cutting edge of what is known about the effects of OPs on the microbiome.

Where possible I used more than one method to show a direction and an association pattern to reduce bias in the conclusions reached. I also applied machine learning techniques to the OTU level molecular taxonomy that became a central and very important aspect of the analysis. Originally I used beta diversity distance metrics that compare ecological compositions by a pairwise difference score between all individuals to identify cluster groups. The analysis of differences in microbiome composition between individuals was initially performed by the Bray-Curtis dissimilarity matrix [184] defined as:

$$BC_{ij} = 1 - (2C_{ij} / (S_i + S_j)) \quad \text{Eq. 5}$$

Where C_{ij} is the sum of values of species in common between individuals ***ij*** and S_i and S_j total count of species in both individuals ***ij***. Bray-Curtis values range from 0 to 1, with 1 indicating the samples are completely different in composition. Using a Multi-Dimensional Scaling (MDS) method of unsupervised machine learning to reduce the dimensions of the Bray-Curtis matrix and project it onto an X-Y coordinate system reveals two primary microbiome-type clusters of individuals. One of the two clusters is highly enriched for individuals with farmworker status and individuals with blood concentration exposure measurements in the summer to the insecticide Azinphos-methyl. This approach is valid and provided good resolution in the clustering of exposed and unexposed individuals, but did not provide a computational framework in which we could directly identify which bacteria are different.

In order to evaluated the contributions of specific taxa, we elected to analyze the data using the centered-log-ratio with principal components analysis of the common OTUs. This allowed us to directly rank the top scoring loadings in PC1 and PC2 to see which microbiota are driving the observed pattern in the composition of exposed individuals. We complemented this with the non-parametric Wilcoxon's Rank Sum Test between those individual with Azinphos-methyl detected in the blood and those with none detected. I also applied the centered-log-ratio to correct for the autocorrelation simplex proportional geometry of the compositional 16S read count data as the genus taxonomic level. I tested hypotheses with the bacterial genera identified at more than 0.5% mean proportional abundance of sequenced reads at the genus taxonomic levels to limit the number of hypotheses to only a much shorter common bacteria list (n ~ 22 in adults) that we have the power to detect in nearly all samples and applied a false discovery rate to adjust for multiple testing.

In the case of blood cholinesterases, the knowledge of genetic variant carrier effects spans back to the 1940's [140], but even the current literature and practice is lacking exploration in specific translational applications of this enzyme system knowledge into the regulatory framework used to ensure public health. The recent comprehensive review of Butyrylcholinesterase by Dr. Oksana Lockridge [140] and the attached email correspondence with her (Appendix 1 Email) show that the Gene by Environment interactions of AChE and BChE minor alleles potential to cause susceptibility to organophosphates was predicted, but had never convincingly been observed in an *in-situ* natural human system. One of the logical lines of evidence that previously would have lent the scientific establishment to think this BChE gene - AChE gene by environment differential buffering susceptibility to be true, is that intravenous dosing of BChE preparations can rescue severely poisoned subjects from death [140]. The gene-gene by environment results I show concerning genetic minor allele contributions to organophosphate pesticide susceptibility took a stepwise approach to the integration of the analysis. In the preliminary analysis, I used t-test's between minor and major allele carrying groups to collectively replicate many of the well-known alleles that have activity differences in the cholinesterase blood phenotype. This utilized an understanding of loss of function phenotypes in a heterozygous context. The genetic findings effects on phenotype by themselves are not novel, but the demonstrated replication usefulness of these variants in a clinical genomics context is relevant as we look at personalized medicine and risk. In this dissertation, I use gene by environment interaction to

improve our understanding of the occupational medicine standard of care for pome fruit orchard workers. My research can inform public health regulators on how genetic minor alleles of blood cholinesterase genes may have to be more readily monitored during pesticide use to identify the changes in cholinesterase activity during environmental exposures. My research also shows that combining the genetic minor allele effects on cholinesterase activity with the Azinphos-methyl pesticide exposure in a linear model by including both genotype and environmental variables improved our understanding of population response variability to organophosphate pesticide exposure. This graduate school experience complements the computational biology and nucleic acid analysis skills that I came to school with very well. If you would like to see example code please feel free to contact me at bard@uw.edu.

5.1. Special Acknowledgements

In the process of this Ph.D. I had several mentors and collaborators that I would like to give special gratitude to. First is Dr. Elaine Faustman, my advisor. Her wide degree of academic experience as well as savvy interactions with the NIH and EPA funding apparatus and scientific establishments has made all this possible. Second I would like to thank Jim Wallace. Jim is a great person with many years of experience in bioinformatics applications. We have spent many collective hours talking through analyses and manuscripts and stratagem of how to make something of this data. I think for a measure of total time spent, he is the most invested person in my Ph.D., besides myself of course. I think we spent it very well and it has been fruitful and fun. Third, I would like to thank Sungwoo Hong, Carly Wilder and Foad Green. They were in the lab generating the microbiome 16S sequence data more than I. I was involved in the lab the first few weeks of the bench efforts with Foad and Sung designing primers, making sure we saw bands in our gels and sequencing the first 16 samples on the Ion Torrent with Jesse Tsai. Then that summer when I had to take my preliminary PhD exam, Carly and Sung powered through the samples and when I returned from the exam there was our data ready for me to start learning the computational ways to analyze the information. The efficiency this lent is amazing. Fourth, I would like to thank Tomomi Workman and Dr. William (Bill) Griffith. These two are the analytical and institutional knowledge pillars of Elaine's lab. They provide the context and data validation that is needed to link this cohort's analysis into a coherent framework. In a similar context Eric Vigoren also fulfilled this role in the first year that I worked with Elaine's group before he moved onto another employment opportunity, I am grateful to him also for that first year of pointers to where in the information is. Dr. Ali Shojaie is a most valuable contributor to this dissertation. His statistical guidance focused the lens I made through which we looked. I would also like to thank Dr. Lianne Sheppard for her valuable guidance in bioethics and her generous support by including me in her training grant funding paradigm. This allowed me to work unfettered for the last two years of graduate school without worry to how I was going to pay my bills and eat. I have the most gratitude to you all, it takes a village.

This research was funded by the National Institute of Child Health and Human Development, Department of Health and Human Services, https://www.nichd.nih.gov/ (Contract No. HHSN267200700023C recipient: EMF), the National Institute of Environmental Health Sciences, http://www.niehs.nih.gov/ (T32ES015459 recipient: IBS, 5P01 ES009601 recipient: EMF, and P30ES007033) and the United States Environmental Protection Agency, http://www.epa.gov/ (RD83451401 recipient: EMF, and RD832733 recipient: EMF).

6. REFERENCES

1. Weldon, B.A., et al., *Urinary microRNAs as potential biomarkers of pesticide exposure.* Toxicol Appl Pharmacol, 2016.
2. Thompson, B., et al., *Variability in the take-home pathway: farmworkers and non-farmworkers and their children.* J Expo Sci Environ Epidemiol, 2014. **24**(5): p. 522-31.
3. Thompson, B., et al., *Variability in the take-home pathway: Farmworkers and non-farmworkers and their children.* J Expos Sci Environ Epidemiol, 2014. **24**(5): p. 522-531.
4. Thompson, B., et al., *Para ninos saludables: a community intervention trial to reduce organophosphate pesticide exposure in children of farmworkers.* Environ Health Perspect, 2008. **116**(5): p. 687-94.
5. Smith, M.N., et al., *Seasonal and occupational trends of five organophosphate pesticides in house dust.* J Expo Sci Environ Epidemiol, 2016.
6. Smith, M.N., et al., *Seasonal variation in cortisol biomarkers in Hispanic mothers living in an agricultural region.* Biomarkers, 2015. **20**(5): p. 299-305.
7. Ou, Y.C., et al., *Induction of the cell cycle regulatory gene p21 (Waf1, Cip1) following methylmercury exposure in vitro and in vivo.* Toxicol Appl Pharmacol, 1999. **157**(3): p. 203-12.
8. Ou, Y.C., et al., *Induction of growth arrest and DNA damage-inducible genes Gadd45 and Gadd153 in primary rodent embryonic cells following exposure to methylmercury.* Toxicol Appl Pharmacol, 1997. **147**(1): p. 31-8.
9. Nonnenmann, M.W., et al., *Utilizing pyrosequencing and quantitative PCR to characterize fungal populations among house dust samples.* J Environ Monit, 2012. **14**(8): p. 2038-43.
10. Holme, F., et al., *The role of diet in children's exposure to organophosphate pesticides.* Environ Res, 2016. **147**: p. 133-40.
11. Coronado, G.D., et al., *Organophosphate pesticide exposure and work in pome fruit: evidence for the take-home pesticide pathway.* Environ Health Perspect, 2006. **114**(7): p. 999-1006.
12. Coronado, G.D., et al., *Do workplace and home protective practices protect farm workers? Findings from the "For Healthy Kids" study.* J Occup Environ Med, 2012. **54**(9): p. 1163-9.
13. Coronado, G.D., et al., *Organophosphate pesticide exposure and residential proximity to nearby fields: evidence for the drift pathway.* J Occup Environ Med, 2011. **53**(8): p. 884-91.
14. Coronado, G.D., et al., *Where's the dust? Characterizing locations of azinphos-methyl residues in house and vehicle dust among farmworkers with young children.* J Occup Environ Hyg, 2010. **7**(12): p. 663-71.
15. Coronado, G., et al., *Organophosphate Pesticide Exposure Among Pome and Non-Pome Farmworkers: A Subgroup Analysis of a Community Randomized Trial.* J Occup Environ Med, 2009. **51**(4): p. 500-9.
16. Das, S. and T.K. Adhya, *Degradation of chlorpyrifos in tropical rice soils.* J Environ Manage, 2015. **152**: p. 36-42.

17. Knuth, M.L., L.J. Heinis, and L.E. Anderson, *Persistence and distribution of azinphos-methyl following application to littoral enclosure mesocosms.* Ecotoxicol Environ Saf, 2000. **47**(2): p. 167-77.
18. Ellman, G.L., et al., *A new and rapid colorimetric determination of acetylcholinesterase activity.* Biochem Pharmacol, 1961. **7**: p. 88-95.
19. Barr, D.B., et al., *A multi-analyte method for the quantification of contemporary pesticides in human serum and plasma using high-resolution mass spectrometry.* J Chromatogr B Analyt Technol Biomed Life Sci, 2002. **778**(1-2): p. 99-111.
20. Wang, Q., et al., *Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy.* Applied and Environmental Microbiology, 2007. **73**(16): p. 5261-5267.
21. Keinan, A. and A.G. Clark, *Recent explosive human population growth has resulted in an excess of rare genetic variants.* Science, 2012. **336**(6082): p. 740-3.
22. Costello, E.K., et al., *Bacterial Community Variation in Human Body Habitats Across Space and Time.* Science (New York, N.Y.), 2009. **326**(5960): p. 1694-1697.
23. Maurice, C.F., H.J. Haiser, and P.J. Turnbaugh, *Xenobiotics shape the physiology and gene expression of the active human gut microbiome.* Cell, 2013. **152**(1-2): p. 39-50.
24. Davari, S., et al., *Probiotics treatment improves diabetes-induced impairment of synaptic activity and cognitive function: behavioral and electrophysiological proofs for microbiome-gut-brain axis.* Neuroscience, 2013. **240**: p. 287-96.
25. Clarke, G., et al., *The microbiome-gut-brain axis during early life regulates the hippocampal serotonergic system in a sex-dependent manner.* Mol Psychiatry, 2013. **18**(6): p. 666-73.
26. Zheng, X., et al., *Thinking outside the brain for cognitive improvement: Is peripheral immunomodulation on the way?* Neuropharmacology, 2015. **96**(Pt A): p. 94-104.
27. Ussar, S., et al., *Interactions between Gut Microbiota, Host Genetics and Diet Modulate the Predisposition to Obesity and Metabolic Syndrome.* Cell Metab, 2015. **22**(3): p. 516-530.
28. Rosenbaum, M., R. Knight, and R.L. Leibel, *The gut microbiota in human energy homeostasis and obesity.* Trends Endocrinol Metab, 2015. **26**(9): p. 493-501.
29. Di Luccia, B., et al., *Rescue of Fructose-Induced Metabolic Syndrome by Antibiotics or Faecal Transplantation in a Rat Model of Obesity.* PLoS One, 2015. **10**(8): p. e0134893.
30. Stilling, R.M., et al., *Friends with social benefits: host-microbe interactions as a driver of brain evolution and development?* Front Cell Infect Microbiol, 2014. **4**: p. 147.
31. Desbonnet, L., et al., *Microbiota is essential for social development in the mouse.* Mol Psychiatry, 2014. **19**(2): p. 146-8.
32. Mayer, B.T., et al., *Rapid and Profound Shifts in the Vaginal Microbiota Following Antibiotic Treatment for Bacterial Vaginosis.* Journal of Infectious Diseases, 2015.
33. Mason, M.R., et al., *The subgingival microbiome of clinically healthy current and never smokers.* ISME J, 2015. **9**(1): p. 268-272.
34. Wang, F., et al., *Gut microbiota community and its assembly associated with age and diet in Chinese centenarians.* J Microbiol Biotechnol, 2015.

35. De Filippo, C., et al., *Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa.* Proceedings of the National Academy of Sciences, 2010. **107**(33): p. 14691-14696.
36. Carmody, R.N. and P.J. Turnbaugh, *Host-microbial interactions in the metabolism of therapeutic and diet-derived xenobiotics.* J Clin Invest, 2014. **124**(10): p. 4173-81.
37. Dassi, E., et al., *Enhanced microbial diversity in the saliva microbiome induced by short-term probiotic intake revealed by 16S rRNA sequencing on the IonTorrent PGM platform.* Journal of Biotechnology, 2014. **190**(0): p. 30-39.
38. Lu, K., et al., *Arsenic exposure perturbs the gut microbiome and its metabolic profile in mice: an integrated metagenomics and metabolomics analysis.* Environ Health Perspect, 2014. **122**(3): p. 284-91.
39. Choi, J.J., et al., *Exercise attenuates PCB-induced changes in the mouse gut microbiome.* Environ Health Perspect, 2013. **121**(6): p. 725-30.
40. Sonnenburg, E.D., et al., *Diet-induced extinctions in the gut microbiota compound over generations.* Nature, 2016. **529**(7585): p. 212-5.
41. Mor, A., et al., *Prenatal exposure to systemic antibacterials and overweight and obesity in Danish schoolchildren: a prevalence study.* Int J Obes (Lond), 2015.
42. Arrieta, M.C., et al., *Early infancy microbial and metabolic alterations affect risk of childhood asthma.* Sci Transl Med, 2015. **7**(307): p. 307ra152.
43. Fukuto, T.R., *Mechanism of action of organophosphorus and carbamate insecticides.* Environ Health Perspect, 1990. **87**: p. 245-54.
44. Horton, M.K., et al., *Does the home environment and the sex of the child modify the adverse effects of prenatal exposure to chlorpyrifos on child working memory?* Neurotoxicol Teratol, 2012. **34**(5): p. 534-41.
45. Rauh, V., et al., *Seven-year neurodevelopmental scores and prenatal exposure to chlorpyrifos, a common agricultural pesticide.* Environ Health Perspect, 2011. **119**(8): p. 1196-201.
46. Lovasi, G.S., et al., *Chlorpyrifos exposure and urban residential environment characteristics as determinants of early childhood neurodevelopment.* Am J Public Health, 2011. **101**(1): p. 63-70.
47. Rauh, V.A., et al., *Impact of prenatal chlorpyrifos exposure on neurodevelopment in the first 3 years of life among inner-city children.* Pediatrics, 2006. **118**(6): p. e1845-59.
48. Rauh, V.A., et al., *Brain anomalies in children exposed prenatally to a common organophosphate pesticide.* Proc Natl Acad Sci U S A, 2012. **109**(20): p. 7871-6.
49. Yan, C., et al., *Repeated exposures to chlorpyrifos lead to spatial memory retrieval impairment and motor activity alteration.* Neurotoxicol Teratol, 2012. **34**(4): p. 442-9.
50. Edgar, R.C., *UPARSE: highly accurate OTU sequences from microbial amplicon reads.* Nat Meth, 2013. **10**(10): p. 996-998.
51. Fernandes, A.D., et al., *Unifying the analysis of high-throughput sequencing datasets: characterizing RNA-seq, 16S rRNA gene sequencing and selective growth experiments by compositional data analysis.* Microbiome, 2014. **2**: p. 15.
52. Aitchison, J., *The statistical analysis of compositional data*. Monographs on statistics and applied probability. 1986, London ; New York: Chapman and Hall. xv, 416 p.

53. Lovell, D., et al., *Proportionality: a valid alternative to correlation for relative data.* PLoS Comput Biol, 2015. **11**(3): p. e1004075.
54. Turnbaugh, P.J., et al., *A core gut microbiome in obese and lean twins.* Nature, 2009. **457**(7228): p. 480-484.
55. Hodges, J.L., Jr. and L.L. Cam, *The Poisson Approximation to the Poisson Binomial Distribution.* The Annals of Mathematical Statistics, 1960. **31**(3): p. 737-740.
56. Cai, L., et al., *Biased Diversity Metrics Revealed by Bacterial 16S Pyrotags Derived from Different Primer Sets.* PLoS ONE, 2013. **8**(1): p. e53649.
57. Ewing, B. and P. Green, *Base-calling of automated sequencer traces using phred. II. Error probabilities.* Genome Res, 1998. **8**(3): p. 186-94.
58. Ewing, B., et al., *Base-calling of automated sequencer traces using phred. I. Accuracy assessment.* Genome Res, 1998. **8**(3): p. 175-85.
59. McMurdie, P.J. and S. Holmes, *phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data.* PLoS ONE, 2013. **8**(4): p. e61217.
60. Boogaart, K.G.v.d. and R. Tolosana-Delgado, *Analyzing compositional data with R*. Use R! 2013, Heidelberg: Springer. xv, 258 pages.
61. Fraley, C. and A.E. Raftery, *Model-based clustering, discriminant analysis, and density estimation.* Journal of the American Statistical Association, 2002. **97**(458): p. 611-631.
62. Hochberg, Y. and Y. Benjamini, *More powerful procedures for multiple significance testing.* Stat Med, 1990. **9**(7): p. 811-8.
63. Di Pierro, F., et al., *Cariogram outcome after 90 days of oral treatment with Streptococcus salivarius M18 in children at high risk for dental caries: results of a randomized, controlled study.* Clin Cosmet Investig Dent, 2015. **7**: p. 107-13.
64. Bao, X., et al., *Streptococcus oligofermentans Inhibits Streptococcus mutans in Biofilms at Both Neutral pH and Cariogenic Conditions.* PLoS One, 2015. **10**(6): p. e0130962.
65. Terai, T., et al., *Screening of Probiotic Candidates in Human Oral Bacteria for the Prevention of Dental Disease.* PLoS One, 2015. **10**(6): p. e0128657.
66. Miyatani, F., et al., *Relationship between Cnm-positive Streptococcus mutans and cerebral microbleeds in humans.* Oral Dis, 2015. **21**(7): p. 886-93.
67. Naka, S., et al., *A specific Streptococcus mutans strain aggravates non-alcoholic fatty liver disease.* Oral Dis, 2014. **20**(7): p. 700-6.
68. Zielnik-Jurkiewicz, B. and A. Bielicka, *Antibiotic resistance of Streptococcus pneumoniae in children with acute otitis media treatment failure.* Int J Pediatr Otorhinolaryngol, 2015. **79**(12): p. 2129-33.
69. Azizi, A., et al., *In Vitro Effect of Zingiber officinale Extract on Growth of Streptococcus mutans and Streptococcus sanguinis.* Int J Dent, 2015. **2015**: p. 489842.
70. Oda, Y., F. Hayashi, and M. Okada, *Longitudinal study of dental caries incidence associated with Streptococcus mutans and Streptococcus sobrinus in patients with intellectual disabilities.* BMC Oral Health, 2015. **15**: p. 102.
71. Shimazu, K., et al., *Effects of surface reaction-type pre-reacted glass ionomer on oral biofilm formation of Streptococcus gordonii.* Odontology, 2015.

72. Chuzeville, S., et al., *Antigen I/II encoded by integrative and conjugative elements of Streptococcus agalactiae and role in biofilm formation.* Microb Pathog, 2015. **88**: p. 1-9.
73. Zbinden, A., N. Bostanci, and G.N. Belibasakis, *The novel species Streptococcus tigurinus and its association with oral infection.* Virulence, 2015. **6**(3): p. 177-82.
74. Asam, D., S. Mauerer, and B. Spellerberg, *Streptolysin S of Streptococcus anginosus exhibits broad-range hemolytic activity.* Med Microbiol Immunol, 2015. **204**(2): p. 227-37.
75. Asam, D. and B. Spellerberg, *Molecular pathogenicity of Streptococcus anginosus.* Mol Oral Microbiol, 2014. **29**(4): p. 145-55.
76. Al-Hebshi, N.N., et al., *Species-level core oral bacteriome identified by 16S rRNA pyrosequencing in a healthy young Arab population.* J Oral Microbiol, 2016. **8**: p. 31444.
77. Hu, J., et al., *Streptococcus pneumoniae and Haemophilus influenzae type b carriage in Chinese children aged 12-18 months in Shanghai, China: a cross-sectional study.* BMC Infect Dis, 2016. **16**(1): p. 149.
78. Kress-Bennett, J.M., et al., *Identification and Characterization of msf, a Novel Virulence Factor in Haemophilus influenzae.* PLoS One, 2016. **11**(3): p. e0149891.
79. Hu, J., et al., *[Multicenter prospective epidemiological studies on Haemophilus influenzae infection among hospitalized children with lower respiratory tract infections].* Zhonghua Er Ke Za Zhi, 2016. **54**(2): p. 119-25.
80. Greene, C.J., et al., *A novel strategy to protect against influenza-induced pneumococcal disease without interfering with commensal colonization.* Infect Immun, 2016.
81. Wang, K., et al., *Preliminary analysis of salivary microbiome and their potential roles in oral lichen planus.* Sci Rep, 2016. **6**: p. 22943.
82. Forstner, C., et al., *Community-acquired Haemophilus influenzae pneumonia - New insights from the CAPNETZ study.* J Infect, 2016. **72**(5): p. 554-63.
83. Kosikowska, U., et al., *Haemophilus parainfluenzae as a marker of the upper respiratory tract microbiota changes under the influence of preoperative prophylaxis with or without postoperative treatment in patients with lung cancer.* BMC Microbiol, 2016. **16**(1): p. 62.
84. Kosikowska, U., et al., *The Association of Chronic Hepatitis C with Respiratory Microbiota Disturbance on the Basis of Decreased Haemophilus Spp. Colonization.* Med Sci Monit, 2016. **22**: p. 625-632.
85. Chen, T., et al., *The Human Oral Microbiome Database: a web accessible resource for investigating oral microbe taxonomic and genomic information.* Database, 2010. **2010**.
86. Byrd, A.L. and J.A. Segre, *Infectious disease. Adapting Koch's postulates.* Science, 2016. **351**(6270): p. 224-6.
87. Bingham, E., B. Cohrssen, and C.H. Powell, *Patty's toxicology*. 5th ed. 2001, New York: Wiley.
88. Liu, G., C.M. Tang, and R.M. Exley, *Non-pathogenic Neisseria: members of an abundant, multi-habitat, diverse genus.* Microbiology, 2015. **161**(7): p. 1297-312.
89. Embree, M., et al., *Networks of energetic and metabolic interactions define dynamics in microbial communities.* Proc Natl Acad Sci U S A, 2015. **112**(50): p. 15450-5.

90. Crisol-Martinez, E., et al., *A low dose of an organophosphate insecticide causes dysbiosis and sex-dependent responses in the intestinal microbiota of the Japanese quail (Coturnix japonica).* PeerJ, 2016. **4**: p. e2002.
91. Joly, C., et al., *Impact of chronic exposure to low doses of chlorpyrifos on the intestinal microbiota in the Simulator of the Human Intestinal Microbial Ecosystem (SHIME) and in the rat.* Environ Sci Pollut Res Int, 2013. **20**(5): p. 2726-34.
92. Joly Condette, C., et al., *Chlorpyrifos Exposure During Perinatal Period Affects Intestinal Microbiota Associated With Delay of Maturation of Digestive Tract in Rats.* J Pediatr Gastroenterol Nutr, 2015. **61**(1): p. 30-40.
93. Joly Condette, C., et al., *Increased gut permeability and bacterial translocation after chronic chlorpyrifos exposure in rats.* PLoS One, 2014. **9**(7): p. e102217.
94. Gao, B., et al., *Gender-Specific Effects of Organophosphate Diazinon on the Gut Microbiome and Its Metabolic Functions.* Environ Health Perspect, 2016.
95. Chao, A. and T.J. Shen, *Nonparametric estimation of Shannon's index of diversity when there are unseen species in sample.* Environmental and Ecological Statistics, 2003. **10**(4): p. 429-443.
96. Siddaramappa, R., K.P. Rajaram, and N. Sethunathan, *Degradation of parathion by bacteria isolated from flooded soil.* Appl Microbiol, 1973. **26**(6): p. 846-9.
97. Sethunathan, N. and T. Yoshida, *A Flavobacterium sp. that degrades diazinon and parathion.* Can J Microbiol, 1973. **19**(7): p. 873-5.
98. Sethunathan, N., *Degradation of parathion in flooded acid soils.* J Agric Food Chem, 1973. **21**(4): p. 602-4.
99. Sethunathan, N. and T. Yoshida, *Parathion degradation in submerged rice soils in the Philippines.* J Agric Food Chem, 1973. **21**(3): p. 504-6.
100. Sethunathan, N., *Microbial degradation of insecticides in flooded soil and in anaerobic cultures.* Residue Rev, 1973. **47**: p. 143-65.
101. Singh, B.K., *Organophosphorus-degrading bacteria: ecology and industrial applications.* Nat Rev Microbiol, 2009. **7**(2): p. 156-64.
102. Below, J.E. and E.J. Parra, *Genome-Wide Studies of Type 2 Diabetes and Lipid Traits in Hispanics.* Curr Diab Rep, 2016. **16**(5): p. 41.
103. Below, J.E., et al., *Meta-analysis of lipid-traits in Hispanics identifies novel loci, population-specific effects, and tissue-specific enrichment of eQTLs.* Sci Rep, 2016. **6**: p. 19429.
104. Kaiser, P., et al., *Neighborhood Environments and Incident Hypertension in the Multi-Ethnic Study of Atherosclerosis.* Am J Epidemiol, 2016.
105. Pflederer, M.C., et al., *Longitudinal Changes in Vascular Risk Markers and Mortality Rates among a Latino Population with Hypertension.* Tex Heart Inst J, 2016. **43**(2): p. 131-6.
106. Rana, J.S., et al., *Accuracy of the Atherosclerotic Cardiovascular Risk Equation in a Large Contemporary, Multiethnic Population.* J Am Coll Cardiol, 2016. **67**(18): p. 2118-30.
107. Qato, D.M., et al., *Statin and Aspirin Use Among Hispanic and Latino Adults at High Cardiovascular Risk: Findings From the Hispanic Community Health Study/Study of Latinos.* J Am Heart Assoc, 2016. **5**(4).

108. Silverman, M.G., et al., *Impact of Race, Ethnicity, and Multimodality Biomarkers on the Incidence of New-Onset Heart Failure With Preserved Ejection Fraction (from the Multi-Ethnic Study of Atherosclerosis).* Am J Cardiol, 2016. **117**(9): p. 1474-81.
109. Meyer, M.L., et al., *Association of glucose homeostasis measures with heart rate variability among Hispanic/Latino adults without diabetes: the Hispanic Community Health Study/Study of Latinos (HCHS/SOL).* Cardiovasc Diabetol, 2016. **15**: p. 45.
110. Castaneda, S.F., et al., *Cardiovascular disease risk factors and psychological distress among Hispanics/Latinos: The Hispanic Community Health Study/Study of Latinos (HCHS/SOL).* Prev Med, 2016. **87**: p. 144-150.
111. Duparc, T., et al., *Hepatocyte MyD88 affects bile acids, gut microbiota and metabolome contributing to regulate glucose and lipid metabolism.* Gut, 2016.
112. Zhu, Z., S. Xiong, and D. Liu, *The Gastrointestinal Tract: an Initial Organ of Metabolic Hypertension?* Cell Physiol Biochem, 2016. **38**(5): p. 1681-94.
113. Emoto, T., et al., *Characterization of gut microbiota profiles in coronary artery disease patients using data mining analysis of terminal restriction fragment length polymorphism: gut microbiota could be a diagnostic marker of coronary artery disease.* Heart Vessels, 2016.
114. Kobyliak, N., O. Virchenko, and T. Falalyeyeva, *Pathophysiological role of host microbiota in the development of obesity.* Nutr J, 2016. **15**: p. 43.
115. Ferguson, J.F., et al., *Nutrigenomics, the Microbiome, and Gene-Environment Interactions: New Directions in Cardiovascular Disease Research, Prevention, and Treatment: A Scientific Statement From the American Heart Association.* Circ Cardiovasc Genet, 2016.
116. Xin, X., H. Junzhi, and Z. Xuedong, *[Oral microbiota: a promising predictor of human oral and systemic diseases].* Hua Xi Kou Qiang Yi Xue Za Zhi, 2015. **33**(6): p. 555-60.
117. Branchereau, M., et al., *Periodontal dysbiosis linked to periodontitis is associated with cardio-metabolic adaptation to high-fat diet in mice.* Am J Physiol Gastrointest Liver Physiol, 2016: p. ajpgi 00424 2015.
118. Tang, W.W. and S.L. Hazen, *Dietary metabolism, gut microbiota and acute heart failure.* Heart, 2016. **102**(11): p. 813-4.
119. Zhang, W., et al., *Deletion of the Toll-Like Receptor 5 Gene Per Se Does Not Determine the Gut Microbiome Profile That Induces Metabolic Syndrome: Environment Trumps Genotype.* PLoS One, 2016. **11**(3): p. e0150943.
120. Emoto, T., et al., *Analysis of Gut Microbiota in Coronary Artery Disease Patients: a Possible Link between Gut Microbiota and Coronary Artery Disease.* J Atheroscler Thromb, 2016.
121. Feng, Q., et al., *Integrated metabolomics and metagenomics analysis of plasma and urine identified microbial metabolites associated with coronary heart disease.* Sci Rep, 2016. **6**: p. 22525.
122. Mazidi, M., et al., *Gut microbiome and metabolic syndrome.* Diabetes Metab Syndr, 2016.
123. Wang, Y., et al., *High Molecular Weight Barley beta-Glucan Alters Gut Microbiota Toward Reduced Cardiovascular Disease Risk.* Front Microbiol, 2016. **7**: p. 129.
124. Gerardi, V., et al., *[The relationship between gut microbiota and cardiovascular diseases].* G Ital Cardiol (Rome), 2016. **17**(1): p. 11-4.

125. Budzynski, J., et al., *Association between Bacterial Infection and Peripheral Vascular Disease: A Review.* Int J Angiol, 2016. **25**(1): p. 3-13.
126. Claus, S.G., H; Ellero-Simatos, S, *The gut microbiota: a major player in the toxicity of environmental pollutants?* Biofilms and Microbiomes, 2016. **2**(16003).
127. Spanogiannopoulos, P., et al., *The microbial pharmacists within us: a metagenomic view of xenobiotic metabolism.* Nat Rev Microbiol, 2016. **14**(5): p. 273-87.
128. Vallès, Y., et al., *Microbial Succession in the Gut: Directional Trends of Taxonomic and Functional Change in a Birth Cohort of Spanish Infants.* PLoS Genet, 2014. **10**(6): p. e1004406.
129. Thaiss, C.A., et al., *The microbiome and innate immunity.* Nature, 2016. **535**(7610): p. 65-74.
130. Bokulich, N.A., et al., *Antibiotics, birth mode, and diet shape microbiome maturation during early life.* Sci Transl Med, 2016. **8**(343): p. 343ra82.
131. Arukwe, A., C.C. Carteny, and T. Eggen, *Lipid peroxidation and oxidative stress responses in juvenile salmon exposed to waterborne levels of the organophosphate compounds tris(2-butoxyethyl)- and tris(2-chloroethyl) phosphates.* J Toxicol Environ Health A, 2016. **79**(13-15): p. 515-25.
132. Nwani, C.D., et al., *Genotoxicity assessment and oxidative stress responses in freshwater African catfish Clarias gariepinus exposed to fenthion formulations.* Drug Chem Toxicol, 2016: p. 1-8.
133. Pearson, J.N. and M. Patel, *The role of oxidative stress in organophosphate and nerve agent toxicity.* Ann N Y Acad Sci, 2016.
134. Smith, A.D., et al., *Model Selection of the Effect of Binary Exposures over the Life Course.* Epidemiology, 2015. **26**(5): p. 719-26.
135. Schwarzer, M., et al., *Lactobacillus plantarum strain maintains growth of infant mice during chronic undernutrition.* Science, 2016. **351**(6275): p. 854-7.
136. Korpela, K., et al., *Intestinal microbiome is related to lifetime antibiotic use in Finnish pre-school children.* Nat Commun, 2016. **7**: p. 10410.
137. Honda, K. and D.R. Littman, *The microbiota in adaptive immune homeostasis and disease.* Nature, 2016. **535**(7610): p. 75-84.
138. Yee, A.L. and J.A. Gilbert, *MICROBIOME. Is triclosan harming your microbiome?* Science, 2016. **353**(6297): p. 348-9.
139. Gilbert, J.A., et al., *Microbiome-wide association studies link dynamic microbial consortia to disease.* Nature, 2016. **535**(7610): p. 94-103.
140. Lockridge, O., *Review of human butyrylcholinesterase structure, function, genetic variants, history of use in the clinic, and potential therapeutic uses.* Pharmacol Ther, 2015. **148**: p. 34-46.
141. Shapira, M., et al., *A transcription-activating polymorphism in the ACHE promoter associated with acute sensitivity to anti-acetylcholinesterases.* Hum Mol Genet, 2000. **9**(9): p. 1273-81.
142. Stitziel, N.O., et al., *Exome sequencing in suspected monogenic dyslipidemias.* Circ Cardiovasc Genet, 2015. **8**(2): p. 343-50.
143. Guerrette, Z.N., *Integrating Biomarkers of Organophosphate Pesticides in an Agriculturally Exposed Population*, in *Department of Environmental and Occupational Health Sciences*. 2012, University of Washington.

144. Gamazon, E.R., et al., *Comprehensive survey of SNPs in the Affymetrix exon array using the 1000 Genomes dataset.* PLoS One, 2010. **5**(2): p. e9366.
145. Gamazon, E.R., et al., *A pharmacogene database enhanced by the 1000 Genomes Project.* Pharmacogenet Genomics, 2009. **19**(10): p. 829-32.
146. Zheng-Bradley, X. and P. Flicek, *Applications of the 1000 Genomes Project resources.* Brief Funct Genomics, 2016.
147. Devuyst, O., *The 1000 Genomes Project: Welcome to a New World.* Perit Dial Int, 2015. **35**(7): p. 676-7.
148. Akiyama, Y., et al., *Immune response-associated gene analysis of 1,000 cancer patients using whole-exome sequencing and gene expression profiling-Project HOPE.* Biomed Res, 2016. **37**(4): p. 233-42.
149. Rosenthal, E.A., et al., *Association Between Absolute Neutrophil Count and Variation at TCIRG1: The NHLBI Exome Sequencing Project.* Genet Epidemiol, 2016. **40**(6): p. 470-4.
150. Kan, M., et al., *Rare variant associations with waist-to-hip ratio in European-American and African-American women from the NHLBI-Exome Sequencing Project.* Eur J Hum Genet, 2016. **24**(8): p. 1181-7.
151. Kim, D.S., et al., *Rare coding variation in paraoxonase-1 is associated with ischemic stroke in the NHLBI Exome Sequencing Project.* J Lipid Res, 2014. **55**(6): p. 1173-8.
152. Duan, Q., et al., *Imputation of coding variants in African Americans: better performance using data from the exome sequencing project.* Bioinformatics, 2013. **29**(21): p. 2744-9.
153. Johnsen, J.M., et al., *Common and rare von Willebrand factor (VWF) coding variants, VWF levels, and factor VIII levels in African Americans: the NHLBI Exome Sequencing Project.* Blood, 2013. **122**(4): p. 590-7.
154. Auer, P.L., et al., *Imputation of exome sequence variants into population- based samples and blood-cell-trait-associated loci in African Americans: NHLBI GO Exome Sequencing Project.* Am J Hum Genet, 2012. **91**(5): p. 794-808.
155. Sun, R., et al., *A W-test collapsing method for rare-variant association testing in exome sequencing data.* Genet Epidemiol, 2016.
156. Timbers, T.A., et al., *Accelerating Gene Discovery by Phenotyping Whole-Genome Sequenced Multi-mutation Strains and Using the Sequence Kernel Association Test (SKAT).* PLoS Genet, 2016. **12**(8): p. e1006235.
157. Haddad, S.A., et al., *An exome-wide analysis of low frequency and rare variants in relation to risk of breast cancer in African American Women: the AMBER Consortium.* Carcinogenesis, 2016. **37**(9): p. 870-7.
158. Wu, B. and J.S. Pankow, *On Sample Size and Power Calculation for Variant Set-Based Association Tests.* Ann Hum Genet, 2016. **80**(2): p. 136-43.
159. Fan, R., et al., *Gene-Based Association Analysis for Censored Traits Via Fixed Effect Functional Regressions.* Genet Epidemiol, 2016. **40**(2): p. 133-43.
160. Zheng, X., et al., *A high-performance computing toolset for relatedness and principal component analysis of SNP data.* Bioinformatics, 2012. **28**(24): p. 3326-8.
161. Staples, J., et al., *PRIMUS: rapid reconstruction of pedigrees from genome-wide estimates of identity by descent.* Am J Hum Genet, 2014. **95**(5): p. 553-64.

162. Scheet, P. and M. Stephens, *A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase.* Am J Hum Genet, 2006. **78**(4): p. 629-44.
163. Chambers, J.M., *Statistical Models in S*, ed. J.M.C.a.T.J. Hastie. 1992: Wadsworth and Brooks/Cole.
164. Zeileis, A., *Econometric Computing with HC and HAC Covariance Matrix Estimators.* Journal of Statistical Software, 2004. **11**(10): p. 1-17.
165. Zeileis, A., *Object-oriented Computation of Sandwich Estimators.* Journal of Statistical Software, 2006. **16**(9): p. 1-16.
166. Box, G., Jenkins, G., Reinsel, G., *Time Series Analysis: Forecasting and Control*. 3rd Edition. 1994: Holden-Day.
167. Wang, Y., et al., *Resistance to organophosphorus agent toxicity in transgenic mice expressing the G117H mutant of human butyrylcholinesterase.* Toxicol Appl Pharmacol, 2004. **196**(3): p. 356-66.
168. Bartels, C.F., et al., *DNA mutation associated with the human butyrylcholinesterase K-variant and its linkage to the atypical variant mutation and other polymorphic sites.* Am J Hum Genet, 1992. **50**(5): p. 1086-103.
169. McGuire, M.C., et al., *Identification of the structural mutation responsible for the dibucaine-resistant (atypical) variant form of human serum cholinesterase.* Proc Natl Acad Sci U S A, 1989. **86**(3): p. 953-7.
170. Nogueira, C.P., et al., *Identification of two different point mutations associated with the fluoride-resistant phenotype for human butyrylcholinesterase.* Am J Hum Genet, 1992. **51**(4): p. 821-8.
171. Sim, X., et al., *Transferability of type 2 diabetes implicated loci in multi-ethnic cohorts from Southeast Asia.* PLoS Genet, 2011. **7**(4): p. e1001363.
172. Valle, A.M., et al., *Naturally occurring variations in the human cholinesterase genes: heritability and association with cardiovascular and metabolic traits.* J Pharmacol Exp Ther, 2011. **338**(1): p. 125-33.
173. Rosenbloom, K.R., et al., *The UCSC Genome Browser database: 2015 update.* Nucleic Acids Res, 2015. **43**(Database issue): p. D670-81.
174. Speir, M.L., et al., *The UCSC Genome Browser database: 2016 update.* Nucleic Acids Res, 2016. **44**(D1): p. D717-25.
175. Bartels, C.F., T. Zelinski, and O. Lockridge, *Mutation at codon 322 in the human acetylcholinesterase (ACHE) gene accounts for YT blood group polymorphism.* Am J Hum Genet, 1993. **52**(5): p. 928-36.
176. Strelitz, J., L.S. Engel, and M.C. Keifer, *Blood acetylcholinesterase and butyrylcholinesterase as biomarkers of cholinesterase depression among pesticide handlers.* Occup Environ Med, 2014. **71**(12): p. 842-7.
177. Harrison, K.L., *Fetal erythrocyte lifespan.* Aust Paediatr J, 1979. **15**(2): p. 96-7.
178. Moser, V.C. and S. Padilla, *Age- and gender-related differences in the time course of behavioral and biochemical effects produced by oral chlorpyrifos in rats.* Toxicol Appl Pharmacol, 1998. **149**(1): p. 107-19.
179. Moser, V.C., *Comparison of aldicarb and methamidophos neurotoxicity at different ages in the rat: behavioral and biochemical parameters.* Toxicol Appl Pharmacol, 1999. **157**(2): p. 94-106.

180. Moon, J., B. Chun, and S. Lee, *Variable response of cholinesterase activities following human exposure to different types of organophosphates.* Hum Exp Toxicol, 2015. **34**(7): p. 698-706.
181. Rubinstein, H.M., et al., *Silent cholinesterase gene: variations in the properties of serum enzyme in apparent homozygotes.* J Clin Invest, 1970. **49**(3): p. 479-86.
182. Eskildsen, K.Z. and M.R. Gatke, *[Danish Cholinesterase Research Unit diagnoses patients with prolonged paralysis after succinylcholine and mivacurium].* Ugeskr Laeger, 2014. **176**(12A).
183. Thomsen, J.L. and M.R. Gatke, *Patient information sheet for plasma cholinesterase deficiency.* Anaesthesia, 2014. **69**(11): p. 1287-8.
184. Bray, J.R. and J.T. Curtis, *An Ordination of the Upland Forest Communities of Southern Wisconsin.* Ecological Monographs, 1957. **27**(4): p. 326-349.

7. APPENDICES

Appendix 1. Correspondence with Oksana Lockridge.

Lockridge, Oksana <olockrid@unmc.edu> 8/14/15
to me

Dear Ian Stanaway,
As far as I know there is no convincing evidence to date that people who are deficient in butyrylcholinesterase are at increased risk of toxicity from organophosphorus pesticides. It makes sense that it should be so, but there is no scientific evidence.

Oksana Lockridge
985950 Nebraska Medical Center
Omaha, NE 68198-5950 USA
Phone 402 559 6032 FAX 402 559-4651
olockrid@unmc.edu

Pharmacology & Therapeutics 148 (2015) 34–46
Contents lists available at ScienceDirect
Pharmacology & Therapeutics
journal homepage: www.elsevier.com/locate/pharmthera
ELSEVIER

Associate editor: C.N. Pope

Review of human butyrylcholinesterase structure, function, genetic variants, history of use in the clinic, and potential therapeutic uses

Oksana Lockridge *
Eppley Institute, University of Nebraska Medical Center, Omaha, NE 68198-5950, USA

NeuroToxicology® 21(1-2): 113-126, 2000
Copyright © 2000 by Intox Press, Inc.

Pesticides and Susceptible Populations: People With Butyrylcholinesterase Genetic Variants May Be At Risk

Oksana Lockridge[1] and Patrick Masson[2]

[1]University of Nebraska Medical Center, Eppley Institute, Omaha, NE 68198-6805, [2]Centre de Recherches du Service de Santé des Armées, Unité d'Enzymologie, BP 87, 38702 La Tronche Cédex, France

Figure 7-1 Email from Oksana Lockridge Concerning BCHE Variants and Risk