January 2014

Novel Approaches Investigating Predisposition For Maternal Group B Streptococcus Colonization

Michelle Lynn Wright

NOVEL APPROACHES INVESTIGATING PREDISPOSITION FOR MATERNAL GROUP B STREPTOCOCCUS COLONIZATION

by

Michelle Lynn Wright
Bachelor of Science, Western Michigan University, 2001
Master of Science, DePaul University, 2004

A Dissertation
Submitted to the Graduate Faculty

of the

University of North Dakota

In partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Grand Forks, North Dakota
May
2014

This dissertation, submitted by Michelle Lynn Wright in partial fulfillment of the requirements for the Degree of Doctor of Philosophy from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done, and is hereby approved.

_____
Dr. Denise M. Korniewicz, Chairperson

_____
Dr. Cindy M. Anderson

_____
Dr. Jody L. Ralph

_____
Dr. Joyce E. Ohm

_____
Dr. David S. Bradley

_____
Dr. Maher M. El-Masri

This dissertation is being submitted by the appointed advisory committee as having met all the requirements of the Graduate School at the University of North Dakota and is hereby approved.

_____
Dr. Wayne E. Swisher
Dean of the School of Graduate Studies

_____
Date

Title	Novel Approaches Investigating Predisposition for Maternal Group B Streptococcus Colonization
Department	Department of Nursing
Degree	Doctor of Philosophy

In presenting this dissertation in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my dissertation work or, in her absence, by the Chairperson of the department of the dean of the Graduate School. It is understood that any copying or publication or other use of this dissertation or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my dissertation.

Michelle Lynn Wright
February 15, 2014

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I wish to express my deep appreciation and gratitude to my advisor, Dr. Denise Korniewicz, for the patient guidance and mentorship she provided to me. Her support from the time of my interview, through transferring schools and research topics has helped me grow immensely as a nurse scientist. Dr. Korniewicz's intellect, strength and forward thinking have left an indelible mark on my analytic skills; and for that I am forever grateful. I am fortunate to have had the opportunity to work with her.

I would also like to thank the other members of my committee for their support and contributions to my intellectual growth. Dr. Cindy Anderson for her unyielding patience and positive reinforcement in my knowledge and abilities. I am eternally grateful she allowed me to complete secondary analyses on samples and data she had collected for a previous study. Her kindness, support, and socialization into research circles are invaluable and much appreciated. I am also appreciative to the staff at the Grand Forks Human Nutrition Research Center for the laboratory analyses they completed. Dr. Jody Ralph also offered me countless hours of support, without which it would have been near impossible for me to complete my educational endeavor. I am eternally grateful for her support, suggestions and mentorship throughout my time at the University of North Dakota. Additionally, I would like to thank Dr. Joyce Ohm for allowing me the opportunity to conduct bench work in her laboratory. Her intellect and kindness in allowing a student from another school on campus participate in some of her own research added to my intellectual development during my years of study at the University of North Dakota. For that, I am thankful. The support and kindness of her laboratory manager, Angela Floden, will also never be forgotten. Lastly, I would like to thank Dr. El-Masri for agreeing to join my committee as an additional member. His statistical expertise and editorial support have helped me grow immensely as a scientist. He has taught me skills I will apply throughout my research career.

I would like to thank Dr. Keith Henry and his assistant Danielle Perley for initially assisting me with R. Additionally, I would like to thank Drs. Kellie Archer and Timothy York at Virginia Commonwealth University and Patrick Barry at University of Alaska for their support coding in R and verifying my code accuracy. I am forever grateful for your knowledge and assistance.

With much love, I would like to thank my family and friends for their never ending support and encouragement. I could not have accomplished this feat without your kind words, shoulders to cry on, and meals (which always fed me more than just food). Thank you, you all mean the world to me.

ABSTRACT

Group B streptococcus (GBS) is the leading cause of infectious neonatal morbidity and mortality in the United States. Maternal GBS colonization is the primary risk factor associated with neonatal infection. However, maternal risk factors for GBS colonization are ambiguous. A conceptual framework of gene-environment interactions guided the approach for this study analyzing DNA methylation, serum cytokines, and vitamin D levels. The purpose of this study was to identify potential maternal biomarkers associated with GBS colonization. Descriptive statistics were conducted to depict sample characteristics (n=42 pregnant women) and identify potential confounding variables including, but not limited to: medical history, race, weight, and infections. A series of repeated measures ANOVAs were performed to compare each of three serum cytokines (TNF-α, IL-6 and IL-10) and vitamin D levels between the two groups in each trimester of pregnancy. All statistical analysis was completed using a two-tailed alpha of < 0.05 or 95% confidence interval. Mean differences of greater than 20% in DNA methylation of maternal white blood cells collected in the first trimester were analyzed using a false discovery rate of 0.05 to determine significance, as well as independent sample t-tests with a p-value of 0.05 using the Illumina Infinium platform and grouped by GBS status (n=9/group) identified in the third trimester. Function of differentially methylated genes was determined using DAVID Bioinformatics software to identify clinically relevant findings. No statistically significant differences in IL-6 $F(2, 80) = 2.99$, $p = 0.056$; IL-10 $F(2, 80) = 0.445$, $p = 0.642$; TNF-α $F(2, 80) = 2.187$, $p = 0.119$; or vitamin D $F(1.380,$

55.218) = 0.882, p = 0.384 were identified between GBS positive and negative women during pregnancy. Analysis of DNA methylation indicates there are no statistically significant differences between GBS positive and GBS negative women using and FDR of 0.05. When a less stringent p-value of 0.05 was applied, 125 CpG sites differed by 20% or more between GBS positive versus negative women and different results are yielded using multiple statistical approaches (GenomeStudio versus R). Functional analysis suggests genes with methylation differences in the cell morphogenesis cluster may be associated with GBS colonization, although the significance is questionable.

CHAPTER I

INTRODUCTION

Group B streptococcus (GBS) is the leading cause of neonatal morbidity and mortality due to infection in the United States (Phares et al., 2008). Maternal GBS colonization is the primary risk factor associated with the development of neonatal GBS sepsis (Verani, McGee, Schrag, & Division of Bacterial Diseases Centers for Disease Control and Prevention (CDC), 2010). While between 10 – 30 percent of pregnant women are colonized with GBS (Schrag et al., 2002), risk factors for maternal colonization are ambiguous and inconclusive in the literature. To prevent transmission of GBS, colonized women are normally given antibiotics during the intrapartum period which significantly reduces the incidence of early onset GBS infections in neonates (Verani et al., 2010). However, current screening techniques have a 10% false negative rate (Towers et al., 2010) and do not prevent preterm labor, miscarriages, and stillbirths caused by GBS colonization; nor do they reduce the incidence of late onset GBS sepsis in infants (Clifford, Garland, & Grimwood, 2011; Jordan et al., 2008; Verani et al., 2010). There is a compelling need to investigate genetic and environmental factors that may help identify biomarkers for colonization because GBS continues to cause poor pregnancy outcomes and is associated with the absence of definitive maternal risk factors for colonization. If genetic and environmental factors can be identified, early screening and effective interventions can be developed and implemented. Preliminary data from our laboratory indicated DNA methylation differences can be measured early in pregnancy

between women with and without late pregnancy GBS colonization. DNA methylation, an epigenetic modification that can result in altered gene expression and related protein production, has the potential to drastically impact health and alter disease susceptibility (Baccarelli, Rienstra, & Benjamin, 2010; Berger, Kouzarides, Shiekhattar, & Shilatifard, 2009; Rodenhiser & Mann, 2006). Differential DNA methylation in genes regulating immunity and inflammation could lead to varied levels of pro-inflammatory tumor necrosis factor-alpha (TNF-α), interleukin-6 (IL-6), and anti-inflammatory interleukin-10 (IL-10). TNF α, IL-6, and IL-10 are proteins that are produced in varying amounts in response to threats to the immune system and could be clinical laboratory indicators for GBS colonization (Berner, Welter, & Brandis, 2002; Fan et al., 2003; Madureira et al., 2011; Maisey, Doran, & Nizet, 2008; Mikamo, Johri, Paoletti, Madoff, & Onderdonk, 2004; Ng et al., 2003; Parameswaran & Patial, 2010; Puliti et al., 2002; Santhanam et al., 1991; Vieira et al., 1991). Additionally, serum markers associated with immune function and vitamin D (25[OH]D) status have previously been identified and utilized as prognostic indicators for infectious disease (Chesney, 2010; Fahey et al., 1990) and may be a cost effective clinical intervention if altered serum vitamin D (25[OH]D) levels are associated with GBS colonization status. The long term goal for investigating different exogenous and endogenous clinical indicators in women with and without GBS colonization is to identify factors that may help later identify a mechanistic explanation for maternal GBS colonization and to develop and implement targeted primary prevention strategies to reduce neonatal sepsis caused by GBS. The specific purpose of this study was to identify variants in maternal blood that are associated with maternal GBS colonization.

Significance

Streptococcus agalactiae, or group B β-hemolytic streptococcus, is a gram positive bacterium that causes a wide spectrum of illness in multiple clinical populations. In 2005, GBS was attributed as the cause of 21,500 infections and 1,700 deaths in the United States (Phares et al., 2008). However, the rates of GBS colonization, independent of infection, were not reported or required to be reported. Primarily a bacterium that only causes invasive disease in patients with altered immune function (Johri et al., 2006), GBS was first associated with neonatal sepsis and maternal infection in the 1970s and continues to be the leading cause of neonatal sepsis (Verani et al., 2010). Maternal colonization with GBS is the strongest predictor for the development of GBS sepsis in the neonate. Currently, there are no effective strategies for preventing maternal GBS colonization because of the inconsistent results obtained from epidemiological studies (Clifford et al., 2011; Kovavisarach, Ying, & Kanjanahareutai, 2007; Zusman, Baltimore, & Fonseca, 2006). Studies investigating the number of pregnancies, maternal age, race, ethnicity, and other maternal characteristic have failed to identify common maternal risk factors. The variability in potential predisposing factors has resulted in research efforts focused on preventing transmission and subsequent development of GBS infection in neonates.

Compounding the issue of GBS infection in neonates is the fact that their risk for infection continues into the first 3 months of life. GBS infections are categorized in neonates by time of disease onset after delivery as either early or late onset. Early onset neonatal GBS sepsis occurs within the first 7 days of life with onset usually occurring within the first 48 hours of life. The time period ascribed to late onset neonatal GBS

sepsis is an infection that develops after the first week of life up until the infant is 3 months of age (Verani et al., 2010). In an attempt to circumvent GBS transmission to the neonate from colonized women, clinical guidelines recommend maternal screening for GBS colonization via cultures obtained from recto-vaginal swabs between 35-37 weeks gestation. If maternal screening tests are positive for GBS colonization, the CDC recommends intravenous antibiotics administration after the start of labor and at least 4 hours prior to delivery to prevent transmission of GBS to the neonate (Verani et al., 2010). Implementation of the CDC guidelines has resulted in a 80% decrease in the incidence of early onset neonatal GBS sepsis, although rates of late onset GBS sepsis have been unaffected (Schrag & Verani, 2013). The continued prevalence of early and late onset neonatal GBS sepsis since the guideline implementation could be a result of the 10% false negative rate associated with maternal screening for GBS colonization (Towers et al., 2010). Of the infants who develop GBS infections, 61-82 percent of infants are born to mothers with negative GBS screening at 35-37 weeks gestation and 6.3% of infants are colonized by GBS despite administration of maternal antibiotic treatment (Lin et al., 2011). This may be in part because colonization with GBS can be transient, intermittent, or persistent. Maternal GBS status could change between the time of screening and delivery. Additionally, the infant may become infected from environmental exposure after delivery (Verani et al., 2010). The transient nature of GBS colonization suggests that environmental factors or immune response may play a critical role in maternal colonization, warranting further investigation into maternal risk factors and more reliable screening.

In order to reach the Healthy People 2020 goal of a 10% reduction in neonatal GBS infections (U. S. Department of Health and Human Services (HHS), 2013), alternative interventions are required for two primary reasons. First, the 10% false negative rate during screening needs to be reduced. Women who screen negative for GBS colonization will not receive antibiotic prophylactic treatment and could contribute to the continued prevalence of GBS disease in neonates. Furthermore, it is unclear why some women initially screen negative and later convert to GBS positive status. The transient nature of some GBS colonization implies there may be environmental factors that contribute to colonization susceptibility. Second, increasing rates of antimicrobial resistance to intrapartum antibiotics are being reported. The antibiotic currently recommended for treatment of GBS colonization during pregnancy is intravenous penicillin (Verani et al., 2010). However, approximately 10% of the general population report having a penicillin allergy (Solensky, 2003). Patients with allergies are administered erythromycin with clindamycin or vancomycin instead. In the US, over 54% of invasive strains of GBS are resistant to erythromycin and 33% are resistant to clindamycin (DiPersio & DiPersio, 2006) and similar rates of resistance are seen worldwide (Bergseng, Rygg, Bevanger, & Bergh, 2008; Janapatla, Ho, Yan, Wu, & Wu, 2008; Uh et al., 2007). Dual resistance to both clindamycin and erythromycin is also increasing, with rates reported as high as 94% of clindamycin resistant isolates also being resistant to erythromycin (Back, O'Grady, & Back, 2011). Additionally, Stoll et al. (2011) found 53% of infants that developed early onset sepsis were born to mothers who had received intrapartum antibiotic treatment. The failure rates associated with current treatment methods are associated with continuing neonatal morbidity and moratlity. New

strategies which are robust at identifying women at risk for GBS colonization may increase the likelihood of developing successful intervention alternatives, further reducing the incidence of neonatal GBS sepsis.

The significance of the problem surrounding maternal GBS colonization is compounded by the fact that current guidelines to prevent neonatal GBS do not prevent poor maternal and fetal outcomes associated with GBS colonization. Administering antibiotics during labor does not prevent stillbirths, miscarriages, chorioamnionitis, or other poor pregnancy outcomes associated with GBS prior to 35 weeks gestation. Furthermore, there has been a 32% increase in GBS infections in non-pregnant adults with no information on colonization rates available (Phares et al., 2008). It is unclear why there has been such a large increase in the incidence of GBS infections.

Ambiguous risk factors for colonization, increasing rates of bacterial resistance to antibiotics, increasing GBS infections in non-pregnant populations, and failure rates associated with current treatment methods are compelling reasons to discover new approaches to identify individuals at risk for GBS colonization. Identification of factors associated with GBS colonization will provide new clinically relevant targets to prevent and treat GBS colonization. Identification of definitive environmental and/or genetic maternal risk factors associated with GBS colonization is a substantively different approach to preventing neonatal GBS sepsis. This contribution is significant and will improve scientific knowledge by identifying key differences in pregnant women with and without GBS colonization. If differences are identified between women with and without GBS colonization, knowledge gained from this study can be used to develop a more

accurate screening tool or prevent maternal GBS colonization thereby improving clinical practice and pregnancy outcome.

Purpose

The purpose of this study was to identify variants in maternal blood that are associated with maternal GBS colonization by investigating the following specific aims:

1. Differentiate serum levels of TNF-α, IL-6, and IL-10 as potential clinical laboratory indicators for GBS colonization longitudinally in pregnant women; and examine circulating 25-hydroxyvitamin D (25[OH]D) (vitamin D (25[OH]D) as a potential covariate of TNF-α, IL-6, and IL-10 serum levels.
2. Determine if DNA methylation are different in pregnant women with and without GBS colonization.
3. Examine the relationship of any differentially methylated genes for association with immune function and inflammatory serum markers in pregnant women colonized with GBS.

Conceptual Framework

Investigations of the interaction between genomic and environmental factors have been suggested as key research avenues in identifying the most effective methods to prevent disease. Cohesive investigations illuminate biochemical explanations for health problems and identify modifiable risk factors that can be controlled or altered to prevent disease (Willett, 2002). The relationship between genetic and environmental factors associated with the development of disease was first suggested in 1902 (Hunter, 2005). Examples of conditions known to be affected by genetic and environmental factors include: halitosis (Bretz et al., 2011), head and neck cancer after human papilloma virus

exposure (Jamaly et al., 2012), sun exposure and skin cancer (Rees, 2004), and an increased susceptibility to human immunodeficiency virus infection based on cytokine profiles (Smith et al., 1997). Hunter (2005) published a conceptual model (Figure 1)

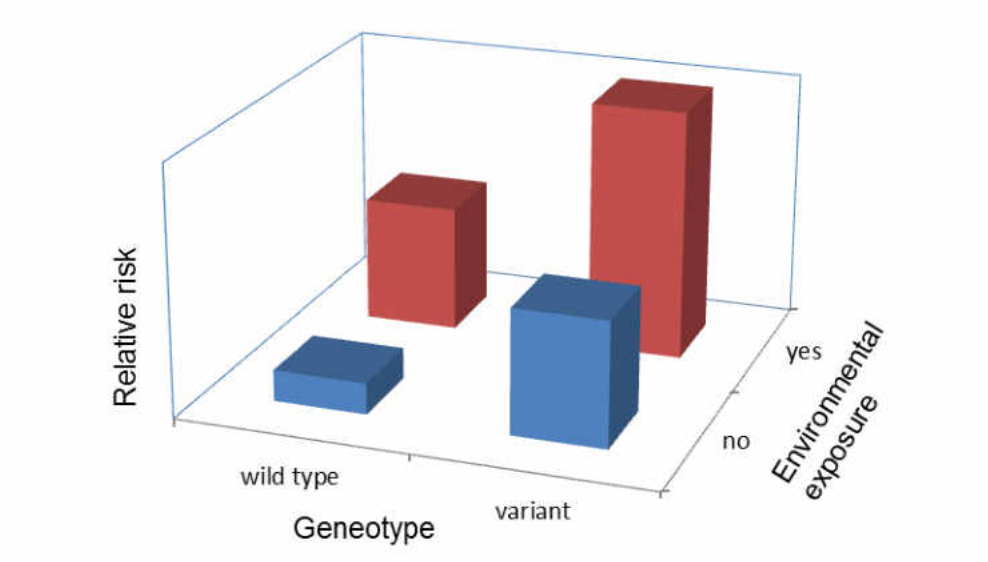


Figure 1. Gene-environment interaction.

illustrating how gene-environment interactions can potentiate disease processes. In the model, either a genetic variant or environmental exposure may result in disease even in the absence of an interaction between the two. The gene-environment model described by Hunter shows relative risk for disease states can be the result of either 1) a genetic variant 2) an environmental factor or 3) the interaction of the two. However, when there is an interaction between a predisposing genetic variant and environmental exposure, the risk for disease may be amplified. Hunter has used the inherited condition xeroderma pigmentosum to illustrate the model. Inheritance of the genetic mutation that causes the disorder greatly increases the risk for skin cancer and sun exposure further elevates risk (Cleaver, 2005; Kraemer, Lee, Andrews, & Lambert, 1994). When both elements are present, it results in a multiplicative effect drastically increasing disease risk (Figure 1). In other words the most basic interpretation of this relationship can be expressed as a

mathematical equation, genetic variants (G) x environmental exposure (E) = change in disease susceptibility (D):

$$G \times E = D$$

Another example is that of emphysema. Individuals with the genetic mutation (G) causing an alpha 1-antitrypsin deficiency are likely to develop emphysema at a young age (DeMeo & Silverman, 2004). Smoking is an environmental exposure (E) linked to development of emphysema. An individual who has the mutation and smokes may develop more severe emphysema (ΔD) at an earlier age than a person with the mutation who does not smoke.

Genetic Variants

The identification of the underlying genetic mechanism for an observable characteristic, referred to as a phenotype, is a common undertaking in the field of genetics. The genetic code contained within the nucleus of the cell is the genotype. The combination of dominant and recessive alleles, commonly called genes, and non-coding regions of the genome make up an individual's genotype. The genotype serves as a blueprint, contributing to the development of the phenotype. Copy number variants, insertions, deletions, and single nucleotide polymorphisms are some examples of how changes in the genetic code can lead to phenotypic changes. Hunter (2005) does not define genetic variants when presenting his model. However, the most simplistic model compares the "genotype" of individuals with or without a given trait to illustrate how environment may impact the observed phenotype based on genotype. The complexity of the model increases when multiple genes contribute to the development of a single phenotype. Eye color is an example of a complex phenotype resulting from multiple

genes (Liu et al., 2009). Hunter (2005) did not include epigenetic modifications in his discussion of genetic variants, although he did acknowledge the plausibility of assessing factors altering gene expression in G x E interactions. Since the introduction of this simple model, the field has advanced to include methods for measuring epigenetic mechanisms that alter gene expression. Considering epigenetic alterations have the ability to augment phenotypic expression; inclusion in, or extension of, the model is logical.

Environmental Variants

The environment has a profound impact on the health of individuals. While Hunter (2005) does not specifically define environmental variants in his model, he does consistently state that "environmental and lifestyle" factors must be assessed, implying a broad definition of environment. Ottman (1996) previously defined environmental exposures in gene-environment studies as:

> The environmental risk factor can be an exposure, either physical (e.g., radiation, temperature), chemical (e.g., polycyclic aromatic hydrocarbons), or biological (e.g., a virus); a behavior pattern (e.g., late age at first pregnancy); or a "life event" (e.g., job loss, injury). This is not intended as an exhaustive taxonomy of risk factors, but indicates as broad a definition as possible of environmental exposures. (p. 764-765)

Further, epidemiologists are experts at identifying associations between environmental exposure and disease processes. Unfortunately, many epidemiological studies fail to collect DNA samples making it impossible to assess the genetic variation in large samples of unrelated subjects. Decreasing the existing disconnect between epidemiologic and genetic analysis is possible when factors identified in epidemiologic studies are

assessed in genetic research involving human subjects. Most genetic studies involving human subjects collect some information about participants. Data may be limited to demographic information or include detailed laboratory values and other potentially relevant risk factors. When designing genetic studies with human subjects, improved assessment and collection of environmental exposure data could offer insight into disease processes (Hunter, 2005).

Gene-environment Interactions

Hunter (2005) describes two possible approaches for interpreting gene-environment interactions that contribute to disease in the model. The interpretation of the interaction is dependent on the statistical model selected, and must be appropriate for the type of clinical question being addressed. When scientists are interested in determining how factors contribute to the relative risk of a disease, they assume a multiplicative interaction where the risk is either increased or decreased when multiple factors contribute to disease development. Results from studies using a multiplicative approach usually report findings in terms of relative risk. Using the emphysema example from before, the probability of developing emphysema early is more likely to occur if a person with the genetic defect smokes. The probability of how likely it is for an outcome to occur is referred to as the relative risk.

If the *a priori* assumption is the interaction is a joint effect, the relationship is additive. To illustrate this type of relationship a clinical example where diagnosis of a disease is usually based clinical presentation will be used as an exemplar. For example, a patient presents to their primary care provider with skin lesions, a fever, and headache and is subsequently diagnoses with varicella by the provider. Each symptom does not

does not cause varicella; it is caused by a virus that results in the presentation of a specific a set of symptoms. The collection of symptoms together results in the clinical presentation caused by the viral infection. Studies utilizing an additive approach methodology usually report findings as rate differences, such as 80% of patients with chicken pox have a headache. Hunter encourages explicitly stating if the relationship assumed for analysis is multiplicative (relative risk) or additive (percent affected) in order to appropriately replicate and compare research studies. Research investigating gene-environment interactions has the potential to improve and individualize patient care by improving understanding of disease susceptibility allowing for development of alternate treatment and prevention strategies.

Modified Conceptual Framework

Epigenomic research has led to a greater understanding of how our genes and environment contribute to complex disease processes. Epigenome adds the Greek prefix "epi" to genome and literally translates to above the genome (epi, n.d). The epigenome is fundamental for normal human development and contributes to what makes individuals unique. Epigenomics is the study of heritable alterations in the chromosomes which do not change the DNA sequence itself, but result in a specific phenotype (Berger et al., 2009). Structural and functional modifications of the epigenome modulate expression of the genes encoded by DNA. Nurses, as members of interdisciplinary teams, can use advances in epigenomic techniques to better assess levels of health and disease risk. Additionally, it is important when conceptualizing environment in the extended model that endogenous and exogenous environmental variants be considered. The environment should be considered anything outside of the DNA because any exposure to the DNA

could alter the epigenomic signature. Epigenomic alterations should be included in studies investigating gene-environment interactions due to the dynamic, and potentially reversible, nature of the epigenome that can be modulated by endogenous and exogenous influences throughout the lifespan (Feinberg, 2008).

Moffitt, Caspi, & Rutter (2006) argue against incorporating epigenetic mechanisms into gene-environment interaction models because the alterations modulate the effects of environment on gene expression and do not represent actual alterations in the genes or the DNA sequence. This is true in some cases, but not all. For example, women with breast cancer type 2 susceptibility protein (BRCA2) mutations have a higher risk for developing ovarian cancer (Kanchi et al., 2014; King, 2003; Welcsh, 2001). Additionally, hypomethylation of BRCA2 in ovarian tumor DNA has been associated with advanced tumor staging (Chan, Ozcelik, Cheung, Ngan, & Khoo, 2002). In other words, there is an existing genetic variant that results in increased risk of a disease and the severity of disease is modulated by epigenetic variants. However, during embryonic development tissue specific DNA methylation patterns are established across the genome (Hajkova et al., 2002) and the epigenetic signature can be altered by environmental exposures. Since epigenetic (eG) modifications can occur as a direct result of environmental exposures (E) resulting in altered susceptibility to disease (ΔD) without a genetic variant being present, the same conceptual model can be applied to disease states without mutations.

Summary of Framework

Since publication of Hunter's (2005) model, methods for quantifying epigenomic modifications that alter gene expression have emerged. Assessing the impact of

environment on epigenetic signatures without gene mutations can be expressed as:

$$eG \times E = \Delta D$$

Since DNA methylation patterns can contribute to the development of complex disease states, epigenetic modifications appear to be a good fit for Hunter's model. The model is useful in guiding research studies analyzing DNA methylation patterns because the research can be initiated by identifying epigenetic and environmental components separately and then considering the interaction between them. When little is known about a disease process or risk factors for a disease are ambiguous, the investigator can initiate inquiry by looking for DNA methylation patterns that may be associated with the disease state. A limitation to this approach, as with any retrospective analysis, is that we cannot establish causation. However, effect size established from a small exploratory study can be used to estimate the sample size needed for a prospective study that incorporates environmental interactions.

Application of multiplicative gene-environment interaction models is useful in designing studies to determine environmental factors that modify epigenomic variants. DNA methylation, an epigenetic process that predominates during development and can be modulated throughout postnatal life, is influenced by factors such as nutrition, body weight and smoking status that are amenable to nursing interventions (Davis & Uthus, 2004; Kargul & Laurent, 2009). By modifying Hunter's (2005) model incorporating epigenome, instead of genotype, it is possible to identify epigenomic signatures, environmental factors, and interaction between the two that result in complex disease states.

The purpose of this study was to identify variants in maternal blood that are associated with maternal GBS colonization. In chapter two, what is known about GBS colonization is reviewed to justify the need for using a substantially different approach to identifying host factors that may be associated with GBS colonization. To date, epidemiological studies have failed to identify consistent maternal risk factors. The approach for this study will be to analyze epigenomic variants of affected individuals, specifically DNA methylation, and evaluate several maternal endogenous environmental variants IL-10, IL-6, TNF-α, and serum vitamin D (25[OH]D) (25[OH]D). Lastly, alterations in DNA methylation patterns will be evaluated to determine if they are involved in functional pathways associated with immune function. Identification of specific alterations in DNA methylation that result in altered immune function would suggest a mechanistic explanation as to why a third of pregnant women are colonized with GBS.

Research Questions

1. Are serum levels of TNF-α, IL-6, IL-10, and vitamin D (25[OH]D) different in pregnant women with GBS colonization than pregnant women without GBS colonization?
2. Are DNA methylation patterns different between pregnant women with GBS colonization and those without GBS colonization?
3. Is there a relationship between differentially methylated genes and immune function in pregnant women colonized with GBS?

Assumptions

The study will be conducted based on the following pre-stated assumptions:

1. Pregnant women presenting with GBS colonization did not have clinical symptoms indicative of infection such as fever, chorioamnionitis, bacturia or preterm labor at the time recto-vaginal swabs for culture were collected. Medical records indicate no other active infections at the time of screenings.
2. Altered levels of circulating TNF-α, IL-6, IL-10, and vitamin D (25[OH]D) are associated with infectious disease susceptibility (Berner et al., 2002; Fan et al., 2003; P Madureira et al., 2011; Maisey et al., 2008; Mikamo et al., 2004; Ng et al., 2003; Parameswaran & Patial, 2010; Puliti et al., 2002; Santhanam et al., 1991; Vieira et al., 1991).
3. Serum vitamin D (25[OH]D) levels are variable related to cytokine production (Hopkins et al., 2011; Shab-Bidar et al., 2012).
4. Environmental influences throughout life have the potential to induce variation in DNA methylation patterns, modulating gene expression that contributes to health and disease states (Baccarelli et al., 2010; Berger et al., 2009; Rodenhiser & Mann, 2006)
5. Differences in maternal DNA methylation patterns during early pregnancy among women with and without GBS colonization represent a biomarker for early risk identification.

CHAPTER II

REVIEW OF LITERATURE

Maternal group B streptococcus (GBS) colonization continues to be a global health burden among pregnant women and neonates despite current treatment and prevention strategies because neither adequately address the underlying cause of disease (Edmond et al., 2012; Verani et al., 2010). In this chapter, evidence supporting the modification of the gene-environment framework, described in chapter 1, to incorporate the inclusion of epigenomic-environment interactions will be described for GBS colonization as the intended research target. The characteristics of GBS and clinical significance of GBS colonization and infection will also be reviewed. GBS continues to be a pathogen of interest due to increasing resistance to antibiotics, continued neonatal infections despite antibiotic prophylaxis, and increasing incidence of invasive disease in non-pregnant adults. Current treatment and prevention protocols have been in place since the 1990s and the incidence of disease has plateaued in the neonatal population, but has risen in historically unaffected populations (Edmond et al., 2012; Phares et al., 2008; Schrag & Verani, 2013). Most recently, multiple scientific teams have begun to develop vaccines against GBS in attempts to mitigate the disease burden caused by GBS (Johri et al., 2006; Schrag & Verani, 2013). However, characteristics of the bacterium and ethical issues have hindered successful vaccine development. The intent of this study was to identify variants in maternal blood that are associated with maternal GBS colonization

that could help identify new intervention targets to prevent GBS colonization and subsequent infection.

Group B Streptococcus

GBS is a gram positive bacterium that causes invasive diseases, such as pneumonia, meningitis, chorioamnionitis, and sepsis, primarily affecting pregnant women, the elderly, and infants. GBS continues to cause significant morbidity and mortality, particularly in neonatal populations, despite current clinical practice designed to prevent the transmission of GBS from mother to neonate during delivery (Phares et al., 2008). Clinical recommendations described in Chapter 1 to circumvent GBS transmission to the neonate, have significantly decreased the incidence of early neonatal sepsis. However, there have been no further decrease in the amount of late or early onset sepsis and new approaches for prevention infection are need to further reduce poor outcomes associated with GBS infection. What is known about GBS infection and colonization will be reviewed, as well as novel new approaches to identify endogenous maternal factors that may be associated with risk for GBS colonization.

Characteristics of GBS bacterium

GBS are facultative anaerobic gram positive cocci that grow in pairs or chains. Initially, GBS was only associated with cattle as the cause of bovine mastitis. Lancefield (1933), first categorized streptococci into five groups based on cell wall carbohydrate antigens, and the groups also differ in laboratory identification techniques, colony morphology, and disease association. The groups are labeled A-E, GBS belongs to Lancefield group B, producing 1-3 mm diameter grayish-white flat mucoid colonies when grown in the laboratory. The colonies have a narrow zone of hemolysis with a

positive CAMP test. CAMP is an acronym for the scientist who developed the test (Christie, Atkinson, Munch, Peterson) for selectively identifying GBS. The additive they developed results in a larger area of clearance around the colonies because the additive increases the hemolytic activity of GBS (CDC, 2010). Group B specific antigen must also be detected when identifying GBS, usually done by latex agglutination. However, molecular identification via rapid polymerase chain reaction (PCR) and other methods is becoming more common (Bergseng et al., 2008; Kong, Gowan, Martin, James, & Gilbert, 2002).

GBS Pathogenesis

The key feature of GBS that allows it to evade the host immune response is the thick capsular polysaccharide layer that surrounds the bacterium. Antigenic differences in the layer allows for differentiation into one of 10 distinct serotypes (Ia, Ib, II-IX) (Lancefield & Freimer, 1966; Slotved, Kong, Lambertsen, Sauer, & Gilbert, 2007). However, some GBS isolates are of indeterminate serotype which is likely because the antigens for the serotype have not yet been identified (Ferrieri, Baker, Hillier, & Flores, 2004). Prevalence of serotypes varies by geographic location (Ippolito et al., 2010), although serotype III causes most cases of disease in infants (48.9%) with serotype Ia (22.9%), serotype Ib (7.0%), serotype II (6.2%), and serotype V (9.1%) accounting for the majority of other cases (K M Edmond et al., 2012). The capsular polysaccharide layer prevents the binding of compliment factor C3 to the surface of GBS, allowing evasion of the immune recognition (Doran & Nizet, 2004; Spellerberg, 2000).

GBS beta-hemolysin/cytolysin (β-h/c) is the second virulence factor that is involved in GBS pathogenicity. GBS β-h/c is a non-immunogenic pore-forming

membrane associated toxin capable of damaging multiple tissues and impacts disease severity (Nizet et al., 1996; Puliti et al., 2000; Ring et al., 2002). In animal models GBS β-h/c resulted in increased bacterial load, pro-inflammatory cytokines IL-6 and IL-1α, and mortality compared to non-hemolytic mutants (Puliti et al., 2000). However, when GBS β-h/c damages host cells it results in the release of IL-8 causing local inflammation and recruitment of neutrophils seen in GBS infections (Doran, Chang, Benoit, Eckmann, & Nizet, 2002). Additionally, macrophages exposed to GBS β-h/c have higher expression of nitric oxide synthase which generates 4 fold more nitric oxide than strains without β-h/c. High levels of nitric oxide are exhibited in septic shock caused by GBS β-h/c (Ring et al., 2002).

The surface protein C5a peptidase plays a key role in adhesion to host cells and is present on all strains and serotypes of GBS (Cheng et al., 2001; G. Y.-H. Liu & Nizet, 2004). C5a peptidase is encoded by the ScpB gene and enables binding to epithelial cells (Brown et al., 2005). However, if the ScpB gene is deleted it does not completely inhibit GBS from adhering to host cells, suggesting that other factors play a role in GBS adherence to host cells (Cheng, Stafslien, Purushothaman, & Cleary, 2002; Lindahl, Stalhammer-Carlemalm, & Areschoug, 2005; Tamura, Hull, Oberg, & Castner, 2006). The following additional surface proteins also contribute to GBS adherence to host cells to varying degrees: pili, α-C protein, Lmb, FbsA, and Rib. The surface proteins interact with fibronectin, fibrinogen, laninin, and integrins that attach to host cells (Doran & Nizet, 2004; Lindahl et al., 2005).

Maternal GBS Colonization

Rates of maternal GBS colonization vary extensively worldwide (Table 1). Rates of GBS colonization have been reported as low as 1.8% in Maputo, Mozambique (de Steenwinkel et al., 2008), up to 65% in non-pregnant women in the United States (Meyn, Krohn, & Hillier, 2009). The gastrointestinal tract is thought to be the primary reservoir associated with maternal GBS colonization; sexual activity and abnormal vaginal microbiota presumably contribute to the development of vaginal GBS colonization (Meyn et al., 2009). GBS also is likely sexually transmitted because sexual partners are frequently colonized with the same strain (Foxman et al., 2006; Manning et al., 2004; Meyn et al., 2009; Meyn, Moore, Hillier, & Krohn, 2002). Maternal factors that have previously been associated with colonization are: young maternal age, black race, and having low levels of GBS-specific anticapsular antibodies (Verani et al., 2010). However, a study by Kovavisarach et al. (2007) identified older maternal age as a risk factor and Zusman et al. (2006) found no association with race or maternal age. Therefore, risk factors associated with maternal colonization based on epidemiological studies appear to differ by geographic location and are ambiguous (Kovavisarach et al., 2007; Phares et al., 2008; Verani et al., 2010; Zusman et al., 2006). Further research is necessary to identify factors associated with maternal GBS colonization so targeted prevention methods can be developed.

GBS Disease in Pregnant Women

The incidence of invasive GBS disease in pregnant women is twofold higher than non-pregnant women (Deutscher et al., 2011). GBS can result in stillbirth, preterm birth, premature rupture of membranes, abortion, bacteremia, endometritis, chorioamnionitis,

Table 1. Global Maternal GBS Colonization Rates

Country	Colonization Rate (%)	Reference
Brazil	17.9	(Zusman et al., 2006)
Central African Republic	17.5	(Brochet, Couvé, Bercion, Sire, & Glaser, 2009)
Germany	16.0	(Brimil et al., 2006)
Korea	8.3	(Kim et al., 2011)
Lebanon	17.7	(Seoud et al., 2010)
Mozambique	1.8	(de Steenwinkel et al., 2008)
Netherlands	21.0	(Valkenburg-van den Berg et al., 2006)
New Zealand	22.0	(Grimwood et al., 2002)
Norway	34.8	(Hakon Bergseng, Bevanger, Rygg, & Bergh, 2007)
Senegal	20.0	(Brochet et al., 2009)
Switzerland	21.0	(Rausch, Gross, Droz, Bodmer, & Surbek, 2009)
Taiwan	6.2	(Yang et al., 2012)
Thailand	18.1	(Kovavisarach et al., 2007)
United Kingdom	21.3	(N. Jones, Oliver, Jones, Haines, & Crook, 2006)
United States	24.2	(Verani et al., 2010)
Uruguay	17.3	(Laufer et al., 2009)
Zimbabwe	31.6	(Moyo, Mudzori, Tswana, & Maeland,2000)

pneumonia, puerperal sepsis, endocarditis, and infections of the genital tract, placenta, and amniotic sac. There is currently no known way to prevent GBS colonization and

GBS screening is not conducted until late in pregnancy. Identification of maternal factors associated with risk for colonization may allow for the development of prevention and treatment strategies to prevent poor early pregnancy outcomes associated with GBS colonization. New prevention and treatment strategies are necessary because 70% of women who have poor pregnancy outcomes due to GBS will also endure poor fetal outcomes for their offspring (Phares et al., 2008; Verani et al., 2010).

Maternal GBS Screening and Treatment

The CDC currently recommends screening for GBS colonization 35-37 weeks into the pregnancy for all pregnant women. Women positive for GBS are treated with antibiotics after they go into labor, preferably at least 4 hours prior to delivering the infant. Penicillin G (5 million units) and ampicillin (2 grams) are first line antibiotics used to prevent neonatal sepsis in infants born to mothers with GBS colonization because they both reach minimum bacteriocidal concentrations in the amniotic fluid, maternal, and fetal circulations (Pacifici, 2006; Verani et al., 2010). The following medications are recommended for GBS prevention for women with severe penicillin allergy: cefazolin, clindamycin, erthyromycin, and vancomycin. However, drug levels of these antibiotics are lower in fetal serum than maternal serum or have variable transfer rates across the placenta (Pacifici, 2006; Philipson, Sabath, & Charles, 1973). Ampicillin administered intravenously exceeds the minimum bactericidal concentration to kill 99.9% of GBS within five minutes of intravenous administration in the maternal and fetal circulation (Bloom, Cox, Bawdon, & Gilstrap, 1996). Penicillin levels 179 times above the minimum concentration required to eliminate GBS have been collected in fetal serum (Barber, Zhao, Buhimschi, & Illuzzi, 2008), indicating intrapartum antibiotics result in significant

maternal and fetal exposure from the intrapartum antibiotics administered. Neonatal serum levels of ampicillin are higher than maternal concentrations after delivery and persist for at least 5.6 hours (Colombo, Lew, Pedersen, Johnson, & Fan-Havard, 2006). Further studies evaluating the persistence of antibiotics in the fetal circulation after delivery are sparse, but likely contribute to the decreased incidence in early onset neonatal GBS sepsis.

Maternal Vaccination: Currently, there are a number of clinical trials underway examining proposed GBS vaccines for immunization of women prior to, or during, pregnancy. The vaccines currently being tested target either GBS capsular carbohydrates or proteins (Heath, 2011; Johri et al., 2006). Immunization with the capsular carbohydrate alone proved not to be sufficiently immunogenic. However, when capsular carbohydrate is combined with tetanus toxoid conjugate vaccine sufficient antibodies are produced against GBS (Baker, Rench, & McInnes, 2003). Baker at al.'s (2003) study only added conjugated type III capsular carbohydrate to the tetanus toxoid and did not result in immunity to other GBS serotypes. Ongoing studies are investigating potential capsular carbohydrate vaccines that are multivalent to ensure broader coverage to prevent infection. Despite the success of generating some immunity to GBS, a number of concerns related to GBS vaccine development remain. For example, vaccines may not be effective globally since there is variation in risk factors in the literature. There are also significant ethical concerns related to testing vaccines on pregnant women that could harm to the fetus (Johri et al., 2006; Paradiso, 2001). Vaccine manufacturing companies directly contribute to a lag in vaccine development for pregnant women because of fears of liability if the exposed child develops health issues later in life (Kaposy & Lafferty,

2012). Liability fears related to administering vaccines during pregnancy may worsen in light of growing evidence illustrating altered epigenomic patterns resulting from chemical exposures occurring *in utero*. For example, permanent alteration in methylation patterns can occur in fetal DNA in response to chemical exposures *in utero,* such as DES described previously. Litigation has also resulted from residual DES effects on the grandchildren of women given the medication during pregnancy (Rothstein, Cai, & Marchant, 2009). In light of this, vaccines and systemic antibiotic treatment to prevent GBS transmission may not be the least harmful approach. Further research is needed to determine the long term impact of vaccine and antibiotic administration in utero and early in development.

GBS Disease in Neonates

GBS remains the leading cause of neonatal infectious morbidity and mortality, despite the administration of antibiotics colonized women to prevent vertical transmission (Clifford et al., 2011; Verani et al., 2010). The incidence of GBS disease in neonates is lower in developed countries, 0.4 - 0.81 per 100 live births, than in developing countries 0.91 - 1.81 per 100 live births (Table 2). Neonatal fatality rates are also disproportionately higher in some developing nations (Table 2), which could be due to variable prevalence of serotypes by geographic location (Johri et al., 2006) and different standards of medical care (Edmond et al., 2012; Heath, 2011).

Meningitis, bacteremia, and pneumonia are the neonatal clinical diagnoses most commonly caused by invasive GBS. In the US, the incidence of invasive neonatal GBS infection is higher in African American infants (Phares et al., 2008). A meta-analysis completed by Edmond et al. (2012) indicated that infants weighing less than 1.5 kg at

Table 2. Global Neonatal GBS Morbidity and Mortality

Country	Incidence/1000	Fatality Rate (%)	Reference
Denmark	0.4	8.0	(Ekelund & Konradsen, 2004)
Jamaica	0.91	3.6	(Trotman & Bell, 2006)
Malawi	1.81	33.0	(Gray, Bennett, French, Phiri, & Graham, 2007)
Netherlands	0.56	12.3	(Trijbels-Smeulders et al., 2007)
Norway	0.66	6.5 (1996-2005) 20.0 (2006)	(Bergseng et al., 2008)
United Kingdom and Ireland	0.72	9.7	(Heath et al., 2004)
United States	0.81-0.68	5.0-9.0	(Phares et al., 2008)

birth are 8 times more likely to develop invasive illness caused by GBS than normal weight infants. Additionally infants between 1.4 - 2.5 kg at birth are three times more likely to develop GBS infection. Neonatal GBS disease is classified by time of invasive disease onset after birth. Early onset GBS disease occurs during the first 7 days of life. Late onset GBS disease occur after the first week of life through the first 90 days after birth (Verani et al., 2010). In the US, the number of cases of early onset (1232) and late onset disease (1036) are nearly equal since the CDC guidelines for intrapartum antibiotics were implemented in the 1990s. However, mortality is higher for infants with early onset GBS disease (83 versus 48 deaths) (Phares et al., 2008).

Early Onset Neonatal GBS Disease

The incidence of early onset GBS disease in neonates has decreased from 1.7 cases per 1,000 live births in 1990 to 0.37 cases per 1,000 live births in 2008 in the US; as a result of widespread implementation of intrapartum antibiotic administration for mothers colonized with GBS. Maternal GBS colonization is the strongest predictor of early onset disease. Other risk factors for early onset neonatal disease are GBS in maternal urine at any point in pregnancy, rupture of membranes greater than 12 hours, delivery before 37 weeks gestation, young maternal age, African American race, infection, low maternal anticapsular antibodies to GBS, prior delivery with GBS, and maternal fever greater than 37.5˚C during labor (Verani et al., 2010). Women with heavy GBS colonization are more likely to infect their infants versus women who have a lower bacterial load (Regan et al., 1996; Yancey, Duff, Kubilis, Clark, & Frentzen, 1996). In other words, women with more GBS present have a higher likelihood of infecting their infants with GBS. A majority of infants that develop early onset disease are full term (77%) and 90% of infants become ill within 12 hours of birth (Phares et al., 2008). The most common presentation of early onset disease are pneumonia and sepsis, and less commonly meningitis (Verani et al., 2010).

Late Onset Neonatal GBS Disease

The incidence of late onset GBS disease in neonates is currently 0.35 cases per 1,000 live births (Jordan et al., 2008). Intrapartum antibiotic administration for GBS colonization has not had any significant effect on the incidence of late onset disease (Berardi et al., 2013; Phares et al., 2008). Late onset disease commonly presents as bacteremia and meningitis, and less commonly pneumonia or local site infections such as

cellulitis (Berardi et al., 2013; Jordan et al., 2008). African American infants are disproportionately affected (Jordan et al., 2008). Unlike the presentation and time of onset of early onset disease, late onset disease is different between term and preterm infants; suggesting that there may be different mechanisms involved. For example, Berardi et al (2013) identified term infants present with late onset disease earlier than preterm infants. Additionally, the etiology may be different from early onset disease because less than 30% of infants who develop late onset disease had mothers with positive GBS screenings. It is been speculated the source may be breast milk, persistent maternal colonization (for the 30% where maternal GBS colonization was present), or healthcare workers but additional studies are needed for validation. Further, Jordan et al. (2013) identified that 47% of infants that developed late onset disease had been exposed to intrapartum antibiotics for either GBS colonization, Cesarean section, or other complications. Late onset disease has not been studied with the same intensity as early onset disease. Now that the incidence of early and late onset neonatal disease is equivalent, the depth of research investigating factors associated with late onset neonatal GBS may improve.

GBS Infection in Non-pregnant Populations

The incidence of GBS infections in non-pregnant adults doubled from 3.6 per 100,000 people in 1990 to 7.3 per 100,000 people in 2007 (Skoff et al., 2009). Traditionally, GBS disease was seen in individuals with compromised immune systems due to advanced age or other underlying conditions. Similar to neonatal GBS disease, bacteremia is a common outcome of invasive GBS in adults followed by skin and/or soft tissue infections, and pneumonia. Since the implementation of intrapartum antibiotics for

GBS colonized mothers, half of case fatalities are adults older than 65 years old (13.2% of those infected perish) and are approaching pneumonia fatality rates (20.6%) for the elderly population (Edwards, Rench, Palazzi, & Baker, 2005; Schrag et al., 2000). Colonization rates are similar to those seen in pregnant women (21.7%) and almost half of the elderly (47.3%) affected are colonized with serotype V (Edwards et al., 2005), whereas over 60% of disease in neonates is caused by serotype III and Ia (K M Edmond et al., 2012). Since this incidence is rising in this population, identification of factors contributing to maternal GBS colonization may also be informative for reducing the incidence in the elderly population as well.

GBS Resistance to Antibiotics

As previously described, infants born to mothers adequately treated with intrapartum antibiotics still develop GBS disease and a large proportion of infants are born to mother with negative GBS screening. Additionally, GBS positive women treated with intrapartum antibiotics are more than 4 times likely to be positive for GBS 6 weeks after delivery than women not given intrapartum antibiotics (Manning et al., 2008). Interestingly, 65% of women in Manning et al.'s (2008) study continue to have GBS colonization despite antibiotic treatment and 18.3% of women that retained GBS positive status were colonized with a different strain of GBS. Perhaps the persistence of GBS colonization in mothers despite antibiotic treatment is a contributing factor in the unchanged incidence in late onset neonatal GBS disease since the implementation of intrapartum antibiotic protocols. Conversely, GBS bacteria may just be resistant to the antibiotic treatment. In recent years, there has been a push in implementing antibiotic stewardship programs in hospital settings to decrease the rates of antibiotic resistance.

The programs should also incorporate GBS prophylaxis recommendations and providers should be aware of the resistance rates in their area. Many studies evaluating antibiotic resistance of GBS assert the strains remain susceptible to penicillin and amoxicillin (Castor et al., 2008; Chohan, Hollier, Bishop, & Kilpatrick, 2006; Garland et al., 2011). However, reduced susceptibility to penicillin by GBS has been identified in Hong Kong and warrants further monitoring for increasing resistance (Chu et al., 2007). Antibiotic sensitivity testing is particularly warranted for women with penicillin allergies, since clindamycin and erythromycin resistance is high (Table 3). Notably in the US, strains that are clindamycin or erythromycin resistance are likely to have dual resistance to both erythromycin (94.3%) and clindamycin (71.5%), respectively. In light of increasing

Table 3. Global GBS Resistance to Antibiotics

Country	E	C	EC	Reference
	Percent (%) Resistant			
Australia	17	22	38	(Garland et al., 2011)
Korea	9.7	6.8	-	(Uh et al., 2007)
Malawi	21	-	-	(Gray et al., 2007)
New Zealand	-	15.4	7.7	(Grimwood et al., 2002)
Norway*	11.9	10.9	25.4	(Bergseng et al., 2008)
Taiwan	44.0	39.0	-	(Janapatla et al., 2008)
United States	54.0	33.0	-	(DiPersio & DiPersio, 2006)
United States (New York)	50.7	38.4	94.3/71.5	(Back et al., 2011)

E = Erythromycin
C = Clindamycin
EC = Both Clindamycin and Erythromycin
*= E & C reported for adult cases EC reported for neonatal disease

antibiotic resistance, identification of definitive environmental and/or genetic maternal risk factors associated with GBS colonization could result in alternate clinical approaches for preventing neonatal GBS sepsis.

Epigenome-Environment Interaction and GBS Colonization

Epidemiologic findings suggest an association between environmental factors and epigenetic alterations, serving as the basis for many complex diseases including obesity, type 2 diabetes, asthma, autism spectrum disorder, and attention deficit hyperactivity disorder (Latham, Sapienza, & Engel, 2012). Environmental influences that range from behavioral conditions, dietary factors, and toxic exposures can be modulated to achieve optimal health outcomes for the populations. Additionally, infections and events occurring within a person can alter epigenetic signatures (Tolg et al., 2011). Because of this, measuring endogenous factors that may contribute to disease development are also valuable in epigenomic-environment interactions studies. Further, if modifiable exogenous environmental factors can be linked to an altered endogenous environment or epigenome, implications for treatment and prevention are ascertained.

Epidemiologists have identified associations between environmental exposures and disease for decades linking poor hand hygiene of healthcare workers with patient sepsis in the 19th century (Gould, 2010), poor nutrition during pregnancy to cardiovascular disease later in life in offspring (Barker & Osmond, 1986; Barker, 1995), and smoking to lung cancer (Doll & Hill, 1950). However, epidemiological studies traditionally do not identify biologic mechanisms that cause disease and are limited to identifying the association between population-level risk factors and disease. Determining the role of DNA methylation in the origin of disease could help identify

epigenetic linkages, not explained by a change in sequence or genetic alteration, that have been elusive for most complex diseases. Factors identified in epidemiologic studies combined with DNA methylation analyses have the potential to identify epigenomic-environment interactions at critical time points across the lifespan that contribute to disease phenotype and inheritance. Analysis of DNA methylation patterns alone will not explain why or how phenotypes are altered; nor will they identify effective primary, secondary or tertiary intervention strategies. Environmental factors (e.g., diet, lifestyle, stress) must be measured to identify interactions that may cause alterations in DNA methylation. Clinical assessments completed by nurses, particularly of lifestyle factors that alleviate or exacerbate symptoms, can help identify modifiable factors that are related to altered DNA methylation patterns. Each person's DNA and environmental exposures are unique and assessing both simultaneously will result in more personalized healthcare. For example, clinical assessments first recognized the association between diethylstilbestrol (DES) administration to prevent miscarriage and development of reproductive tract anomalies in offspring. *In vitro* and in adults, DES did not cause alteration in methylation patterns. However, DES causes hypermethylation of a gene responsible for reproductive tract development to organisms exposed *in utero* that persisted into adulthood (Bromer, Wu, Zhou, & Taylor, 2009). Therefore, it is necessary to perform detailed assessments of patient, and family, medical history to identify how exposures may alter gene expression because the associations are not always obvious.

To illustrate the utility of evaluating epigenomic changes in the context of gene-environment interaction models as described in Chapter 1, the concept of DNA methylation and how perturbations in DNA methylation patterns can alter gene

expression and disease susceptibility will be reviewed. DNA methylation is the best understood epigenetic mechanism that modulates gene expression (Baccarelli, Rienstra, & Benjamin, 2010). Differential methylation induced by endogenous or exogenous influences can lead to both genome instability and inappropriate gene transcription, contributing to pathology. DNA methylation patterns are often specific to lineage, organ, and cell-type (Cedar & Bergman, 2012). For example, all the human cells in an individual contain the exact same DNA. The epigenomic signature of a cell will program it to differentiate into a heart, eye, or skin cell. Abnormal loss or gain of methylation at key DNA sites may result in inappropriate expression of a gene. When specific patterns of methylation are associated with a phenotype, such as risk for disease, the patterns can be used to identify those at risk for poor health outcomes and provide the basis for new treatments. The investigation of epigenetic markers to identify biological mechanisms of complex disease processes, such as atherosclerosis(Guay et al., 2013; Lund et al., 2004; Zaina, Lindholm, & Lund, 2005), schizophrenia(Auta et al., 2013; Costa et al., 2002; Sharma, 2005), and lupus(Absher et al., 2013; Li, Gorelik, Strickland, & Richardson, 2014; Sekigawa et al., 2003), have been increasing in recent years (Bergman & Cedar, 2013; Petronis, 2010; Rakyan, Down, Balding, & Beck, 2011; Rodenhiser & Mann, 2006). Furthermore, DNA methylation is heritable during cellular reproduction, and likely from one generation to another. This means methylation signature can pass from both cell to cell and transfer from parent to offspring (Guerrero-Bosagna & Skinner, 2012; Rodenhiser & Mann, 2006). Therefore, DNA methylation patterns passed across generations may provide an explanation for the transmission of complex disease susceptibility among families that is modulated by environmental exposures (Figure 2).

Figure 2 illustrates how alterations in DNA methylation by exogenous and endogenous factors can contribute to level of susceptibility to complex disease in individuals, with transmission of the pattern throughout future generations (Wright, Ralph, Ohm, & Anderson, 2013).

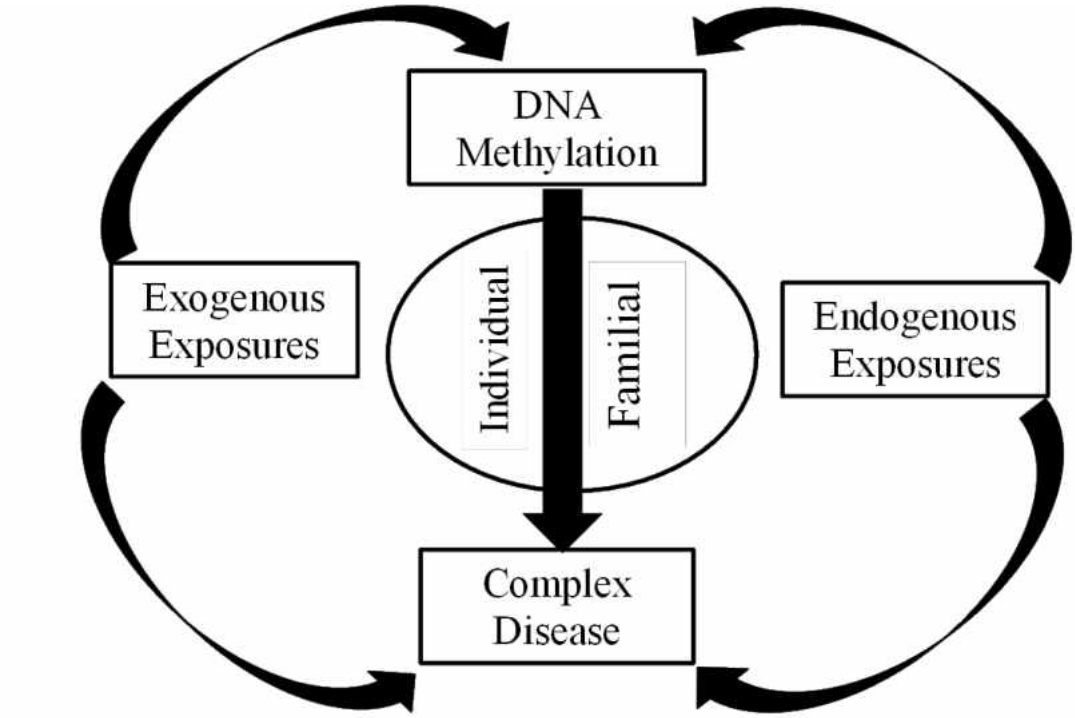


Figure 2. Influence of DNA Methylation in Complex Disease States.

DNA methylation patterns associated with GBS colonization could not be identified in the literature. However, altered DNA methylation patterns are known modulators of immune function and alterations have been associated with other infectious disease processes (Table 4). Notably, T-lymphocyte cell function and cytokine expression are altered by methylation patterns present on T cells (Fitzpatrick, Shirley, & Kelso, 1999). Additionally, DNA methylation patterns are altered by bacterial infection (Mikovits et al., 1998; Tolg et al., 2011) and different methylation patterns associated

with the development of disease have been identified in animal models when they are colonized with altered bacterial populations after birth (Olszak et al., 2012). Since DNA methylation vitally contributes to cell differentiation in the immune system and programming memory in immune cells, it is an excellent candidate for identifying unknown mechanisms that may be associated with infectious diseases susceptibilities.

Table 4. DNA Methylation and Immune Function

Immune Function or Alteration	Reference
Bacterial infection induces hypermethylation	(Tolg et al., 2011)
Discriminates between regulatory and conventional T cells	(Baron et al., 2007)
Downregulation of IFN-γ in fetus, helps prevent fetal loss	(White, Watt, Holt, & Holt, 2002)
IgE production	(Liu, Ballaney, Al-alem, & Quan, 2008)
Maintenance of T cell memory and cytokine expression pattern	(Fitzpatrick et al., 1999)
Maternal bacterial infection promotes fetal hypermethylation	(Bobetsis et al., 2007)
Number and function of regulatory T cells	(Schaub et al., 2009)
Viral infection increases DNA methylation	(Mikovits et al., 1998)

DNA Methylation

DNA methylation is an epigenetic modification that may result in gene silencing, gene activation resulting in chromosome instability, inappropriate gene expression, and inability to carry the epigenomic signature to future cell lines (Bergman & Cedar, 2013; Cedar & Bergman, 2012; Jones, 2012; Rodenhiser & Mann, 2006). Epigenetic changes are heritable alterations in the chromosome that do not change the DNA sequence that result in a specific phenotype, which are observable characteristics (Berger et al., 2009;

He, Chen, & Zhu, 2011; Tost, 2010). DNA methylation is a specific epigenetic alteration in which a methyl group attaches to a cytosine (C) residue in DNA that is followed by a guanine (G) residue connected by a phosphate bond, commonly referred to as a CpG dinucleotide, and is currently the most well understood epigenetic mechanism (Baccarelli et al., 2010; Cedar & Bergman, 2012; Chen & Riggs, 2011; P. A. Jones, 2012). Areas with dense concentrations of CpG dinucleotides are located in promoter regions of genes, which is where transcription factors bind to initiate the reading of a DNA sequence so a gene will be expressed. Methylation of cytosines located in the promoter region can alter gene expression by blocking transcription of the DNA, resulting in gene silencing. Conversely, a loss of methylation at these sites may result in inappropriate expression of a gene.

Between 60-90 percent of cytosines are methylated in human DNA (Ehrlich et al., 1982) and the establishment of "normal" DNA methylation patterns are necessary for embryonic development. Normal patterns of methylation are required for the differentiation of cell types. For example, every cell in an individual's body has the same DNA sequence and methylation patterns present on the DNA sequence is specific to tissue type(Cedar & Bergman, 2012; Jones, 2012; Laird, 2010). In other words, the methylation pattern for a cell in the heart will be different from the methylation pattern on a cell in the eye, even though the DNA sequence is exactly the same. Alterations in methylation patterns also explain some of the processes that occur in complex disease states, such as delayed onset disease or situations where only one identical twin develops cancer (Boks et al., 2009; Fraga et al., 2005; Kaminsky et al., 2009; Petronis, 2001). Furthermore, there is evidence that DNA methylation patterns are altered by

environmental exposures (Table 5), implying that we have the ability to intervene to prevent disease or promote desired health outcomes. Patterns of methylation associated

Table 5. Environmental Exposures that Alter DNA Methylation

Environmental factor	Reference
Alcohol	(Choi et al., 1999)
Bisphenol-A (BPA)	(Bromer, Zhou, Taylor, Doherty, & Taylor, 2010)
Diethylstilbestrol (DES)	(Bromer, Wu, Zhou, & Taylor, 2009)
Exercise	(Barrès et al., 2012)
Fear	(Miller & Sweatt, 2007)
Hydralazine	(Cornacchia et al., 1988)
Maternal Care	(Weaver et al., 2004)
Maternal Diet	(Wolff, Kodell, Moore, & Cooney, 1998)
Microbiome	(Olszak et al., 2012)
Procainamide	(Cornacchia et al., 1988b)
Smoking	(Toyooka et al., 2003)
Traffic pollution	(Baccarelli et al., 2009)

with a specific phenotype, such as susceptibility to infection, have potential for use in identifying people at risk for developing conditions and as treatment targets. Leukemia treatments based on epigenetic markers have been approved and used successfully in the clinical setting (Rodriguez-Paredes & Esteller, 2011). Identifying differences in DNA methylation patterns during early pregnancy in women with GBS colonization could represent a biomarker for early risk identification or develop methods to prevent colonization. Furthermore, because there is a strong relationship between DNA

methylation patterns and the function of cells in the immune system (Table 4); differential DNA methylation in genes associated with immune function could offer mechanistic insight as to why certain bacteria, like GBS, colonize some individuals and not others.

Environmentally Induced Modification

Environmental influences throughout life have the potential to induce variation in DNA methylation patterns, modulating gene expression that contributes to health and disease states (Baccarelli et al., 2010; Berger et al., 2009; Rodenhiser & Mann, 2006). For example, permanent alteration in methylation patterns can occur in fetal DNA in response to chemical exposures *in utero*. Diethylstilbestrol (DES) is a synthetic estrogen that was administered to pregnant women to prevent spontaneous abortions prior to the mid 1970's. DES causes hypermethylation of homeobox protein Hox-A10 (HOXA10), a gene that controls uterine organ development, resulting in reproductive tract anomalies that persist into adulthood (Bromer et al., 2009). Furthermore Bromer et al., determined the hypermethylation of HOXA10 was specific to the fetus and did not occur laboratory experiments using cell line or the pregnant women who received DES.

Assessment of DNA methylation patterns in disease states where the mechanism that alters gene expression are unknown can help identify etiology of disease. For example, in 1915, Kendall determined the microbes that colonize the gastrointestinal tract at birth are involved in normal development of the immune system. However, the biological mechanism of how this occurred remained unknown for years. Olszak et al. (2012) recently published a study suggesting that the type of bacteria that colonize the gastrointestinal tract in the neonatal period has an effect on the function of cells in the

immune system. This study was performed on germ-free and specific-pathogen free mice. Further, Olszak's study suggests that microbial exposure alters gene expression in specific tissues. The authors of the study noted hypermethylation of CpG sites in colonand lung tissues of the chemokine (C-X-C motif) ligand 16 (CXCL16) gene occurred when specific pathogens were not present during development. The CXCL16 gene encodes a chemokine receptor on invariant natural killer T cells (iNKT), resulting in higher accumulations of iNKT cells that are involved with inflammatory processes. Additionally, the higher accumulation of the iNKTs only occurred in the colon and lung when specific bacteria were not present. The authors hypothesize an environmental exposure later in life triggers various inflammatory disease processes programmed by the methylation changes in the bowel and lungs, like asthma and irritable bowel syndrome. Therefore, exposure to bacteria early in development affects the programming of the immune system, in mice, by causing perturbation in DNA methylation patterns. They concluded the findings could be extrapolated to humans because the mouse model used is similar to human cells. Further studies investigating alterations of methylation patterns in humans may be needed to demonstrate the effects of environmental exposures on immune function. This study aims to identify differential DNA methylation patterns in pregnant women colonized with GBS. If differential patterns are identified, future investigations will be focused on identifying causes of the altered methylation patterns.

Endogenous Maternal Environment

Other clinical indicators could potentially be used to identify pregnant women at increased risk for GBS colonization and have not been discussed in published literature. Clinical indicators, such as serum cytokine levels that reflect immune system functioning,

could also be associated with altered DNA methylation patterns. Variability in serum markers associated with immune function and vitamin D (25[OH]D) status have previously been identified and utilized as prognostic indicators of disease states, such as respiratory infections (Chesney, 2010), human immunodeficiency virus infections (Fahey et al., 1990), pancreatitis (Pezzilli et al., 1995), and depression (Kiecolt-Glaser & Glaser, 2002). There has been a plethora of research investigating the role of vitamin D (25[OH]D) as an immune function modulator in recent years (Figure 3) and could offer a cost effective intervention target if low serum vitamin D (25[OH]D) levels are associated with GBS colonization . In this section, a brief review of immune function during pregnancy, relevant cytokines, and vitamin D (25[OH]D) will be discussed.

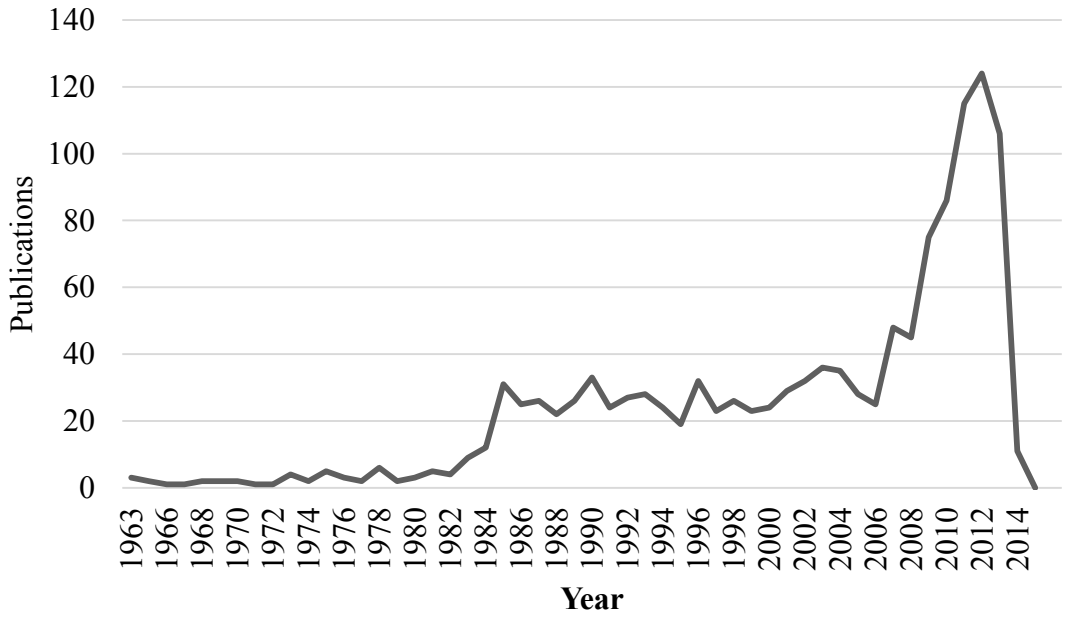


Figure 3. PubMed results for “Vitamin D” and “Immune Function”.

Immune System During Pregnancy

Multiple alterations in immune function are necessary during pregnancy to prevent the mother’s body from recognizing the developing fetus as a foreign pathogen.

Serum markers of immune function are increasingly being evaluated in pregnancy because alterations occur throughout normal pregnancy and inappropriate levels contribute to the development of pathology during pregnancy (Ponsonby, Lucas, Lewis, & Halliday, 2010). Alterations in the maternal immune system vary during pregnancy and can be effected by the overall health of the mother. For example, if a woman has a preexisting autoimmune disorder, like rheumatoid arthritis, she may experience remission from symptoms during the pregnancy as a result of altered immune functioning that protects the fetus. However, pregnancy does not result in the same altered state throughout the entire pregnancy because there are three distinct phases. First, in the early stages of pregnancy the environment is that of an invasion (Ashkar, Di Santo, & Croy, 2000; Dekel, Gnainsky, Granot, & Mor, 2010; Shimada et al., 2006). Implantation occurs, the placenta develops and an inflammatory environment allows the establishment of these entities in the maternal system. The initial pro-inflammatory stage enables vasculature to develop, removes the cellular byproducts of implantation and results in the clinical manifestation of "morning sickness" that is seen in the early stages of pregnancy(Mor & Cardenas, 2010). In the second phase, the initial inflammatory state resolves and an anti-inflammatory state begins to predominate; creating a safe environment for fetal growth. In the final phase, a pro-inflammatory environment is induced in order to deliver the fetus (Romero et al., 2006). Increased inflammatory markers contribute to the promotion of the rupture of membranes, uterine contractions, and delivery (Mor & Cardenas, 2010).

Cytokines

Cytokines are proteins produced by cells that are generally classified as pro-inflammatory or anti-inflammatory (Denney et al., 2011). There is redundancy of function between cytokines. Multiple cytokines illicit the same action, and have multiple target cells, so it can be difficult to attribute an action to a specific cytokine (Miyajima, Hara, & Kitamura, 1992). Cytokines are secreted by, and activate, various cells involved in the host immune response including activated phagocytes, epithelial cells, and T cells (Abbas, Lichtman, & Pillai, 2012). In general, cytokines produced by T-helper 1 (*Th1*) cells are pro-inflammatory and cytokines produced by T-helper 2 (*Th2*) cells are anti-inflammatory and some cytokines exhibit properties of both (Brogin Moreli, Cirino Ruocco, Vernini, Rudge, & Calderon, 2012). An imbalance of pro versus anti-inflammatory cytokines produces inflammation or muted immune responses. This can be attributed to the magnification of the normal synergistic or antagonistic effects that cytokines exhibit (Abbas, Lichtman, & Pillai, 2012). During normal pregnancy there is a shift to upregulate expression of *Th2* cells and suppression of *Th1* cells to prevent abortion of the fetus during development (Thellin & Heinen, 2003). Disruption of this altered balance of the immune system during pregnancy has been associated with gestational diabetes, preeclampsia, preterm labor, abortion and infection (Brogin Moreli et al., 2012; Fichorova et al., 2011). However, studies investigating levels of cytokines throughout pregnancy have presented dissimilar cytokine values during normal pregnancy (Curry et al., 2008; Makhseed et al., 2000; Vassiliadis, Ranella, Papadimitriou, Makrygiannakis, & Athanassakis, 1998). Variations in results could be explained by

different methodologies used to measure the cytokines and disparate sample sizes (30 versus 1200).

In a recent study, Fichorova et al. (2011) identified patterns in immune function markers that were specific to the type of bacteria present in the vaginal mucosa and placenta of pregnant women. They found that TNF-α, IL-8 and ICAM-1 were elevated in the presence of *Gardnerella*, which is the most common causative pathogen of bacterial vaginosis. When multiple organisms associated with bacterial vaginosis were present, pro-inflammatory cytokine (TNF-α, IL-1β, IL-6), chemokine (IL-8), and acute phase marker (CRP and serum amyloid A) levels were elevated. Furthermore, *Lactobacillus*, which colonize the vaginal mucosa and are not pathogenic, suppress pathogenic strains and downregulate pro-inflammatory cytokines (Donato, Gareau, Wang, & Sherman, 2010; Othman, Neilson, & Alfirevic, 2007; Zeuthen et al., 2010). However, patterns of immune function serum markers were not analyzed related to GBS colonization or infection in any identified studies. For this study, TNF-α, IL-6, IL-10 and vitamin D (25[OH]D) status were selected specifically for analysis because of their identified association with infectious diseases and action during pregnancy and because serum levels for these specific cytokines were available to for secondary analysis for this study.

TNF-α: TNF-α is a serum immune function marker of interest because it is involved with coordination of the cytokine cascade and regulation of macrophage biology, which are both needed to fight infection. Alteration in TNF-α serum levels contribute to the development of various disease states, including sepsis and autoimmune conditions (Parameswaran & Patial, 2010). Macrophages are the primary producers of TNF-α in non-pregnant populations. However, in pregnant women the placenta

contributes to increased TNF-α levels throughout pregnancy (Brogin Moreli et al., 2012). The elevated TNF-α levels increase insulin resistance, which could contribute to the development of hyperglycemia or gestational diabetes during pregnancy (Kirwan et al., 2002). There have been no human studies analyzing TNF-α in pregnant women related to GBS infection or colonization. However, in vitro experiments show an increase in TNF-α production in cells exposed to GBS (Berner et al., 2002; Mikamo et al., 2004). Additionally, TNF-α and IL-6 levels in mice increase systemically when inoculated with GBS when IL-10 production is decreased resulting in 60% mortality (Puliti et al., 2000). Because TNF-α levels increase in laboratory and animal studies as a result from GBS exposure, serum TNF-α levels may be increased in pregnant women colonized with GBS. Furthermore, due to the inverse relationship TNF-α and IL-6 have with IL-10 in response to GBS exposure in animal models, IL-6 and IL-10 will also be evaluated in this study.

IL-6: IL-6 is a serum immune function marker that is a pro-inflammatory cytokine involved in the acute phase in the immune response. IL-6 is elevated in amniotic fluid of pregnant women with premature rupture of membranes due to intrauterine infections and also increases during active labor (Santhanam et al., 1991). Elevated serum IL-6 levels have also been identified in other inflammatory conditions experienced during pregnancy, like preeclampsia (A. Sharma, Satyam, & Sharma, 2007). It is unclear if a similar elevation in IL-6 occurs in women with GBS colonization. However, IL-6 production does increases in vitro and in animal studies with exposure to GBS (Berner et al., 2002; Mikamo et al., 2004; Puliti et al., 2002). The increase in IL-6 in response to GBS exposure has not been verified or validated in human studies. Therefore, in this

study we will evaluate the level of serum IL-6 to determine if a similar increase in IL-6 production occurs in pregnant women in response to GBS colonization.

IL-10: IL-10 is a serum immune function marker that inhibits the synthesis of pro-inflammatory cytokines, such as IL-6 and TNF-α, and stimulates the production of B cells and their differentiation into antibodies (Vieira et al., 1991). Additionally, IL-10 levels vary during pregnancy and may be involved in the maintenance of a viable pregnancy (Denney et al., 2011; Hashii et al., 1998). Early in pregnancy IL-10 is protective because it inhibits secretion of inflammatory IL-6, TNFα, and INF-γ allowing the fetus and placenta to develop without being rejected by the maternal system. As the pregnancy progresses, the level of IL-10 decreases and the resulting increase in inflammatory cytokines allows the initiation of labor (Brogin Moreli et al., 2012).

Reduced IL-10 levels are associated with fetal loss in the first trimester, preeclamsia, gestational diabetes, and preterm birth (Brogin Moreli et al., 2012). Serum levels of IL-10, as related to GBS colonization in pregnant women, have not been evaluated. Although, Madureira et al. (2011) conducted a study using a murine model and found reduced levels of IL-10 in animals that carry GBS antibodies, which conferred immunity to the bacteria in offspring. Conversely, Bebien et al. (2012) found that the βh/c component of GBS induced IL-10 production. The higher production of IL-10 inhibited IL-12 production, which is involved in inducing immune responses, resulting in GBS being able to escape host cell detection and survive. When Bebien administered recombinant IL-10 to the GBS infected mice to test the effect of IL-10, the number of GBS in the mice increased significantly. If results from the animal studies described can be translated to humans, it is expected that there will be some elevation in serum IL-10

levels later in pregnancy in women colonized with GBS. Elevated IL-10 levels could be the result of GBS presence, like Madureira's study, or because the normal IL-10 elevation in early pregnancy enables GBS colonization in exposed women, like Bebien's study. Regardless of the mechanism, elevated IL-10 could be a clinical laboratory indicator for GBS colonization in pregnant women during the third trimester.

Vitamin D (25[OH]D) Status

Circulating vitamin D (25[OH]D) deficiencies have been associated with susceptibility to infectious diseases such as influenza, tuberculosis, and pneumonia (Chesney, 2010). Vitamin D3 is the form of vitamin D that is produced by the skin in response to sunlight and primarily the form of vitamin D consumed from foods containing vitamin D. Vitamin D3 is then converted to Vitamin D (25[OH]D) in the liver. Vitamin D (25[OH]D) is the circulating form of the vitamin that is traditionally measured to identify vitamin D status in clinical populations (Hollis, 2005, 2008, 2012). It is generally accepted that serum vitamin D (25[OH]D) levels must be above 20 ng/ml to maintain normal physiologic processes and fetal development during pregnancy. Recently, vitamin D (25[OH]D) experts suggest that levels greater than 32 ng/ml are necessary to support all physiologic processes that require vitamin D (25[OH]D) for optimal functioning, like preventing infections (ACOG, 2011; Holick, 2011; Hollis, 2012). Therefore, the normal increase in serum vitamin D (25[OH]D) levels during pregnancy may be involved in preventing infection and colonization with pathogenic bacteria during normal pregnancy. There is a possibility that colonized women may have lower serum vitamin D (25[OH]D) levels, predisposing them to group B streptococcus colonization. To date, no randomized control trials evaluating the effects of vitamin D

(25[OH]D) supplementation during pregnancy related to maternal colonization or infection have been completed (De-Regil Luz, Palacios, Ansary, Kulier, & Peña-Rosas Juan, 2012). However, a meta- analysis conducted by Thorne-Lyman and Fawzi (2012) suggests that it is unknown how vitamin D (25[OH]D) relates to maternal infections since the relationship between vitamin D (25[OH]D) and immunity has only recently been established. There is currently no documented relationship between serum vitamin D (25[OH]D) levels and maternal GBS colonization.

Implications for the Nursing Discipline

Person, health, environment, and nursing are the metaparadigm concepts that remain the pillars of the nursing discipline. Perception of these concepts constantly evolves to incorporate new knowledge gained through practice, research, education, and exposure to other disciplines. Since most human disease processes are multifactorial in nature and nurses interact with individuals throughout the illness-wellness continuum, it is imperative nurses understand how gene-environment interactions impact health. Research investigating exogenous and endogenous, like DNA methylation and serum cytokines and how they respond to environmental exposures, continues to generate data that improves our understanding of factors associated with complex disease processes. As the state of the science continues to evolve, nurses must begin to incorporate new data into their own research to deliver the best possible care to patients.

Investigations designed to identify biologic mechanisms explaining how epidemiologically defined risk factors result in complex disease processes lag behind data generated by epidemiologic studies (Dempfle et al., 2008; Hunter, 2005; Khoury, Davis, Gwinn, Lindegren, & Yoon, 2005; Martino & Prescott, 2011). Nurses, as members of

interdisciplinary teams, can utilize advances in epigenomic techniques to better assess levels of health and disease risk that may help explain findings identified in epidemiological studies. Nurses are well suited to investigate factors that may contribute to epigenomic variation because by nature of the profession, nurses continuously bridge science and technology to patient populations (Clark, Adamian, & Taylor, 2013; Loescher & Merkle, 2005). This study aimed to identify DNA methylation patterns and serum immune system markers associated with maternal GBS colonization. If vitamin D (25[OH]D) plays a role in the modulation of serum immune system markers, it would be a cost effective clinical intervention that could be introduced into practice to reduce GBS colonization rates and an alternative to antibiotic treatment. This study exemplifies how translational research can be initiated by hypotheses from the bedside, examined at the bench and brought back to the bedside to improve health outcomes.

CHAPTER III

METHODS

Design

An exploratory secondary data analysis was completed using acquired quantitative data and maternal peripheral blood samples that were previously collected in a prospective longitudinal cohort study of nulliparous pregnant women. The data was initially collected to evaluate differences in women with and without preeclampsia. All women who had preeclampsia were excluded from this analysis because women with preeclampsia have different methylation patterns than women without preeclampsia (Anderson, Ralph, Wright, Linggi, & Ohm, 2013). The participants were enrolled in the primary study during the first trimester of their pregnancy and were followed through the time of childbirth. Eligibility criteria for the primary study included; English speaking, no previous births after 20 weeks gestation, age >18 years, and singleton pregnancy. At enrollment between 10-14 weeks gestation, baseline demographic information and venipuncture to collect blood for genome-wide DNA methylation analysis and other serum blood tests including TNF-α, IL-6, IL-10 and circulating vitamin D (25[OH]D were completed. In subsequent trimesters (second trimester 22-26 weeks gestation and third trimester 32-36 weeks gestation), venipuncture to collect blood for serum blood tests including TNF-α, IL-6, IL-10, and circulating vitamin D (25[OH]D) was completed. A subgroup of 6 women (n=2 GBS positive; n=4 GBS negative) were selected as control

samples for genome-wide DNA methylation analysis in the primary study. The methylation analysis was completed on peripheral blood samples obtained in the first trimester using the Infinium bead-based array platform (Illumina, Inc., San Diego, CA). For this study, only women (n=42) with documented uncomplicated (no preeclampsia, gestational diabetes, chorioaminoitis, etc) pregnancies in the primary study were included in the cytokine and vitamin D analysis. A subset of 18 women (n=9/group) were evaluated for differential genome wide methylation differences.

Procedures

Collection of Physiologic Data

Prior to commencement of data collection, this study was approved by the University of North Dakota Institutional Review Board (Appendix A). The primary study was approved by the University of North Dakota Institutional Review Board (Appendix B) and Altru Health System (Appendix C) prior to enrollment of participants into the study. Informed consent was completed by participants after verbal and written descriptions of the parent study were given. Participants were informed that blood samples would be taken and laboratory studies would be conducted on DNA extracted from these samples. Serum collected for cytokine and vitamin D analysis was collected via venipuncture. Whole blood was collected in a red top vacutainer (B-D) blood tube, retained at room temperature and allowed to clot for 1 hour followed by centrifugation at 1,000 x g at 4 degrees centigrade for 10minutes. Serum (500 µl) was placed in separate vials and frozen at -80 degrees centigrade until they were analyzed for the parent study. Cytokine laboratory analysis were completed at the United States Department of Agriculture, Agricultural Research Service (USDA ARS) Grand Forks Human Nutrition

Center and vitamin D laboratory analysis was completed in the Brown-Borg laboratory at the University of North Dakota, School of Medicine and Health Sciences for the parent study. Only de-identified data was utilized in the current study. Medical history, GBS colonization status, and information about the participants had previously been extracted from the medical record for the purpose of the primary study using a standard data abstraction form (Appendix D). Descriptive statistics were used to elicit information about the study population including: age, race/ethnicity, ethnicity, weight, sex of the infant, gestational age at birth, and co-morbidities.

Cytokine Laboratory Analysis

Cytokine analysis that was previously completed on serum samples for TNF-α, IL-6, and IL-10 (Bio-Plex, Millipore, Fountain Hills, AZ) at each of the three pregnancy trimesters. In brief, the target protein antibody was coupled to dual beads and incubated with sample. The protein of interest was captured, combined with a biotinylated antibody for a different epitope, and detected using a dual-laser flow based reader. For the cytokine analyses, the serum was diluted one volume of sample to three volumes of Bio-Plex human serum sample diluent. Next, 50 µl of assay diluent was added to each well in the 96 well plate. Then, 200 µl of cytokine standard, control, or sample was added to each plate and allowed to incubate at room temperature for two hours. All fluid from the wells was aspirated and each well washed a total of four times. 200 µl of conjugate was then added to each well and allowed to incubate a room temperature for two hours. After aspirating all fluid from the wells and washing the wells four times, 200 µl of substrate solution was added to all the wells, covered with foil and incubated at room temperature

for 20 minutes. Lastly, 50 µl of stop solution was added to each well and read the plate at 450nm within 30 minutes with the wavelength correction set at 540 or 570nm.

Vitamin D (25[OH]D) Laboratory Analysis

Circulating vitamin D (25[OH]D) levels were previously analyzed from serum as follows using the Immunodiagnostic Systems Ltd (IDS) 25-Hydroxyvitamin D enzymeimmunoassay (EIA) kit. Per the manufacturers assay protocol, the assay allows for 39 samples to be run in duplicate with 2 built in controls. The assay has good correlation with other methods for determining vitamin D (25[OH]D) in both serum or plasma with excellent sensitivity and specificity (5.3 - 7.4% variability within assay, and 5.3 - 11.7% between assays (Hyppönen, Turner, Cumberland, Power, & Gibb, 2007). Briefly, the procedure for vitamin D (25[OH]D) EIA analysis was completed by adding 25 µl of serum to 1 ml of a propriety buffer reagent that dissociates vitamin D from protein. The diluted samples were then incubated at room temperature in a vitamin D antibody coated plate for 2 hours. Enzymes that bind selectively to the vitamin D biotin complex was then added and the samples were then washed. Next, a chromogenic substrate was added to the samples. The reaction was then stopped by adding a hydrochloric acid solution so the intensity can be measure of the treated vitamin D biotin labeled complex using a microtitre plate reader.

DNA Methylation Laboratory Analysis

To complete the DNA methylation analysis portion of this study, the following steps were completed to examine the hypothesis that "differences in maternal DNA methylation patterns during early pregnancy are present in women with and without GBS colonization." DNA methylation was previously quantified in peripheral white blood

cells, collected in the first trimester, in a subset of six participants in the primary study using the Infinium bead-based array platform at the University of Minnesota's Biomedical Genomics Center and Illumina Core laboratory. Of those participants, four screened negative for GBS and two screened GBS positive in the third trimester.

In order to identify trends and potential differences in methylation between GBS positive and negative women, 12 additional samples were analyzed for this study (n=5 GBS negative; n=7 GBS positive). Samples were randomly selected from the remaining samples from the primary study. Genome-wide DNA methylation was determined using the Infinium platform (Illumina, Inc., San Diego, CA) at the University of Minnesota's Biomedical Genomics Center and Illumina Core Laboratory. Results from these samples were pooled with the results previously obtained to evaluate changes in DNA methylation between groups using the GenomeStudio Data Analysis Software that is compatible with output from Illumina. DNA methylation analysis was also completed using R 3.0.2 Statistical Environment for Windows.

DNA methylation at over 485,000 individual CpG dinucleotides in peripheral blood collected from women with and without GBS colonization (n=9 GBS negative; n= 9 GBS positive) was quantified. Samples underwent bisulfite conversion, which turns unmethylated cytosines to uracil and leaves methylated cytosines unaffected. The change allowed two query probes to detect differences in methylation between samples for predefined segments of DNA based on binding of specific nucleotides. This process is known to have greater than 99% conversation efficiency, yielding highly sensitive single-nucleotide resolution of methylation status. The quantitative amount of methylation at each CpG loci was reported as a beta value (β = methylated sites / (unmethylated gene +

methylated sites + 100)) (Bibikova et al., 2011). Beta values range from 0-1 and represents the percentage of DNA methylation (0-100) present at a given site. The Infinium platform processes 12 samples per plate.

Individual CpG dinucleotides that differed in DNA methylation between women with and without GBS colonization by +/- 20% as a percentage of total methylation, were designated as differentially methylated. Statistical significance of these changes between groups was then determined by *t*-tests (using a two-tailed, alpha of 0.05) (see Data Analysis section). In order to determine the potential for clinical relevance of differentially methylated CpG sites, functional analysis of genes with differentially methylated CpGs was also completed (see Functional Analysis section).

The specific function of individual genes was verified utilizing GeneCards and Pubmed. In order to validate the array, six genes were selected for validation of methylation patterns via bisulfite sequencing by Genwiz, Inc. (South Plainfield, NJ), based on greatest mean methylation differences. Two of the six gene sites also underwent a cloning step to assist in identifying methylation on specific CpG sites since the quality of DNA may not have been sufficient to obtain accurate reading without magnification because it is a clinical blood sample with multiple cell types present. After the samples were bisulfite treated, sequencing was used to determine the DNA sequence and validate the results from the Illumina Infinium array. The next sections describe the process in greater detail.

Extraction of White Blood Cells from Blood

For the initial study, whole blood was obtained via venipuncture and collected in a purple top vacutainer (B-D) blood tube containing ethylenediaminetetraacetic acid

(EDTA). Collection tubes were immediately placed on ice during transport to the GFHNRC, followed by centrifugation at 1,000 x g at 4˚C for 10 minutes. Plasma was removed and buffy coat collected and aliquoted into 1ml vials and frozen at -80 degrees centigrade until they were analyzed. The WBCs separate from the rest of the sample into a "buffy coat" layer as a result of being mixed with the EDTA. All other cell types are lysed from the EDTA solution and the WBCs form a pellet in the bottom of the tube when centrifuged. After centrifuging, the liquid components above the pellet were pipetted off leaving the WBC pellet. The remaining debris was washed from the WBC pellet by resuspending the pellet in EDTA solution, centrifuging, and discarding the fluid.

DNA Extraction from WBCs

Pelleted cells from the previous step were resuspended and homogenized by adding 270 µl of lysis buffer to the sample in preparation for DNA extraction. Next, 30µl of proteinase K (10 mg/ml from Invitrogen) was added to each sample tube and the tubes were vortexed to thoroughly mix the samples. The samples were then incubated overnight at 60˚ C. The next day, the samples were incubated at 100˚C to deactivate the enzyme. The samples were then transferred into 2ml phase lock microcentrifuge tubes and 300 µl phenol:choloroform:isoamyl alcohol (25:24:1) was added to each sample. The tubes were then centrifuged and the upper aqueous phase was transferred into a fresh microcentrifuge tube. Next 600 µl of 100% ethanol and 30 µl of 7.5 M ammonium acetate were added to each tube containing the aqueous phase and samples were incubated for 3 hours at -20˚C. The samples were then centrifuged, the fluid was discarded and the remaining pellet was washed with 300 µl of 75% ethanol. The samples were centrifuged again and the remaining DNA pellets were allowed to air dry.

DNA Quantification

All of the DNA sample pellets from the previous procedure were then individually re-hydrated in 100 µl of nuclease free water and the DNA concentrations of each sample were determined using the Epoch micro-volume spectrophotometer system. After each well of the spectrophotometer was calibrated by running all 16 plate wells with 2 µl DNAse and RNase free water, 2 µl of a DNA sample was added to each well to determine the amount of DNA in the sample. For this study, 30µl of each of the 12 samples were plated and sent to the University of Minnesota for genome wide methylation analysis. The remaining samples volumes were used to validate DNA methylation patterns identified by the Illumina platform.

Bisulfite Conversion of DNA

To verify the DNA methylation patterns identified by the Illumina Infinium array, remaining DNA from six participants (3 GBS positive and 3 GBS negative) were treated for bisulfite conversion using the EZ DNA Methylation Kit (Zymo Research). Per the kit specifications, the protocol has greater than 99% conversion efficiency converting unmethylated cytosine residues to uracil and greater than 99% protection of methylated cytosines. Samples were prepared, based on the concentration of DNA in each sample after being quantified using the Epoch micro-volume spectrophotometer system, for an input DNA amount of 500ng/sample. For example, if the amount of DNA in a sample was 59.9 ng,

500 ng (1µl/59.5ng) = 8 µl of the sample with DNA was used for next step

The appropriate volume of DNA from each sample for 500ng/sample along with 5 µl of M-dilution buffer and nuclease free water was added to a microcentrifuge tube to

yield a total volume of 50 µl for each sample. For example, 8µl of DNA sample + 5 µl M-dilution buffer + 37 µl water = 50 µl. Samples were then incubated for 15 minutes at 37˚ C. Next, 100 µl of CT conversion reagent was mixed with each sample and the samples were then incubated in the dark for 14 hours at 50˚ C.

The samples were then incubated at 0-4˚C (on ice) for 10 minutes. While the incubation was ongoing, 400 µl of M-binding buffer was added to a Zymo-spin IC column tube for each of the six samples, then the samples were added to the Zymo-spin IC column tubes and mixed by inversion after incubation was complete. The tubes were then centrifuged at full speed (≥ 10,000 x g) for 30 seconds. The flow-through was discarded and 100 µl of M-wash buffer was added and the tubes were again centrifuged at full speed for 30 seconds. Next, 200 µl of M-desulphonation buffer was added to each sample tube and the samples were incubated at room temperature for 15-20 minutes. The samples were then centrifuged at full speed for 30 seconds and then washed with 200 µl of M-wash buffer. The samples were then centrifuged at full speed for 30 seconds and washed again with 200 µl of M-wash buffer and centrifuged at full speed for 30 seconds. The Zymo-spin IC columns for each sample were then placed into a fresh 1.5 ml microcentrifuge tube and 10 µl of M-elution buffer was added directly to the column matrix and centrifuged for 30 seconds. The samples were then stored at -20˚C until the primers were designed to complete polymerase chain reaction (PCR) and validation sequencing.

Primer Design

Primers were designed to cut the DNA and validate methylation patterns in six sites identified as differentially methylated between women with and without GBS colonization. The sites were selected based on greatest difference in methylation between groups that were known genes (3 sites with loss of methylation, 3 with methylation gain). Two of the six sites were also be amplified using TOPO cloning reactions in anticipation that the amplification of the DNA samples via PCR alone may not yield adequate quality samples for sequencing. The DNA sequences for the six sites were determined utilizing University of California, Santa Cruz (UCSC) Genome Browser (Kent et al., 2002). The chromosome number and map information for the genes, from the Illumina Infinium output, were entered into the search box. After selecting "DNA" from the "view" dropdown a new window opens allowing selection of how many nucleotides away from the CpG of interest you would like the software to search for a suitable region to cut the DNA for sequencing. Segments 250 bases upstream and 250 bases downstream were entered as criteria to ensure the region of interest would be in the resultant product after PCR was complete. The actual DNA sequence was then copied and pasted into The Li Lab Department of Urology, University of California, San Francisco MethPrimer design tool (Li & Dahiya, 2002). The identified primer sequence was then used to order the primers from Integrated DNA Technologies (IDT) (http://www.IDTDNA.com) and utilized in the PCR step. DNA site sequences, primer sequences and properties of the primers supplied by IDT are located in Appendix E. When the primers arrived from IDT, all were rehydrated into a stock solution of 1µg/µl. For example, one vial was 0.27 mg/vial, therefore 270 µl of DNAse/RNAse free water was added to rehydrate the

sample. From the stock solutions, a working concentration of 200 ng/µl for each primer was created by adding 20 µl of stock and 80 µl of DNAse/RNAse free water to a 1.5 ml microcentrifuge tube. Stock and working concentrations for each primer were then stored at -20˚C.

Polymerase Chain Reaction

To complete the PCR step for each primer, the *ZymoTaq*™ PreMix protocol was completed using the bisulfite treated DNA (described previously) from each of the six participants (3 GBS positive, 3 GBS negative) for all six primers. The following reaction set up was used for each primer for each sample of DNA: 25 µl *Zymo Taq*™ PreMix, 1 µl forward primer, 1 µl reverse primer, 1 µl bisulfite treated DNA, 22 µl DNase/RNase free water. The following conditions were used for the PCR reaction using a hot start: initial denaturation at 95˚C for 10 minutes, followed by 35 cycles of denaturing at 95˚C for 30 seconds, annealing at 54ºC for 35 seconds, and extension at 72˚C for 45 seconds. After the 35 cycles were completed, final extension at 72˚C for seven minutes was done followed by a hold at 4˚C for 30 minutes. The PCR products were visualized by electrophoresis in agarose gel (0.5 gm agarose, 50ml TAEx1 buffer, with 5µl DNA star) at 100V with a 100bp ladder. 13.5 µl of each sample was loaded into the gel with 1.5 µl x10 loading dye. Products were produced for all primers (Appendix F).

Methylation Validation

In order to validate the methylation patterns identify by the Illumina Infinium array, the PCR products that were produced in the previous step were sent to Genewiz, Inc. (South Plainfield, NJ) for sequencing. A portion of the PCR products from ANXA2

and RHPN1 were cloned using the Topo TA-cloning kit; all remaining products were purified using the QIAquick® PCR purification kit described in the following sections.

PCR Clean-up: First, five volumes of PB buffer were added for every volume of PCR product for each sample. The samples were then transferred to QIAquick spin column tubes and centrifuged for 60 second. After discarding the flow through, the samples were washed with 0.75ml of PE buffer, then centrifuged 60 seconds. After removing the flow through, samples were centrifuged again for 60 seconds and the columns were then placed in a clean 1.5 ml microcentrifuge tube. Then, 50 µl of EB buffer was added to the center of the column and centrifuged for one minute to elute the DNA. All samples were then sent to Genewiz, Inc. (South Plainfield, NJ) for sequencing (Appendix G). Since the samples were not of adequate quality for sequencing, cloned samples were also sent to validate the array.

TOPO TA-cloning: PCR products from ANXA2 and RHPN1 were used in the cloning protocol to optimize segments for sequencing using the TOPO TA Cloning Kit for Sequencing (Invitrogen, Life Technologies). Briefly, 4 µl of fresh PCR product was incubated at room temperature for 30 minutes with 1 µl salt solution (1.2M NaCl, 0.06 M $MgCl_2$), and 1 µl TOPO vector, then placed on ice. Next, 2 µl from the completed reactions were added to a vial of One Shot® Chemically Competent *E. coli,* which were then incubated on ice for 30 minutes. The samples were then heat shocked at 42 degrees C for 30 seconds, then placed back on ice. Next, 25 µl of room temperature Super Optimal broth with Catabolite repression (S.O.C. medium) was added to each sample and the samples were shaken horizontally at 200 rpm for one hour at 37 degrees C. Kanamycin selective agar plates were prepared and two different concentrations (20 µl

and 50 µl) were plated for each sample to ensure adequately spaced colonies. The plates were then incubated at 37 degrees C overnight. Then, 10 white colonies were selected and suspended in individual test tubes with 3ml Luria Broth (LB broth) containing 50 µg/mL kanamycin and incubated overnight.

The plasmid DNA was then extracted using the QIAprep® Spin Miniprep Kit High-Yield Protocol. First, sample tubes were centrifuged at 8,000 rpm for three minutes at room temperature to pellet the bacteria with plasmid products. The pellets were then resuspended in 250 µl P1 buffer and placed into a microcentrifuge tube. Next P2 buffer was added to the samples and mixed by inverting the tube 10 times. Then 350 µl of N3 buffer was added to each tube, mixed by inverting 10 times and then centrifuged at full speed for 10 minutes. The supernatant was then pipetted into QIAprep spin columns and centrifuged at full speed for 60 seconds and the flow through was discarded. Next, the samples were washed with 500 µl PB buffer and centrifuged again at full speed for 60 seconds. Then 750 µl of PE buffer was added to each sample and centrifuged for 60 seconds. Samples were centrifuged for an additional minute to ensure all wash buffers had been removed. The spin columns were then placed into a clean microcentrifuge tube and 60 µl of EB buffer was added to each tube. Samples were incubated at room temperature for 60 seconds and were then centrifuged for one minute to elute the DNA. Samples were then sent to Genewiz, Inc. for validation; and sequencing was successfully matched (Appendix G).

Data Analyses

Statistical Analysis

All cytokine and vitamin D (25[OH]D) data analysis procedures were performed using Statistical Package for Social Sciences (SPSS) version 22. All variables were examined for violations of statistical assumptions, including missing values and outliers, with SPSS Frequencies, and Explore. Upon completion of data screening to ensure all univariate and multivariate assumptions were met, descriptive statistics were completed to describe the sample characteristics and compare the groups (GBS positive and negative). The normality of the distribution of cytokines was determined by evaluating the skewness, kurtosis, and Kolmogorov-Smirnov statistics for each variable. Pearson correlation tests were completed to determine if cytokines and vitamin D (25[OH]D) levels co-varied throughout pregnancy. Since there was no significant correlation among these variables, a series of repeated measures ANOVAs were performed to compare each of the three cytokines (TNF-α, IL-6 and IL-10) and vitamin D (25[OH]D) between the two groups (n=16 GBS positive and n=26 GBS negative) from each of the three trimesters of pregnancy. Mauchly's test was used to identify violations in sphericity and bonferroni corrections were applied for multiple comparisons. The Greenhouse-Geisser correction for a violation in sphericity was used for vitamin D (25[OH]D) analysis. Given the restrictions of performing a secondary analysis, it was not feasible to change the sample size. However, performing a compromised post-hoc power analysis using "G*power" software indicated that the current sample size allowed 69% power and a medium effect size of 0.3. This particular type of power analysis was developed for studies where the sample size cannot be altered and equates the risk of committing a

Type I and Type II error (Faul, Erdfelder, Buchner, & Lang, 2009; Faul, Erdfelder, Lang, & Buchner, 2007). This approach to analysis helps in identifying trends and effect size so sample size can be determined in future investigations that have adequate power.

Mean values for serum cytokine and vitamin D (25[OH]D) levels were utilized to conduct a one-way multivariate analysis of variance (MANOVA) to correct for confounding resulting from data being collected over time, determine observed effect size and observed power. The results obtained were then used in G*power to determine the number of participants that would be required to see an effect with the power and effect size calculated from the study sample. All data analyses procedures in the study were performed using a two-tailed alpha of 0.05.

For DNA methylation analysis, fluorescent background intensities were normalized using the Genome Studio software. A series of negative controls are embedded into the assay that the software utilized to generate the detection p values for each probe. The specific normalization process applied when using GenomeStudio software is propriety information, and a limitation of using this software for methylation analysis (Gentleman et al., 2004; Hansen et al., 2013, Smyth, 2005). Probes that had detection *p* values greater than 0.05 were considered not to have reliable signal intensities and eliminated from analysis. Independent sample *t*-test comparisons were conducted to examine differences in mean DNA methylation at individual CpG sites between GBS positive and negative groups. The selection of independent sample *t*-test was driven by the fact the methylation testing was limited to 18 participants and DNA methylation was reported as a percentage of methylation at each CpG site. Findings from this analysis will inform the researcher of the effect size so that proper sampling can be determined for

future research. All statistical techniques were performed using a 2-tailed alpha of < 0.05 or 95% confidence interval. Statistical analysis of raw data provided by Illumina was also completed using the minfi package (Hansen, Ayree, & Irizarry, 2013) in the R programming environment (R Core Team, 2013) to first normalize the data and identify differentially methylated regions by using the limma package (Smyth, 2005). Significance testing was also completed using Benjamini and Hochberg's false discovery rate to correct for multiple comparisons (Benjamini & Hochberg, 1995) because it is the recommended standard and used for a majority of studies investigating specific clinical outcomes in human samples (Allison, Cui, Page, & Sabripour, 2006; Benjamini & Hochberg, 1995; Michels et al., 2013; Storey, 2003; Tusher, Tibshirani, & Chu, 2001; Wilhelm-Benartzi et al., 2013; York, 2003). R code used for analysis with normalization figures are located in Appendix I. Established work flows with detailed explanations of computations and code are available in open source, published workflows and user manuals for minfi (Hansen, Ayree & Irizarry, 2013) and limma (Smyth, 2005) packages.

Briefly, raw data were preprocessed using a series of minfi commands (Appendix I) to normalize the intensity for methylated and unmethylated channels (G. K. Smyth, Yang, & Speed, 2003). For the statistical analysis of DNA methylation, M-values were used because they are logit transformed β values and result in data that more closely follows a normal distribution and preferred for statistical methods such as a *t*-test. Beta values exhibit severe heteroscedasticity for extremes of methylation (highly methylated or unmethylated). Because the logit transformation corrects for this, M-values outperform beta values in terms of detection rate and true positives for detecting differences of CpG sites on the extremes of percent methylated. Dedeurwaerder et al.'s (2011) conversion

method for beta to M-values was utilized for analysis because a peak correction can be performed on the M-values to adjust the Infinium II probes. Infinium claims there are no differences between type I and II detection probes that affect detection of differential methylation (Bibikova et al., 2011). However, design I type probes have been shown to be more stable and have a more extensive detection range than type II probe, therefore require correction for statistical analysis (Dedeurwaerder et al., 2011). Using M-values results in more accurate analysis of the data, although are difficult to clinically interpret. Therefore, after using the logit transformed values (M-values) for statistical analysis, the data were back transformed to yield peak-corrected beta values for reporting and to allow for easier clinical interpretation of findings because beta values are biologically meaningful (Du et al., 2010). Sites with known single nucleotide polymorphisms (SNPs) were excluded from the analysis as these sites can disturb the accuracy of the Infiniuim probes (Bibikova et al., 2011b; Dedeurwaerder et al., 2011). Additionally, SNP sites were excluded to remove confounding effects that could result from having two different modifications at the same site. Furthermore, SNP analysis is beyond the scope of this project and the original consent form was not written to include this type of analysis.

After all the raw files were normalized to ensure the peaks of the Infinium II design bead type probes were comparable to the peak locations of the Infinium I bead bead type probes and converted to M-values, differential methylation analysis was conducted using the limma package. This package was selected because it performs additional normalization and offers commands that are ideal for use with small data sets. Essentially, Symth (2005) designed the program to strengthen analysis for small data sets by "borrowing information across genes" (p. 4) using empirical Bayesian methods to help

control for variance; which results in less false discoveries for small data sets than other programs. All normalization and analytics commands were designed for two color channel arrays, including the Infinium platform used for this study (Gentleman et al., 2004; Smyth, Yang, & Speed, 2003; Smyth, 2004; Smyth, 2005). The series of R commands fit each CpG site into a linear model using least squares fitting, then a contrast matrix was created to compare values between GBS positive and GBS negative women and in the final normalization command, the empirical Bayesian command incorporates the use of array weights to improve identification of CpG sites that are most likely to be different (Ritchie et al., 2006; G. Smyth, 2005). This correction for multiple comparisons increases the likelihood of identifying clinically meaningful results and reduces the chance of identifying false positives.

Functional Analysis

In order to evaluate the last research question, that differential DNA methylation in genes associated with immune function and inflammation, a functional analysis of genes associated with differential DNA methylation was completed using Database for Annotation, Visualization, and Integrated Discovery (DAVID) Bioinformatics Resource version v6.7 (Huang, Sherman, & Lempicki, 2009a; Huang, Sherman, & Lempicki, 2009b). Functional analysis reveals potential mechanistic underpinnings of GBS colonization or other clinically relevant information. Gene lists with significant differentially methylated CpG sites were uploaded into the DAVID v6.7 database. Sites with a gain of methylation in women with GBS colonization were uploaded as one group and those with a loss of methylation were entered as a separate group, to investigate changes in function based on differences in methylation at those sites. DAVID

determines functional categories of genes and classifies them with “high” stringency selected by the investigator for analysis (Huang, Sherman, & Lempicki, 2009a; Jiao et al., 2012). Using a classification stringency of “high” to determine functional annotation clusters identifies a “tight, clean, and smaller numbers of clusters” (Huang et al., 2009a, p. 47) that are more likely to be associated with biologically meaningful results. Output generated by the DAVID software also includes enrichment scores for each cluster, with higher numbers indicating there may greater involvement in the disease state being studied. Pathways with enrichment scores greater than 1.3 are likely most important in the functional or disease process, although lower scoring clusters could also offer insight into biological mechanisms associated with the disease process. Statistical calculations presented related to the function of the genes are determined utilizing conservative correction methodologies for multiple comparisons (Bonferroni; Benjamini, and Hochberg) (Huang et al., 2009). The specific function of individual genes was verified utilizing GeneCards® and Pubmed and reviewed in the discussion chapter.

CHAPTER IV

RESULTS

This chapter presents the results by first providing general descriptive statistics of the sample characteristics, followed by presentation of the statistical analyses that were completed to answer the research questions as presented in chapter one.

Description of Sample

A total of 42 women had an outcome of uncomplicated pregnancy in the parent study and were included in the analysis for this study. Among those women, 38% were GBS positive (n = 16) and 62% were GBS negative (n = 26), 9 participants from each group had DNA methylation analyzed. The sample population was primarily comprised of Caucasian women (84%), reflecting the demographics of the area from which the subjects were recruited. Race was self-identified by participants upon entry to the parent study. Other demographic data of the participants are presented in Table 6, which shows that no statistically significant differences were found (age, weight, gestational age at delivery, or infant gender) between GBS positive and negative women. Notably, there are no significant differences in co-morbidities, infections, or antimicrobial usage. Infection data and antimicrobial data were extracted from the medical record for usage at any point during the pregnancy because it may alter the vaginal microbial composition. Currently, there is no information in the literature describing how long and when vaginal microbiota returns to a woman's baseline composition or how it may contribute to GBS status.

Table 6. Maternal Characteristics Compared between GBS Positives and Negatives

Variable	Total Sample (N=42)	GBS + (n=16)	GBS – (n=26)	t /χ^2	p
Maternal Age (years) [M(SD)]	26.6 (4.10)	25.1 (3.05)	27.5 (4.46)	-1.806	0.194
Maternal Prenatal Weight (pounds) [M(SD)]	168.7 (35.5)	161.2 (31.6)	173.3 (37.5)	-1.074	0.573
Gestational Age at Birth (weeks) [M(SD)]	39.1 (1.25)	38.7 (1.45)	39.2 (1.08)	-1.488	0.763
Race [n(%)]					
Caucasian	38 (90.5)	14 (87.5)	24 (92.3)	1.679	0.432
Multi-racial	3 (7.1)	1 (6.25)	2 (7.7)		
Native Hawaiian/ Pacific Islander	1 (2.4)	1 (6.25)	0 (0)		
Infant Gender [n(%)]					
Female	22 (52.4)	12 (75)	10 (38.5)	5.480	0.065
Male	19 (45.2)	4 (25)	15 (57.7)		
Missing	1 (2.4)	0 (0)	1 (3.8)		
Maternal Asthma [n(%)]	7 (16.7)	2 (12.5)	5 (19.2)	0.323	0.570
Infection [n(%)]					
None	28 (66.7)	10 (62.5)	18 (69.2)	3.998	0.677
Respiratory	3 (7.1)	1 (6.2)	2 (7.7)		
Urinary Tract	3 (7.1)	1 (6.2)	2 (7.7)		
Chlamydia	1 (2.4)	0	1 (3.8)		
Herpes	1 (2.4)	1 (6.2)	0		
Tuberculosis	1 (2.4)	0	1 (3.8)		
Unknown	5 (11.9)	3 (18.8)	2 (7.7)		
Antimicrobial [n(%)]					
Yes	13 (31)	5 (31.2)	8 (30.8)	0.001	0.974
No	29 (69)	11 (68.8)	18 (69.2)		

P indicates significance level based on a two-tailed alpha of 0.05

Serum Cytokine and Vitamin D (25[OH]D) Analysis

Research Question 1: Are serum levels of TNF-α, IL-6, IL-10, and vitamin D (25[OH]D) different in pregnant women with GBS colonization than pregnant women without GBS colonization?

Cytokines and vitamin D (25[OH]D) levels were not normally distributed for the 42 participants (Table 7). However, ANOVA is considered robust to departures in normality. Kepple and Wickens (2004) argue when sample size is greater than approximately 12 ANOVA analyses can be successfully completed if normality has been violated. Further, the laboratory values analyzed in this study were highly variable and were not normally distributed in similar studies that evaluated cytokine levels (Curry et al., 2008; Makhseed et al., 2000). Because cytokine levels are variable between individuals (not normally distributed, making it impossible to identify outliers) all participant data were included in the analysis in order to assess if evaluation of serum cytokine levels and vitamin D (25[OH]D) would be a useful clinically relevant measure

Table 7. Tests of Normality for Serum Cytokines and Vitamin D (25[OH]D)

Variable*	**Skewness**	**Kurtosis**	**Kolmogorov-Smirnov**			**Shapiro-Wilk**		
			Stat	*df*	*P*	*Stat*	*df*	*P*
TNF-α	1.484	3.989	0.141	42	<0.001	0.851	42	<0.001
IL-6	4.999	26.755	0.478	42	0.036	0.898	42	0.001
IL-10	0.462	-0.059	0.107	42	<0.001	0.273	42	<0.001
(25[OH]D)	1.918	5.797	0.196	42	0.200	0.947	42	0.052

Stat = statistic
df = degrees of freedom
P = significance 2-tailed 0.05
**mean laboratory values across all trimesters*

for predicting GBS colonization susceptibility. Serum cytokine and vitamin D (25[OH]D) levels were described by time during pregnancy with the following notations: T1 = first trimester (10- 14 weeks); T2 = second trimester (22-26 weeks); and T3 = third trimester (32-36 weeks). Correlation tests indicate vitamin D (25[OH]D) levels were not correlated with serum TNF-α, IL-6, or IL-10 levels during pregnancy, except T1vitamin D (25[OH]D) and T3 IL-10 (Table 8). Therefore, vitamin D (25[OH]D) was not used as a covariate and repeated measures ANOVA was completed on serum vitamin D (25[OH]D), TNF-α, IL-6 and IL-10 independently.

Table 8. Serum Vitamin D (25[OH]D) and Cytokine Correlations

Variable	T1 Vitamin D (25[OH]D)		T2 Vitamin D (25[OH]D)		T3 Vitamin D (25[OH]D)	
	r	*P*	*R*	*P*	*r*	*P*
T1 TNF-α	-0.111	0.483	-0.265	0.090	-0.277	0.076
T2 TNF-α	-0.001	0.994	-0.005	0.975	-0.101	0.523
T3 TNF-α	-0.018	0.091	-0.031	0.846	-0.067	0.674
T1 IL-6	-0.192	0.224	-0.140	0.375	-0.162	0.305
T2 IL-6	-0.224	0.154	-0.147	0.352	-0.186	0.238
T3 IL-6	-0.215	0.172	-0.162	0.306	-0.183	0.245
T1 IL-10	0.142	0.368	0.074	0.642	0.063	0.693
T2 IL-10	0.007	0.963	0.093	0.557	0.018	0.911
T3 IL-10	**0.310***	**0.046***	0.214	0.173	0.129	0.415

r = Pearson correlation
P=significance 2-tailed 0.05

TNF-α

TNF-α levels were compared between GBS positive (n=16) and negative (n=26) women across three trimesters of pregnancy. Mean TNF-α levels were calculated for each group in each trimester (Table 9, Figure 4a). A repeated measures ANOVA revealed no significant difference in TNF-α between GBS positive and negative women, $F(2, 80) = 2.187$, $p = 0.119$, $\eta^2 = 0.052$ with an observed power of 0.453.

IL-6

IL-6 levels were compared between GBS positive (n=16) and negative (n=26) women across three trimesters of pregnancy. Mean IL-6 levels were calculated for each group in each trimester (Table 10, Figure 4b). A repeated measures ANOVA revealed no significant difference in IL-6 between GBS positive and negative women, $F(2, 80) = 2.991$, $p = 0.056$, $\eta^2 = 0.070$ with an observed power of 0.566.

Table 9. Serum TNF-α Levels

	Mean (SD)	
	GBS +	*GBS -*
T1 TNF-α	6.20 (2.77)	6.44 (1.92)
T2 TNF-α	6.21 (2.74)	7.41 (4.75)
T3 TNF-α	8.71 (6.19)	7.46 (2.54)

Table 10. Serum IL-6 Levels

	Mean (SD)	
	GBS +	*GBS -*
T1 IL-6	1.06 (2.98)	1.28 (5.71)
T2 IL-6	1.67 (4.21)	1.04 (5.31)
T3 IL-6	1.25 (3.58)	1.31 (6.70)

IL-10

IL-10 levels were compared between GBS positive (n=16) and negative (n=26) women across three trimesters of pregnancy. Mean IL-10 levels were calculated for each group in each trimester (Table 11, Figure 4c). A repeated measures ANOVA revealed no significant difference in IL-10 between GBS positive and negative women, $F(2, 80) = 0.445$, $p = 0.642$, $\eta^2 = 0.011$ with an observed power of 0.120.

Vitamin D (25[OH]D)

Vitamin D (25[OH]D) levels were compared between GBS positive (n=16) and negative (n=26) women across three trimesters of pregnancy. Mean vitamin D(25[OH]D) levels were calculated for each group in each trimester (Table 12, Figure 4d). Mauchly's test indicated a violation in the assumption of sphericity, $W(2) = 0.551$, $p < 0.05$. Based

on a Greenhouse-Geisser correction for the sphericity violation, a repeated measures ANOVA revealed no significant difference in vitamin D (25[OH]D) between GBS positive and negative women, $F(1.380, 55.218) = 0.882$, $p = 0.384$, $\eta^2 = 0.022$ with an observed power of 0.169.

Table 11. Serum IL-10 Levels

	Mean (SD)	
	GBS +	*GBS -*
T1 IL-10	4.23 (3.27)	4.23 (3.06)
T2 IL-10	5.29 (3.48)	4.35 (4.55)
T3 IL-10	4.64 (4.56)	4.08 (2.85)

Table 12. Vitamin D (25[OH]D) Levels

	Mean (SD)	
	GBS +	*GBS -*
T1 Vit. D	27.22 (9.36)	23.76(6.21)
T2 Vit. D	30.18(10.94)	25.14(6.42)
T3 Vit. D	30.08(14.09)	24.89(7.06)

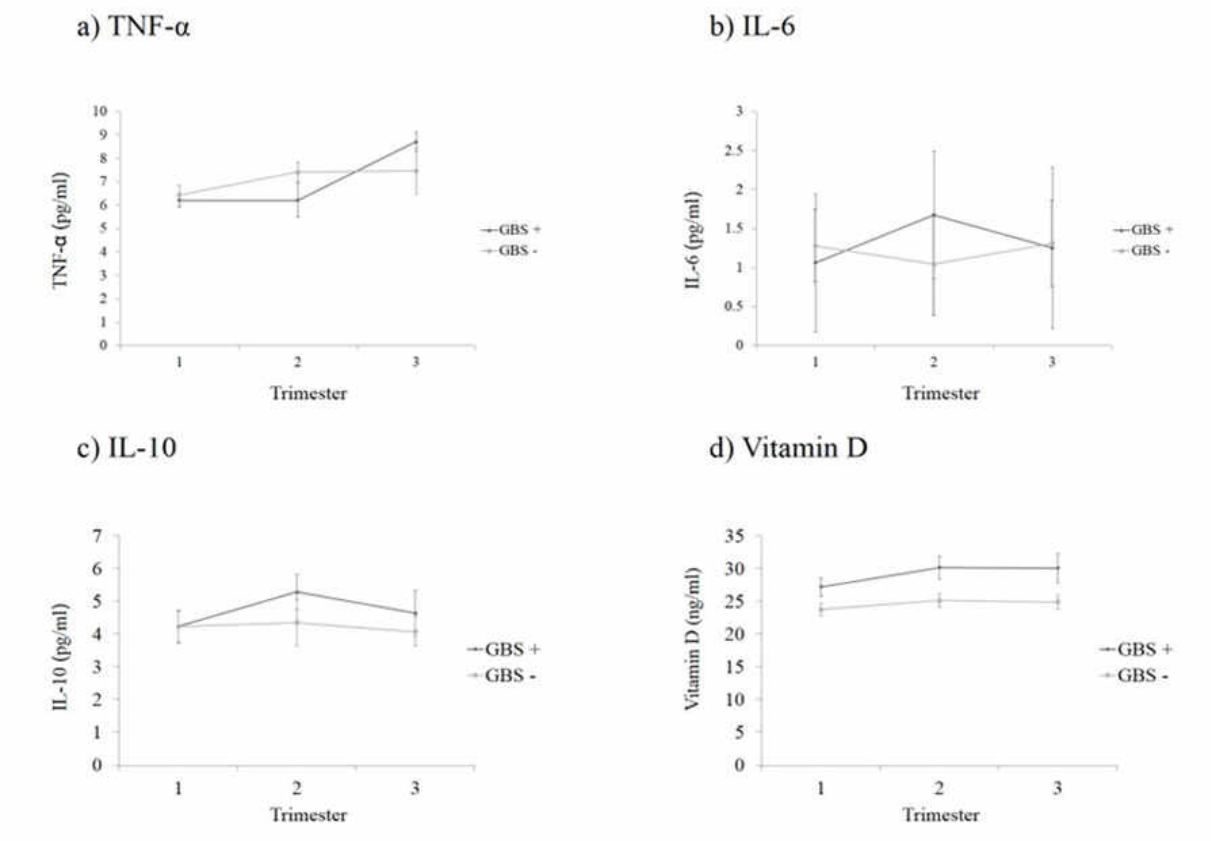


Figure 4. Cytokines and Vitamin D (25[OH]D) Across Pregnancy.

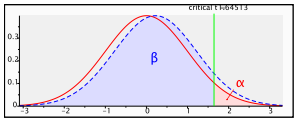
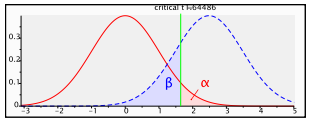
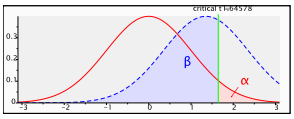
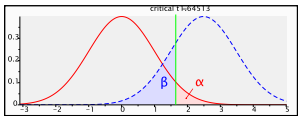
Multivariate Analysis

Mean values for serum cytokine and vitamin D levels were utilized to conduct a one-way multivariate analysis of variance (MANOVA) to correct for the confounding resulting from data being collected over time, determine effect size and observed power (Table 13). MANOVA revealed that serum cytokines and vitamin D levels do not significantly combine to affect GBS colonization status Wilks' $\Lambda = 0.927$, $F(4, 37) = 0.725$, $p = 0.581$, partial $\eta^2 = 0.073$. Further analysis comparing the sample based power and power needed to see a significant effect of cytokines and vitamin D on GBS status, indicates large sample sizes are required to see significant difference (Table 14)

Table 13. Multivariate Results

	Mean ± SD		**Partial η^2**	**Observed Power**
	GBS +	*GBS -*		
TNF-α (pg/ml)	7.04 ± 3.34	7.11 ± 2.75	<0.001	0.051
IL-6 (pg/ml)	1.33 ± 3.57	1.21 ± 5.89	<0.001	0.051
IL-10 (pg/ml)	4.72 ± 3.19	4.22 ± 3.16	0.006	0.077
(25[OH]D)(ng/ml)	29.16 ± 11.13	24.59 ± 6.24	0.068	0.384

Table 14. Comparison of Sample Based Power to 80% Power with Required N

Effect Size		Observed Power	Required Power
IL-10	0.006		
		0.077 GBS + = 2,228 : GBS - = 3,342 Total N = 5,570	0.8 GBS + = 286,230 : GBS - = 429,346 Total N = 715,576
Vit. D (25[OH]D)	0.068		
		0.384 GBS + = 658 : GBS - = 986 Total N = 1,644	0.8 GBS + = 2,230 : GBS - = 3,344 Total N = 5,574

DNA Methylation Analysis

Research Question 2: Are DNA methylation patterns different between pregnant women with GBS colonization and those without GBS colonization?

GenomeStudio Statistical Analysis

Analyses comparing the pooled mean methylation of GBS positive and negative women (n = 9/group) identified a total of 141 CpG dinucleotides that were differentially methylated in maternal peripheral white blood cells between women who had a positive screening for GBS and those who had a negative screening. Figure 5, depicts the average beta scores of the 141 CpG dinucleotides that were identified as differentially methylated by greater than 20% betweem GBS positive and GBS negative women and statistically significant with a p value of less than 0.05. Of the 141 CpG sites, 62 sites had a gain of

methylation and 79 sites had a loss in methylation in women with positive GBS screening compared to GBS negative women. CpG sites with differential methylation were distributed across the chromosomes (Figures 6a and 7a), associated with known genes (62%) and located predominantly in the body (66%) of genes (Figures 6b and 7b). Approximately half of the CpG sites (49%) were not associated with CpG islands (Figures 6c and 7c).

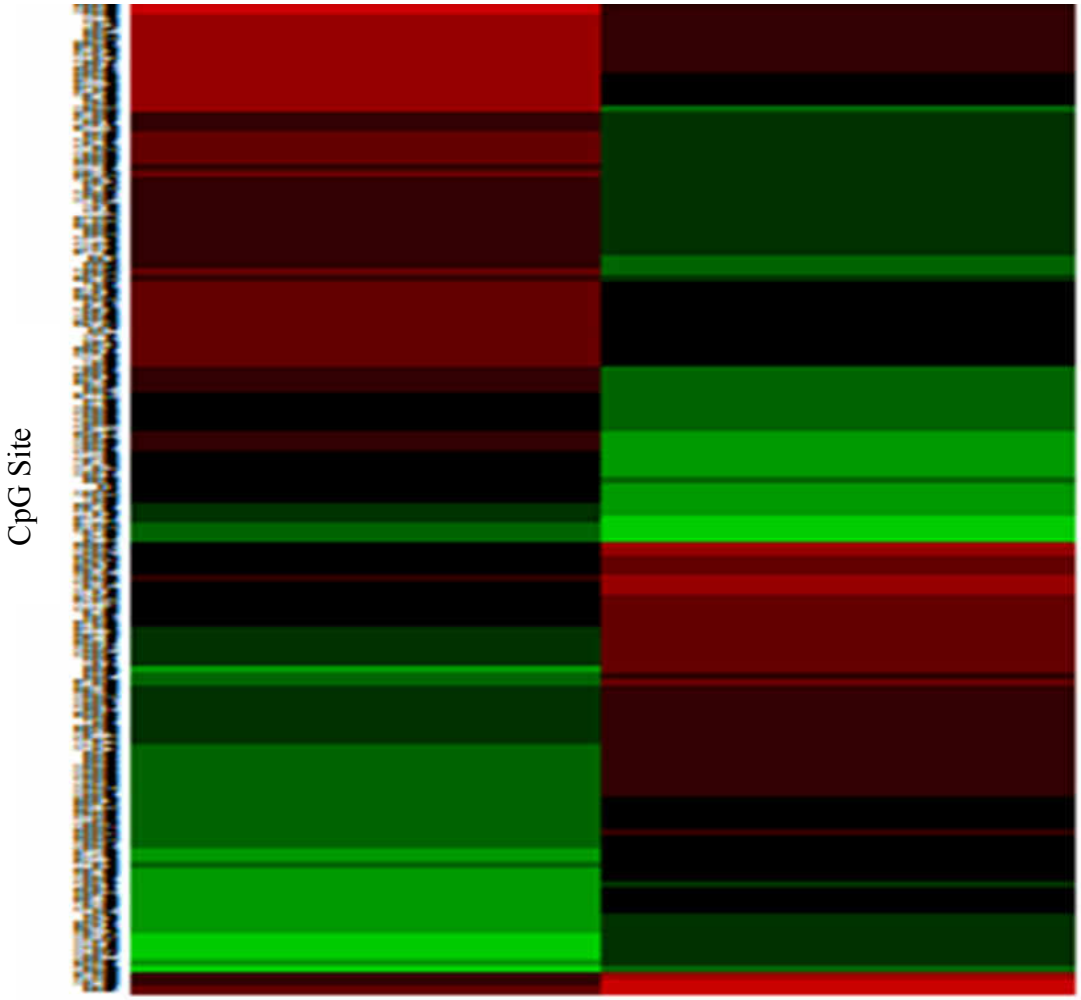
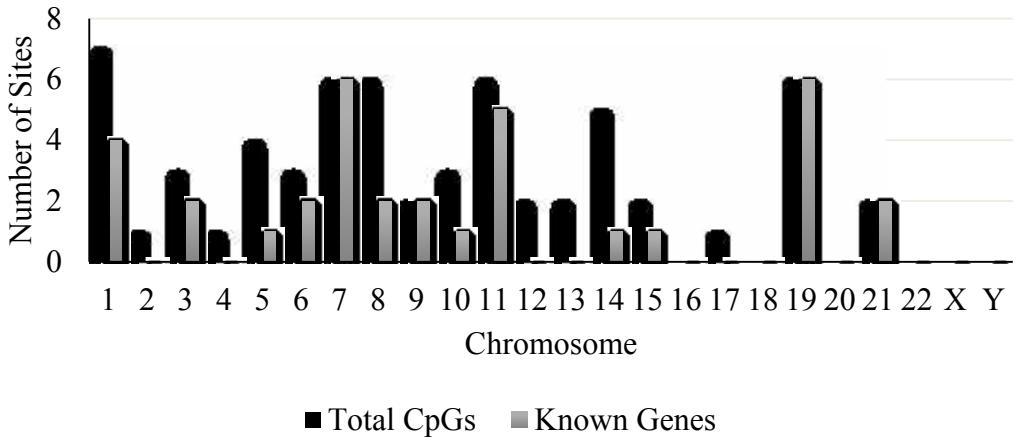
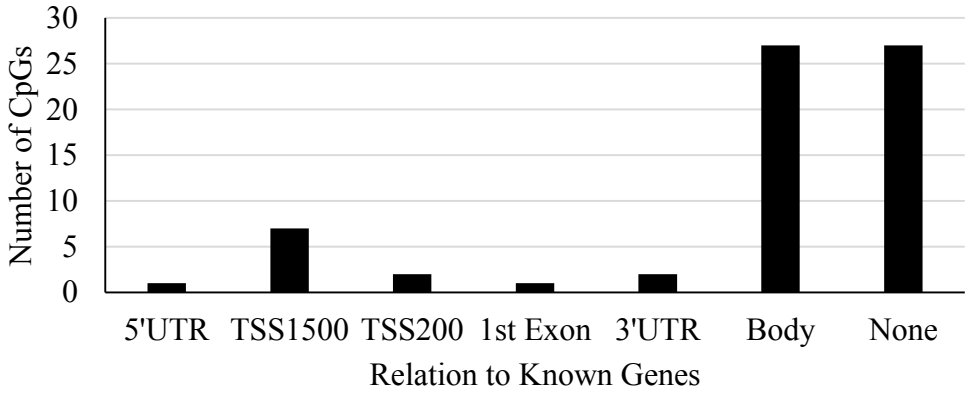


Figure 5. Differential Methylation with GBS Colonization (GenomeStudio).

a) Chromosomal Distribution



b) Proximity to Gene



c) Proximity to CpG Island

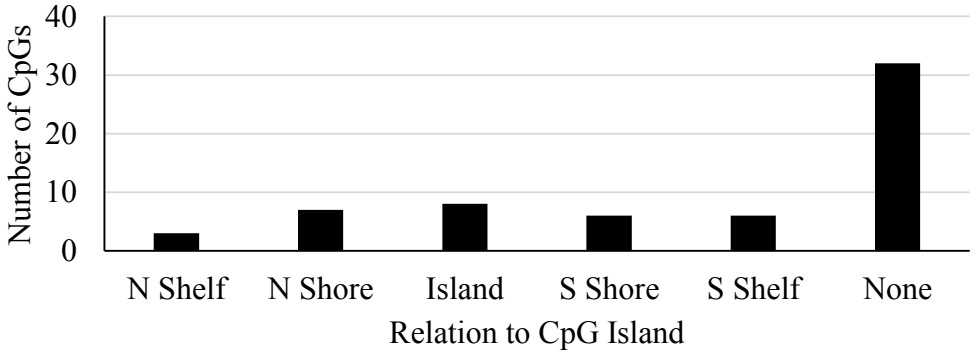
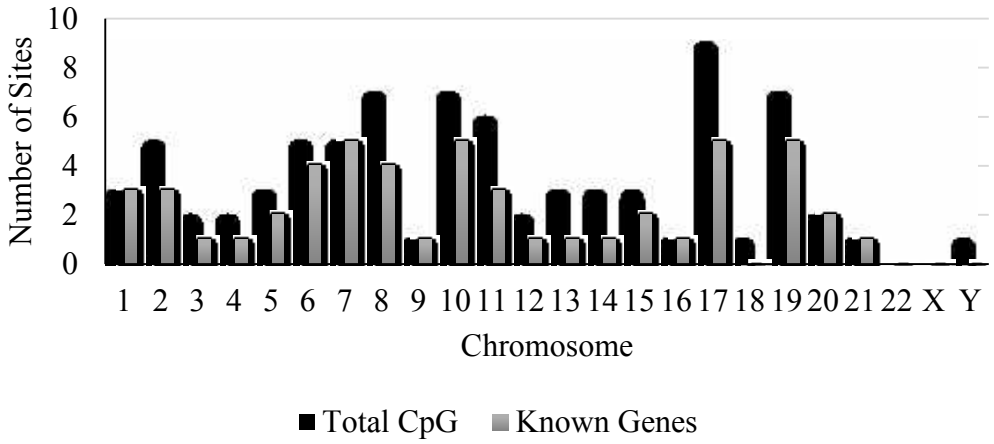
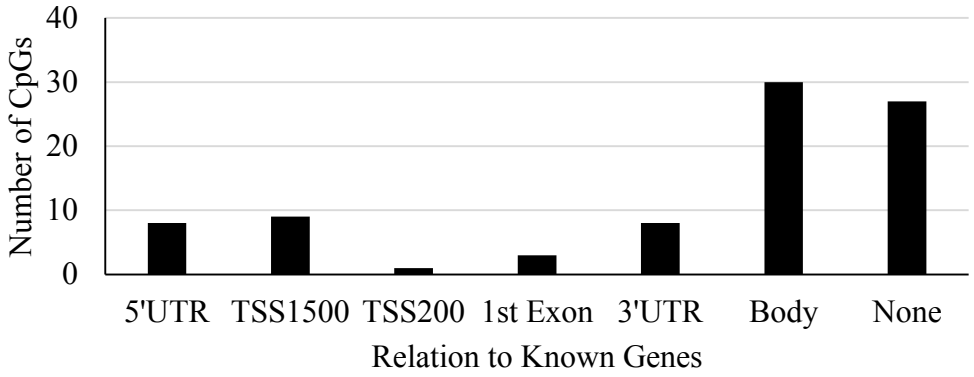


Figure 6. DNA Methylation Gain Distributions with GBS (GenomeStudio).

a) Chromosome Distribution



b) Proximity to Genes



c) Proximity to CpG Island

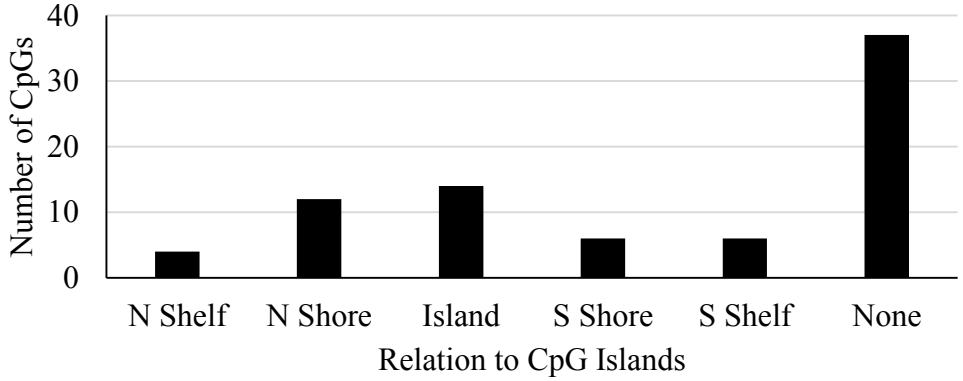


Figure 7. DNA Methylation Loss Distributions with GBS Colonization (GenomeStudio).

R Statistical Environment Analysis

Analyses conducted in R comparing the pooled mean methylation of GBS positive and negative women (n = 9/group) identified no statistically significant differences in methylation at any CpG site when a 20% change in beta value when FDR was used to determine significance at any level is used for analysis (adjusted p= 0.001, 0.05, 0.01). Figure 8 is a volcano plot presenting all CpG sites with parameters set to highlight blue any sites with greater than a 20% difference in methylation between GBS positive and GBS negative women that met significance using an FDR of 0.05. No blue CpG sites are present on the graph, indicating there is no statistically significantly different methylation at any CpG site by GBS status for this sample (R code and mathematical code returning zero results available in Appendix I). The plot is designed using the log values for more concentrated clear visualization of the data. In order to compare the results using the R programming environment to the results produced using the GenomeStudio software, the analysis was also conducted using the same unadjusted p value to determine significance at 0.05. A total of 125 CpG dinucleotides that were identified as having greater than 20% difference in methylation in maternal peripheral white blood cells between GBS positive and negative women. Of the 125 CpG sites, 54 sites had a gain in methylation and 71 sites had a loss in methylation for women with positive GBS screening at 37 weeks compared to GBS negative women when alpha is 0.05. CpG sites with differential methylation were distributed across the chromosomes (Figures 9a and 10a), associated with known genes (62%) and located predominantly in the body (59%) of genes (Figures 9b and 10b). Approximately half of the CpG sites (51%) were not associated with CpG islands (Figures 9c and 10c).

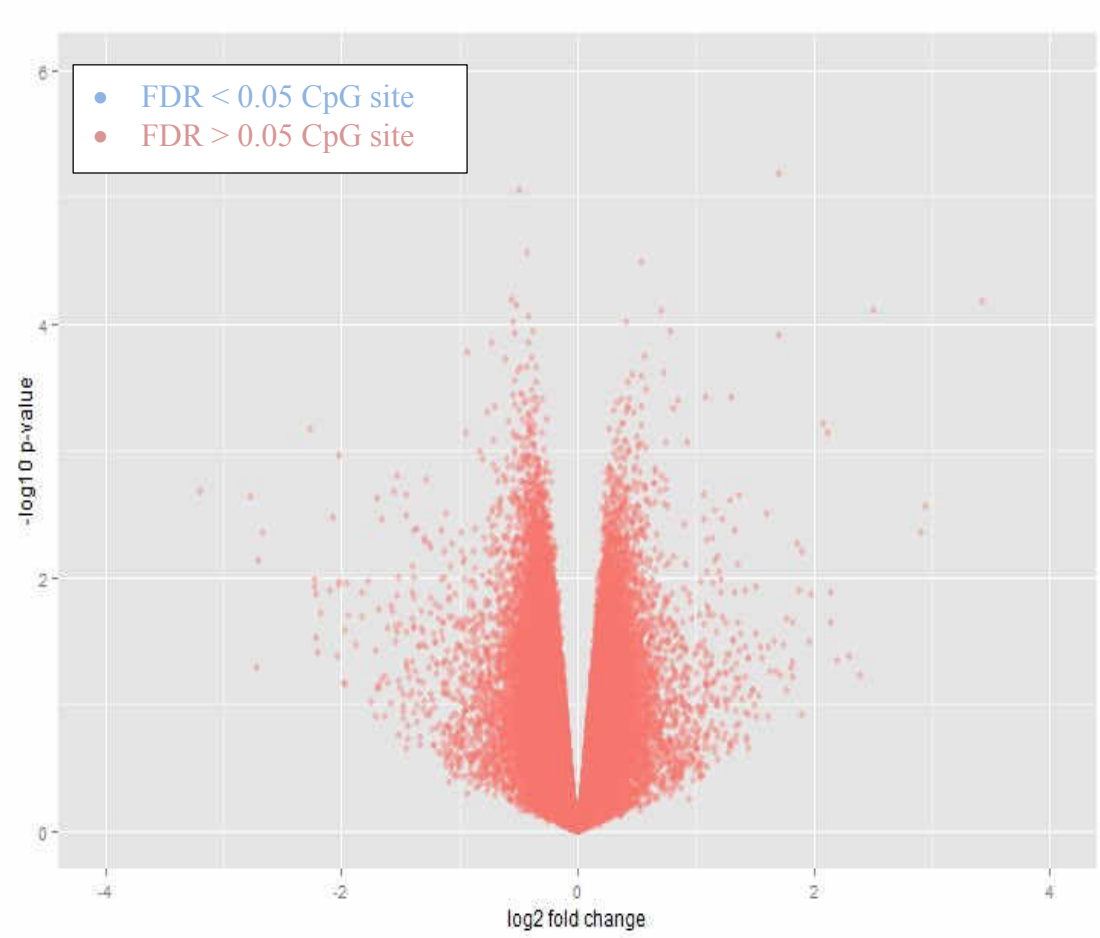
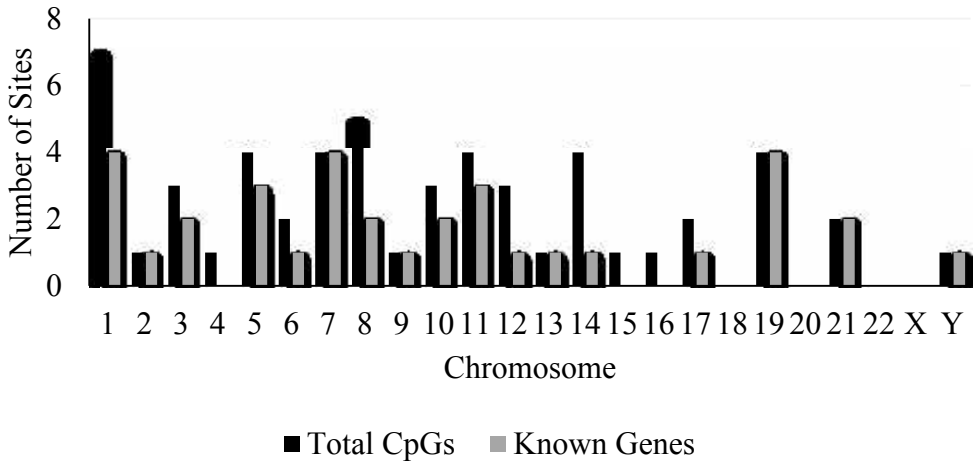
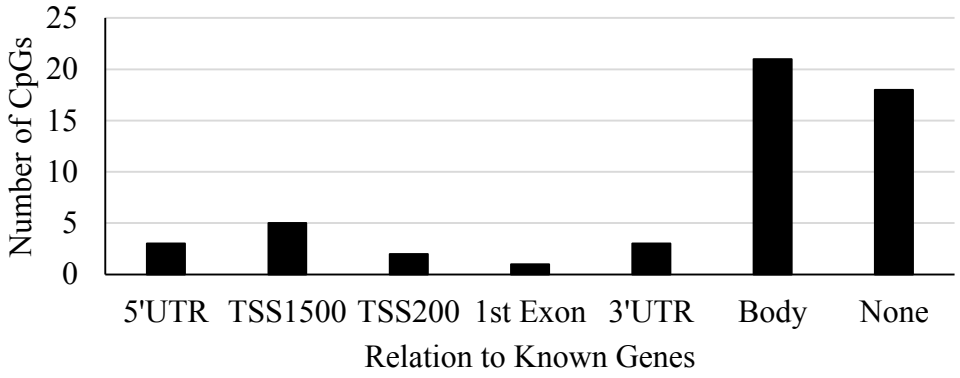


Figure 8. Volcano Plot of CpG Methylation Differences by GBS Status (R)

a) Chromosomal Distribution



b) Proximity to Gene



c) Proximity to CpG Island

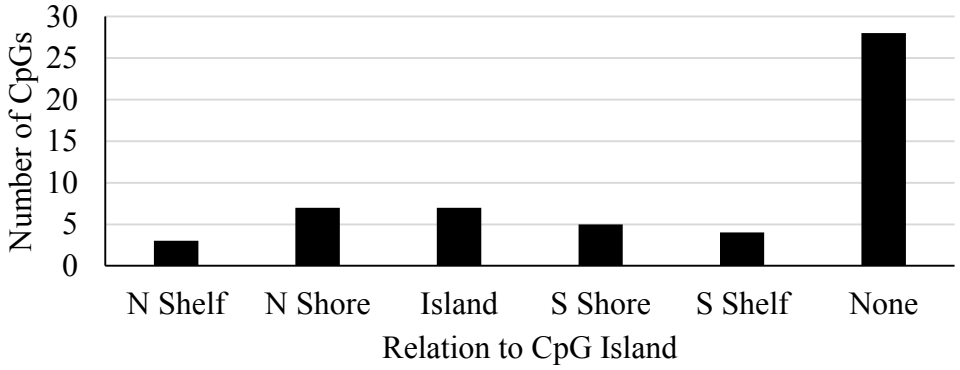
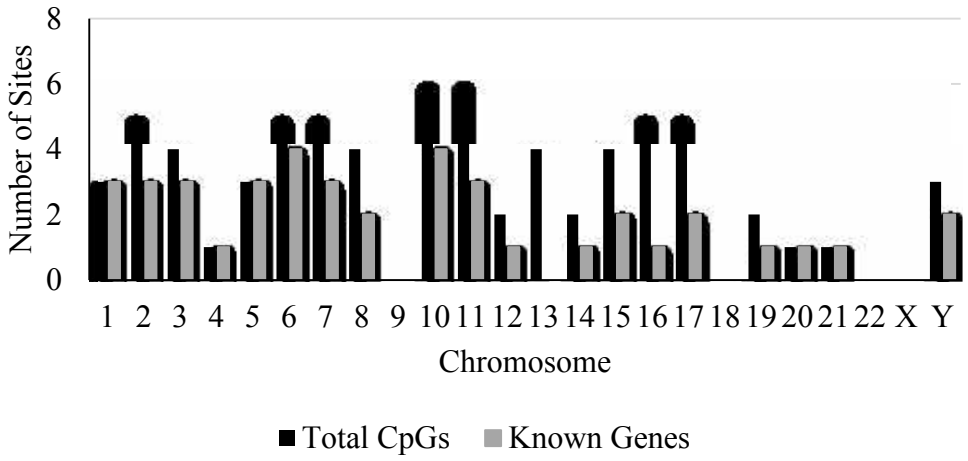
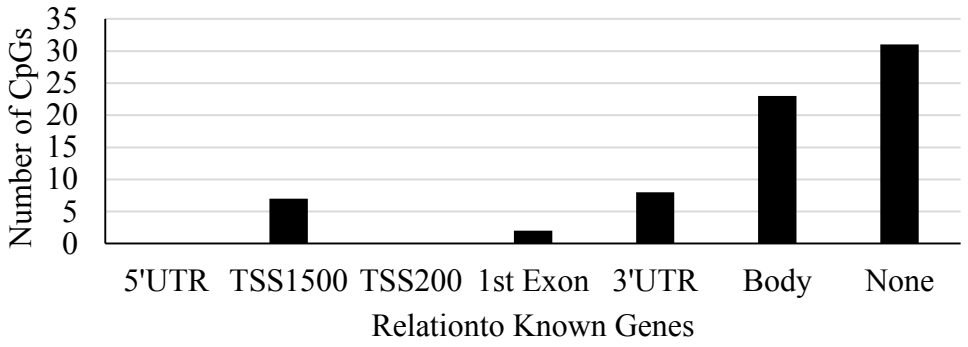


Figure 9. DNA Methylation Gain Distributions with GBS (R).

a) Chromosomal Distribution



a) Proximity to Gene



b) Proximity to CpG Island

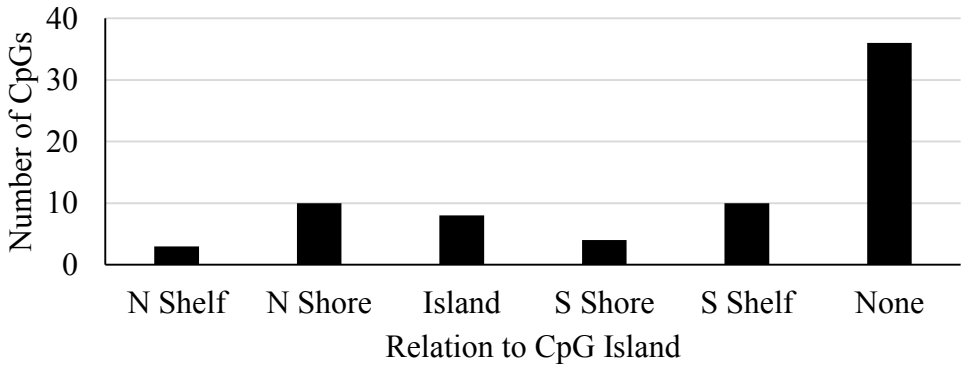


Figure 10. DNA Methylation Loss with GBS (R).

Gene Comparison

Using two different statistical software applications to identify significant differentially methylated CpG sites yielded a discordant number of CpG sites, 141 using GenomeStudio software and 125 using R programming environment. Of the known genes associated with the CpG sites, 52% of the genes identified were common in the results using both packages, 29% were specific to analysis with the GenomeStudio platform and 19% were unique to analysis conducted using the R platform (Table 15). Highlighted

Table 15. Differentially Methylated Genes by Statistical Approach

Statistical Approach	Genes
	Methylation Gain
GenomeStudio	CASD1, HLA-DRB6, LGALS8, SPTBN4, TUBAL3, ANAPC2, NAT14, OCA2, SYT8, TNNT3, ZNF628
R	AHRR, BAZ2B, C1orf192, CLECL1, HRNBP3, RASA3, RBMY1F, RBMY2FP, SPINK5, TIAL1, TTC22, TUBB8
Both Approaches	ATP8B3, BTNL9, C21orf29, C3orf50, CCS, CNST, COPB1, DCAF11, INSC, KRTAP12-3, KRTAP12-4, LMX1B, MAGI2, MIB2, MRI1, NFIC, PIWIL2, RELN, RHPN1, SNX26, STAG3L4, TAGLN3, TAS2R60, TP73, UST
	Methylation Loss
GenomeStudio	CLPTM1, FAM120B, FAM69B, KIAA1199, MTUS2, TIMP2, VIPR2, ZNF137, BMP8B, FLJ37201, FLJ43860, FOXK2, JRK, KCNH6, MRGPRX2, PPIE, SND1, SULF2, ZNF490, ZNF665
R	ADORA3, BRMS1, CCDC50, MYO10, NCRNA00052, OPRM1, RIN1, TTTY12, TTTY18
Both Approaches	ACTN3, AKR1C2, ANXA2, ARID1B, B4GALNT3, C2orf69, CAMK1D, COL11A2, CUL3, DEFB128, DMBX1, FAM124B,FRMD4A, GAP43,GCK, JPH3, KCNK7, KRTAP27-1, LAMB1, MAPK10, MGMT, MORN1, PRKCA, PTPRN2, RAB11B, SAMD4A, SERPINF2, SLC39A14, TOP1MT, WDR36, ZMAT2

genes in the table are associated with SNPs and had been excluded from analysis in the R environment because SNPs disturb the accuracy of the Infiniuim probes as described in the methods section (Bibikova et al., 2011b; Dedeurwaerder et al., 2011).

Validation of Methylation

In order to validate the methylation identified by the Illuminia Infinium 450K array, 6 CpG sites (3 = gain of methylation; 3 = loss of methylation) identified as differentially methylated by GenomeStudio analysis were selected. Sites that were at the extremes of differential methylation, associated with known genes and that primers could be developed were used (Table 16). Six participants were selected at random (3 GBS positive; 3 GBS negative) to validate the array. Participant samples that were sent immediately for sequencing after bisulfite treatment were too poor of quality

Table 16. Primer Sequences

UCSCREFGENE	Sequence
RHPN1fwd	GGATGTATTTTTTTTAGTGGTTGG
RHPN1rev	CCTCACCCAAATAAACCCTACT
HLA-DRB6fwd	TATTTTAGGATGGATTAGGAGAAAAA
HLA-DRB6rev	CAAAAATTTATAAACACTTCAACAATAC
MRI1fwd	AATTTTTGATTTTAAGTGATTTGTT
MRI1rev	AAACTATTTCTAAACCATTTTCTACTC
ANXA2fwd	TTGAGGAAAAATAATAAAGAGTTATTAGAT
ANXA2rev	AACCTAAACAATACCATTCAAAACAA
GAP43fwd	TTTAGGTGTGTGTTTATTTTTAGGA
GAP43rev	TAACCTTATCTAATTTATCATTTTAACAAC
CUL3fwd	TAGGGGAAAATTGAGGTTATAAGAAG
CUL3rev	TCCTCCTACAATACTAAAATTACAAAC

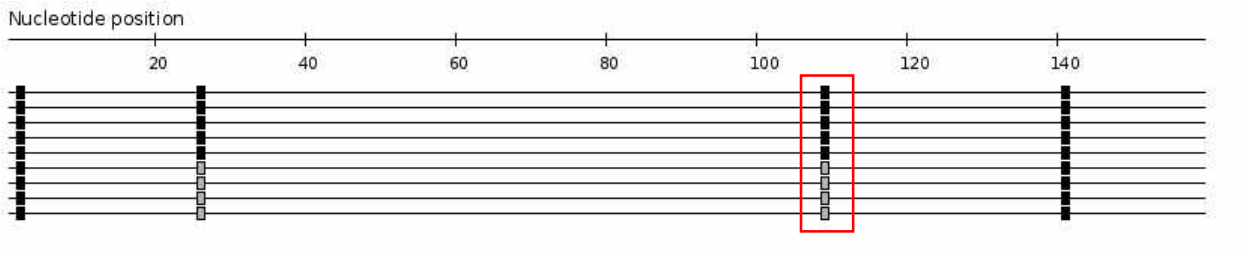
for sequencing as anticipated (Appendix G). The CpG sites associated with ANXA2 (loss of methylation in GBS positive women) and RHPN1 (gain of methylation in GBS positive women) were successfully cloned and were of sufficient quality for sequencing. The DNA used for primers designed with ANXA2 and RHPN1 CpG sites are in Figure 9 below, as the product for these sites were successfully cloned and sequenced Figure 10. In Figure 9, CpG sites within the primer products are blue text and the CpGs identified as differentially methylated by GBS status are highlighted yellow. The gray area the DNA added to identify primers, as described in the methods chapter, and the primer sequence.

ANXA2>hg19_dna range=chr15:60643907-60644407 5'pad=250 3'pad=250 strand=+ repeatMasking=none
GGCCACATTCACTTACCCAGGTTCAGGAAAGCATTTTCCAGGTCTCCTTT
AACCTCTTTCCTGATGCTTTCCAACATGTCATAAGGGCTGTAACTCTTGT
ACCTATCAAATACTGAGGAAAAACAACAAAGAGTTATCAGATCCGAGCCA
CTAGTCAAAGCTGTCAACGATCACCCACCTAGTTTTATGCACCATAATTT
TTTTAAAAATTGAGGATGATCACAGCATCCTAGGAGCTTAGAGGTTACCA
CGGTGACCAGAGCCAACATTGGCCAAGTTTGTCGTGGAACAGCCATACCA
CCTGTCCTGAATGGCACTGCCCAGGCCACATATTTGGACCATCTCTATCT
CCCCTGAGTGGAACCCATTCCATCCGAAAACCATAGGAAACAGTACAGAG
CATGCACCAAAGTCCACTACTTCAACAAATAATGGCAAGACCAAATGATC
ATCAAACAAGAAGGAGCTGCAGAATAAAGCACCAAATGCAGAAACTATTT
G

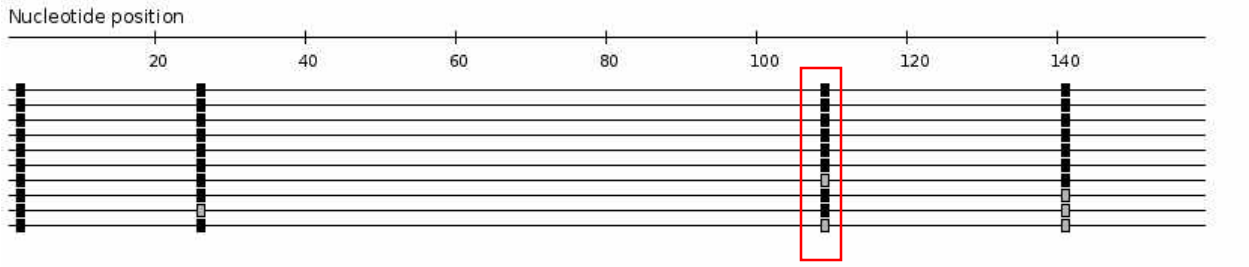
RHPN1>hg19_dna range=chr8:144457427-144457927 5'pad=250 3'pad=250 strand=+ repeatMasking=none
TCAGTCTGGCTTCTGGTGTCCTTGGCAGGTGCCAGCCTCCCCCGCTGACCCCC
ATCACGAGTCAGCAGCTTACCCCACCGACCACGTCCTTCTGCATTGACTGCC
TCCTGTCCTGCTCTGGCCAGGCCTGTGTTCACACTAGTTCTGTCCAGCCC
CTCCCTGTGAGGCCAGCTCCAGCCCCAGCGCATGGTGACCATCCCGTTAC
CCATGGGCAGGATGCACTCCTCGGCTGGCGAGGCGCAGCCTGGTG
CGGGCGCCACGGGGTCGGGCTGTGATCGCCTGTGGCCTCCCTGCAGGGCT
GTGACTCCCTGACGCAGATCCAGTGCGGCCAGCTGCAGAGCCGCAGGGCC
CAGATTCACCAGCAGATTGACAAGGAGCTGCAGATGCGGACGGGCGCTGA
GAACCTCTACAGGTCAGTGCTTGAGACTGCCCGGCCCCGGGAGCAGGGCC
CACCTGGGTGAGGGGGGCAGGACAGCCACGCAGGCAGATGTCTGCCCCAT
G

Figure 9. ANXA2 and RHPN1 DNA, Primer, and Product Sequences

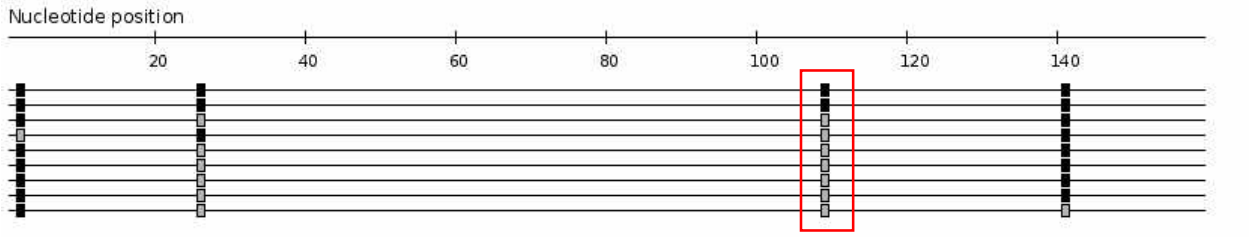
ANXA2 Participant 10 GBS+ beta value Raw = 0.5709; GS = 0.5638; R= 0.5969



ANXA2 Participant 2 GBS- beta value Raw = 0.8634; GS = 0.8490; R = 0.9388



ANXA2 Participant 14 GBS+ beta value Raw = 0.5409; GS = 0.5344; R = 0.5589



ANXA2 Participant 6 GBS – beta value Raw = 0.8833; GS = 0.8723; R = 9471

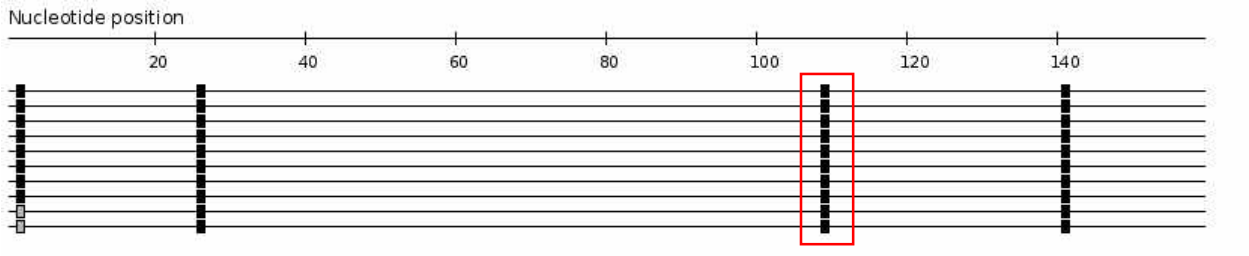
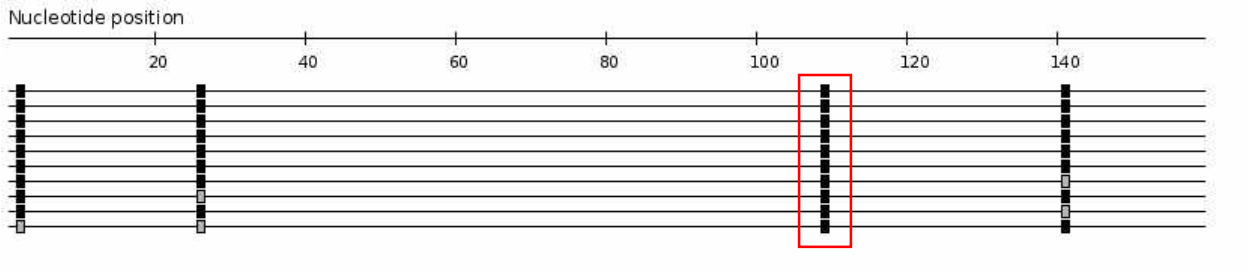
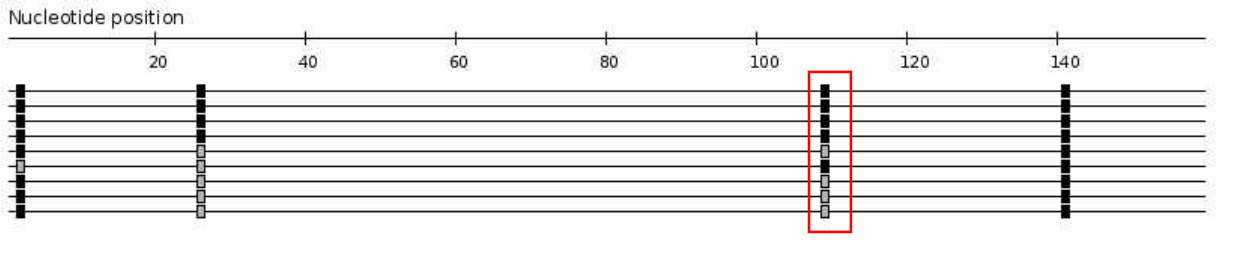


Figure 10. Methylation at ANXA2 and RHPN1 CpG Sites of Cloned Sequences

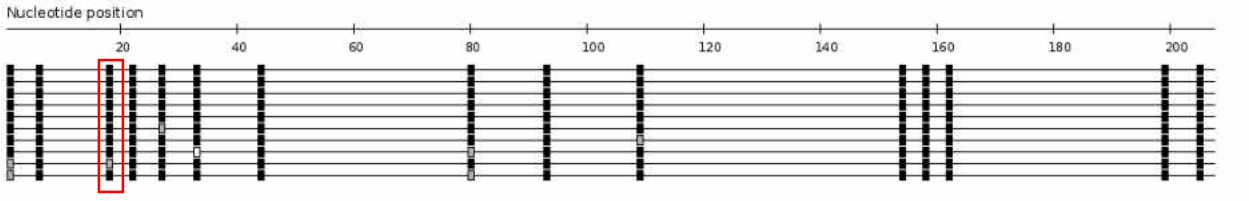
ANXA2 Participant 8 GBS- beta value Raw = 0.8901; GS = 0.8768; R = 0.9539



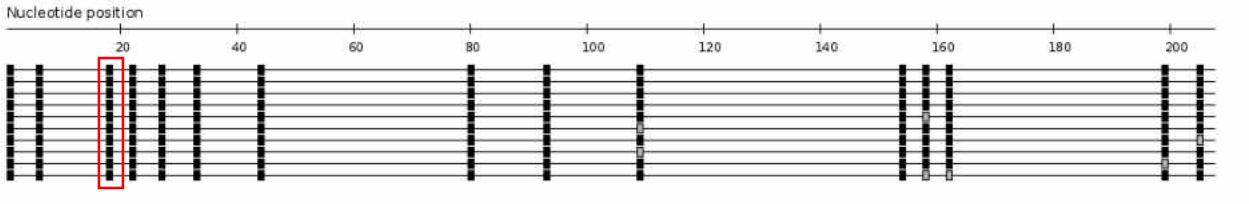
ANXA2 Participant 16 GBS + beta value Raw = 0.4998; GS = 0.4939; R = 0.4997



RHPN1 Participant 10 GBS+ beta value Raw = 0.8229; GS = 0.8186; R = 0.8920



RHPN1 Participant 2 GBS- beta value Raw = 0.4076; GS = 0.4058; R = 0.3845



RHPN1 Participant 14 GBS+ beta value Raw = 0.8049; GS = 0.8005; R = 0.8855

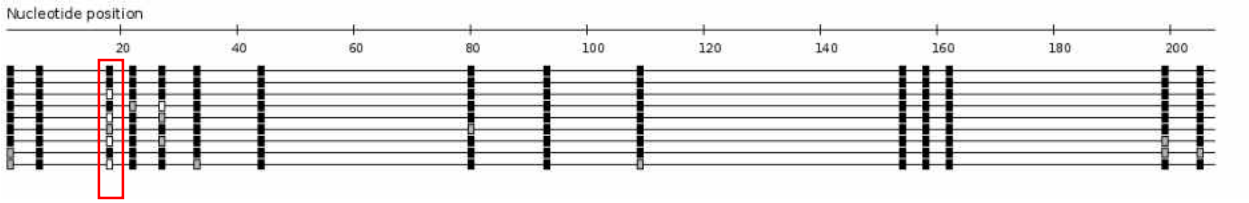
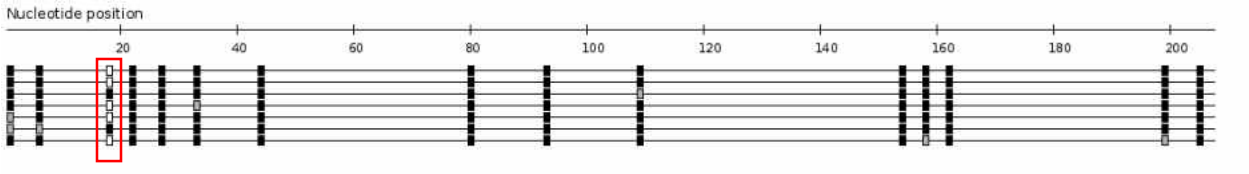
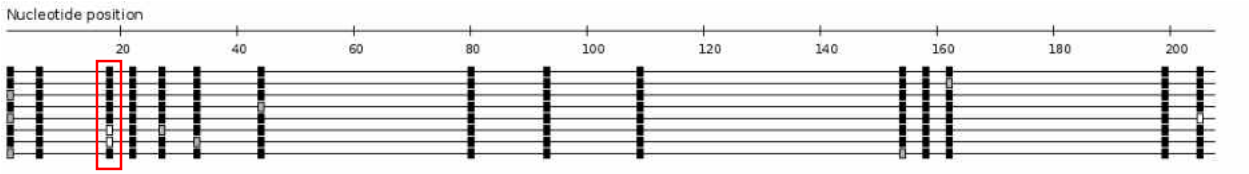


Figure 10. Cont.

RHPN1 Participant 8 GBS- beta value Raw = 0.3858; GS = 0.3842; R = 0.3629



RHPN1 Participant 6 GBS – beta value Raw = 0.3915; GS = 0.4073; R = 0.3684



RHPN1 Participant 16 GBS + beta value Raw = 0.7804; GS = 0.7764; R = 0.8620

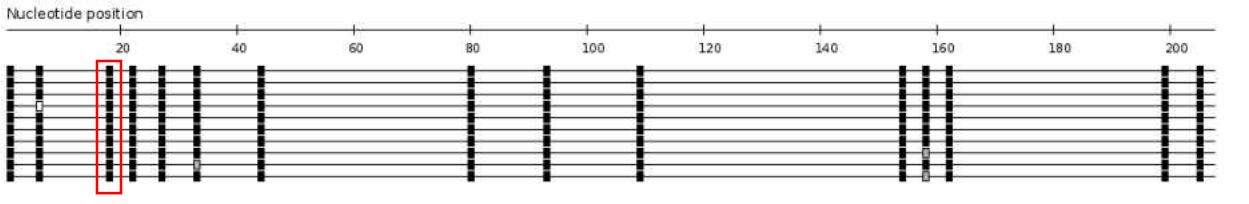


Figure 10. Cont.

In Figure 10, the CpC sites that were identified as having significantly different methylation have a red box around the CpG site in the DNA product sequence. Each figure are cloned sequences for one participant. The notation above the each participant figure indicates the primer by name, participant, GBS status, raw Infinum beta value, and normalized betas calculated by GenomeStudio software and the R programming environment. Black squares indicate the CpG site was methylated, gray squares indicated an unmethylated CpG and white squares indicate it could not be determined. The percent of cloned CpG site methylation was manually calculated by dividing the number of methylated CpGs for the specific site by total number of CpGs that methylation could be determined for the site for each of the participant clones (Table 16). Methylation percent

Table 16. Percent Methylation at Cloned CpG sites

				ANXA2				RHPN1			
Participant	**Sentrix_ID**	**Position**	**GBS**	**Raw beta**	**GS beta**	**R beta**	**Clones**	**Raw beta**	**GS beta**	**R beta**	**Clones**
1	5806636022	R01C02	Negative	0.5228	0.5133	0.5339		0.8029	0.7975	0.8897	
2	5806636023	R01C02	Negative	0.8634	0.8490	0.9388	0.8000	0.4076	0.4058	0.3845	1.0000
3	6229017070	R02C02	Negative	0.9026	0.8906	0.9599		0.4405	0.0439	0.4194	
4	6229017070	R04C02	Negative	0.9072	0.8966	0.9539		0.1235	0.1231	0.0750	
5	9379082114	R03C01	Negative	0.0766	0.0755	0.0473		0.4089	0.4073	0.3906	
6	9379082114	R05C01	Negative	0.8833	0.8723	0.9471	1.0000	0.3915	0.4073	0.3684	1.0000
7	9379082114	R01C02	Negative	0.5122	0.5034	0.5184		0.0621	0.0619	0.0400	
8	9379082114	R02C02	Negative	0.8901	0.8768	0.9539	1.0000	0.3858	0.3842	0.3629	1.0000
9	9379082114	R03C02	Negative	0.9273	0.9142	0.9711		0.4361	0.4344	0.4219	
	Average beta GBS negative			**0.7206**	**0.7102**	**0.7583**		**0.3843**	**0.3406**	**0.3725**	
10	5806636023	R03C01	Positive	0.5709	0.5638	0.5969	0.5556	0.8229	0.8186	0.8920	0.9000
11	6229017070	R01C02	Positive	0.8768	0.8632	0.9458		0.8318	0.8267	0.9111	
12	9379082114	R01C01	Positive	0.0718	0.0708	0.0491		0.7830	0.7779	0.8781	
13	9379082114	R02C01	Positive	0.0460	0.0454	0.0244		0.3940	0.3924	0.3727	
14	9379082114	R04C01	Positive	0.5409	0.5344	0.5589	0.2222	0.8049	0.8005	0.8855	0.8000
15	9379082114	R06C01	Positive	0.0876	0.0862	0.0497		0.8336	0.8299	0.9032	
16	9379082114	R04C02	Positive	0.4998	0.4939	0.4997	0.5556	0.7804	0.7764	0.8620	1.0000
17	9379082114	R05C02	Positive	0.0985	0.0973	0.0632		0.7999	0.7960	0.8797	
18	9379082114	R06C02	Positive	0.5116	0.5053	0.5163		0.8388	0.8352	0.9102	
	Average beta GBS positive			**0.3671**	**0.3623**	**0.3671**		**0.7655**	**0.7615**	**0.8327**	
	Delta beta			**-0.3535**	**-0.3479**	**-0.3912**		**0.3812**	**0.4209**	**0.4602**	

GS = Genome Studio; R = R statistical environment

in the cloned sequences was congruent with percent methylation determined by the Illumia Infinium array (Table 16). Participants that had CpG sites where methylation status could not be determined (RHPN1 participant 6, 2 undetermined; RHPN1 Patrticpant 8, 5 undetermined; RHPN1 particpant 14, 4 undetermined) did not correlated as strongly as other sites. CpG sites that were identified as significantly different by only one of the software packages had similar beta values, and greater delta beta values did not correlate with significance. Table 17 presents beta results for all participants for one site

Table 17. Beta Value Similarity of Significant Results

	Genome Studio			R environment		
	SND1			RASA3		
Participant	**Raw beta**	**GS beta**	**R beta**	**Raw beta**	**GS beta**	**R beta**
1	0.4466	0.4435	0.4343	0.3979	0.3958	0.3750
2	0.0689	0.0683	0.0365	0.8440	0.8376	0.9241
3	0.8832	0.8766	0.9710	0.4294	0.4273	0.4045
4	0.8767	0.8711	0.9312	0.4414	0.4397	0.4251
5	0.4216	0.4193	0.4058	0.8268	0.8219	0.9052
6	0.8424	0.8372	0.9160	0.4147	0.4130	0.3963
7	0.8430	0.8358	0.9268	0.4256	0.4233	0.4131
8	0.8263	0.8196	0.9055	0.8222	0.8165	0.9019
9	0.8606	0.8548	0.9252	0.8553	0.8501	0.9209
Average Negaive	**0.6744**	**0.6696**	**0.7169**	**0.6064**	**0.6028**	**0.6296**
10	0.4428	0.4405	0.4270	0.8536	0.8485	0.9186
11	0.0519	0.0516	0.0320	0.8801	0.8735	0.9480
12	0.8383	0.8312	0.9263	0.8177	0.8119	0.9097
13	0.0532	0.0529	0.0293	0.8271	0.8216	0.9060
14	0.4477	0.4453	0.4365	0.8244	0.8197	0.9031
15	0.8578	0.8527	0.9235	0.8430	0.8388	0.9113
16	0.0532	0.0529	0.0293	0.8362	0.8315	0.9134
17	0.8666	0.8608	0.9362	0.8343	0.8299	0.9106
18	0.0688	0.0685	0.0373	0.4285	0.4269	0.4111
Average Positive	**0.4089**	**0.4063**	**0.4197**	**0.7939**	**0.7891**	**0.8591**
Delta beta	**-0.2655**	**-0.2633**	**-0.2972**	**0.1878**	**0.1863**	**0.2295**

GS= GenomeStudio; R = R statistical environment

Identified as significant using GenomeStudio software (SND1) or the R environment (RASA3). The delta beta for both sites is higher after the R normalization, however only the RASA3 site was identified using the R analysis pipeline as significantly different and SND1 was not. The SND1 site was identified as being significantly different by Genome studio, and RASA3 was not. This type of incongruence is similar at all discordant significant results.

Functional Analysis of Methylation

Research Question 3: Is there a relationship between differentially methylated genes and immune function in pregnant women colonized with GBS?

Since no CpG sites were identified as differentially methylated when FDR is applied to determine significance, results produced using the less stringent alpha of 0.05 for significance were used for functional analysis. Cluster analysis using DAVID revealed differential methylation in women with GBS is primarily related to basic cellular processes. DAVID analysis was performed as described in the methods section independently for DNA methylation results from GenomeStudio (Table 16), R (Table 17), and genes common to both analytical approaches (Table 18). However, none of the functional clusters were significant using FDR of 0.05 and the cell morphogenesis cluster was the only functional cluster that was significant at an alpha level of 0.05 for data output from GenomeStudio, R and for common genes identified by both approaches. The results of the functional analysis using the data generated with this sample yielded different functional clusters from the preliminary data analysis used when designing and conceptualizing this study that clustered more specifically to immune function such as major histocompatibility complex and antigen presentation (Appendix H).

Table 16. Biological Pathways Associated with GBS (GenomeStudio data)

Functional Cluster	Enrichment Score	Genes	*p-value*	FDR
		Methylation Gain		
Keratin	1.26	KRTAP12-3, KRTAP12-4, ZNF628	**0.04**	0.98
Cellular metabolic and biosynthetic processes	1.15	LMX1B, ANAPC2, PIWIL2, NFIC, TP73	**0.04**	0.97
Protein kinase and phosphorylation	0.87	ANAPC2, RELN, TP73	0.10	0.99
Regulation of transcription	0.67	LMX1B, NFIC, TP73	0.17	1.00
Cell cycle	0.64	ANAPC2, PIWIL2, TP73	0.17	1.00
Cytoplasmic membrane-bounded vesicle	0.53	ATP8B3, COPB1, SYT8	0.26	1.00
Transcription factor activity	0.42	LMX1B, NAT14, NFIC, TP73, ZNF628	0.22	1.00
Ion binding	0.31	ATP8B3, LMX1B, CCS, MIB2, RELN, TP73, ZNF628	0.40	1.00
Membrane	0.25	ATP8B3, CASD1, NAT14, BTNL9, CNST, COPB1, PIWIL2, MAGI2, OCA2,	0.53	1.00
		Methylation loss		
Extracellular matrix	1.58	TIMP2, ANXA2, COL11A2, LAMB1	**0.02**	0.96
Cell morphogenesis	0.84	CUL3, GAP43, LAMB1, PRKCA	**0.03**	1.00
Regulation of apoptosis	0.75	MGMT, ACTN3, CAMK1D, CUL3, PRKCA	0.18	1.00

Table 16. Cont.

Functional Cluster	Enrichment Score	Genes	*p-value*	FDR
Cytoplasmic membrane-bounded vesicle	0.57	RAB11B, ANXA2, SERPINF2, SND1	0.22	0.98
Phosphorus metabolic process	0.56	CAMK1D, GCK, MAPK10, PRKCA, PTPRN2	0.27	1.00
Protein dimerization activity	0.56	MTUS2, ACTN3, COL11A2, DMBX1	0.18	1.00
Ion transport	0.53	JPH3, KCNK7, KCNH6, SLC39A14	0.13	1.00
Cellular homeostasis	0.46	COL11A2, GCK, JPH3, PRKCA	0.28	1.00
Transcription regulation	0.39	ARID1B, DMBX1, FAM120B, FOXK2, SND1, ZNF490, ZNF665	0.36	0.98
Ion binding	0.34	MGMT, ACTN3, ANXA2, CAMK1D, FOXK2, KCNK7, KCNH6, PRKCA, SLC39A14, SULF2, ZNF490, ZNF665, ZMAT2	0.44	1.00
Nucleotide binding	0.33	RAB11B, CAMK1D, GCK, MAPK10, PPIE, PRKCA, TOP1MT	0.58	1.00
Regulation metabolic and biosynthetic processes	0.28	ARID1B, GCK, PRKCA, SAMD4A	0.42	1.00
Biological adhesion	0.25	ACTN3, COL11A2, LAMB1	0.52	1.00
Zinc ion binding	0.18	MGMT, TIMP2, PRKCA, SLC39A14, ZNF490, ZNF665, ZMAT2	0.43	0.97

Table 16. Cont.

Functional Cluster	Enrichment Score	Genes	*p-value*	FDR
Membrane-enclosed lumen	0.16	ARID1B, MGMT, DMBX1, GCK, SERPINF2, TOP1MT	0.65	1.00
Zinc finger, C2H2-like	0.14	ZNF490, ZNF665, ZMAT2	0.62	1.00
Transmembrane	0.07	MRGPRX2, B4GALNT3, CLPTM1, FAM69B, JPH3, KCNK7, KCNH6, PTPRN2, SLC39A14, VIPR2	0.87	1.00

Table 17. Biological Pathways Associated with GBS (R data)

Functional Cluster	Enrichment Score	Genes	*p-value*	FDR
		Methylation Gain		
RNA recognition motif, RNP-1	1.19	RBMY1F, TIAL1, HRNBP3	0.06	1.00
Cellular metabolic and biosynthetic processes	0.97	LMX1B, PIWIL2, NFIC, TP73	0.08	1.00
Induction of apoptosis	0.71	TIAL1, AHRR, TP73	0.09	1.00
Metal-binding	0.65	ATP8B3, LMX1B, RASA3, BAZ2B, CCS, MIB2, RELN, TP73	0.17	1.00
Regulation of transcription	0.64	LMX1B, TIAL1, AHRR, BAZ2B, NFIC, TP73	0.11	1.00
Cytoplasmic membrane-bounded vesicle	0.61	ATP8B3, COPB1, SPINK5	0.21	1.00
Transcription factor activity	0.60	LMX1B, NFIC, TP73	0.18	0.99
Ion binding	0.44	ATP8B3, LMX1B, RASA3, AHRR, BAZ2B, CCS, MIB2, RELN, TP73	0.43	1.00

Table 17. Cont.

Functional Cluster	Enrichment Score	Genes	*p-value*	FDR
Transcription regulation	0.29	LMX1B, TIAL1, AHRR, BAZ2B, NFIC, TP73	0.42	1.00
Intrinsic to membrane	0.06	ATP8B3, CLECL1, RASA3, BTNL9, CNST, PIWIL2, TAS2R60, UST	0.83	1.00
		Methylation loss		
Cell morphogenesis	1.63	CUL3, GAP43, LAMB1, PRKCA	**0.01**	1.00
Neuron development	0.85	GAP43, LAMB1, PRKCA	0.07	1.00
Phosphorus metabolic process	0.84	CAMK1D, GCK, MAPK10, PRKCA, PTPRN2	0.11	1.00
Nucleotide binding	0.73	RAB11B, CAMK1D, GCK, MAPK10, MYO10, PRKCA, TOP1MT	0.10	0.79
Serine/threonine protein kinase	0.68	CAMK1D, MAPK10,PRKCA	0.09	1.00
Cellular homeostasis	0.67	GCK, JPH3, PRKCA	0.16	1.00
G-protein coupled receptor protein signaling pathway	0.64	ADORA3, GAP43, OPRM1	0.65	1.00
Ion transport	0.52	JPH3, KCNK7, SLC39A14	0.22	1.00
Regulation biosynthetic processes	0.42	ARID1B, GCK, SAMD4A	0.36	1.00
Biological adhesion	0.41	ACTN3, COL11A2, LAMB1	0.39	1.00
Cytoplasmic membrane-bounded vesicle	0.36	RAB11B, ANXA2, SERPINF2	0.39	0.99

Table 17. Cont.

Functional Cluster	Enrichment Score	Genes	*p-value*	FDR
Membrane-enclosed lumen	0.28	ARID1B, MGMT, DMBX1, GCK, SERPINF2, TOP1MT	0.46	0.98
Receptor	0.18	ADORA3, OPRM1, PTPRN2	0.82	1.00
Ion binding	0.09	MGMT, ACTN3, ANXA2, CAMK1D, KCNK7, PRKCA, SLC39A14, ZMAT2	0.81	1.00
Zinc ion binding	0.06	MGMT, PRKCA, SLC39A14, ZMAT2	0.81	1.00
Integral to membrane	0.06	ADORA3, B4GALNT3, JPH3, NCRNA00052, OPRM1, KCNK7, PTPRN2, SLC39A14	0.99	1.00

Table 18. Biological Pathways Common between R and GenomeStudio

Functional Cluster	Enrichment Score	Genes	*p-value*	FDR
		Methylation Gain		
Cellular metabolic and biosynthetic processes	1.46	LMX1B, PIWIL2, NFIC, TP73	**0.03**	1.00
Metal-binding	0.51	ATP8B3, LMX1B, CCS, MIB2, RELN, TP73	0.28	0.99
Regulation of transcription	0.50	LMX1B, NFIC, TP73	0.08	0.99
Intrinsic to membrane	0.06	ATP8B3, BTNL9, CNST, PIWIL2, TAS2R60, UST	0.74	1.00
		Methylation loss		
Cell morphogenesis	1.77	CUL3, GAP43, LAMB1, PRKCA	**0.01**	0.97
Regulation of apoptosis	1.34	MGMT, ACTN3, CAMK1D, CUL3, PRKCA	0.05	0.98
Extracellular matrix	1.14	ANXA2, COL11A2, LAMB1	0.05	0.98

Table 18. Cont.

Functional Cluster	Enrichment Score	Genes	*p-value*	FDR
Neuron development	0.94	GAP43, LAMB1, PRKCA	0.07	1.00
Serine/threonine protein kinase	0.79	CAMK1D, MAPK10,PRKCA	0.07	1.00
Cellular homeostasis	0.75	GCK, JPH3, PRKCA	0.14	0.99
Phosphorus metabolic process	0.74	RAB11B, CAMK1D, GCK, MAPK10, PRKCA, PTPRN2	0.12	0.83
Metal ion transport	0.59	JPH3, KCNK7, SLC39A14	0.19	0.99
Cellular biosynthetic processes	0.49	ARID1B, GCK, SAMD4A	0.31	1.00
Biological adhesion	0.49	ACTN3, COL11A2, LAMB1	0.31	0.98
Cytoplasmic membrane-bounded vesicle	0.46	RAB11B, ANXA2, SERPINF2	0.31	0.98
Membrane-enclosed lumen	0.43	ARID1B, MGMT, DMBX1, GCK, SERPINF2, TOP1MT	0.31	0.99
Ion binding	0.19	MGMT, ACTN3, ANXA2, CAMK1D, KCNK7, PRKCA, SLC39A14, ZMAT2	0.64	1.00
Zinc ion binding	0.10	MGMT, PRKCA, SLC39A14, ZMAT2	0.68	1.00
Integral to membrane	0.02	B4GALNT3, JPH3, KCNK7, PTPRN2, SLC39A14	0.80	1.00

Summary of Results

In summary, the results from this study indicate that there are no significant differences in DNA methylation between women with and without GBS colonization when FDR is used to determine significance of DNA methylation differences that are

greater than 20% between groups. However, if a less stringent p-value of 0.05 is used there are a small number of CpG sites that have significant differences with greater than a 20% difference in methylation. The number of significantly different CpG sites identified using different software for analysis varies (141 versus 125). No significant association was found between serum TNF-α, IL-6, IL-10 or vitamin D (25[OH]D) levels and maternal GBS colonization status. Lastly, analysis of functional pathways did not find a correlation between differentially methylated genes and genes directly related to cytokine production or specific immune pathways. No functional clusters were significant when applying FDR for significance; and only the cell morphology functional cluster was significant for output from both GenomeStudio and R data when an alpha of 0.05 was applied.

CHAPTER V

DISCUSSION

GBS sepsis continues to be the leading cause of infectious neonatal morbidity and mortality despite current practice guidelines to prevent the transmission of GBS from mothers to their infants (Phares et al., 2008). The primary purpose of this exploratory study was to identify variants in maternal blood that are associated with maternal GBS colonization in order to assist with the development of more accurate screening tools and/or assist in identifying to targets to prevent maternal GBS colonization. It is unknown why GBS selectively colonizes one third of pregnant women, placing the health of women and their offspring at risk. Currently in the US, all pregnant women are screened for GBS colonization between 35-37 weeks gestation (Verani et al., 2010). However, significant false negative screening results and infant illness despite maternal antibiotic treatment during labor requires further investigation to identify a biological reasons as to why certain women are preferentially colonized with GBS (Lin et al., 2011; Towers et al., 2010). For this study, variants that can be measure in the serum that are increasingly used for monitoring and diagnosing other clinical conditions, were investigated to determine if there was any association between serum levels and maternal GBS colonization status. This contribution to science is significant because the relationship between DNA methylation patterns, TNF-α, IL-6, IL-10, and vitamin D (25[OH]D) with maternal GBS colonization status have not previously been reported.

Maternal Serum Cytokines and Vitamin D (25[OH]D)

Research Question 1: Are serum levels of TNF-α, IL-6, IL-10, and vitamin D (25[OH]D) different in pregnant women with GBS colonization than pregnant women without GBS colonization?

This study was designed to capitalize on data that was collected for a previous study. Other clinical indicators could potentially be used to identify pregnant women at increased risk for GBS colonization that have not previously been considered. Clinical indicators, such as serum cytokine levels that reflect immune system functioning, could also be associated with altered DNA methylation patterns. Variability in serum markers associated with immune function and vitamin D (25[OH]D) have been previously reported and utilized as a prognostic indicator of disease states, such as respiratory infections (Chesney, 2010), human immunodeficiency virus infections (Fahey et al., 1990), pancreatitis (Pezzilli et al., 1995), and depression (Kiecolt-Glaser & Glaser, 2002). Research investigating the role of vitamin D (25[OH]D) as an immune function modulator has increased dramatically in recent years (Figure 3) and could offer a cost effective intervention target if low serum vitamin D (25[OH]D) levels are associated with GBS colonization.

Multiple alterations in immune function are necessary during pregnancy to prevent the mother's body from perceiving the developing fetus as a foreign pathogen. Serum markers of immune function are being increasingly evaluated in pregnancy because alterations occur throughout normal pregnancy and inappropriate levels contribute to the development of pathology during pregnancy (Ponsonby et al., 2010). In a recent study, Fichorova et al. (2011) identified patterns in immune function markers

that were specific to the type of bacteria present in the vaginal mucosa and placenta of pregnant women. Notably, they found TNF-α, IL-1β, IL-6, IL-8 and ICAM-1 were elevated when pathogenic organisms associated with bacterial vaginosis were present. Furthermore, *Lactobacillus*, which colonize the vaginal mucosa and are not pathogenic, suppress pathogenic strains and downregulate pro-inflammatory cytokines (Donato et al., 2010; Othman et al., 2007; Zeuthen et al., 2010). However, no evidence could be found in the literature to suggest that patterns of immune function serum markers have been explored with regards to GBS colonization or infection. Additionally vitamin D (25[OH]D), a known modulator of immune function and deficiencies, has been associated with susceptibility to infectious diseases (Chesney, 2010). It is generally accepted that serum vitamin D (25[OH]D) levels must be above 20 ng/ml during pregnancy to maintain normal physiologic processes and fetal development. Recently, experts have noted vitamin D (25[OH]D) levels greater than 32 ng/ml are necessary to support all physiologic processes (e.g. infection prevention) that require vitamin D (25[OH]D) for optimal functioning (ACOG, 2011; Holick et. al., 2011). For this study, TNF-α, IL-6, IL-10 and vitamin D (25[OH]D) were selected specifically for analysis because of the reported association with infectious diseases and action during pregnancy. The serum cytokine and vitamin D (25[OH]D) levels were evaluated throughout the pregnancy in order to identify any difference in how the levels change throughout pregnancy, as well as independently during each trimester. Laboratory results were also evaluated to identify any direct correlation between serum cytokine and vitamin D (25[OH]D) levels.

First, correlation tests indicated that vitamin D (25[OH]D) levels were not correlated with serum TNF-α, IL-6, or IL-10 levels during pregnancy. The initial hypothesis of the study was that vitamin D (25[OH]D) would be correlated with TNF-α, IL-6, or IL-10 levels because vitamin D (25[OH]D) modulates immune function. Because there was no correlation between the cytokines and vitamin D (25[OH]D); vitamin D (25[OH]D) was not used as a covariate and repeated measures ANOVA was completed using serum vitamin D (25[OH]D), TNF-α, IL-6 and IL-10 independently instead of the originally planned repeated measures ANCOVA (with vitamin D (25[OH]D) as the covariate). In addition to there being no correlation between serum vitamin D (25[OH]D) levels and the serum cytokines, no significant association between serum TNF-α or IL-6 levels and maternal GBS colonization status was identified. In vitro experiments have indicated there is an increase in TNF-α production in neonatal and adult peripheral mononuclear cells (Berner et al., 2002) and human epithelial cells (Mikamo et al., 2004) exposed to GBS. Additionally, TNF-α and IL-6 levels increase systemically in murine models when the mice are inoculated with GBS resulting in high mortality rates (Puliti et al., 2000). Because TNF-α levels increase in laboratory and animal studies as a result of GBS exposure, this investigator hypothesized that TNF-α levels may be elevated in pregnant women who are GBS positive. In this study, TNF-α levels were only elevated in pregnant women with GBS during the third trimester and the difference was not statistically significant. Furthermore, results from the one-way MANOVA indicated TNF-α has no effect (partial $\eta^2 < 0.001$) on GBS colonization status in pregnant women. Given this very low η^2, it is unlikely that any significant difference would be identified in

TNF-α values in pregnant women with and without GBS colonization; even with a larger sample size.

There is an increased IL-6 production in vitro, and in animal models, with exposure to GBS (Berner et al., 2002; Mikamo et al., 2004; Puliti et al., 2002). The increase in IL-6 in response to GBS exposure has not been verified or validated in human studies. This study evaluated the level of serum IL-6 to determine if a similar increase in IL-6 production occurs in pregnant women in response to GBS colonization and found no statically significant effect. Results from the one-way MANOVA indicated IL-6 has no effect (partial $\eta^2 < 0.001$) on GBS colonization status in pregnant women. It is unlikely that even with a larger sample size any significant difference would be identified in IL-6 values in pregnant women with and without GBS colonization.

Two previous studies in murine models identified an increase in IL-10 levels related to GBS exposure or infection (Bebien et al., 2012; Madureira et al., 2011). No statistically significant elevation in IL-10 levels in pregnant women who were GBS positive were observed at any point during pregnancy. Because of the small sample size, we calculated the mean serum level values and conduct a one-way MANOVA to correct for confounding resulting from data being collected over time, determine effect size and observed power. Because we found a negligible effect (partial $\eta^2 = 0.006$) of IL-10 on GBS colonization status, it may be warranted to repeat the study with a larger sample size.

Serum vitamin D (25[OH]D) levels had a small effect on GBS colonization status. The normal increase in serum vitamin D (25[OH]D) levels during pregnancy that may be involved in preventing infection and colonization with pathogenic bacteria during normal

pregnancy was not evident in this study (De-Regil Luz et al., 2012). Surprisingly, multivariate analysis revealed that GBS positive women are more likely to have higher serum vitamin D (25[OH]D) levels. However, both GBS positive (29.16 ng/ml) and GBS negative women (24.59 ng/ml) had serum vitamin D (25[OH]D) levels below the 32 ng/ml recommended for optimal physiologic functioning during pregnancy. Both groups of women likely had low vitamin D (25[OH]D) levels because of the fact that the primary study was undertaken in the high northern latitude. Unfortunately, previous studies evaluating vitamin D (25[OH]D) levels related to GBS colonization status could not be found in the literature. Therefore, it is unclear how these results should be interpreted. A meta-analysis conducted by Thorne-Lyman and Fawzi (2012) revealed it is unknown how vitamin D (25[OH]D) relates to maternal infections since the relationship between vitamin D (25[OH]D) and immunity has only recently been established. Vitamin D (25[OH]D) levels across pregnancy will be evaluated in a larger cohort of women in an upcoming follow-up study to determine if the same results persist. Repeating the analysis with a larger cohort may result in better understanding of the significance of these study results and how vitamin D (25[OH]D) is related to GBS status during pregnancy. Post-hoc power analysis indicated that replicating the study with a larger sample size may yield significant differences in vitamin D (25[OH]D) levels associated with GBS colonization status because serum vitamin D (25[OH]D) levels appear to have some effect on GBS colonization status (partial $\eta^2 = 0.068$).

Maternal DNA Methylation

Research Question 2: Are DNA methylation patterns different between pregnant women with GBS colonization and those without GBS colonization?

DNA methylation is currently the most well understood epigenetic modification and is increasingly being integrated into clinical nursing research. In this study, the investigator used previously collected peripheral blood samples in order to complete the study in a timely and cost effective manner. The exploratory nature of this genome wide DNA methylation analysis allowed the investigator to quantify methylation of individual CpG sites. Specifically, DNA methylation in peripheral white blood cells that were collected during the first trimester of pregnancy in women colonized with GBS in late pregnancy were compared to women who screened negative for GBS in the third trimester of pregnancy. Preliminary analysis of a subset of individuals by the investigator (n=6), initially indicated that over 1,000 potential early pregnancy DNA methylation differences existed between women with and without late pregnancy GBS colonization. Since DNA methylation is an epigenetic modification that can result in altered gene expression, with the potential to impact health and disease susceptibility, this study was designed to see if the preliminary differences identified in a very small sample (2 GBS positive, 4 GBS negative) would persist in a larger sample. Previous studies have identified potentially useful, clinically relevant DNA methylation biomarkers for preeclampsia using a sample size of n=6/group (Anderson et al., 2013). For this study, the sample size was increased to n=9/group, which reduced the number of statistically significant differentially methylated sites from over 1,000 genes between women with and without GBS colonization to 141 CpG sites using the GenomeStudio software. Due to the small number of significant genes and the drastic decrease in potentially different CpG sites, it is premature to assume these sites may be an early biomarker for GBS colonization. Support for this assessment will be evident in the discussion of the variation

in the number and actual genes associated with differentially methylated CpG sites using different software for statistical analysis. Prior to this study, no other papers could be identified in the literature investigating host DNA methylation patterns related to colonization with a certain microbe. As discussed in chapter 2, this is an area of inquiry that will likely increase over time because of the role that methylation plays with cell differentiation and memory in immune cells (Bobetsis et al., 2007; Torsten Olszak et al., 2012; Schaub et al., 2009; Tolg et al., 2011).

Of the 18 women included in the methylation analysis for this study, data extracted from the medical record indicated no difference in baseline characteristics between GBS positive and GBS negative women. Further, no significant differences in maternal co-morbidities that may indicate, or cause, altered immune function; such as asthma, infections or antimicrobial usage were found. Because some of the normalizations features incorporated into the analytic component of GenomeStudio software are proprietary, it has limited functionality in assessing and presenting differentially methylated data and may limit reproducibility (Gentleman et al., 2004; Smyth, 2005). Data analysis pipelines for biology and bioinformatics were design because "the primary motivations for an open-source computing environment for statistical genomics are transparency, pursuit of reproducibility and efficiency of development" (Gentleman et al., 2004, p. R80.2). Therefore, statistical analyses were also completed using the R statistical environment to ensure accurate normalization and interpretation and reproducibility of analysis of the raw data. Additionally, Hansen et al. (2013) noted that it was unclear as to what process is used for normalization by GenomeStudio because it is not explicitly publicized. Due to the lack of transparency of

normalization methods used in GenomeStudio analytic software that limits reproducibility of analysis, biostatisticians and R developers recommend not using GenomeStudio to conduct statistical analysis. However, R and GenomeStudio are both used extensively to conduct methylation analysis. Because both programs are frequently used, the results produced using GenomeStudio and R to conduct statistical analysis were included in the results section. Some of the differences in the results are likely a direct result of different normalization methods of the raw data in each platform. Each result can be reproduced using the specified statistical software, although results differ between software packages.

For the actual statistical analysis of DNA methylation using the R statistical environment, M- values were used, instead of beta values because M-values are homoscedastic across different levels of methylation. Given that beta values exhibit severe heteroscedasticity at the methylation extremes, M-values are the superior choice for conducting valid differential and statistical analysis (Du, Kibbe, & Lin, 2008). However, M-values are difficult to interpret clinically for relevance because the transformed negative values are not biologically interpretable. Therefore, after using the logit transformed values (M values) for statistical analysis, the data were reported using the original untransformed beta values to allow for clinical interpretation of the findings. In fact, beta values are most frequently reported because beta values represent a percentage of methylation, and therefore are more biologically meaningful (Du et al., 2010).

Regardless of the statistical platform used for the data analysis, we did find over 100 CpG sites with significantly different methylation between women who were GBS

positive or GBS negative when applying an alpha of 0.05. Of the sites identified by both GenomeStudio and R statistical environment, there were 56 common genes. However, it is expected that statistically by chance 24,250 CpG sites would be identified as differentially methylated when using an alpha of 0.05 (485,000 sites in the array * 0.05 = 24,250). In other words, basic statistics indicate an arbitrary investigator is more likely to find a significantly different methylated CpG site between GBS positive and negative women, than the sites identified using an alpha of 0.05 actually being significant. Similar analytic approaches for identifying biomarkers, such as the NIMBL package for Matlab, likely also lack the power to identify sub-sample heterogeneity or reasonably identify biomarkers in a sample of this size when not using FDR to determine significance. Therefore, conducting additional testing assessing for significant in this sample would not yield valuable biomarker information because no sites are significantly different when FDR is applied to determine significance. It is imperative to generate empirical estimates of test statistics (and p-values) via bootstrapping methods for small sample sizes, coupled with FDR to appropriately safeguard against over-interpretation of microarray data (Benjamini & Hochberg, 1995; Tusher, Tibshirani, & Chu, 2001; York, 2003). Furthermore, recommendations for design and analysis of epigenome-wide association studies, such as this one, include using multiple-testing adjustments and validating the methylation in a similar but different cohort using a different laboratory methodology (Michels et al., 2013). Analysis of methylation on this study cohort using FDR indicate no CpG sites have significantly different methylation using any FDR cutoff value (adjusted *p* value= 0.99992 for all values, Appendix I). The lack of significance maybe attributed to small sample size and the use of peripheral blood instead of vaginal

epithelial cells. Michels et al. (2013) stated "As natural variation affects DNA methylation, larger sample sizes will typically be required for EWAS than for GWAS for any given phenotype, even when the most technically sophisticated assays are used" (p. 952), therefore it is premature to assume that any of the CpG sites identified in this study with 18 participants will yield reliable biomarker results regardless of approach. In future studies, the same analysis will be completed using a larger sample size of GBS positive and GBS negative women who have already had genome-wide methylation analysis completed. This investigator suspects that repeating the analysis on a larger cohort of women may further decrease the number of statistically significant differentially methylated sites identified, or result in completely different findings between the two groups of women. Such replication will also enable the investigator to determine if genes identified by only one of the approaches (GenomeStudio or R) are no longer significant. Further analysis with a larger sample or statistical simulations, such as bootstrapping that are beyond the scope of this dissertation, may be able to better identify why the results differ when using different software to perform statistical analysis.

Additional studies will also be needed to determine the significance of DNA methylation on gene expression. RNA was not collected in the parent study, and therefore gene expression studies using RNA could not be completed for this study. However, based on the location of differential DNA methylation, it is possible that the differentially methylated CpG sites associated with genes may have an effect on gene expression. CpG islands, dense regions of cytosine and guanine dinucleotides, contribute to the regulation of gene transcription and subsequent gene expression (Deaton & Bird, 2011). Approximately 72% of known gene promoter regions are associated with CpG islands

(Saxonov, Berg, & Brutlag, 2006). Results from the analysis conducted in the R environment identified 15 CpG sites that were within CpG islands and 44 CpG sites were located in the regions flanking the islands that may also result in altered gene expression patterns (Doi et al., 2009). Upon follow-up studies methylation at these sites will be assessed, since they are most likely related to gene expression. Appropriate samples will also be collected to perform RNA and protein analysis in the next study.

If a selective and specific biomarker panel for GBS colonization based on differential methylation patterns can be developed after repeating the analysis with a larger cohort, it could be useful for identifying women at risk for poor pregnancy outcomes (e.g. miscarriage, preterm birth, premature rupture of membranes and maternal infections) that occur as a result of GBS prior to 35-37 weeks of pregnancy. Since DNA methylation vitally contributes to programming memory in immune cells, altered methylation patterns in women with GBS could represent a novel target for designing novel treatment and prevention modalities.

Biologic Functions Associated with Altered Methylation

Research Question 3: Is there a relationship between differentially methylated genes and immune function in pregnant women colonized with GBS?

Although no CpG sites were differentially methylated when applying FDR for significance, functional analysis was conducted using DAVID bioinformatics data base to determine if the differentially methylated genes identified using an alpha of 0.05 may be related to immune function or inflammation. A recently published manuscript by Laayouni et al. (2014), identified alterations in 20 genes associated with immune function in populations exposed to *Yersinia pestis* that have persisted over time and resulted in

altered predisposition for autoimmune disorders in individual of European descent. Laayouni's team found variant SNPs in genes that alter how the immune system responds to *Y. pestis.* They found that the production of pro-inflammatory cytokines is increased in response to *Y. pestis*. The increased inflammatory response allowed for some of the population to be resistant to, or heal from, the Black Plague. These SNP variations are not seen in populations that were out of the endemic area during the Black Plague. Laayouni's group suggested that the SNP variations, driven by pathogenic exposure, contribute to the increased prevalence of autoimmune disorders in populations of European descent where Black Plague was endemic. The SNP variants are not present in populations not exposed to *Y. pestis* and the populations also exhibit lower prevalence of autoimmune disorders. Lauyouni's study further supports that genetic variation can be driven by, and contribute to, pathogen specific immune response that persists for generations. Multiple studies have identified altered DNA methylation patterns that have occurred in response to, or as a result of, exposure to specific pathogens (Bobetsis et al., 2007; Mikovits et al., 1998; Tolg et al., 2011). This study was the first to evaluate DNA methylation patterns in women with GBS colonization and functional analysis reveals these changes may play contribute to colonization susceptibility.

Genes incorporated into significant functional clusters using the DAVID bioinformatics software were independently searched within the GeneCards® database, The Human Gene Compendium Encyclopedia (http://www.genecards.org/). Functional cluster analysis was completed using DAVID bioinformatics software. None of the functional categories were significant when FDR is applied to determine significance of the clusters identified. However, for genes identified by using both GenomeStudio and R

software as being differentially methylated, the cell morphogenesis functional cluster has potential to offer mechanistic insight into GBS colonization and was significant using and alpha of 0.05 (p=0.01). The functional cluster was associated with methylation loss in GBS positive women. The cell morphogenesis cluster includes four differentially methylated genes and has an enrichment score of 1.77, which is above the suggested 1.3 cutoff value indicating likely involvement in biological process (Huang et al., 2009a). What makes this particular cluster interesting is that the genes within the cluster are associated with various immune functions and pathways. Gómez et al., (2010) determined that there is an association with PRKCA and bacterial vaginosis. This gene is associated with abnormal bacteria in the vagina and a similar association may exist with GBS susceptibility. The CUL3 gene is in a SuperPaths specifically related to antigen processing and the adaptive immune system (Andérica-Romero, González-Herrera, Santamaría, & Pedraza-Chaverri, 2013; Pintard, Willems, & Peter, 2004; Singer, Gurian-West, Clurman, & Roberts, 1999). It is possible that the loss in methylation in women with GBS alters antigen presentation and how the body responds to GBS (e.g. allowing colonization or clearing the bacteria). The GAP43 gene has been associated with inflammatory disease processes including contact dermatitis (El-Nour et al., 2006) and cutaneous malignant melanoma (Reed, Finnerty, & Albino, 1999). Since the gene is hypomethylated in women with GBS, it is possible there is increased cutaneous inflammation which is damaging to normal flora and creates a niche for GBS to colonize. Additional research investigating expression levels, protein products, and associated clinical outcomes, could be beneficial for the genes in this functional cluster.

There were two additional functional clusters with enrichment scores above the 1.3 threshold for biological significance: cellular metabolic and biosynthetic processes (1.46, p= 0.03) associated with genes that were hypermethylation and regulation of apoptosis (1.34, p = 0.05) associated with genes that were hypomethylated. Upon searching in GeneCards and PubMed, there does not appear to be any literature directly associating the genes examined in this study to with immune variations related to infections in either of the aforementioned pathways. There are 1,166 publications associating the genes identified with various cancers and neurological ailments, but none of the studies directly pertaining to infectious disease processes. Therefore, it seems unlikely that the genes in this cluster will offer any mechanistic insight unless there are indirect linkages to inflammatory processes. However, genes identified in this cluster may still be useful as clinical biomarkers in the future for identifying carriers, or women at risk for colonization in early pregnancy to prevent preterm labor or other poor health outcomes if the results are replicated in a larger cohort.

Two of the genes (CUL3 and PRCKA) in the apoptosis cluster are the same as in the cell morphogenesis cluster. The remaining genes in the apoptosis functional cluster do not appear to have any direct linkages to immune processes related to infection. There were 1,462 publications identified for the three genes that did not overlap with the morphogenesis cluster that were associated primarily with tumors, cancers, and muscular dystrophies. Similar to the metabolic and biosynthetic cluster, it seems unlikely that this cluster will offer any significant mechanistic understanding to GBS colonization. However, these genes may also be useful clinical biomarkers in the future after additional analysis is conducted for identifying GBS carriers or individuals susceptible to GBS

colonization. Further evaluation in a larger cohort and gene expression data may assist in determining if these genes are useful as clinical biomarkers for GBS colonization. Future studies investigating the relationship between PRKCA, CUL3, and GAP43 genes and GBS colonization may offer mechanistic insight and provide targets for future GBS treatment or to develop colonization prevention strategies.

Nursing Implications

Advances in epigenomic research are beginning to contribute significantly to scientific understanding of how environmental factors may contribute to various disease processes. This study is the first to assess laboratory values that are increasingly being used for nursing research (cytokines, vitamin D and DNA methylation) related to GBS colonization status in pregnant women. Existing gene-environment interaction models (Figure 1) were adapted (Figure 2) in order to guide the approach for this nursing research study to investigate if quantitative variations in cytokines or DNA methylation levels identified in in pregnant women's could be used to identify women at risk for GBS colonization. Although this study did not identify any significant differences between women colonized with GBS and women without GBS colonization, this study adds a novel model and approach method that can be used in future nursing research that can be modified as research methodology continues to evolve. Expanding on existing research models to bridge the gap for nurses to conduct translational research to improve outcomes will have implications for nursing research, practice, education and policy.

Nursing Research Implications

Nurses are uniquely poised to accelerate the translational arm of epigenomic research to better assist clinical populations of interest to attain and maintain optimal health functioning. Conducting research that critically examines environmental exposures

and unique epigenomic signatures will allow for the discovery of new treatment targets and the ability to create new disease prevention strategies ranging from diet modification to driving forward policy chance to protect public health and well-being. This is the first study investigating cytokines, vitamin D, or DNA methylation levels in peripheral blood to identify potential associations with GBS colonization susceptibility. Historically, much of the epigenomic literature is dominated by cancer studies. However, studies investigating the relationship between epigenomic alterations and complex diseases, other than cancer, have been increasing in recent years. It is important to note that identifying aberrant DNA methylation patterns alone will not explain why or how it was altered, how to intervene, or help the patients avoid acquiring abnormal patterns.

Nursing Practice Implications

A holistic approach for investigating the impact of epigenomic alterations on health status is a necessity if science intends on using epigenomic information to improve health. Other lifestyle patterns (diet, lifestyle, stress) and exposures must be assessed to identify interactions that may be causing the altered DNA methylation pattern. Nursing clinical assessments can help illuminate human-environment interactions, the endogenous and exogenous factors in the model developed to guide this study (Figure 2), that may be altering DNA methylation patterns that cannot be identified by studies using cell or animal models. Studies involving actual clinical populations are needed, specifically for disease processes that are inflammatory in nature because the immune response in murine models does not correlate with the human inflammatory response (Seok et al., 2013). Although it is near impossible to select populations to eliminate all confounding variables, perhaps it is time to embrace studies that acknowledge and address

confounding variables. The environment has an undeniable impact on health and disease states, therefore to completely eliminate all confounding variables from bench studies may explain some of the barriers encountered when translating research from bench to bedside. Humans do not live in a well-controlled, isolated environment; and their environment will alter how they respond to treatment and environmental exposures. Most complex diseases are the results of a culmination of genetic and environmental factors unique to an individual. Each person's DNA and environmental exposures are unique and assessing both (Figure 2), as well as the interaction of the two, will result in more personalized healthcare. Because nurses are educated on how to assess patients and the environment holistically; nurses are well poised to drive translational research and include information obtained from these assessments to investigating the interaction of environment and epigenomic signatures. Strong communication between bedside nurses and nurse scientists are needed to reconcile the gap between bench studies and what nurses find applicable and useful in the clinical setting.

Nursing Policy Implications

Investigation of clinical values that may be directly altered by environmental exposures has implications for nursing policy as well as general public health policies. Person, health, environment, and nursing are the metaparadigm concepts that remain the pillars of the nursing discipline. Perception of these concepts constantly evolves to incorporate new knowledge gained through practice, research, education, and exposure to other disciplines. Since most human disease processes are multifactorial in nature and nurses interact with individuals throughout the illness-wellness continuum, it is imperative nurses become involved with policy development to protect public health

based on research findings that are evident in clinical populations. For example, if the results from this study had indicated that women without GBS colonization had significantly higher serum vitamin D levels additional studies would be needed to validate that the findings were accurate and then to determine if vitamin D supplementation could be protective. After further studies, a significant effort would be required to incite policy change incorporating vitamin D supplementation to prevent GBS colonization into practice. As the state of the science continues to rapidly evolve, it is important to being considering how public health and nursing policies can be addressed and updated to reflect current methodologies and research findings.

Nursing Education Implications

Incorporating research models, like the one developed for this study, could have implications for nursing education. Students are usually required to complete basic science courses prior to acceptance and entry into a nursing program (nutrition, chemistry, anatomy, physiology). The base knowledge is required in order to understand how biological phenomena can be utilized, manipulated, and applied to nursing practice. Since the sequencing of the human genome was completed, striking advances in genetics and genomics have occurred and nurses at all levels will be expected to be able to communicate these findings to patients and be able to identify how the gene-environment interactions affect health and illness (Consensus Panel on Genetic/Genomic Nursing Competencies, 2009). Per the recommendations from the Consensus Panel on Genetic/Genomic Nursing Competencies, all programs of nursing should be incorporating education on genetics and genomics for entry level nurses. Essential content that should be taught includes: incorporating genetic and genomic knowledge

into nursing assessment and care, what to do with the information obtained or where to refer patients, understanding how personal opinions of genetic and genomic testing and interventions may affect practice. Translation of genetics into treatment is already occurring and very prevalent in certain areas of nursing, such as maternal-child health and oncology and will become more pronounce in other areas as the state of the science continues to evolves (Kirk, Calzone, Arimori, & Tonkin, 2011). It is imperative that nurses incorporate this knowledge into practice to properly care for patients. Further, "there is a growing abundance of genomic resources already available in a range of formats that cover most teaching environments and learning approaches. For many topic areas, particularly bioscience, there is no need to reinvent the wheel and develop new resources" (Tonkin, Calzone, Jenkins, Lea, & Prows, 2011, p. 336). Content on the epigenetics and appropriate models for analysis should be incorporated into the genomics content due to the increasing number of research studies investigating epigenetic mechanisms contributions to alter health status.

Limitations

Many of the limitations of this study are a direct result of being a secondary data analysis. Therefore, it is difficult to eliminate the possibility of misclassification bias. The first major limitation of the study is the small sample size. The intent of this exploratory study was to identify significant group differences and determine effect size so that sample size could be determined in future investigations. There have been no previous studies evaluating DNA methylation patterns, serum cytokine levels or vitamin D (25[OH]D) status in pregnant women with and without GBS colonization. Given the high cost of conducting these types of laboratory analyses, conducting a study with a small

sample size to determine feasibility and potential clinical utility is the most economically sound option. Additionally, the intent of using the DAVID database was to identify additional clinically relevant findings that could be utilized as pilot data for a subsequent grant proposal. As outlined in the methods section in the approach for analyzing DNA methylation, both gain and loss of methylation was determined and evaluated. The functional significance varied depending on the direction of change in methylation. We did find significant pathways for methylation gains and losses. However, future studies focusing on genes identified in the methylation loss pathways may prove to have some clinical utility. As discussed previously, the genes involved in the cell morphogenesis and regulation of apoptosis pathways make clinical sense and could potentially be epigenetic factors that contribute to GBS colonization susceptibility.

Another limitation of the study is that only peripheral blood samples were available for analysis and all other information had been previously extracted from the medical record. For example, it would have been more ideal to run the genome-wide DNA methylation analysis on epithelial cells taken from the recto-vaginal swabs at the time of GBS screening because that is when colonization status is determined and that is the reservoir site for neonatal infection. However, we had no access to the swabs or peripheral blood samples at the exact same time point. Since we had maternal DNA samples from the first trimester in pregnancy, any differences associated with GBS colonization could be a clinical indicator for susceptibility. It may also allow for a more targeted screening approach for preventing poor pregnancy outcomes associated with GBS colonization that are currently not prevented or screened. Additionally, a breakdown of the composition of cell types in the peripheral blood samples was not available.

Statistical corrections can be completed to correct for the heterogeneity of cell types found in peripheral blood (Houseman et al., 2012). However, Houseman et al., (2012) states that current statistical correction strategies are "a computationally difficult task that would have extreme vulnerability to model mis-specification." (p. 10). Given the exploratory nature, small sample size, difficulty in establishing model fit, and the fact that the variable cell types that may have an impact on results only make up 2-3% of the cell population, corrections for cell type were not incorporated into the analysis. However, in future studies with a larger sample size consideration will be given to separate cell types prior to analysis of DNA methylation to avoid this type of confounding in the analysis.

Conclusions

As the state of the science continues to evolve, it is imperative for nurses to incorporate advances in science into their program of research. Nurses are trained to translate information to people and populations with all levels of understanding. It is a natural fit for nurse scientists to step into a translational role and design studies to assess how the environment interacts with the individual in order to improve health outcomes. Investigation of epigenomic and genomic alterations related to complex disease processes has the potential to identify biologic mechanisms that contribute to the development of disease. Overall, gene-environment interaction models are useful for guiding nursing research investigating DNA methylation patterns because it allows for a holistic approach that clinical assessment data can be incorporated into. Additionally, DNA methylation patterns are readily measurable and offer insight into how environmental interaction can impact health by causing changes in gene expression. A number of standard laboratory protocols and bioinformatics tools can be utilized to complete exploratory studies. Since

DNA methylation patterns can now be identified by laboratory techniques, clinicians and research scientists must learn to decipher what the patterns mean and what the implications are for health. As health care progresses to incorporate more patient centered approaches, identification of altered DNA methylation patterns will improve nurses' ability to provide optimal care for patients. With an understanding of the impact of DNA methylation patterns, personalized, individual interventions can be developed to improve health based on research findings. Ultimately, this will improve care at the level of the individual.

APPENDICES

Appendix A

Study Institutional Review Board Approval

UNIVERSITY OF UND NORTH DAKOTA

INSTITUTIONAL REVIEW BOARD
c/o RESEARCH DEVELOPMENT AND COMPLIANCE
DIVISION OF RESEARCH
TWAMLEY HALL ROOM 106
264 CENTENNIAL DRIVE STOP 7134
GRAND FORKS ND 58202-7134
(701) 777-4279
FAX (701) 777-6708

October 17, 2012

Michelle Lynn Wright
Department of Nursing
Stop 9025

Dear Ms. Wright:

We are pleased to inform you that your project titled, "Inflammation and DNA Methylation as Group B Streptococcus Colonization Biomarkers" (IRB-201210-092) has been reviewed and approved by the University of North Dakota Institutional Review Board (IRB). The expiration date of this approval is December 31, 2014.

As principal investigator for a study involving human participants, you assume certain responsibilities to the University of North Dakota and the UND IRB. Specifically, any adverse events or departures from the protocol that occur must be reported to the IRB immediately. It is your obligation to inform the IRB in writing if you would like to change aspects of your approved project, prior to implementing such changes.

When your research, including data analysis, is completed, you must submit a Research Project Termination form to the IRB office so your file can be closed. A Termination Form has been enclosed and is also available on the IRB website.

If you have any questions or concerns, please feel free to call me at (701) 777-4279 or e-mail michelle.bowles@research.und.edu.

Sincerely,

Michelle L. Bowles, M.P.A., CIP
IRB Coordinator

MLB/jle

Enclosures

REPORT OF ACTION: EXEMPT/EXPEDITED REVIEW

University of North Dakota Institutional Review Board

Date: 10/16/2012 **Project Number:** IRB-201210-092

Principal Investigator: Wright, Michelle

Department: Nursing

Project Title: Inflammation and DNA Methylation as Group B Streptococcus Colonization Biomarkers

The above referenced project was reviewed by a designated member for the University's Institutional Review Board on 10/16/2012 and the following action was taken:

☐ Project approved. **Expedited Review** Category No. _____
Next scheduled review must be before: _____

☐ **Copies of the attached consent form with the IRB approval stamp dated _____ must be used in obtaining consent for this study.**

☒ Project approved. **Exempt Review** Category No. 4
This approval is valid until December 31, 2014 as long as approved procedures are followed. No periodic review scheduled unless so stated in the Remarks Section.

☐ **Copies of the attached consent form with the IRB approval stamp dated N/A must be used in obtaining consent for this study.**

☐ Minor modifications required. The required corrections/additions must be submitted to RDC for review and approval. **This study may NOT be started UNTIL final IRB approval has been received.**

☐ Project approval **deferred. This study may not be started until final IRB approval has been received.**
(See Remarks Section for further information.)

☐ Disapproved claim of exemption. This project requires Expedited or Full Board review. The Human Subjects Review Form must be filled out and submitted to the IRB for review.

☐ Proposed project is not human subjects research as defined under Federal regulations 45 CFR 46 or 21 CFR 50 and does not require IRB review.

☐ Not Research ☐ Not Human Subject

PLEASE NOTE: Requested revisions for student proposals MUST include adviser's signature. All revisions MUST be highlighted and submitted to the IRB within 90 days of the above review date.

☒ Education Requirements Completed. (Project cannot be started until IRB education requirements are met.)

cc: Denise Korniewicz, PhD, RN, FAAN; Cindy Anderson, PhD, WHNP-BC, FAAN

Michelle L. Wright 10/16/2012
Signature of Designated IRB Member Date
UND's Institutional Review Board

If the proposed project (clinical medical) is to be part of a research activity funded by a Federal Agency, a special assurance statement or a completed 310 Form may be required. Contact RDC to obtain the required documents.

(Revised 10/2006)

Appendix B

Parent Study UND IRB Approval

Research Project Review and Progress Report
University of North Dakota Institutional Review Board

RESEARCH DEVELOPMENT AND COMPLIANCE MAR 20 2012

DATE: March 5, 2012 **DEPARTMENT:** Family and Community Nursing

PRINCIPAL INVESTIGATOR: Anderson, Cindy

PROJECT TITLE: Vitamin D Status During Preeclampsia: Mechanisms Underlying Placental Vascular Alterations

PROPOSAL NUMBER: IRB-200809-045

IF MEDICAL COMPONENT, PLEASE GIVE PHYSICIAN'S NAME:

IRB USE ONLY

☐ FULL BOARD REVIEW REQUIRED, EVEN THOUGH ORIGINAL APPROVAL WAS EXPEDITED

☒ CONTINUED APPROVAL, "EXPEDITED" CATEGORY 2, 3, 7

☒ NEXT REVIEW REQUIRED BEFORE: March 22, 2013

☐ CONTINUED APPROVAL, BASED ON FULL BOARD REVIEW

☐ NEXT REVIEW REQUIRED BEFORE:

☐ SUSPEND APPROVAL, PENDING INVESTIGATION

☐ APPROVAL TERMINATED

COMMENTS OF REVIEWER:

Signature of Chair/Vice Chair or Designee:

cc: Chair, Family and Community Nursing Approval Date: 3-23-12

1. Is project complete? Yes ☐ No ☒

2. Is project ongoing? Yes ☒ No ☐

If No, explain below and indicate if continued approval and continuing review is desired.

3. How many subjects have been enrolled in the research project?

7 since the date of last approval, and

65 since the initial approval

4. Is the research permanently closed to the enrollment of new subjects? Yes ☒ No ☐

Have all subjects completed all research-related interventions? Yes ☒ No ☐

Does the research remain active only for long-term follow-up of subjects? Yes ☒ No ☐

5. Is data analysis complete? Yes ☐ No ☒

*** If the research is permanently closed to the enrollment of new subjects, all subjects have completed all research-related interventions, the research does not need to remain active for long-term follow-up of subjects, and all data analysis is complete, please sign here that you would like the IRB to terminate approval for this project, and finish filling out the rest of this form.

Please terminate IRB approval for this research project _____ _____
Signature of Principal Investigator Date

6. Has any additional grant money been awarded for this project in the past year? Yes ☐ No ☒
If yes, submit a copy of the grant along with this completed form.

7. Describe any adverse events and/or unanticipated problems involving risks to subjects or others that have occurred since the last approval. If you did not report the adverse event or unanticipated problem previously, a separate Unanticipated Problem/Adverse Event Form must be submitted to RD&C with this form.

No adverse events or unanticipated problems

8. Have any additional risks with this research been identified? Yes ☐ No ☒
Describe all benefits experienced by participants, and include a current risk/benefit assessment based on study results.

Participants expressed satisfaction about their contributions to generation of new knowledge to improve pregnancy outcomes

9. Have there been any changes or deviations from the approved protocol since the most recent approval? Yes ☐ No ☒ If Yes, elaborate below, and submit a separate Protocol Change Form to the RD&C indicating proposed protocol changes.

   a. Have any of these changes been implemented already? Yes ☐ No ☐
   If yes, please describe fully.

   b. Are any protocol changes being planned for later implementation? Yes ☐ No ☒
   If yes, please describe fully. A separate Protocol Change Form must be submitted to RD&C for approval before the proposed protocol changes can be implemented.

10. Have any subjects withdrawn from the research? Yes ☒ No ☐
If yes, state how many have withdrawn and describe the circumstances.

4 have withdrawn due to move away from study site and pregnancy termination. None withdrawn since last continuing review

11. Have there been any complaints about the research since the last IRB review? Yes ☐ No ☒
If yes, please report and summarize the complaints and your response/action.

12. Summarize any multi-site trial reports relevant to your research. N/A

13. Summarize any recent literature, findings, or other information relevant to your research, especially information about risks associated with the research.
No risks reported. Increasing evidence regarding health risks of vitamin D deficiency.

14. Have all PI's involved with the research completed the IRB Educational Requirements?
Yes ☑ No ☐ (Educational requirements must be completed before the IRB can grant continued approval for the research project.)

15. On a separate piece of paper, provide a thorough protocol summary (approximately 300 words) giving a concise summary of the protocol's progress to date and the reasons for continuing the study or reasons for asking the IRB to terminate approval. The summary should include, for instance, an explanation of any complaints about the research, relevant multi-site trial reports, participant benefits, or a current risk-benefit assessment based on study results. Sufficient information is required in the summary so that the IRB can determine whether the proposed research continues to fulfill the criteria for approval.
attached

16. A copy of the current informed consent document(s) (with the IRB Approval stamp), as well as a clean copy of the consent document(s) (with no IRB Approval stamp) must be submitted with this report.
current copy included.
clean copy included

17. Have there been any changes in the conflict of interest statement or situation for the Principal Investigators, research staff involved in the study, or each individual's respective family members in the last 12 months?
Yes ☐ No ☒ If yes, please describe fully on a separate sheet of paper.

Signature of Principal Investigator Cindy M Anderson Date 3/19/12

Current email address: cindy.anderson@email.und.edu

Current Address: Stop 9025

This completed form should be returned to the IRB, University of North Dakota, 264 Centennial Drive Stop 7134, Grand Forks, ND 58202-7134.

Appendix C

Parent Study Altru IRB Approval

Altru HEALTH SYSTEM

Institutional Review Board (IRB) Continuing Review

Cps to Cindy 5/24/12

Revised 5/10/11

Date: 5/16/12 IRB # ST-54

Principal Investigator: Cindy Anderson

Department: Nursing Phone # 777-4354

Research Coordinator: _____ Phone # _____

Project Title: Vitamin D Status during Preeclampsia: Mechanisms Underlying Placental Vascular Alterations.

INVESTIGATOR COMPLETE:

1. Is it an Expedited Review? ..... X Yes ___ No
2. Date of Original Approval: 9/8/08
3. Date of Last Approval: (if applicable) 6/21/11
4. Is project completed? (all research done, enrollment close & patients follow-up complete) ___ Yes X No

   if yes, complete questions 7-13. If no, complete questions 5-13.

5. Is the research permanently closed to the enrollment of new subjects?..... X Yes ___ No
6. Have all participants completed all research-related interventions? ..... X Yes ___ No
7. Give a brief statement of the progress of the project (i.e., number of subjects entered, length of time since study initiated, length of time projected to go on, etc.)

   Recruitment closed on 8/30/11. A total of 67 participants provided data. Analysis is expected to extend to June 2013.

8. Have any subjects withdrawn from the research?..... ___ Yes ___ No

   If yes, state how many have withdrawn and describe the circumstances:

   nine withdrawn - 8 did not complete data collection 2 miscarried, 1 changed practice provider, 4 moved, 1 lost to follow-up
9. Have there been any complaints/concerns about the research since the last IRB review?.... ___ Yes ✓ No

   If yes, please report the complaints/concerns and your response/action.

10. Any unanticipated problems involving risks not detailed in your original application? ..... ___ Yes ✓ No
11. Any serious adverse events (SAE) at your site? ..... ___ Yes ✓ No
12. Any SAE, safety update, action letter, IND or other at another site? ..... ___ Yes ✓ No
13. Any unexpected mortality or morbidity? ..... ___ Yes ✓ No

If yes to any of the above (7-13), please provide a detailed summary below. Need to write a more detailed summary about the adverse events, problems, etc. (i.e., the types of adverse events, numbers, trends).

*(**Example:** There were 6 adverse events during the past year, four of which occurred in patients at other sites and 2 occurred at this site. Three of the events were constipation thought to be related to the study drug; two were shortness*

of breath thought not to be related to the study drug; and one was a petechial rash thought to be possibly related.)

IRB Continuing Review

Also, make certain adverse events/mortality reports have been submitted and are on file in the IRB office.

14. Any protocol changes, consent form changes or amendments? ..... ___ Yes ✓ No
If yes, list below (make certain protocol changes/amendments have been submitted and are on file in the IRB office). No protocol changes since last review.

15. Number of subjects enrolled since last review: 0

16. Total number of subjects enrolled since project initiated: 67

17. Have all PIs involved with the research completed the IRB Education Requirements? ..... ✓ Yes ___ No
(Education requirements must be completed before the IRB can grant continued approval for the research project).

18. A copy of the current informed consent document (with, if possible, the IRB Approval dated stamp) must be submitted with this report.

Cindy Anderson
Signature of Principal Investigator

5/16/12
Date

****

IRB USE ONLY:

☒ Continue approval based on Expedited Review

θ Continue approval based on Full Board Review

θ Suspend approval, pending investigation

W Mabb
Signature of IRB Chairperson/Designee

5/23/12
Date

☒ Next Review required before 5/22/13

Form last updated 5/10/11

Completed by: _____

Appendix D

Medical Record Abstraction Form

Study 608 **ID Number:**

Reviewed by: _____ On: ____/____/____

Instructions

Complete this medical chart abstraction form for each participant enrolled in the study.

A. Detailed Past Obstetrical History

LMP: ___/___/___ **EDC:** ___/___/___ **Date of First Prenatal Visit:** ___/___/___ **Maternal Age at Time of Delivery:** _____

Gravida Status

G: ____	P: ____	T: ____	A: ____	L: ____

Pregnancy Outcome	Type of Delivery	Birth Weight	Maternal Description	Code	Fetal/Infant Description	Code	If LB & Child NOT Living
1. Date: __ __/__ __/__ __ Time: _____ *mm/dd / yy* Outcome: __| __| GA: __| __| / __| __| *wks days* GA code: __| __| Gender: __| __|	__| __|	1) __ __/ __ __ *lbs / oz* ***OR*** 2) _____ *grams*	1)_____ 2)_____ 3)_____	1) __| __| 2) __| __| 3) __| __|	1)_____ 2)_____ 3)_____	1) __| __| 2) __| __| 3) __| __|	Date of Death: __ __/__ __/__ __ *mm /dd / yy* Reason for death: _____ Code: __| __|

Outcome	GA (Gestational Age) code	Gender	Maternal Complications	Fetal/Infant Complications	Child NOT Living (Reason for Death)
LB=Live birth SB=Stillbirth SA=Spontaneous Abortion TA=Therapeutic Abortion EP=Ectopic Pregnancy MP=Molar Pregnancy	Enter time in weeks/days if documented. If number of weeks is not documented, enter: FT=full term(≥ 37+ weeks) NT=near term (32-36 wks) PT=early preterm (20-31 wks) ET=early termination (<20 wks)	01=male 02=female **Type of Delivery** 01=vaginal 02=op vaginal (forceps or vacuum) 03=C-section	00=None noted 01=Preeclampsia 02- Gestational Hypertention 03=Placenta previa 04=Placental abruption 05=Pre-gestational diabetes – Type I 06=Pre-gestational diabetes – Type II 07=Gestational diabetes 08=PPH, requiring blood transfusion 09=PTL requiring treatment 10=PPROM 99=Other	00=None noted 01-Intra-Uterine Growth Restriction 02- Small for Gestational Age 03- Large for Gestational Age 04-Respiratory Distress Syndrome 05-Meconium Aspiration 06=Other defect 07=Other genetic disease 08-Observe for sepsis 09-Hypoglycemia 99=Other	01=Cardiovascular 02=Congenital Defect 03=Respiratory 04-Prematurity 05-Sepsis 06-Unknown 99=Other

Coding Key:
-4 = Temporarily Missing (not currently available) -7 = Don't Know
-5 = Multiple Responses (needs review) -8 = Refused to Answer
-6 = Permanently Missing (not documented in chart) -9 = Does Not Apply

B. History of Medical Conditions

	Medical Condition	(circle one for each) Present	Absent	Receiving Medication	Not Receiving Medication
1.	Asthma	1	2	1	2
2.	Seizure disorder	1	2	1	2
3.	Chronic hypertension	1	2	1	2
4.	Diabetes mellitus (type 1 & 2)	1	2	1	2
5.	Hyperthyroidism	1	2	1	2
6.	Hypothyroidism	1	2	1	2
7.	Valvular heart disease	1	2	1	2
8.	Other structural heart disease	1	2	1	2
9.	Coronary artery disease/ congestive heart failure	1	2	1	2
10.	Nephropathy/Nephrotic syndrome/ Glomerulonephritis	1	2	1	2
11.	Renal insufficiency/ renal failure	1	2	1	2
12.	Sickle cell anemia	1	2	1	2
13.	Thrombocytopenia	1	2	1	2
14.	Lupus erythematosus	1	2	1	2
15.	Antiphospholipid antibody syndrome	1	2	1	2
16.	Rheumatoid arthritis	1	2	1	2
17.	Ulcerative colitis/Crohn's disease	1	2	1	2
19.	Malignancy Specify _____ (please specify)	1	2	1	2
20.	Hepatitis B	1	2	1	2
21.	Hepatitis C	1	2	1	2
22.	Psychiatric Disorder Specify _____ (please specify)	1	2	1	2
23.	Other Specify _____	1	2	1	2

Medical Condition	(circle one for each) Present	Absent	Receiving Medication	Not Receiving Medication
24. Other Specify _____ (please specify)	1	2	1	2
25. Other Specify _____ (please specify)	1	2	1	2
26. Other Specify _____ (please specify)	1	2	1	2

C. Surgical Conditions

1. Has the woman had any surgeries:

(circle one)

No 0

Yes..... 1

IF YES

List the surgery and the year:

	Surgery	Year
a.	_____	____ ____ ____ ____
b.	_____	____ ____ ____ ____
c.	_____	____ ____ ____ ____
d.	_____	____ ____ ____ ____
e.	_____	____ ____ ____ ____
f.	_____	____ ____ ____ ____
g.	_____	____ ____ ____ ____
h.	_____	____ ____ ____ ____

Coding Key:
-4 = Temporarily Missing (not currently available)
-5 = Multiple Responses (needs review)
-6 = Permanently Missing (not documented in chart)
-7 = Don't Know
-8 = Refused to Answer
-9 = Does Not Apply

D. Diabetes

1. Did the woman enter pregnancy with a diagnosis of diabetes:

(circle one)

No 0

Yes 1

IF YES

a. At what age was diabetes first diagnosed: _____ years
(please specify)

b. What type of diabetes:

(circle one)

Type 1 DM 1

Type 2 DM 2

Unspecified 3

c. Which of the following treatments were used at the time of LMP/conception:

		(circle one for each) NO	YES
1.	Diet	0	1
2.	Insulin injections	0	1
3.	Insulin pump	0	1
4.	Oral hypoglycemic agent	0	1
5.	Oral insulin sensitizing agent	0	1

d. Were any of the following complications of diabetes noted:

		(circle one for each) *Present, at entry to prenatal care*	*Present, found during pregnancy*	*Absent*
1.	Nephropathy/renal insufficiency/renal failure	0	1	2
2.	Retinopathy/blindness...	0	1	2
3.	Peripheral neuropathy....	0	1	2
4.	Coronary artery disease.....	0	1	2

Coding Key:
-4 = Temporarily Missing (not currently available)
-5 = Multiple Responses (needs review)
-6 = Permanently Missing (not documented in chart)
-7 = Don't Know
-8 = Refused to Answer
-9 = Does Not Apply

E. Prescription Medication, Vitamins and Vaccines

	Medication, Vitamin, Vaccine			Trimester (01=Taken 02=Not taken)			Postpartum
	Name	Code*	other	1st	2nd	3rd	
1.							
2.							
3.							
4.							
5.							
6.							
7.							
8.							
9.							
10.							

Medications (100- 300 series)					**Vitamins (500 series)**	**Vaccines (600 series)**
Analgesics 101=Narcotic 102=NSAID 103=Aspirin 104=Acetaminophen	110=Antibiotics 120=Anticoagulants 130=Antidepressants 140=Anticonvulsants 150=Antihistamines	Antihypertensives 161= Aldomet 162= Labetolol 163= Ca-Channel Blockers 164= Beta-Blockers 165= Ace-Inhibitor 169 = Other antihypertensive	170=Antiemetics 180=Antipsychotics 190=Antivirals 200=Birth control pills 210=Chemotherapeutics 220=Diuretics 230=GI agents 240=Progesterone 260=Rhogam 270= Sleep Aide 280=Steroids	Thyroid Agents 251=Antithyroids (overactive) 252=Thyroid Replacement (under active) 399=Other Medication	510=Multi – vitamin 520=Iron 530=Folate 540=Calcium + D 599=Other vitamin	610=Influenza 620=Hepatitis B 630=Rubella 640=Varicella-zoster immune globulin (VZIG) 699=Other vaccine

Coding Key:
-4 = Temporarily Missing (not currently available)
-5 = Multiple Responses (needs review)
-6 = Permanently Missing (not documented in chart)
-7 = Don't Know
-8 = Refused to Answer
-9 = Does Not Apply

F. Prenatal Care Visits

Instructions

Complete the Prenatal Care Flow sheet on all women enrolled in the study using medical records. Complete one row for each prenatal visit.

Visit	Date mm/ dd / yy	Weight lbs	Highest Blood Pressure mm Hg	Fundal Height cm	Fetal Heart Rate 01=present 02=absent	Fetal Movement 00=normal 01=decreased 02=absent	Urine Dipstick Proteinuria 01=Negative 02=Trace 03=+1 04=+2 05=+3 06=+4	Urine Dipstick Glucosuria 01=Negative 02=Trace 03= ≥ +1	Other Condition Specify
1.	__ __/__ __/__ __		1. Sys __ __ __ 2. Dia __ __ __						
2.	__ __/__ __/__ __		1. Sys __ __ __ 2. Dia __ __ __						
3.	__ __/__ __/__ __		1. Sys __ __ __ 2. Dia __ __ __						
4.	__ __/__ __/__ __		1. Sys __ __ __ 2. Dia __ __ __						
5.	__ __/__ __/__ __		1. Sys __ __ __ 2. Dia __ __ __						
6.	__ __/__ __/__ __		1. Sys __ __ __ 2. Dia __ __ __						
7.	__ __/__ __/__ __		1. Sys __ __ __ 2. Dia __ __ __						
8.	__ __/__ __/__ __		1. Sys __ __ __ 2. Dia __ __ __						
9.	__ __/__ __/__ __		1. Sys __ __ __ 2. Dia __ __ __						

Coding Key:
-4 = Temporarily Missing (not currently available)
-5 = Multiple Responses (needs review)
-6 = Permanently Missing (not documented in chart)
-7 = Don't Know
-8 = Refused to Answer
-9 = Does Not Apply

Visit	Date mm/ dd / yy	Weight lbs	Highest Blood Pressure mm Hg	Fundal Height cm	Fetal Heart Rate 01=present 02=absent	Fetal Movement 00=normal 01=decreased 02=absent	Urine Dipstick Proteinuria 01=Negative 02=Trace 03=+1 04=+2 05=+3 06=+4	Urine Dipstick Glucosuria 01=Negative 02=Trace 03= ≥ +1	Other Condition Specify
10.	__/__/__		1. Sys ___ 2. Dia ___						
11.	__/__/__		1. Sys ___ 2. Dia ___						
12.	__/__/__		1. Sys ___ 2. Dia ___						
13.	__/__/__		1. Sys ___ 2. Dia ___						
14.	__/__/__		1. Sys ___ 2. Dia ___						
15.	__/__/__		1. Sys ___ 2. Dia ___						
16.	__/__/__		1. Sys ___ 2. Dia ___						
17.	__/__/__		1. Sys ___ 2. Dia ___						
18.	__/__/__		1. Sys ___ 2. Dia ___						

Coding Key:
-4 = Temporarily Missing (not currently available) -7 = Don't Know
-5 = Multiple Responses (needs review) -8 = Refused to Answer
-6 = Permanently Missing (not documented in chart) -9 = Does Not Apply

G. General Prenatal Labs

1. **Mother's Blood Type:** *(circle one)*
   - A.....1
   - B.....2
   - O.....3
   - AB.....4

2. **RH Factor:** *(circle one)*
   - Positive.....1
   - Negative.....2

   a. Antibody screen: ***(circle one)***
   - Positive.....1
   - Negative.....2

   b. Date of most recent RhIG given: (mm/dd/yyyy)

   ___ ___ / ___ ___ / ___ ___ ___ ___

IF SCREEN POSITIVE

	Date Identified mm/ dd / yy	Type(s) Identified	Titer	Type of Antibody: Enter alpha code (upper & lower case):
1.	__ __/ __ __/ __ __	1. _ _ _ _ _ 2. _ _ _ _ _ 3. _ _ _ _ _	1. _ : _ _ _ 2. _ : _ _ _ 3. _ : _ _ _	
2.	__ __/ __ __/ __ __	1. _ _ _ _ _ 2. _ _ _ _ _ 3. _ _ _ _ _	1. _ : _ _ _ 2. _ : _ _ _ 3. _ : _ _ _	
3.	__ __/ __ __/ __ __	1. _ _ _ _ _ 2. _ _ _ _ _ 3. _ _ _ _ _	1. _ : _ _ _ 2. _ : _ _ _ 3. _ : _ _ _	
4.	__ __/ __ __/ __ __	1. _ _ _ _ _ 2. _ _ _ _ _ 3. _ _ _ _ _	1. _ : _ _ _ 2. _ : _ _ _ 3. _ : _ _ _	

Type	Code
D	D
Kell	Kell
E	E
e	e
C^w	Cw
C	C
Ce	Ce
Kp^a	Kpa
Kp^b	Kpb
cE	cE
k	k
Jk	Jk
s	s
Wr^a	Wra
Fy^a	Fya
M	M
Other	Oth

Coding Key:
-4 = Temporarily Missing (not currently available)
-5 = Multiple Responses (needs review)
-6 = Permanently Missing (not documented in chart)
-7 = Don't Know
-8 = Refused to Answer
-9 = Does Not Apply

3. **HCT/HgB:**

	Date mm/ dd / yy	Result HgB g/dl	HCT %
a.	__/__/__	__.__	__.__
b.	__/__/__	__.__	__.__
c.	__/__/__	__.__	__.__

4. **PAP smear:** *(circle one)*

Normal.....1
Abnormal, ASCUS.....2
Abnormal, LGSIL.....3
Abnormal, HGSIL.....4
Abnormal, CIS/invasive disease.....5
Unsatisfactory/insufficient for evaluation.....6

5. **Rubella:** *(circle one)*

Immune.....1
Nonimmune.....2

6. **GBS:**

	Date mm/ dd / yy	Result 01= Positive 02= Negative
a.	__/__/__	__
b.	__/__/__	__

7. **RPR/VDRL:**

	Date mm/ dd / yy	Result 01=Reactive 02=Nonreactive	Titer	FTA Result 01=Positive 02=Negative
a.	__/__/__	__	__:__	__
b.	__/__/__	__	__:__	__

8. **Urinalysis:**

	Date mm/ dd / yy	Result(s) 01=Negative 02=Protein 03=Ketones 04=Bacteria	If Other, specify	00 = Absent, 1+, 2+, 3+ or 4+, trace (5)
a.	__/__/__			__
b.	__/__/__			__
c.	__/__/__			__
d.	__/__/__			__

9. **Urine culture:**

	Date mm/ dd / yy	Result 01=Positive 02=Negative	If Positive, specify organism 01=GBS 02=E. coli 03=Other	If Other, specify
a.	__/__/__	__	__	Other: _____
b.	__/__/__	__	__	Other: _____
c.	__/__/__	__	__	Other: _____
d.	__/__/__	__	__	Other: _____

Coding Key:
-4 = Temporarily Missing (not currently available)
-5 = Multiple Responses (needs review)
-6 = Permanently Missing (not documented in chart)
-7 = Don't Know
-8 = Refused to Answer
-9 = Does Not Apply

10. Hepatitis B Surface Antigen:

	Date mm/ dd / yy	Result 01= Positive 02= Negative	If other, hep B serology: (use codes below)
a.	__/__/__	__ __	__ __
b.	__/__/__	__ __	__ __

Other hep B serology codes:

01=HbeAg ((Hepatitis E antigen)
02=Anti HBc IgM (IgM core antibody)
03=Anti-HBs (Surface antibody)
04=Anti–Hbe (e antibody)
05=HBV DNA

11. PPD: *(circle one)*

Positive..... 1
Negative..... 2

12. Chlamydia:

	Date mm/ dd / yy	Result 01= Positive 02= Negative
a.	__/__/__	__ __
b.	__/__/__	__ __
c.	__/__/__	__ __

13. Gonorrhea:

	Date mm/ dd / yy	Result 01= Positive 02= Negative
a.	___/___/___	___ ___
b.	___/___/___	___ ___
c.	___/___/___	___ ___

14. Diabetes screen:

	Date mm/ dd / yy	Blood Sugar Result Fasting	1 Hour	Random
		_____.___	_____.___	_____.___
		☐ mg/dl	☐ mg/dl	☐ mg/dl
a.	___/___/___	☐ mmol/L	☐ mmol/L	☐ mmol/L
		_____.___	_____.___	_____.___
		☐ mg/dL	☐ mg/dl	☐ mg/dl
b.	___/___/___	☐ mmol/L	☐ mmol/L	☐ mmol/L

15. GTT:

	Date mm/ dd / yy	Blood Sugar Result Fasting	1 hr	2 hr	3 hr
		___ ___ ___ . ___	___ ___ ___ . ___	___ ___ ___ . ___	___ ___ ___ . ___
		☐ mg/dl	☐ mg/dl	☐ mg/dl	☐ mg/dl
a.	___/___/___	☐ mmol/L	☐ mmol/L	☐ mmol/L	☐ mmol/L
		___ ___ ___ . ___	___ ___ ___ . ___	___ ___ ___ . ___	___ ___ ___ . ___
		☐ mg/dl	☐ mg/dl	☐ mg/dl	☐ mg/dl
b.	___/___/___	☐ mmol/L	☐ mmol/L	☐ mmol/L	☐ mmol/L

H. Specialized Testing

Instructions

Complete the Specialized Prenatal Labs on all women enrolled prospectively into the study who have results of specialized testing in their charts. All questions refer to the mother. If the number of tests done of a type exceeds the number of rows provided, enter the results from the first prenatal lab tests on the first row. Use the remaining rows to enter the most recent results chronologically.

3.

1. **Specialized testing:**

(circle one)

No..... 0

Yes..... 1

2. **Hemoglobin (HgB) A1C:**

	Date mm/ dd / yy	Result %
a.	__ __/ __ __/ __ __	__ __ . __
b.	__ __/ __ __/ __ __	__ __ . __
c.	__ __/ __ __/ __ __	__ __ . __
d.	__ __/ __ __/ __ __	__ __ . __

3. **Twenty-four hour urine protein:**

	Date mm/ dd / yy	Result mg/24 hr
a.	__ __/ __ __/ __ __	__ __ __ __ __
b.	__ __/ __ __/ __ __	__ __ __ __ __
c.	__ __/ __ __/ __ __	__ __ __ __ __

4.

4. **Pre-Eclampsia Labs:**

5.

Test	Date Done mm/dd/yy	Result	Units
LDH	/ /		IU
	/ /		IU
	/ /		IU
	/ /		IU
AST (SGOT)	/ /		IU
	/ /		IU
	/ /		IU
	/ /		IU
ALT (SGPT)	/ /		IU
	/ /		IU
	/ /		IU
	/ /		IU

Uric Acid	/ /		mg/dL
	/ /		mg/dL
	/ /		mg/dL
	/ /		mg/dL
BUN	/ /		mg/dL
	/ /		mg/dL
	/ /		mg/dL
	/ /		mg/dL
Creatinine	/ /		mg/dL
	/ /		mg/dL
	/ /		mg/dL
	/ /		mg/dL
Hemoglobin	/ /		g/dL
	/ /		g/dL
	/ /		g/dL
	/ /		g/dL

Platelets	/ /		K/uL
	/ /		K/uL
	/ /		K/uL
	/ /		K/uL
Urine Prot/ Creat Ratio	Date mm/dd/yy	Ratio Result	
	/ /		urine total protein ___ mg/dL urine creatinine, random ___ mg/dL
	/ /		urine total protein ___ mg/dL urine creatinine, random ___ mg/dL
	/ /		urine total protein ___ mg/dL urine creatinine, random ___ mg/dL
	/ /		urine total protein ___ mg/dL urine creatinine, random ___ mg/dL
24- hour urine protein	/ /	Result	mg/24 hr
	/ /		mg/24 hr
	/ /		mg/24 hr
	/ /		mg/24 hr

5. Other Lab Tests:

	Specify Test	Date Done mm / dd / yy	Result 01=Normal 02= Abnormal
a.		__ __ / __ __ / __ __	
b.		__ __ / __ __ / __ __	
c.		__ __ / __ __ / __ __	
d.		__ __ / __ __ / __ __	

6. Fetal Surveillance: Non-Stress Tests (NSTs):

	Date mm/ dd / yy	Result 01= Reactive 02= Non-reactive, reassuring 03= Nonreactive, nonreassuring: requires further testing or delivery
a.	__ __/ __ __/ __ __	__ __
b.	__ __/ __ __/ __ __	__ __
c.	__ __/ __ __/ __ __	__ __
d.	__ __/ __ __/ __ __	__ __
e.	__ __/ __ __/ __ __	__ __
f.	__ __/ __ __/ __ __	__ __
g.	__ __/ __ __/ __ __	__ __
h.	__ __/ __ __/ __ __	__ __
i.	__ __/ __ __/ __ __	__ __
j.	__ __/ __ __/ __ __	__ __

I. Prenatal Ultrasound

Instructions

Complete the Prenatal Ultrasound Diagnosis on all women enrolled prospectively into the study using medical records. Complete one row for each ultrasound the woman received. If the number of ultrasounds is more than eight, note the first 4 ultrasounds on lines #1-4 and last 4 on lines #5-8. For each ultrasound record up to three results.

	Date mm/dd/yy	Measurements CRL=Crown-rump length BPD=Biparietal diameter HC=Head circumference AC=Abdominal circumference FL=Femur length	Result* Code (see codes on page 17)	Specify	AFI (Amniotic Fluid Index) cm	*BioPhysical Profile*	
1.	__/__/__	Consistent with: 1. CRL _____ cm _____ wks gestation 2. BPD _____ cm _____ wks gestation 3. HC _____ cm _____ wks gestation 4. AC _____ cm _____ wks gestation 5. FL _____ cm _____ wks gestation _____ wks gestation	1. ___ 2. ___ 3. ___ 4. ___ 5. ___		___.___ ☐ Check if measurement is normal.	Fetal Tone Gross body movements Fetal breathing movements Amniotic fluid volume	_____/2 _____/2 _____/2 _____/2
2.	__/__/__	Consistent with: 1. CRL _____ cm _____ wks gestation 2. BPD _____ cm _____ wks gestation 3. HC _____ cm _____ wks gestation 4. AC _____ cm _____ wks gestation 5. FL _____ cm _____ wks gestation _____ wks gestation	1. ___ 2. ___ 3. ___ 4. ___ 5. ___		___.___ ☐ Check if measurement is normal.	Fetal Tone Gross body movements Fetal breathing movements Amniotic fluid volume	_____/2 _____/2 _____/2 _____/2
3.	__/__/__	Consistent with: 1. CRL _____ cm _____ wks gestation 2. BPD _____ cm _____ wks gestation 3. HC _____ cm _____ wks gestation 4. AC _____ cm _____ wks gestation 5. FL _____ cm _____ wks gestation _____ wks gestation	1. ___ 2. ___ 3. ___ 4. ___ 5. ___		___.___ ☐ Check if measurement is normal.	Fetal Tone Gross body movements Fetal breathing movements Amniotic fluid volume	_____/2 _____/2 _____/2 _____/2
4.	__/__/__	Consistent with: 1. CRL _____ cm _____ wks gestation 2. BPD _____ cm _____ wks gestation 3. HC _____ cm _____ wks gestation 4. AC _____ cm _____ wks gestation 5. FL _____ cm _____ wks gestation _____ wks gestation	1. ___ 2. ___ 3. ___ 4. ___ 5. ___		___.___ ☐ Check if measurement is normal.	Fetal Tone Gross body movements Fetal breathing movements Amniotic fluid volume	_____/2 _____/2 _____/2 _____/2

	Date mm/dd/yy	Measurements CRL=Crown-rump length BPD=Biparietal diameter HC=Head circumference AC=Abdominal circumference FL=Femur length	Result* Code (see codes on page 17)	Specify	AFI (Amniotic Fluid Index) cm	*BioPhysical Profile*
4.	__/__/__	Consistent with: 1. CRL _____ cm _____ wks gestation 2. BPD _____ cm _____ wks gestation 3. HC _____ cm _____ wks gestation 4. AC _____ cm _____ wks gestation 5. FL _____ cm _____ wks gestation _____ wks gestation	1. ___ 2. ___ 3. ___ 4. ___ 5. ___	_____ _____ _____ _____	___.___ ☐ Check if measurement is normal.	Fetal Tone _____/2 Gross body movements _____/2 Fetal breathing movements _____/2 Amniotic fluid volume _____/2
5.	__/__/__	Consistent with: 1. CRL _____ cm _____ wks gestation 2. BPD _____ cm _____ wks gestation 3. HC _____ cm _____ wks gestation 4. AC _____ cm _____ wks gestation 5. FL _____ cm _____ wks gestation _____ wks gestation	1. ___ 2. ___ 3. ___ 4. ___ 5. ___	_____ _____ _____ _____	___.___ ☐ Check if measurement is normal.	Fetal Tone _____/2 Gross body movements _____/2 Fetal breathing movements _____/2 Amniotic fluid volume _____/2
6.	__/__/__	Consistent with: 1. CRL _____ cm _____ wks gestation 2. BPD _____ cm _____ wks gestation 3. HC _____ cm _____ wks gestation 4. AC _____ cm _____ wks gestation 5. FL _____ cm _____ wks gestation _____ wks gestation	1. ___ 2. ___ 3. ___ 4. ___ 5. ___	_____ _____ _____ _____	___.___ ☐ Check if measurement is normal.	Fetal Tone _____/2 Gross body movements _____/2 Fetal breathing movements _____/2 Amniotic fluid volume _____/2
7.	__/__/__	Consistent with: 1. CRL _____ cm _____ wks gestation 2. BPD _____ cm _____ wks gestation 3. HC _____ cm _____ wks gestation 4. AC _____ cm _____ wks gestation 5. FL _____ cm _____ wks gestation _____ wks gestation	1. ___ 2. ___ 3. ___ 4. ___ 5. ___	_____ _____ _____ _____	___.___ ☐ Check if measurement is normal.	Fetal Tone _____/2 Gross body movements _____/2 Fetal breathing movements _____/2 Amniotic fluid volume _____/2

***Results/Interpretation of Ultrasound**

100=No structural abnormalities 101=Dates consistent with EDC

OB complications:
110=IUGR/SGA (Specify growth %-tile)
120=Macrosomia (Specify growth %-tile)
130=Oligohydramnios (AFI ≤ 5)
140=Polyhydramnios (AFI ≥ 25)
150=Cervical shortening (≤ 25 mm at ≤28 weeks)
160=Fetal arrhythmia
170=Fetal demise

Multiple gestations:
200=Twins - di/di
210=Twins - mono/di
220=Twins - mono/mono
230=Triplets or higher order gestation
240=Concordant growth (multiple gestation) -All ≥10%
250=Concordant growth multiple gestation) -All SGA (<10%)
260=Discordant growth (multiple gestation) – At least one SGA (<10%)
270=Percent discordant(Specify %)
(calculated by Big twin-Little Twin Weight/Big twin)

Structural anomalies:

Markers
301=Choroid Plexus Cyst (CPC)
302=Pyelectasis (4 -10 mm)
303=Echogenic Intra-Cardiac Focus (ECF)
304=Ventriculomegaly (≥10 mm)
305=Bowel hyperechoic
306=Shortened femur or humerus
307= Increased nuchal translucency (first trimester)
308= Thickened nuchal fold (>6 mm second trimester)

Congenital Anomaly
310=Congenital Anomaly
Under the specify column, note the code from the **congenital anomaly code list on the following page.**

Structural anomalies (Continued)
Abnormal placenta
321=Previa
322=Abruption / sub-chorionic hemorrhage
329=Other (specify)

Abnormal Cord insertion
331=Marginal
332=Velamentous
339=Other (specify)

Abnormal Umbilical cord
341=2VC
342=Nuchal
343=Varix
349=Other (specify)

350 = Amniotic Band

Other
499=Other (specify)

For Result Code 310: Specify Congenital Anomaly Code

100=Cystic Hygroma

Central Nervous System Defects
201=Open neural tube defect (Meningomyelocele, Spina bifida)
202=Anencephaly
203=Hydranencephaly
204=Hydrocephalus
205=Holoprosencephaly
299=Other (Specify)

Thorax Defects
301=Congenital Diaphragmatic Hernia
302=Cystic Adenomatoid Malformation (CAM)
303=Pulmonary Sequestration
399=Other (Specify)

Congenital Heart Defects
401=Atrioseptal defect (ASD)
402=Ventricular Septal Defect (VSD)
403=Atrioventricular Canal Defect (Endocardial Cushion Defect)
404=Transposition of the Great Vessels
405=Tetralogy of Fallot
499=Other Congenital Heart Defect (Specify)

Gastrointestinal defects
501=Cleft Palate
502=Gastroschisis
503=Omphalocele
504=Duodenal Atresia
599=Other Gastro-Intestinal Defect (Specify)

Genitourinary defects
601=Hydronephrosis/ureteropelvic junction obstruction
602=Autosomal Recessive Polycystic Kidney Disease
603=Mutlicystic/Dysplastic Kidney
604=Posterior Urethral valves (PUV)
605=Renal Agenesis
699=Other GU defect (Specify)

Skeletal defects
701=Skeletal dysplasia
702=Club feet
799=Other limb abnormalities (Specify)

Other
899=Other anomaly (Specify)

J. In-Patient Stay

Date	Highest Blood Pressures	On $MgSO_4$? (circle one)		
__/__/__	1. Sys ___ 2. Dia ___	yes no	**If yes:** Start date: Start time:	End date: End time
__/__/__	1. Sys ___ 2. Dia ___	yes no	**If yes:** Start date: Start time:	End date: End time
__/__/__	1. Sys ___ 2. Dia ___	yes no	**If yes:** Start date: Start time:	End date: End time
__/__/__	1. Sys ___ 2. Dia ___	yes no	**If yes:** Start date: Start time:	End date: End time

Appendix E

DNA and Primer Sequences for Validation

CpG Site DNA Sequence Data for Primer Design

RHPN1>hg19_dna range=chr8:144457427-144457927 5'pad=250 3'pad=250 strand=+ repeatMasking=none
CTGGCTTCTGGTGTCCTTGGCAGGTGCCAGCCTCCCCCGCTGACCCCCATCAC
GAGTCAGCAGCTTACCCCACCGACCACGTCCTTCTGCATTGACTGCCTCCTGT
CCTGCTCTGGCCAGGCCTGTGTTCACACTAGTTCTGTCCAGCCCCTCCCTGTG
AGGCCAGCTCCAGCCCCAGCGCATGGTGACCATCCCGTTACCCATGGGCAGG
ATGCACTCCTCTCAGTGGCTGGCGAGGCGCAGCCTGGTGCGGGCGCCACGGG
GTCGGGCTGTGATCGCCTGTGGCCTCCCTGCAGGGCTGTGACTCCCTGACGCA
GATCCAGTGCGGCCAGCTGCAGAGCCGCAGGGCCCAGATTCACCAGCAGATT
GACAAGGAGCTGCAGATGCGGACGGGCGCTGAGAACCTCTACAGGTCAGTG
CTTGAGACTGCCCGGCCCCGGGAGCAGGGCCCACCTGGGTGAGGGGGGCAG
GACAGCCACGCAGGCAGATGTCTGCCCCATG

HLA-DRB6>hg19_dna range=chr6:32522622-32523122 5'pad=250 3'pad=250 strand=+ repeatMasking=none
TCCTGACCATTCTGGAACCACCTGACTTTAATGCTGCCTGGATAGAAACCACT
CACAGAGCCGACCAGGGGGTTGCGGTGATGCAGGGGCTGGGTCTTTGCAGGA
TACACAGTCACCTTAGGATGGACTAGGAGAAAAAAAGGTAGAGAGAATGAA
TCAGGAAGTTAGAGTCTCGTTGTTCAGCTGTTTGTATGCTTCTCTGTAAACCC
AGGCTCTGGCCTCGACCAGGCCTCCAGCACAGCTGGCCATACGCCCTCACAG
TGTCATCGGCCTGGAATTTAATCGTGATAGTGTGGACCTATCAGATTTGAGAG
ATGTTATAAAAAATTTTATTTGTTTCTTCATAGCTTGAAATTGTCACGCATTGT
TGAAGTGTTTACAAATCTCTGAAAGTACAGTGTGTATTAATTAAAACTGATAC
CTGAGCCAGGTTGCCTGGTTCAAATCCAAGGTCTGCCTTTTACTGGTTGATCC
TGGAAGAGTTTTTTGATTCTTTTGTGT

ANXA2>hg19_dna range=chr15:60643907-60644407 5'pad=250 3'pad=250 strand=+ repeatMasking=none
GGCCACATTCACTTACCCAGGTTCAGGAAAGCATTTTCCAGGTCTCCTTTAAC
CTCTTTCCTGATGCTTTCCAACATGTCATAAGGGCTGTAACTCTTGTACCTATC
AAATACTGAGGAAAAACAACAAAGAGTTATCAGATCCGAGCCACTAGTCAA
AGCTGTCAACGATCACCCACCTAGTTTTATGCACCATAATTTTTTTAAAAATT
GAGGATGATCACAGCATCCTAGGAGCTTAGAGGTTACCACGGTGACCAGAGC
CAACATTGGCCAAGTTTGTCGTGGAACAGCCATACCACCTGTCCTGAATGGC
ACTGCCCAGGCCACATATTTGGACCATCTCTATCTCCCCTGAGTGGAACCCAT
TCCATCCGAAAACCATAGGAAACAGTACAGAGCATGCACCAAAGTCCACTAC
TTCAACAAATAATGGCAAGACCAAATGATCATCAAACAAGAAGGAGCTGCA
GAATAAAGCACCAAATGCAGAAACTATTTG

MRI1>hg19_dna range=chr19:13874611-13875611 5'pad=500 3'pad=500 strand=+ repeatMasking=none

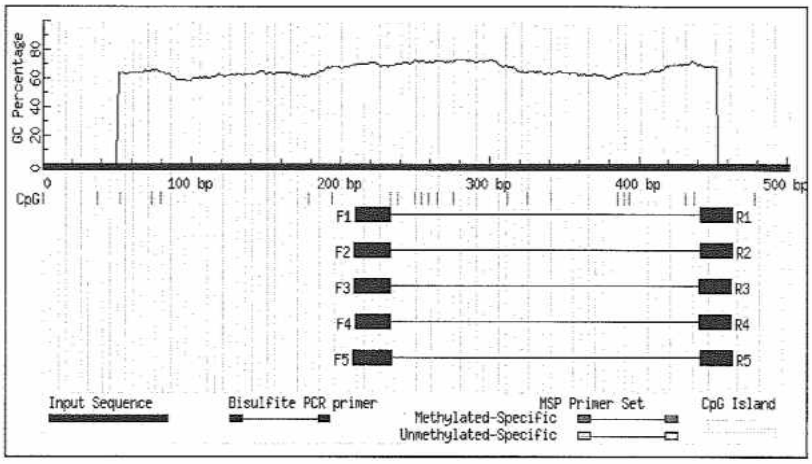
CCTCAGCCTCCCGAGCAGCTGGGACTACGGGTGTGCGCCACCACGCCCAGCT
AATTTTTTGTATTTTTAGTAGAGACTGGGTTTCACCATGTTGGCCAGGCTGGT
CTCGAACTCCTGACCTCAAGTGATCTGCCCGACTCGGCCTCCCAAAATGCTGG
GATTCCAGGCGTGAGCCACAGCGCCTGTCCTGCATGTTACTTTTGAATGAAAC
CGAGCAGAAAATGGCCCAGAAACAGCCTTGCATCCATCAAGGGGCACACGA
CCCCCCACTACCTCCCCCTCAACCTTGGAAGATCATTTAACAAATTCTTTGGT
TTGAACACTTGATGTTACCTTGCCACTGGGGATACATCCCTAACTCTAGACAG
CAGGTTGTTAAACACGGGGCCTGGTATCCACTAGGCGTCCCATAAATGCTGC
CACTTTTGTGGTTCCGAGGAGGCGGCTCACTCCGTTCGGGCTTGGCAGGAGTC
GTGGAGTGGGTTCGGCCACGTGGAATCCGCGTCCTGGGAACCCGTGGAATCC
GCGTCCTGGGAACCCGTGGAATCCGCCTCCTGGGAACCCGTGGAATCCGCCT
CCTGGGAACCCGTGGAATCCGCCTCCTGGGAACCCGTGGAATCCGCCTCCTG
GGAACCCGTGGAATCCGCCTCCTGGGAACCCGTGGAATCGGGTTGGATGCGC
ATGTGCGTGTCTCTTTTTCCGGGGGAGGCTCCGCCCACGGCCCCGCCCCGCTC
CCAAGTGCGCGCGGACCCCTAGCTCCCTCTGAGTTGCGCTGGGCTTGGCTGCT
GCACCATGACCCTGGAGGCGATCCGCTACTCGCGGGGCTCCCTGCAGATCCT
AGACCAGCTGCTGCTGCCCAAGCAGAGCCGCTACGAGGCGGTGGGCTCGGTG
CACCAGGCCTGGGAGGCCATCCGCGCCATGAAGGTGCAGCGGGGCGGCGGG
GCGGCGGGGCGGCGGGGCGGCGGGGCGGCGGGGCGGCGGGGCGGCGGGGC
GGCGGGGCG

GAP43>hg19_dna range=chr3:115376099-115376599 5'pad=250 3'pad=250 strand=+
repeatMasking=none
AGTGTAGGAGAGGTGAGTTGCTTAGGTCTAAGGAGAAAGACTGCTTAGGTGT
GTGTTCACCCCCAGGACGAAGAAAGGAACACTGGGTGAGATTTTGTTCAACT
ACCCATAGTTACCACCAGATGGTGAAACTGATCCCGGGCCTCTTGGGTATTG
ATCAGTTTATGGGGAGATGGGGAGAAGACTATCTTTCACTTGTTAATTCATTA
ATTTCTTTCGCAAATATTTTTTCAGTACCTGCTAAGTCCCACGGACTATGCTA
GGAGCTGCTGTTAAAATGACAAACCAGATAAGGTCACTGCCCTTAATCAACT
TACAGTTGGGTGAGAAGCTATCAGGTACAAGTATGGCCCTAGAACAAATTAG
TCTTTTCTAGTTAATAATCTTATGTGATGAGATTTGGCCTTGCTCCTTTGGTGA
CTTGCCTCAAGGAGCCCCAGGCAAAACCAATGTAACATATATTAATAATATA
TGAAATAATATATTTTGTAGACACAATTG

CUL3>hg19_dna range=chr2:225441582-225442082 5'pad=250 3'pad=250 strand=+
repeatMasking=none
ATCATATATGAACTTCTGTTTTTGAAGCCACCCCTCAAGAGCCAACAGGATTC
TTTAAGTATCCCAGTGGTACTAAACCCATATCCTTTGAGAATGCTTTCTTAGA
ACGATTCACAAACTGGCTCTGCAGGCTTTTCAAAACTTAAGTTCTAGAAGTTG
TACAATAAAATGACAGGATCGCTAAAATAAGTGTATGGCATTCGATGTAACT
GCTTGCAAAAACAACATCCAATTTTAATATTGGCCTAATCGTGGCTAAATATT
GGTATAATAATAGTTAACTATCTGCTAAGTTCTATTTTAAAGCTTTATTTTATT
TATCACACAACAAACCTGTAAGGTGGGTCTCATTAGCAGTCTCATTTTACACA
GGGGAAAACTGAGGCTACAAGAAGTAACTTGTTAAAGGTTATGCAGCTAGAG
GCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTGTAGGAGGATGAGGC
AGACGGATCACGAGGTCAGGAGATCG

MethPrimer result

Please cite MethPrimer: Li LC and Dahiya R. MethPrimer: designing primers for methylation PCRs. Bioinformatics. 2002 Nov;18(11):1427-31.
PMID: 12424112

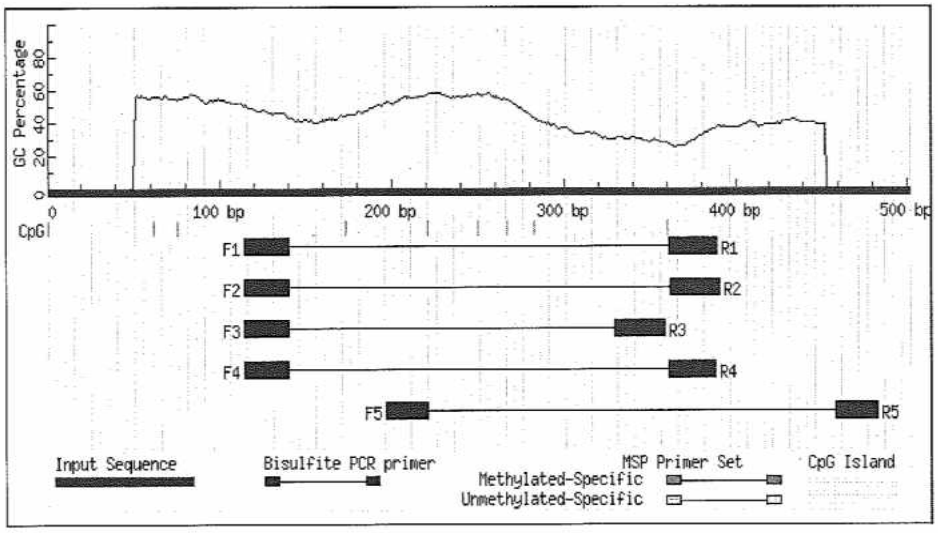



```
Sequence Name:
Sequence Length: 501

CpG island prediction results
(Criteria used: Island size > 100, GC Percent > 50.0, Obs/Exp > 0.6)
No CpG islands were found in your sequence
Primer picking results for bisulfite sequencing (or restriction) PCR

   Primer            Start Size  Tm      GC%   'C's  Sequence
1  Left  primer       210   24   58.03   62.50   7  GGATGTATTTTTTTTAGTGGTTGG
   Right primer       463   22   58.45   72.73   6  CCTCACCCAAATAAACCCTACT
   Product size: 254, Tm: 69.8, CpGs in product: 15

2  Left  primer       209   25   58.88   60.00   7  AGGATGTATTTTTTTTAGTGGTTGG
   Right primer       463   22   58.45   72.73   6  CCTCACCCAAATAAACCCTACT
   Product size: 255, Tm: 69.9, CpGs in product: 15

3  Left  primer       209   25   58.88   60.00   7  AGGATGTATTTTTTTTAGTGGTTGG
   Right primer       462   22   56.82   72.73   6  CTCACCCAAATAAACCCTACTC
   Product size: 254, Tm: 69.8, CpGs in product: 15

4  Left  primer       210   24   58.03   62.50   7  GGATGTATTTTTTTTAGTGGTTGG
   Right primer       462   22   56.82   72.73   6  CTCACCCAAATAAACCCTACTC
   Product size: 253, Tm: 69.7, CpGs in product: 15

5  Left  primer       208   26   58.92   61.54   8  TAGGATGTATTTTTTTTAGTGGTTGG
   Right primer       463   22   58.45   72.73   6  CCTCACCCAAATAAACCCTACT
   Product size: 256, Tm: 70.1, CpGs in product: 15


    1 CTGGCTTCTGGTGTCCTTGGCAGGTGCCAGCCTCCCCCGCTGACCCCCATCACGAGTCAG
      :|||:||:||||||::||||:|||||::||::|::::++:|||:::::|||:++|||:||
    1 TTGGTTTTTGGTGTTTTTGGTAGGTGTTAGTTTTTTTCGTTGATTTTTATTACGAGTTAG


   61 CAGCTTACCCCACCGACCACGTCCTTCTGCATTGACTGCCTCCTGTCCTGCTCTGGCCAG
```

```
      :||:|||::::|:++|::|++|::||:|:|||||:||::|::|||::||:|:|||::||
   61 TAGTTTATTTTATCGATTACGTTTTTTTGTATTGATTGTTTTTTGTTTTGTTTTGGTTAG


  121 GCCTGTGTTCACACTAGTTCTGTCCAGCCCCTCCCTGTGAGGCCAGCTCCAGCCCCAGCG
      |::||||||:|:|:|||||:|||::||::::|:::|||||||::||:|::||::::||++
  121 GTTTGTGTTTATATTAGTTTTGTTTAGTTTTTTTTTTGTGAGGTTAGTTTTAGTTTTAGCG


  181 CATGGTGACCATCCCGTTACCCATGGGCAGGATGCACTCCTCTCAGTGGCTGGCGAGGCG
      :||||||::||::++|||:::|||||:||||||:|:|::|:|:|||||:|||++|||++
  181 TATGGTGATTATTTCGTTATTTATGGGTAGGATGTATTTTTTTTAGTGGTTGGCGAGGCG
                                  >>>>>>>>>>>>>>>>>>>>>>>>


  241 CAGCCTGGTGCGGGCGCCACGGGGTCGGGCTGTGATCGCCTGTGGCCTCCCTGCAGGGCT
      :||::|||||++||++::|++||||++||:||||||++::|||||::|:::||:||||:|
  241 TAGTTTGGTGCGGGCGTTACGGGGTCGGGTTGTGATCGTTTGTGGTTTTTTTGTAGGGTT


  301 GTGACTCCCTGACGCAGATCCAGTGCGGCCAGCTGCAGAGCCGCAGGGCCCAGATTCACC
      ||||:|:::|||++:||||::|||||++||::||:||:||||:++:||||:::|||||:|::
  301 GTGATTTTTTGACGTAGATTTAGTGCGGTTAGTTGTAGAGTCGTAGGGTTTAGATTTATT


  361 AGCAGATTGACAAGGAGCTGCAGATGCGGACGGGCGCTGAGAACCTCTACAGGTCAGTGC
      ||:|||||||:||||||:||:|||||++||++||++:||||||::|:||:||||:||||:
  361 AGTAGATTGATAAGGAGTTGTAGATGCGGACGGGCGTTGAGAATTTTTATAGGTTAGTGT


  421 TTGAGACTGCCCGGCCCCGGGAGCAGGGCCCACCTGGGTGAGGGGGGCAGGACAGCCACG
      ||||||:||::++|:::++||||:||||:::|::||||||||||||||:||||:||::|++
  421 TTGAGATTGTTCGGTTTCGGGAGTAGGGTTTATTTGGGTGAGGGGGGTAGGATAGTTACG
                              <<<<<<<<<<<<<<<<<<<<<<


  481 CAGGCAGATGTCTGCCCCATG
      :|||:||||||:||::::|||
  481 TAGGTAGATGTTTGTTTTATG


**********************************************************************
*  Explanations                                                      *
*--------------------------------------------------------------------*
*  Upper row: Original sequence                                      *
*  Lower row: Bisulfite modified sequence                            *
*             (For display, assume all CpG sites are methylated)     *
*  ++         CpG sites                                              *
*  ::::       Non-CpG 'C' converted to 'T'                           *
*  >>>>>>     Left  primer                                           *
*  <<<<<<     Right primer                                           *
*                                                                    *
**********************************************************************



MethPrimer v1.1 beta

Li Lab, Department of Urology, UCSF
```

MethPrimer result

Please cite MethPrimer: Li LC and Dahiya R. MethPrimer: designing primers for methylation PCRs. Bioinformatics. 2002 Nov;18(11):1427-31.
PMID: 12424112




```
Sequence Name:
Sequence Length: 501

CpG island prediction results
(Criteria used: Island size > 100, GC Percent > 50.0, Obs/Exp > 0.6)
No CpG islands were found in your sequence
Primer picking results for bisulfite sequencing (or restriction) PCR

   Primer           Start Size  Tm      GC%    'C's  Sequence
1  Left  primer      114   26   57.95   42.31   4   TATTTTAGGATGGATTAGGAGAAAAA
   Right primer      389   28   56.17   35.71   4   CAAAAATTTATAAACACTTCAACAATAC
   Product size: 276, Tm: 69.9, CpGs in product: 6

2  Left  primer      114   26   57.95   42.31   4   TATTTTAGGATGGATTAGGAGAAAAA
   Right primer      390   29   57.60   34.48   4   TCAAAAATTTATAAACACTTCAACAATAC
   Product size: 277, Tm: 69.9, CpGs in product: 6

3  Left  primer      114   26   57.95   42.31   4   TATTTTAGGATGGATTAGGAGAAAAA
   Right primer      359   30   57.71   26.67   4   AACAATTTCAAACTATAAAAAAACAAATAA
   Product size: 246, Tm: 69.6, CpGs in product: 5

4  Left  primer      114   26   57.95   42.31   4   TATTTTAGGATGGATTAGGAGAAAAA
   Right primer      388   27   53.96   33.33   4   AAAAATTTATAAACACTTCAACAATAC
   Product size: 275, Tm: 69.9, CpGs in product: 6

5  Left  primer      197   25   56.79   56.00   9   TTTTTGTAAATTTAGGTTTTGGTTT
   Right primer      482   25   59.79   44.00   4   CTCTTCCAAAATCAACCAATAAAAA
```

Product size: 286, Tm: 69.8, CpGs in product: 5

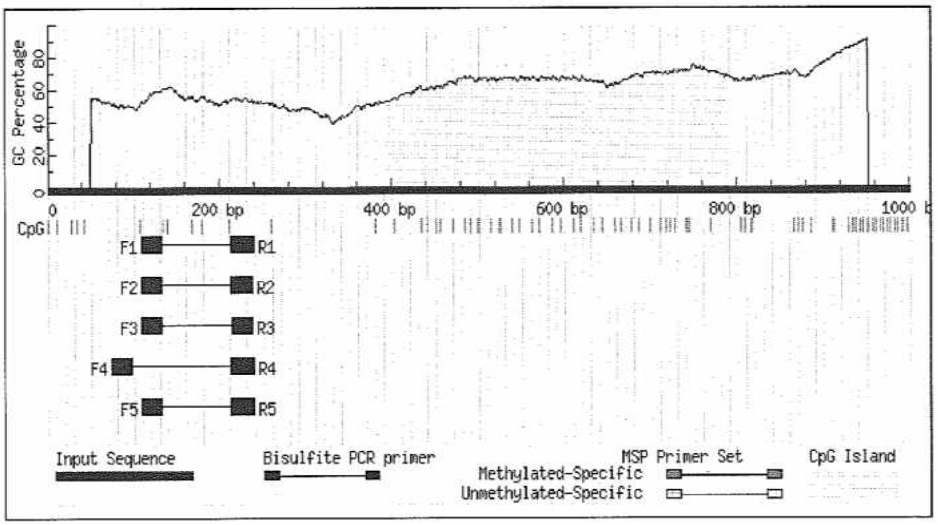
```
  1 TCCTGACCATTCTGGAACCACCTGACTTTAATGCTGCCTGGATAGAAACCACTCACAGAG
    |::|||::|||:|||||::|::|||:|||||||:||::|||||||||||::|:|:|:|||
  1 TTTTGATTATTTTGGAATTATTTGATTTTAATGTTGTTTGGATAGAAATTATTTATAGAG

 61 CCGACCAGGGGGTTGCGGTGATGCAGGGGCTGGGTCTTTGCAGGATACACAGTCACCTTA
    :++|::|||||||||++||||||:|||||:|||||:||||:||||||:|:|||:|::|||
 61 TCGATTAGGGGGTTGCGGTGATGTAGGGGTTGGGTTTTTGTAGGATATATAGTTATTTTA
                                                       >>>>>>>

121 GGATGGACTAGGAGAAAAAAAGGTAGAGAGAATGAATCAGGAAGTTAGAGTCTCGTTGTT
    |||||||:|||||||||||||||||||||||||||:||||||||||||||:|++|||||
121 GGATGGATTAGGAGAAAAAAAGGTAGAGAGAATGAATTAGGAAGTTAGAGTTTCGTTGTT
    >>>>>>>>>>>>>>>>>>>

181 CAGCTGTTTGTATGCTTCTCTGTAAACCCAGGCTCTGGCCTCGACCAGGCCTCCAGCACA
    :||:|||||||||||:||:|:||||||:::|||:|:|||::|++|::|||::|::||:|:|
181 TAGTTGTTTGTATGTTTTTTTGTAAATTTAGGTTTTGGTTTCGATTAGGTTTTTAGTATA

241 GCTGGCCATACGCCCTCACAGTGTCATCGGCCTGGAATTTAATCGTGATAGTGTGGACCT
    |:|||::|||++:::|:|:|||||:||++|::|||||||||||++|||||||||||::|
241 GTTGGTTATACGTTTTTATAGTGTTATCGGTTTGGAATTTAATCGTGATAGTGTGGATTT

301 ATCAGATTTGAGAGATGTTATAAAAAATTTTATTTGTTTCTTCATAGCTTGAAATTGTCA
    ||:||||||||||||||||||||||||||||||||:||:||||:||||||||||:|
301 ATTAGATTTGAGAGATGTTATAAAAAATTTTATTTGTTTTTTTATAGTTTGAAATTGTTA

361 CGCATTGTTGAAGTGTTTACAAATCTCTGAAAGTACAGTGTGTATTAATTAAAACTGATA
    ++:|||||||||||||||:||||:|:||||||||:|||||||||||||||||:|||||
361 CGTATTGTTGAAGTGTTTATAAATTTTTGAAAGTATAGTGTGTATTAATTAAAATTGATA
     <<<<<<<<<<<<<<<<<<<<<<<<<<<<<

421 CCTGAGCCAGGTTGCCTGGTTCAAATCCAAGGTCTGCCTTTTACTGGTTGATCCTGGAAG
    ::||||::||||||::||||||:||||::|||||:|||::|||||:||||||||:|||||
421 TTTGAGTTAGGTTGTTTGGTTTAAATTTAAGGTTTGTTTTTTATTGGTTGATTTTGGAAG

481 AGTTTTTTGATTCTTTTGTGT
    ||||||||||||:||||||||
481 AGTTTTTTGATTTTTTTGTGT
```

```
**********************************************************************
*  Explanations                                                      *
*--------------------------------------------------------------------*
*  Upper row: Original sequence                                      *
*  Lower row: Bisulfite modified sequence                            *
*             (For display, assume all CpG sites are methylated)     *
*  ++         CpG sites                                              *
*  ::::       Non-CpG 'C' converted to 'T'                           *
*  >>>>>>     Left  primer                                           *
```

```
*  <<<<<<     Right primer                                              *
*                                                                       *
************************************************************************


MethPrimer v1.1 beta

Li Lab, Department of Urology, UCSF
```

MethPrimer result

Please cite MethPrimer: Li LC and Dahiya R. MethPrimer: designing primers for methylation PCRs. Bioinformatics. 2002 Nov;18(11):1427-31.
PMID: 12424112




```
Sequence Name:
Sequence Length: 1001

CpG island prediction results
(Criteria used: Island size > 100, GC Percent > 50.0, Obs/Exp > 0.6)
 1 CpG island(s) were found in your sequence
           Size        (Start - End)
 Island 1   399 bp       (393 - 791)

Primer picking results for bisulfite sequencing (or restriction) PCR

   Primer           Start Size  Tm      GC%   'C's  Sequence
1  Left  primer       110   25   54.83   52.00   9  AATTTTTGATTTTAAGTGATTTGTT
   Right primer       239   27   54.57   51.85   7  AAACTATTTCTAAACCATTTTCTACTC
   Product size: 130, Tm: 67.3, CpGs in product: 5

2  Left  primer       110   25   54.83   52.00   9  AATTTTTGATTTTAAGTGATTTGTT
   Right primer       238   26   53.27   53.85   7  AACTATTTCTAAACCATTTTCTACTC
   Product size: 129, Tm: 67.3, CpGs in product: 5

3  Left  primer       110   25   54.83   52.00   9  AATTTTTGATTTTAAGTGATTTGTT
   Right primer       237   25   51.86   52.00   6  ACTATTTCTAAACCATTTTCTACTC
   Product size: 128, Tm: 67.4, CpGs in product: 5

4  Left  primer        75   25   57.24   56.00   6  GATTGGGTTTTATTATGTTGGTTAG
   Right primer       239   27   54.57   51.85   7  AAACTATTTCTAAACCATTTTCTACTC
   Product size: 165, Tm: 68.2, CpGs in product: 6
```

```
5  Left  primer      110   25   54.83   52.00   9  AATTTTTGATTTTAAGTGATTTGTT
   Right primer      240   28   55.76   50.00   7  AAAACTATTTCTAAACCATTTTCTACTC
   Product size: 131, Tm: 67.2, CpGs in product: 5


    1 CCTCAGCCTCCCGAGCAGCTGGGACTACGGGTGTGCGCCACCACGCCCAGCTAATTTTTT
      ::|:||::|::++||:||:||||||:||++||||||++::|::|++:::||:|||||||||
    1 TTTTAGTTTTTCGAGTAGTTGGGATTACGGGTGTGCGTTATTACGTTTAGTTAATTTTTT


   61 GTATTTTTAGTAGAGACTGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCCTGACC
      |||||||||||||||||:|||||||:|::|||||||::|||:||||:|++||:|::|||::
   61 GTATTTTTAGTAGAGATTGGGTTTTATTATGTTGGTTAGGTTGGTTTCGAATTTTTGATT
                                                          >>>>>>>>>>>


  121 TCAAGTGATCTGCCCGACTCGGCCTCCCAAAATGCTGGGATTCCAGGCGTGAGCCACAGC
      |:|||||||:||::++|:|++|::|:::||||||:|||||||::|||++||||::|:||+
  121 TTAAGTGATTTGTTCGATTCGGTTTTTTAAAATGTTGGGATTTTAGGCGTGAGTTATAGC
      >>>>>>>>>>>>>


  181 GCCTGTCCTGCATGTTACTTTTGAATGAAACCGAGCAGAAAATGGCCCAGAAACAGCCTT
      +::|||::||:||||||:|||||||||||||:++||:|||||||||:::|||||:|||::||
  181 GTTTGTTTTGTATGTTATTTTTGAATGAAATCGAGTAGAAAATGGTTTAGAAATAGTTTT
                                      <<<<<<<<<<<<<<<<<<<<<<<<<<<<


  241 GCATCCATCAAGGGGCACACGACCCCCCACTACCTCCCCCTCAACCTTGGAAGATCATTT
      |:||::||:||||||:|:|++|:::::::|:||::|:::::|:||::|||||||||:||||
  241 GTATTTATTAAGGGGTATACGATTTTTTATTATTTTTTTTTTAATTTTGGAAGATTATTT


  301 AACAAATTCTTTGGTTTGAACACTTGATGTTACCTTGCCACTGGGGATACATCCCTAACT
      ||:||||||:|||||||||||:|:|:|||||||||::|||::|:|||||||||:||:::|||:|
  301 AATAAATTTTTTGGTTTGAATATTTGATGTTATTTTGTTATTGGGGATATATTTTTAATT


  361 CTAGACAGCAGGTTGTTAAACACGGGGCCTGGTATCCACTAGGCGTCCCATAAATGCTGC
      :||||:||:|||||||||||:|++|||::||||||::|:||||++|:::|||||||:||:
  361 TTAGATAGTAGGTTGTTAAATACGGGGTTTGGTATTTATTAGGCGTTTTATAAATGTTGT


  421 CACTTTTGTGGTTCCGAGGAGGCGGCTCACTCCGTTCGGGCTTGGCAGGAGTCGTGGAGT
      :|:|||||||||:++||||||++|:|:|:++||++||:|||||:||||||++||||||
  421 TATTTTTGTGGTTTCGAGGAGGCGGTTTATTTCGTTCGGGTTTGGTAGGAGTCGTGGAGT


  481 GGGTTCGGCCACGTGGAATCCGCGTCCTGGGAACCCGTGGAATCCGCGTCCTGGGAACCC
      |||||++|::|++||||||:++++|::|||||||::++|||||:++++|::|||||||::+
  481 GGGTTCGGTTACGTGGAATTCGCGTTTTGGGAATTCGTGGAATTCGCGTTTTGGGAATTC


  541 GTGGAATCCGCCTCCTGGGAACCCGTGGAATCCGCCTCCTGGGAACCCGTGGAATCCGCC
      +||||||:++::|::||||||::++||||||:++::|::||||||::++||||||:++::
  541 GTGGAATTCGTTTTTTGGGAATTCGTGGAATTCGTTTTTTGGGAATTCGTGGAATTCGTT


  601 TCCTGGGAACCCGTGGAATCCGCCTCCTGGGAACCCGTGGAATCCGCCTCCTGGGAACCC
```

```
     |::||||||::++||||||:++::|::||||||::++||||||:++::|::||||||::+
601 TTTTGGGAATTCGTGGAATTCGTTTTTTGGGAATTCGTGGAATTCGTTTTTTGGGAATTC


661 GTGGAATCGGGTTGGATGCGCATGTGCGTGTCTCTTTTTCCGGGGGAGGCTCCGCCCACG
    +||||||++|||||||||++:|||||++|||:|:|||||:++|||||||:|:++:::|++
661 GTGGAATCGGGTTGGATGCGTATGTGCGTGTTTTTTTTTTCGGGGGAGGTTTCGTTTACG


721 GCCCCGCCCCGCTCCCAAGTGCGCGCGGACCCCTAGCTCCCTCTGAGTTGCGCTGGGCTT
    |:::++:::++:|:::|||||++++++||::::|||:|:::|:|||||||++:||||:||
721 GTTTCGTTTCGTTTTTAAGTGCGCGCGGATTTTTAGTTTTTTTTGAGTTGCGTTGGGTTT


781 GGCTGCTGCACCATGACCCTGGAGGCGATCCGCTACTCGCGGGGCTCCCTGCAGATCCTA
    ||:||:||:|::||||:::||||||++||:++:||:|++++|||:|:::||:||||::||
781 GGTTGTTGTATTATGATTTTGGAGGCGATTCGTTATTCGCGGGGTTTTTTGTAGATTTTA


841 GACCAGCTGCTGCTGCCCAAGCAGAGCCGCTACGAGGCGGTGGGCTCGGTGCACCAGGCC
    ||::||:||:||:||:::|||:||||:++:||++|||++|||||:|++|||:|::|||::
841 GATTAGTTGTTGTTGTTTAAGTAGAGTCGTTACGAGGCGGTGGGTTCGGTGTATTAGGTT


901 TGGGAGGCCATCCGCGCCATGAAGGTGCAGCGGGGCGGCGGGGCGGCGGGGCGGCGGGGC
    |||||||::||:++++::|||||||||:|||++|||++|++|||++|++|||++|++|+
901 TGGGAGGTTATTCGCGTTATGAAGGTGTAGCGGGGCGGCGGGGCGGCGGGGCGGCGGGGC


961 GGCGGGGCGGCGGGGCGGCGGGGCGGCGGGGCGGCGGGGCG
    +|++|||++|++|||++|++|||++|++|||++|++|||++
961 GGCGGGGCGGCGGGGCGGCGGGGCGGCGGGGCGGCGGGGCG
```

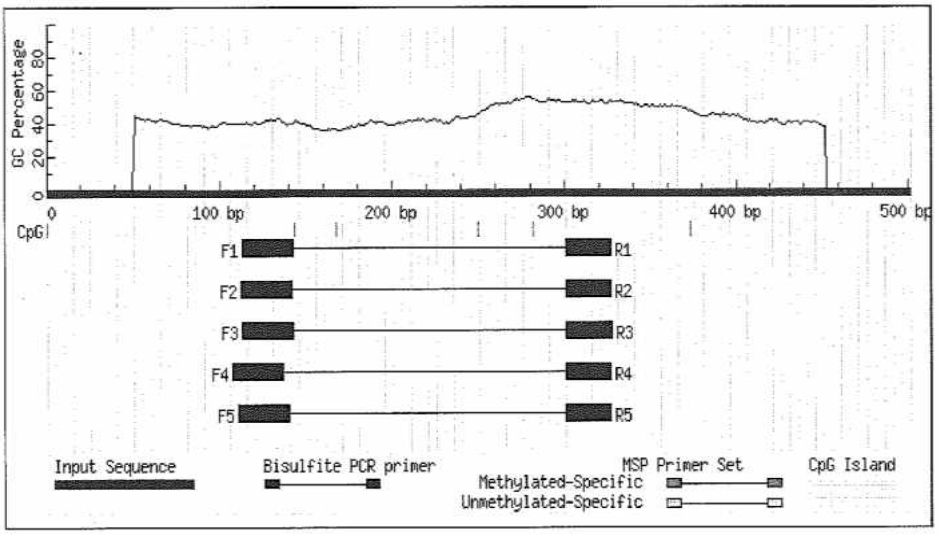
```
**********************************************************************
*  Explanations                                                      *
*--------------------------------------------------------------------*
*  Upper row: Original sequence                                      *
*  Lower row: Bisulfite modified sequence                            *
*             (For display, assume all CpG sites are methylated)     *
*  ++         CpG sites                                              *
*  ::::       Non-CpG 'C' converted to 'T'                           *
*  >>>>>>     Left  primer                                           *
*  <<<<<<     Right primer                                           *
*                                                                    *
**********************************************************************
```

MethPrimer v1.1 beta

Li Lab, Department of Urology, UCSF

MethPrimer result

Please cite MethPrimer: Li LC and Dahiya R. MethPrimer: designing primers for methylation PCRs. Bioinformatics. 2002 Nov;18(11):1427-31.
PMID: 12424112




```
Sequence Name:
Sequence Length: 501

CpG island prediction results
(Criteria used: Island size > 100, GC Percent > 50.0, Obs/Exp > 0.6)
No CpG islands were found in your sequence
Primer picking results for bisulfite sequencing (or restriction) PCR

   Primer            Start Size  Tm      GC%    'C's  Sequence
1  Left  primer       113   30   56.56   33.33   4   TTGAGGAAAAATAATAAAGAGTTATTAGAT
   Right primer       327   26   59.23   65.38  10   AACCTAAACAATACCATTCAAAACAA
   Product size: 215, Tm: 64.0, CpGs in product: 4

2  Left  primer       112   30   56.56   33.33   4   ATTGAGGAAAAATAATAAAGAGTTATTAGA
   Right primer       327   26   59.23   65.38  10   AACCTAAACAATACCATTCAAAACAA
   Product size: 216, Tm: 64.0, CpGs in product: 4

3  Left  primer       113   30   56.56   33.33   4   TTGAGGAAAAATAATAAAGAGTTATTAGAT
   Right primer       328   27   59.26   62.96  10   TAACCTAAACAATACCATTCAAAACAA
   Product size: 216, Tm: 64.0, CpGs in product: 4

4  Left  primer       107   30   55.61   30.00   4   TAAATATTGAGGAAAAATAATAAAGAGTTA
   Right primer       327   26   59.23   65.38  10   AACCTAAACAATACCATTCAAAACAA
   Product size: 221, Tm: 64.1, CpGs in product: 4

5  Left  primer       111   30   55.34   33.33   4   TATTGAGGAAAAATAATAAAGAGTTATTAG
   Right primer       327   26   59.23   65.38  10   AACCTAAACAATACCATTCAAAACAA
```

```
Product size: 217, Tm: 63.9, CpGs in product: 4


  1 GGCCACATTCACTTACCCAGGTTCAGGAAAGCATTTTCCAGGTCTCCTTTAACCTCTTTC
    ||::|:|||:|:|||:::|||||:|||||||:|||||::||||:|::|||||::|:|||:
  1 GGTTATATTTATTTATTTAGGTTTAGGAAAGTATTTTTTAGGTTTTTTTTAATTTTTTTT


 61 CTGATGCTTTCCAACATGTCATAAGGGCTGTAACTCTTGTACCTATCAAATACTGAGGAA
    :|||||:|||::||:||||:|||||||:|||||:|:|||||::|||:|||||:|||||||
 61 TTGATGTTTTTTAATATGTTATAAGGGTTGTAATTTTTGTATTTATTAAATATTGAGGAA
                                                      >>>>>>>>


121 AAACAACAAAGAGTTATCAGATCCGAGCCACTAGTCAAAGCTGTCAACGATCACCCACCT
    |||:||:||||||||||:||||:++||::|:||||:||||:|||:||++||:|:::|::|
121 AAATAATAAAGAGTTATTAGATTCGAGTTATTAGTTAAAGTTGTTAACGATTATTTATTT
    >>>>>>>>>>>>>>>>>>>>


181 AGTTTTATGCACCATAATTTTTTTAAAAATTGAGGATGATCACAGCATCCTAGGAGCTTA
    |||||||||:|::|||||||||||||||||||||||||||:|:||:||::||||||:|||
181 AGTTTTATGTATTATAATTTTTTTTAAAAATTGAGGATGATTATAGTATTTTAGGAGTTTA


241 GAGGTTACCACGGTGACCAGAGCCAACATTGGCCAAGTTTGTCGTGGAACAGCCATACCA
    |||||||::|++||||::||||::||:|||||::||||||||++|||||:||::|||::|
241 GAGGTTATTACGGTGATTAGAGTTAATATTGGTTAAGTTTGTCGTGGAATAGTTATATTA


301 CCTGTCCTGAATGGCACTGCCCAGGCCACATATTTGGACCATCTCTATCTCCCCTGAGTG
    ::|||::|||||||:|:||:::|||::|:|||||||||::||:|:|||:|::::|||||
301 TTTGTTTTGAATGGTATTGTTTAGGTTATATATTTGGATTATTTTTATTTTTTTTGAGTG
     <<<<<<<<<<<<<<<<<<<<<<<<<<<


361 GAACCCATTCCATCCGAAAACCATAGGAAACAGTACAGAGCATGCACCAAAGTCCACTAC
    |||:::|||::||:++||||::|||||||||:||||:||||:|||:|::||||||::|:||:
361 GAATTTATTTTATTCGAAAATTATAGGAAATAGTATAGAGTATGTATTAAAGTTTATTAT


421 TTCAACAAATAATGGCAAGACCAAATGATCATCAAACAAGAAGGAGCTGCAGAATAAAGC
    ||:||:|||||||||:||||::|||||||:||:|||:|||||||||:||:||||||||:
421 TTTAATAAATAATGGTAAGATTAAATGATTATTAAATAAGAAGGAGTTGTAGAATAAAGT


481 ACCAAATGCAGAAACTATTTG
    |::|||||:|||||:||||||
481 ATTAAATGTAGAAATTATTTG


**********************************************************************
*  Explanations                                                      *
*--------------------------------------------------------------------*
*  Upper row: Original sequence                                      *
*  Lower row: Bisulfite modified sequence                            *
*             (For display, assume all CpG sites are methylated)     *
*  ++         CpG sites                                              *
*  ::::       Non-CpG 'C' converted to 'T'                           *
*  >>>>>>     Left  primer                                           *
```

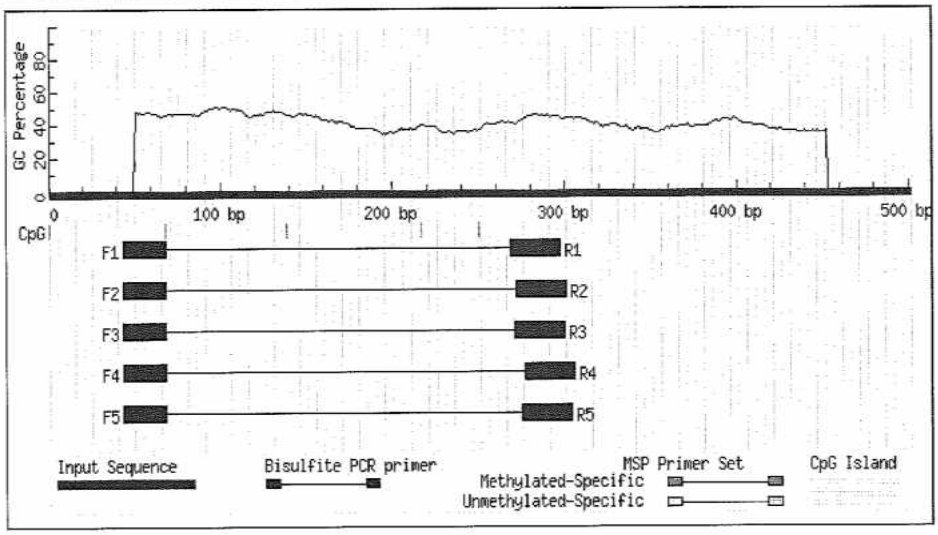
```
*  <<<<<<     Right primer                                                *
*                                                                         *
***************************************************************************
```

MethPrimer v1.1 beta

Li Lab, Department of Urology, UCSF

MethPrimer result

Please cite MethPrimer: Li LC and Dahiya R. MethPrimer: designing primers for methylation PCRs. Bioinformatics. 2002 Nov;18(11):1427-31.
PMID: 12424112




```
Sequence Name:
Sequence Length: 501

CpG island prediction results
(Criteria used: Island size > 100, GC Percent > 50.0, Obs/Exp > 0.6)
No CpG islands were found in your sequence
Primer picking results for bisulfite sequencing (or restriction) PCR

   Primer            Start Size  Tm      GC%   'C's  Sequence
1  Left  primer       44   25   56.36   56.00   7   TTTAGGTGTGTGTTTATTTTTAGGA
   Right primer      298   30   55.74   36.67   5   TAACCTTATCTAATTTATCATTTTAACAAC
   Product size: 255, Tm: 68.8, CpGs in product: 4

2  Left  primer       44   25   56.36   56.00   7   TTTAGGTGTGTGTTTATTTTTAGGA
   Right primer      301   30   55.41   36.67   5   CAATAACCTTATCTAATTTATCATTTTAAC
   Product size: 258, Tm: 68.7, CpGs in product: 4

3  Left  primer       44   25   56.36   56.00   7   TTTAGGTGTGTGTTTATTTTTAGGA
   Right primer      300   30   55.41   33.33   5   AATAACCTTATCTAATTTATCATTTTAACA
   Product size: 257, Tm: 68.7, CpGs in product: 4

4  Left  primer       44   25   56.36   56.00   7   TTTAGGTGTGTGTTTATTTTTAGGA
   Right primer      306   30   56.35   43.33   8   AAAAACAATAACCTTATCTAATTTATCATT
   Product size: 263, Tm: 68.8, CpGs in product: 4

5  Left  primer       44   25   56.36   56.00   7   TTTAGGTGTGTGTTTATTTTTAGGA
   Right primer      304   30   56.35   43.33   8   AAACAATAACCTTATCTAATTTATCATTTT
```

Product size: 261, Tm: 68.8, CpGs in product: 4

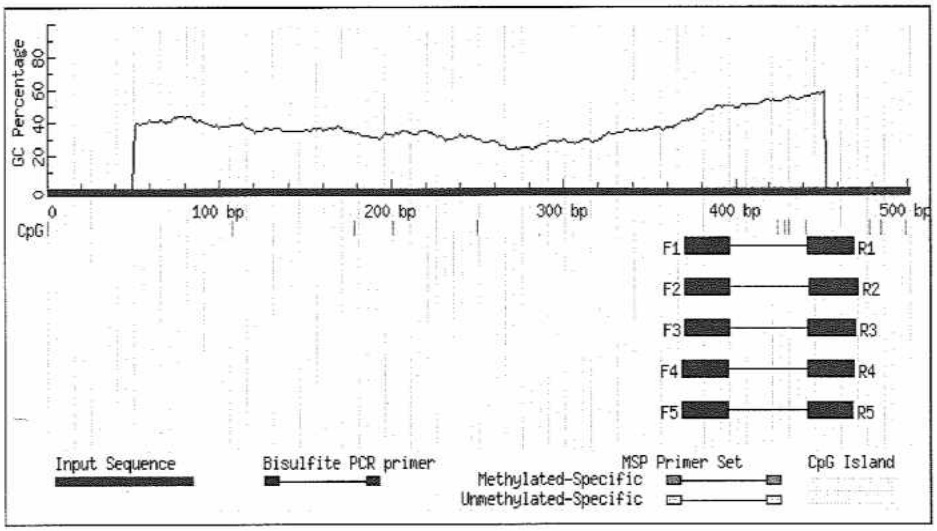
```
  1 AGTGTAGGAGAGGTGAGTTGCTTAGGTCTAAGGAGAAAGACTGCTTAGGTGTGTGTTCAC
    ||||||||||||||||||:||||||:||||||||||||:|!:|||||||||||||:|:
  1 AGTGTAGGAGAGGTGAGTTGTTTAGGTTTAAGGAGAAAGATTGTTTAGGTGTGTGTTTAT
                                                >>>>>>>>>>>>>>>>>

 61 CCCCAGGACGAAGAAAGGAACACTGGGTGAGATTTTGTTCAACTACCCATAGTTACCACC
    ::::||||++|||||||||||:|:|||||||||||||||:||:||:::|||||||::|::
 61 TTTTAGGACGAAGAAAGGAATATTGGGTGAGATTTTGTTTAATTATTTATAGTTATTATT
    >>>>>>>>

121 AGATGGTGAAACTGATCCCGGGCCTCTTGGGTATTGATCAGTTTATGGGGAGATGGGGAG
    ||||||||||:||||::++||::|:||||||||||||:||||||||||||||||||||
121 AGATGGTGAAATTGATTTCGGGTTTTTTGGGTATTGATTAGTTTATGGGGAGATGGGGAG

181 AAGACTATCTTTCACTTGTTAATTCATTAATTTCTTTCGCAAATATTTTTTCAGTACCTG
    ||||:|||:|||:|:||||||||:||||||||:|||++:|||||||||||:||||::||
181 AAGATTATTTTTTATTTGTTAATTTATTAATTTTTTTCGTAAATATTTTTTTTAGTATTTG

241 CTAAGTCCCACGGACTATGCTAGGAGCTGCTGTTAAAATGACAAACCAGATAAGGTCACT
    :|||||:::|++||:||||:|||||:||:||||||||||||:|||::||||||||||:|:|
241 TTAAGTTTTACGGATTATGTTAGGAGTTGTTGTTAAAATGATAAATTAGATAAGGTTATT
                                  <<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<

301 GCCCTTAATCAACTTACAGTTGGGTGAGAAGCTATCAGGTACAAGTATGGCCCTAGAACA
    |:::||||||:||:|||:|||||||||||||||:|||:|||||:||||||||:::||||||:|
301 GTTTTTAATTAATTTATAGTTGGGTGAGAAGTTATTAGGTATAAGTATGGTTTTAGAATA

361 AATTAGTCTTTTCTAGTTAATAATCTTATGTGATGAGATTTGGCCTTGCTCCTTTGGTGA
    |||||||:||||:||||||||||||:|||||||||||||||||||::|||:|::||||||||
361 AATTAGTTTTTTTTAGTTAATAATTTTATGTGATGAGATTTGGTTTTGTTTTTTTGGTGA

421 CTTGCCTCAAGGAGCCCCAGGCAAAACCAATGTAACATATATTAATAATATATGAAATAA
    :|||::|:||||||::::|||:||||::|||||||:|||||||||||||||||||||||
421 TTTGTTTTAAGGAGTTTTAGGTAAAATTAATGTAATATATATTAATAATATATGAAATAA

481 TATATTTTGTAGACACAATTG
    ||||||||||||||:|:||||||
481 TATATTTTGTAGATATAATTG
```

```
**********************************************************************
*  Explanations                                                      *
*--------------------------------------------------------------------*
*  Upper row: Original sequence                                      *
*  Lower row: Bisulfite modified sequence                            *
*             (For display, assume all CpG sites are methylated)     *
*  ++         CpG sites                                              *
*  ::::       Non-CpG 'C' converted to 'T'                           *
*  >>>>>>     Left  primer                                           *
```

```
*  <<<<<<     Right primer                                              *
*                                                                       *
************************************************************************


MethPrimer v1.1 beta

Li Lab, Department of Urology, UCSF
```

MethPrimer result

Please cite MethPrimer: Li LC and Dahiya R. MethPrimer: designing primers for methylation PCRs. Bioinformatics. 2002 Nov;18(11):1427-31.
PMID: 12424112




```
Sequence Name:
Sequence Length: 501

CpG island prediction results
(Criteria used: Island size > 100, GC Percent > 50.0, Obs/Exp > 0.6)
No CpG islands were found in your sequence
Primer picking results for bisulfite sequencing (or restriction) PCR

   Primer            Start Size  Tm      GC%    'C's  Sequence
1  Left  primer       370   26   59.41   50.00   4   TAGGGGAAAATTGAGGTTATAAGAAG
   Right primer       469   27   55.81   55.56   7   TCCTCCTACAATACTAAAATTACAAAC
   Product size: 100, Tm: 60.1, CpGs in product: 4

2  Left  primer       370   26   59.41   50.00   4   TAGGGGAAAATTGAGGTTATAAGAAG
   Right primer       471   29   58.29   55.17   7   CATCCTCCTACAATACTAAAATTACAAAC
   Product size: 102, Tm: 60.0, CpGs in product: 4

3  Left  primer       370   26   59.41   50.00   4   TAGGGGAAAATTGAGGTTATAAGAAG
   Right primer       470   28   56.18   53.57   7   ATCCTCCTACAATACTAAAATTACAAAC
   Product size: 101, Tm: 60.1, CpGs in product: 4

4  Left  primer       369   27   59.64   48.15   4   ATAGGGGAAAATTGAGGTTATAAGAAG
   Right primer       469   27   55.81   55.56   7   TCCTCCTACAATACTAAAATTACAAAC
   Product size: 101, Tm: 60.1, CpGs in product: 4

5  Left  primer       369   27   59.64   48.15   4   ATAGGGGAAAATTGAGGTTATAAGAAG
   Right primer       468   26   54.24   57.69   7   CCTCCTACAATACTAAAATTACAAAC
```

```
Product size: 100, Tm: 60.1, CpGs in product: 4


  1 ATCATATATGAACTTCTGTTTTTGAAGCCACCCCTCAAGAGCCAACAGGATTCTTTAAGT
    ||:|||||||||:||:||||||||||||::|::::|:||||||::||:|||||||:|||||||
  1 ATTATATATGAATTTTTGTTTTTGAAGTTATTTTTTAAGAGTTAATAGGATTTTTTAAGT


 61 ATCCCAGTGGTACTAAACCCATATCCTTTGAGAATGCTTTCTTAGAACGATTCACAAACT
    ||:::|||||||:||||:::||||::|||||||||||:|||:||||||++|||:|:|||:|
 61 ATTTTAGTGGTATTAAATTTATATTTTTTTGAGAATGTTTTTTTAGAACGATTTATAAATT


121 GGCTCTGCAGGCTTTTCAAAACTTAAGTTCTAGAAGTTGTACAATAAAATGACAGGATCG
    ||:|:|||:|||:||||:||||:|||||||:|||||||||||:|||||||||||:|||||++
121 GGTTTTGTAGGTTTTTTAAAATTTAAGTTTTAGAAGTTGTATAATAAAATGATAGGATCG


181 CTAAAATAAGTGTATGGCATTCGATGTAACTGCTTGCAAAAACAACATCCAATTTTAATA
    :|||||||||||||||:|||++||||||:||:|||:|||||:||:||::||||||||||
181 TTAAAATAAGTGTATGGTATTCGATGTAATTGTTTGTAAAAATAATATTTAATTTTAATA


241 TTGGCCTAATCGTGGCTAAATATTGGTATAATAATAGTTAACTATCTGCTAAGTTCTATT
    ||||::||||++|||:|||||||||||||||||||||||||:|||:||:||||||:||||
241 TTGGTTTAATCGTGGTTAAATATTGGTATAATAATAGTTAATTATTTGTTAAGTTTTATT


301 TTAAAGCTTTATTTTATTTATCACACAACAAACCTGTAAGGTGGGTCTCATTAGCAGTCT
    ||||||:|||||||||||||||:|:|:||:|||::|||||||||||:|:|||||:|||:|
301 TTAAAGTTTTATTTTATTTATTATATAATAAATTTGTAAGGTGGGTTTTATTAGTAGTTT


361 CATTTTACACAGGGGAAAACTGAGGCTACAAGAAGTAACTTGTTAAAGGTTATGCAGCTA
    :||||||:|:|||||||||:|||||:||:|||||||||:|||||||||||||||:||:||
361 TATTTTATATAGGGGAAAATTGAGGTTATAAGAAGTAATTTGTTAAAGGTTATGTAGTTA
             >>>>>>>>>>>>>>>>>>>>>>>>>>>


421 GAGGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTGTAGGAGGATGAGGCAGACG
    ||||:++||++++||||:|:|++::||||||:::||:|:|||||||||||||||:|||++
421 GAGGTCGGGCGCGGTGGTTTACGTTTGTAATTTTAGTATTGTAGGAGGATGAGGTAGACG
                          <<<<<<<<<<<<<<<<<<<<<<<<<<<<<


481 GATCACGAGGTCAGGAGATCG
    |||:|++||||:|||||||++
481 GATTACGAGGTTAGGAGATCG


**********************************************************************
*  Explanations                                                      *
*--------------------------------------------------------------------*
*  Upper row: Original sequence                                      *
*  Lower row: Bisulfite modified sequence                            *
*             (For display, assume all CpG sites are methylated)     *
*  ++         CpG sites                                              *
*  ::::       Non-CpG 'C' converted to 'T'                           *
*  >>>>>>     Left  primer                                           *
```

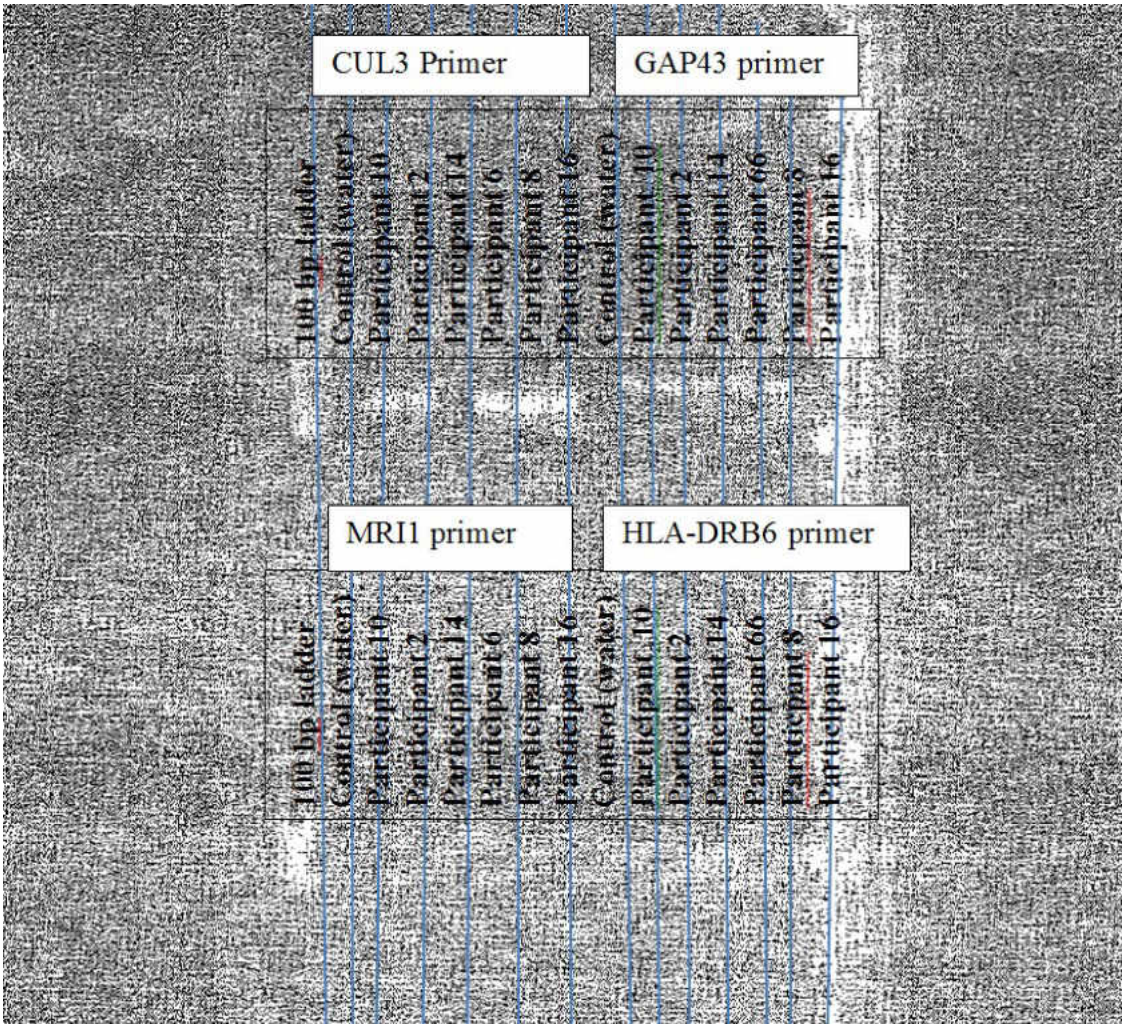
```
*  <<<<<<     Right primer                                              *
*                                                                       *
*************************************************************************
```

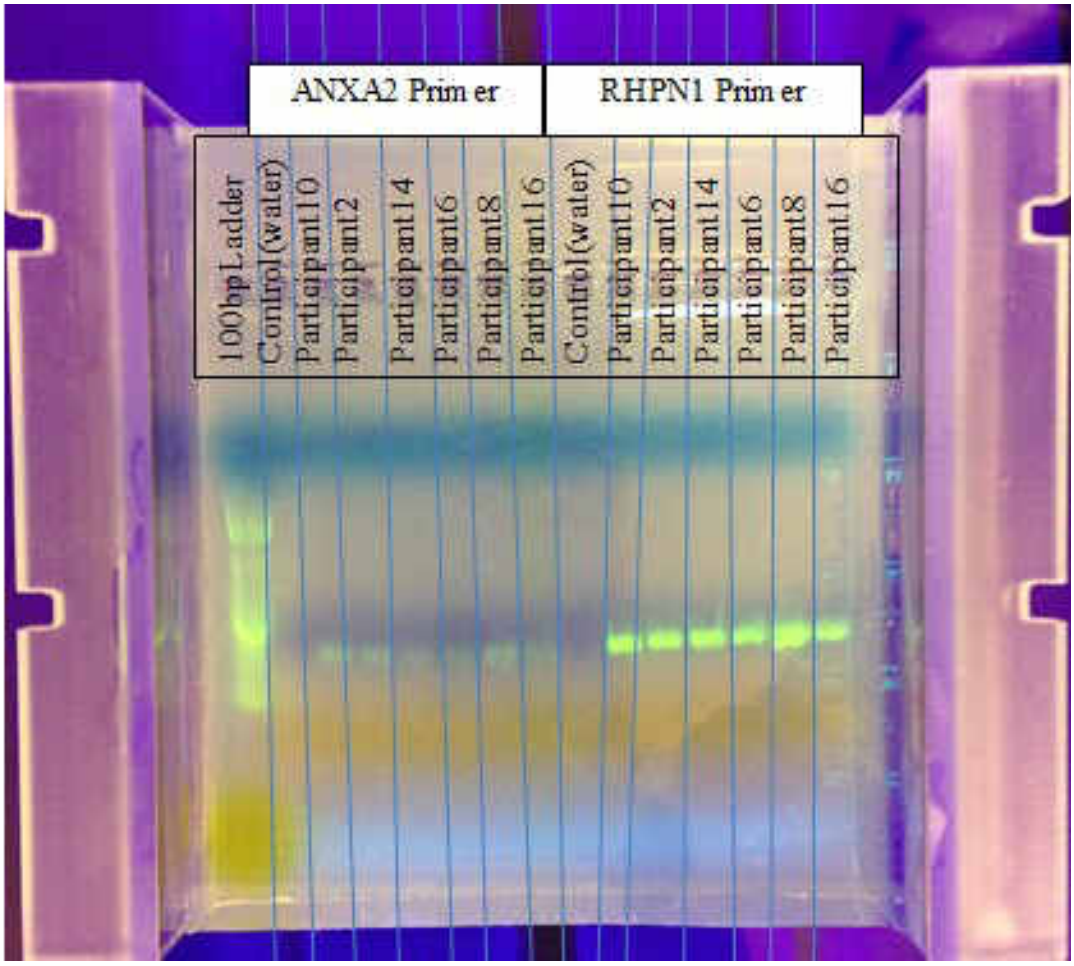
MethPrimer v1.1 beta

Li Lab, Department of Urology, UCSF

Appendix F

PCR product Images


CUL3 Primer
GAP43 primer
100 bp Ladder
Control (water)
Participant 10
Participant 2
Participant 14
Participant 6
Participant 8
Participant 16
Control (water)
Participant 10
Participant 2
Participant 14
Participant 66
Participant 8
Participant 16
MRI1 primer
HLA-DRB6 primer
100 bp Ladder
Control (water)
Participant 10
Participant 2
Participant 14
Participant 6
Participant 8
Participant 16
Control (water)
Participant 10
Participant 2
Participant 14
Participant 66
Participant 8
Participant 16


ANXA2 Primer
RHPN1 Primer
100bpLadder
Control(water)
Participant10
Participant2
Participant14
Participant6
Participant8
Participant16
Control(water)
Participant10
Participant2
Participant14
Participant6
Participant8
Participant16

Appendix G

Samples of Inadequate Quality for Sequencing

RHPN1

CGAGGCGCAGCCTGGTGCGGGCGCCACGGGGTCGGGCTGTGATCGCCTGTGG
CCTCCCTGCAGGGCTGTGACTCCCTGACGCAGATCCAGTGCGGCCAGCTGCA
GAGCCGCAGGGCCCAGATTCACCAGCAGATTGACAAGGAGCTGCAGATGCG
GACGGGCGCTGAGAACCTCTACAGGTCAGTGCTTGAGACTGCCCGGCCCCGG

excluded: RHPN16216FWD-RHPN1FWD Sequence identity: 89.2%
excluded: RHPN16216REV-RHPN1REV Sequence identity: 89.2%
excluded: RHPN16217FWD-RHPN1REV Sequence identity: 87.2%
excluded: RHPN16228FWD-RHPN1FWD Sequence identity: 87.7%
excluded: RHPN16228REV-RHPN1REV Sequence identity: 89.5%
excluded: RHPN16235FWD-RHPN1FWD Sequence identity: 88.6%
excluded: RHPN16235REV-RHPN1REV Sequence identity: 15.4%
excluded: RHPN16252FWD-RHPN1FWD Sequence identity: 88.2%
excluded: RHPN16259FWD-RHPN1FWD Sequence identity: 76.6%
excluded: RHPN16259REV-RHPN1REV Sequence identity: 86.3%
excluded: RHPN16217FWD-RHPN1FWD Sequence identity: 90.2%; Conversion rate: 100.0%; N-sites at non CpG cytosines positions: 5.9%; N-sites at CpG positions: 28.6%; Gaps: 1.4%

1 out of **12** uploaded sequencing results pass the quality criteria when compared to the reference sequence. At the next step all sequences are compared against all others to detect clonal amplifications as described in the manual.

Sequence Alignment of the sequences included:

RHPN16252REV-RHPN1RE

ANXA

CCGAGCCACTAGTCAAAGCTGTCAACGATCACCCACCTAGTTTTATGCACCAT
AATTTTTTTAAAAATTGAGGATGATCACAGCATCCTAGGAGCTTAGAGGTTAC
CACGGTGACCAGAGCCAACATTGGCCAAGTTTGTCGTGGAACAGCCATACCA
C

excluded: ANXA26216FWD-ANXA2FWD Sequence identity: 65.0%
excluded: ANXA26216FWD-ANXA2FWD_R Sequence identity: 71.5%
excluded: ANXA26216REV-ANXA2REV Sequence identity: 80.9%
excluded: ANXA26216REV-ANXA2REV_R Sequence identity: 64.6%
excluded: ANXA26217FWD-ANXA2FWD Sequence identity: 68.8%
excluded: ANXA26217FWD-ANXA2REV Sequence identity: 79.7%
excluded: ANXA26228FWD-ANXA2FWD Sequence identity: 67.1%
excluded: ANXA26228REV-ANXA2REV Sequence identity: 69.7%
excluded: ANXA26235FWD-ANXA2FWD Sequence identity: 78.5%
excluded: ANXA26235REV-ANXA2REV Sequence identity: 67.7%

excluded: ANXA26252FWD-ANXA2FWD Sequence identity: 65.6%
excluded: ANXA26252REV-ANXA2REV Sequence identity: 79.6%
excluded: ANXA26259FWD-ANXA2FWD Sequence identity: 57.0%
excluded: ANXA26259REV-ANXA2REV Sequence identity: 51.3%

No sequence passed the chosen sequence identity threshold of 90%

HLA-DRB6

AAGGTAGAGAGAATGAATCAGGAAGTTAGAGTCTCGTTGTCAGCTGTTTGTA
TGCTTCTCTGTAAACCCAGGCTCTGGCCTCGACCAGGCCTCCAGCACAGCTGG
CCATACGCCCTCACAGTGTCATCGGCCTGGAATTTAATCGTGATAGTGTGGAC
CTATCAGATTTGAGAGATGTTATAAAAAATTTTATTTGTTTCTTCATAGCTTG
AAATTGTCAC

excluded: HLADBR66216FWD-HLADBR6FWD Sequence identity: 36.5%
excluded: HLADBR66216REV-HLADBR6REV Sequence identity: 23.4%
excluded: HLADBR66217FWD-HLADBR6FWD Sequence identity: 18.3%
excluded: HLADBR66217REV-HLADBR6REV Sequence identity: 18.0%
excluded: HLADBR66228FWD-HLADBR6FWD Sequence identity: 13.6%
excluded: HLADBR66228REV-HLADBR6REV Sequence identity: 26.3%
excluded: HLADBR66235FWD-HLADBR6FWD Sequence identity: 17.2%
excluded: HLADBR66235REV-HLADBR6REV Sequence identity: 13.2%
excluded: HLADBR66252FWD-HLADBR6FWD Sequence identity: 44.2%
excluded: HLADBR66252REV-HLADBR6REV Sequence identity: 33.3%
excluded: HLADBR66259FWD-HLADBR6FWD Sequence identity: 33.3%
excluded: HLADBR66259REV-HLADBR6REV Sequence identity: 27.2%

No sequence passed the chosen sequence identity threshold of 90%

MRI1

CGACTCGGCCTCCCAAAATGCTGGGATTCCAGGCGTGAGCCACAGCGCCTGT
CCTGCATGTTACTTTTGAATGAAACC

excluded: MRI16216FWD-MRI1FWD Sequence identity: 37.8%
excluded: MRI16216FWD-MRI1FWD_R Sequence identity: 40.3%
excluded: MRI16216REV-MRI1REV Sequence identity: 37.7%
excluded: MRI16216REV-MRI1REV_R Sequence identity: 55.8%
excluded: MRI16217FWD-MRI1FWD Sequence identity: 33.8%
excluded: MRI16217REV-MRI1REV Sequence identity: 44.2%
excluded: MRI16228FWD-MRI1FWD Sequence identity: 51.9%
excluded: MRI16228REV-MRI1REV Sequence identity: 64.4%
excluded: MRI16235FWD-MRI1FWD Sequence identity: 50.6%
excluded: MRI16235REV-MRI1REV Sequence identity: 35.1%

excluded: MRI16252FWD-MRI1FWD Sequence identity: 30.6%
excluded: MRI16252REV-MRI1REV Sequence identity: 36.0%
excluded: MRI16259FWD-MRI1FWD Sequence identity: 44.6%
excluded: MRI16259REV-MRI1REV Sequence identity: 68.4%

No sequence passed the chosen sequence identity threshold of 90%

GAP43

CGAAGAAAGGAACACTGGGTGAGATTTTGTTCAACTACCCATAGTTACCACC
AGATGGTGAAACTGATCCCGGGCCTCTTGGGTATTGATCAGTTTATGGGGAG
ATGGGGAGAAGACTATCTTTCACTTGTTAATTCATTAATTTCTTTCGCAAATA
TTTTTTCAGTACCTGCTAAGTCCCACGGACTATGCTAGGAGCT

excluded: GAP436216FWD-GAP43FWD Sequence identity: 54.8%
excluded: GAP436216REV-GAP43REV Sequence identity: 44.3%
excluded: GAP436217FWD-GAP43FWD Sequence identity: 70.1%
excluded: GAP436217REV-GAP43REV Sequence identity: 31.2%
excluded: GAP436228FWD-GAP43FWD Sequence identity: 37.8%
excluded: GAP436228REV-GAP43REV Sequence identity: 74.6%
excluded: GAP436235FWD-GAP43FWD Sequence identity: 76.6%
excluded: GAP436235REV-GAP43REV Sequence identity: 36.0%
excluded: GAP436252FWD-GAP43FWD Sequence identity: 80.3%
excluded: GAP436252REV-GAP43REV Sequence identity: 48.9%
excluded: GAP436259FWD-GAP43FWD Sequence identity: 60.7%
excluded: GAP436259REV-GAP43REV Sequence identity: 67.9%

No sequence passed the chosen sequence identity threshold of 90%

CUL3

TAACTTGTTAAAGGTTATGCAGCTAGAGGCCGGGCGCGGTGGCTCAC

excluded: CUL36216FWD-CUL3FWD Sequence identity: 45.7%
excluded: CUL36216REV-CUL3REV Sequence identity: 60.9%
excluded: CUL36217FWD-CUL3FWD Sequence identity: 56.5%
excluded: CUL36217REV-CUL3REV Sequence identity: 54.3%
excluded: CUL36228FWD-CUL3FWD Sequence identity: 50.0%
excluded: CUL36228REV-CUL3REV Sequence identity: 60.9%
excluded: CUL36235FWD-CUL3FWD Sequence identity: 45.7%
excluded: CUL36235REV-CUL3REV Sequence identity: 52.2%
excluded: CUL36252FWD-CUL3FWD Sequence identity: 54.3%
excluded: CUL36252REV-CUL3REV Sequence identity: 47.8%
excluded: CUL36259FWD-CUL3FWD Sequence identity: 54.3%
excluded: CUL36259REV-CUL3REV Sequence identity: 58.7%

No sequence passed the chosen sequence identity threshold of 90%

Appendix I

R Code and Normalization Figures and Output Table

```
library(BiocInstaller)
biocValid()
source("http://bioconductor.org/biocLite.R")
biocLite("limma")
biocLite("illuminaio")
biocLite("minfi")
biocLite("minfiData")
biocLite() ##installs every package on bioconductor
biocLite("IlluminaHumanMethylation450kmanifest")
require("minfi")
require("minfiData")

##generate a Red Green data set
baseDir<-("C:\\Users\\Michelle\\Dropbox\\dissertation\\iDat_all")
targets <- read.450k.sheet(baseDir)
RGset<-read.450k.exp(base=baseDir, targets=targets)
pd<-pData(RGset)
qcReport(RGset, sampNames = pd$Sample_ID,sampGroups=pd$GBS,pdf =
"qcReport.pdf")
densityPlot(RGset, sampGroups = pd$GBS,main = "Beta", xlab = "Beta")##shows
density plot in R workspace
par(oma=c(2,10,1,1))
densityBeanPlot(RGset, sampGroups = pd$GBS,sampNames = pd$Sample_ID)
##Methylation sets
MSet.raw <- preprocessRaw(RGset)
MSet.norm <- preprocessIllumina(RGset, bg.correct = TRUE, normalize = "controls",
reference = 2) ##normalizes methylation data
controlStripPlot(RGset, controls="BISULFITE CONVERSION II",sampNames =
pd$Sample_ID)##control bisulfite conversion II
annot<-read.csv("http://supportres.illumina.com/documents/myillumina/b78d361a-def5-
4adb-ab38-e8990625f053/humanmethylation450_15017482_v1-2.csv",
skip=7,head=TRUE)
getMeth(MSet.raw)[1:4,1:3]##returns raw methylation values use
getUnmeth(MSet.raw)[1:4,1:3]##returns raw unmethylated values
getBeta(MSet.raw, type = "Illumina")[1:4,1:3] ##returns raw beta values
MSet.raw<-preprocessRaw(RGset)
Beta.raw<-getBeta(MSet.raw)
rownames(Beta.raw)[1:5]
annot$IlmnID[1:5]
annot<-annot[match(rownames(Beta.raw),annot$IlmnID),,drop=FALSE] ##annotates
betafile with illumina IDs
class(rownames(Beta.raw))
class(annot$IlmnID)
all.equal(as.character(annot$IlmnID),rownames(Beta.raw))
names(annot)
head(annot)
```

```
table(annot$Probe_SNPs)
##first removed CpG sites that were known to be associated with SNP
any.SNPs<-ifelse(annot$Probe_SNPs!="",1,0)
raw.wo<-Beta.raw[any.SNPs!=1,]
dim(raw.wo)
annot.wo<-annot[any.SNPs!=1,]
dim(annot.wo)
all.equal(as.character(annot.wo$IlmnID),rownames(raw.wo))
sum(raw.wo==1,na.rm=TRUE)
sum(raw.wo==0,na.rm=TRUE)
## To avoid errors when applying the logit transformation that would occur if beta=0 or
beta=1,
## we imputed 0.999 when beta=1 and 0.001 when beta=0
raw.wo[raw.wo==1]<-0.999
raw.wo[raw.wo==0]<-0.001
sum(raw.wo==1,na.rm=TRUE)
sum(raw.wo==0,na.rm=TRUE)
#After imputation, the logit transformation was applied to the beta values
logit<-log(raw.wo/(1-raw.wo)) ## Logit transformed beta values
hist(logit[1,])
hist(raw.wo[1,])
Peak.correction<-
function(exprs,annot) {
       for (i in 1:dim(exprs)[2]) {

      dens.I<-density(exprs[annot$Infinium_Design_Type=="I", i],na.rm=T)
       dens.II<-density(exprs[annot$Infinium_Design_Type=="II", i],na.rm=T)
       sigma.uII<- -dens.II$x[dens.II$x<0][which.max(dens.II$y[dens.II$x<0])]
       sigma.mII<-dens.II$x[dens.II$x>0][which.max(dens.II$y[dens.II$x>0])]
       sigma.uI<- -dens.I$x[dens.I$x<0][which.max(dens.I$y[dens.I$x<0])]
       sigma.mI<-dens.I$x[dens.I$x>0][which.max(dens.I$y[dens.I$x>0])]
       exprs[,i]<- ifelse(annot$Infinium_Design_Type=="II" & exprs[,i]<0,
         exprs[,i]/sigma.uII*sigma.uI,

         ifelse(annot$Infinium_Design_Type=="II" & exprs[,i]>0,
               exprs[,i]/sigma.mII*sigma.mI, exprs[,i]))
 }

       exprs

      }
all.equal(as.character(annot.wo$IlmnID),rownames(logit)
##The logit transformed beta values were then peak corrected to ensure the peaks of the
Infinium II
##design beadtype were comparable to the peak locations of the Infinum I beadtype
correct.methyl<-Peak.correction(logit,annot.wo)
```

```
pData(MSet.raw)
plot(density(logit[annot.wo$Infinium_Design_Type=="I",1]), col="black", xlab="M-
values",main="Sample 1", ylim=c(0,0.3))
lines(density(logit[annot.wo$Infinium_Design_Type=="II",1]), col="red")
lines(density(logit[annot.wo$Infinium_Design_Type=="II",1],na.rm=TRUE), col="red")
plot(density(correct.methyl[annot.wo$Infinium_Design_Type=="I",1]), col="black",
xlab="M-values",main="Peak-corrected Sample 1", ylim=c(0,0.3))
lines(density(correct.methyl[annot.wo$Infinium_Design_Type=="II",1],na.rm=TRUE),
col="red")
plot(density(logit[annot.wo$Infinium_Design_Type=="I",4]), col="black", xlab="M-
values",main="Sample 4", ylim=c(0,0.3))
lines(density(logit[annot.wo$Infinium_Design_Type=="II",4],na.rm=TRUE), col="red")
plot(density(correct.methyl[annot.wo$Infinium_Design_Type=="I",4]), col="black",
xlab="M-values",main="Peak-corrected Sample 4", ylim=c(0,0.3))
lines(density(correct.methyl[annot.wo$Infinium_Design_Type=="II",4],na.rm=TRUE),
col="red")
plot(density(logit[annot.wo$Infinium_Design_Type=="I",7]), col="black", xlab="M-
values",main="Sample Unmethy", ylim=c(0,0.3))
lines(density(logit[annot.wo$Infinium_Design_Type=="II",7],na.rm=TRUE), col="red")
plot(density(correct.methyl[annot.wo$Infinium_Design_Type=="I",7]), col="black",
xlab="M-values",main="Peak-corrected Sample unmeth", ylim=c(0,0.3))
lines(density(correct.methyl[annot.wo$Infinium_Design_Type=="II",7],na.rm=TRUE),
col="red")
##this is unnecessary because GBS is either positive or negative so a subset does not
need to be created;
##however other data sets may need subsets created to carry out analysis and here for
your reference
Msubset<-correct.methyl[,pData(MSet.raw)$GBS=="Negative"|
  pData(MSet.raw)$GBS=="Positive"]
##Notice results for GBS will be the same as the above using
Msubset2<-correct.methyl
all.equal(Msubset,Msubset2)
dim(Msubset)
class(Msubset)

##if you need a group vector, this is how it should be formatted
group<-pData(MSet.raw)[pData(MSet.raw)$GBS=="Negative"|
   pData(MSet.raw)$GBS=="Positive","GBS"]
##However, I don't need to use one because of how my data is structured;
##instead I could use group<-pData(MSet.raw)$GBS as the vector in design
pData(MSet.raw)
library(limma)
design<-model.matrix(~as.factor(group)-1)
design
colnames(design)<-c("Negative","Positive")
design
```

```
fit<-lmFit(Msubset, design) ##least squares fitting off linear model for each gene
fit
contr.matrix<-makeContrasts(GBS.status=Positive-Negative, levels=design)
contr.matrix
fit2<-contrasts.fit(fit, contr.matrix) #makes contras to compare by GBS status
fit3<-eBayes(fit2) ##use of array weights increases the significance of top genes
results<-topTable(fit3, coef="GBS.status",  number=dim(Msubset)[1],
  sort.by="none", adjust="BH")
### Warning, the sort.by="none" and number=dim()[1] parameters are important to
ensure your results are aligned with the annotation data
head(results)
results<-data.frame(ID=rownames(results),results)
head(results)
head(Msubset)
dim(results)
dim(Msubset)
all.equal(results$ID, rownames(Msubset$ID))
class(results$ID)
class(Msubset)
class(Msubset[,1])
all.equal(comp[,1],comp[,2])
all.equal(results$ID,Msubset[,1])
### Note that without specifying sort.by="none" the annotation information is misaligned
##correct.beta is the back-transformed peak corrected logitvalues
correct.beta<-exp(Msubset)/(1+exp(Msubset))
dim(correct.beta)
head(correct.beta)
sum(results$adj.P.Val<0.05)

mu.Negative<-apply(correct.beta[,group=="Negative"], 1, mean)
mu.Positive<-apply(correct.beta[,group=="Positive"], 1, mean)
delta.beta<-mu.Positive-mu.Negative
mu.Negative[1:10]
mu.Positive[1:10]
delta.beta[1:10]
all.equal(results$ID,as.character(annot.wo$IlmnID))

class(results$ID)
class(annot.wo$IlmnID)
all.equal(comp[,1],comp[,2])
all.equal(results$ID,annot.wo$IlmnID)
class(results$ID)==class(annot.wo$llmnID)
head(annot.wo)
head(results)
results<-data.frame(ID=rownames(results),results)
```

```
### Note that final.results and annot.wo are matched so we can append annot.wo to final
results;

## using generally accepted FDR cut off standards
final.FDR.001results<-data.frame(Probe=rownames(results),
beta.Negative=mu.Negative, beta.Positive=mu.Positive, delta.beta=delta.beta,
p.value=results$P.Value, FDR=results$adj.P.Val, annot.wo)
sign.results<-final.results[final.results$FDR<0.001 & abs(final.results$delta.beta)>0.2,]
dim(sign.FDR.001results) # yields 0 rows = zero significantly different CpG sites
final.FDR.05results<-data.frame(Probe=rownames(results), beta.Negative=mu.Negative,
beta.Positive=mu.Positive, delta.beta=delta.beta, p.value=results$P.Value,
FDR=results$adj.P.Val, annot.wo)
sign.results<-final.results[final.results$FDR<0.05 & abs(final.results$delta.beta)>0.2,]
dim(sign.FDR.05results) # yields 0 rows = zerp significantly differnt CpG sites
final.FDR.01results<-data.frame(Probe=rownames(results), beta.Negative=mu.Negative,
beta.Positive=mu.Positive, delta.beta=delta.beta, p.value=results$P.Value,
FDR=results$adj.P.Val, annot.wo)
sign.results<-final.results[final.results$FDR<0.01 & abs(final.results$delta.beta)>0.2,]
dim(sign.results) # yields 0 rows = zero significantly differnt CpG sites

##Volcano plot, showing no significant differences. Signifincat CpG sites would be blue
install.packages("ggplot2")
require(ggplot2)
##Highlight CpGs that have an absolute fold change > 2 and a FDR < 0.05
FDRcut = as.factor(abs(results$logFC) > 2 & results$adj.P.Val < 0.05)
sum(abs(results$logFC) > 2 & results$adj.P.Val < 0.05)
##Construct the plot object
g = ggplot(data=results, aes(x=logFC, y=-log10(P.Value), colour= FDRcut)) +
  geom_point(alpha=0.4, size=1.75) +
  theme(legend.position = "none") +
  xlim(c(-4, 4)) + ylim(c(0, 6)) +
  xlab("log2 fold change") + ylab("-log10 p-value") ##returns no blue (significant) dots,
all pink CpGs


### using generic p-value 0.05 cutoff that Joyce uses
final.results<-data.frame(Probe=results$ID, beta.Negative=mu.Negative,
beta.Positive=mu.Positive, delta.beta=delta.beta, p.value=results$P.Value,
FDR=results$adj.P.Val, annot.wo)
sign.results<-final.results[final.results$p.value<.05& abs(final.results$delta.beta)>0.2,]
dim
### To output in order to email your findings to an investigator, use write.table
write.table(sign.results,"significantRrun.csv",sep=",",row.names=FALSE)
### You could also remove some of the annotation fields using -c() in square bracket
notation; investigators may not be interested in all fields.
```

bisulfitesequencing validation code

getBeta(MSet.raw)["cg11164659",] ##RHPN1 raw uncorrected beta value
getBeta(MSet.raw)["cg09785377",] ##ANXA2 raw uncorrected beta value

cpgsD<-c("cg09785377","cg11164659")
cpgsD
plotCpg(Msubset, cpg=cpgsD[1], pheno=pData(Msubset)$GBS) #ANXA2 plot of raw beta values by GBS status for each participant
plotCpg(Msubset, cpg=cpgsD[2], pheno=pData(Msubset)$GBS) ##RHPN1 plot of raw beta values by GBS status for each participant

(correct.beta)["cg11164659",] ##RHPN1 normalized beta value
(correct.beta)["cg09785377",] ##ANXA2 normalized beta value

##raw and corrected beats for CpG site only significant using R platform compared to genome studio significant values
getBeta(MSet.raw)["cg23947138",] ##RASA3 raw beta value

(correct.beta)["cg23947138",] ##RASA3 normalized beta value

getBeta(MSet.raw)["cg00540295",] ##FAM69B raw beta value

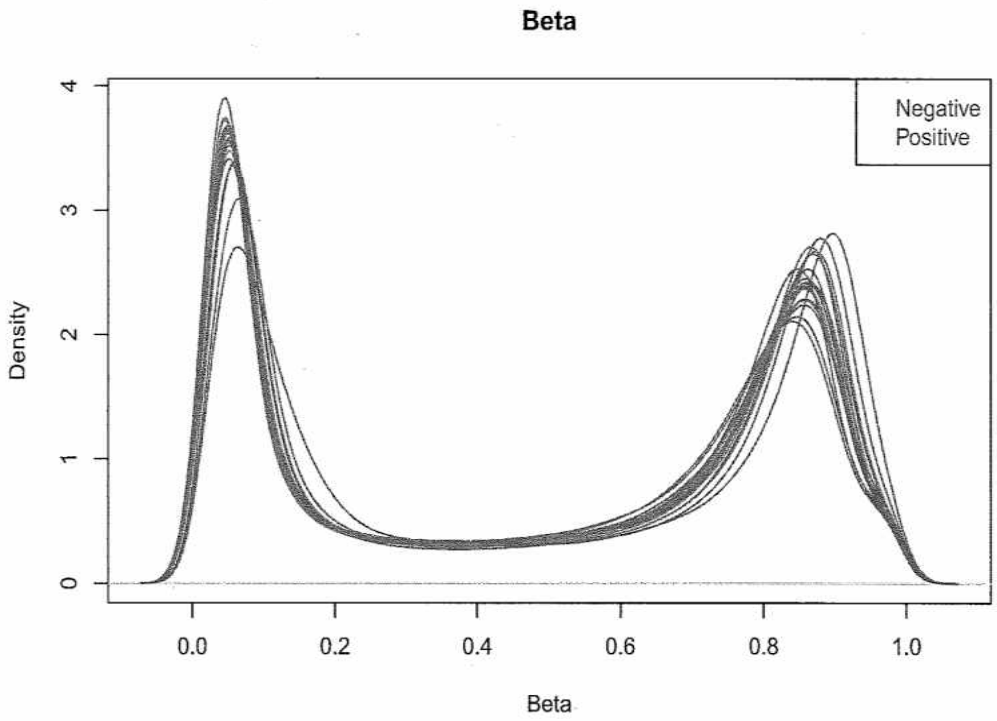
(correct.beta)["cg00540295",] ##FAM69B - snp

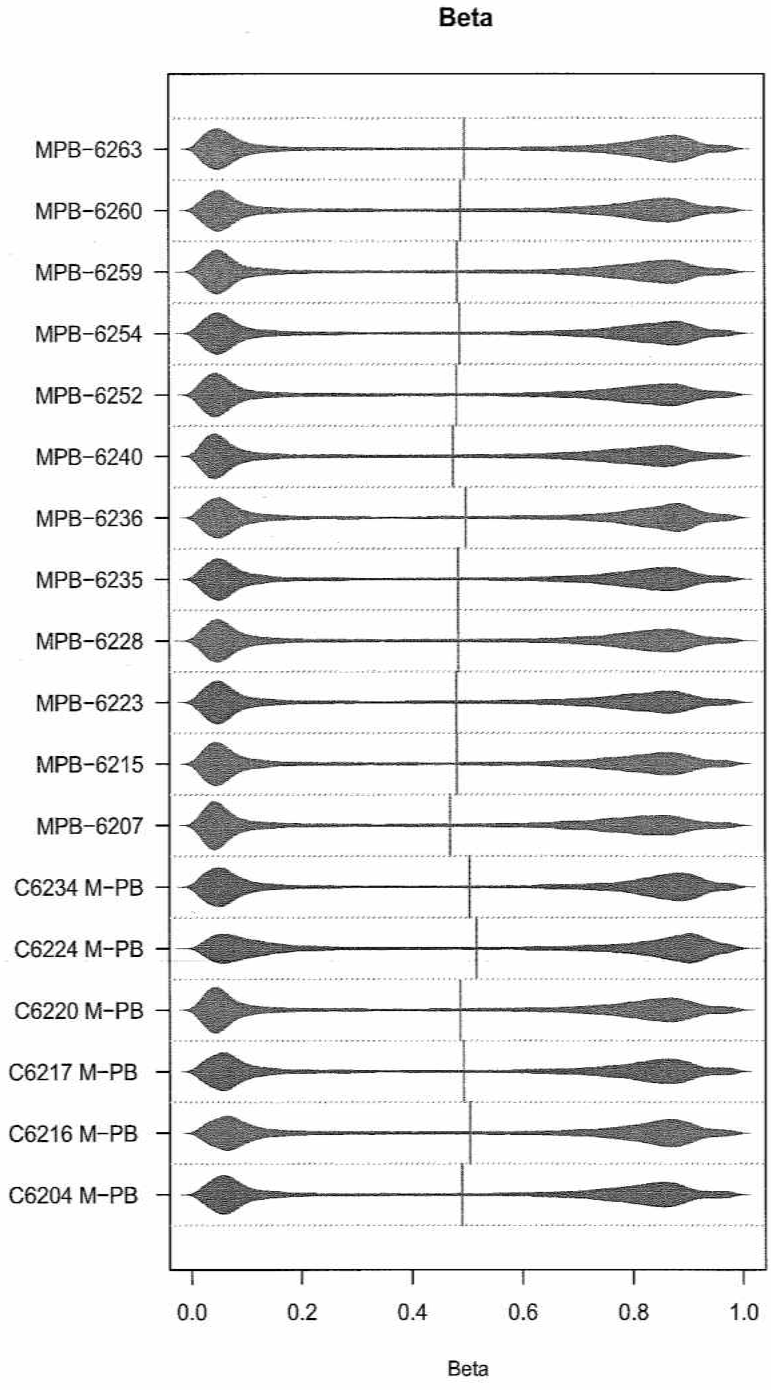
getBeta(MSet.raw)["cg01270299",] ##ZNF137 - snp
(correct.beta)["cg01270299",]

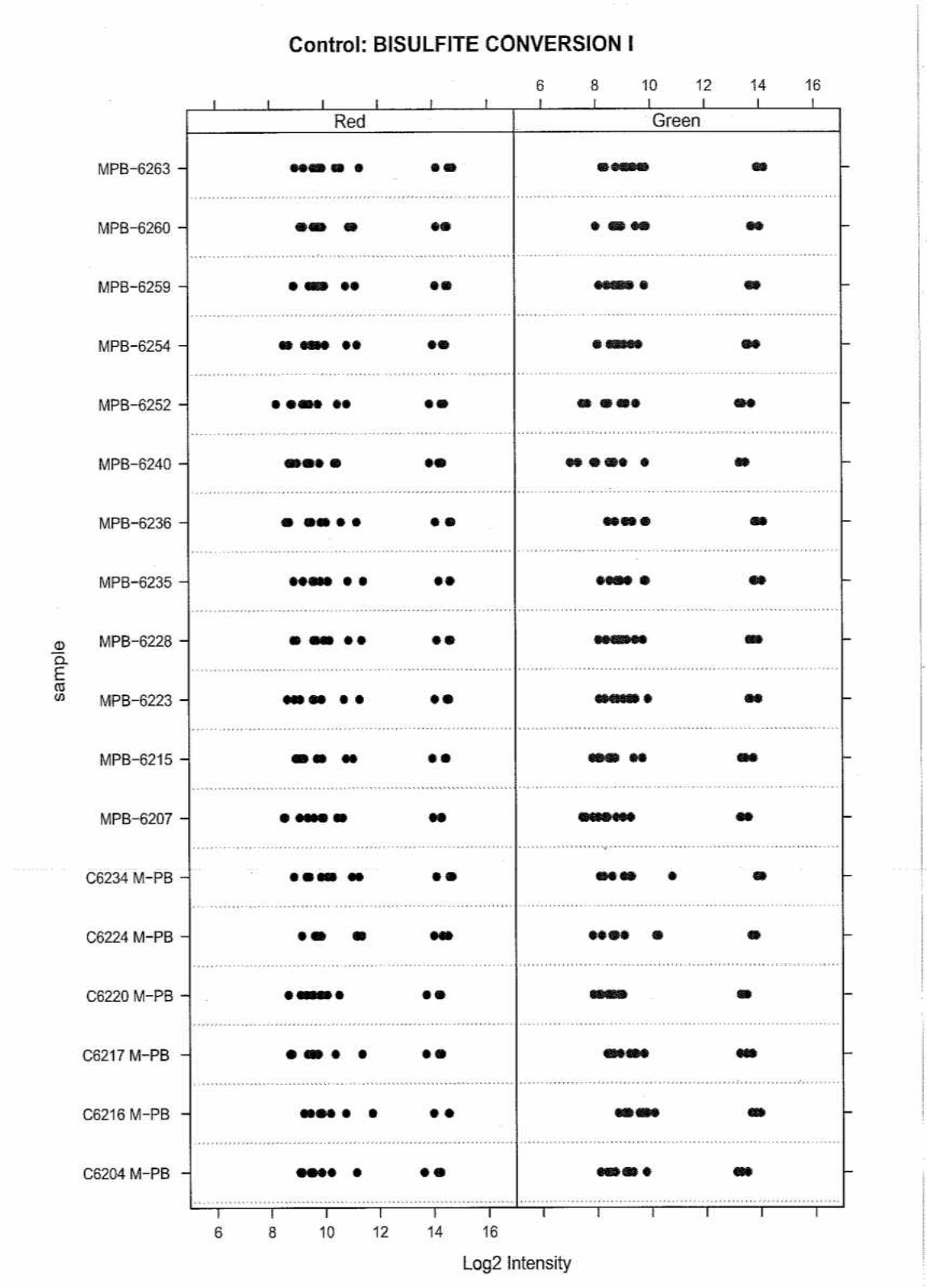
getBeta(MSet.raw)["cg25909532",] ##VIPR2 - snp
(correct.beta)["cg25909532",]

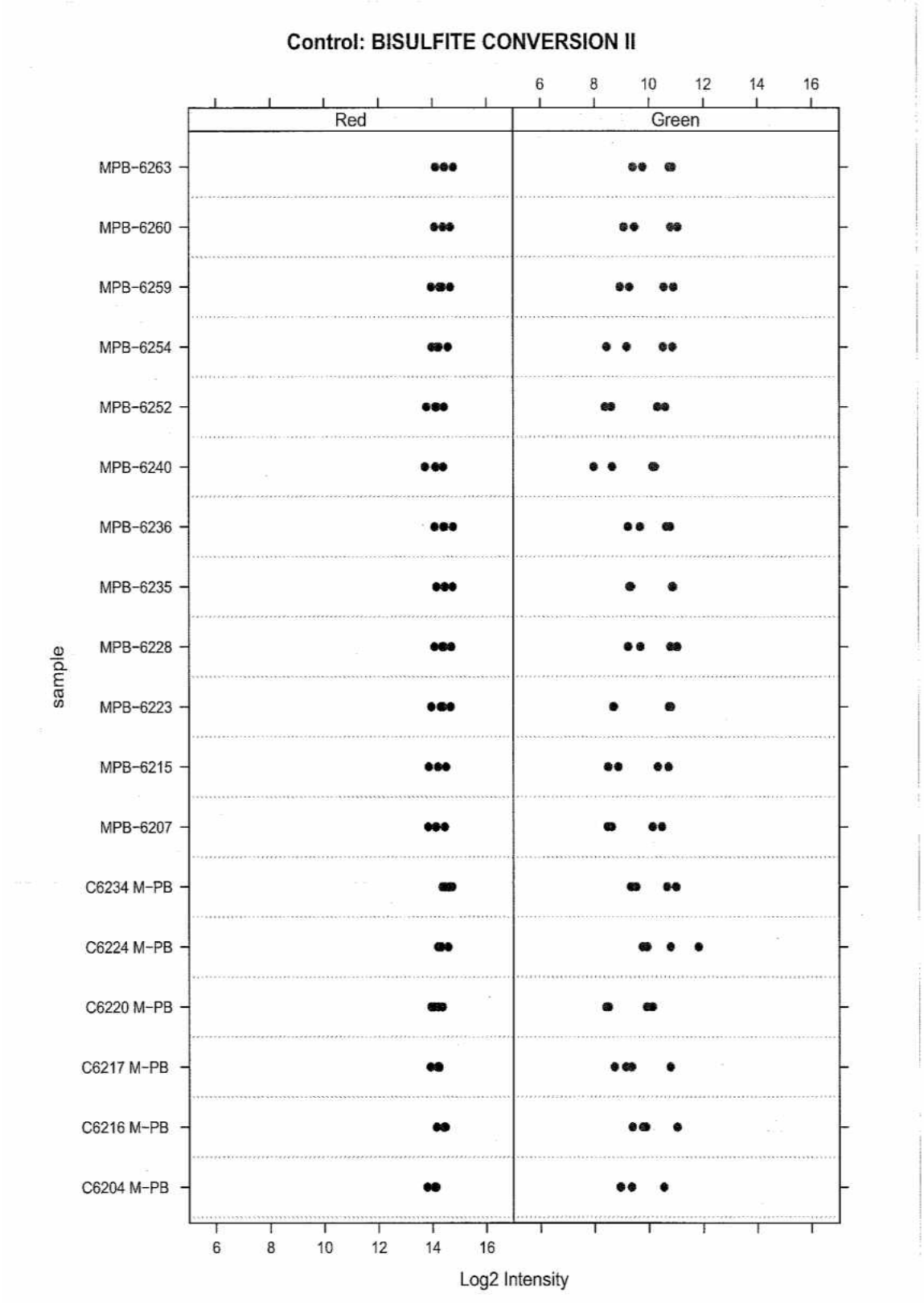
getBeta(MSet.raw)["cg06688803",] ##CLPTM1 - snp
(correct.beta)["cg06688803",]
getBeta(MSet.raw)["cg07304760",] ##SND1 - reads
(correct.beta)["cg07304760",]
getBeta(MSet.raw)["cg15290312",] ##TIMP2 - snp
(correct.beta)["cg15290312",]
getBeta(MSet.raw)["cg10058204",] ##FLJ37201 - reads
(correct.beta)["cg10058204",]
getBeta(MSet.raw)["cg05331763",] ##FOxK2 - reads
(correct.beta)["cg05331763",]
getBeta(MSet.raw)["cg12434901",] ##KCNH6 - reads
(correct.beta)["cg12434901",]
getBeta(MSet.raw)["cg24634471",] ##JRK - reads
(correct.beta)["cg24634471",]
getBeta(MSet.raw)["cg01421902",] ##ZNF665 - reads

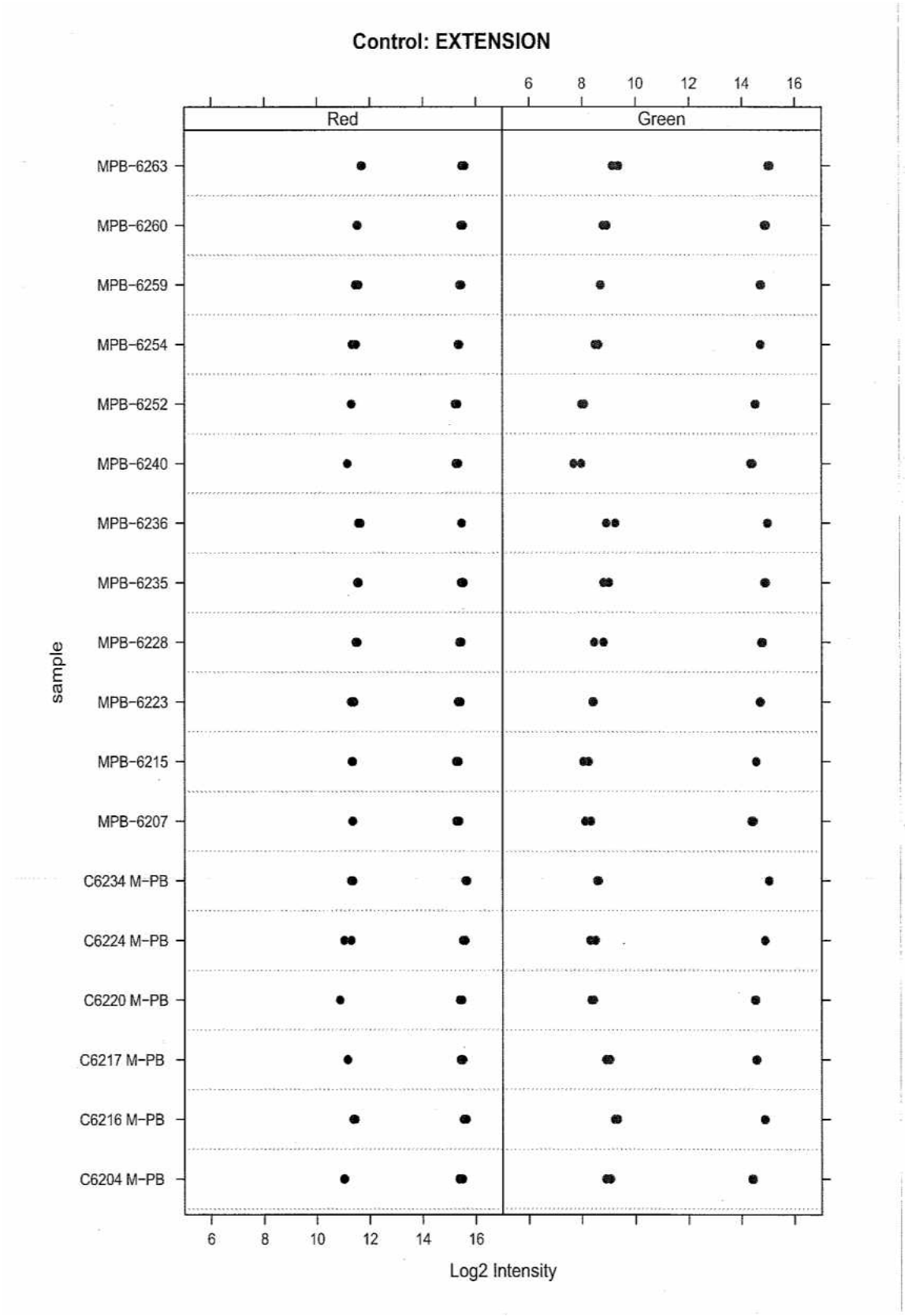
```
(correct.beta)["cg01421902",]
getBeta(MSet.raw)["cg04388792",] ##ZNF490 - reads
(correct.beta)["cg04388792",]
getBeta(MSet.raw)["cg13506281",] ##MTUS2 - snp
(correct.beta)["cg04388792",]
getBeta(MSet.raw)["cg07703391",] ##BMP8B/PPIE -reads
(correct.beta)["cg07703391",]
getBeta(MSet.raw)["cg10632770",] ##KIAA1199 -snp
(correct.beta)["cg10632770",]
getBeta(MSet.raw)["cg20479209",] ##FLJ43860 -reads
(correct.beta)["cg20479209",]
getBeta(MSet.raw)["cg13066461",] ##MRGPRX2-reads
(correct.beta)["cg13066461",]
getBeta(MSet.raw)["cg21130926",] ##SULF2 -reads
(correct.beta)["cg21130926",]
getBeta(MSet.raw)["cg04922606",] ##FAM120B -SNP
(correct.beta)["cg04922606",]
getBeta(MSet.raw)["cg03292225",] ##TNNT3 -reads
(correct.beta)["cg03292225",]
getBeta(MSet.raw)["cg10890644",] ##TUBAL3- snp
(correct.beta)["cg10890644",]
getBeta(MSet.raw)["cg17671604",] ##SPTBN4- snp
(correct.beta)["cg17671604",]
getBeta(MSet.raw)["cg09307883",] ##ANAPC2- reads
(correct.beta)["cg09307883",]
getBeta(MSet.raw)["cg10528424",] ##SYT8 - reads
(correct.beta)["cg10528424",]
getBeta(MSet.raw)["cg07480176",] ##CASD1 - snp
(correct.beta)["cg07480176",]
getBeta(MSet.raw)["cg14252149",] ##LGALS8 -snp
(correct.beta)["cg14252149",]
getBeta(MSet.raw)["cg12155450",] ##NAT14/ZNF628 -reads
(correct.beta)["cg12155450",]
getBeta(MSet.raw)["cg20976286",] ##OCA2 - reads
(correct.beta)["cg20976286",]
getBeta(MSet.raw)["cg10995422",] ##HLA-DRB6 - snp
(correct.beta)["cg10995422",]
```

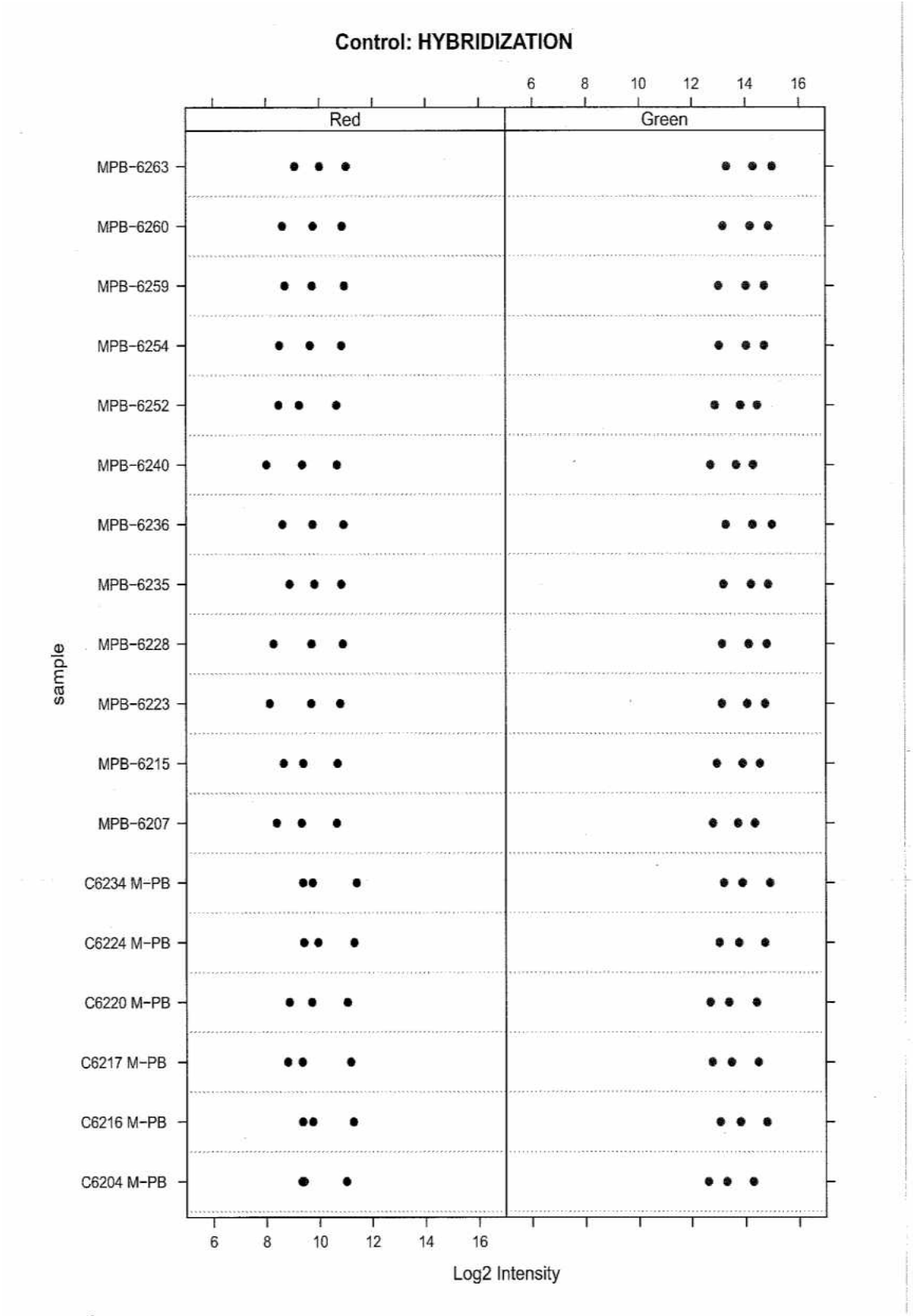

Beta
Negative
Positive
Density
0
1
2
3
4
0.0
0.2
0.4
0.6
0.8
1.0
Beta

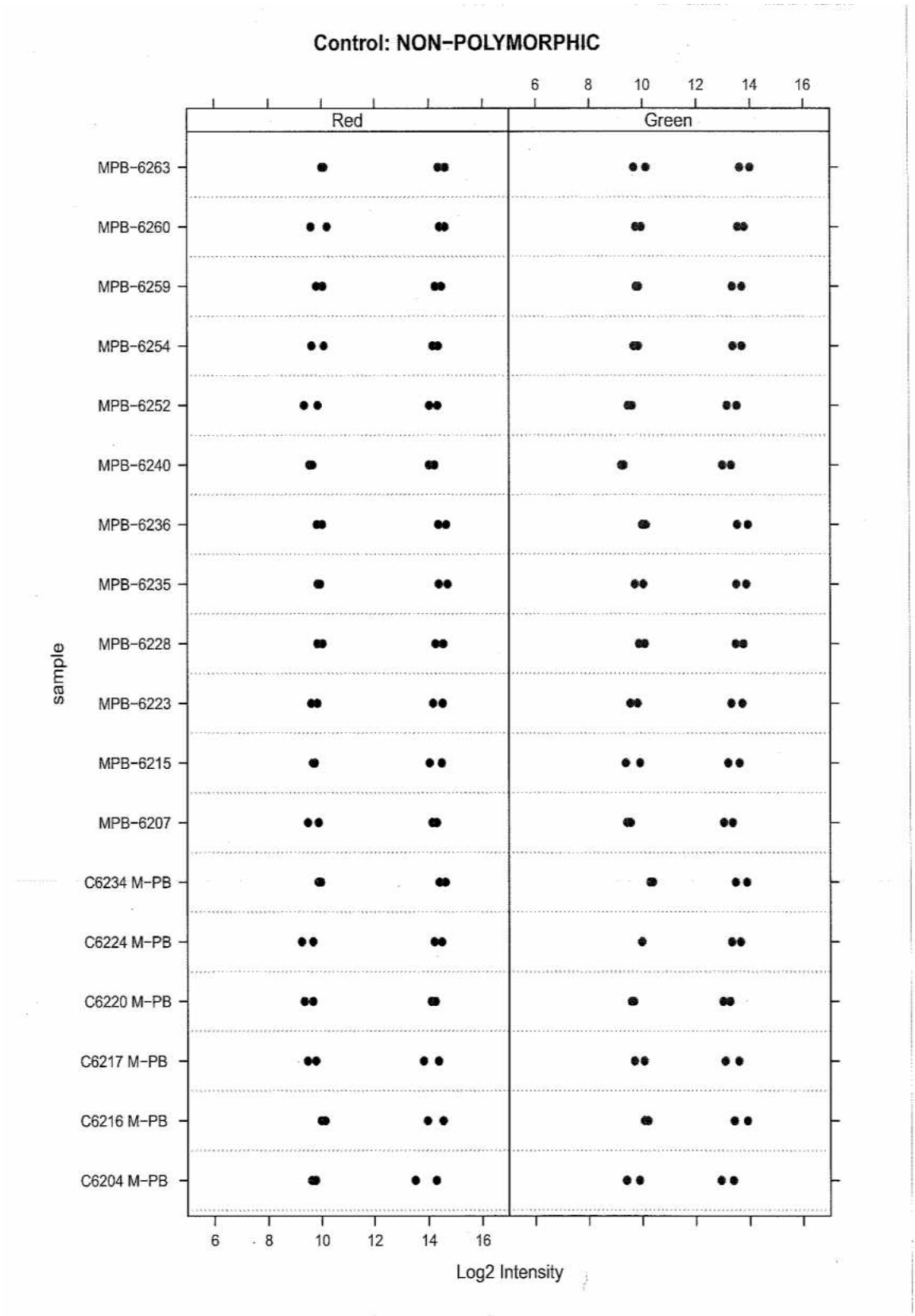

Beta
MPB-6263
MPB-6260
MPB-6259
MPB-6254
MPB-6252
MPB-6240
MPB-6236
MPB-6235
MPB-6228
MPB-6223
MPB-6215
MPB-6207
C6234 M-PB
C6224 M-PB
C6220 M-PB
C6217 M-PB
C6216 M-PB
C6204 M-PB
0.0
0.2
0.4
0.6
0.8
1.0
Beta

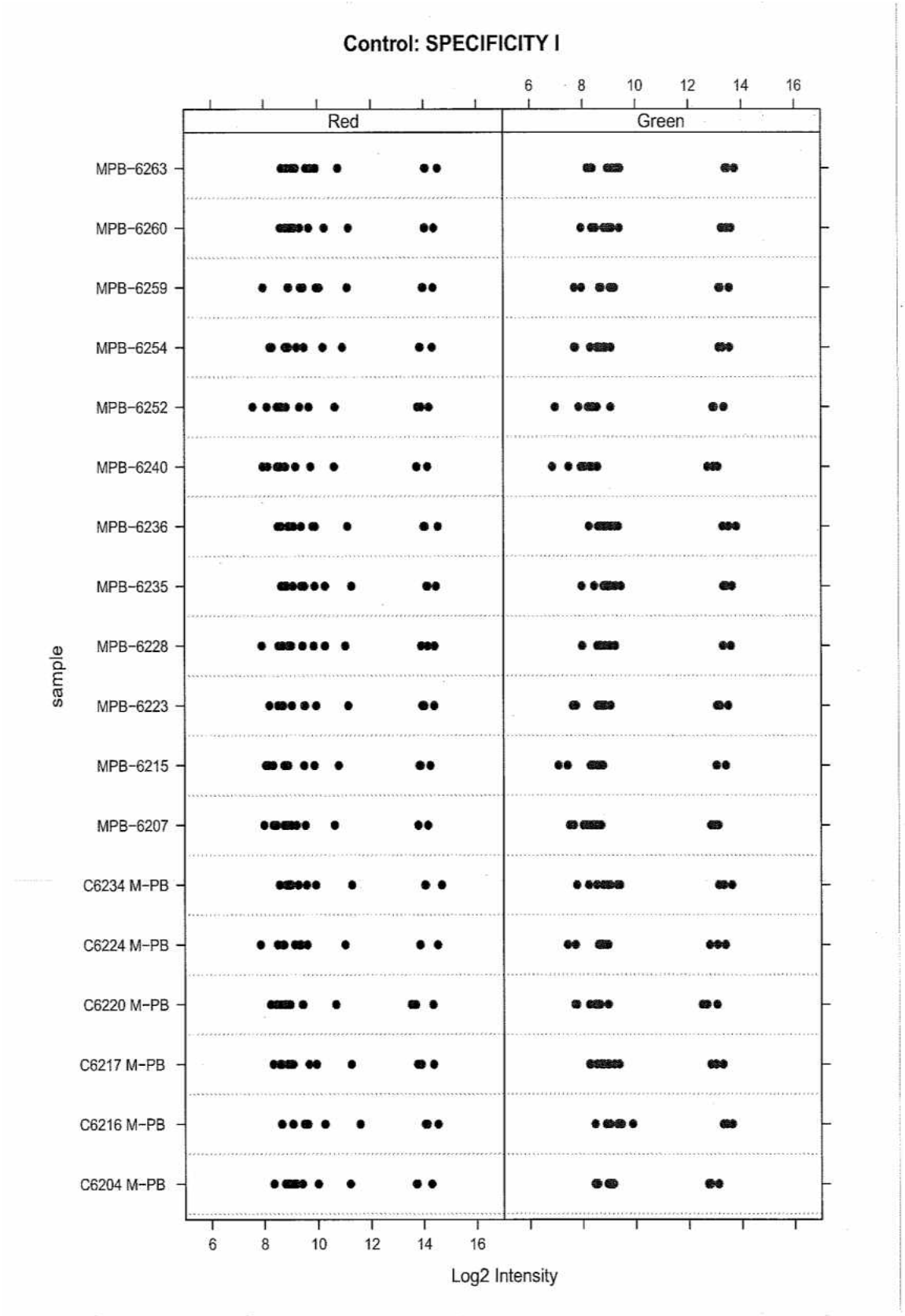

Control: BISULFITE CONVERSION I
Red
Green
MPB-6263
MPB-6260
MPB-6259
MPB-6254
MPB-6252
MPB-6240
MPB-6236
MPB-6235
MPB-6228
MPB-6223
MPB-6215
MPB-6207
C6234 M-PB
C6224 M-PB
C6220 M-PB
C6217 M-PB
C6216 M-PB
C6204 M-PB
sample
6
8
10
12
14
16
Log2 Intensity

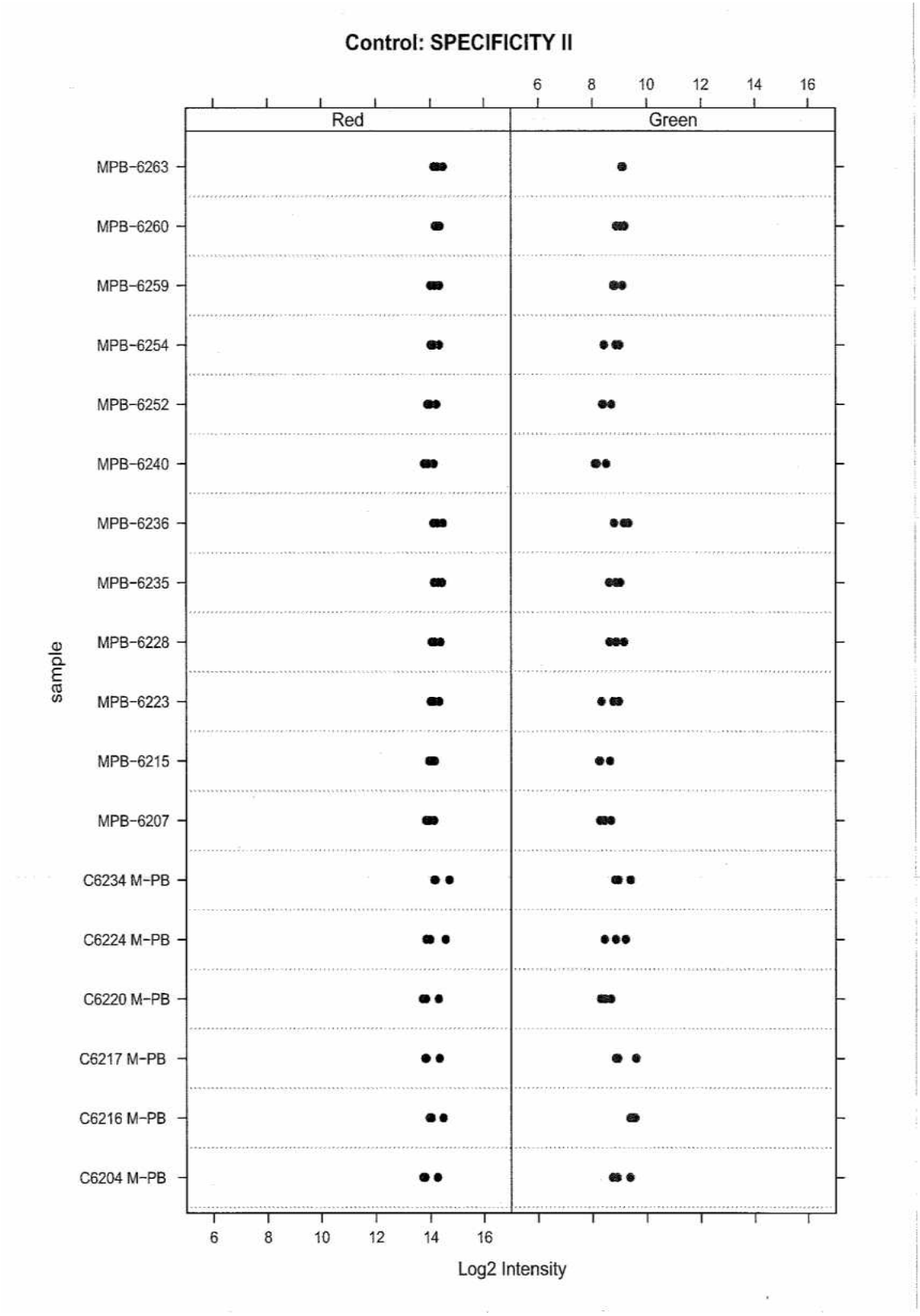

Control: BISULFITE CONVERSION II
Red
Green
6
8
10
12
14
16
MPB-6263
MPB-6260
MPB-6259
MPB-6254
MPB-6252
MPB-6240
MPB-6236
MPB-6235
MPB-6228
MPB-6223
MPB-6215
MPB-6207
C6234 M-PB
C6224 M-PB
C6220 M-PB
C6217 M-PB
C6216 M-PB
C6204 M-PB
sample
Log2 Intensity

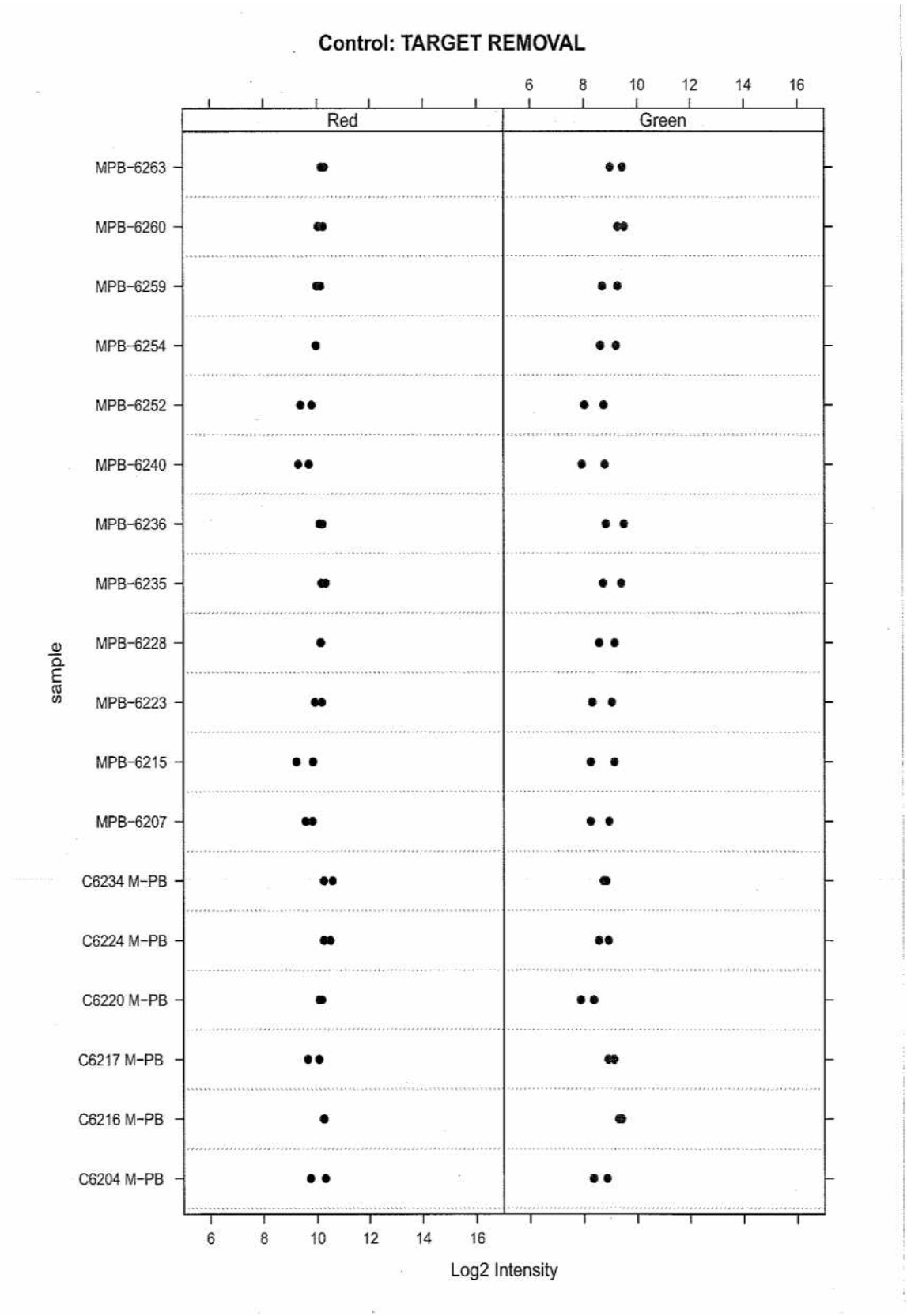

Control: EXTENSION
Red
Green
MPB-6263
MPB-6260
MPB-6259
MPB-6254
MPB-6252
MPB-6240
MPB-6236
MPB-6235
MPB-6228
MPB-6223
MPB-6215
MPB-6207
C6234 M-PB
C6224 M-PB
C6220 M-PB
C6217 M-PB
C6216 M-PB
C6204 M-PB
sample
6
8
10
12
14
16
Log2 Intensity


Control: HYBRIDIZATION
Red
Green
6
8
10
12
14
16
MPB-6263
MPB-6260
MPB-6259
MPB-6254
MPB-6252
MPB-6240
MPB-6236
MPB-6235
MPB-6228
MPB-6223
MPB-6215
MPB-6207
C6234 M-PB
C6224 M-PB
C6220 M-PB
C6217 M-PB
C6216 M-PB
C6204 M-PB
sample
Log2 Intensity


Control: NON-POLYMORPHIC
Red
Green
6
8
10
12
14
16
MPB-6263
MPB-6260
MPB-6259
MPB-6254
MPB-6252
MPB-6240
MPB-6236
MPB-6235
MPB-6228
MPB-6223
MPB-6215
MPB-6207
C6234 M-PB
C6224 M-PB
C6220 M-PB
C6217 M-PB
C6216 M-PB
C6204 M-PB
sample
Log2 Intensity


Control: SPECIFICITY I
Red
Green
MPB-6263
MPB-6260
MPB-6259
MPB-6254
MPB-6252
MPB-6240
MPB-6236
MPB-6235
MPB-6228
MPB-6223
MPB-6215
MPB-6207
C6234 M-PB
C6224 M-PB
C6220 M-PB
C6217 M-PB
C6216 M-PB
C6204 M-PB
sample
6
8
10
12
14
16
Log2 Intensity


Control: SPECIFICITY II
Red
Green
6
8
10
12
14
16
MPB-6263
MPB-6260
MPB-6259
MPB-6254
MPB-6252
MPB-6240
MPB-6236
MPB-6235
MPB-6228
MPB-6223
MPB-6215
MPB-6207
C6234 M-PB
C6224 M-PB
C6220 M-PB
C6217 M-PB
C6216 M-PB
C6204 M-PB
sample
Log2 Intensity


Control: TARGET REMOVAL
Red
Green
6
8
10
12
14
16
MPB-6263
MPB-6260
MPB-6259
MPB-6254
MPB-6252
MPB-6240
MPB-6236
MPB-6235
MPB-6228
MPB-6223
MPB-6215
MPB-6207
C6234 M-PB
C6224 M-PB
C6220 M-PB
C6217 M-PB
C6216 M-PB
C6204 M-PB
sample
Log2 Intensity

Probe	beta.Negat	beta.Positi	delta.beta	p.value	FDR	IlmnID	Name	AddressA_I
cg0446763	0.25843	0.800366	-0.54194	6.67E-05	0.999992	cg0446763	cg0446763	63703470
cg1116465	0.372497	0.832657	-0.46016	7.79E-05	0.999992	cg1116465	cg1116465	27622355
cg0322139	0.329288	0.73649	-0.4072	0.002769	0.999992	cg0322139	cg0322139	23660495
cg0852247	0.210945	0.581815	-0.37087	0.004461	0.999992	cg0852247	cg0852247	53649408
cg1097852	0.466372	0.806391	-0.34002	0.01313	0.999992	cg1097852	cg1097852	23654464
cg2413629	0.160883	0.4966	-0.33572	0.013453	0.999992	cg2413629	cg2413629	58712385
cg1563391	0.597114	0.921117	-0.324	0.012817	0.999992	cg1563391	cg1563391	46646406
cg0897731	0.219708	0.543341	-0.32363	0.006153	0.999992	cg0897731	cg0897731	31717420
cg0424670	0.370418	0.682364	-0.31195	0.045569	0.999992	cg0424670	cg0424670	15809316
cg2524308	0.149963	0.459685	-0.30972	0.000125	0.999992	cg2524308	cg2524308	55626484
cg1255190	0.481229	0.786936	-0.30571	0.022987	0.999992	cg1255190	cg1255190	54759409
cg0595694	0.534448	0.839984	-0.30554	0.00312	0.999992	cg0595694	cg0595694	57628353
cg0851419	0.577935	0.876401	-0.29847	0.031578	0.999992	cg0851419	cg0851419	72705460
cg2485165	0.346336	0.644359	-0.29802	0.031454	0.999992	cg2485165	cg2485165	22743423
cg2138833	0.239326	0.531601	-0.29227	0.012766	0.999992	cg2138833	cg2138833	69638509
cg2487392	0.488609	0.780509	-0.2919	0.000374	0.999992	cg2487392	cg2487392	29785495
cg2616263	0.620694	0.910751	-0.29006	0.000618	0.999992	cg2616263	cg2616263	35653353
cg0992568	0.461025	0.747655	-0.28663	0.027623	0.999992	cg0992568	cg0992568	36637470
cg0120191	0.543413	0.827772	-0.28436	0.043415	0.999992	cg0120191	cg0120191	51627476
cg1316715	0.659629	0.9436	-0.28397	0.021125	0.999992	cg1316715	cg1316715	52796390
cg0112760	0.213465	0.497024	-0.28356	0.042571	0.999992	cg0112760	cg0112760	24715331
cg2453864	0.356838	0.639609	-0.28277	0.00224	0.999992	cg2453864	cg2453864	56634317
cg2742395	0.120616	0.402233	-0.28162	6.47E-06	0.999992	cg2742395	cg2742395	30689485
cg0668491	0.501861	0.781294	-0.27943	0.038384	0.999992	cg0668491	cg0668491	70649349
cg2129430	0.336143	0.61348	-0.27734	0.03464	0.999992	cg2129430	cg2129430	45781502
cg0663177	0.333922	0.608484	-0.27456	0.048309	0.999992	cg0663177	cg0663177	24644310
cg2453678	0.429272	0.700238	-0.27097	0.002489	0.999992	cg2453678	cg2453678	11672327
cg2575542	0.162803	0.426214	-0.26341	0.031468	0.999992	cg2575542	cg2575542	14653467
cg0559388	0.306954	0.56904	-0.26209	0.045719	0.999992	cg0559388	cg0559388	15800414
cg1417593	0.363199	0.624853	-0.26165	0.011818	0.999992	cg1417593	cg1417593	47755467
cg1880516	0.302811	0.564398	-0.26159	0.004211	0.999992	cg1880516	cg1880516	51795502
cg1287021	0.463043	0.723859	-0.26082	0.00802	0.999992	cg1287021	cg1287021	72752369
cg0375488	0.445066	0.705665	-0.2606	0.022821	0.999992	cg0375488	cg0375488	45731391
cg1783975	0.229936	0.490444	-0.26051	0.022773	0.999992	cg1783975	cg1783975	70682329
cg2667988	0.236455	0.487702	-0.25125	0.048664	0.999992	cg2667988	cg2667988	58642406
cg2758679	0.429769	0.678816	-0.24905	0.035195	0.999992	cg2758679	cg2758679	12784395
cg1536550	0.61212	0.859914	-0.24779	0.034901	0.999992	cg1536550	cg1536550	35642301
cg1642309	0.454119	0.695547	-0.24143	0.042855	0.999992	cg1642309	cg1642309	42782411
cg0777476	0.545257	0.786343	-0.24109	0.007408	0.999992	cg0777476	cg0777476	44748386
cg2497656	0.677242	0.9144	-0.23716	0.000726	0.999992	cg2497656	cg2497656	24665479
cg0450634	0.562057	0.799043	-0.23699	0.010389	0.999992	cg0450634	cg0450634	47741327
cg1934487	0.717966	0.948952	-0.23099	0.027555	0.999992	cg1934487	cg1934487	68703407
cg2394713	0.629578	0.859095	-0.22952	0.022554	0.999992	cg2394713	cg2394713	40775474
cg1907951	0.568012	0.794152	-0.22614	0.043828	0.999992	cg1907951	cg1907951	66718372
cg1337008	0.696183	0.919452	-0.22327	0.035625	0.999992	cg1337008	cg1337008	51629325
cg0482944	0.629061	0.850403	-0.22134	0.020011	0.999992	cg0482944	cg0482944	43621482

cg2369827	0.632711	0.842081	-0.20937	0.032221	0.999992	cg2369827	cg2369827	70810465
cg1367916	0.579942	0.787173	-0.20723	0.000851	0.999992	cg1367916	cg1367916	58762470
cg0448683	0.734228	0.939273	-0.20504	0.04044	0.999992	cg0448683	cg0448683	61780346
cg2357685	0.505282	0.709808	-0.20453	0.012287	0.999992	cg2357685	cg2357685	74615430
cg1541173	0.71671	0.921222	-0.20451	0.039182	0.999992	cg1541173	cg1541173	29723310
cg2693049	0.705217	0.909556	-0.20434	0.003532	0.999992	cg2693049	cg2693049	36600316
cg1704224	0.377571	0.581196	-0.20362	0.026914	0.999992	cg1704224	cg1704224	19652463
cg0382729	0.459861	0.660819	-0.20096	0.01673	0.999992	cg0382729	cg0382729	74720388
cg1467135	0.575907	0.375792	0.200115	0.027063	0.999992	cg1467135	cg1467135	33748352
cg0316740	0.848672	0.648112	0.20056	0.02498	0.999992	cg0316740	cg0316740	41609406
cg2201856	0.954853	0.753163	0.201689	0.043565	0.999992	cg2201856	cg2201856	14803393
cg2729222	0.826475	0.623284	0.203191	0.016989	0.999992	cg2729222	cg2729222	41607394
cg0457918	0.825164	0.62187	0.203293	0.046919	0.999992	cg0457918	cg0457918	11749430
cg1681031	0.65546	0.451768	0.203693	0.035013	0.999992	cg1681031	cg1681031	57768347
cg2420858	0.244437	0.040272	0.204165	0.002108	0.999992	cg2420858	cg2420858	52723466
cg1584536	0.27708	0.069441	0.207639	0.026749	0.999992	cg1584536	cg1584536	68737498
cg0146123	0.825087	0.617264	0.207823	0.01631	0.999992	cg0146123	cg0146123	31654378
cg0979888	0.547037	0.336846	0.210191	0.004471	0.999992	cg0979888	cg0979888	52724391
cg1233913	0.948301	0.738073	0.210228	0.001594	0.999992	cg1233913	cg1233913	55756448
cg0230902	0.783653	0.57016	0.213493	0.014514	0.999992	cg0230902	cg0230902	33758486
cg0027464	0.841629	0.626229	0.2154	0.043587	0.999992	cg0027464	cg0027464	47758438
cg0191436	0.948436	0.731168	0.217267	0.014629	0.999992	cg0191436	cg0191436	64666443
cg0768464	0.907695	0.690243	0.217451	0.03079	0.999992	cg0768464	cg0768464	38691370
cg0865722	0.386748	0.16906	0.217688	0.031716	0.999992	cg0865722	cg0865722	67723380
cg0352965	0.628531	0.409944	0.218588	0.034733	0.999992	cg0352965	cg0352965	54618309
cg0255387	0.878973	0.658326	0.220648	0.022832	0.999992	cg0255387	cg0255387	12768430
cg2676476	0.754143	0.533338	0.220806	0.010795	0.999992	cg2676476	cg2676476	37790457
cg1679183	0.858235	0.63495	0.223285	0.02141	0.999992	cg1679183	cg1679183	37644405
cg1327992	0.828004	0.602627	0.225377	0.015159	0.999992	cg1327992	cg1327992	49683476
cg0129687	0.570192	0.340315	0.229877	0.024778	0.999992	cg0129687	cg0129687	29677416
cg0233894	0.959124	0.725499	0.233625	0.026863	0.999992	cg0233894	cg0233894	24753391
cg0365105	0.808251	0.574017	0.234234	0.024905	0.999992	cg0365105	cg0365105	21645475
cg0381217	0.783778	0.548839	0.23494	0.046879	0.999992	cg0381217	cg0381217	53764329
cg1921470	0.504721	0.269661	0.23506	0.047232	0.999992	cg1921470	cg1921470	51609466
cg1303102	0.662808	0.427737	0.235071	0.013885	0.999992	cg1303102	cg1303102	61688421
cg0901915	0.453633	0.21724	0.236393	0.026767	0.999992	cg0901915	cg0901915	73687380
cg2068168	0.809503	0.57275	0.236752	0.006181	0.999992	cg2068168	cg2068168	21696495
cg1188103	0.709616	0.468427	0.241189	0.034223	0.999992	cg1188103	cg1188103	24635354
cg0212689	0.896029	0.652959	0.24307	0.014633	0.999992	cg0212689	cg0212689	21603365
cg0093943	0.928675	0.68282	0.245855	0.003286	0.999992	cg0093943	cg0093943	70718348
cg0919744	0.918631	0.670628	0.248003	0.035546	0.999992	cg0919744	cg0919744	63655325
cg1314387	0.390549	0.141706	0.248843	0.002399	0.999992	cg1314387	cg1314387	24637395
cg0877964	0.8074	0.557188	0.250213	0.021568	0.999992	cg0877964	cg0877964	14604388
cg2360399	0.529474	0.277221	0.252253	0.011315	0.999992	cg2360399	cg2360399	14792375
cg1400175	0.749682	0.491235	0.258447	0.018637	0.999992	cg1400175	cg1400175	11758404
cg1301702	0.947897	0.689225	0.258672	0.003452	0.999992	cg1301702	cg1301702	52689403
cg0938768	0.922428	0.663113	0.259315	0.03727	0.999992	cg0938768	cg0938768	47648483

cg0014624	0.815754	0.554807	0.260947	0.0246	0.999992	cg0014624	cg0014624	44697466
cg2746787	0.693033	0.431724	0.261309	0.02545	0.999992	cg2746787	cg2746787	14632431
cg0289023	0.845479	0.581761	0.263718	0.049368	0.999992	cg0289023	cg0289023	64667483
cg1650756	0.936238	0.672067	0.264171	0.018053	0.999992	cg1650756	cg1650756	31743395
cg0999331	0.6496	0.372745	0.276855	0.039451	0.999992	cg0999331	cg0999331	14642304
cg0205655	0.723422	0.445196	0.278226	0.013002	0.999992	cg0205655	cg0205655	22681329
cg0600268	0.726262	0.447336	0.278926	0.023372	0.999992	cg0600268	cg0600268	24632477
cg1462190	0.745655	0.465995	0.27966	0.018646	0.999992	cg1462190	cg1462190	48627421
cg0183592	0.361672	0.078214	0.283457	0.001109	0.999992	cg0183592	cg0183592	21638313
cg0875045	0.599339	0.310648	0.288692	0.024058	0.999992	cg0875045	cg0875045	65787351
cg0332735	0.74072	0.450433	0.290287	0.01305	0.999992	cg0332735	cg0332735	15736380
cg1193653	0.925907	0.634749	0.291158	0.033668	0.999992	cg1193653	cg1193653	24658402
cg1083223	0.855719	0.55978	0.295939	0.016736	0.999992	cg1083223	cg1083223	57802488
cg0724084	0.83221	0.53567	0.29654	0.025943	0.999992	cg0724084	cg0724084	35619467
cg1158502	0.713611	0.410317	0.303295	0.011149	0.999992	cg1158502	cg1158502	68619457
cg0810398	0.746795	0.442793	0.304002	0.020788	0.999992	cg0810398	cg0810398	29755444
cg2516514	0.775347	0.471297	0.304051	0.004235	0.999992	cg2516514	cg2516514	69721466
cg1220863	0.674749	0.366219	0.30853	0.041976	0.999992	cg1220863	cg1220863	27767440
cg2158700	0.748968	0.428555	0.320413	0.019786	0.999992	cg2158700	cg2158700	66612314
cg0580958	0.579108	0.258094	0.321014	0.019395	0.999992	cg0580958	cg0580958	47748455
cg2135833	0.758876	0.437735	0.321141	0.010715	0.999992	cg2135833	cg2135833	14702477
cg0044354	0.735279	0.402488	0.332791	0.012894	0.999992	cg0044354	cg0044354	12725423
cg0003321	0.541816	0.207825	0.33399	0.000666	0.999992	cg0003321	cg0003321	73611364
cg0461002	0.800062	0.46058	0.339482	0.013614	0.999992	cg0461002	cg0461002	16698312
cg1478255	0.92087	0.580606	0.340264	0.003323	0.999992	cg1478255	cg1478255	50621445
cg1122971	0.498728	0.154903	0.343825	0.011955	0.999992	cg1122971	cg1122971	42796477
cg0105569	0.728802	0.379156	0.349646	0.030341	0.999992	cg0105569	cg0105569	63796358
cg0509381	0.924233	0.574417	0.349816	0.004392	0.999992	cg0509381	cg0509381	68717455
cg0468351	0.66586	0.314219	0.35164	0.010373	0.999992	cg0468351	cg0468351	51686418
cg0978537	0.758268	0.367113	0.391156	0.007481	0.999992	cg0978537	cg0978537	41742374
cg2670559	0.641306	0.215406	0.4259	0.002307	0.999992	cg2670559	cg2670559	44686317
cg1234250	0.670831	0.242	0.428831	0.002112	0.999992	cg1234250	cg1234250	18752444
NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA

AlleleA_Prc	AddressB_I	AlleleB_Prc	Infinium_D	Next_Base	Color_Char	Forward_S	Genome_B	CHR
TAAACCTCCCTATTCTATTCCCTAAA	II					TTTAAAGT	37	11
CATAAACAAAATACACTCCTCTCAA	II					ATCCCGTT	37	8
TCTAATAAACATCTTCCCRAAATAA	II					TGAAATTC	37	1
TCTAATCACCTACAACCTTCTATCAA	II					GCTCCCTT	37	3
TTTATCCCTCCAACTACCCCAAACRC	II					TTGACAAT	37	7
ACATTTCACAAAAACATACAAAAAT	II					TACCCTCA	37	11
AAAACTAACRCRAAAACTTTCCTTTC	II					ACAGCCAA	37	10
CCAAACAAATTTCTACAATATACAA	II					AATCTCTC	37	3
TTTACTCTAACACTAAACACAAAATT	II					GTAAACAG	37	1
TAACAAAT. 43698461 TAACAAAT.	I	A		Red		TTTTCACA	37	4
TAAATAACTATATTATTTTTTTACTC	II					CTTTTTCAT	37	7
TCCCAATATAAATTTCCCCCTAAAAT	II					CCCAGGGA	37	8
ATACCAAACTATAATTATCTCATTCA	II					GGTCTAGC	37	21
TTTCTTAAACAACAAATAAAAACTTC	II					CCCTCAGC	37	11
TTCAAACRACCCAACAAAAAATAAC	II					TACAATGG	37	1
AACACAAATAAAACCTAAACTCATT	II					ACAGCTTG	37	11
AAATTTCTATCAAATTTATATTAATA	II					CATGGTTT	37	7
AAATTCATTCTATTACTCAACAAACA	II					TGGATCCA	37	19
TTTCAATCACAAATTAAAAACATCCA	II					TAAAAAAA	37	15
ACCCAAAAAAAATAACCCAAAACACC	II					ACAAGCTT	37	1
TACTTCCRAAATCRAAAACTATACTA	II					CCACCAGC	37	9
ACACAAAA 29671432 ACACGAAA	I	A		Red		GGGTCAGC	37	8
AATAAAATTAATCTTCTTCCRTTTTT	II					AGGTCTGA	37	3
TACTATTAACTAAAAAAACAACAAA	II					CTGTCACC	37	19
TAAATCTAAACAAAACCTCTAAAAA	II					GCTTAATT	37	1
AATATTTA 56680385 AATATTTA	I	A		Red		CCTGCCTC	37	12
ACAAAAAA 35767456 ACGAAAAA	I	A		Red		CTGTGCCC	37	8
CATTCAAA 70711339 CGTTCGAA	I	C		Grn		GGCGGCTC	37	19
ACTATATACTCCTTTCTACCTTCATA	II					CTGGGATA	37	7
TTCAAAACCTACCCAAAACTAAATA	II					GGAAACAC	37	14
CAAAAAAC 49795427 CGAAAAAC	I	A		Red		TATATGAG	37	19
TCCTAAAACAAATCAATTTCTAAACA	II					GCCTGATA	37	14
CTCCAATTTCTTTCTTTTATAACTAA	II					CCCACCAC	37	8
ATTATTACCTCAACATCTTCCTTAAT	II					GCTCAAAA	37	21
CAAATATACRATCATCAATAATCACC	II					CACTCGAC	37	10
AAAACTTTCTTAACATAACCCTTAAA	II					TGTACTGG	37	5
TAATCAATAAACCTCCTTCATTACAC	II					TTGTGCCA	37	6
TAATATAATCTATAACCACAATTTTC	II					TTCCATTA	37	14
ACCATAAA 21733390 ACCGTAAA	I	C		Grn		GCGATTTT	37	5
TACACAAA 61624432 TACACAAA	I	T		Red		CTGATCTT	37	14
TTCCTTCRAATAAATACAAAAAACA	II					CTGGGATT	37	2
TTTCCAATTCCAACTAATAAATAATT	II					GATGACCA	37	12
TTCAAAATAACATACAAATAATACA	II					CGATGGCC	37	13
TCTACTTCTCTATTCATATATATTTC	II					GTTCTCAC	37	17
ACCCCATATAACCTACCTACATCATA	II					AAACAGGC	37	1
CRACTAAATTTTTTCCTTTTTAATAA	II					TCAACACT	37	5

ACTCCTACCATCTCTTACRACTAATT II			CCTTTCCAT	37	10
AACTAAAT 60697445 AACTAAAT I	T	Red	TCCGCAGC	37	6
ATATTTCTAAACCTCAACTAAATTAA II			GTTCCAGG	37	1
CRATTAACAAAATACTAACAAAACA II			GAGGTTCT	37	5
AATCCCTCCTATAAAAACCAACTTCT II			TGGCATTT,	37	12
CACRTAAAATATAACCCACCCAAAA II			ACACCTCT(	37	16
TCTTAAAAAAAATACCACCATACTA II			ATGGCCAA	37	17
AACRAAAAAACCTCAAAACTCRCCC II			AGGCCCGC	37 Y	
ACATAAATATTACAAACCTCTCTAA/ II			TTGTATGA(	37 Y	
ACAAATCTATAATACTTAAAACATC/ II			GCTGTATG	37	2
TTAAAAACAACCTTACTTAAATATA/ II			TGAGATGT	37	1
TATAATCCCAAAAATCCCTAAACCA/ II			TCTTCAGA(	37	6
AACAACACTAAAAAATTAATAACTC II			GGTTCTTT(	37	15
AACCRAAATCAAAACCACATCCTCT II			GATAATCA	37	11
AAAAAACTAATTATCTTTTAAAAAA II			GAGGTCTC	37	7
AATCTTCCCAATACAAAACTTAACA II			CCAGCCAC	37	5
ACCAACAATCAAAACTCCCTACTTC(II			TTCTTTTCT	37	3
AAAACCAA 33707303 AAAACCGA I	T	Red	TGGGGGG(	37	12
AACATAAACCACCRCACCTAACCTA, II			CTGTGCGA	37	15
ATATTAAT(13769337 ATATTAAT(I	C	Grn	GAAGGCA(	37	10
AAACCCTTCRAAAAAACCRAAACAT II			GAGGAAA/	37	16
ATATTTAATCTTCTCACACAAATTTT II			GAGAAGA/	37	3
CTACCAAACTACTATAATACRCCRA/ II			CAGGCTGC	37	16
AATTATATAAAATAATCCTATCTTAT II			CGTGGGCC	37	20
TCCCTATCTTAATAAATCRATTCTAT II			CTATTGCA,	37	3
AAACCTATTCCTATCTCTTTCTAAAA II			TTGCAGCT	37	16
CAAACCAC 47742450 CAAACCGC I	A	Red	GGCAGGG(	37	16
AAATCCATACAAAACCAAAACACCC II			GTAGAATC	37	16
ATTCAAAA 59679461 ATTCGAAA I	A	Red	TAGGGCCC	37	7
TCTTACTCCTAATCCTCTAAAAATAC II			CAGGCCCC	37	1
CATTCRATTTTTTTCCCTATATATTA/ II			CGGAACAA	37	2
CRATATCTAAACAACRAAATATCTT(II			CCCACTGT(	37	13
ACTCACTCAAAACTTTAATACATTTC II			CAGGGAA(	37	7
AAAACCAAAATCTTACTCAAAAATT, II			TTTCTCCA(	37	7
AAAAAAAACATCACTAAAACAAAA(II			GATGAGA(	37	7
TCACACAACACCCTCTTAAAACAAA II			TTCCCACA	37	8
TTATCCCTACAAACCCTAACTAAAC/ II			CGCTCACC(	37	13
CRAATTAAATCCATTTTCAAACAAA(II			TCCACATC(	37	6
CAAAAACAAAAAAACAAAACCCCA/ II			TGTAGAGC	37	11
AAACAAACAACAAACCTACTCTATA II			GTAGCCAG	37	14
RAATCRTAAAAAATAAATTTCTCCA I II			GAGCGACT	37 Y	
TCCCTTCTTATCTCTCTAATAACTTT/ II			ATGAGTTC	37	2
TAAAACTACTAAATAAAACCAAAAT II			TTTCTCTCC	37	13
TCCTTACCTACTAACTCCACCTACAT II			TCCGTGCT	37	6
TCRCTTTTACTTATATATTAAATTAA, II			TTTATTCTT	37	17
ACCCAATCAATACCATTTATTATACA II			GATGCAAA	37	5
AACTTCCTCRCCTTTACCRTAACATT II			ACGCCAGC	37	10

AATACTCCACRATAAACCACTATTA						GGCACAGA		37	15
TACTACTCCTCACTAAATACCTCACC						GGCTGAGC		37	8
AAATCTACTTTCTCATAATCTAATTC						CCCCAGGT		37	19
CACTCTAATAAAATCCACAAATATA						GAGATGAC		37	6
TAAACTCTAAAATCRTATTCTAACRT						TGGCCTCC		37	10
AACATAAAAACATTATTCTCCCACT						AATGAAAC		37	Y
AACTATTC 52784353 GACTATTC			A	Red		TGCGTGAC		37	8
AAAACCTTAAACCCCCCATAACAAA						GTATCGGC		37	11
TTCTATACCAACRCTTCAACTAACA						GTGAAGAC		37	11
TTCCTACRTATTTTCTATCCAAATCC						TTGATTCC		37	17
AATAAAATAACCAACATTCCTACCT						TCTGGGAA		37	1
RACTACTTCTAATCACTAAATCATTA						TTGGTGGA		37	11
AAAAAAAAACATAAATCTAAACTAA						GACTTTAC		37	14
CAACAACTATTACTAATTCAACTTA						CCTTCTCC		37	10
AAATCAATCTCCTTCAACATCCTAA						CCACTTCA		37	5
CTCCCTCT 67753457 CTCCCTCT			C	Grn		CGGAAACC		37	17
CCCCAATAAAATATATCCACTAAAA						ACTGAGCG		37	12
TTCAAAACCTAAAAAACRATAAATT						GAAAGGC		37	11
AAAAAATCAAAAAACCTAATACCRT						AAGACCTA		37	10
AAAACTAAAAAACACAAACCAACCC						TTCTGATTC		37	21
ATTCCTTTTAAAAATAAAAATCTACC						AGGCCCGC		37	17
AAATTCTAAAAACTAACCTAAAACT						AGCCTGAT		37	17
AAAAATACRATAAACRCCRAAATAA						AGCCCTGC		37	8
TAAAACAACCCTACTCTAAAAACCA						CTTGCCCC		37	19
AAAATCCRTCCTACTAAATTATCAA						CTGGTGTT		37	6
CRATATAACTACTTACAAAAACAAC						TGTATGGC		37	2
CAATAACCTTATCTAATTTATCATTT						TTCACTTG		37	3
AAACATCAAAAATCAATTCCCCAAT						ATACTAAG		37	4
AAAAAAACTTAAATTTTACRTAAAA						GAAACAAA		37	10
ATAATATAACTATTCCACRACAAAC						ACCATAAT		37	15
TAATCRCTATCAAATATACRCCAAA						CTTGGCCT		37	13
CCAATAAAACAACAATCRCTATCAT						GGGCATGA		37	2
NA	NA	NA	NA	NA	NA	NA	NA	NA	
NA	NA	NA	NA	NA	NA	NA	NA	NA	
NA	NA	NA	NA	NA	NA	NA	NA	NA	
NA	NA	NA	NA	NA	NA	NA	NA	NA	
NA	NA	NA	NA	NA	NA	NA	NA	NA	
NA	NA	NA	NA	NA	NA	NA	NA	NA	

MAPINFO	SourceSeq	Chromosor	Coordinate	Strand	Probe_SNP	Probe_SNP	Random_L	Methyl27_
14495049	CGGCCTAA	11	14451625	F		rs5566583	NA	NA
1.44E+08	ATGGGCAC	8	1.45E+08	R			NA	NA
2.48E+08	CGGCTCTC	1	2.46E+08	R			NA	NA
1.12E+08	CGGCTGGA	3	1.13E+08	R			NA	NA
1.43E+08	TTGTCCCT(	7	1.43E+08	F		rs5827052	NA	NA
15180922	CATTTCAC/	11	15137498	R			NA	NA
3179703	CGCTTGCT	10	3169703	R		rs4881107	NA	NA
1.68E+08	CGGCCTTT(	3	1.7E+08	R			NA	NA
2.47E+08	CGGTTCTA	1	2.45E+08	F		rs3129547	NA	NA
40267141	TGACAAGT	4	39943536	F			NA	NA
66785137	GGATAGCT	7	66422572	F		rs9443	NA	NA
22133375	CGGACAGC	8	22189320	F		rs6982089	NA	NA
46075092	CGGCCACC	21	44899520	R		rs2838613	NA	NA
66362959	TTCTTAGG(	11	66119535	R			NA	NA
3606550	TCAGACGG	1	3596410	R		rs3765725	NA	NA
14431708	CGCCACAT	11	14388284	F		rs1102321	NA	NA
1.03E+08	CGGTGACC	7	1.03E+08	R			TRUE	NA
3373819	CGGCATGC	19	3324819	F		rs1297412	NA	NA
28644585	CGCCTAGG	15	26318180	F		rs6200697	NA	NA
1562535	CGCCACCT(	1	1552398	R			NA	NA
1.29E+08	ACTTCCGG	9	1.28E+08	R			NA	NA
216453	ACACGGGC	8	206453	R		rs5709037	NA	NA
1.27E+08	CGGGAGG(	3	1.28E+08	F			NA	NA
1792217	GCTGTTGG	19	1743217	F		rs7258947	NA	NA
8120055	AAGTCTGA	1	8042642	F		rs7263424	NA	NA
1.33E+08	GATGTTTG	12	1.31E+08	F		rs4076044	NA	NA
216659	GCGGAGG(	8	206659	R			NA	NA
13875111	CGTTCGGG	19	13736111	R		rs3620361	NA	NA
77827379	CTGTGTGC	7	77665315	R			NA	NA
23018807	CGGTGAG/	14	22088647	R			NA	NA
36265700	CGAGAAAC	19	40957540	R			NA	NA
26862142	CGGCTAAG	14	25931982	F		rs7359792	NA	NA
1.25E+08	CGGAAATT	8	1.25E+08	R			NA	NA
46077562	TTATTGCC1	21	44901990	R			NA	NA
94136	CGCTTGGC	10	84136	R		rs4607995	NA	NA
13664584	GGGCTTTC	5	13717584	R		rs870546	NA	NA
1.49E+08	GGTCAATG	6	1.49E+08	F		rs4897076	NA	NA
22279816	CGTGATCT	14	21349656	R		rs2856411	NA	NA
1.8E+08	ACCGTGAG	5	1.8E+08	F		rs1043480	NA	NA
24587638	CGGAAACC	14	23657478	R		rs7784824	NA	NA
1.6E+08	TCCTTCGG/	2	1.6E+08	F		rs1020453	NA	NA
1.31E+08	CGGCATGT	12	1.3E+08	R			NA	NA
1.15E+08	CGACGAG/	13	1.14E+08	R			NA	NA
14936230	CTGCTTCT(	17	14876955	R			NA	NA
55260950	CCCCATGT(	1	55033538	R			NA	NA
1.47E+08	GGCTAAAT	5	1.47E+08	F		rs7700488	NA	NA

1.21E+08	CTCCTGCC/	10	1.21E+08	R		NA	NA
1.61E+08	CGGGAAG(	6	1.61E+08	R		NA	NA
1.61E+08	TATTTCTG(	1	1.6E+08	R		NA	NA
373299	GGTTGGCA	5	426299	R		NA	NA
9886905	CGGTGCTC	12	9778172	R	rs1084463:	NA	NA
71458897	ACGTGGAA	16	70016398	R	rs7629132:	NA	NA
77136834	CTTGGGGA	17	74648429	R	rs2707040	NA	NA
24454927	GCGGAGG(	Y	22864315	R		NA	NA
7676811	CGTGAGTA	Y	7736811	R		NA	NA
2.41E+08	CAGGTCTG	2	2.41E+08	R	rs2352821	NA	NA
1.12E+08	TGAGGGCA	1	1.12E+08	R	rs7737	NA	NA
30980847	GTGATCCC	6	31088826	R	rs7527209(	NA	NA
88119834	CGGGCCAC	15	85920838	R		NA	NA
66104993	GCCGAAGT	11	65861569	F	rs3116068	NA	NA
22122872	GGGGACT(	7	22089397	R		NA	NA
16785399	ATCTTCCC/	5	16838399	R		NA	NA
1.91E+08	CCAGCAAT	3	1.93E+08	R	rs7900178(	NA	NA
1.08E+08	CGGGGCCA	12	1.07E+08	F		NA	NA
90927939	CGATTGAA	15	88728943	F	rs9796504	NA	NA
855060	CGGGGCC(	10	845060	R		NA	NA
88238863	CGTTGGCT	16	86796364	F	rs1244663	NA	NA
32511650	CGCTGTAT.	3	32486654	R	rs421653	NA	NA
88296994	TGCCAGGC	16	86854495	R	rs2873329(	NA	NA
170641	GTTGTGTG	20	118641	R		NA	NA
1.04E+08	CGCCCAAT.	3	1.06E+08	R	rs6226057(	NA	NA
28273096	GGCCTATT	16	28180597	F	rs4788054	NA	NA
87682142	CAGGCCGC	16	86239643	R		NA	NA
73102243	AATCCATA(	16	71659744	F	rs1183967	NA	NA
1.57E+08	GTTCGGGA	7	1.57E+08	R		NA	NA
2274955	CGAGGGG(	1	2264815	R		NA	NA
2.25E+08	ATTCGGTT	2	2.25E+08	R	rs1686590:	NA	NA
50194643	GGTGTCTG	13	49092644	F		NA	NA
44184403	CTCACTCA(	7	44150928	F	rs1330638(	NA	NA
3157722	CGCCTCCC/	7	3124248	F		NA	NA
1.08E+08	GAGGAAG(	7	1.07E+08	R		NA	NA
19616280	CACACAGC	8	19660560	R		NA	NA
1.13E+08	TGTCCCTG(	13	1.12E+08	R		NA	NA
1.54E+08	CGCATGGT	6	1.54E+08	R		NA	NA
13938802	CGTGCATA	11	13895378	R	rs7131580	NA	NA
55151579	GGCAGGCA	14	54221329	R		NA	NA
14107195	CGTGCAGC	Y	12617195	F		NA	NA
2.01E+08	CCCTTCTTA	2	2E+08	R		NA	NA
50194554	CGCCTCTG(	13	49092555	R		NA	NA
1.57E+08	CCTTGCCT(	6	1.57E+08	R	rs7582559(	NA	NA
64302651	CGGCCTAA	17	61733113	R		NA	NA
1.4E+08	CCCAGTCA	5	1.4E+08	F	rs6849	NA	NA
13688165	ACTTCCTC(	10	13728171	R		NA	NA

62899159	GTGCTCCA		15	60686451	R		NA	NA
22266134	GCTGCTCC		8	22322079	R		NA	NA
55735946	GGTCTGCT		19	60427758	F	rs1261080	TRUE	NA
33158020	ACTCTGAT		6	33265998	F	rs7471695	NA	NA
1.32E+08	GAGCTCTG		10	1.31E+08	R		NA	NA
8552374	CGGGAATC	Y		8612374	F		NA	NA
80783997	CGGGGATT		8	80946552	R		NA	NA
65363274	GGGCCTTA		11	65119850	R		NA	NA
1.22E+08	TCTGTGCC		11	1.21E+08	F	rs1944694	NA	NA
6558815	TCCTGCGT		17	6499539	R		NA	NA
46979222	CGGGCCTC		1	46751809	R	rs4660355	NA	NA
1.34E+08	CGTCAAAG		11	1.33E+08	F	rs7303206	NA	NA
1.06E+08	GAAGGAG(		14	1.05E+08	R		NA	NA
12438782	CGGGGAG(		10	12478788	R		NA	NA
1.1E+08	CGGGCAGT		5	1.1E+08	R		NA	NA
6558365	CGGAGCCA		17	6499089	R		NA	NA
670974	CGCCAAGC		12	541235	F	rs3497962	NA	NA
66317822	TCAGGACC		11	66074398	R		NA	NA
839609	GAGAATCA		10	829609	R		NA	NA
31709690	AGGCTGAC		21	30631561	R		NA	NA
6558440	CGCCTGAG		17	6499164	F		NA	NA
1645410	CGGGCTGC		17	1592160	F	rs6209005	NA	NA
1.44E+08	CGCACAGA		8	1.44E+08	R		NA	NA
8464538	GGGACAG(		19	8370538	R		NA	NA
33131893	CGGCTTTT		6	33239871	F	rs9405002	NA	NA
2.25E+08	GATGTAAC		2	2.25E+08	R		NA	NA
1.15E+08	CGGACTAT		3	1.17E+08	F	rs1093430	NA	NA
87282697	CGACAAAC		4	87501721	R		NA	NA
5047487	CGGCAGGT		10	5037487	R		NA	NA
60644157	CGGTGACC		15	58431449	F	rs1163365	NA	NA
1.12E+08	GGTCGCTG		13	1.11E+08	R		NA	NA
8530521	CGGAGGA(		2	8447972	R		NA	NA
NA	NA	NA		NA	NA	NA	NA	NA	NA
NA	NA	NA		NA	NA	NA	NA	NA	NA
NA	NA	NA		NA	NA	NA	NA	NA	NA
NA	NA	NA		NA	NA	NA	NA	NA	NA
NA	NA	NA		NA	NA	NA	NA	NA	NA
NA	NA	NA		NA	NA	NA	NA	NA	NA

UCSC_RefG	UCSC_RefG	UCSC_RefG	UCSC_CpG_	Relation_to	Phantom	DMR	Enhancer	HMM_Islar
COPB1;COI	NM_01645	Body;Body;Body					NA	
RHPN1	NM_05292	Body	chr8:14445	S_Shore			NA	8:1445288(
			chr1:2478(	S_Shore			NA	1:2458702(
TAGLN3;TA	NM_00100	Body;Body;Body					TRUE	
TAS2R60	NM_17743	TSS200					NA	
INSC;INSC	NM_00103	Body;Body					TRUE	
			chr10:318(	N_Shore			NA	10:316969:
C3orf50	NR_02148	Body					TRUE	
CNST;CNST	NM_00113	Body;Body	chr1:2467E	S_Shore			TRUE	
							TRUE	
STAG3L4	NM_02290	3'UTR					NA	
PIWIL2;PIW	NM_01806	5'UTR;1stE	chr8:22132	S_Shore			NA	
KRTAP12-4	NM_19869	TSS1500;Body					NA	
CCS	NM_00512	Body	chr11:663E	S_Shelf			NA	
TP73;TP73	NM_00112	TSS1500;TS	chr1:3607(	N_Shore			NA	
							TRUE	
RELN;RELN	NM_17305	Body;Body					NA	
NFIC;NFIC	NM_00559	Body;Body	chr19:3369	S_Shelf			NA	
			chr15:2864	N_Shelf			NA	
MIB2;MIB2	NM_00117	Body;Body	chr1:1563E	N_Shore			NA	1:1551913-
LMX1B	NM_00231	Body	chr9:12937	Island		DMR	TRUE	9:1284157
			chr8:21635	Island			NA	8:206391-2
							TRUE	
ATP8B3	NM_13881	Body	chr19:1795	N_Shelf			NA	
							TRUE	
			chr12:1325	Island			NA	12:131482:
			chr8:21635	Island			NA	8:206391-2
MRI1;MRI1	NM_00103	TSS1500;TS	chr19:1387	Island			NA	19:137360!
MAGI2	NM_01230	Body					TRUE	
							TRUE	
SNX26	NM_05294	TSS1500	chr19:362E	N_Shore			NA	
							TRUE	
							TRUE	
C21orf29;K	NM_14499	Body;TSS1500					NA	
TUBB8;TUE	NM_00116	Body;Body	chr10:9452	N_Shore			NA	10:83977-8
							TRUE	
UST	NM_00571	Body					TRUE	
							NA	
BTNL9	NM_15254	3'UTR	chr5:18048	S_Shore			NA	5:1804187(
DCAF11;DC	NM_18135	Body;Body	chr14:2458	S_Shelf			NA	
BAZ2B	NM_01345	5'UTR					TRUE	
			chr12:1313	N_Shelf			NA	
RASA3	NM_00736	Body	chr13:1147	Island			NA	13:113800(
							TRUE	
TTC22;TTC	NM_00111	Body;Body					TRUE	
SPINK5;SPI	NM_00112	Body;Body;Body					NA	

TIAL1;TIAL:	NM_00103	Body;Body			TRUE	
					NA	6:1611784:
C1orf192	NM_00101	3'UTR			NA	
AHRR	NM_02073	Body	chr5:37384	N_Shore	TRUE	
CLECL1	NM_17200	TSS1500			NA	
			chr16:7145	N_Shore	NA	
HRNBP3	NM_00108	5'UTR	chr17:7713	S_Shelf	NA	
RBMY1F;RI	NM_15258	TSS200;TSS	chrY:24454	Island	NA	Y:2286417
TTTY12	NR_00155:	Body			NA	
			chr2:24126	N_Shore	NA	
ADORA3;A	NM_00108	3'UTR;3'UTR			NA	
					NA	
NCRNA000	NR_02686	TSS1500			NA	
BRMS1;BRI	NM_01539	3'UTR;3'UT	chr11:661(	S_Shelf	NA	
			chr7:22122	Island	TRUE	7:2208893
MYO10	NM_01233	Body			TRUE	
CCDC50;CC	NM_17833	Body;Body	chr3:19104	S_Shelf	NA	
			chr12:1082	S_Shore	NA	12:106763:
			chr15:9093	N_Shelf	NA	
			chr10:8556	N_Shore	NA	
			chr16:8823	S_Shore	NA	16:867963!
			chr3:3250!	S_Shelf	NA	
			chr16:8825	S_Shelf	NA	16:868542:
DEFB128	NM_00103	TSS1500			NA	
					TRUE	
			chr16:2827	S_Shelf	TRUE	
JPH3	NM_02065	Body	chr16:8767	S_Shelf	NA	16:862395(
			chr16:7305	S_Shore	TRUE	
PTPRN2;PT	NM_00284	Body;Body;Body			NA	7:1571566
MORN1	NM_02484	Body	chr1:22763	N_Shore	NA	
FAM124B;I	NM_02478	3'UTR;3'UTR			NA	
					NA	13:490924:
GCK;GCK;G	NM_03350	3'UTR;3'UT	chr7:44184	N_Shore	NA	
					TRUE	
LAMB1	NM_00229	Body			TRUE	
			chr8:19614	S_Shore	CDMR NA	
			chr13:1128	Island	NA	13:111896!
OPRM1;OP	NM_00114	Body;Body;Body;1stExon;Body;Body;5'UTR;Body;Body;B			NA	
					TRUE	
SAMD4A;S.	NM_00116	Body;Body			TRUE	
			chrY:14107	Island	NA	Y:1261713
C2orf69	NM_15368	Body	chr2:20077	S_Shelf	NA	
					NA	13:490924:
ARID1B;AR	NM_01751	Body;Body;Body			TRUE	
PRKCA	NM_00273	Body	chr17:642!	S_Shelf	NA	
ZMAT2	NM_14472	3'UTR			NA	
FRMD4A	NM_01802	3'UTR		low-CpG:13728022-13	NA	

							TRUE	
SLC39A14;S	NM_01535	Body;Body;Body;Body					NA	8:2232190
							NA	19:604274
COL11A2;C	NM_08067	Body;Body	chr6:33159	N_Shore			NA	
MGMT	NM_00241	Body					TRUE	
TTTY18	NR_00155	TSS1500					NA	
							NA	
KCNK7;KCN	NM_03334	1stExon;5'U	chr11:6536	N_Shelf			NA	
							TRUE	
			chr17:6559	N_Shore			NA	17:649869
DMBX1;DN	NM_14719	3'UTR;3'UTR					NA	
			chr11:1337	N_Shelf		RDMR	NA	
			chr14:1061	N_Shore			NA	
CAMK1D;C	NM_02039	Body;Body					TRUE	
WDR36	NM_13928	TSS1500	chr5:11042	N_Shore		CDMR	NA	
			chr17:6558	Island			NA	17:649869
B4GALNT3	NM_17359	3'UTR					NA	12:541125
ACTN3	NM_00110	Body	chr11:6631	S_Shelf			NA	
			chr10:8396	Island			NA	
KRTAP27-1	NM_00107	1stExon					NA	
			chr17:6558	Island			NA	17:649869
SERPINF2;S	NM_00093	TSS1500;TSS1500;TSS1500					NA	
TOP1MT	NM_05296	Body	chr8:14439	Island			NA	
RAB11B	NM_00421	Body	chr19:8464	N_Shore			NA	
COL11A2;C	NM_08067	Body;Body	chr6:33129	S_Shelf			NA	
CUL3	NM_00359	Body					TRUE	
GAP43;GAI	NM_00204	Body;5'UTI	chr3:11537	N_Shore			NA	
MAPK10;M	NM_13898	TSS1500;TSS1500;5'UTR					NA	
AKR1C2;AK	NM_00113	TSS1500;5'UTR;5'UTR					NA	
ANXA2;AN	NM_00113	Body;Body;Body;Body					TRUE	
			chr13:1117	Island			NA	13:110543
							TRUE	
NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA

Regulatory	Regulatory	DHS
		NA
)3-144528896		NA
50-245870330		NA
		NA
		NA
11:151807	Unclassifie	NA
1-3169822		TRUE
		NA
		NA
4:4026692	Unclassifie	NA
		NA
		NA
		NA
		NA
		NA
		NA
		NA
		NA
1552467		NA
32-128417915		NA
06774		NA
		NA
		NA
		NA
196-131483014		NA
06774		NA
53-13736951		NA
		NA
		NA
19:362656	Unclassifie	NA
		NA
		NA
21:460773	Unclassifie	NA
4317		NA
		TRUE
		NA
		NA
51-180419744		NA
		NA
		NA
		NA
13:114782	Gene_Asso	TRUE
		NA
		TRUE
		NA

	NA
11-161178494	NA
	NA
	NA
	NA
16:714587	Unclassifie NA
	NA
1-22864502	NA
	NA
	NA
	NA
	NA
	NA
	NA
7:2212206	Promoter_ NA
	NA
	NA
202-106763970	NA
	NA
	NA
56-86796487	TRUE
	TRUE
17-86854496	NA
	NA
	NA
	NA
05-86239722	TRUE
	TRUE
7:1574641	Unclassifie NA
	NA
	NA
10-49092680	NA
	NA
	NA
	NA
8:1961507	Promoter_ NA
577-111897215	NA
	NA
	NA
	NA
3-12617480	NA
	NA
10-49092680	NA
	NA
	NA
	NA
	NA

		NA
3-22322173		NA
)8-60427866		NA
		NA
		NA
		NA
		NA
11:653626	Unclassifie	TRUE
		NA
17:655871	Unclassifie	NA
1:4697902	Unclassifie	NA
		NA
		TRUE
		NA
5:1104272	Promoter_	NA
17:655834	Unclassifie	NA
-541330		NA
		NA
		NA
		NA
17:655834	Unclassifie	NA
		NA
		TRUE
		NA
		NA
		NA
		NA
		NA
		NA
		NA
554-110543657		NA
		NA
NA	NA	NA
NA	NA	NA
NA	NA	NA
NA	NA	NA
NA	NA	NA
NA	NA	NA

Appendix H

Preliminary Data

Differential DNA Methylation in Genes Associated with Immune Function in Pregnant Women with Group B Streptococcus Colonization

Michelle L. Wright[1], Cindy M. Anderson[2], and Joyce E. Ohm[1]
[1]University of North Dakota, Grand Forks, ND; [2]The Ohio State University, Columbus, OH

Introduction

- Group B streptococcus (GBS) is the leading cause of infectious neonatal morbidity and mortality in the United States
- Maternal GBS colonization is the primary risk factor associated with neonatal infection
  - Ambiguous maternal risk factors for colonization
  - 10-30% of pregnant women are colonized with GBS
- DNA methylation is an epigenomic modification that can be altered by environmental exposures resulting in:
  - Altered gene expression
  - Predisposition to disease
- Purpose:
  - Determine feasibility of evaluating DNA methylation patterns as a contributing factor to GBS colonization

Methods

- DNA methylation in maternal peripheral white blood cells collected in first trimester of pregnancy were analyzed using Illumina Infinium®
- Group differences in DNA methylation patterns in women with and without GBS colonization were determined with:
  - Delta-beta criteria >0.2 or <-0.2 to determine change in gain or loss in methylation, respectively.
  - Significance determined by t-test of mean methylation difference, $p<0.05$
- Functional relevance of differentially methylated genes was determined by:
  - GeneCards®, DAVID Bioinformatics Resources v6.7 (Huang, et al., 2009), PubMed
  - Utilized "high" classification stringency
  - Reference: Huang, D. W., Sherman, B. T., & Lempicki, R. a. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols*, *4*(1), 44–57. doi:10.1038/nprot.2008.211

Preliminary Results

- Results are from a secondary analysis of DNA methylation were used to determine feasibility for a larger study
- Preliminary results of differential DNA methylation resulted in >1,000 CpG (phosphate linked cytosine-guanine base pairs) sites with >20% difference in methylation in GBS positive (n=2) versus GBS negative (n=4) women
  - 350 CpG sites had >30% difference in methylation **(Figure 1)**
  - Differential methylation most strongly associated with immune function pathways **(Table 1)**

Figure 1. Potential DNA methylation biomarkers associated with maternal GBS status. Average beta scores from 350 maternal DNA CpG (phosphate linked cytosine-guanine base pairs) dinucleotides with potential differential methylation in maternal peripheral blood delta beta > 0.3 or -0.3 are shown for 4 GBS negative and 2 GBS positive women

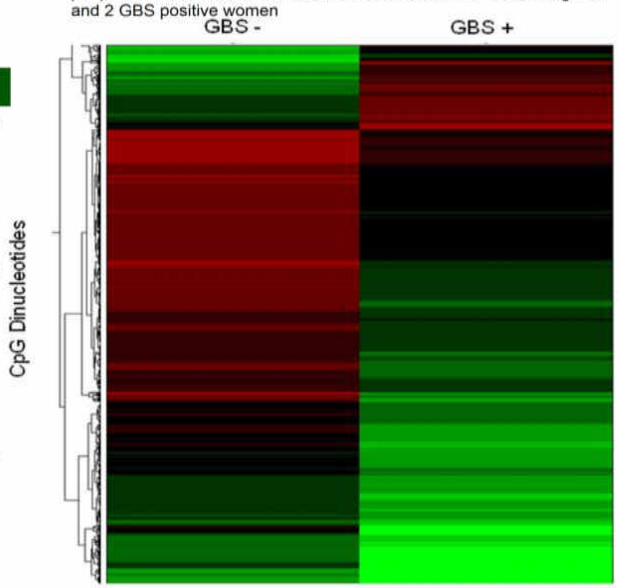


Table 1. Functional analysis of differential DNA methylation

Cluster	*Enrichment Score*	*Gene members in Functional cluster*
Methylation Loss		
1	3.96	MHC protein complex
		Antigen processing and presenting
		Immunoglobulin C1-set
		Immunoglobulin/MHC , conserved site
		Immune response
2	3.79	MHC class II receptor activity
		MHC class II, alpha/beta chain, N-terminal
		MHC class II protein complex
		Antigen processing and presenting
		Intestinal immune network for IgA production
		Cell adhesion molecules
Methylation Gain		
1	2.41	Pleckstrin homology
2	2.31	MHC class II receptor activity
		MHC class II, alpha/beta chain, N-terminal
		MHC protein complex
		Immunoglobulin C1-set
		MHC class II protein complex
		Antigen processing and presenting
		Cell adhesion molecules
		Intestinal immune network for IgA production
		Immune response

Conclusions

- Preliminary data suggests:
  - Differentially methylated genes between women with and without GBS colonization are tightly clustered in immune function pathways
  - Differentially methylated CpG dinucleotides associated with immune function may contribute to GBS colonization susceptibility

Acknowledgements

- Funding for this project provided by:
  - Robert Wood Johnson Nurse Faculty Scholar Award (CMA, 64202)
  - International Society of Nurses in Genetics Research Grant (MLW)
  - Midwest Nursing Research Society Dissertation Grant (MLW)

References

Abbas, A. K., Lichtman, A. H. H., & Pillai, S. (2012). *Cellular and Molecular Immunology* (7th ed.). Philadelphia PA: Elsevier Saunders Inc.

Absher, D. M., Li, X., Waite, L. L., Gibson, A., Roberts, K., Edberg, J., … Kimberly, R. P. (2013). Genome-wide DNA methylation analysis of systemic lupus erythematosus reveals persistent hypomethylation of interferon genes and compositional changes to CD4+ T-cell populations. *PLoS Genetics*, *9*(8), e1003678. doi:10.1371/journal.pgen.1003678

ACOG, C. on O. P. (2011). ACOG Committee Opinion No. 495: Vitamin D: screening and supplementation during pregnancy. *Obstet Gynecol*, *118*(1), 197–8. doi:10.1097/AOG.0b013e318227f06b

Allison, D. B., Cui, X., Page, G. P., & Sabripour, M. (2006). Microarray data analysis: from disarray to consolidation and consensus. *Nature Reviews. Genetics*, *7*(1), 55–65. doi:10.1038/nrg1749

Andérica-Romero, A. C., González-Herrera, I. G., Santamaría, A., & Pedraza-Chaverri, J. (2013). Cullin 3 as a novel target in diverse pathologies. *Redox Biology*, *1*(1), 366–372. doi:10.1016/j.redox.2013.07.003

Anderson, C. M., Ralph, J. L., Wright, M. L., Linggi, B., & Ohm, J. E. (2013). DNA methylation as a biomarker for preeclampsia. *Biological Research for Nursing*. doi:10.1177/1099800413508645

Ashkar, A. A., Di Santo, J. P., & Croy, B. A. (2000). Interferon gamma contributes to initiation of uterine vascular modification, decidual integrity, and uterine natural killer cell maturation during normal murine pregnancy. *The Journal of Experimental Medicine*, *192*(2), 259–70. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2193246&tool=pmcentrez&rendertype=abstract

Baccarelli, A., Rienstra, M., & Benjamin, E. J. (2010). Cardiovascular epigenetics: basic concepts and results from animal and human studies. *Circulation.Cardiovascular genetics*, *3*(6), 567–573. doi:10.1161/CIRCGENETICS.110.958744

Baccarelli, Andrea, Wright, R. O., Bollati, V., Tarantini, L., Litonjua, A. a, Suh, H. H., Zanobetti, A., et al. (2009). Rapid DNA methylation changes after exposure to traffic particles. *American journal of respiratory and critical care medicine*, *179*(7), 572–8. doi:10.1164/rccm.200807-1097OC

Back, E. E., O'grady, E. J., & Back, J. D. (2011). High rates of perinatal Group B Streptococcus clindamycin and erythromycin resistance in an Upstate NY hospital. *Antimicrobial Agents and Chemotherapy*. doi:10.1128/AAC.05794-11

Baker, C., Rench, M., & McInnes, P. (2003). Immunization of pregnant women with group B streptococcal type III capsular polysaccharide-tetanus toxoid conjugate vaccine. *Vaccine*, *21*(24), 3468–3472. doi:10.1016/S0264-410X(03)00353-0

Barber, E. L., Zhao, G., Buhimschi, I. A., & Illuzzi, J. L. (2008). Duration of intrapartum prophylaxis and concentration of penicillin G in fetal serum at delivery. *Obstetrics and gynecology*, *112*(2 Pt 1), 265–270. doi:10.1097/AOG.0b013e31817d0246

Barker, D. J. (1995). Fetal origins of coronary heart disease. *BMJ (Clinical research ed.)*, *311*(6998), 171–174.

Barker, D. J. P., & Osmond, C. (1986). INFANT MORTALITY, CHILDHOOD NUTRITION, AND ISCHAEMIC HEART DISEASE IN ENGLAND AND WALES. *The Lancet*, *327*(8489), 1077–1081.

Baron, U., Floess, S., Wieczorek, G., Baumann, K., Grützkau, A., Dong, J., Thiel, A., et al. (2007). DNA demethylation in the human FOXP3 locus discriminates regulatory T cells from activated FOXP3(+) conventional T cells. *European journal of immunology*, *37*(9), 2378–89. doi:10.1002/eji.200737594

Barrès, R., Yan, J., Egan, B., Treebak, J. T., Rasmussen, M., Fritz, T., Caidahl, K., et al. (2012). Acute exercise remodels promoter methylation in human skeletal muscle. *Cell metabolism*, *15*(3), 405–11. doi:10.1016/j.cmet.2012.01.001

Bebien, M., Hensler, M., Davanture, S., & Hsu, L. (2012). The Pore-Forming Toxin beta hemolysin/cytolysin Triggers p38 MAPK-Dependent IL-10 Production in Macrophages and Inhibits Innate Immunity. *PLoS pathogens*, *8*(7), e1002812. doi:10.1371/journal.ppat.1002812

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B ...*, *57*(1), 289–300. Retrieved from http://www.jstor.org/stable/2346101

Berardi, A., Rossi, C., Lugli, L., Creti, R., Bacchi Reggiani, M. L., Lanari, M., Memo, L., et al. (2013). Group B Streptococcus Late-Onset Disease: 2003-2010. *Pediatrics*. doi:10.1542/peds.2012-1231

Berger, S. L., Kouzarides, T., Shiekhattar, R., & Shilatifard, A. (2009). An operational definition of epigenetics. *Genes & development*, *23*(7), 781–783. doi:10.1101/gad.1787609

Bergman, Y., & Cedar, H. (2013). DNA methylation dynamics in health and disease. *Nature Structural & Molecular Biology*, *20*(3), 274–81. doi:10.1038/nsmb.2518

Bergseng, H., Rygg, M., Bevanger, L., & Bergh, K. (2008). Invasive group B streptococcus (GBS) disease in Norway 1996-2006. *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology*, *27*(12), 1193–1199. doi:10.1007/s10096-008-0565-8

Bergseng, Hakon, Bevanger, L., Rygg, M., & Bergh, K. (2007). Real-time PCR targeting the sip gene for detection of group B Streptococcus colonization in pregnant women at delivery. *Journal of medical microbiology*, *56*(Pt 2), 223–8. doi:10.1099/jmm.0.46731-0

Berner, R., Welter, P., & Brandis, M. (2002). Cytokine expression of cord and adult blood mononuclear cells in response to Streptococcus agalactiae. *Pediatric research*, *51*(3), 304–309.

Bibikova, M., Barnes, B., Tsan, C., Ho, V., Klotzle, B., Le, J. M., … Shen, R. (2011). High density DNA methylation array with single CpG site resolution. *Genomics*, *98*(4), 288–95. doi:10.1016/j.ygeno.2011.07.007

Bloom, S. L., Cox, S. M., Bawdon, R. E., & Gilstrap, L. C. (1996). Ampicillin for neonatal group B streptococcal prophylaxis: how rapidly can bactericidal concentrations be achieved? *American Journal of Obstetrics and Gynecology*, *175*(4 Pt 1), 974–976.

Bobetsis, Y. a, Barros, S. P., Lin, D. M., Weidman, J. R., Dolinoy, D. C., Jirtle, R. L., Boggess, K. a, et al. (2007). Bacterial infection promotes DNA hypermethylation. *Journal of dental research*, *86*(2), 169–74. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17251518

Boks, M. P., Derks, E. M., Weisenberger, D. J., Strengman, E., Janson, E., Sommer, I. E., … Ophoff, R. A. (2009). The relationship of DNA methylation with age, gender and genotype in twins and healthy controls. *PloS One*, *4*(8), e6767. doi:10.1371/journal.pone.0006767

Bretz, W. A., Biesbrock, A., Corby, P. M., Corby, A. L., Bretz, W. G., Wessel, J., & Schork, N. J. (2011). Environmental and genetic contributions to indicators of oral malodor in twins. *Twin research and human genetics : the official journal of the International Society for Twin Studies*, *14*(6), 568–572.

Brimil, N., Barthell, E., Heindrichs, U., Kuhn, M., Lütticken, R., & Spellerberg, B. (2006). Epidemiology of Streptococcus agalactiae colonization in Germany. *International journal of medical microbiology : IJMM*, *296*(1), 39–44. doi:10.1016/j.ijmm.2005.11.001

Brochet, M., Couvé, E., Bercion, R., Sire, J.-M., & Glaser, P. (2009). Population structure of human isolates of Streptococcus agalactiae from Dakar and Bangui. *Journal of clinical microbiology*, *47*(3), 800–3. doi:10.1128/JCM.01103-08

Brogin Moreli, J., Cirino Ruocco, A. M., Vernini, J. M., Rudge, M. V. C., & Calderon, I. M. P. (2012). Interleukin 10 and tumor necrosis factor-alpha in pregnancy: aspects of interest in clinical obstetrics. *ISRN obstetrics and gynecology*, *2012*, 230742. doi:10.5402/2012/230742

Bromer, J G, Wu, J., Zhou, Y., & Taylor, H. S. (2009). Hypermethylation of homeobox A10 by in utero diethylstilbestrol exposure: an epigenetic mechanism for altered developmental programming. *Endocrinology*, *150*(7), 3376–3382. doi:10.1210/en.2009-0071

Bromer, Jason G, Zhou, Y., Taylor, M. B., Doherty, L., & Taylor, H. S. (2010). Bisphenol-A exposure in utero leads to epigenetic alterations in the developmental programming of uterine estrogen response. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, *24*(7), 2273–80. doi:10.1096/fj.09-140533

Brown, C. K., Gu, Z.-Y., Matsuka, Y. V, Purushothaman, S. S., Winter, L. a, Cleary, P. P., Olmsted, S. B., et al. (2005). Structure of the streptococcal cell wall C5a peptidase. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(51), 18391–6. doi:10.1073/pnas.0504954102

Castor, M. L., Whitney, C. G., Como-Sabetti, K., Facklam, R. R., Ferrieri, P., Bartkus, J. M., Juni, B. a, et al. (2008). Antibiotic resistance patterns in invasive group B streptococcal isolates. *Infectious diseases in obstetrics and gynecology*, *2008*, 727505. doi:10.1155/2008/727505

CDC. (2010). Group B Strep Laboratory Photo Examples. *National Center for Immunization and Respiratory Diseases, Division of Bacterial Diseases* (March 16, 2012). Retrieved from http://www.cdc.gov/groupbstrep/lab/lab-photos.html

Cedar, H., & Bergman, Y. (2012). Programming of DNA methylation patterns. *Annual Review of Biochemistry*, *81*, 97–117. doi:10.1146/annurev-biochem-052610-091920

Chan, K. Y., Ozcelik, H., Cheung, A. N., Ngan, H. Y., & Khoo, U. S. (2002). Epigenetic factors controlling the BRCA1 and BRCA2 genes in sporadic ovarian cancer. *Cancer research*, *62*(14), 4151–4156.

Cheng, Q. I., Carlson, B., Pillai, S. U. B., Eby, R. O. N., Edwards, L., Olmsted, S. B., Cleary, P., et al. (2001). Antibody against Surface-Bound C5a Peptidase Is Opsonic and Initiates Macrophage Killing of Group B Streptococci. *Infection and immunity*, *69*(4), 2302–2308. doi:10.1128/IAI.69.4.2302

Cheng, Q., Stafslien, D., Purushothaman, S. S., & Cleary, P. (2002). The Group B Streptococcal C5a Peptidase Is Both a Specific Protease and an Invasin. *Infection and immunity*, *70*(5), 2408–2413. doi:10.1128/IAI.70.5.2408

Chen, Z., & Riggs, A. D. (2011). DNA methylation and demethylation in mammals. *The Journal of Biological Chemistry*, *286*(21), 18347–53. doi:10.1074/jbc.R110.205286

Chesney, R. W. (2010). Vitamin D and the Magic Mountain: The Anti-Infectious Role of the Vitamin. *The Journal of pediatrics*, *156*(5), 698–703. doi:10.1016/j.jpeds.2010.02.002

Chohan, L., Hollier, L. M., Bishop, K., & Kilpatrick, C. C. (2006). Patterns of antibiotic resistance among group B streptococcus isolates: 2001-2004. *Infectious diseases in obstetrics and gynecology*, *2006*, 57492. doi:10.1155/IDOG/2006/57492

Choi, S., Stickel, F., Baik, H., Kim, Y., Seitz, H., & Mason, J. (1999). Chronic alcohol consumption induces genomic but not p53-specific DNA hypomethylation in rat colon. *The Journal of ...*, *129*(11), 1945–1950. Retrieved from http://jn.nutrition.org/content/129/11/1945.short

Chu, Y. W., Tse, C., Tsang, G. K.-L., So, D. K.-S., Fung, J. T.L., & Lo, J. Y.C. (2007). Invasive group B Streptococcus isolates showing reduced susceptibility to penicillin in Hong Kong. *The Journal of antimicrobial chemotherapy*, *60*(6), 1407–9. doi:10.1093/jac/dkm390

Clark, A. E., Adamian, M., & Taylor, J. Y. (2013). An overview of epigenetics in nursing. *The Nursing Clinics of North America*, *48*(4), 649–59. doi:10.1016/j.cnur.2013.08.004

Cleaver, J. E. (2005). Cancer in xeroderma pigmentosum and related disorders of DNA repair. *Nature Reviews. Cancer*, *5*(7), 564–73. doi:10.1038/nrc1652

Clifford, V., Garland, S. M., & Grimwood, K. (2011). Prevention of neonatal group B streptococcus disease in the 21st century. *Journal of paediatrics and child health*. doi:10.1111/j.1440-1754.2011.02203.x; 10.1111/j.1440-1754.2011.02203.x

Colombo, D. F., Lew, J. L., Pedersen, C. A., Johnson, J. R., & Fan-Havard, P. (2006). Optimal timing of ampicillin administration to pregnant women for establishing bactericidal levels in the prophylaxis of Group B Streptococcus. *American Journal of Obstetrics and Gynecology*, *194*(2), 466–470. doi:10.1016/j.ajog.2005.07.098

Cornacchia, E., Golbus, J., Maybaum, J., Strahler, J., Hanash, S., & Richardson, B. (1988a). Hydralazine and procainamide inhibit T cell DNA methylation and induce autoreactivity. *Journal of immunology (Baltimore, Md. : 1950)*, *140*(7), 2197–200. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3258330

Cornacchia, E., Golbus, J., Maybaum, J., Strahler, J., Hanash, S., & Richardson, B. (1988b). Hydralazine and procainamide inhibit T cell DNA methylation and

induce autoreactivity. *Journal of immunology (Baltimore, Md. : 1950)*, *140*(7), 2197–200.

Consensus Panel on Genetic/Genomic Nursing Competencies (2009). *Essentials of Genetic and Genomic Nursing: Competencies, Curricula Guidelines, and Outcome Indicators, 2nd Edition*. Silver Spring, MD: American Nurses Association.

Curry, a E., Vogel, I., Skogstrand, K., Drews, C., Schendel, D. E., Flanders, W. D., Hougaard, D. M., et al. (2008). Maternal plasma cytokines in early- and mid-gestation of normal human pregnancy and their association with maternal factors. *Journal of reproductive immunology*, *77*(2), 152–60. doi:10.1016/j.jri.2007.06.051

Davis, C., & Uthus, E. (2004). DNA methylation, cancer susceptibility, and nutrient interactions. *Experimental biology and medicine*, *229*(10), 988–995. Retrieved from http://ebm.rsmjournals.com/content/229/10/988.short

Deaton, A. M., & Bird, A. (2011). CpG islands and the regulation of transcription. *Genes & Development*, *25*(10), 1010–22. doi:10.1101/gad.2037511

Dedeurwaerder, S., Defrance, M., Calonne, E., Denis, H., Sotiriou, C., & Fuks, F. (2011). Evaluation of the Infinium Methylation 450K technology. *Epigenomics*, *3*(6), 771–784. doi:10.2217/epi.11.105

Dekel, N., Gnainsky, Y., Granot, I., & Mor, G. (2010). Inflammation and implantation. *American Journal of Reproductive Immunology (New York, N.Y. : 1989)*, *63*(1), 17–21. doi:10.1111/j.1600-0897.2009.00792.x

DeMeo, D. L., & Silverman, E. K. (2004). Alpha1-antitrypsin deficiency. 2: genetic aspects of alpha(1)-antitrypsin deficiency: phenotypes and genetic modifiers of emphysema risk. *Thorax*, *59*(3), 259–64. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1746953&tool=pmcentrez&rendertype=abstract

De-Regil Luz, M., Palacios, C., Ansary, A., Kulier, R., & Peña-Rosas Juan, P. (2012). *Vitamin D supplementation for women during pregnancy*. *Cochrane Database of Systematic Reviews*. John Wiley & Sons, Ltd. doi:10.1002/14651858.CD008873.pub2

De Steenwinkel, F. D. O., Tak, H. V, Muller, A. E., Nouwen, J. L., Oostvogel, P. M., & Mocumbi, S. M. (2008). Low carriage rate of group B streptococcus in pregnant women in Maputo, Mozambique. *Tropical medicine & international health : TM & IH*, *13*(3), 427–9. doi:10.1111/j.1365-3156.2008.02018.x

Denney, J. M., Nelson, E. L., Wadhwa, P. D., Waters, T. P., Mathew, L., Chung, E. K., Goldenberg, R. L., et al. (2011). Longitudinal modulation of immune system cytokine profile during pregnancy. *Cytokine*, *53*(2), 170–177. doi:10.1016/j.cyto.2010.11.005

Deutscher, M., Lewis, M., Zell, E. R., Taylor, T. H., Van Beneden, C., & Schrag, S. (2011). Incidence and severity of invasive Streptococcus pneumoniae, group A Streptococcus, and group B Streptococcus infections among pregnant and postpartum women. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, *53*(2), 114–23. doi:10.1093/cid/cir325

DiPersio, L. P., & DiPersio, J. R. (2006). High rates of erythromycin and clindamycin resistance among OBGYN isolates of group B Streptococcus. *Diagnostic microbiology and infectious disease*, *54*(1), 79–82. doi:10.1016/j.diagmicrobio.2005.07.003

Doi, A., Park, I., Wen, B., Murakami, P., Aryee, M., Irizarry, R., … Feinberg, A. P. (2009). Differential methylation of tissue-and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nature Genetics*, *41*(12), 1350–1353. doi:10.1038/ng.471.Differential

Donato, K. A., Gareau, M. G., Wang, Y. J., & Sherman, P. M. (2010). Lactobacillus rhamnosus GG attenuates interferon-{gamma} and tumour necrosis factor-alpha-induced barrier dysfunction and pro-inflammatory signalling. *Microbiology (Reading, England)*, *156*(Pt 11), 3288–3297. doi:10.1099/mic.0.040139-0

Du, P., Kibbe, W. A., & Lin, S. M. (2008). "lumi: a Bioconductor package for processing Illumina microarray" *Bioinformatics*, *24*(13), 1547-1548.

Du, P., Zhang, X., Huang, C.-C., Jafari, N., Kibbe, W. a, Hou, L., & Lin, S. M. (2010). Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics*, *11*(1), 587. doi:10.1186/1471-2105-11-587

Edmond, K. M., Kortsalioudaki, C., Scott, S., Schrag, S. J., Zaidi, A. K., Cousens, S., & Heath, P. T. (2012). Group B streptococcal disease in infants aged younger than 3 months: systematic review and meta-analysis. *Lancet*. doi:10.1016/S0140-6736(11)61651-6

Edwards, M. S., Rench, M. a, Palazzi, D. L., & Baker, C. J. (2005). Group B streptococcal colonization and serotype-specific immunity in healthy elderly persons. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, *40*(3), 352–7. doi:10.1086/426820

Ehrlich, M., Gama-Sosa, M. A., Huang, L. H., Midgett, R. M., Kuo, K. C., McCune, R. A., & Gehrke, C. (1982). Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic acids research*, *10*(8), 2709–2721.

Ekelund, K., & Konradsen, H. B. (2004). Invasive group B streptococcal disease in infants: a 19-year nationwide study. Serotype distribution, incidence and recurrent infection. *Epidemiology and infection*, *132*(6), 1083–90. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2870199&tool=pmcentrez&rendertype=abstract

El-Nour, H., Lundeberg, L., Al-Tawil, R., Granlund, a, Lonne-Rahm, S.-B., & Nordlind, K. (2006). Upregulation of the axonal growth and the expression of substance P and its NK1 receptor in human allergic contact dermatitis. *Immunopharmacology and Immunotoxicology*, *28*(4), 621–31. doi:10.1080/08923970601067136

Emeritus, G. K. P., & Wickens, T. D. (2004). *Design and Analysis: A Researcher's Handbook (4th Edition)* (p. 624). Pearson. Retrieved from http://www.amazon.com/Design-Analysis-Researchers-Handbook-Edition/dp/0135159415

epi. (n.d.). *Collins English Dictionary - Complete & Unabridged 10th Edition*. Retrieved March 30, 2014, from Dictionary.com website:http://dictionary.reference.com/browse/epi

Fahey, J. L., Taylor, J. M., Detels, R., Hofmann, B., Melmed, R., Nishanian, P., & Giorgi, J. V. (1990). The prognostic value of cellular and serologic markers in infection with human immunodeficiency virus type 1. *The New England journal of medicine*, *322*(3), 166–172. doi:10.1056/NEJM199001183220305

Fan, H., Teti, G., Ashton, S., Guyton, K., Tempel, G. E., Halushka, P. V, & Cook, J. A. (2003). Involvement of G(i) proteins and Src tyrosine kinase in TNFalpha production induced by lipopolysaccharide, group B Streptococci and Staphylococcus aureus. *Cytokine*, *22*(5), 126–133.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–60. doi:10.3758/BRM.41.4.1149

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–91. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17695343

Feinberg, A. P. (2008). Epigenetics at the epicenter of modern medicine. *JAMA : the journal of the American Medical Association*, *299*(11), 1345–50. doi:10.1001/jama.299.11.1345

Ferrieri, P., Baker, C. J., Hillier, S. L., & Flores, a E. (2004). Diversity of surface protein expression in group B streptococcal colonizing & invasive isolates. *The Indian*

journal of medical research, *119 Suppl*(May), 191–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15232193

Fichorova, R. N., Onderdonk, A. B., Yamamoto, H., Delaney, M. L., DuBois, A. M., Allred, E., Leviton, A., et al. (2011). Maternal microbe-specific modulation of inflammatory response in extremely low-gestational-age newborns. *mBio*, *2*(1), e00280–10. doi:10.1128/mBio.00280-10

Fitzpatrick, D. R., Shirley, K. M., & Kelso, A. (1999). Cutting edge: stable epigenetic inheritance of regional IFN-gamma promoter demethylation in CD44highCD8+ T lymphocytes. *Journal of immunology (Baltimore, Md.: 1950)*, *162*(9), 5053–5057.

Foxman, B., Gillespie, B., Manning, S. D., Howard, L. J., Tallman, P., Zhang, L., & Marrs, C. F. (2006). Incidence and duration of group B Streptococcus by serotype among male and female college students living in a single dormitory. *American journal of epidemiology*, *163*(6), 544–51. doi:10.1093/aje/kwj075

Fraga, M. F., Ballestar, E., Paz, M. F., Ropero, S., Setien, F., Ballestar, M. L., … Esteller, M. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(30), 10604–9. doi:10.1073/pnas.0500398102

Garland, S. M., Cottrill, E., Markowski, L., Pearce, C., Clifford, V., Ndisang, D., Kelly, N., et al. (2011). Antimicrobial resistance in group B streptococcus: the Australian experience. *Journal of medical microbiology*, *60*(Pt 2), 230–235. doi:10.1099/jmm.0.022616-0

Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., … Zhang, J. (2004). Bioconductor: open software development for computational

biology and bioinformatics. *Genome Biology*, *5*(10), R80. doi:10.1186/gb-2004-5-10-r80

Gómez, L. M., Sammel, M. D., Appleby, D. H., Elovitz, M. a, Baldwin, D. a, Jeffcoat, M. K., … Parry, S. (2010). Evidence of a gene-environment interaction that predisposes to spontaneous preterm birth: a role for asymptomatic bacterial vaginosis and DNA variants in genes that control the inflammatory response. *American Journal of Obstetrics and Gynecology*, *202*(4), 386.e1–6. doi:10.1016/j.ajog.2010.01.042

Gould, I. M. (2010). Alexander Gordon, puerperal sepsis, and modern theories of infection control--Semmelweis in perspective. *The Lancet infectious diseases*, *10*(4), 275–8. doi:10.1016/S1473-3099(09)70304-4

Gray, K. J., Bennett, S. L., French, N., Phiri, A. J., & Graham, S. M. (2007). Invasice Group B Streptococcal Infection in Infants , Malawi. *Emerging infectious diseases*, *13*(2), 223–229. Retrieved from http://wwwnc.cdc.gov/eid/article/13/2/pdfs/06-0680.pdf

Grimwood, K., Stone, P. R., Gosling, I. a, Green, R., Darlow, B. a, Lennon, D. R., & Martin, D. R. (2002). Late antenatal carriage of group B Streptococcus by New Zealand women. *The Australian & New Zealand journal of obstetrics & gynaecology*, *42*(2), 182–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12069147

Guerrero-Bosagna, C., & Skinner, M. K. (2012). Environmentally induced epigenetic transgenerational inheritance of phenotype and disease. *Molecular and Cellular Endocrinology*, *354*(1-2), 3–8. doi:10.1016/j.mce.2011.10.004

Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., … Surani, M. A. (2002). Epigenetic reprogramming in mouse primordial germ cells. *Mechanisms of development*, *117*(1-2), 15–23.

Hansen, K.D., Ayree, M., and Irizarry, R. A. (2013). Minfi: Analyze Illumina's 450k methylation arrays. R package version 1.8.9.

Hashii, K., Fujiwara, H., Yoshioka, S., Kataoka, N., Yamada, S., Hirano, T., Mori, T., et al. (1998). Peripheral blood mononuclear cells stimulate progesterone production by luteal cells derived from pregnant and non-pregnant women: possible involvement of interleukin-4 and interleukin-10 in corpus luteum function and differentiation. *Human reproduction (Oxford, England)*, *13*(1O), 2738–2744.

Heath, P T. (2011). An update on vaccination against group B streptococcus. *Expert review of vaccines*, *10*(5), 685–694. doi:10.1586/erv.11.61

Heath, Paul T, Balfour, G., Weisner, A. M., Efstratiou, A., Lamagni, T. L., Tighe, H., O'Connell, L. a F., et al. (2004). Group B streptococcal disease in UK and Irish infants younger than 90 days. *Lancet*, *363*(9405), 292–4. doi:10.1016/S0140-6736(03)15389-5

He, X.-J., Chen, T., & Zhu, J.-K. (2011). Regulation and function of DNA methylation in plants and animals. *Cell Research*, *21*(3), 442–65. doi:10.1038/cr.2011.23

Holick, M. F. (2011). Vitamin D: evolutionary, physiological and health perspectives. *Current Drug Targets*, *12*(1), 4–18. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20795941

Hollis, B. (2005). Circulating 25-hydroxyvitamin D levels indicative of vitamin D sufficiency: implications for establishing a new effective dietary intake

recommendation for vitamin D. *The Journal of Nutrition*, *135*(2), 317–322. Retrieved from http://jn.nutrition.org/content/135/2/317.short

Hollis, B. W. (2008). Assessment of vitamin D status and definition of a normal circulating range of 25-hydroxyvitamin D. *Current Opinion in Endocrinology, Diabetes, and Obesity*, *15*(6), 489–94. doi:10.1097/MED.0b013e328317ca6c

Hollis, B. W. (2012). Assessment and interpretation of circulating 25-hydroxyvitamin D and 1,25-dihydroxyvitamin D in the clinical environment. *Rheumatic Diseases Clinics of North America*, *38*(1), 29–44. doi:10.1016/j.rdc.2012.03.005

Hopkins, M. H., Owen, J., Ahearn, T., Fedirko, V., Flanders, W. D., Jones, D. P., & Bostick, R. M. (2011). Effects of supplemental vitamin D and calcium on biomarkers of inflammation in colorectal adenoma patients: a randomized, controlled clinical trial. *Cancer prevention research (Philadelphia, Pa.)*, *4*(10), 1645–1654. doi:10.1158/1940-6207.CAPR-11-0105

Huang, D. W., Sherman, B. T., & Lempicki, R. a. (2009a). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, *4*(1), 44–57. doi:10.1038/nprot.2008.211

Huang D. W., Sherman B.T., & Lempicki R.A. (2009b). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research. 37*(1):1-13.

Hunter, D. J. (2005). Gene-environment interactions in human diseases. *Nature reviews.Genetics*, *6*(4), 287–298. doi:10.1038/nrg1578

Hyppönen, E., Turner, S., Cumberland, P., Power, C., & Gibb, I. (2007). Serum 25-hydroxyvitamin D measurement in a large population survey with statistical

harmonization of assay variation to an international standard. *The Journal of Clinical Endocrinology and Metabolism*, *92*(12), 4615–22. doi:10.1210/jc.2007-1279

Ippolito, D. L., James, W. a, Tinnemore, D., Huang, R. R., Dehart, M. J., Williams, J., Wingerd, M. a, et al. (2010). Group B streptococcus serotype prevalence in reproductive-age women at a tertiary care military medical center relative to global serotype distribution. *BMC infectious diseases*, *10*(1), 336. doi:10.1186/1471-2334-10-336

Jamaly, S., Khanehkenari, M. R., Rao, R., Patil, G., Thakur, S., Ramaswamy, P., … Sahoo, R. (2012). Relationship between p53 overexpression, human papillomavirus infection, and lifestyle in Indian patients with head and neck cancers. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine*, *33*(2), 543–550. doi:10.1007/s13277-011-0295-x

Janapatla, R. P., Ho, Y. R., Yan, J. J., Wu, H. M., & Wu, J. J. (2008). The prevalence of erythromycin resistance in group B streptococcal isolates at a University Hospital in Taiwan. *Microbial drug resistance (Larchmont, N.Y.)*, *14*(4), 293–297. doi:10.1089/mdr.2008.0853

Jiao, X., Sherman, B. T., Huang da, W., Stephens, R., Baseler, M. W., Lane, H. C., & Lempicki, R. A. (2012). DAVID-WS: a stateful web service to facilitate gene/protein list analysis. *Bioinformatics (Oxford, England)*, *28*(13), 1805–1806. doi:10.1093/bioinformatics/bts251

Johri, A. K., Paoletti, L. C., Glaser, P., Dua, M., Sharma, P. K., Grandi, G., & Rappuoli, R. (2006). Group B Streptococcus: global incidence and vaccine development. *Nature reviews.Microbiology*, *4*(12), 932–942. doi:10.1038/nrmicro1552

Jones, N., Oliver, K., Jones, Y., Haines, a, & Crook, D. (2006). Carriage of group B streptococcus in pregnant women from Oxford, UK. *Journal of clinical pathology*, *59*(4), 363–6. doi:10.1136/jcp.2005.029058

Jones, P. A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Reviews. Genetics*, *13*(7), 484–92. doi:10.1038/nrg3230

Jordan, H. T., Farley, M. M., Craig, A., Mohle-Boetani, J., Harrison, L. H., Petit, S., … Active Bacterial Core Surveillance (ABCs)/Emerging Infections Program Network, C. D. C. (2008). Revisiting the need for vaccine prevention of late-onset neonatal group B streptococcal disease: a multistate, population-based analysis. *The Pediatric infectious disease journal*, *27*(12), 1057–1064. doi:10.1097/INF.0b013e318180b3b9

Kaminsky, Z. A., Tang, T., Wang, S.-C., Ptak, C., Oh, G. H. T., Wong, A. H. C., … Petronis, A. (2009). DNA methylation profiles in monozygotic and dizygotic twins. *Nature Genetics*, *41*(2), 240–5. doi:10.1038/ng.286

Kanchi, K. L., Johnson, K. J., Lu, C., McLellan, M. D., Leiserson, M. D. M., Wendl, M. C., … Ding, L. (2014). Integrated analysis of germline and somatic variants in ovarian cancer. *Nature Communications*, *5*, 3156. doi:10.1038/ncomms4156

Kaposy, C., & Lafferty, L. (2012). Overcoming liability concerns in vaccine trials involving pregnant women. *Accountability in research*, *19*(3), 156–74. doi:10.1080/08989621.2012.678686

Kargul, J., & Laurent, G. J. (2009). Epigenetics and human disease. *The international journal of biochemistry & cell biology*, *41*(1), 1. doi:10.1016/j.biocel.2008.09.017

Kent W. J., Sugnet C. W., Furey T. S., Roskin K. M., Pringle T. H., Zahler A. M., Haussler D. (2002).The human genome browser at UCSC. *Genome Research, 12*(6):996-1006.

Kiecolt-Glaser, J. K., & Glaser, R. (2002). Depression and immune function: central pathways to morbidity and mortality. *Journal of psychosomatic research*, *53*(4), 873–876.

Kim, E. J., Oh, K. Y., Kim, M. Y., Seo, Y. S., Shin, J.-H., Song, Y. R., Yang, J.-H., et al. (2011). Risk factors for group B streptococcus colonization among pregnant women in Korea. *Epidemiology and health*, *33*, e2011010. doi:10.4178/epih/e2011010

King, M.-C. (2003). Breast and Ovarian Cancer Risks Due to Inherited Mutationsin BRCA1 and BRCA2. *Science. 10/24/2003*, *302*(5645), 643–646.

Kirk, M., Calzone, K., Arimori, N., & Tonkin, E. (2011).Genetics-genomics competencies and nursing regulation. *Journal of Nursing Scholarship, 43(2), 107-116. doi:10.1111/j.1547-5069.2011.01388.x*

Kirwan, J. P., Mouzon, S. H., Lepercq, J., Challier, J., Huston-presley, L., Friedman, J. E., Kalhan, S. C., et al. (2002). TNF-alpha is a predictor of insulin resistance in human pregnancy. *Diabetes*, *51*(7), 2207–2213.

Kong, F., Gowan, S., Martin, D., James, G., & Gilbert, G. L. (2002). Serotype Identification of Group B Streptococci by PCR and Sequencing. *Journal of clinical microbiology*, *40*(1), 216–226. doi:10.1128/JCM.40.1.216

Kovavisarach, E., Ying, W. S., & Kanjanahareutai, S. (2007). Risk factors related to group B streptococcal colonization in pregnant women in labor. *Journal of the Medical Association of Thailand = Chotmaihet thangphaet*, *90*(7), 1287–1292.

Kraemer, K. H., Lee, M. M., Andrews, A. D., & Lambert, W. C. (1994). The role of sunlight and DNA repair in melanoma and nonmelanoma skin cancer. The xeroderma pigmentosum paradigm. *Archives of Dermatology*, *130*(8), 1018–21. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8053698

Laayouni, H., Oosting, M., Luisi, P., Ioana, M., Alonso, S., Ricaño-Ponce, I., … Netea, M. G. (2014). Convergent evolution in European and Rroma populations reveals pressure exerted by plague on Toll-like receptors. *Proceedings of the National Academy of Sciences* . doi:10.1073/pnas.1317723111

Lancefield, R C, & Freimer, E. H. (1966). Type-specific polysaccharide antigens of group B streptococci. *The Journal of hygiene*, *64*(2), 191–203. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2134722&tool=pmcentrez&rendertype=abstract

Lancefield, Rebecca C. (1933). A Serological Differentiation of Human and Other Groups of Hemolytic Streptococci. *The Journal of Experimental Medicine*, *57*(4), 571–595.

Latham, K. E., Sapienza, C., & Engel, N. (2012). The epigenetic lorax: gene-environment interactions in human health. *Epigenomics*, *4*(4), 383–402. doi:10.2217/epi.12.31

Laufer, J., Scasso, S., Sosa, C. G., Rodríguez-Cuns, G., Alonso, J., & Pons, J. E. (2009). Group B streptococcus colonization among pregnant women in Uruguay. *International journal of gynaecology and obstetrics: the official organ of the*

International Federation of Gynaecology and Obstetrics, *104*(3), 242–3. doi:10.1016/j.ijgo.2008.10.020

Li L. C., & Dahiya R. (2002) MethPrimer: designing primers for methylation PCRs. *Bioinformatics*.*18*(11):1427-31.

Lin, F. Y., Weisman, L. E., Azimi, P., Young, A. E., Chang, K., Cielo, M., … Robbins, J. B. (2011). Assessment of intrapartum antibiotic prophylaxis for the prevention of early-onset group B Streptococcal disease. *The Pediatric infectious disease journal*, *30*(9), 759–763. doi:10.1097/INF.0b013e31821dc76f

Lindahl, G., Stalhammer-Carlemalm, M., & Areschoug, T. (2005). Surface Proteins of Streptococcus agalactiae and Related Proteins in Other Bacterial Pathogens. *Clinical microbiology reviews*, *18*(1), 102–127. doi:10.1128/CMR.18.1.102

Liu, F., van Duijn, K., Vingerling, J. R., Hofman, A., Uitterlinden, A. G., Janssens, A. C. J. W., & Kayser, M. (2009). Eye color and the prediction of complex phenotypes from genotypes. *Current Biology*, *19*(5), R192–R193. doi:10.1016/j.cub.2009.01.027

Li, Y., Gorelik, G., Strickland, F. M., & Richardson, B. C. (2014). Oxidative Stress, T Cell DNA Methylation and Lupus. *Arthritis & Rheumatology (Hoboken, N.J.)*. doi:10.1002/art.38427

Loescher, L. J., & Merkle, C. J. (2005). Focus on Genomics The Interface of Genomic Technologies and Nursing, (c), 111–119.

Lund, G., Andersson, L., Lauria, M., Lindholm, M., Fraga, M. F., Villar-Garea, A., … Zaina, S. (2004). DNA methylation polymorphisms precede any histological sign

of atherosclerosis in mice lacking apolipoprotein E. *The Journal of Biological Chemistry*, *279*(28), 29147–54. doi:10.1074/jbc.M403618200

Madureira, P, Andrade, E. B., Gama, B., Oliveira, L., Moreira, S., Ribeiro, A., … Ferreira, P. (2011). Inhibition of IL-10 production by maternal antibodies against Group B Streptococcus GAPDH confers immunity to offspring by favoring neutrophil recruitment. *PLoS pathogens*, *7*(11), e1002363. doi:10.1371/journal.ppat.1002363

Maisey, H. C., Doran, K. S., & Nizet, V. (2008). Recent advances in understanding the molecular basis of group B Streptococcus virulence. *Expert reviews in molecular medicine*, *10*, e27. doi:10.1017/S1462399408000811

Makhseed, M., Raghupathy, R., Azizieh, F., Farhat, R., Hassan, N., & Bandar, a. (2000). Circulating cytokines and CD30 in normal human pregnancy and recurrent spontaneous abortions. *Human Reproduction (Oxford, England)*, *15*(9), 2011–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10967006

Manning, S. D., Lewis, M. a, Springman, a C., Lehotzky, E., Whittam, T. S., & Davies, H. D. (2008). Genotypic diversity and serotype distribution of group B streptococcus isolated from women before and after delivery. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, *46*(12), 1829–37. doi:10.1086/588296

Manning, S. D., Neighbors, K., Tallman, P. a, Gillespie, B., Marrs, C. F., Borchardt, S. M., Baker, C. J., et al. (2004). Prevalence of group B streptococcus colonization and potential for transmission by casual contact in healthy young men and

women. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, *39*(3), 380–8. doi:10.1086/422321

Meyn, L. a, Krohn, M. a, & Hillier, S. L. (2009). Rectal colonization by group B Streptococcus as a predictor of vaginal colonization. *American journal of obstetrics and gynecology*, *201*(1), 76.e1–7. doi:10.1016/j.ajog.2009.02.011

Meyn, L. a, Moore, D. M., Hillier, S. L., & Krohn, M. a. (2002). Association of sexual activity with colonization and vaginal acquisition of group B Streptococcus in nonpregnant women. *American journal of epidemiology*, *155*(10), 949–57. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11994235

Michels, K. B., Binder, A. M., Dedeurwaerder, S., Epstein, C. B., Greally, J. M., Gut, I., … Irizarry, R. a. (2013). Recommendations for the design and analysis of epigenome-wide association studies. *Nature Methods*, *10*(10), 949–55. doi:10.1038/nmeth.2632

Mikamo, H., Johri, A. K., Paoletti, L. C., Madoff, L. C., & Onderdonk, A. B. (2004). Adherence to, invasion by, and cytokine production in response to serotype VIII group B Streptococci. *Infection and immunity*, *72*(8), 4716–4722. doi:10.1128/IAI.72.8.4716-4722.2004

Mikovits, J. a, Young, H. a, Vertino, P., Issa, J. P., Pitha, P. M., Turcoski-Corrales, S., Taub, D. D., et al. (1998). Infection with human immunodeficiency virus type 1 upregulates DNA methyltransferase, resulting in de novo methylation of the gamma interferon (IFN-gamma) promoter and subsequent downregulation of IFN-gamma production. *Molecular and cellular biology*, *18*(9), 5166–77.

Miller, C. a, & Sweatt, J. D. (2007). Covalent modification of DNA regulates memory formation. *Neuron*, *53*(6), 857–69. doi:10.1016/j.neuron.2007.02.022

Miyajima, A., Hara, T., & Kitamura, T. (1992). Common subunits of cytokine receptors and the functional redundancy of cytokines. *Trends in Biochemical Sciences*, *17*(10), 378–382. doi:10.1016/0968-0004(92)90004-S

Moffitt, T. E., Caspi, A., & Rutter, M. (2006). Measured Gene-Environment Interactions in Psychopathology: Concepts, Research Strategies, and Implications for Research, Intervention, and Public Understanding of Genetics. *Perspectives on Psychological Science*, *1*(1), 5–27. Retrieved from http://www.jstor.org.ezproxy.undmedlibrary.org/stable/40184195

Mor, G., & Cardenas, I. (2010). The immune system in pregnancy: a unique complexity. *American journal of reproductive immunology (New York, N.Y. : 1989)*, *63*(6), 425–33. doi:10.1111/j.1600-0897.2010.00836.x

Moyo, S. R., Mudzori, J., Tswana, S. A., & Maeland, J. A. (2000). Prevalence, capsular type distribution, anthropometric and obstetric factors of group B Streptococcus (Streptococcus agalactiae) colonization in pregnancy. *The Central African journal of medicine*, *46*(5), 115–20. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11210331

Ng, P. C., Li, K., Wong, R. P., Chui, K., Wong, E., Li, G., & Fok, T. F. (2003). Proinflammatory and anti-inflammatory cytokine responses in preterm infants with systemic infections. *Archives of disease in childhood.Fetal and neonatal edition*, *88*(3), F209–13.

Nizet, V., Gibson, R. L., Chi, E. Y., Framson, P. E., Hulse, M., & Rubens, C. E. (1996). Group B streptococcal beta-hemolysin expression is associated with injury of lung epithelial cells. *Infection and immunity*, *64*(9), 3818–26. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=174298&tool=pmcentrez&rendertype=abstract

Olszak, T., An, D., Zeissig, S., Vera, M. P., Richter, J., Franke, A., Glickman, J. N., et al. (2012). Microbial exposure during early life has persistent effects on natural killer T cell function. *Science (New York, N.Y.)*, *336*(6080), 489–493. doi:10.1126/science.1219328; 10.1126/science.1219328

Ottman, R. (1996). Gene–Environment Interaction: Definitions and Study Design. *Preventive medicine*, *25*(6), 764–770. doi:10.1006/pmed.1996.0117

Othman, M., Neilson, J. P., & Alfirevic, Z. (2007). Probiotics for preventing preterm labour. *Cochrane database of systematic reviews (Online)*, *(1)*(1), CD005941. doi:10.1002/14651858.CD005941.pub2

Pacifici, G. M. (2006). Placental transfer of antibiotics administered to the mother: a review. *International journal of clinical pharmacology and therapeutics*, *44*(2), 57–63.

Paradiso, P. R. (2001). Maternal immunization: the influence of liability issues on vaccine development. *Vaccine*, *20 Suppl 1*, S73–4. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11587817

Parameswaran, N., & Patial, S. (2010). Tumor necrosis factor-alpha signaling in macrophages. *Critical reviews in eukaryotic gene expression*, *20*(2), 87–103.

Petronis, A. (2010). Epigenetics as a unifying principle in the aetiology of complex traits and diseases. *Nature*, *465*(7299), 721–7. doi:10.1038/nature09230

Petronis, A. (2001). Human morbid genetics revisited: relevance of epigenetics. *Trends in Genetics : TIG*, *17*(3), 142–146.

Petronis, A., Gottesman, I. I., Kan, P., Kennedy, J. L., Basile, V. S., Paterson, A. D., & Popendikyte, V. (2003). Monozygotic twins exhibit numerous epigenetic differences: clues to twin discordance? *Schizophrenia Bulletin*, *29*(1), 169–78. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12908672

Pezzilli, R., Billi, P., Miniero, R., Fiocchi, M., Cappelletti, O., Morselli-Labate, A. M., Barakat, B., et al. (1995). Serum interleukin-6, interleukin-8, and beta 2-microglobulin in early assessment of severity of acute pancreatitis. Comparison with serum C-reactive protein. *Digestive diseases and sciences*, *40*(11), 2341–2348.

Phares, C. R., Lynfield, R., Farley, M. M., Mohle-Boetani, J., Harrison, L. H., Petit, S., … Network, A. B. C. surveillance/Emerging I. P. (2008). Epidemiology of invasive group B streptococcal disease in the United States, 1999-2005. *JAMA : the journal of the American Medical Association*, *299*(17), 2056–2065. doi:10.1001/jama.299.17.2056

Philipson, A., Sabath, L. D., & Charles, D. (1973). Transplacental passage of erythromycin and clindamycin. *The New England journal of medicine*, *288*(23), 1219–1221. doi:10.1056/NEJM197306072882307

Pintard, L., Willems, A., & Peter, M. (2004). Cullin-based ubiquitin ligases: Cul3-BTB complexes join the family. *The EMBO Journal*, *23*(8), 1681–7. doi:10.1038/sj.emboj.7600186

Ponsonby, A. L., Lucas, R. M., Lewis, S., & Halliday, J. (2010). Vitamin D status during pregnancy and aspects of offspring health. *Nutrients*, *2*(3), 389–407. doi:10.3390/nu2030389

Puliti, M., Nizet, V., Von Hunolstein, C., Bistoni, F., Mosci, P., Orefici, G., & Tissi, L. (2000). Severity of group B streptococcal arthritis is correlated with beta-hemolysin expression. *The Journal of infectious diseases*, *182*(3), 824–32. doi:10.1086/315773

Puliti, M., Von Hunolstein, C., Verwaerde, C., Bistoni, F., Orefici, G., & Tissi, L. (2002). Regulatory role of interleukin-10 in experimental group B streptococcal arthritis. *Infection and immunity*, *70*(6), 2862–2868.

Rausch, A.-V., Gross, A., Droz, S., Bodmer, T., & Surbek, D. V. (2009). Group B Streptococcus colonization in pregnancy: prevalence and prevention strategies of neonatal sepsis. *Journal of perinatal medicine*, *37*(2), 124–9. doi:10.1515/JPM.2009.020

Rakyan, V. K., Down, T. A., Balding, D. J., & Beck, S. (2011). Epigenome-wide association studies for common human diseases. *Nature Reviews. Genetics*, *12*(8), 529–41. doi:10.1038/nrg3000

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/

Reed, J. a, Finnerty, B., & Albino, a P. (1999). Divergent cellular differentiation pathways during the invasive stage of cutaneous malignant melanoma progression. *The American Journal of Pathology*, *155*(2), 549–55. doi:10.1016/S0002-9440(10)65150-4

Rees, J. L. (2004). The genetics of sun sensitivity in humans. *American Journal of Human Genetics*, *75*(5), 739–751. doi:10.1086/425285

Regan, J. a, Klebanoff, M. a, Nugent, R. P., Eschenbach, D. a, Blackwelder, W. C., Lou, Y., Gibbs, R. S., et al. (1996). Colonization with group B streptococci in pregnancy and adverse outcome. VIP Study Group. *American journal of obstetrics and gynecology*, *174*(4), 1354–60. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8623869

Ring, A., Braun, J. S., Pohl, J., Nizet, V., Stremmel, W., & Shenep, J. L. (2002). Group B streptococcal beta-hemolysin induces mortality and liver injury in experimental sepsis. *The Journal of infectious diseases*, *185*(12), 1745–53. doi:10.1086/340818

Ritchie, M. E., Diyagama, D., Neilson, J., van Laar, R., Dobrovic, A., Holloway, A., & Smyth, G. K. (2006). Empirical array quality weights in the analysis of microarray data. *BMC Bioinformatics*, *7*, 261. doi:10.1186/1471-2105-7-261

Rodenhiser, D., & Mann, M. (2006). Epigenetics and human disease: translating basic biology into clinical applications. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne*, *174*(3), 341–348. doi:10.1503/cmaj.050774

Rodriguez-Paredes, M., & Esteller, M. (2011). Cancer epigenetics reaches mainstream oncology. *Nature medicine*, *17*(3), 330–339. doi:10.1038/nm.2305

Romero, R., Espinoza, J., Gonçalves, L. F., Kusanovic, J. P., Friel, L. A., & Nien, J. K. (2006). Inflammation in preterm and term labour and delivery. *Seminars in Fetal & Neonatal Medicine*, *11*(5), 317–26. doi:10.1016/j.siny.2006.05.001

Rothstein, M., Cai, Y., & Marchant, G. (2009). The ghost in our genes: legal and ethical implications of epigenetics. *Health Matrix*, *19*(1), 1–43. Retrieved from http://works.bepress.com/context/gary_marchant/article/1001/type/native/viewcontent/

Santhanam, U., Avila, C., Romero, R., Viguet, H., Ida, N., Sakurai, S., & Sehgal, P. B. (1991). Cytokines in normal and abnormal parturition: elevated amniotic fluid interleukin-6 levels in women with premature rupture of membranes associated with intrauterine infection. *Cytokine*, *3*(2), 155–163.

Saxonov, S., Berg, P., & Brutlag, D. L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(5), 1412–7. doi:10.1073/pnas.0510310103

Schaub, B., Liu, J., Höppler, S., Schleich, I., Huehn, J., Olek, S., Wieczorek, G., et al. (2009). Maternal farm exposure modulates neonatal immune mechanisms through regulatory T cells. *The Journal of allergy and clinical immunology*, *123*(4), 774–82.e5. doi:10.1016/j.jaci.2009.01.056

Schrag, S. J., & Verani, J. R. (2013). Intrapartum antibiotic prophylaxis for the prevention of perinatal group B streptococcal disease: experience in the United States and implications for a potential group B streptococcal vaccine. *Vaccine*, *31 Suppl 4*, D20–6. doi:10.1016/j.vaccine.2012.11.056

Schrag, S. J., Zell, E. R., Lynfield, R., Roome, A., Arnold, K. E., Craig, A. S., … Schuchat, A. (2002). A population-based comparison of strategies to prevent early-onset group B streptococcal disease in neonates. *The New England journal of medicine*, *347*(4), 233–9. doi:10.1056/NEJMoa020205

Seoud, M., Nassar, A. H., Zalloua, P., Boghossian, N., Ezeddine, J., Fakhoury, H., Abboud, J., et al. (2010). Prenatal and neonatal Group B Streptococcus screening and serotyping in Lebanon: incidence and implications. *Acta obstetricia et gynecologica Scandinavica*, *89*(3), 399–403. doi:10.3109/00016340903560008

Shab-Bidar, S., Neyestani, T. R., Djazayery, A., Eshraghian, M. R., Houshiarrad, A., Kalayi, A., … Gharavi, A. (2012). Improvement of vitamin D status resulted in amelioration of biomarkers of systemic inflammation in the subjects with type 2 diabetes. *Diabetes/metabolism research and reviews*, *28*(5), 424–430. doi:10.1002/dmrr.2290; 10.1002/dmrr.2290

Sharma, A., Satyam, A., & Sharma, J. B. (2007). Leptin, IL-10 and inflammatory markers (TNF-alpha, IL-6 and IL-8) in pre-eclamptic, normotensive pregnant and healthy non-pregnant women. *American journal of reproductive immunology (New York, N.Y.: 1989)*, *58*(1), 21–30. doi:10.1111/j.1600-0897.2007.00486.x

Shimada, S., Nishida, R., Takeda, M., Iwabuchi, K., Kishi, R., Onoé, K., … Yamada, H. (2006). Natural killer, natural killer T, helper and cytotoxic T cells in the decidua from sporadic miscarriage. *American Journal of Reproductive Immunology (New York, N.Y. : 1989)*, *56*(3), 193–200. doi:10.1111/j.1600-0897.2006.00417.x

Singer, J. D., Gurian-West, M., Clurman, B., & Roberts, J. M. (1999). Cullin-3 targets cyclin E for ubiquitination and controls S phase in mammalian cells. *Genes & Development*, *13*(18), 2375–87.

Skoff, T. H., Farley, M. M., Petit, S., Craig, A. S., Schaffner, W., Gershman, K., Harrison, L. H., et al. (2009). Increasing burden of invasive group B streptococcal disease in nonpregnant adults, 1990-2007. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, *49*(1), 85–92. doi:10.1086/599369

Slotved, H.-C., Kong, F., Lambertsen, L., Sauer, S., & Gilbert, G. L. (2007). Serotype IX, a Proposed New Streptococcus agalactiae Serotype. *Journal of clinical microbiology*, *45*(9), 2929–36. doi:10.1128/JCM.00117-07

Smith, M. W., Dean, M., Carrington, M., Winkler, C., Huttley, G. A., Lomb, D. A., … O'Brien, S. J. (1997). Contrasting genetic influence of CCR2 and CCR5 variants on HIV-1 infection and disease progression. Hemophilia Growth and Development Study (HGDS), Multicenter AIDS Cohort Study (MACS), Multicenter Hemophilia Cohort Study (MHCS), San Francisco City Cohort. *Science (New York, N.Y.)*, *277*(5328), 959–965.

Smyth, G. K., Yang, Y. H., & Speed, T. (2003). Statistical issues in cDNA microarray data analysis. *Methods in Molecular Biology (Clifton, N.J.)*, *224*, 111–36. doi:10.1385/1-59259-364-X:111

Smyth, G. K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology*, *3*(1), Article3. doi:10.2202/1544-6115.1027

Smyth, GK (2005). Limma: linear models for microarray data. In: 'Bioinformatics and Computational Biology Solutions using R and Bioconductor'. R. Gentleman, V. Carey, S. Dudoit, R. Irizarry, W. Huber (eds), Springer, New York, pages 397-420.

Solensky, R. (2003). Hypersensitivity reactions to beta-lactam antibiotics. *Clinical reviews in allergy & immunology*, *24*(3), 201–220. doi:10.1385/CRIAI:24:3:201

Spellerberg, B. (2000). Pathogenesis of neonatal Streptococcus agalactiae infections. *Microbes and infection / Institut Pasteur*, *2*(14), 1733–42. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11137046

Stoll, B. J., Hansen, N. I., Sanchez, P. J., Faix, R. G., Poindexter, B. B., Van Meurs, K. P., … Network, E. K. S. N. I. of C. H. and H. D. N. R. (2011). Early onset neonatal sepsis: the burden of group B Streptococcal and E. coli disease continues. *Pediatrics*, *127*(5), 817–826. doi:10.1542/peds.2010-2217

Storey, J. (2003). The positive false discovery rate: A Bayesian interpretation and the q-value. *Annals of Statistics*, *31*(6), 2013–2035. Retrieved from http://www.jstor.org/stable/3448445

Tamura, G. S., Hull, J. R., Oberg, M. D., & Castner, D. G. (2006). High-affinity interaction between fibronectin and the group B streptococcal C5a peptidase is unaffected by a naturally occurring four-amino-acid deletion that eliminates peptidase activity. *Infection and immunity*, *74*(10), 5739–46. doi:10.1128/IAI.00241-06

Thellin, O., & Heinen, E. (2003). Pregnancy and the immune system: between tolerance and rejection. *Toxicology*, *185*(3), 179–84.

Thorne-Lyman, A., & Fawzi, W. W. (2012). Vitamin D During Pregnancy and Maternal, Neonatal and Infant Health Outcomes: A Systematic Review and Meta-analysis. *Paediatric and perinatal epidemiology*, *26*, 75–90. doi:10.1111/j.1365-3016.2012.01283.x

Tolg, C., Sabha, N., Cortese, R., Panchal, T., Ahsan, A., Soliman, A., Aitken, K. J., et al. (2011). Uropathogenic E. coli infection provokes epigenetic downregulation of CDKN2A (p16INK4A) in uroepithelial cells. *Laboratory investigation; a journal of technical methods and pathology*, *91*(6), 825–36. doi:10.1038/labinvest.2010.197

Tonkin, E., Calzone, K., Jenkins, J., Lea, D., & Prows, C. (2011). Genomic Education Resources for Nursing Faculty. *Journal of Nursing Scholarship*, *43*(4), 330–340. doi:10.1111/j.1547-5069.2011.01415.x

Tost, J. (2010). DNA methylation: an introduction to the biology and the disease-associated changes of a promising biomarker. *Molecular Biotechnology*, *44*(1), 71–81. doi:10.1007/s12033-009-9216-2

Towers, C. V, Rumney, P. J., Asrat, T., Preslicka, C., Ghamsary, M. G., & Nageotte, M. P. (2010). The accuracy of late third-trimester antenatal screening for group B streptococcus in predicting colonization at delivery. *American Journal of Perinatology*, *27*(10), 785–790. doi:10.1055/s-0030-1254237

Towers, C. V, Rumney, P. J., Asrat, T., Preslicka, C., Ghamsary, M. G., & Nageotte, M. P. (2010). The accuracy of late third-trimester antenatal screening for group B streptococcus in predicting colonization at delivery. *American Journal of Perinatology*, *27*(10), 785–790. doi:10.1055/s-0030-1254237

Toyooka, S., Maruyama, R., Toyooka, K. O., McLerran, D., Feng, Z., Fukuyama, Y., Virmani, A. K., et al. (2003). Smoke exposure, histologic type and geography-related differences in the methylation profiles of non-small cell lung cancer. *International journal of cancer. Journal international du cancer*, *103*(2), 153–60. doi:10.1002/ijc.10787

Trijbels-Smeulders, M., De Jonge, G. a, Pasker-de Jong, P. C. M., Gerards, L. J., Adriaanse, a H., Van Lingen, R. a, & Kollee, L. a a. (2007). Epidemiology of neonatal group B streptococcal disease in the Netherlands before and after introduction of guidelines for prevention. *Archives of Disease in Childhood - Fetal and Neonatal Edition*, *92*(4), F271–F276. doi:10.1136/adc.2005.088799

Trotman, H., & Bell, Y. (2006). Neonatal group B streptococcal infection at the University Hospital of the West Indies, Jamaica: a 10-year experience. *Annals of tropical paediatrics*, *26*(1), 53–7. doi:10.1179/146532806X90619

Tusher, V. G., Tibshirani, R., & Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(9), 5116–21. doi:10.1073/pnas.091062498

Uh, Y., Hwang, G. Y., Jang, I. H., Kwon, O., Kim, H. Y., & Yoon, K. J. (2007). Antimicrobial susceptibility patterns and macrolide resistance genes of beta-hemolytic viridans group streptococci in a tertiary Korean hospital. *Journal of Korean medical science*, *22*(5), 791–794.

U.S. Department of Health and Human Services, (2013, August 28) Healthy People 2020 [Web site]. Retrieved from http://www.healthypeople.gov/2020/default.aspx

Valkenburg-van den Berg, A. W., Sprij, A. J., Oostvogel, P. M., Mutsaers, J. a E. M., Renes, W. B., Rosendaal, F. R., & Joep Dörr, P. (2006). Prevalence of colonisation with group B Streptococci in pregnant women of a multi-ethnic population in The Netherlands. *European journal of obstetrics, gynecology, and reproductive biology*, *124*(2), 178–83. doi:10.1016/j.ejogrb.2005.06.007

Vassiliadis, S., Ranella, A., Papadimitriou, L., Makrygiannakis, A., & Athanassakis, I. (1998). Serum levels of pro- and anti-inflammatory cytokines in non-pregnant women, during pregnancy, labour and abortion. *Mediators of inflammation*, *7*(2), 69–72. doi:10.1080/09629359891199

Verani, J. R., McGee, L., Schrag, S. J., & Division of Bacterial Diseases Centers for Disease Control and Prevention (CDC), N. C. for I. and R. D. (2010). Prevention of perinatal group B streptococcal disease--revised guidelines from CDC, 2010. *MMWR.Recommendations and reports : Morbidity and mortality weekly report.Recommendations and reports / Centers for Disease Control*, *59*(RR-10), 1–36.

Vieira, P., de Waal-Malefyt, R., Dang, M. N., Johnson, K. E., Kastelein, R., Fiorentino, D. F., … Moore, K. W. (1991). Isolation and expression of human cytokine synthesis inhibitory factor cDNA clones: homology to Epstein-Barr virus open reading frame BCRFI. *Proceedings of the National Academy of Sciences of the United States of America*, *88*(4), 1172–1176.

Welcsh, P. L. (2001). BRCA1 and BRCA2 and the genetics of breast and ovarian cancer. *Human Molecular Genetics*, *10*(7), 705–713. doi:10.1093/hmg/10.7.705

White, G. P., Watt, P. M., Holt, B. J., & Holt, P. G. (2002). Differential patterns of methylation of the IFN-gamma promoter at CpG and non-CpG sites underlie differences in IFN-gamma gene expression between human neonatal and adult CD45RO- T cells. *Journal of immunology (Baltimore, Md. : 1950)*, *168*(6), 2820–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11884451

Wilhelm-Benartzi, C. S., Koestler, D. C., Karagas, M. R., Flanagan, J. M., Christensen, B. C., Kelsey, K. T., … Brown, R. (2013). Review of processing and analysis methods for DNA methylation array data. *British Journal of Cancer*, *109*(6), 1394–402. doi:10.1038/bjc.2013.496

Willett, W. C. (2002). Balancing life-style and genomics research for disease prevention. *Science (New York, N.Y.)*, *296*(5568), 695–8. doi:10.1126/science.1071055

Wright, M. L., Ralph, J. L., Ohm, J. E., & Anderson, C. M. (2013). DNA methylation in complex disease: applications in nursing research, practice, and policy. *Nursing Outlook*, *61*(4), 235–241.e4. doi:10.1016/j.outlook.2013.04.010

Yancey, M. K., Duff, P., Kubilis, P., Clark, P., & Frentzen, B. H. (1996). Risk Fators for Neonatal Sepsis. *Obstetrics and gynecology*, *87*(2), 188–194.

Yang, M.-J. J., Sun, P.-L. L., Wen, K.-C. C., Chao, K.-C. C., Chang, W.-H. H., Chen, C.-Y. Y., & Wang, P.-H. H. (2012). Prevalence of maternal group B streptococcus colonization and vertical transmission in low-risk women in a single institute. *Journal of the Chinese Medical Association : JCMA*, *75*(1), 25–28. doi:10.1016/j.jcma.2011.10.011

York, T. P. (2003). Approaches to the Statisitcal-Genetic Analysis of Association and Microarray Data. (Doctoral dissertation). Available from ProQuest Dissertation and Theses database. (UMI No. 3076861)

Zaina, S., Lindholm, M. W., & Lund, G. (2005). Nutrition and aberrant DNA methylation patterns in atherosclerosis: more than just hyperhomocysteinemia? *The Journal of Nutrition*, *135*(1), 5–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15623824

Zeuthen, L. H., Fink, L. N., Metzdorff, S. B., Kristensen, M. B., Licht, T. R., Nellemann, C., & Frokiaer, H. (2010). Lactobacillus acidophilus induces a slow but more sustained chemokine and cytokine response in naive foetal enterocytes compared to commensal Escherichia coli. *BMC immunology*, *11*, 2. doi:10.1186/1471-2172-11-2

Zusman, A. S., Baltimore, R. S., & Fonseca, S. N. (2006). Prevalence of maternal group B streptococcal colonization and related risk factors in a Brazilian population. *The Brazilian journal of infectious diseases : an official publication of the Brazilian Society of Infectious Diseases*, *10*(4), 242–246.