Cardiovascular Disease Risk Markers in Young Adults: 9p21 Genotype, Plasma Proteomics and Dietary Patterns

by

Sara Mahdavi

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Department of Nutritional Sciences
Faculty of Medicine
University of Toronto

© Copyright by Sara Mahdavi 2018
Cardiovascular Disease Risk Markers in Young Adults: 9p21 Genotype, Plasma Proteomics and Dietary Patterns

Sara Mahdavi
Doctor of Philosophy

Department of Nutritional Sciences
Faculty of Medicine
University of Toronto
2018

Abstract

**Background:** 9p21 single nucleotide polymorphisms (SNPs) have been associated with cardiovascular disease (CVD) and to a lesser extend insulin sensitivity. Mechanisms by which 9p21 contributes to CVD development remain unknown since no annotated proteins are present. Previous studies have been in older, diseased-populations, and few have examined gene-diet interactions or the plasma proteome.

**Objective:** To determine the association of 9p21 SNPs with traditional and emerging biomarkers of CVD risk in a diverse ethnocultural group of young-adults and assess if habitual dietary patterns modify these associations.

**Methods:** Subjects were 1,639 young adults, from the Toronto Nutrigenomics and Health Study. Five 9p21 SNPs were genotyped. Traditional and emerging biomarkers of CVD risk were
assessed from fasting blood. Dietary patterns were determined using principal components analysis with a 196-item food frequency questionnaire. Linear regression was used to assess all associations.

**Results:** 9p21 high-risk alleles were associated with higher fasting insulin, especially in women not on hormonal contraceptives (HC). Consuming a Prudent diet was associated with a lower mean fasting insulin in homozygous carriers of risk alleles, but no associations were found with diet in non-risk alleles carriers. Gene-diet interactions were associated with lower HOMA-IR with five SNPs and lower HOMA-Beta with one SNP. When analyzing the plasma proteome; in Caucasians, risk alleles from four SNPs were associated with higher Apolipoprotein-E and two with higher Haptoglobin; while in East Asians, three SNPs were associated with higher α-1b-Glycoprotein concentrations; two with higher Apolipoprotein-B-100 and one with higher Apolipoprotein-E and Haptoglobin and in South Asians, one was associated with higher Apolipoprotein-B-100.

**Conclusion:** 9p21 genotypes were associated with markers of insulin sensitivity and plasma proteins, which varied by ethnicity, sex and HC use. The 9p21 genotype associations with insulin markers were modifiable by a Prudent dietary pattern in the high-genetic risk group only.
Acknowledgments

Completion of this thesis would not have been possible without the guidance and continuous support of key individuals in my personal and professional life, who I would like to thank here.

First and most important, I would like to sincerely thank my mentor and PhD supervisor, Dr. Ahmed El-Sohemy, from whom I have learned the most in this journey. These learnings have well extended beyond becoming an effective scientist, but also how to be a mentor, colleague and an entrepreneur. Despite his many impressive scholastic accomplishments at such a young age, family commitments and endless international leadership engagements in the field of personalized medicine, Dr. El-Sohemy always found time for me as his graduate student. Furthermore, he has never hesitated to provide me additional opportunities to explore the vast array of experiences that undoubtedly contributed to the making of a well-rounded new scientist. However, what has aided me the most in the completion of this journey, was his skillfulness in creating a perfect balance between nurturing independent thinking while providing critical feedback. In addition, his caring approach made easier the most challenging times of pursuing a doctorate degree in one of the most prestigious universities in the world. Words cannot express my gratitude to him for all that he has instilled in me.

To my Doctorate Advisory Committee, each of whom uniquely contributed to my learnings and completion of my thesis in record time, I thank you! I am so incredibly humbled to have had a legendary scientist and physician, Dr. David Jenkins, as my PhD Co-Supervisor. His deep wisdom and vast knowledge along with his kind demeanor in mentorship, has encouraged me to work harder and think more critically about my research. I am also very grateful to Dr. John Sievenpiper for his insightful comments and recommendations that have shaped my thinking of Cardiovascular Disease as a clinician. Likewise, I am also very thankful to Dr. Anthony Hanley for being one of my professors, a mentor in Clinical Epidemiology and key member of my advisory committee. His kind and honest feedback made me gain confidence in my research.

Lastly, I would like to thank my small family and many friends who have always encouraged me to be my best self. Without their unconditional love and believe in me, I would have never been able to achieve so many of my dreams, including this degree. This is especially true of my husband and life partner who has been a pillar of strength, particularly in the final stages of my PhD.
I would like to dedicate this thesis with endless love to Sofia and Caspian, who are to be born shortly after submission of this thesis! I am so excited to meet the two of you and thank you for making today, August 8, 2018, the most extraordinary day of my life.
# Table of Contents

Acknowledgments........................................................................................................ iv
Table of Contents ............................................................................................................ vi
List of Tables .................................................................................................................. viii
List of Figures .................................................................................................................. ix
Chapter 1 Introduction ....................................................................................................1
  1 Introduction .................................................................................................................2
Chapter 2 Review of Literature .....................................................................................4
  2 Background: Cardiovascular Disease .......................................................................4
    2.1 Definition, Incidence and Prevalence ....................................................................4
  Figure 2.1: Overview of Cardiovascular Disease and its Risk Factors .......................5
    2.2 Pathophysiology .................................................................................................7
  Figure 2.2: Insulin Resistance and Beta-cell Dysfunction ...........................................8
    2.3 Risk Factors ......................................................................................................9
    2.4 Biomarkers .......................................................................................................10
      2.4.1 Traditional Markers ..................................................................................10
  Figure 2.3: Results of Random-Effect Meta-Analysis Comparing Cardiovascular Disease
    Risk for an Increase of 1-SD ..................................................................................12
      2.4.2 Insulin Sensitivity Markers .....................................................................13
      2.4.3 New Biomarkers ..........................................................................................14
      2.4.4 Proteomics Markers ..................................................................................15
    2.5 Genetics .............................................................................................................17
      2.5.1 Common polymorphisms in 9p21 Region ..................................................18
  Figure 2.4: Signal-Intensity Plots Showing the Association of Single-Nucleotide
    Polymorphisms (SNPs) with Coronary Artery Disease in the Genome Wide Association
    Analysis for the Trust Case Control Consortium (WTCCC). ...................................20
  2.6 Diet .......................................................................................................................21
      2.6.1 Prudent Dietary Pattern ............................................................................22
      2.6.2 Principal Component Analysis ..................................................................23
      2.6.3 Prudent Diet and 9p21 Genotype Interaction .............................................25
Chapter 3 Rationale and Objectives .............................................................................27
  3 Rationale and Objectives ..........................................................................................28
    3.1 Rationale ...........................................................................................................28
    3.2 Objectives .......................................................................................................29
Chapter 4 9p21 Genotype and CVD Biomarkers in Young-Adults .................................30
  4 9p21 Genotype and CVD Biomarkers in Young-Adults .............................................31
    4.1 Abstract ............................................................................................................31
    4.2 Introduction .......................................................................................................32
    4.3 Methods ...........................................................................................................34
      Study Population .................................................................................................34
      DNA Analysis and Ethnocultural Determination ...............................................34
      Statistical Analysis .............................................................................................35
    4.4 Results .............................................................................................................37
  Figure 4.1: Fasting Serum Insulin by 9p21 Genotype (rs10757278) in Women (A) and Men
    (B) .......................................................................................................................42
List of Tables

Table 2.1: 54-Plasma Proteins Proteomics Analysis with Possible Physiological Roles .......... 16

Table 2.2: Principal Component Analysis Loading Factors for Top Ten Food Items of Three Toronto Nutrigenomics and Health Study Dietary Patterns ............................................ 24

Table 4.1: Subject Characteristics ......................................................................................... 37

Table 4.2: Biomarkers of CVD Risk by Rs10757278 Genotype and Ethnocultural Group .... 38

Table 4.3: Interaction of 9p21 SNPs with Sex and Ethnocultural Groups ............................ 41

Table 4.4: Fasting Insulin Associated with rs10757278 according to HC use in Women ....... 43

Table 5.1: Subject Characteristics According to 9p21 Genotype (rs1333049) ...................... 62

Table 5.2: Biomarkers of CVD Risk by 9p21 Genotype Risk (rs1333049) in Subgroups ......... 63

Table 5.3: Fasting Insulin by 9p21 Genotype, and Prudent Dietary Categories ....................... 65

Table 5.4: Interaction Between Dietary Components and 9p21 SNPs on HOMA-IR .............. 66

Table 6.1: Subject Anthropometric Characteristics, According to 9p21 Genotype .............. 81

Table 6.2: Prevalence of Risk Alleles in Four 9p21 SNPs in Each Ethnocultural Group ........ 82

Table 6.3: Plasma Protein Concentrations of α-1b-Glycoprotein, Apolipoprotein B-100, Apolipoprotein E, and Haptoglobin β-chain in Four 9p21 SNP and Ethnocultural Group ........ 83
List of Figures

Figure 2.1: Overview of Cardiovascular Disease and its Risk Factors.......................................................... 5

Figure 2.2: Insulin Resistance and Beta-cell Dysfunction.............................................................................. 8

Figure 2.3: Results of Random-Effect Meta-Analysis Comparing Cardiovascular Disease Risk for an Increase of 1-SD......................................................................................................................... 12

Figure 2.4: Signal-Intensity Plots Showing the Association of Single-Nucleotide Polymorphisms (SNPs) with Coronary Artery Disease in the Genome Wide Association Analysis for the Trust Case Control Consortium (WTCCC)............................................................................................................ 20

Figure 4.1: Fasting Serum Insulin by 9p21 Genotype (rs10757278) in Women (A) and Men (B) ............................................................................................................................................................................. 42

Figure 5.1: Homeostatic Model Assessment of Insulin Resistance (HOMA-IR) by Prudent Dietary Score Categories and 9p21 Genotype (Rs10757274). ......................................................................................................................... 67

Figure 5.2: Homeostatic Model Assessment of Beta Cell Dysfunction (HOMA-Beta) by Prudent Dietary Score Categories and 9p21 Genotype (Rs10757274). ........................................................................................................... 68
Chapter 1
Introduction
1 Introduction

Cardiovascular disease (CVD) is a leading cause of death worldwide \(^1\) and its pathogenesis involves both environmental and genetic factors \(^2-^4\). In 2007, genome wide association studies (GWAS) identified an 8-kilobase interval of DNA in chromosome 9 that is highly associated with CVD. 9p21 is located near tumor suppressor genes \(CDKN2A\) and \(CDNK2B\), but contains no annotated genes \(^5,^6\), hence mechanisms by which it increases CVD risk are not known. To date, 9p21, is the most robust genetic predictor of CVD with a population attributable risk of 20% for CVD events in homozygous risk carriers versus non-carriers. This risk ratio is increased to 30% for early-onset cases. In one study, homozygous risk carriers were twice as likely as likely to suffer from myocardial infarction compared to non-carriers where the risk allele frequency of 9p21 occurs in 40-50% of most populations \(^7\).

Previous research has not been able to link 9p21 risk alleles to any traditional biomarkers of CVD such as dyslipidemia \(^8\) in the populations studied, which have often been middle-aged or older. Meanwhile, preventing the development of CVD may require identifying biomarkers of early risk in young adults, well before the onset of disease-associated processes have caused tissue damage \(^9,^10\). Proteomics technologies increasingly permit the evaluation of systemic changes in protein expression in various tissues, and have been applied to early biomarker discovery in CVD \(^11\). Plasma levels of several novel protein markers of CVD appear to differ across different populations \(^12\). As such, a panel of 54 abundant plasma proteins has been shown to be associated with distinct characteristics, some of which are associated with cardiometabolic disease, including dietary patterns \(^13\), hormonal contraceptive use \(^14\), gluten intake \(^15\), and vitamin C \(^16\), D \(^17\) and, E \(^18\) status in an ethnoculturally diverse group of young adults. This panel may also be useful in
revealing specific protein patterns associated with 9p21 genotypes in a young ethnoculturally diverse group of adults.

In addition to identifying biomarkers of disease, CVD related deaths are preventable through lifestyle modification 19,20, including altering dietary habits 1,21. Given the long lag times between exposure to environmental risks and CVD outcomes, more emphasis on preventive approaches to modify CVD at an earlier age may be more advantageous than treatment later in life 9,10,22. Combining dietary recommendations with genetic risk stratification, also known as personalized nutrition, has been shown to be more effective in aiding young adults change their dietary habits than dietary advice alone 23. In addition, in order to improve cardiometabolic outcomes, a recent review highlights the importance of considering combing traditional and emerging factors in precision nutrition including microbiota, metabolomics, deep phenotyping, food behaviour, physical activity and dietary habits in future studies 24.

The main aim of this thesis is to assess the association between 9p21 genetic risk with traditional and emerging biomarkers of CVD risk, and to determine if dietary factors interact with these associations in a group of young adults.
Chapter 2
Review of Literature

2 Background: Cardiovascular Disease

2.1 Definition, Incidence and Prevalence

Cardiovascular disease (CVD) is a leading cause of mortality in Canada and worldwide \(^1\) and its pathogenesis involves both modifiable and non-modifiable factors, such as nutrition and genetics, respectively \(^2-4\) (Figure 2.1). CVD is defined as a collection of ailments that affect the heart and blood vessels with intermediate phenotypes such as atherosclerosis or endothelial dysfunction. CVD risk can be measured as intermediate outcomes using surrogate markers such as blood lipoproteins, blood pressure or plasma level of select hormones and enzymes \(^20\). CVD events occur as hard outcomes such as Cerebral Vascular Accident (CVA), Coronary Heart Disease (CHD) Myocardial Infarction (MI), Peripheral Vascular Disease (PVD), and aneurysms \(^20\).

In Canada, approximately 40% of adult men and women are at risk of CVD with approximately 25% of the population having a family history of heart disease and or stroke \(^25\). According to the Public Health Agency of Canada’s Canadian Chronic Disease Surveillance System 2012/2013 report, 2.4 million Canadians live with diagnosed heart disease and they are four to six times more likely to die prematurely than those without heart disease \(^26\). Men are twice as likely to suffer from a heart attack and on average are ten years younger than women when they are newly diagnosed (Men’s age range at first diagnosis 55-64 years versus 65-74 in women) \(^26\).
Figure 2.1: Overview of Cardiovascular Disease and its Risk Factors

Much effort has been spent in new drug discovery and surgical techniques in treating CVD. However, despite medical advancements in treatment of CVD, it remains the most prevalent disease worldwide, although, the majority of CVD events are preventable. It is estimated that up to 80% of CVD related deaths are preventable through lifestyle modification, including through altering dietary habits. Given the long lag times between exposure to environmental risks and CVD outcomes, more emphasis on preventive approaches to modify CVD at an earlier age may be more advantageous than treatment later in life. Although CVD is often diagnosed after an acute ischemic CVD event in middle or late-life, its pathogenesis starts decades before.
In a landmark autopsy study of children, youth and young adults who had died of secondary causes, evidence of atherosclerosis was found as early as the second decade of life.\textsuperscript{28} Atherosclerotic regions became progressively more prominent with each decade of life and with the presence of each additional modifiable CVD risk factor.\textsuperscript{28} A recent study of factors differentiating biological versus chronological aging processes in young-adults also showed that several factors, including genetics and diet determine rate of biological aging and susceptibility to chronic disease process.\textsuperscript{29} and perhaps these findings could be extrapolated to atherosclerosis. The study found that during early adulthood, dietary and genetic factors, in addition to others, either delayed or accelerated biological aging up to three times that of the chronological age (up to three years of biological aging could be accounted for every 12 months of chronological aging).\textsuperscript{29} Similarly, these factors may play a role in differences in susceptibility to chronic disease development such as CVD; providing a unique opportunity to combine genetic and dietary factors for early prevention of atherosclerosis that is personalized for those at the highest risk of disease.

CVD is a complex disease with many manifestations, risk factors and underlying mechanisms. Of importance to this thesis, a review of well-known pathophysiology, traditional biomarkers and new and upcoming biomarkers using plasma proteomics are presented here.
2.2 Pathophysiology

Early CVD signs are often presented as intermediate phenotypes, including dyslipidemia, hypertension, and hyperinsulinemia. Abnormal biomarkers can be clinically detected by comparing results to normal population ranges standardized for age, sex and other risk factors. At this stage, the underlying tissue damage might be reversible by early pharmaceutical and/or lifestyle intervention. Having one or more of these phenotypes however, increases the likelihood of suffering a CVD event sooner, often defined as heart disease and stroke. However, other major organ damage such as kidney disease, type 2 Diabetes Mellitus, non-alcoholic steatohepatitis and abdominal aneurysms, are among other serious ailments that could also result from those same CVD intermediate phenotypes, if they are not managed early on.

The most common underlying syndrome of CVD, atherosclerosis, is a complex disease process, which involves lipid metabolism, inflammation, endothelial dysfunction, and other pathophysiological aspects. Atherosclerosis is also influenced by genetics and nutrition, amongst other factors, however, our understanding of the interplay of underlying mechanisms and risk factors is still evolving.

In addition, insulin resistance is a major contributor to CVD. Tissue-specific insulin resistance in adipose, liver and skeletal muscle, together with central adiposity are often the hallmark of systemic insulin resistance and a major predictor of CVD and T2DM. One theory is that prolonged elevated plasma free fatty acids, which are often stored in these three tissues, compromises their ability to respond to circulating insulin. In response, the pancreas produces more insulin, which may increase uptake of more nutrients by the muscle, adipose and liver,
further increasing their resistance to circulating insulin (Figure 2.2). Abnormal and sustained increased concentration of glucose in plasma is a distinguishing feature of DM. Untreated hyperglycaemia promotes several CVD risk factors leading to micro and macro vascular damage \(^{38-40}\), as well as coagulation abnormalities through non-enzymatic glycation and increased oxidative stress \(^{41}\). In additional to glucose-mediated metabolic pathways, insulin also directly interacts with vascular tissue, contributing to endothelial integrity \(^{42}\), vascular inflammation \(^{43}\), coronary artery remodelling \(^{44}\), as well as mediating atherosclerotic plaque formation and stability \(^{45}\).

**Figure 2.2: Insulin Resistance and Beta-cell Dysfunction**

Increased fatty acids in the circulation activate glucose production in the liver and decrease glucose uptake in adipose and muscle tissue, requiring additional insulin to stimulate uptake (insulin resistance). Beta-cells initially increase insulin production (hyperinsulinemia), however, if prolonged insulin resistance continues, and beta-cells cannot meet the increasing demand, or insulin production decreases, leading to excess glucose circulating (hyperglycemia) then type 2 diabetes develops. *Reprinted from reference* \(^{46}\) *with permission.*
2.3 Risk Factors

Risk factors for CVD often require years of exposure to develop into local and systemic pathology\textsuperscript{20}. Intermediate phenotypes of CVD, before major organ damage, present an opportunity to slow down its progression, and in some instances, reverse tissue injury\textsuperscript{22}. The long lag time between exposure to risk and disease development, however, also translates to long-term interventions and life-long behavior modifications to adhere to in order to improve outcomes. Many of these lifestyle modifications, are also in direct contradiction to modern lifestyles, requiring not only changes in how CVD is treated on an individual level, but also from a clinical and public health perspective. Given the commitment required to make these changes, new technologies and approaches to improve personal motivation as well as systematic approaches to CVD prevention and management are needed. Personalized medicine has been an innovative approach presenting an opportunity to improve personal motivation to adhere to interventions needed for chronic disease management\textsuperscript{47}. In addition, clinical risk prediction of some chronic diseases have improved when genetic susceptibility has been added to traditional risk factors\textsuperscript{48}. From a public health perspective, a more targeted approach to disease prevention initiatives may also be more effective in motivating families and communities than generalized health claims and recommendations\textsuperscript{49}. This approach has been shown to be effective in dietary sodium reduction efforts\textsuperscript{50}, where dietary adherence measures were compared between public health dietary guidelines and DNA-based dietary recommendations. In this randomized control trial, subjects either received general healthy eating guidelines or specific dietary advice according to their DNA, with disclosing genetic susceptibility to develop hypertension, targeting those at the highest risk of harm by non-adherence to the recommendations of strict sodium reduction\textsuperscript{50}. 
2.4 Biomarkers

2.4.1 Traditional Markers

Biomarkers are measurable biological factors such as lipids, enzymes, hormones, proteins, and metabolites, often used as a surrogate indicator of a normal or pathological state. In CVD, these typically include plasma circulating biochemical components that mark a specific or several biological pathways. However, other anthropometric and functional tests can also be used as surrogate “markers” of CVD risk such as blood pressure, central obesity and body mass index. Most CVD predictive clinical scoring takes into account several of these markers to estimate a 10-year likelihood of CVD events, which can then be used for clinical decision making.

Commonly used CVD risk biomarkers are used in clinical settings. These include serum levels of low-density lipoproteins (LDL), high-density lipoproteins (HDL), and triglycerides (TG), and at times include glycemic measure such as fasting glucose, glycosylated hemoglobin (HbA1C), and fasting insulin. Many clinicians may also use measures of inflammatory markers such as C-reactive protein (CRP) after results of the JUPITER trial showing a significant reduction in CVD events by lowering CRP using Rosuvastatin therapy in those with normal LDL, while others argue that the clinical utility of CRP in CVD treatment is still questionable. Despite recent advances in CVD treatment, prediction of cardiovascular events still relies on dyslipidemia although additional factors such as family history, smoking, age and anthropometrics have been added to more recent versions of the original Framingham Risk Score. A recent RCT (CANDOS trial) demonstrated that using an antiinflammatory medication, canakinumab targeting interleukin-1-Beta, reduced CRP and fatal CVD events without any changes to LDL. However, risk algorithms, such as the Framingham risk score, that are mainly based on traditional biomarkers
have lower predictive power for some populations \(^{56}\), although in many others \(^{57}\) it has been an effective tool to predict and prevent CVD events. In addition, the strongest factor in calculating of CVD event risk in the Framingham Score is age; therefore, its predictive power is lower for early onset of CVD. Given the long lag times between exposure to environmental risks and CVD outcomes, emphasis on preventive approaches to modify CVD risk at an earlier age may be more advantageous than treatment later in life \(^{9,10,22}\).

Most studies of CVD risks to date are conducted in older, diseased adults. Although CVD is often diagnosed after an acute ischemic event in middle or late-life, its pathogenesis starts decades earlier \(^{27}\), when clinical symptoms of CVD may be absent according to traditional markers. Many individuals develop CVD in the absence of conventional risk factors \(^{58}\), and as many as 80% of CVD cases are only diagnosed after a myocardial infarction (MI). The pathogenesis of CVD often starts decades before an ischemic event \(^{27}\), when clinical signs and symptoms may be absent according to traditional markers. Therefore, refining methods for early detection of CVD risk are needed to implement preventive measures that will reduce the global incidence of disease.

Insulin resistance and dysglycemia have also been associated with higher CVD incidences \(^{59,60}\), however inclusion of these measures in predicting CVD risk among non-diabetics is subject to debate \(^{61–63}\). One of the algorithms used to assess CVD risk, the Reynolds Risk Score \(^{64}\), incorporates HbA1C, however, it is only used in those with diagnosed diabetes. Overall, both the Framingham Risk Score and the Reynolds Risk Score have lower predictability of CVD hard outcomes in those without pre-existing diabetes and or CVD \(^{65}\). Fasting insulin has been shown to be a better predictor of 10-year risk of CVD in non-diabetic women compared to other glycemic and lipid measures \(^{66}\). In fact, in a cohort of 3,246 older women higher insulin
resistance as measured by homeostasis model assessment for insulin sensitivity (HOMA-IR) was the only accurate determinant of CVD events (CHD or and stroke) where as baseline LDL, HDL, TG, fasting glucose and HgA1C did not affect the outcomes \(^{66}\). A meta-analysis confirmed pooled relative risk of CVD events was directly linked with increasing HOMA-IR in non-diabetic populations \(^{67}\) (Figure 2.3).

<table>
<thead>
<tr>
<th>Number of Participants</th>
<th>Pooled relative risk per 1 SD (95% CI)</th>
<th>(I^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CHD</strong> Glucose</td>
<td>23 (140,721)</td>
<td>1.21 (1.13, 1.30)</td>
</tr>
<tr>
<td>Insulin</td>
<td>9 (32,104)</td>
<td>1.04 (0.96, 1.12)</td>
</tr>
<tr>
<td>HOMA-IR</td>
<td>7 (17,452)</td>
<td>1.46 (1.26, 1.69)</td>
</tr>
<tr>
<td><strong>CVD</strong> Glucose</td>
<td>44 (450,487)</td>
<td>1.19 (1.14, 1.23)</td>
</tr>
<tr>
<td>Insulin</td>
<td>16 (46,236)</td>
<td>1.13 (1.05, 1.22)</td>
</tr>
<tr>
<td>HOMA-IR</td>
<td>17 (51,161)</td>
<td>1.25 (1.16, 1.35)</td>
</tr>
</tbody>
</table>

| **CVD** Glucose men    | 22 (183,802)                          | 1.13 (1.08, 1.18) | 29.3% |
| women                  | 14 (51,527)                           | 1.25 (1.11, 1.41) | 65.0% |
| Insulin men            | 10 (18,411\(^a\))                    | 1.06 (0.97, 1.16) | 60.4% |
| women                  | 4 (5,082\(^a\))                      | 1.24 (1.08, 1.44) | 18.5% |
| HOMA-IR men            | 6 (9,768)                             | 1.41 (1.12, 1.77) | 66.5% |
| women                  | 3 (5,049)                             | 1.37 (1.05, 1.80) | 33.6% |

\(^a\)1 study did not specify sex-specific numbers. SD, standard deviation; 95% CI, 95% confidence interval; \(I^2\), measure of heterogeneity; CHD, coronary heart disease; CVD, cardiovascular disease (myocardial infarction, angina pectoris, hemorrhagic stroke, ischemic stroke, arrhythmias, congestive heart failure or sudden cardiac death); HOMA-IR, Homeostasis Model Assessment Insulin Resistance. \textit{Reproduced from reference} \(^{67}\) \textit{with credit to Gast et al 2012}.
2.4.2 Insulin Sensitivity Markers

Higher fasting insulin levels are associated with numerous adverse risk factors in young adults as well, increasing the risk of atherosclerosis and subsequent type 2 diabetes. Even in nondiabetic individuals, hyperinsulinemia is associated with decreases in insulin-mediated glucose uptake, as well as with a number of clinical symptoms associated with insulin resistance. Prolonged and untreated insulin resistance and hyperinsulinemia are known to be associated with hypertriglyceridemia and low concentrations of HDL, hypertension, and coronary artery calcification, which are all risk factors for CVD. Hyperinsulinemia has also independently been associated with ischemic heart disease.

Alterations in several physiological pathways, such as innate immunity and lipid metabolism, lead to inflammation, thrombosis, calcification, vascular remodeling, oxidative stress and cell death, which contribute to the development of CVD. Insulin resistance is likely a subclinical marker of some of these compromised metabolic pathways, signaling reduced insulin-receptor mediated uptake of nutrients in an attempt to sustain homeostasis. Insulin resistance is often accompanied by one or more tissue impairments such as increased adiposity, reduced skeletal muscle mass and quality, and non-alcoholic fatty liver disease, independent of the metabolic syndrome. In addition to its major role in macronutrient metabolism, insulin also directly interacts with vascular tissue including endothelial tissue, coronary artery remodelling, and atherosclerotic plaque.
2.4.3 New Biomarkers

Additional pathways involved in the development of CVD have yet to be elucidated, and new biomarkers are needed to predict early dysregulation of both traditional and novel pathways in order to implement more successful CVD preventive strategies. These markers must be reproducible, reliable relatively easy to measure and cost efficient to have clinical utility. With advancement in laboratory techniques, the capacity to accurately and efficiently analyze disease biomarkers has significantly improved. This has also expanded our capacity for earlier, faster and more comprehensive CVD risk detection and diagnostic measures. In addition, the plasma proteome is a promising new and emerging tool with utility as biomarkers of CVD with improved disease onset predictability such as in earlier subclinical detection of CVD. Newer highly-multiplexed Multiple Reaction Monitoring-based (MRM) assays are now easy to use, robust, sensitive, and have high-throughput capabilities through short analysis time and complete automated sample preparation. Some of these assays are specifically designed for CVD determination, and classification using over 65 plasma proteins. There is evidence that predictive power increases with the number of biomarkers measured. Since larger collection of biomarkers might better portray a collection metabolic abnormalities, these new assays are therefore well suited for CVD biomarker validation and discovery in using the plasma proteome where multiple biological process could be accounted for simultaneously. Since CVD is a complex disease with multiple pathogenic pathways, new biomarkers over the past decade have shown to be predictive of these additional pathways including trimethylamine N-oxide (TMAO), small dense LDL, serum amyloid A, and lipoprotein A, and, interleukin-6. Similarly, adding the plasma proteome has the potential to better predict CVD risk when added to the traditional handful of biomarkers and anthropometric measures currently used.
2.4.4 Proteomics Markers

In the past, accurately measuring the concentration of more than a few proteins at the same time has been difficult. However, novel proteomic assays now allow for the simultaneous measurement of multiple proteins along diverse physiological pathways. This allows testing combinations of multiple biomarkers and increases their predictive power. Recent proteomics technologies allow evaluation of thousands of proteins and their post-translationally modified protein isoforms at once, permitting better potential understanding of atherosclerosis at the cellular level. This could potentially allow for discovery of new biomarkers of CVD earlier and in younger individuals, since these populations have less confounding factors present than their older counterparts.

Amongst proteomic panels, plasma proteins are considered to be the largest and most diverse collection of the human proteome. Although upwards of 3000 different proteins have been identified in human plasma, a relatively small number of them account for the majority of the total protein mass. The composition of plasma proteins is thought to differ under various physiological conditions, and most metabolic abnormalities are reflected in these differences. A range of 20 to 55 highly abundant plasma proteins have been speculated to be of metabolic importance and might be useful biomarkers of disease. Some of these plasma proteins have been identified as biomarkers for systemic inflammation and have been used as indicators of risk for cardiovascular disease. Multiple reaction monitoring (MRM) proteomic assays have been used to measure levels of 45 plasma proteins involved in inflammation and endothelial cell function, with further expansion to 54 plasma proteins (Table 2.1). More recent techniques are expanding the validity and reliability of these proteomic assays for clinical utility.
Table 2.1: 54-Plasma Proteins Proteomics Analysis with Possible Physiological Roles

<table>
<thead>
<tr>
<th>Positive Acute phase reactants</th>
<th>Negative Acute Phase reactants</th>
<th>Innate Immunity</th>
<th>Coagulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Afamin</td>
<td>Albumin</td>
<td>Complement C3</td>
<td>Fibrinogen α₂ chain</td>
</tr>
<tr>
<td>Angiotensinogen</td>
<td>Antithrombin-III</td>
<td>Complement C4 β chain</td>
<td>Fibrinogen β chain</td>
</tr>
<tr>
<td>Apolipoprotein A-I</td>
<td>Apolipoprotein A-IV</td>
<td>Complement C4 γ chain</td>
<td>Fibrinogen γ chain</td>
</tr>
<tr>
<td>Apolipoprotein A-II</td>
<td>Apolipoprotein C-I</td>
<td>Complement C9</td>
<td>Fibrinopeptide A</td>
</tr>
<tr>
<td>Apolipoprotein B-100</td>
<td>Apolipoprotein E</td>
<td>Complement factor B</td>
<td>Fibronectin</td>
</tr>
<tr>
<td>Apolipoprotein C-I</td>
<td>Complement C1 inactivator</td>
<td>Complement factor H</td>
<td></td>
</tr>
<tr>
<td>Apolipoprotein C-III</td>
<td>Gelsolin, isoform 1</td>
<td>Haptoglobin β chain</td>
<td></td>
</tr>
<tr>
<td>Apolipoprotein L1</td>
<td>Histidine-rich glycoprotein</td>
<td>Serum amyloid PC</td>
<td></td>
</tr>
<tr>
<td>Ceruloplasmin</td>
<td>Transthyretin</td>
<td>α₁-Acid glycoprotein 1</td>
<td></td>
</tr>
<tr>
<td>Clusterin</td>
<td>α₂-Antiplasmin</td>
<td>α₁-Antichymotrypsin</td>
<td></td>
</tr>
<tr>
<td>Coagulation factor XIIa</td>
<td>α₂-Macroglobulin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complement C3</td>
<td>β₂-Glycoprotein I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hemopexin</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heparin cofactor II</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inter-α-trypsin inhibitor HC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kininogen-1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plasminogen</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prothrombin</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Retinol-binding protein</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transferrin</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vitamin-D-binding protein</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vitronectin</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α₁-Antitrypsin</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α₂-Antiplasmin</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α₂-HS-glycoprotein</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Adapted from references* 16,97
2.5 Genetics

Family history of CVD has long been a known risk factor for the disease \(^{101}\), and contributes to increase the risk of CVD independently of the other well-known risk factors \(^{102}\). The genetic variants predisposing to CVD span from rare mutations, such as those that cause familial hypercholesterolemia, to common polymorphisms that modulate the predisposition to complex diseases \(^{103}\). CVD is a complex disease with multiple genetic and environmental components contributing to its development. Since several causal pathways might lead to CVD it has been challenging to identify its genetic causes, other than those rare mutations with severe and early onset phenotypes. Traditional genetic risk factors for CVD are located in coding or regulatory regions of the DNA and may affect the protein sequence or the level of gene expression, such as that in APOE and PCSK9 \(^{104}\). However, although these factors have strong associations with CVD outcomes, they are rare variants, occurring in less than 1% of most populations. Meanwhile, premature CVD deaths accounted for 31% of global death in 2015, indicating other possible genetic causes of CVD are likely at play, amongst other risk factors.

Genome-wide association studies (GWAS) have been able to identify several loci for common single nucleotide polymorphisms (SNPs) that are associated with CVD events \(^{104}\). The majorities of these polymorphisms are located in non-coding regions of DNA, and have high prevalence in most populations. However, since these polymorphisms are not in regions with annotated genes, they are often difficult to directly connect to phenotypic impacts. This further complicates identifying mechanisms by which these variations affect CVD development, hindering efforts to discover therapeutic agents and biomarkers to prevent CVD events associated with them.
Adding these common DNA variations to well-known CVD risk factors however, might be beneficial in improving the accuracy of predicting CVD risk profile in order to adopt preventive or therapeutic measures.

2.5.1 Common polymorphisms in 9p21 Region

Genetic factors play an important role in CVD events as demonstrated by several large-scale GWAS. One such discovery was in chromosome 9 arm p section 21 (9p21) as markers for CVD risk (figure 2.4) with one risk allele prevalence of 40-50%. The 9p21 risk variants are independent CVD risk factors associated with MI, stroke and abdominal aortic aneurism. Multiple case-control studies in several ethnicities have confirmed these findings. However, despite the wealth of replication studies supporting a role for SNPs in 9p21 in CVD, the risk locus contains no protein coding sequences or known microRNAs, so its functional effects remain unknown. Furthermore, 9p21 risk alleles appear to be independent of traditional risk factors, including elevated lipid levels, high blood pressure, and obesity, suggesting alternative pathogenic mechanisms to those known to date.

An 8-kilobase interval in the 9p21 region contains SNPs that were originally associated with an increased risk of CVD in at least six independent populations involving more than 23,000 participants. These variants were in linkage disequilibrium (LD) and now have been shown to span a 58-kb region that has multiple neighboring genes (CDKN2A, CDKN2B, and MTAP), without annotating to any single protein sequence. SNPs in this region have also been associated with type 2 diabetes mellitus (T2DM) and insulin sensitivity, although the evidence is not as strong as those seen for CVD. Risk alleles of 9p21 are common in most populations with an estimated 30% increased risk of CVD in risk allele carriers compared to non-carriers. Four
of the most frequently reported 9p21 SNPs associated with CVD (rs10757274, rs10757278, rs4977574 and rs2383206) have an A>G substitution while another commonly reported SNP (rs1333049) has a G>C substitution. CVD risk increases with each copy of 9p21 risk alleles when compared to those who are homozygous for the ancestral alleles. The risk of CVD associated with 9p21 variants is even stronger in those 55 years and younger. In a meta-analysis of genetic risk factors for CVD, those who were heterozygous for the 9p21 risk allele had a 25% increased risk, while those who were homozygous for the risk allele had a 45% increased risk compared to those had no risk alleles. These associated risks were greater among those diagnosed with CVD before 55 years of age compared to those diagnosed before 75 years of age, suggesting that 9p21 plays a bigger role in earlier cases of CVD where environmental risk factors might not have had enough time to impact the development of CVD as much as in those older.
Figure 2.4: Signal-Intensity Plots Showing the Association of Single-Nucleotide Polymorphisms (SNPs) with Coronary Artery Disease in the Genome Wide Association Analysis for the Trust Case Control Consortium (WTCCC).

The −log P values are for the association of each SNP with coronary artery disease, from two-sided Cochran–Armitage tests for trend. The signals within each chromosome are shown on the x-axis, the data are plotted from the p-trend. Reproduced with permission from reference 119, Copyright Massachusetts Medical Society.
2.6 Diet

Lifestyle interventions, such as dietary modification, are effective in preventing CVD risk factors in many cases\textsuperscript{21,120,121}. Dietary factors have been shown to affect CVD outcomes\textsuperscript{21} and predictive biomarkers in dyslipidemia, endothelial dysfunction\textsuperscript{122,123}, inflammation\textsuperscript{124} and hyperinsulinemia\textsuperscript{125}. Consumption of specific food groups such as fruits and vegetables\textsuperscript{126,127}, whole grains\textsuperscript{128,129}, nuts and legumes\textsuperscript{130} as well as combining these foods to provide a “cardio-protective diet”, such as the Portfolio\textsuperscript{121} and Prudent Dietary Pattern\textsuperscript{123} have been shown to reduce risk factors for CVD.

In contrast to the traditional analytical approach used in nutritional epidemiology, dietary pattern analysis considers overall diet rather than individual nutrients or foods. Although traditional approaches to dietary analysis have important features, such as linking specific nutrients to disease pathways, they have several limitations in studying complex diseases such as CVD\textsuperscript{131}. These limitations include; not accounting for complex and multifactorial exposure to foods consumed, such as nutrient-nutrient interactions, nutritional bioavailability, and many confounding nutrients and foods that would be unaccounted for in free-living individuals. In addition, the effect of a single nutrient may be too small to detect, but the cumulative effects of multiple nutrients included in a dietary pattern may be sufficiently large to be detectable\textsuperscript{132}. Meanwhile, excess or lack of a single nutrient may just be a surrogate for a lifestyle. For example, high potassium intake may be a marker for veganism or high sodium intake might be a surrogate for frequent consumption of prepared foods, therefore confounding the link of the nutrients to complex diseases such as hypertension. Instead of looking at individual nutrients or foods, dietary pattern analyses examine the effects of overall diet which would represent a broader picture of food and nutrient consumption, and may thus be more predictive of disease risk than individual foods or nutrients\textsuperscript{133}.
2.6.1 Prudent Dietary Pattern

The Prudent Dietary Pattern was first described by Willett et al in 2000, where dietary patterns, rather than single-food or nutrients, were analyzed using a 131-item food frequency questionnaire (FFQ) in a prospective cohort to determine their association with CVD in the Health Professional Follow-up Study. This diet was defined using an *a posteriori* approach and therefore the exact factors that constitute a Prudent Dietary Pattern may vary depending on the population studied.

The Prudent Dietary Pattern was generally characterized by higher intake of vegetables, fruit, legumes, whole grains, fish, and poultry, compared to the "Western Pattern," which was characterized by higher intake of red meat, processed meat, refined grains, sweets and dessert, French fries, and high-fat dairy products. Those who had the highest scores in consuming the Prudent Dietary Pattern had relative risk reduction of 0.70 (95% CI: 0.56, 0.86) whereas those who had the highest Western Dietary Pattern score had an increase relative risk of 1.64 (95% CI: 1.24, 2.17) in coronary heart disease, after correcting for other CVD risk factors. Both Prudent and Western Dietary Patterns predicted CHD independent of all other CVD risk factors. Others have replicated similar findings in predicting CVD risk factors in younger adults as well as older adults although the definition and nutrient content of these dietary patterns might vary depending on the foods each population consumes and the items included in the food frequency questionnaires used to collect dietary data in the study. Dietary patterns have been more informative in linking diet to CVD giving a more complete picture of food consumption. Providing dietary intervention is also more effective using dietary patterns since most nutrients are not consumed in isolation, but rather as part of whole foods and meals patterns.
2.6.2 Principal Component Analysis

Principal Component Analysis (PCA) is an empirical statistical technique that is used to assess dietary patterns from a set of dietary data such as an FFQ. PCA utilizes preexisting correlations between different foods in the FFQ, and combines frequently consumed foods together in a linear function. The correlation matrix of these food intake variables are then used to identify common patterns of food consumption within the data in order to account for the largest amount of variation in the diet. Major components or distinct dietary patterns are identified using an eigenvalue and breaks on a scree plot. Eigenvalue and breaks on a scree plot are used to decide how many distinct dietary patterns are observed and should be retained in the study population. For each dietary factor, a cut-off factor loading value could be established in order to include or exclude the food in a pattern (this can range from factor loading of 0.20-0.40). The factor loading might also be used for post-hoc interpretations of the derived patterns to determine specific characteristics of the dietary patterns observed. Factor scores are then calculated for each of the derived patterns by summing the products of the observed consumption frequency and the factor loading for each of the significant food pattern. Each subject in the study will then have a composite score of the major components (dietary patterns) identified that can be used in the study analyses as their specific dietary score of the particular pattern or used as a surrogate marker of their adherence to a particular pattern. Examples of three common dietary patterns from the Toronto Nutrigenomics and Health (TNH) Study are shown in Table 2.2 with the top 10 food items included in each pattern and corresponding factor loading value of each item.
Table 2.2: Principal Component Analysis Loading Factors for Top Ten Food Items of Three Toronto Nutrigenomics and Health Study Dietary Patterns

<table>
<thead>
<tr>
<th>Food Item</th>
<th>Prudent Pattern Loading Factor</th>
<th>Western Pattern Loading Factor</th>
<th>Eastern Pattern Loading Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carrots</td>
<td>0.45</td>
<td>Pizza</td>
<td>White Rice</td>
</tr>
<tr>
<td>Zucchini</td>
<td>0.44</td>
<td>Hamburger</td>
<td>Liver</td>
</tr>
<tr>
<td>Peppers</td>
<td>0.44</td>
<td>French Fries</td>
<td>Pork</td>
</tr>
<tr>
<td>Strawberries</td>
<td>0.41</td>
<td>Doughnuts</td>
<td>Crustacean</td>
</tr>
<tr>
<td>Legumes</td>
<td>0.40</td>
<td>Fish Cakes</td>
<td>Beef &amp; Lamb</td>
</tr>
<tr>
<td>Apples</td>
<td>0.40</td>
<td>Lean Hamburger</td>
<td>Organ Meats</td>
</tr>
<tr>
<td>Red Peppers</td>
<td>0.40</td>
<td>Low Calorie Cola</td>
<td>Cabbage</td>
</tr>
<tr>
<td>Squash</td>
<td>0.39</td>
<td>Regular Colas</td>
<td>Spinach</td>
</tr>
<tr>
<td>Blueberries</td>
<td>0.39</td>
<td>Ketchup</td>
<td>Mixed Vegies</td>
</tr>
<tr>
<td>Lettuce</td>
<td>0.39</td>
<td>Mac &amp; Cheese</td>
<td>Fish Filet</td>
</tr>
</tbody>
</table>

Top 10 food items and their factor loading that were derived from a 196-item food frequency questionnaire to create three distinct dietary patterns. Loading Factor cut off was set at 0.25 for each food item to be included. Foods with higher factor loading scores for a particular pattern have a higher degree of similarity in intake response with the other foods in the pattern. The Prudent pattern was the largest pattern and consequently most predictive, explaining 6.8% of the between-person variation in dietary intake. Together, these three patterns explained 15.6% of the total between-person variation in reported food intake. *Adapted from reference* 136

Although dietary approaches have been identified as the cornerstone of CVD prevention in population studies 51 the magnitude of effectiveness of dietary practices in preventing CVD can vary significantly between individuals 142,143. Some of these variations have been attributed to intra-individual differences such as genetic differences 142, including SNPs in the 9p21 chromosomal region 144,145.
2.6.3 Prudent Diet and 9p21 Genotype Interaction

The risks of CVD as measured by CVD events have been associated with high-risk alleles in 9p21 \(^6,^{146-150}\). This risk has been shown to be modified by as much as two-fold with prudent dietary habits \(^{144,145}\).

One study reported a significant interaction \((p=4.0 \times 10^{-4})\) between rs2383206 in 9p21 and a factor-analysis-derived “Prudent” diet pattern score in the INTERHEART study \(^{145}\). A major component of the Prudent dietary pattern from that study was raw vegetables. An effect of 9p21 on MI was observed in the group with a low prudent diet score \((\text{OR}=1.32, \ p=6.82 \times 10^{-7})\), but the effect was diminished in a step-wise fashion in the medium \((\text{OR}=1.17, \ p=4.9 \times 10^{-3})\) and high prudent diet scoring groups \((\text{OR}=1.02, \ p=0.68)\) \((p=0.014\) for difference). The combination of the least prudent diet and two copies of the risk allele was associated with a 2-fold increase in risk for MI \((\text{OR}=1.98, \ p=2.11 \times 10^{-9})\). This report also analyzed the results of another large prospective epidemiological study, the FINRISK study, which used a closely related dietary variable and rs4977574. The risk allele associated with a larger effect on CVD risk in the groups with diets low or average for fresh vegetables, fruits, and berries \((\text{hazard ratio [HR]}=1.22, \ p=3.0 \times 10^{-4}, \ \text{and } \text{HR}=1.35, \ p=4.1 \times 10^{-3},\) respectively) compared to the group with high consumption of these foods \((\text{HR}=0.96, \ p=0.73)\) \((p=0.0011\) for difference). The combination of the least prudent diet and two copies of the risk allele was associated with a 1.66-fold increase in risk for CVD in the FINRISK study \((\text{HR}=1.66, \ p=0.0026)\) \(^{145}\).

Another report from the FINRISK study showed higher vegetable intake \((\text{hazard ratio (HR)}, 0.95 [\text{CI}: 0.91–0.996])\), wine intake \((\text{HR}, 0.91 [\text{CI}: 0.86–0.96])\), and total alcohol consumption \((\text{HR}, 0.92 [\text{CI}: 0.86–0.98])\) were associated with lower CVD incidence. The increased CVD incidence
by the G allele was restricted to individuals with medium or high vegetable intake (P for interaction = 0.043), and to non- and low consumers of wine (P for interaction = 0.029). Although rs4977574 did not associate with any known risk markers, stratification by vegetable intake and smoking suggested an interaction with rs4977574 on glycated hemoglobin and high-density lipoprotein cholesterol (P for interaction = 0.015 and 0.049, respectively)\(^{144}\).

However, the specific dietary factors that are “cardio-protective” for individuals with high-risk 9p21 genotypes have not been tested in an intervention trial. Furthermore, it is not known if the same dietary factors would be associated with similar cardiovascular biomarkers in a young cohort with high-risk 9p21 genotypes. In addition, early CVD biomarkers could aid in early identification of those at risk of CVD at a younger age would allow a more targeted approach in preventive interventions.
Chapter 3
Rationale and Objectives
3 Rationale and Objectives

3.1 Rationale

CVD is a leading cause of death and disability worldwide. Many CVD risk factors are modifiable, including dietary patterns. Genetic variability is a major non-modifiable CVD risk factor, with high-risk genotypes in 9p21 widely prevalent globally however mechanisms by which this DNA region modifies CVD risk are largely unknown. Two more recent studies demonstrated high Prudent dietary patterns might modify the risk of CVD events and glycemic measures by two fold in those who inherit high-risk genotypes of 9p21. However, it is not clear if these findings can be extended to younger cohorts and those without prior CVD events. Exposure to risk factors may take decades to manifest as clinical CVD events because of the long lag-time in subclinical disease development prior to clinical diagnosis of CVD. Investigation of early CVD biomarkers as well as discovery of new biomarkers that could predict subclinical CVD disease processes earlier in life would provide more evidence in linking 9p21 to CVD risk and possible modification through diet for earlier prevention.

This study will help fill the gap in the literature linking 9p21 genetic predisposition to CVD through early biomarkers of CVD risk in a young-adult cohort with a diverse ethnocultural background. In addition, this study will also investigate and validate the role of dietary patterns in modifying 9p21 genetic risk at a young age, providing evidence for more targeted intervention using gene-diet interaction.
3.2 Objectives

The overall objective of this study was to identify if there was an association between early biomarkers of CVD risk and 9p21 genotype in a young multiethnic population and if dietary factors modify these associations.

The specific objectives were to examine:

1. The association between 9p21 genotype and traditional CVD markers, and to determine whether ethnicity, sex and hormonal contraceptive use modify these associations.

2. The association between 9p21 genotype and insulin sensitivity, and to determine whether ethnicity, sex, hormonal contraceptive use, and dietary patterns modify these association.

3. The association between 9p21 genotype and abundant plasma proteins, and to determine whether ethnicity, sex, hormonal contraceptive use, and diet modify this association.
Chapter 4
9p21 Genotype and CVD Biomarkers in Young-Adults

This chapter is adapted from an article submitted to PLoS ONE May 2018. The original article is the following:

Mahdavi S, Jenkins JA, El-Sohemy, A. Genetic Variation in 9p21 is associated with Fasting Insulin in Women but not Men. Submitted May 2018
4 9p21 Genotype and CVD Biomarkers in Young-Adults

4.1 Abstract

**Background:** Single nucleotide polymorphisms (SNPs) in the 9p21 region have been associated with cardiovascular disease (CVD), but previous studies have focused on older populations. The objective of this study was to determine the association between 9p21 genotypes and biomarkers of CVD risk in young adults from different ethnocultural groups.

**Methods:** Subjects were 1,626 participants aged 20-29 years from the Toronto Nutrigenomics and Health Study. Fasting blood was collected to measure glucose, insulin, c-reactive protein, serum lipids, and DNA for genotyping subjects for five SNPs in 9p21. Analyses were conducted for the entire population and separately for women (n=1,109), men (n=517), Caucasians (n=771), East Asians (n=561) South Asians (n=175) and Others (n=119). ANOVA and ANCOVA were used to examine if 9p21 genotypes were associated with biomarkers of CVD risk.

**Results:** In the entire group, the risk alleles of rs10757278 and rs2383206 were associated with higher mean insulin (p=0.01). Risk alleles for rs4977574, rs10757278, rs2383206, rs1333049 and rs10757274 were associated with higher serum insulin in women (p=0.008, p=0.008, p=0.0003, p=0.002, and p=0.001, respectively), but not in men (p=0.41, p=0.13, p=0.31, p=0.34, and 0.35, respectively). The association between 9p21 and insulin remained significant among women not taking hormonal contraceptives, but was not significant among women taking it. **Conclusion:** Our findings suggest that 9p21 genotypes may play a role in the development of insulin resistance in early adulthood among women, but not men, and the effects appear to be attenuated by HC use.
4.2 Introduction

Cardiovascular disease (CVD) is a leading cause of mortality in Canada and worldwide and its pathogenesis involves both environmental and genetic factors. CVD risk is often associated with a cluster of intermediate phenotypes including glucose intolerance, dyslipidemia, hypertension and abdominal obesity, collectively referred to as markers of cardiometabolic disease (CMD), which are predictive of higher rates of CVD events later in life. Genetic factors play an important role in CVD events as demonstrated by several large-scale genome-wide association studies (GWAS). An 8-kilobase interval in the 9p21 region contains single nucleotide polymorphisms (SNPs) that have been associated with an increased risk of CVD in at least six different populations involving more than 23,000 participants. These variants were in linkage disequilibrium (LD) and now have been shown to span a 58-kb region that has multiple neighboring genes (CDKN2A, CDNK2B, and MTAP), without annotating to any single protein sequence. SNPs in this region have also been associated with type 2 diabetes mellitus (T2DM) and insulin sensitivity, although the evidence is not as strong as those seen for CVD. Risk alleles of 9p21 are common in most populations and an estimated 50% of Caucasians carry at least one copy with an estimated 30% increased risk of CVD compared to non-carriers. Four of the most frequently reported 9p21 SNPs associated with CVD (rs10757274, rs10757278, rs4977574 and rs2383206) have an A>G substitution while another commonly reported SNP (rs1333049) has a G>C substitution. CVD risk increases with each copy of 9p21 risk alleles when compared to those who are homozygous for the ancestral alleles. The risk of CVD associated with 9p21 variants is even stronger in those 55 years and younger. In a meta-analysis of genetic risk factors for CVD, those who are heterozygous for the 9p21 risk allele have a 25% increased risk, while those who are homozygous for the risk allele have a 45% increased risk.
These associated risks were greater among those diagnosed with CVD before 55 years of age compared to those diagnosed before 75 years of age, suggesting that 9p21 plays a bigger role in earlier cases of CVD where environmental risk factors might not have had enough time to impact the development of CVD. Given the long lag times between exposure to environmental risks and CVD outcomes, emphasis on preventive approaches to modify CVD risk at an earlier age may be more advantageous than treatment later in life, however, most studies are conducted in older, diseased adults. Although CVD is often diagnosed after an acute ischemic event in middle or late-life, its pathogenesis starts decades earlier, when clinical symptoms of CVD may be absent according to traditional markers.

In a landmark autopsy study of children, youth and young adults (ages 15 to 36) who had died of trauma (motor vehicle accidents, gun violence, other accidental injuries), evidence of atherosclerosis was found as early as the second decade of life. Atherosclerotic lesions became progressively more prominent with each decade of life and with the presence of each additional modifiable CVD risk factor. A recent study of factors differentiating biological versus chronological aging processes in a cohort of 26-38 year-olds showed that several factors, including genetics and diet determine rate of biological aging and susceptibility to chronic disease. Similarly, both diet and genetics likely play a role in early development of CVD.

We aimed to test the hypothesis that 9p21 risk variants are associated with biomarkers of CVD risk in a population of young adults from different ethnocultural groups.
4.3 Methods

Study Population

Participants were from the Toronto Nutrigenomics and Health Study (TNHS), which has been described elsewhere. In brief, the TNHS is a cross-sectional study that aims to explore the link between diet, genes and biomarkers of chronic disease and was approved by the University of Toronto Research Ethics Board. Study participants were aged 20-29 years from various ethnocultural backgrounds and were recruited from the University of Toronto campus between 2004 and 2010. Anthropometric measurements including height, weight, waist circumference, and blood pressure were recorded according to standard procedures. Subjects provided a fasting blood sample for DNA isolation and plasma was separated for measuring biomarkers of CVD risk including blood lipids, inflammatory markers and insulin.

DNA Analysis and Ethnocultural Determination

DNA was extracted from whole blood samples using standard procedures and analyzed for the following SNPs in 9p21: rs2383206, rs10757274, rs10757278 and rs1333049. These SNPs have been shown most consistently to be associated with CVD risk in several populations and ethnicities globally. Genotyping of 9p21 SNPs was completed at the Clinical Genomics Centre in the Princess Margaret Hospital, University Health Network, using the iPLEX Gold assay with mass spectrometry-based detection (Sequenom MassARRAY platform; Sequenom Inc) for all subjects. Ethnocultural status was determined by asking subjects in an open-ended format to self-report the ethnocultural group(s) they identify with. Subjects were then categorized into the three most commonly reported ethnocultural groups based on self-reports and all others into a separate group.
Caucasians included those who considered themselves European, Middle Eastern, or White-Hispanic. East Asians included Chinese, Japanese, Koreans, Filipinos, Vietnamese, Thai, and Cambodians. South Asians consisted of Bangladeshi, Indians, Pakistani, and Sri Lankans. The Others category included individuals who reported belonging to 2 or more ethnocultural groups not included in the same category, as well as Aboriginal Canadians, Africans or Afro-Caribbean.

**Statistical Analysis**

Statistical Analysis Software v.9.4 (SAS Institute Inc, Cary, NC) was used for all analyses. Subjects with missing information on ethnicity, genotype or biomarkers of interest as well as any subjects who broke their 12-hour fast before blood sample collection were excluded. A total of 1,626 subjects with complete data were included in the analyses. Genotypes for 9p21 were examined for Hardy-Weinberg equilibrium. Chi-square test was used to assess 9p21 risk allele frequency across different ethnocultural groups. The distributions of all continuous variables were tested for normality and were log-transformed as needed, however, untransformed means and spreads were reported to facilitate interpretation of the data. The \( \alpha \)-error was set at 0.05, and p-values presented are two-sided. Initially, all analyses were unadjusted, and then adjusted for several covariates. Only those variables that were statistically significant in most models or materially altered the outcomes were retained in the model. The variables in each model were also tested for multicollinearity with tolerance level set at <0.4. No multicollinearity was detected amongst the variables selected for the final models. Using analysis of covariance (ANCOVA), differences in mean biomarker concentrations were examined across 9p21 genotypes, initially in the whole group. In subsequent analyses, data were stratified by four ethnocultural groups, sex as well as users and non-users of hormonal contraceptives (HC) in women. The final models included
adjustments for sex, ethnocultural group, serum glucose, diastolic blood pressure, log body mass index, log waist circumference, and hormonal contraceptive use by women, unless the outcome measure was stratified by one of these variables (such as sex, hormonal contraceptives or ethnocultural groups), in which case it was not included in the adjustment for that model. Post-hoc pair-wise differences were analyzed using Tukey’s test. The Benjamini-Yekutieli (B-Y) procedure \[=\alpha/\sum (1/i), \text{ where } i \text{ varies from one to the total number of tests conducted}\] was applied to adjust for multiple testing (adjusted \(p < 0.02\), calculated for testing the primary hypothesis of the study with six biomarkers of CVD as separate predictors with \(\alpha <0.05\) for each test). The B-Y method was selected because it allows for potential dependence between tests\(^{153}\), and many of the markers used here are biologically related. Adjustments were not made for analyses of five 9p21 SNPs as they were in linkage disequilibrium with each other.
4.4 Results

Subject characteristics are shown in Table 4.1. All five SNPs were in LD (ie >0.8). When the analyses where stratified by sex and four self-identified ethnocultural groups, risk allele frequencies of the five SNPs in each group were similar to those reported elsewhere and ranged from 46-55%. The only exception to this was in the smallest and most heterogeneous groups of Others, that were from a mixed ethnocultural origin. In the Others group, the frequency of the risk alleles for rs10757274, rs10757278 and rs1333049 was lower (34-36%). East Asians also had a slightly lower frequency of the risk allele (G) in rs2383206 (46%), compared to the whole group, however, this prevalence was within the range of what has been reported in populations elsewhere

Table 4.1: Subject Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Women (n = 1109)</th>
<th>Men (n = 517)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>22.5 ± 2.4</td>
<td>22.8 ± 2.4</td>
<td>0.08</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>163.2 ± 6.5</td>
<td>175.9 ± 7.2</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>59.7 ± 10.7</td>
<td>73.1 ± 12.7</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Body mass index (kg/m²)</td>
<td>22.4 ± 3.5</td>
<td>23.6 ± 3.5</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Waist circumference* (cm)</td>
<td>71.3 ± 7.6</td>
<td>80.1 ± 9.1</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Systolic Blood Pressure (mmHg)</td>
<td>109 ± 0</td>
<td>123 ± 1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Diastolic Blood Pressure (mmHg)</td>
<td>68.4 ± 0.4</td>
<td>72.1 ± 0.8</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Ethnocultural group (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>48</td>
<td>48</td>
<td>0.84</td>
</tr>
<tr>
<td>East Asian</td>
<td>36</td>
<td>31</td>
<td>0.14</td>
</tr>
<tr>
<td>South Asian</td>
<td>9</td>
<td>14</td>
<td>0.005</td>
</tr>
<tr>
<td>Other</td>
<td>8</td>
<td>7</td>
<td>0.87</td>
</tr>
</tbody>
</table>

Values are mean ± standard error, p-values are for comparison between three genotypes using unadjusted linear regression model. Chi-square test was used to test for differences between genotypes in categorical variables. *Variables were log-transformed to normalize distribution before use in regression
Traditional biomarkers of CVD were not significantly different when compared between genotypes, except for fasting insulin (Table 4.2). Mean fasting insulin for the group was 47.4 ± 0.9 pmol/L. Mean fasting insulin was 8% higher in those who are heterozygous for the risk allele (AG) in rs10757278 and 16% higher for those who are homozygous (GG) (p = 0.01).

**Table 4.2: Biomarkers of CVD Risk by Rs10757278 Genotype and Ethnocultural Group**

<table>
<thead>
<tr>
<th>Group</th>
<th>CVD Biomarker</th>
<th>Genotype</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Glucose (mmol/L)</td>
<td>AA</td>
<td>AG</td>
</tr>
<tr>
<td>All (n= 1,626)</td>
<td>4.77 ± 0.02</td>
<td>4.79 ± 0.01</td>
<td>4.79 ± 0.02</td>
</tr>
<tr>
<td></td>
<td>Insulin* (pmol/L)</td>
<td>44.0 ± 1.6</td>
<td>47.7 ± 1.3</td>
</tr>
<tr>
<td></td>
<td>CRP* (mmol/L)</td>
<td>1.35 ± 0.14</td>
<td>1.21 ± 0.08</td>
</tr>
<tr>
<td></td>
<td>TG* (mmol/L)</td>
<td>0.98 ± 0.14</td>
<td>0.98 ± 0.08</td>
</tr>
<tr>
<td></td>
<td>HDL* (mmol/L)</td>
<td>1.55 ± 0.02</td>
<td>1.54 ± 0.01</td>
</tr>
<tr>
<td></td>
<td>LDL (mmol/L)</td>
<td>2.27 ± 0.03</td>
<td>2.30 ± 0.02</td>
</tr>
<tr>
<td>Caucasians (n= 771)</td>
<td>Glucose (mmol/L)</td>
<td>4.74 ± 0.02</td>
<td>4.77 ± 0.03</td>
</tr>
<tr>
<td></td>
<td>Insulin* (pmol/L)</td>
<td>40.3 ± 1.5</td>
<td>45.9 ± 2.0</td>
</tr>
<tr>
<td></td>
<td>CRP* (mmol/L)</td>
<td>1.45 ± 0.19</td>
<td>1.44 ± 0.13</td>
</tr>
<tr>
<td></td>
<td>TG* (mmol/L)</td>
<td>1.00 ± 0.05</td>
<td>0.97 ± 0.03</td>
</tr>
<tr>
<td></td>
<td>HDL* (mmol/L)</td>
<td>1.56 ± 0.04</td>
<td>1.55 ± 0.01</td>
</tr>
<tr>
<td></td>
<td>LDL (mmol/L)</td>
<td>2.30 ± 0.01</td>
<td>2.24 ± 0.05</td>
</tr>
<tr>
<td></td>
<td>East Asians (n=561)</td>
<td>South Asians (n=175)</td>
<td>Others (n=119)</td>
</tr>
<tr>
<td>-------------------------</td>
<td>---------------------</td>
<td>----------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Glucose (mmol/L)</td>
<td>4.77 ± 0.04</td>
<td>4.78 ± 0.06</td>
<td>4.83 ± 0.04</td>
</tr>
<tr>
<td></td>
<td>0.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insulin* (pmol/L)</td>
<td>39.6 ± 1.7</td>
<td>44.6 ± 1.6</td>
<td>47.5 ± 2.9</td>
</tr>
<tr>
<td></td>
<td>0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRP* (mmol/L)</td>
<td>0.99 ± 0.22</td>
<td>0.69 ± 0.12</td>
<td>0.81 ± 0.17</td>
</tr>
<tr>
<td></td>
<td>0.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TG* (mmol/L)</td>
<td>0.97 ± 0.08</td>
<td>0.93 ± 0.05</td>
<td>1.05 ± 0.08</td>
</tr>
<tr>
<td></td>
<td>0.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HDL* (mmol/L)</td>
<td>1.53 ± 0.03</td>
<td>1.60 ± 0.03</td>
<td>1.57 ± 0.08</td>
</tr>
<tr>
<td></td>
<td>0.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LDL (mmol/L)</td>
<td>2.23 ± 0.05</td>
<td>2.29 ± 0.06</td>
<td>2.31 ± 0.08</td>
</tr>
<tr>
<td></td>
<td>0.17</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Women (n=1109)

<table>
<thead>
<tr>
<th></th>
<th>Glucose (mmol/L)</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4.70 ± 0.02</td>
<td>4.74 ± 0.01</td>
<td>4.71 ± 0.02</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>Insulin* (pmol/L)</td>
<td>43.4 ± 1.4</td>
<td>49.0 ± 1.6</td>
<td>55.7 ± 2.9</td>
<td>&lt;0.01</td>
<td></td>
</tr>
<tr>
<td>CRP* (mmol/L)</td>
<td>1.32 ± 0.15</td>
<td>1.35 ± 0.11</td>
<td>1.4 ± 0.18</td>
<td>0.91</td>
<td></td>
</tr>
<tr>
<td>TG* (mmol/L)</td>
<td>0.94 ± 0.02</td>
<td>0.94 ± 0.02</td>
<td>0.98 ± 0.05</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td>HDL* (mmol/L)</td>
<td>1.64 ± 0.02</td>
<td>1.65 ± 0.02</td>
<td>1.63 ± 0.02</td>
<td>0.94</td>
<td></td>
</tr>
<tr>
<td>LDL (mmol/L)</td>
<td>2.25 ± 0.03</td>
<td>2.27 ± 0.03</td>
<td>2.28 ± 0.04</td>
<td>0.81</td>
<td></td>
</tr>
</tbody>
</table>

### Men (n=517)

<table>
<thead>
<tr>
<th></th>
<th>Glucose (mmol/L)</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4.91 ± 0.03</td>
<td>4.92 ± 0.02</td>
<td>4.95 ± 0.04</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>Insulin* (pmol/L)</td>
<td>45.5 ± 3.9</td>
<td>44.8 ± 1.9</td>
<td>42.4 ± 2.5</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td>CRP* (mmol/L)</td>
<td>1.43 ± 0.34</td>
<td>0.9 ± 0.1</td>
<td>0.96 ± 0.17</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>TG* (mmol/L)</td>
<td>1.06 ± 0.07</td>
<td>1.05 ± 0.05</td>
<td>1.02 ± 0.05</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>HDL* (mmol/L)</td>
<td>1.34 ± 0.02</td>
<td>1.29 ± 0.02</td>
<td>1.31 ± 0.03</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>LDL (mmol/L)</td>
<td>2.29 ± 0.06</td>
<td>2.36 ± 0.04</td>
<td>2.24 ± 0.06</td>
<td>0.22</td>
<td></td>
</tr>
</tbody>
</table>

Values are mean ± standard error, p-values are for comparison between three genotypes using unadjusted linear regression model. *Variables were log-transformed to normalize distribution for model building. CRP, C-reactive protein; HDL, high-density lipoproteins; LDL, low density lipoproteins; TG, triglycerides.

Fasting insulin was then assessed across all five 9p21 SNPs and similar patterns of a stepwise increase in fasting insulin were seen with each additional risk allele for all five SNPs. However, only differences in genotypes of rs10757278 and rs2383206 were statically significant. Mean fasting insulin for Caucasians was 44.6 ± 1.3 pmol/L, for East Asians 43.9 ± 1.1 pmol/L, for South Asians 63.4 ± 4.0 pmol/L, and for the Others 57.0 ± 3.7 pmol/L. Among those in the Others group, the presence of two risk alleles in rs10757278 was associated with approximately 30 pmol/L higher mean insulin when compared to no risk alleles (pairwise comparison, p=0.031).
When the study population was stratified by sex, the association between 9p21 genotype and fasting insulin was significant for women with all five SNPs (p-value between 0.008 to 0.0003) but not for men. Mean insulin for men was 44.4 ± 1.5 and for women 48.8 ± 1.1. Fasting serum insulin remained relatively constant across genotypes in men, but there was a 13% increase with one risk allele and a 28% increase with two risk alleles in women. Furthermore, there was a significant sex-genotype interaction for all five SNPs analyzed (Table 4.3).

Table 4.3: Interaction of 9p21 SNPs with Sex and Ethnocultural Groups

<table>
<thead>
<tr>
<th>9p21 SNP</th>
<th>P</th>
<th>p</th>
<th>p-interaction</th>
<th>p-interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Ethnicity</td>
<td>Sex</td>
<td>Ethnicity x gene</td>
</tr>
<tr>
<td>rs10757274</td>
<td>0.16</td>
<td>&lt;0.001</td>
<td>0.89</td>
<td>0.009</td>
</tr>
<tr>
<td>rs10757278</td>
<td>0.11</td>
<td>&lt;0.001</td>
<td>0.92</td>
<td>0.004</td>
</tr>
<tr>
<td>rs2383206</td>
<td>0.06</td>
<td>&lt;0.001</td>
<td>0.82</td>
<td>0.004</td>
</tr>
<tr>
<td>rs1333049</td>
<td>0.18</td>
<td>&lt;0.001</td>
<td>0.97</td>
<td>0.012</td>
</tr>
<tr>
<td>rs4977574</td>
<td>0.43</td>
<td>&lt;0.001</td>
<td>0.68</td>
<td>0.019</td>
</tr>
</tbody>
</table>

*p-values are for linear regression models adjusted for: ethnocultural group, sex, hormonal contraceptives in women, log-body mass index, diastolic blood pressure, log-waist circumference and plasma glucose*

Despite variation in mean fasting insulin between ethnocultural groups among women, increasing fasting insulin was consistently associated with the presence of risk alleles in a stepwise manner across all groups. These associations were significance for East Asians and South Asians, but not for Caucasians and Others (Figure 4.1).
Figure 4.1: Fasting Serum Insulin by 9p21 Genotype (rs10757278) in Women (A) and Men (B)

ANCOVA was adjusted for: log-body mass index, diastolic blood pressure, log-waist circumference and plasma glucose. Fasting Insulin (pmol/L) was log-transformed, but expressed as normal mean to facilitate interpretation. Different letters, a, b and c indicate statistically significant differences in insulin concentrations between genotypes (pairwise differences) within a group which were tested only if overall ANCOVA for the group was statistically significant. p-value at the top of each group is indicative of overall differences in each group for comparing all three genotypes within the group. CA; Caucasians, EA; East Asians, SA; South Asians.
Among the women, 30% were on HCs and had higher mean fasting insulin levels than those not taking HCs (50.0 pmol/L ± 1.5 versus 48.3 pmol/L ± 1.4, p= 0.01, Table 4.4).

Table 4.4: Fasting Insulin Associated with rs10757278 according to HC use in Women in Each Ethnocultural Group

<table>
<thead>
<tr>
<th>Genotype</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td></td>
</tr>
<tr>
<td>AG</td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td></td>
</tr>
</tbody>
</table>

All Women

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No HC (n=784)</td>
<td>41.7 ± 1.7</td>
<td>48.5 ± 2.1</td>
<td>56.5 ± 3.9</td>
</tr>
<tr>
<td>On HC (n=325)</td>
<td>47.5 ± 2.6</td>
<td>50.1 ± 2.1</td>
<td>53.6 ± 3.9</td>
</tr>
</tbody>
</table>

Caucasians

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No HC (n=295)</td>
<td>37.7 ± 2.9</td>
<td>50.2 ± 4.3</td>
<td>53.2 ± 7.6</td>
</tr>
<tr>
<td>On HC (n=223)</td>
<td>46.2 ± 2.9</td>
<td>49.1 ± 2.5</td>
<td>51.1 ± 4.5</td>
</tr>
</tbody>
</table>

East Asians

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No HC (n=346)</td>
<td>37.4 ± 2.0</td>
<td>44.2 ± 2.2</td>
<td>47.8 ± 3.8</td>
</tr>
<tr>
<td>On HC (n=52)</td>
<td>48.7 ± 6.1</td>
<td>50.2 ± 5.3</td>
<td>61.8 ± 10.2</td>
</tr>
</tbody>
</table>

South Asians

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No HC (n=87)</td>
<td>53.8 ± 5.6</td>
<td>60.4 ± 7.1</td>
<td>77.1 ± 9.2</td>
</tr>
<tr>
<td>On HC (n=18)</td>
<td>69.0 ± 15.9</td>
<td>54.4 ± 10.4</td>
<td>46.5 ± 21.9</td>
</tr>
</tbody>
</table>

Others

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No HC (n=54)</td>
<td>40.3 ± 5.3</td>
<td>55.5 ± 6.3</td>
<td>60.3 ± 11.5</td>
</tr>
<tr>
<td>On HC (n=28)</td>
<td>58.8 ± 6.9</td>
<td>54.6 ± 7.3</td>
<td>85.5 ± 33.2</td>
</tr>
</tbody>
</table>

ANCOVA was adjusted for: log-body mass index, diastolic blood pressure, log-waist circumference and plasma glucose. Fasting Insulin was log-transformed to normalize distribution, but expressed as normal mean to facilitate interpretation. HC, hormonal contraceptives.

When women were stratified according to HC use, the association between 9p21 genotype and fasting insulin remained statistically significant for some ethnocultural groups, but no associations were observed among women taking HCs for any ethnocultural group (Table 4.4). This was most striking in the Caucasians, the largest group, where mean fasting insulin in women who were not
taking HCs was 1.5 times higher in those who were homozygous for the G allele compared to those who were homozygous for the A allele (p=0.0002, Table 4.4). In the Caucasian HC users, on the other hand, no differences in fasting insulin means were observed when comparing the same genotypes (p=0.63, Table 4.4).

In East Asians, similar findings were observed as those seen in Caucasians, where those taking HCs had a mean fasting insulin of 53.0 ± 4.1 pmol/L versus 43.1± 1.5 pmol/L for those not taking HCs (p=0.007 for pairwise comparison). In this group of East Asian women, fasting insulin was only associated with genotype in those not using HCs (p=0.02, Table 4.4). South Asian women and women from the Others group had a much higher mean concentration of fasting insulin (63.0 ± 4.1 pmol/L and 57.2 ± 4.1 pmol/L, respectively) even though the pattern of association was similar to all other ethnocultural groups (p=0.018 and p=0.05, Figure 4.1). The association between 9p21 risk alleles and fasting insulin was the same in women not on HC from all ethnocultural groups, despite significantly different mean fasting insulin in each ethnocultural group. Among the women who were not taking HCs, serum insulin was highest in South Asians (64.2 ± 4.5 pmol/L) compared to Caucasians (47.4 ± 2.8 pmol/L) and East Asians (43.1 ± 1.5 pmol/L) (p<0.001, all pairwise comparisons). Although the South Asian women not on HCs had the highest mean fasting insulin levels for all three genotypes, the mean differences among the genotypes were significantly different (p=0.02, Table 4.4).
4.5 Discussion

Variations in 9p21 are the most robust genetic markers of CVD. We sought to examine the associations between 9p21 genotype with early biomarkers of CVD risk in over 1,600 young-adults. Other studies on 9p21 are almost exclusively conducted in older adults\textsuperscript{108}, where CVD risk confounders are likely present in addition to genetic risk of 9p21. Therefore, despite this region being the subject of many studies over the past 15 years, the biological mechanisms by which it is linked to CVD development are still largely unknown. To our knowledge, this study is the first to examine 9p21 genotypes and biomarkers of CVD risk in young adults in their early 20’s from different ethnocultural groups.

We found that risk alleles of five most commonly studied SNPs in the 9p21 region were associated with higher levels of fasting serum insulin in this cohort, with rs10757278 showing the strongest association. This suggests that pathways involving insulin might be involved in mechanisms by which the 9p21 region is linked with CVD. Higher fasting insulin levels are associated with numerous adverse risk factors in young adults, increasing the risk of atherosclerosis\textsuperscript{68} and subsequent type 2 diabetes\textsuperscript{69}. Even in individuals without diabetes, hyperinsulinemia is associated with decreases in insulin-mediated glucose uptake\textsuperscript{70}, as well as with a number of clinical symptoms associated with insulin resistance\textsuperscript{71,72}. Prolonged and untreated insulin resistance and hyperinsulinemia are known to be associated with hypertriglyceridemia\textsuperscript{73} and low concentrations of high-density lipoprotein cholesterol\textsuperscript{73,74}, hypertension\textsuperscript{75}, and coronary artery calcification\textsuperscript{76}, which are all risk factors for CVD. Hyperinsulinemia has also independently been associated with ischemic heart disease\textsuperscript{72}. 
The synergism of risk alleles in the 9p21 locus and hyperglycemia have also been demonstrated in coronary artery disease (CAD) development and accelerated CVD mortality in those with type 2 diabetes \(^{154}\); where those homozygous for the risk alleles had worse glycemic control and higher incidents of CAD than others. Although 9p21 does not contain annotated genes, rs10757278 is in high linkage disequilibrium with two known genes CDKN2A and CDKN2B. These genes code for three known proteins p15\(^{\text{INK4b}}\), p16\(^{\text{INK4b}}\) and ARF, that have been linked to the presence of hereditary atherosclerosis \(^{155,156}\) where they inhibit cyclin-dependent kinases controlling cell proliferation and apoptosis in endothelial tissue of vascular smooth muscle, among other tissue.

One study found that antisense noncoding RNA in the INK4 locus (ANRIL) expression were significantly higher in carriers of the 9p21 risk alleles, with expression directly correlating with atherosclerotic severity \(^{157}\). The findings in our study suggest that insulin-dependent pathways might be one of the earliest signs of metabolic dysregulation, linking 9p21 risk alleles to CVD risk factors, evident in early adulthood.

We observed a significant genotype-sex interaction on fasting insulin, which has not been previously reported. The associations were observed in women, but not in men. Similarly, in the ADVANCE study, where genotypic odds ratio (OR) was stratified by sex, a trend towards higher ORs in women rather than men was observed \(^{110}\). The authors however concluded that this was likely due to admix of genotype present in the population, since the risk allele frequency differed between men and women in that cohort \(^{110}\). More recently, GG genotype of rs10757278 was also shown to be significantly and independently associated with carotid plaque in women only \(^{158}\).

Insulin resistance, hormonal glucose intolerance and type 2 diabetes increase the risk of CVD and mortality, however women are more affected as it increases the risk of CVD death by about four times in women verses two-fold in men \(^{159}\). In other studies, women have been identified as being
distinctly different than men with regard to insulin action, susceptibility to develop insulin resistance, and response to stimuli that are known to enhance or impair sensitivity to the effects of insulin\textsuperscript{160,161}. In the present study, we report a sex-specific genetic association between 9p21 and fasting insulin that might contribute to our understanding of how 9p21 impacts the development of CVD.

We also demonstrated varying magnitudes of these associations between the four ethnocultural groups of women (Figure 4.1). In the present study, South Asian and East Asian women had the largest differences in their mean fasting insulin between genotypes, and the direction of associations was consistent amongst all groups. These findings of differences in ethnocultural groups are in agreement with others where South Asians have been shown to have an increased risk of developing CVD compared with Caucasian populations\textsuperscript{162}, while East Asians have a lower incidence of CVD compared to Caucasians\textsuperscript{163}. Furthermore, the high prevalence of insulin resistance and type 2 diabetes in South Asians has been attributed as a major cause for their elevated CVD risk at a younger age\textsuperscript{162,164} compared to other ethnocultural groups. Similarly, in the present study, the mean insulin for South Asians was significantly higher than mean serum insulin in all other ethnocultural groups, suggesting additional factors might contribute to the observed higher insulin levels this population. A novel aspect of the present study is the young age of the cohort (22.7 ± 2.4 years), which highlights the possibility of very early onset of these ethnocultural differences in disease risk. Furthermore, genotype associations remained consistent in women, irrespective of ethnocultural group. In other studies, although variants at the 9p21 locus were previously associated with risk of myocardial infarction in South Asians, the strength of the associations were weaker in South Asians than in Europeans\textsuperscript{165}. However, this weaker association might be explained by the fact that the analyses of that study were not stratified by sex. In the
present study, similar patterns were observed when men and women from the South Asian group were analyzed together versus the marked differences observed when comparing men and women separately in the entire cohort.

Additional analyses in the present study demonstrated that when women were grouped according to HC use, those on HCs had significantly higher mean fasting insulin with no significant association between genotype and fasting insulin. In contrast, those who were not taking HCs had lower mean fasting insulin, which was significantly associated with 9p21 genotype (Table 4.4). HC use has been associated with insulin resistance, manifested by reduced peripheral tissue insulin sensitivity \(^{166}\), however this association has not been shown to be related to 9p21 genotype elsewhere. In the present study, Caucasian women formed the largest ethnocultural group, where the number of those on HC versus those without was similar (Table 4.4). When comparing mean fasting insulin according to genotype, a protective effect seems to be present in carriers of the A allele for rs10757278, whereas in those taking HCs this protection was not observed, despite having the same genotype and ethnocultural group.

Currently in clinical settings, biomarkers of CVD risk, including elevated CRP, and dyslipidemia determine intermediate phenotypic presentation of CVD and predict 10-year risk of CVD events. However, in the present study, similar to others \(^{108}\), we found no meaningful associations between other traditional CVD biomarkers including lipid profile, BMI, or hypertension suggesting that insulin-related associations observed might be either independent of those involved in cardiometabolic disease or the earliest dysregulations seen in the cascade of phenotypic events. Although the mechanisms by which 9p21 SNPs influence CVD risk are not well defined, a link between CVD genetic susceptibility and the response to inflammatory signaling has been reported.
No clear relationship between inflammatory markers and 9p21 genotype was observed in the present study, and this could be due to the young age of the study participants and the observational nature of the study design. However, others have reported long-distance interactions of CDKN2A/B, the MTAP gene, and interval downstream of IFNA21 human vascular endothelial cells. Interferon-γ (IFNγ) activation seems to affect the structure of chromatin, therefore, transcriptional regulation in the 9p21 locus, including STAT1-binding, likely plays a role as a long-range enhancer, altering expression of neighbouring genes. One study identified 33 enhancers in the 9p21 region, with rs10757278 located in one of these enhancers, where the presence of the risk allele (G), disrupted a binding site for STAT1. The link between this disturbance and insulin resistance was demonstrated by another study where IFNγ was shown to induce insulin resistance in mature human adipocytes. In the present study, we demonstrated that 9p21 risk alleles were associated with higher serum insulin levels, particularly with rs10757278. This finding may be explained by a possible mechanism whereby the A>G substitution in rs10757278 disrupts STAT1 binding site and modifies IFNA21 gene expression. IFNγ then possibly induces insulin insensitivity seen as higher insulin to be an early marker of endothelial dysfunction and increased susceptibility to CVD later in life. However, given the cross-sectional nature of the present study, causality cannot be determined from these results.

Early detection of CVD is necessary to improve health outcomes, however, most studies in this area have been in older populations. Currently, biomarkers of CVD risk are used as a tool for early intervention in clinical settings; however, these interventions are not personalized to individual risk factors such as genetic susceptibility, sex, ethnicity or environmental factors. Here we demonstrate that biomarkers of CVD risk that signify early metabolic disturbances related to CVD are present in early adulthood, and differ by genetic risk factors, sex and hormonal contraceptive
use in women. This approach can potentially have an added value to the traditional biomarkers used to treat or prevent CVD in high-risk populations. In one study, adding the 9p21 allele to traditional risk factors was associated with improved predictably of CVD hazard ratio of incident coronary heart disease of 1.2 per allele (p<0.000003) \(^{169}\). The currently study provides further support in better defining at risk groups that might have otherwise been missed using traditional markers only.
4.6 Conclusion

In summary, the present study is the first to demonstrate that common risk variants in the 9p21 region are associated with elevated fasting insulin in young adult women, but not in men, and not in women taking HCs. The role of 9p21 in insulin signaling and glucose metabolism in young women and it is possible links to CVD risk warrants further investigation.
Chapter 5

9p21 Genotype, Insulin Sensitivity and Dietary Patterns
5 9p21 Genotype, Insulin Sensitivity and Dietary Patterns

5.1 Abstract

**Background:** Single nucleotide polymorphisms (SNPs) in the 9p21 region have been associated with cardiovascular disease (CVD) and to a lesser extent insulin sensitivity. Previous studies have focused on older populations and few have examined the impact of gene-diet interactions. The objective of this study was to determine the interaction between dietary patterns and 9p21 genotypes on insulin sensitivity in young adults from different ethnocultural groups.

**Methods:** Subjects were 1,333 participants aged 20-29 years from the Toronto Nutrigenomics and Health Study (405 men and 928 women; 776 Caucasians and 557 East Asians). Fasting blood was collected to measure glucose, insulin, c-reactive protein and serum lipids, as well as to isolate DNA for genotyping subjects for five SNPs in 9p21 (rs10757274, rs10757278, rs1333049, rs2383206, and rs4977574). Insulin resistance (HOAM-IR) and beta-cell dysfunction (HOMA-Beta) were calculated from fasting insulin and glucose concentrations. The Toronto-modified Willett 196-item semi-quantitative food frequency questionnaire was used to measure dietary intake over one month and principal components analysis was used to identify three dietary patterns (Prudent, Western and Eastern). ANOVA and ANCOVA were used to examine gene-diet interactions on markers of insulin sensitivity.

**Results:** Among those who were homozygous for the 9p21 risk allele (rs1333049), fasting insulin was 40% higher in those who with a low-prudent dietary score compared to those with a high-prudent dietary score (p<0.05). No differences were observed between those with a low- versus high-prudent dietary score among those who did not carry a risk allele. Significant gene-diet
interactions on insulin sensitivity using HOMA-IR were observed with all five SNPs, which remained significant after adjusting for covariates (p<0.05). Similar findings were observed with HOMA-Beta, however, the association was only significant for rs10757274 (p=0.04).

**Conclusion:** Our findings suggest that a prudent dietary pattern may protect against the association of 9p21 risk genotypes with insulin resistance.
Introduction

Insulin resistance (IR) and glucose intolerance in young adults are risk factors for cardiovascular disease (CVD) in middle and late adulthood, even in those without diabetes. IR and endothelial dysfunction (ED) have some common biological mechanisms that stem from similar lifestyle and genetic factors. In addition to traditional pathways, IR has also been linked to atherosclerosis acting through pro-inflammatory pathways on vascular and immune cells. Genetic predisposition as well as dietary intake can modify insulin resistance, however, in most research and clinical settings, these factors are not considered together.

Single nucleotide polymorphism (SNPs) in the 9p21 region increase risk of CVD events, with a population attributable risk of 20-30% for myocardial infarction. SNPs in 9p21 have also been associated with type 2 diabetes (DM2) and IR, suggesting a possible common genetic predisposition. The mechanisms by which 9p21 risk variants impact the pathogenesis of CVD and DM2 remain unknown since the region does not contain annotated genes. However, alteration in response to inflammatory signaling in human vascular endothelial cells has been observed in those with risk alleles.

One of the most important environmental exposures that modify risk of IR and CVD, is dietary intake. Some dietary patterns such as the Mediterranean diet, portfolio diet and prudent dietary patterns, have been shown to significantly reduce CVD risk independent of other lifestyle factors. However, few studies have reported on the interaction between diet and genetic risks, such as those in 9p21. In one study that analyzed gene-diet interactions in several distinct populations, some food groups were identified to modify CVD risk related to 9p21 genotype. This interaction was observed in risk allele carriers but no effect was observed among
those without the risk alleles\textsuperscript{16}. However, the study was conducted in older adults with CVD events and matched controls. To date, no studies have reported on 9p21 gene-diet interactions in younger adults. Given the long lag times between exposure to risks and CVD outcomes, emphasis on preventive approaches to modify CVD at an earlier age may be more advantageous than later in life\textsuperscript{9,10,22}. Since CVD risks take several decades to develop into CVD events, studying gene-diet interactions in younger adults might be more advantageous in identifying biological mechanisms of CVD risk with fewer confounders. We previously observed an association between 9p21 genotype and fasting insulin. The current study aims to extend those findings to determine if 9p21 risk variants interact with dietary patterns to impact insulin resistance in a population of young adults from different ethnocultural groups.
5.3 Methods

Study Population

Participants were from the Toronto Nutrigenomics and Health Study (TNHS), which has been described elsewhere\textsuperscript{97}. In brief, the TNHS is a cross-sectional study that aims to explore the link between diet, genes and biomarkers of chronic disease and was approved by the University of Toronto Research Ethics Board. Study participants were aged 20-29 years from various ethnocultural backgrounds and were recruited from the University of Toronto campus between 2004 and 2010. Anthropometric measurements including height, weight, waist circumference, and blood pressure were recorded according to standard procedures\textsuperscript{97}. Subjects provided a fasting blood sample for DNA isolation and plasma was separated for measuring biomarkers of CVD risk including blood lipids, inflammatory markers and insulin. The current study included 1,333 subjects from the two largest ethnocultural groups (Caucasian and East Asian) who had complete dietary and genetic data available for analyses.

Clinical Characteristics

Anthropometric measurements including height, weight, waist circumference, and blood pressure were measured using standard procedures described elsewhere\textsuperscript{97}. Subjects also answered a general health and lifestyle questionnaire, indicating any medication use, including contraceptive use\textsuperscript{14}. Physical activity was assessed using a validated physical activity questionnaire and expressed as metabolic equivalent (MET) hours per week\textsuperscript{177}. Subjects provided a 12-hour overnight fasting blood sample for DNA isolation and plasma was separated for measuring biomarkers of interest such as glycemic measures, lipids, inflammatory markers and insulin described in more detail.
Homeostatic measure of insulin resistance (HOMA-IR) and Beta cell function (HOMA-Beta) were calculated using validated mathematical formulae; HOMA-IR=(insulin X glucose)/22.5 and HOMA-Beta=(20 X insulin)/(glucose - 3.5) that are described elsewhere 179.

Dietary Analyses

Dietary intake was assessed using the Toronto-modified Willett 196-item semi-quantitative food frequency questionnaire (FFQ) described elsewhere 178. In brief, each subject was given instructions on how to complete the FFQ by using visual aids of portion sizes to improve the measurement of self-reported food intake. This estimated the quantity and frequency of the food items and supplements consumed over one month. Subject responses to the individual foods were converted into daily number of servings for each item 136 and supplement use. Dietary patterns are better at predicting overall disease risk than individual nutrients or foods 132, therefore, dietary patterns were assessed using principal components analysis (PCA). Individual food items in the FFQ were used as the basis of the PCA, and to better describe the pattern structure of clusters, explained in detail elsewhere 136.

The formation of dietary patterns for this cohort have been described previously 136. In brief, three main dietary patterns were identified as “Prudent” “Western,” and “Eastern” patterns. Prudent pattern consisted of fruit, vegetables, nuts, lentils, beans, whole grains, and water. Eastern pattern consisted of vegetables, seafood, rice, and organ meat. Western pattern consisted of processed foods, high-saturated fat, salty and sugary foods, refined grain products, and sugary beverages. Each of the three dietary patterns had a composite score indicating habitual dietary adherence of the subjects to the diets over a one-month period. These scores were converted into three categories for each pattern where “low” indicated a score in the 25th or lower percentile of group
score distribution, “medium” within 26th-74th percentile, and “high” indicating 75th or higher percentile for each of the three dietary patterns.

**Genotyping**

DNA was analyzed for the following 5 SNPs in 9p21: rs133304, rs10757278, rs2383206, rs10757274, and 4977574 that have been shown most consistently to be associated with CVD risk elsewhere. Genotyping of the SNPs in 9p21 was completed at the Clinical Genomics Centre in the Princess Margaret Hospital, University Health Network, using the iPLEX Gold assay with mass spectrometry-based detection (Sequenom MassARRAY platform; Sequenom Inc) for all subjects. Since the presence of two risk alleles in the 9p21 region have been associated with the highest risk of CVD, subjects were grouped in two groups according to having two copies of the risk alleles versus having one or more ancestral allele. For examples, rs1333049 ancestral allele is G hence low CVD risk carriers were those with GG+CG genotype versus high CVD risk carriers with CC genotype. For the other four SNPs, a G>A substitution was indication of the risk allele, therefore, the AA+AG carriers were considered at lower CVD risk, versus those with the GG genotype that were considered higher CVD risk.

**Statistical Analysis**

Statistical Analysis Software v.9.4 (SAS Institute Inc, Cary, NC) was used for all analyses. Subjects with missing information on ethnicity, genotype or biomarkers of interest as well as any subjects who did not fast for at least 12 hours before blood sample collection were excluded. A total of 1,333 subjects were included in the initial analyses, however, an additional 24 subjects did not have complete genetic information for all five SNPs. These subjects were only included in the analyses where the genetic information of interest was available.
Genotypes for 9p21 were examined for Hardy-Weinberg equilibrium. Chi-square test was used to assess 9p21 risk allele frequency in different ethnocultural groups. The distributions of all continuous variables were tested for normality and were log-transformed as needed, however, untransformed means and spreads were reported to facilitate interpretation of the data.

Principal component analysis of food intake scores was used to identify food consumption patterns. These methods have been described in detail elsewhere. Briefly, individual food items in the FFQ were used as the basis of the analysis, and the patterns were obtained through an orthogonal rotation with the varimax rotation function in SAS 9.4, using 0.25 as a loading criterion, with consideration parameters of eigenvalues >1, the scree test, and the qualitative interpretability of patterns. This provided a component structure with three independent patterns. Percentage of variance explained by each pattern was not part of the selection criteria. Pattern scores were standardized and normally distributed.

The α-error was set at 0.05, and p-values presented are two-sided. Initially, all analyses were unadjusted, and then adjusted for several covariates. Only those variables that were statistically significant in most models or materially altered the outcomes were retained in the model. The variables in each model were also tested for multicollinearity with tolerance level set at <0.4. No multicollinearity was detected amongst the variables selected for the final models.

Using analysis of covariance (ANCOVA), mean of subject characteristics (i.e., BMI, age etc.) was examined across 9p21 genotypes (ie high versus low risk groups). In the unadjusted model, multiple linear regression was used to examine if 9p21 genotype, dietary patterns and gene-diet interactions were associated with insulin sensitivity and other biomarkers of CVD among the
different 9p21 SNPs. In the final models, analyses were adjusted for age, sex, hormonal contraceptive use, physical activity and log of body mass index.

For all analyses, reported p-values were 2-sided with statistical significance set at less than 0.05. Since outcome variables were selected based on an *a priori* hypothesis that genetic predisposition to CVD may be associated with altered levels of insulin sensitivity, analyses were not adjusted for multiple comparisons. Furthermore, since the markers of insulin sensitivity assessed are correlated and all five SNPs are in high LD, accounting for multiple testing by treating these as independent tests would lead to an inflated type II error rate.
5.4 Results

Subject characteristics are summarized in Table 5.1 according to rs1333049 where those who were homozygous for the risk alleles C (increased risk of CVD) versus carriers of one or two copies of the ancestral allele (lower risk of CVD). Both groups had a similar ratio of Caucasians to East Asians ethnocultural diversity as well as distribution of men and women. Both groups had a similar ratio of Caucasians to East Asians as well as distribution of men and women. Both groups had similar adherence to Prudent, Western and Eastern style dietary patterns, according to the PCA generated scores. Non-risk allele frequency distribution was also similar among Caucasians and East Asians as well as in men and women.

Table 5.1: Subject Characteristics According to 9p21 Genotype (rs1333049)

<table>
<thead>
<tr>
<th></th>
<th>GG + CG (n= 1035)</th>
<th>CC (n= 298)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>22.7 ± 0.1</td>
<td>22.6 ±0.1</td>
<td>0.19</td>
</tr>
<tr>
<td>Body mass index (kg/m²)</td>
<td>22.8 ± 0.1</td>
<td>22.5 ± 0.2</td>
<td>0.11</td>
</tr>
<tr>
<td>Waist circumference* (cm)</td>
<td>73.9 ± 0.3</td>
<td>73.9 ± 0.5</td>
<td>0.93</td>
</tr>
<tr>
<td>Systolic Blood Pressure (mmHg)</td>
<td>113.9 ± 0.4</td>
<td>114.1 ± 0.7</td>
<td>0.83</td>
</tr>
<tr>
<td>Diastolic Blood Pressure (mmHg)</td>
<td>69.2 ± 0.2</td>
<td>69.6 ± 0.5</td>
<td>0.40</td>
</tr>
<tr>
<td>Prudent Diet Score</td>
<td>2.98 ± 0.05</td>
<td>2.94 ± 0.09</td>
<td>0.38</td>
</tr>
<tr>
<td>Eastern Diet Score</td>
<td>1.23 ± 0.03</td>
<td>1.22 ± 0.05</td>
<td>0.27</td>
</tr>
<tr>
<td>Western Diet Score</td>
<td>1.15 ± 0.03</td>
<td>1.12 ± 0.04</td>
<td>0.52</td>
</tr>
<tr>
<td>Females (%)</td>
<td>70</td>
<td>66</td>
<td>0.31</td>
</tr>
<tr>
<td>Males (%)</td>
<td>30</td>
<td>34</td>
<td>0.22</td>
</tr>
<tr>
<td>Caucasian (%)</td>
<td>57</td>
<td>59</td>
<td>0.59</td>
</tr>
<tr>
<td>East Asian (%)</td>
<td>43</td>
<td>41</td>
<td>0.58</td>
</tr>
</tbody>
</table>

Values are mean ± standard error, p-values are for comparison between two genotypes using unadjusted linear regression model. Chi-square test was used to test for differences between genotypes in categorical variables. *Variables were log-transformed to normalize distribution before use in regression.
All five SNPs were in high linkage disequilibrium (>80%). When traditional biomarkers of CVD risk were analyzed according to rs1333049 genotype for the group, fasting insulin was the only outcome that was significantly different between the high-risk versus the low risk groups (p=0.04). No other associations were observed in the subgroups or with any other CVD risk biomarkers, including CRP (Table 5.2).

Table 5.2: Biomarkers of CVD Risk by 9p21 Genotype Risk (rs1333049) in Subgroups

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>CVD Biomarker</th>
<th>Genotype</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>GG +GC</td>
<td>CC</td>
</tr>
<tr>
<td>All (n=1333)</td>
<td>Glucose (mmol/L)</td>
<td>4.77 ± 0.01</td>
<td>4.77 ± 0.02</td>
</tr>
<tr>
<td></td>
<td>Insulin* (pmol/L)</td>
<td>43.7 ± 0.9</td>
<td>46.8 ± 2.1</td>
</tr>
<tr>
<td></td>
<td>CRP* (mmol/L)</td>
<td>1.17 ± 0.08</td>
<td>1.18 ± 0.14</td>
</tr>
<tr>
<td></td>
<td>TG* (mmol/L)</td>
<td>0.97 ± 0.02</td>
<td>0.98 ± 0.03</td>
</tr>
<tr>
<td></td>
<td>HDL* (mmol/L)</td>
<td>1.56 ± 0.01</td>
<td>1.55 ± 0.02</td>
</tr>
<tr>
<td></td>
<td>LDL (mmol/L)</td>
<td>2.27 ± 0.02</td>
<td>2.24 ± 0.04</td>
</tr>
<tr>
<td>Caucasians (n= 773)</td>
<td>Glucose (mmol/L)</td>
<td>4.76 ± 0.01</td>
<td>4.74 ± 0.03</td>
</tr>
<tr>
<td></td>
<td>Insulin* (pmol/L)</td>
<td>44.1 ± 1.5</td>
<td>46.7 ± 3.0</td>
</tr>
<tr>
<td></td>
<td>CRP* (mmol/L)</td>
<td>1.45 ± 0.11</td>
<td>1.42 ± 0.18</td>
</tr>
<tr>
<td></td>
<td>TG* (mmol/L)</td>
<td>0.98 ± 0.02</td>
<td>0.98 ± 0.05</td>
</tr>
<tr>
<td></td>
<td>HDL* (mmol/L)</td>
<td>1.55 ± 0.02</td>
<td>1.54 ± 0.03</td>
</tr>
<tr>
<td></td>
<td>LDL (mmol/L)</td>
<td>2.27 ± 0.03</td>
<td>2.19 ± 0.05</td>
</tr>
<tr>
<td></td>
<td>East Asians (n= 560)</td>
<td>Females (n=920)</td>
<td>Males (n=413)</td>
</tr>
<tr>
<td>----------------------</td>
<td>-----------------------------</td>
<td>--------------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>Glucose (mmol/L)</td>
<td>4.78 ± 0.02</td>
<td>4.71 ± 0.02</td>
<td>4.9 ± 0.04</td>
</tr>
<tr>
<td>Insulin* (pmol/L)</td>
<td>43.1 ± 1.2</td>
<td>50.8 ± 2.9</td>
<td>38.9 ± 2.4</td>
</tr>
<tr>
<td>CRP* (mmol/L)</td>
<td>0.79 ± 0.1</td>
<td>1.3 ± 0.18</td>
<td>0.98 ± 0.14</td>
</tr>
<tr>
<td>TG* (mmol/L)</td>
<td>0.96 ± 0.04</td>
<td>0.96 ± 0.04</td>
<td>1.03 ± 0.04</td>
</tr>
<tr>
<td>HDL* (mmol/L)</td>
<td>1.57 ± 0.02</td>
<td>1.66 ± 0.03</td>
<td>1.33 ± 0.02</td>
</tr>
<tr>
<td>LDL (mmol/L)</td>
<td>2.27 ± 0.03</td>
<td>2.26 ± 0.04</td>
<td>2.3 ± 0.04</td>
</tr>
</tbody>
</table>

Values are mean ± standard error, p-values are for comparison between three genotypes using unadjusted linear regression model. *Variables were log-transformed to normalize distribution for model building. CRP, C-reactive protein; HDL, high-density lipoproteins; LDL, low density lipoproteins; TG, triglycerides.
**Gene-Diet Interactions**

Prudent dietary pattern was significantly associated with fasting insulin when the subjects where stratified according to genotype in all five SNPs (P<0.001), while Western and Eastern dietary patterns showed no significant associations (p>0.05 for all associations). In the high-risk group (homozygous for 9p21 risk alleles), fasting insulin was on average 1.4 times for those consuming a low-prudent diet compared to those consuming a high-prudent diet (p<0.05, Table 5.3).

**Table 5.3: Fasting Insulin by 9p21 Genotype, and Prudent Dietary Categories**

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Prudent Dietary Score</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>Medium</td>
</tr>
<tr>
<td>rs1333049</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GC + CG</td>
<td>44.5 ± 23.4</td>
<td>44.8 ± 36.4</td>
</tr>
<tr>
<td>CC</td>
<td>58.5 ± 47.0</td>
<td>42.3 ± 24.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs10757278</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA + AG</td>
<td>44.6 ± 23.4</td>
<td>44.6 ± 36.4</td>
</tr>
<tr>
<td>GG</td>
<td>58.8 ± 47.5</td>
<td>42.8 ± 24.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs2383206</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA + AG</td>
<td>44.4 ± 23.1</td>
<td>44.1 ± 36.6</td>
</tr>
<tr>
<td>GG</td>
<td>57.3 ± 45.3</td>
<td>44.5 ± 24.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs10757274</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA + AG</td>
<td>45.2 ± 28.1</td>
<td>44.4 ± 36.4</td>
</tr>
<tr>
<td>GG</td>
<td>56.7 ± 37.3</td>
<td>43.8 ± 24.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs4977574</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA + AG</td>
<td>45.0 ± 28.0</td>
<td>44.3 ± 36.4</td>
</tr>
<tr>
<td>GG</td>
<td>56.4 ± 37.5</td>
<td>43.7 ± 24.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Values are mean fasting insulin ± standard error listed in order of increasing prudent dietary score; p-values are for linear regression models with log- fasting insulin as the dependent variable and prudent dietary pattern and binary genotype as the main determinants of interest as well as diet X gene as the interaction term. All models were adjusted for: sex, hormonal contraceptives in females, physical activity, log-body mass index, log-waist circumference. In all SNPs an A>G substitution was noted as the risk allele, however for rs1333049 a G<C substitution signified risk allele as C, therefore the order for the genotype of this SNP is listed as GG, GC, CC; A, adenine; C; cytosine G, Guanine; SNP, single nucleotide polymorphism.
These dietary associations with fasting insulin were not observed in those carrying non-risk alleles. Significant gene-diet interactions on insulin sensitivity using HOMA-IR were observed with all five SNPs, which remained significant after adjusting for covariates (age, sex, hormonal contraceptive use in females, physical activity and log of body mass index) (p<0.05, Table 5.4).

Table 5.4: Interaction Between Dietary Components and 9p21 SNPs on HOMA-IR

<table>
<thead>
<tr>
<th>9p21 SNP</th>
<th>Dietary Patterns</th>
<th>Food Groups</th>
<th>Fresh Fruits</th>
<th>Raw Vegetables</th>
<th>Cooked Vegetables</th>
<th>Whole Grains</th>
<th>White Grains</th>
<th>Other Meat</th>
<th>Processed Meat</th>
<th>Total Dairy</th>
<th>Total Yogurt</th>
<th>Sweets</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1333049</td>
<td>0.031</td>
<td></td>
<td>0.239</td>
<td>0.716</td>
<td>0.890</td>
<td>0.239</td>
<td>0.263</td>
<td>0.258</td>
<td>0.796</td>
<td>0.765</td>
<td>0.068</td>
<td>0.396</td>
</tr>
<tr>
<td>rs2383206</td>
<td>0.007</td>
<td></td>
<td>0.141</td>
<td>0.999</td>
<td>0.476</td>
<td>0.109</td>
<td>0.172</td>
<td>0.371</td>
<td>0.484</td>
<td>0.701</td>
<td>0.145</td>
<td>0.550</td>
</tr>
<tr>
<td>rs1075727</td>
<td>0.009</td>
<td></td>
<td>0.253</td>
<td>0.999</td>
<td>0.759</td>
<td>0.234</td>
<td>0.275</td>
<td>0.244</td>
<td>0.799</td>
<td>0.573</td>
<td>0.081</td>
<td>0.396</td>
</tr>
<tr>
<td>rs10757274</td>
<td>0.035</td>
<td></td>
<td>0.397</td>
<td>0.958</td>
<td>0.410</td>
<td>0.025</td>
<td>0.340</td>
<td>0.279</td>
<td>0.680</td>
<td>0.822</td>
<td>0.143</td>
<td>0.254</td>
</tr>
<tr>
<td>rs4977574</td>
<td>0.038</td>
<td></td>
<td>0.279</td>
<td>0.968</td>
<td>0.354</td>
<td>0.022</td>
<td>0.319</td>
<td>0.667</td>
<td>0.740</td>
<td>0.844</td>
<td>0.198</td>
<td>0.264</td>
</tr>
</tbody>
</table>

Distribution of dietary components and SNPs were assessed for normality and log-transformed if significantly deviated from a normal distribution. Interaction tests were performed using general linear regression adjusted for main effects of SNP and dietary component. Significance was set at p-value of <0.05 and minimum sample size included for analyses was n=1309.

This association was most significant in rs2383206, where a Prudent dietary score was inversely associated with insulin resistance in the high-genetic risk group but there was no association with insulin resistance as measured by HOMA-IR in the low risk group (p=0.002, and p=0.08 respectively, Figure 5.1).
Figure 5.1: Homeostatic Model Assessment of Insulin Resistance (HOMA-IR) by Prudent Dietary Score Categories and 9p21 Genotype (Rs10757274)

Data are means ± SEM. Statistical significance was calculated using unadjusted ANOVA initially for the interaction term of gene-diet in all six groups (n=1309, p=0.031), and then separately for prudent dietary score in carriers of AA + AG genotype (n=985, p=0.076) and carriers of GG genotype separately (n=324, p=0.002). In GG group those with high (75th percentile) and medium (26th-74th percentile) prudent dietary score had a significantly lower HOMA-IR than those with low (25th percentile) prudent diet score (p=0.004 and p=0.014 respectively). In the AA + AG group HOMA-IR did not differ between medium and low nor high and low prudent dietary score groups (p=0.067 and p=0.386, respectively). Different letters (a versus b) indicate pairwise differences between group means that were statistically different while same letters indicate pairwise comparison were not statistically different.

Similar findings were observed when beta-cell dysfunction was estimated via HOMA-Beta (Figure 5.2), however, the association was only significant for one SNP rs1333049 (p=0.044) while the other four SNPs only trended towards significance (p= 0.06 to p=0.10). Adjusting for co-variables did not meaningfully change these associations with HOMA-Beta.
**Figure 5.2: Homeostatic Model Assessment of Beta Cell Dysfunction (HOMA-Beta) by Prudent Dietary Score Categories and 9p21 Genotype (Rs10757274).**

Data are means ± SEM. Statistical significance was calculated using unadjusted ANOVA initially for the interaction term of gene-diet in all six groups (n=1,309, p=0.044), and then separately for prudent dietary score in carriers of AA + AG genotype (n=985, p=0.360) and carriers of GG genotype separately (n=324, p=0.030). In GG group those with high (75th percentile) and medium (26th-74th percentile) prudent dietary score had a significantly lower HOMA-Beta than those with low (25th percentile) prudent diet score (p=0.013 and p=0.023 respectively). In the AA + AG group HOMA-Beta did not differ between medium and low nor high and low prudent dietary score groups (p=0.080 and p=0.235, respectively). Different letters (a versus b) indicate pairwise differences between group means that were statistically different while same letters indicate pairwise comparison were not statistically different.

In addition to dietary patterns, some foods groups were also analyzed for interactions with genotype. Raw vegetables significantly interacted with two 9p21 SNPs (rs10757272 and rs4977574) on HOMA-IR (p=0.025 and p=0.022 respectively, Table 5.4). Other food groups did not show any significant interactions with any of the SNPs on insulin related outcomes.
5.5 Discussion

Variations in 9p21 are the most robust genetic markers of CVD outcomes, however, few studies have assessed if gene-environment interactions modify these risks, particularly in young adults. Since environmental factors play an important role in determining CVD risk, study designs that analyze both modifiable and non-modifiable risk factors may be more informative than each approach alone. Therefore, we sought to assess the interactions between three common dietary patterns with the five most widely studied 9p21 SNPs on insulin sensitivity and beta-cell dysfunction outcomes in over 1,300 young-adults. To our knowledge, this study is the first to examine these associations in a cohort of young adults.

We found that the 9p21 high-risk genetic group had higher levels of fasting plasma insulin than the low-risk group (p=0.04, Table 5.2). In addition, HOMA-IR and HOMA-Beta were also elevated among those homozygous for the risk alleles, particularly in those consuming a low prudent diet (Figures 5.1 and 5.2). IR earlier in life is often predictive of CVD in middle and late adulthood, even in those without diabetes. In addition to traditional pathways, IR has also been linked to atherosclerosis through pro-inflammatory activity on vascular and immune cells. Mechanisms by which 9p21 risk alleles are associated with CVD are largely unknown, however, there is some evidence that endothelial dysfunction might be one of the pathways. In addition, same mechanisms identified that induces endothelial dysfunction such as interferon-γ activation, have also been shown to induce IR in human adipose tissue. Therefore, our findings on a macro scale, are in agreement with these cellular discoveries and might aid in further understanding of the role of 9p21 SNPs in CVD risk.
Mean HOMA-IR was 150% higher and HOMA-Beta was 30% higher in those with a low Prudent dietary pattern score versus those with a high Prudent dietary pattern from the same high-genetic risk group (p=0.004, p=0.013, respectively). Higher Prudent dietary pattern scores appeared protective against IR and beta-cell dysfunction in the genetically high-risk group (Figure 5.1 and Figure 5.2) and this protection was in a step-wise fashion with HOMA-IR. Meanwhile in the low-genetic risk group, Prudent dietary patterns were not associated with mean HOMA-IR nor HOMA-Beta (p=0.08 and p=0.36, respectively). Although associations of fasting insulin and 9p21 have not been previously reported in the literature, others have reported on 9p21 gene-diet interactions with related outcomes including glycemic control \(^{144,154}\), lipid profile \(^{181}\) and CVD events \(^{144,145}\). Our findings are in agreement with all previous findings and add to the evidence that CVD risk factors may be modifiable in young adults with high risk 9p21 SNPs. It is also important to note that these findings could explain previous inconsistencies linking one-size fits all dietary interventions \(^{182}\) and might justify a more personalized approach to dietary interventions in prevention of CVD.

Higher Prudent pattern scores in the current study had strong correlations with increased vitamin and mineral levels (vitamins A, B1, B2, B6, C, D, and K, magnesium, and iron) and this was likely due to the relatively large presence of fruits and vegetables in this dietary pattern \(^{136}\). Higher intake of fruits and vegetables have been shown to be protective towards CVD risks in several dietary patterns and populations \(^{183}\). Several mechanisms have been proposed by which these foods seem to be protective and include; via bioactive components with antioxidant, anti-inflammatory, and electrolyte properties, as well as functional properties, such as low glycemic load and low energy density \(^{183}\). In the current study, high fruits and vegetable intake was a distinguishing feature of both the Prudent dietary patterns and the Eastern dietary patterns, however, no associations were
found with the Eastern dietary patterns and 9p21 genetic risk on insulin sensitivity. One of the reasons for this lack of association with the Eastern pattern might have been that the majority of fruits and vegetables in this dietary pattern was cooked or processed whereas the fruits and vegetables in the Prudent dietary patterns were often freshly consumed with minimal processing. This association was evident when food groups were analyzed and consumption fresh vegetables were shown to interact with two 9p21 SNPs on HOMA-IR (p=0.022 and 0.025, Table 5.4). Our finding is in agreement with two other studies where fresh vegetables intake were protective against CVD in high-risk genotype in 9p21\textsuperscript{144,145}.

Higher Western dietary patterns in the current population were correlated with high sodium, fat, saturated fat and energy intake, likely because of the high proportion of processed foods included in this pattern. Interestingly, this dietary pattern was not associated with insulin markers and 9p21 genotype. This finding is also in agreement with another study that found while Prudent dietary patterns attenuate the 9p21 genetic risk, Western dietary patterns did not affect CVD events\textsuperscript{145}. This is contrary to other findings were Western dietary patterns have been linked to CVD and diabetes among many other chronic conditions\textsuperscript{184,185}. Although one large systematic review found strong evidence to support a causal relationship between Prudent/Mediterranean dietary factors and reduction in CVD, they concluded that there was insufficient evidence against components of a Western dietary pattern (including saturated fats) to support a causal relationship with increased CVD\textsuperscript{186}. In the same systematic review, however, there was evidence for trans-fats being a contributing factor towards CVD and Western diets are typically high in trans-fats\textsuperscript{186}. In the current study, however, our dietary patterns did not characterize trans-fats as major component of any of the three dietary patterns\textsuperscript{136}. This might have been a limitation of the FFQ in measuring
trans-fat content of the foods reported or that the study population did not have a high trans-fat dietary intake associated with the three dietary patterns.

5.6 Conclusion

In summary, our findings suggest that habitual consumption of a high prudent diet, one especially high in raw vegetables, might be associated with protection against the genetic risk of higher fasting insulin, insulin resistance and to a lesser degree, beta-cell dysfunction amongst young adults who are homozygous for the 9p21 risk alleles. Dietary intervention trials would be needed to conclude a causal relationship.
Chapter 6

9p21 Genotype and the Plasma Proteome

This chapter is adapted from an article submitted to the Journal of Proteome Research February 2018. The original article is the following:

6  9p21 Genotype and the Plasma Proteome

6.1  Abstract

Single nucleotide polymorphisms (SNPs) in the non-coding region of 9p21 have been associated with cardiovascular disease (CVD), but the mechanisms by which these genetic variants contribute to the pathogenesis of CVD remain unknown since no annotated proteins are present in this region of DNA. The objective of the current study was to determine if 9p21 genotypes are associated with distinct plasma proteins in young adults. Subjects were 1,611 young adults aged 20-29 years from the Toronto Nutrigenomics and Health Study (1,098 females and 513 males). DNA was isolated from fasting blood to analyze four SNPs in 9p21 (rs2383206, rs10757274, rs10757278 and rs1333049). High abundant plasma proteins (n=54) were measured using a LC/MRM-MS. ANCOVA was used to determine differences in plasma proteins between genotypes. In Caucasians (n=524), four SNPs were associated with Apolipoprotein E and two with Haptoglobin-β-Chain concentration. In East Asians (n=388), three SNPs were associated with α-1b-Glycoprotein, two with Apolipoprotein B-100, and one with Apolipoprotein E and Haptoglobin-β-Chain concentration. In South Asians (n=117), one SNP was associated with Apolipoprotein B-100 concentration. Our findings suggest that 9p21 genotypes may play a role in various pathophysiological pathways that contribute to risk of CVD in early adulthood that might be distinct amongst different ethnocultural groups.
6.2 Introduction

Cardiovascular disease (CVD) is a leading cause of death worldwide \(^1\) and its pathogenesis involves both environmental and genetic factors \(^2\text{–}^4\). Despite recent advances in CVD treatment, prediction of cardiovascular events still relies on traditional biomarkers such as dyslipidemia \(^54\). Risk algorithms, such as the Framingham risk score, based on these traditional biomarkers have poor predictive power. Indeed, one study showed that the Framingham algorithm correctly predicted only 11% of the CVD events in 10 years \(^56\). Many individuals develop CVD in the absence of conventional risk factors \(^58\), and as many as 80% of CVD cases are only diagnosed after a myocardial infarction (MI). The pathogenesis of CVD often starts decades before an ischemic event \(^27\), when clinical signs and symptoms may be absent according to traditional markers. Therefore, refining methods for early detection of CVD is needed to implement preventive measures that will reduce the global incidence of this disease.

Alterations in several physiological pathways, such as innate immunity and lipid metabolism, lead to inflammation, thrombosis, calcification, vascular remodeling, oxidative stress and cell death, which contribute to the development of CVD \(^30\text{,}32\text{,}77\text{–}79\). Additional pathways involved in the development of CVD have yet to be elucidated, and new biomarkers are needed to predict early dysregulation of both traditional and novel pathways and implement more successful CVD preventive strategies. Genome wide association studies (GWAS) focusing on CVD have identified single nucleotide polymorphisms (SNPs) in chromosome 9 arm p section 21 (9p21) as markers for CVD risk \(^8\text{,}107\text{,}108\text{,}105\). The 9p21 risk variants are independent CVD risk factors associated with MI, stroke and abdominal aortic aneurism \(^109\). Multiple case-control studies in several ethnicities have confirmed these findings \(^4\text{,}107\text{–}113\). However, despite the wealth of replication studies supporting a role for SNPs in 9p21 in CVD, the risk locus contains no protein coding sequences or known
microRNAs, so its functional effects remain unknown. Furthermore, 9p21 risk alleles appear to be independent of traditional risk factors, including elevated lipid levels, high blood pressure, obesity, and diabetes, suggesting alternative pathogenic mechanisms to those known to date.

Proteomic technologies increasingly permit the evaluation of systemic changes in protein expression in various tissues, and have been applied to biomarker discovery in CVD. Plasma levels of several novel protein markers of CVD appear to differ across ethnic groups. Furthermore, the populations under study have often been middle-aged or older, but preventing the development of CVD may require identifying biomarkers of early risk in young adults, well before the onset of disease-associated processes. A panel of 54 abundant plasma proteins has been shown to be associated with distinct characteristics including dietary patterns, hormonal contraceptive use, gluten intake, and vitamin C, D, and E status. This panel may also be useful in revealing specific protein patterns associated with 9p21 genotypes in a multiethnic group of young adults.

Examining whether genetic variation in 9p21 affects the plasma proteome could help identify novel pathways implicated in the development of CVD, and lead to the discovery of new risk biomarkers. The aim of the current study was to investigate the association between four SNPs in 9p21 and a panel of 54 proteins abundant in plasma in a multiethnic group of young adults, and to determine whether these associations differed across ethnic groups.

6.3 Methods

Participants were from the Toronto Nutrigenomics and Health Study (TNHS), which has been described elsewhere. In brief, the TNHS is a cross-sectional study that aims to explore the link
between diet, genes and biomarkers of chronic disease and was approved by the University of Toronto Research Ethics Board. Study participants (n=1,639, aged 20-29 years) from various ethnocultural backgrounds were recruited from the University of Toronto campus between 2004 and 2010. Ethnocultural status was determined by asking subjects in an open-ended format to self-report the ethnocultural group(s) they identified with. Subjects were then grouped into four categories based on their self-reported status as Caucasian, East Asian, South Asian, or Other. Caucasians included those who considered themselves European, Middle Eastern, or Hispanic. East Asians included Chinese, Japanese, Koreans, Filipinos, Vietnamese, Thai, and Cambodians. South Asians consisted of Bangladeshi, Indians, Pakistani, and Sri Lankans. Anthropometric measurements including height, weight, waist circumference, and blood pressure were also recorded.

**Genomic Analysis**

Subjects provided a fasting blood sample for DNA isolation. DNA was extracted from whole blood samples using standard procedures. DNA was analyzed for the four SNPs in 9p21 that have been shown most consistently to be associated with CVD risk \(^{142}\) (rs2383206, rs10757274, rs10757278 and rs1333049). Genotyping was completed using iPLEX Gold assay with mass spectrometry-based detection (Sequenom MassARRAY platform; Sequenom Inc).

**Proteomic Analysis**
Blood samples were collected between 0800 and 1030, after a minimum 12-h overnight fast, at LifeLabs Laboratories (Toronto, Canada). Approximately 44 mL of blood was drawn from each subject’s antecubital vein. Plasma was obtained from blood samples and frozen at -80°C. Frozen plasma samples were shipped to the University of Victoria–Genome British Columbia Proteomics Centre (Victoria, Canada), where concentrations of common plasma proteins were measured by using a liquid chromatography multiple reaction monitoring mass spectrometry (LC/MRM-MS) without enrichment or affinity depletion as described elsewhere. Briefly, 5 µL of each plasma sample was denatured with deoxycholate, reduced with TCEP, and digested with trypsin. Digestion was stopped by acidifying the sample, and a concentration-balanced stable-isotope labeled internal standard (SIS)-peptide mixture was added. Each sample was then desalted by SPE, eluted, lyophilized, and stored at -80°C. Shortly before analysis, samples were reconstituted with 0.1% formic acid to give a concentration of ~1 µg/µL of the original protein, and a 1-µL aliquot was injected onto an Eksigent NanoLC-1Dplus nano-scale HPLC system, equipped with a Magic C18AQ column operated at 300 nL/min, and interfaced to a Sciex QTRAP 4000 mass spectrometer equipped with Analyst 1.5 software, as described previously. The 65 transitions monitored, targeting 63 proteins, are shown in Supporting Information Table S1. Endogenous proteins were quantified against calibration curves made from the SIS standards (a “reverse-curve” approach) spiked into a standard plasma sample. The SIS peptide concentrations were balanced to match the natural abundance of the proteins they represent. Inter-assay reproducibility was determined by processing a standard plasma sample on 13 separate days by performing ten separate tryptic digests of a standard plasma sample on 10 separate days and analyzing them on three different days. Intra-assay reproducibility was assessed by calculating Coefficient of Variation (CV), measured by performing 12 LC/MRM-MS analyses on each SIS peptide. Of the
63 proteins measured, six were below the detection limit and three had inter-assay CVs >20% and were excluded from the statistical analyses. Of the 54 proteins quantified, 50 had CVs <10% and four had CVs of 10%-14% and were retained for further analyses in the TNHS. The concentrations for most of the 54 proteins included in the original assay was within the range of reported values, or within a factor of two from previously reported values. These methods have been validated in an intra-lab investigation.

**STATISTICAL ANALYSIS**

Statistical Analysis Software v.9.4 (SAS Institute Inc, Cary, NC) was used for all analyses. The α-error was set at 0.05, and p-values presented are two-sided. Variables that were not normally distributed were either loge- or square root–transformed before analysis to improve normality. p-values from models using the transformed protein concentrations are reported, but untransformed means and measures of spread are also reported to facilitate interpretation.

The 9p21 genotypes were examined for Hardy-Weinberg equilibrium, and chi-square test was used to test for differences in the prevalence of 9p21 gene variants across the different ethnocultural groups. Initial study population was 1639 however only 1090 had proteomic data available. Subjects with missing information on ethnicity (or anyone outside of the three self-identified groups), missing 9p21 genotype, or any subjects who broke their 12-hour fast before blood sample collection and three persons with diabetes were excluded. Complete data on 1,031 participants were included in the initial analyses. Using analysis of covariance (ANCOVA), differences in mean plasma protein concentrations were examined across 9p21 genotypes, stratified by the three ethnic groups: Caucasians (n=524), East Asians (n=388), South Asians (n=117). These analyses were adjusted for sex, hormonal contraceptive use in females, physical activity, log Body Mass
Index, and age. Post-hoc pair-wise differences were analyzed using Tukey’s test. Several covariates were considered; however, only those that were significant in most models or materially altered the results were retained. The final model included adjustments for sex, ethnocultural group, hormonal contraceptive use among women, age, log body mass index, and physical activity. The Benjamini-Yekutieli (B-Y) procedure \( = \alpha/\sum (1/i) \), where \( i \) varies from one to the total number of tests conducted] was applied to account for multiple testing (\( p < 0.01 \), calculated for testing 54 proteins \( \alpha = 0.05 \)). The B-Y method was selected because it allows for potential dependence between tests, and many of the proteins in the proteomics panel are biologically related \(^{153}\).
6.4 Results

Subject characteristics were similar across genotypes for all four SNPs, and the sex distribution was similar in each ethnocultural group (Tables 6.1 and 6.2). All SNPs were in linkage disequilibrium (ie >0.8).

Table 6.1: Subject Anthropometric Characteristics, According to 9p21 Genotype

<table>
<thead>
<tr>
<th>SNP</th>
<th>WC (cm)</th>
<th>BMI (kg/m²)</th>
<th>BPSYS (mmHg)</th>
<th>BPDIA (mmHg)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs10757274</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Risk allele G</td>
<td>73.5 ± 9.4</td>
<td>22.6 ± 3.6</td>
<td>113.7 ± 11.4</td>
<td>69.2 ± 8.1</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>74.8 ± 9.1</td>
<td>23.1 ± 3.7</td>
<td>114.2 ± 11.9</td>
<td>69.4 ± 7.6</td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>74.5 ± 9.7</td>
<td>22.8 ± 3.4</td>
<td>114.1 ± 11.4</td>
<td>69.6 ± 8.7</td>
<td>0.81</td>
</tr>
<tr>
<td>rs10757278</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Risk allele G</td>
<td>73.6 ± 9.2</td>
<td>22.7 ± 3.5</td>
<td>113.7 ± 11.6</td>
<td>69.0 ± 8.1</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>74.8 ± 9.1</td>
<td>23.1 ± 3.7</td>
<td>114.1 ± 11.7</td>
<td>69.4 ± 7.6</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>74.4 ± 9.7</td>
<td>22.7 ± 3.5</td>
<td>114.1 ± 11.6</td>
<td>69.9 ± 8.8</td>
<td>0.91</td>
</tr>
<tr>
<td>rs1333049</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Risk allele C</td>
<td>74.1 ± 9.5</td>
<td>22.7 ± 3.5</td>
<td>113.9 ± 11.5</td>
<td>69.8 ± 8.7</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>74.8 ± 9.1</td>
<td>23.1 ± 3.7</td>
<td>114.1 ± 11.8</td>
<td>69.4 ± 7.6</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>73.7 ± 9.3</td>
<td>22.7 ± 3.5</td>
<td>113.9 ± 11.7</td>
<td>69.1 ± 8.1</td>
<td>0.37</td>
</tr>
<tr>
<td>rs2383206</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Risk allele G</td>
<td>73.5 ± 9.4</td>
<td>22.6 ± 3.5</td>
<td>113.7 ± 11.8</td>
<td>69.2 ± 8.2</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>74.5 ± 8.8</td>
<td>23.1 ± 3.7</td>
<td>113.9 ± 11.6</td>
<td>69.3 ± 7.5</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>74.9 ± 10.1</td>
<td>22.9 ± 3.5</td>
<td>114.7 ± 11.7</td>
<td>69.7 ± 8.7</td>
<td>0.43</td>
</tr>
</tbody>
</table>

Values are mean ± SD. BMI; body mass index, BPSYS; systolic blood pressure, BPDIA; diastolic blood pressure, WC; waist circumference
* homozygous for the ancestral allele (lowest genetic risk), ** homozygous for the risk allele (highest genetic risk)

The risk-allele frequencies for rs10757274, rs10757278 and rs1333049 were similar in Caucasians and East Asians. East Asians had a lower frequency of the G allele of rs2383206 compared to Caucasians (p= 0.007) and South Asians (p= 0.004) (Table 6.2). Risk-allele frequency of the four SNPs in each group were like those reported elsewhere, and ranged from 40-48% except for East Asians who had a lower frequency of the G allele of rs2383206.
Table 6.2: Prevalence of Risk Alleles in Four 9p21 SNPs in Each Ethnocultural Group

<table>
<thead>
<tr>
<th>Genetic Marker</th>
<th>Risk Allele</th>
<th>Prevalence (%)</th>
<th>All</th>
<th>Caucasians</th>
<th>East Asians</th>
<th>South Asians</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs10757274</td>
<td>G</td>
<td>47</td>
<td>48</td>
<td>47</td>
<td>52</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>rs10757278</td>
<td>G</td>
<td>47</td>
<td>48</td>
<td>47</td>
<td>52</td>
<td>0.26</td>
<td></td>
</tr>
<tr>
<td>rs1333049</td>
<td>C</td>
<td>47</td>
<td>48</td>
<td>47</td>
<td>52</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>rs2383206</td>
<td>G</td>
<td>50</td>
<td>52</td>
<td>46</td>
<td>55</td>
<td>0.04</td>
<td></td>
</tr>
</tbody>
</table>

Out of 54 plasma proteins analyzed, four proteins (Haptoglobin β Chain, α-1b-Glycoprotein, Apolipoprotein (APO) E, and APO B-100) had significantly different mean concentrations across genotypes in at least one 9p21 SNP and one ethnocultural group (Table 6.3). Mean concentrations of these four proteins were significantly different from each other in some ethnocultural groups before addition of genotype. However, the association with genotype once added was in the direction of risk alleles for all except those of South Asians group. The presence of risk alleles from two SNPs was associated with increased plasma concentrations of Haptoglobin-β-Chain in Caucasians (rs2383206, p=0.01, and rs10757278, p=0.03), while one SNP exhibited a trend toward a statistically significant association (rs1333049, p=0.05, Table 6.3). Similar trends were observed with risk allele G in rs10757274 in Caucasians; however, this did not reach statistical significance (p=0.10). In East Asians, these associations were also observed, but the differences were only statistically significant for one SNP (rs2383206, p=0.01, Table 6.3). In South Asians, the opposite patterns were observed compared to the other two groups. Haptoglobin β-Chain concentrations were, overall, 30% higher in South Asians compared to the other two ethnicities (p-value <0.0001), but the mean Haptoglobin β-Chain plasma concentrations were not significantly different across genotypes in South Asians (p>0.20 for all three comparisons).
Table 6.3: Plasma Protein Concentrations of α-1b-Glycoprotein, Apolipoprotein B-100, Apolipoprotein E, and Haptoglobin β-chain in Four 9p21 SNP and Ethnocultural Group

<table>
<thead>
<tr>
<th>Plasma Protein</th>
<th>Caucasians</th>
<th>East Asians</th>
<th>South Asians</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rs10757274</td>
<td></td>
<td></td>
</tr>
<tr>
<td>α-1b-Glycoprotein</td>
<td>AA (133) 1.69 ± 0.05</td>
<td>AG (261) 1.75 ± 0.03</td>
<td>GG (129) 1.73 ± 0.05</td>
</tr>
<tr>
<td></td>
<td>p 0.08</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Apolipoprotein B-100</td>
<td>AA (110) 1.49 ± 0.04</td>
<td>AG (191) 1.61 ± 0.03</td>
<td>GG (87) 1.69 ± 0.05</td>
</tr>
<tr>
<td></td>
<td>p 0.02</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>Apolipoprotein E</td>
<td>AA (29) 1.78 ± 0.11</td>
<td>AG (57) 1.62 ± 0.07</td>
<td>GG (31) 1.60 ± 0.08</td>
</tr>
<tr>
<td></td>
<td>p 0.37</td>
<td>0.53</td>
<td>0.19</td>
</tr>
<tr>
<td>Haptoglobin β-chain</td>
<td>AA (139) 9.67 ± 0.42</td>
<td>AG (259) 10.9 ± 0.3</td>
<td>GG (125) 11.2 ± 0.5</td>
</tr>
<tr>
<td></td>
<td>p 0.16</td>
<td>0.03</td>
<td>0.10</td>
</tr>
<tr>
<td>rs10757278</td>
<td>AA (110) 1.49 ± 0.04</td>
<td>AG (191) 1.61 ± 0.03</td>
<td>GG (87) 1.69 ± 0.05</td>
</tr>
<tr>
<td></td>
<td>p 0.02</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>Apolipoprotein B-100</td>
<td>AA (29) 1.78 ± 0.11</td>
<td>AG (57) 1.62 ± 0.07</td>
<td>GG (31) 1.60 ± 0.08</td>
</tr>
<tr>
<td></td>
<td>p 0.37</td>
<td>0.53</td>
<td>0.19</td>
</tr>
<tr>
<td>Apolipoprotein E</td>
<td>AA (139) 9.67 ± 0.42</td>
<td>AG (259) 10.9 ± 0.3</td>
<td>GG (125) 11.2 ± 0.5</td>
</tr>
<tr>
<td></td>
<td>p 0.16</td>
<td>0.03</td>
<td>0.10</td>
</tr>
<tr>
<td>Haptoglobin β-chain</td>
<td>AA (139) 9.67 ± 0.42</td>
<td>AG (259) 10.9 ± 0.3</td>
<td>GG (125) 11.2 ± 0.5</td>
</tr>
<tr>
<td></td>
<td>p 0.16</td>
<td>0.03</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Protein units are reported in μmol/L. Analyses were adjusted for age, sex, hormonal contraceptive use, physical activity, and log body mass index.

*log-transformed plasma protein, **square-root transformed plasma protein

The presence of 9p21 risk alleles (C, G, and G) was associated with higher α-1b-Glycoprotein concentrations in East Asians from three SNPs (rs1333049, p=0.05, rs10757274 p=0.01, and rs10757278, p=0.04). In Caucasians and South Asians, however, no significant associations were observed (p>0.05 for all comparisons, Table 6.3). The presence of 9p21 risk alleles was associated with higher APO E concentrations in Caucasians and East Asians. In Caucasians, risk alleles (G, C, G, and G) from all four SNPs were associated with higher concentrations of APO E (Table 6.3).
In East Asians, similar patterns were observed for mean APO E concentrations, however, the differences were only significant for one SNP (Table 3). Similar patterns of association were observed in South Asians, but none reached statistical significance (Table 6.3). Mean plasma concentrations of APO B-100 varied across ethnicities and genotype. In East Asians, presence of risk alleles G from two SNP was associated with higher mean plasma concentrations of APO B-100 (Table 6.3). No other associations were found between the SNPs and plasma proteins in East Asians. In South Asians, an opposite association was observed, where the presence of risk alleles G from one SNP (rs2383206) was associated with lower mean plasma concentrations of APO B-100 (P=0.02) (Table 6.3). A trend of increased risk alleles with higher mean plasma concentrations of APO B-100 was observed with all SNPs in Caucasians, but none reached statistical significance (Table 6.3). After correction for multiple testing using B-Y formula (p<0.01), only the mean differences in APO B-100 in East Asians remained significant.
6.5 Discussion

We examined the association between the four most commonly studied SNPs in 9p21 that are associated with CVD risk and high-abundance plasma proteins in Caucasian, East Asian and South Asian young adults living in Canada. Our finding suggests novel and distinct associations of plasma proteins with 9p21 risk alleles that varied with each ethnocultural group. To our knowledge, this study is the first to examine the associations between the SNPs in 9p21 and the plasma proteome.

In the present study, Caucasians and East Asians who where carriers of risk alleles in common 9p21 SNPs had a higher plasma Haptoglobin-β-Chain concentration in a dose dependent fashion. Elevated levels of Haptoglobin have been deemed as independent predictor of CVD \(^{188}\) mortality \(^{189}\) and CVD in populations with diabetes \(^{190}\). Furthermore, haptoglobin gene polymorphisms have been associated with altered lipid metabolism, with high risk alleles contributing to elevated serum triglycerides \(^{191}\) and lower high density lipoprotein (HDL) \(^{192}\). Haptoglobin binds to Apoprotein A1 in the same location as lecithin-cholesterol acyltransferase, subsequently decreasing its activity and therefore limiting HDL maturation \(^{192}\). This inhibits reverse cholesterol transport, causing HDL to become proatherogenic \(^{192}\). In addition, the securing of Haptoglobin to HDL via the HP-ApoA1 allows the oxidation of HDL and its acquisition of proatherogenic and proinflammatory properties \(^{193,194}\). Haptoglobin is also a positive acute-phase plasma glycoprotein that binds to freely circulating hemoglobin for removal from circulation \(^{195}\). Excess hemoglobin metabolism leads to the release of extra-erythrocytic hemoglobin, with potentially severe consequences for health, including the propagation of ROS generation, and the disruption of cell membranes and cell function \(^{195}\). Haptoglobin clears extra-erythrocytic hemoglobin from the circulation by
modulating inflammatory and anti-inflammatory cytokines produced by macrophages that have been exposed to free hemoglobin \textsuperscript{196}. One study has also established a link between the genetic susceptibility to CVD and the response to inflammatory signalling in human vascular cell type, via gene enhancers identified in the 9p21 region, including variation in rs10757278 which is now known to be located in one of the enhancers \textsuperscript{197}. We have demonstrated here that the rs10757278 risk allele is associated with significantly higher serum haptoglobin in Caucasians. This might suggest a possible interaction between haptoglobin and 9p21 SNPs leading to a modulation of inflammatory responses that are associated with atherosclerotic plaque formation that lead to CVD events later in life.

In the present study, a higher plasma Alpha-1β Glycoprotein concentration was associated with risk alleles on 9p21 SNPs in East Asians only. This protein is a plasma glycoprotein of unknown function that belongs to the immunoglobulin supergene family with a similar structure as the receptor for trans-epithelial transport of IgA and IgM \textsuperscript{198}, suggesting its possible involvement in inflammation and immune response. This protein seems to be associated with 9p21 SNPs in a unique way in East Asians.

In the current study, the presence of 9p21 risk alleles was also associated with higher APO-E concentrations in Caucasians and East Asians. Plasma lipoproteins are important determinants of atherosclerosis and their specific apolipoprotein constituents have been identified as predictors of CVD events \textsuperscript{199}. APO-E has a variety of proposed functions including acting as a ligand for lipoprotein receptors in up-take of chylomicrons and intermediate density lipoproteins by liver \textsuperscript{200}. 86
In addition to its role in lipid metabolism, APO-E is also involved in antioxidation, inhibition of platelet aggregation, and immunomodulation. Our findings suggest one or more of these pathways may be associated with the presence of 9p21 SNPs. In one previous study, inflammatory signalling in human vascular cell types has been shown to be influenced by 9p21 genotype, thus the current study further supports the exploration of inflammatory pathways of CVD through 9p21.

Finally, we found that, in East Asians, the concentration of Apolipoprotein B-100 increases with the presence of risk alleles in two 9p21 SNPs. APO B-100 plays a central role in lipid metabolism. APO B-100 is a major protein of Very Low Density Lipoproteins and Intermediate Density Lipoproteins, and sole protein of LDL where it functions as a ligand on the surface of LDL for cellular uptake of cholesterol. Higher plasma levels of APO B-100 correlate with higher total levels of atherogenic lipoproteins, which have been associated with higher risk of coronary heart disease. In the current study, for the first time, we demonstrate that higher APO B-100 is associated with the presence of 9p21 risk alleles, further suggesting that 9p21 may modulate CVD development via multiple pathways, including traditional and novel mechanisms that might differ based on ethnocultural origin.

It is also important to highlight that the South Asians in the present study showed the opposite direction of association of risk alleles with at least three of the four proteins identified, compared to those seen in Caucasians and East Asians. For example, carriers of risk alleles had lower plasma Haptoglobin β Chain concentrations, an opposite trend that was seen in the other two groups, although these differences did not reach statistical significance. These patterns might be simply
due to chance, given the smaller sample size of this group, or might indicate other possible contributing factors, such as distinct dietary patterns or other ethnocultural and genetic factors that are unique to South Asians. South Asians have been known to have a higher risk of CVD as well as lower HDL functionality compared to Caucasians; therefore, different mechanisms by which 9p21 risk alleles contribute to CVD may be a possibility. This finding may also be due to the smaller sample size for the population of South Asians in the present study, compared to Caucasians and East Asians.

This study is the first to show an association between 9p21, one of the most robust genetic predictors of CVD, and a panel of plasma proteins that are involved in several physiological and metabolic pathways. The targeted mass spectrometry, with a stable-isotope labeled internal standard peptide (SIS peptide) corresponding to each targeted peptide is the gold standard for highly abundant proteome quantification. However, the present study has some limitations, and future studies could use these findings to build new knowledge in discovery of mechanisms of action in how risk alleles in 9p21 determine CVD risk. Future studies could consider other ‘omics’ technologies such as metabolomics, lipidomics or unselected large-scale proteomics panels for new biomarker discovery. In addition, given the young age of the population in the present study, follow up data from this cohort could be informative in analyzing plasma protein changes that could be related to development of disease later in life. One of the limitations of the present study was the cross-sectional design, which precludes establishing causality for any of the observed associations. In addition, the smaller sample size of the East Asian and South Asian groups may have led to a lack of sufficient statistical power to adequately assess the effect of 9p21 risk alleles on proteomics in these groups. However, the young age of the cohort and lack of other
chronic diseases limits possibilities confounding factors, unlike other older and diseased cohorts. Another limitation is the proteomics assay that was based on the measurement of peptides that are surrogates of specific proteins, assuming 100% digestion efficiency. Incomplete digestion may result in slight inaccuracies. However, as indicated by its low intra- and inter-assay variability, the MRM assay is very reproducible and allows for precise comparisons between samples 209. Furthermore, although commonly used immunoassay methods have limitations such as poor antibody specificity, our values for most proteins are comparable to published values 96,209. Finally, although we adjusted for several possible covariates, residual confounding may have affected some of the observed results, as well as incidental findings due to multiple testing. However, these factors are unlikely to have differed between genotypes.
6.6 Conclusion

Our findings suggest that 9p21 genotypes may play a role in known as well as unique pathophysiological pathways that contribute to potential CVD biomarkers in early adulthood. These contributions might be distinct amongst different ethnocultural groups.
Chapter 7

Discussion, Limitations and Future Directions
7 General Discussion

7.1 Thesis Discussion

The present thesis is the first to examine the relationship between genetic polymorphisms in 9p21 and traditional as well as emerging proteomics biomarkers of CVD risk. We identified an association between fasting insulin and 9p21 genotype that was unique to females not on hormonal contraceptives. In addition, we identified four abundant plasma proteins that were associated with 9p21 genotype, varying across ethno-cultural groups, but not modified by sex or hormonal contraceptive use. Lastly, we found that Prudent dietary patterns modified 9p21 genetic risk of higher insulin in Caucasians and East Asians who were heterozygous for the risk alleles, while no associations were found between Western and Eastern dietary patterns or between Prudent diet and other SNPs.

Higher fasting insulin levels are associated with numerous adverse risk factors in young adults, increasing the risk of atherosclerosis \(^{68}\) and type 2 diabetes \(^{69}\). Even in individuals without diabetes, hyperinsulinemia is associated with decreases in insulin-mediated glucose uptake \(^{70}\), as well as with a number of clinical symptoms associated with insulin resistance \(^{71,72}\). Prolonged and untreated insulin resistance and hyperinsulinemia are known to be associated with hypertriglyceridemia \(^{73}\) and low concentrations of high-density lipoprotein cholesterol \(^{73,74}\), hypertension \(^{75}\), and coronary artery calcification \(^{76}\), which are all risk factors for CVD. Hyperinsulinemia has also independently been associated with ischemic heart disease \(^{72}\). The five most commonly studied SNPs in the 9p21 region (rs10757272, rs10757278, rs1333049, rs2383206
and rs4977574) were associated with higher levels of fasting serum insulin in this cohort, with rs10757278 showing the strongest association. This suggests that pathways involving insulin might be involved in mechanisms by which the 9p21 region is linked with CVD.

The synergism of risk alleles in the 9p21 locus and hyperglycemia have also been demonstrated in coronary artery disease (CAD) development and accelerated CVD mortality in those with type 2 diabetes \textsuperscript{154}; where those homozygous for the risk alleles had worse glycemic control and a higher incidence of CAD than others. Although 9p21 does not contain annotated genes, rs10757278 is in high linkage disequilibrium with two known genes, CDKN2A and CDKN2B. These genes code for three known proteins \( p_{15}^{INK4b}, p_{16}^{INK4b} \) and ARF, that have been linked to the presence of hereditary atherosclerosis \textsuperscript{155,156} where they inhibit cyclin-dependent kinases controlling cell proliferation and apoptosis in endothelial tissue of vascular smooth muscle, among other tissues. One study found that antisense noncoding RNA expression in the INK4 locus (ANRIL) was significantly higher in carriers of the 9p21 risk alleles, with expression directly correlating with atherosclerotic severity \textsuperscript{157}. The findings from this thesis suggest that insulin-dependent pathways might be one of the earliest signs of metabolic dysregulation, linking 9p21 risk alleles to CVD risk factors, evident in early adulthood.

We observed a significant genotype-sex interaction on fasting insulin, which has not been previously reported. The associations were observed in women, but not in men. Similarly, in the ADVANCE study, where genotypic odds ratio (OR) was stratified by sex, a trend towards higher ORs in women rather than men was observed \textsuperscript{110}. The authors, however, concluded that this was likely due to admixture of genotype present in the population, since the risk allele frequency differed between men and women in that cohort \textsuperscript{110}. More recently, GG genotypes of rs10757278
were also shown to be significantly and independently associated with carotid plaque in women only. Insulin resistance, hormonal glucose intolerance and type 2 diabetes increase the risk of CVD and mortality, however, women are more affected as it increases the risk of CVD death by about four times in women verses two-fold in men. In other studies, women have been identified as being distinctly different from men with regard to insulin action, susceptibility to develop insulin resistance, and response to stimuli that are known to enhance or impair sensitivity to the effects of insulin. Our findings show a sex-specific genetic association between 9p21 and fasting insulin that might contribute to our understanding of how 9p21 impacts the development of CVD.

Additional analyses in the present study demonstrated that when women were grouped according to HC use, those on HCs had significantly higher mean fasting insulin with no significant association between genotype and fasting insulin. In contrast, those who were not taking HCs had lower mean fasting insulin, which was significantly associated with 9p21 genotype. HC use has been associated with insulin resistance, manifested by reduced peripheral tissue insulin sensitivity, however, this association has not been shown to be related to 9p21 genotype elsewhere. In the present study, Caucasian women formed the largest ethnocultural group, where the number of those on HCs versus those without was similar. When comparing mean fasting insulin according to genotype, a protective effect seems to be present in carriers of the A allele for rs10757278, whereas in those taking HCs this protection was not observed, despite having the same genotype and ethnocultural group.

In additional to the stratified analyses according to ethnocultural group, sex and HC, we found that those who were homozygous for the risk alleles of 9p21 (rs1333049) had higher levels of fasting
plasma insulin than the lower risk group (p=0.04). In addition, HOMA-IR and HOMA-Beta were also elevated among the same group, particularly in those consuming a low prudent diet. Insulin resistance earlier in life is often predictive of CVD in middle and late adulthood, even in those without diabetes. In addition to traditional pathways, insulin resistance has also been linked to atherosclerosis through pro-inflammatory activity on vascular and immune cells. Mechanisms by which, 9p21 risk alleles are associated with CVD are largely unknown, however, there is some evidence that endothelial dysfunction might be one of the pathways. In addition, the same mechanism identified that induces endothelial dysfunction, interferon-γ activation, has also been shown to induce IR in human adipose tissue. Therefore, our findings are in agreement with these cellular discoveries and might aid in further understanding of the role of 9p21 SNPs in CVD risk.

Mean HOMA-IR was 150% higher and HOMA-Beta was 30% higher in those with a low Prudent dietary pattern versus those with a high Prudent dietary pattern from the same high-genetic risk group (p=0.004, p=0.013, respectively). Higher Prudent dietary pattern scores appeared protective against IR and Beta-cell dysfunction in the genetically high-risk group in a step-wise fashion. Meanwhile in the low-genetic risk group, Prudent dietary patterns were not associated with mean HOMA-IR nor HOMA-Beta (p=0.08 and p=0.36, respectively). Although associations of fasting insulin and 9p21 have not been previously reported, others have reported on 9p21 gene-diet interaction with related outcomes including; glycemic control (plasma glucose and glycosylated hemoglobin), lipid profile and CVD events. Our findings are in agreement with previous findings and add to the evidence that CVD risk factors may be modifiable in young adults with high risk 9p21 SNPs. It is also important to note that these findings could potentially explain previous inconsistencies between studies linking one-size fits all dietary interventions and
might justify a more personalized approach to dietary interventions in prevention of CVD. A recent GWAS suggested interplay of genes involved in the metabolic response to dietary patterns on obesity, glucose metabolism and food-induced response in the brain in the adoption of dietary patterns. These factors could additionally be explored in the future to identify underlying relationships between gene-environment implications of CVD risk in 9p21 studies.

Higher Prudent pattern scores in the current study had strong correlations with increased vitamin and mineral levels (vitamins A, B1, B2, B6, C, D, and K, magnesium, and iron) and this was likely due to the relatively large presence of fruits and vegetables in this dietary pattern. Higher intake of fruits and vegetables have been shown to be protective towards CVD risks in several dietary patterns and populations. Several mechanisms have been proposed by which these foods seem to be protective and include; via bioactive components with antioxidant, anti-inflammatory, and electrolyte properties, as well as functional properties, such as low glycemic load and low energy density. In the current study, high fruits and vegetable intake was a distinguishing feature of both the Prudent dietary patterns and the Eastern dietary patterns, however, no associations were found with the Eastern dietary patterns and 9p21 genetic risk on insulin sensitivity. One of the reasons for this lack of association with the Eastern pattern might have been that the majority of fruits and vegetables in this dietary pattern was cooked or processed whereas the fruits and vegetables in the Prudent dietary patterns were often freshly consumed with minimal processing. This association was evident when food groups were analyzed and consumption fresh vegetables were shown to interact with two 9p21 SNPs on HOMA-IR (p=0.022 and 0.025). Our finding is in agreement with two other studies where fresh vegetables intake were protective against CVD in high-risk genotype in 9p21.
Higher Western dietary patterns in the current population were correlated with high sodium, fat, saturated fat and energy intake, likely because of the high proportion of processed foods included in this pattern. The Western dietary pattern was not associated with insulin markers and 9p21 genotype. This finding is also in agreement with another study that found while Prudent dietary patterns attenuate the 9p21 genetic risk, Western dietary patterns in relation to 9p21 did not affect CVD events. This is contrary to other findings were Western dietary patterns have been linked to CVD and diabetes among many other chronic conditions. Although one large systematic review found strong evidence to support a causal relationship between Prudent/Mediterranean dietary factors and reduction in CVD, they concluded that there was insufficient evidence against components of a Western dietary pattern (including saturated fats) to support a casual relationship with increased CVD. In the same systematic review, however, there was strong evidence for trans-fats being a contributing factor towards CVD and Westerns diets are typically high in trans-fats. In the current thesis, however, our dietary patterns did not characterize trans-fats as major components of any of the three dietary patterns. This might have been a limitation of the FFQ in measuring trans-fat content of the foods reported or that the study population did not have a high trans-fat dietary intake associated with the three dietary patterns.

Currently in clinical settings, biomarkers of CVD risk, including elevated CRP, and dyslipidemia determine intermediate phenotypic presentation of CVD and predict 10-year risk of CVD events. However, in the present thesis, similar to others, we found no meaningful associations between 9p21 genotype and traditional CVD biomarkers including lipid profile, BMI, or hypertension suggesting that insulin-related associations observed might be either independent of those involved in cardiometabolic disease or the one of the earliest dysregulations seen in the cascade of phenotypic events. Although the mechanisms by which 9p21 SNPs influence CVD risk are not
well defined, a link between CVD genetic susceptibility and the response to inflammatory signaling has been reported. No clear relationship between inflammatory markers and 9p21 genotype was observed in the present thesis, and this could be due to the young age of the study participants and the observational nature of the study design.

Others have reported long-distance interactions of CDKN2A/B, the MTAP gene, and interval downstream of IFNA21 human vascular endothelial cells. Interferon-γ (IFNγ) activation seems to affect the structure of chromatin, therefore, transcriptional regulation in the 9p21 locus, including STAT1-binding, likely plays a role as a long-range enhancer, altering expression of neighbouring genes. One study identified 33 enhancers in the 9p21 region, with rs10757278 located in one of these enhancers, where the presence of the risk allele (G), disrupted a binding site for STAT1. The link between this disturbance and insulin resistance was demonstrated by another study where IFNγ was shown to induce insulin resistance in mature human adipocytes.

In the present thesis, we demonstrated that 9p21 risk alleles were associated with higher serum insulin levels, particularly with rs10757278. This finding may be explained by a possible mechanism whereby the A>G substitution in rs10757278 disrupts STAT1 binding site and modifies IFNA21 gene expression. IFNγ then possibly induces insulin insensitivity seen as higher insulin to be an early marker of endothelial dysfunction and increased susceptibility to CVD later in life.

Caucasians and East Asians who were carriers of risk alleles in common 9p21 SNPs had a higher plasma Haptoglobin-β-Chain concentration in a dose dependent fashion. Elevated levels of Haptoglobin have been identified as independent predictors of CVD mortality and CVD in populations with diabetes. Furthermore, haptoglobin gene polymorphisms have been associated
with altered lipid metabolism, with high risk alleles associated with elevated serum triglycerides \(^{191}\) and lower high density lipoprotein (HDL) \(^{192}\). Haptoglobin binds to Apoprotein A1 in the same location as lecithin-cholesterol acyltransferase, subsequently decreasing its activity and therefore limiting HDL maturation \(^{192}\). This inhibits reverse cholesterol transport, causing HDL to become proatherogenic \(^{192}\). In addition, the securing of Haptoglobin to HDL via the HP-ApoA1 allows the oxidation of HDL and its acquisition of proatherogenic and proinflammatory properties \(^{193,194}\). Haptoglobin is also a positive acute-phase reactant, a plasma glycoprotein that binds to freely circulating hemoglobin for removal from the circulation \(^{195}\). Excess hemoglobin metabolism leads to the release of extra-erythrocytic hemoglobin, with potentially severe consequences for health, including the propagation of ROS generation, and the disruption of cell membranes and cell function \(^{195}\). Haptoglobin clears extra-erythrocytic hemoglobin from the circulation by modulating inflammatory and anti-inflammatory cytokines produced by macrophages that have been exposed to free hemoglobin \(^{196}\). One study has also established a link between the genetic susceptibility to CVD and the response to inflammatory signaling in human vascular cell type, via gene enhancers identified in the 9p21 region, including variation in rs10757278 which is now known to be located in one of the enhancers \(^{197}\). We have demonstrated here that the rs10757278 risk allele is associated with significantly higher serum haptoglobin in Caucasians. This might suggest a possible association between haptoglobin and 9p21 SNPs leading to a modulation of inflammatory responses that are associated with atherosclerotic plaque formation that lead to CVD events later in life.

Higher plasma Alpha-1β Glycoprotein concentration was associated with risk alleles on 9p21 SNPs in East Asians only. This protein is a plasma glycoprotein of unknown function that belongs to the immunoglobulin supergene family with a similar structure as the receptor for trans-epithelial
transport of IgA and IgM\textsuperscript{198}, suggesting its possible involvement in inflammation and immune response. This protein seems to be associated with 9p21 SNPs in a unique way in East Asians.

Presence of 9p21 risk alleles was also associated with higher APO-E concentrations in Caucasians and East Asians. Plasma lipoproteins are important determinants of atherosclerosis and their specific apolipoprotein constituents have been identified as predictors of CVD events\textsuperscript{199}. APO-E has a variety of proposed functions including acting as a ligand for lipoprotein receptors in up-take of chylomicrons and intermediate density lipoproteins by the liver\textsuperscript{200}. Common genetic polymorphisms of APO-E, including APO-E2, 3 and 4 have functional effects on lipoprotein metabolism mediated through hepatic binding, uptake, and catabolism of chylomicrons, chylomicron remnants, VLDL, and HDL subspecies\textsuperscript{211}. APO-E2 has lower receptor binding affinity that results in delayed clearance of APO-E2-bearing lipoprotein particles from plasma\textsuperscript{212}. APO-E4 is distributed differently from APO-E3 between VLDL and HDL and is degraded more rapidly than APO-E3. This may enhance the catabolism of E4-bearing particles, leading to elevated levels of LDL\textsuperscript{212}. In the current thesis, however, although APO-E genetic polymorphism was not investigated, differences in total plasma APO-E concentration by 9p21 potentially present a novel link to APO-E and LDL metabolism. In addition to its role in lipid metabolism, APO-E is also involved in antioxidation\textsuperscript{201}, inhibition of platelet aggregation\textsuperscript{202}, and immunomodulation\textsuperscript{203}. Our findings suggest one or more of these pathways may be associated with the presence of 9p21 SNPs.

In a previous study, inflammatory signaling in human vascular cell types has been shown to be influenced by 9p21 genotype\textsuperscript{197}, thus the current thesis further supports the exploration of inflammatory pathways of CVD through 9p21 through different mechanisms.
In East Asians, the concentration of Apolipoprotein B-100 increases with the presence of risk alleles in two 9p21 SNPs. APO B-100 plays a central role in lipid metabolism. APO B-100 is a major protein of VLDL and IDL, and sole protein of LDL where it functions as a ligand on the surface of LDL for cellular uptake of cholesterol. Higher plasma levels of APO B-100 correlate with higher total levels of atherogenic lipoproteins, which have been associated with higher risk of coronary heart disease. In the current thesis, for the first time, we demonstrate that higher APO B-100 is associated with the presence of 9p21 risk alleles, further suggesting that 9p21 may modulate CVD development via multiple pathways, including traditional and novel mechanisms that might differ based on ethnocultural origin.

Since lipid metabolism is modulated through insulin-regulated mechanisms, it is important to explore a possible link between 9p21 genotype, insulin resistance and variation in apolipoproteins observed this thesis. In the fed state, high insulin levels act on the adipocyte to promote triglyceride uptake and inhibit free fatty acid release whereas in fasting, or insulin resistant state, this process is reversed. Free fatty acids are released from the adipocyte and delivered to the liver where they are re-esterifies to triglycerides and secreted as VLDL. In addition, insulin promotes uptake of LDL through several mechanisms, including by increasing LDL-receptor mRNA in hepatocytes. Given that insulin in part regulates the LDL-receptor binding APO-E and APO-B100, key determinants of plasma lipoprotein concentration, perhaps insulin resistance might be hypothesized as the underlying mechanism by which fasting plasma insulin, APO-E and APO-100 were observed to be higher in risk allele carriers of 9p21.

It is also important to highlight that the South Asians in the present study showed the opposite direction of association of risk alleles with at least three of the four proteins identified, compared
to those seen in Caucasians and East Asians. For example, carriers of risk alleles had lower plasma Haptoglobin β Chain concentrations, an opposite trend that was seen in the other two groups, although these differences did not reach statistical significance. These patterns might be simply due to chance, given the smaller sample size of this group, or might indicate other possible contributing factors, such as distinct dietary patterns or other ethnocultural and genetic factors that are unique to South Asians. South Asians have been known to have a higher risk of CVD as well as lower HDL functionality compared to Caucasians; therefore, different mechanisms by which 9p21 risk alleles contribute to CVD may be a possibility. This finding may also be due to the smaller sample size for the population of South Asians in the present study, compared to Caucasians and East Asians.
7.2 Summary of Findings

The overall goal of this thesis was to determine the association between 9p21 polymorphisms and early biomarkers of CVD risk in a young cohort of adults, including biomarkers of glycemic dysregulation and the plasma proteome, as well as to examine whether other factors modify these associations.

Objective 1: To examine the association between 9p21 genotype and traditional CVD markers, and to determine whether ethnicity, sex and HC use modify these associations.

Results: In the entire group, the risk alleles of rs10757278 and rs2383206 were associated with higher mean insulin (p=0.01) (Thesis Chapter 4). Risk alleles for rs4977574, rs10757278, rs2383206, rs1333049 and rs10757274 were associated with higher serum insulin in women (p=0.008, p=0.008, p=0.0003, p=0.002, and p=0.001, respectively), but not in men (p=0.41, p=0.13, p=0.31, p=0.34, and 0.35, respectively). The association between 9p21 and insulin remained significant among women not taking hormonal contraceptives (HC), but was not significant among women taking HCs. No other associations were found with other CVD risk biomarkers. These findings suggest that 9p21 genotypes may play a role in the development of insulin resistance in early adulthood among women, but not men, and the associations appear to be attenuated by HC use.
**Objective 2:** To examine the association between 9p21 genotype and insulin sensitivity, and to determine whether ethnicity, sex, HC use and dietary patterns modify these association.

**Results:** Among those who were homozygous for the 9p21 risk allele (rs1333049), fasting insulin was 40% higher in those who were consuming a low-prudent diet compared to those consuming a high-prudent diet (p<0.05) (Thesis Chapter 5). No differences were observed between those following a low- versus high-prudent diet among those who did not carry a risk allele. Significant gene-diet interactions on insulin sensitivity using HOMA-IR were observed with all five SNPs, which remained significant after adjusting for covariates (p<0.05). Similar findings were observed with HOMA-Beta, however, the association was only significant for rs10757274 (p=0.04). These findings suggest that a Prudent dietary pattern may protect against the effects of 9p21 risk genotypes on insulin sensitivity.
**Objective 3:** To examine the association between 9p21 genotype and the plasma proteome, and to determine whether ethnicity, sex, and HC use modify this association.

**Results:** In Caucasians (n=524), four SNPs were associated with Apolipoprotein E and two with Haptoglobin-β-Chain concentration. In East Asians (n=388), three SNPs were associated with α-1b-Glycoprotein, two with Apolipoprotein B-100, and one with Apolipoprotein E and Haptoglobin-β-Chain concentration. In South Asians (n=117), one SNP was associated with Apolipoprotein B-100 concentration (Thesis Chapter 6). Sex and HC use did not significantly modify these associations. These findings suggest that 9p21 genotypes may play a role in various pathophysiological pathways that contribute to risk of CVD in early adulthood and these variations might be distinct amongst different ethnocultural groups.
7.2.1 9p21 Relationship with Cardiovascular Disease Risk Markers

Overall, the results of the present thesis indicate that 9p21 SNPs are not associated with biomarkers of CVD including, LDL, HDL, TC, or CRP, other than fasting insulin, in young adults (Chapter 4). In addition, we found no relationship between 9p21 and anthropometric measures often associated with CVD risk predictions including blood pressure, BMI, and/or waist circumference.

Currently in clinical settings, biomarkers of CVD risk, including elevated CRP, and dyslipidemia determine intermediate phenotypic presentation of CVD and predict 10-year risk of CVD events. However, in the present study, similar to others, we found no meaningful associations between these risk factors and 9p21 in this group of young adults.

Although the mechanisms by which 9p21 SNPs influence CVD risk are not well defined, a link between CVD genetic susceptibility and the response to inflammatory signaling has been reported. No clear relationship between inflammatory markers and 9p21 genotype was observed in the present thesis, and this could be due to the young age of the study participants and the observational nature of the study design. However, others have reported long-distance interactions of CDKN2A/B, the MTAP gene, and interval downstream of IFNA21 human vascular endothelial cells possibly through 33 long-range enhancers in the 9p21 region where a substitution of risk alleles seems to disrupt binding sites. The link between this disturbance and insulin resistance was demonstrated by another study where IFNγ was shown to induce insulin resistance in mature human adipocytes. In the present thesis, we demonstrated that 9p21 risk alleles were associated with higher serum insulin levels, particularly with rs10757278. This finding may be explained by a possible mechanism whereby the A>G substitution in rs10757278 disrupts STAT1 binding site and modifies IFNA21 gene expression. IFNγ then possibly induces insulin insensitivity seen as
higher insulin to be an early marker of endothelial dysfunction and increased susceptibility to CVD later in life.

Early detection of CVD is necessary to improve health outcomes, however, most studies in this area have been in older populations and do not use a comprehensive approach of combining genetic risk factors with other risk factors. This approach can potentially have an added value to the traditional biomarkers used to manage CVD in high-risk populations. In one study, adding the 9p21 risk allele to traditional risk factors was associated with improved predictably of CVD hazard ratio of incident coronary heart disease of 1.2 per allele (p<0.000003)\textsuperscript{169}. The currently study provides further support in better defining at risk groups that might have otherwise been missed using traditional markers only.

In summary, the present study is the first to demonstrate that common risk variants in the 9p21 region are associated with elevated fasting insulin in young adult women, but not in men, and not in women taking HCs. The role of 9p21 in insulin signaling and glucose metabolism in young women and its possible links to CVD risk warrants further investigation.
7.2.2 9p21 Relationship with Insulin Sensitivity and Diet

We demonstrated an association between 9p21 and insulin markers, including fasting insulin, insulin sensitivity and pancreatic beta-cell function (Chapter 4 and 5). HOMA-IR and HOMA-Beta were significant in Caucasian and East Asian men and women who were homozygous for 9p21 risk alleles. However, following a habitual Prudent dietary pattern was inversely associated with insulin markers in this group, but no other gene-diet associations were found between the other dietary patterns nor the other genotypes.

Higher Prudent pattern scores in the current thesis had strong correlations with increased some vitamin and mineral levels because of the relatively large presence of fruits and vegetables in this pattern. Several mechanisms have been proposed by which these foods seem to be protective and include; via bioactive components with antioxidant, anti-inflammatory, and electrolyte properties, as well as functional properties, such as low glycemic load and low energy density. High fruits and vegetable intake was also distinguishing feature of the Eastern dietary patterns, however, no associations were found with the Eastern dietary patterns and 9p21 genetic risk on insulin sensitivity. One of the reasons for this lack of association might have been that the majority of fruits and vegetables in the Eastern dietary patterns were cooked or processed whereas the fruits and vegetables in the Prudent dietary patterns were often consumed raw with minimal processing. This difference in association was evident when food groups were analyzed and consumption fresh vegetables were shown to interact with two 9p21 SNPs on HOMA-IR (p=0.022 and 0.025). Our finding is in agreement with two other studies where fresh vegetables intake were protective against CVD in high-risk genotype in 9p21.
Higher Western dietary patterns in the current population were correlated with high sodium, fat, saturated fat and caloric intake, likely because of the high proportion of processed foods included in this pattern. This dietary pattern was not associated with insulin markers and 9p21 genotype. This finding is also in agreement with another study that found while Prudent dietary patterns attenuate the 9p21 genetic risk, Western dietary patterns did not effect CVD events. This is contrary to other epidemiological findings were Western dietary patterns have been linked to CVD and diabetes among many other chronic conditions. Although one large systematic review found strong evidence to support a causal relationship between Prudent/Mediterranean dietary factors and insufficient evidence against components of a Western dietary pattern (including saturated fats). Another reason for this lack of association could have been due to the general dietary habits of the group, given that they were young university students, mostly form the Nutritional Sciences studies, hence, the “Western” dietary pattern they followed was not as potent as those consumed in other populations. This concept can also be extrapolated from the results of the PCA analysis, where the Western dietary pattern only explained less than 5% of the between-person variation in dietary intake.

In summary, the present thesis is the first to demonstrate that common risk variants in the 9p21 region are that were shown to be associated with elevated fasting insulin in young adult earlier, are modifiable by a high-Prudent dietary pattern. The role of dietary interventions in modifying 9p21 risk in insulin sensitivity and its possible links to CVD risk warrants further investigation.
7.2.3 9p21 Genotype and the Plasma Proteome

Overall, there were only a limited number of associations with the 9p21 genotype and the 54 high-abundance plasma proteins (chapter 6). These associations varied by ethnicity, however, were not meaningfully affected by HC use or sex. In Caucasians, four SNPs were associated with Apolipoprotein E and two with Haptoglobin-β-Chain concentration. In East Asians, three SNPs were associated with α-1b-Glycoprotein, two with Apolipoprotein B-100, and one with Apolipoprotein E and Haptoglobin-β-Chain concentration. In South Asians, one SNP was associated with Apolipoprotein B-100 concentration.

The finding of this thesis suggests a possible interaction between haptoglobin and 9p21 SNPs in Caucasians, which might indicate modulation of inflammatory responses that are associated with atherosclerotic plaque formation that lead to CVD events later in life in this group. One study has also established a link between the genetic susceptibility to CVD and the response to inflammatory signaling in human vascular cell type, via gene enhancers identified in the 9p21 region, including variation in rs10757278 which is now known to be located in one of the enhancers 197.

APO-E has a variety of proposed functions including acting as a ligand for lipoprotein receptors in up-take of chylomicrons and intermediate density lipoproteins by liver 200. Our findings suggest one or more of these pathways may be associated with the presence of 9p21 SNPs. In one previous study, inflammatory signaling in human vascular cell types has been shown to be influenced by 9p21 genotype 197, thus the current study further supports the exploration of inflammatory pathways of CVD through 9p21.
APO B-100 also plays a central role in lipid metabolism. In the current study, for the first time, we demonstrate that higher APO B-100 is associated with the presence of 9p21 risk alleles, further suggesting that 9p21 may modulate CVD development via multiple pathways, including traditional and novel mechanisms that might differ based on ethnocultural origin.

Lastly, higher plasma Alpha-1β Glycoprotein concentration was associated with risk alleles on 9p21 SNPs in East Asians only. This protein is a plasma glycoprotein of unknown function that belongs to the immunoglobulin supergene family with a similar structure as the receptor for trans-epithelial transport of IgA and IgM, suggesting its possible involvement in inflammation and immune response. This protein seems to be associated with 9p21 SNPs in a unique way in East Asians.
7.3 Limitations

7.3.1 Insulin Sensitivity

The gold standard for measuring insulin resistance and β-cell dysfunction are the hyperinsulinemic euglycemic clamp and the hyperglycemic clamp, respectively.\textsuperscript{170}

HOMA models correlate well with the clamp methods, however, they are surrogate insulin resistance markers generated using an estimated likelihood of the fasting plasma\textsuperscript{179}. Therefore, they might not be valid in some individuals, particularly those of non-Caucasian origin since validation studies have been mostly conducted in Caucasians of European decent\textsuperscript{179,216}.

7.3.2 Dietary Assessment

The food frequency questionnaire (FFQ) used in this study was a semi-quantitative detailed questionnaire containing 196 questions, which is a modified form of the Willet FFQ that has been well validated to reflect nutrient intake over a one-year period\textsuperscript{217}. The original FFQ, however, was only 61\textsuperscript{217} and then 131-items\textsuperscript{134} in the more recent versions, compared to the Toronto modified version used here and this expansion of the FFQ might falsely indicate a higher intake of some foods and nutrients represented in the study. Larger FFQ’s are often more comprehensive in reflecting foods and nutrients consumed by individuals in epidemiological studies, however, when used in culturally diverse groups, they might not represent specific food items or meal preparations used in different cultures, hence not accurately reflecting nutritional intake of certain groups as well as others. In the current thesis, this might have been the case for the non-Caucasians populations\textsuperscript{136}. 
Dietary patterns that were generated based on the FFQ in the current thesis also have limitations in their application to the study population. First, the Prudent dietary pattern was the most robust pattern, containing the largest number of food items that were above the 0.25 cut off for factor loading. This pattern was also the most likely dietary pattern to be consumed by Caucasians and female participants, who were the largest population of the study. The Western and Eastern dietary patterns were represented by a substantially smaller number of foods, which could indicate that the FFQ did not represent foods typically consumed in these dietary patterns as well as the food items in the Prudent dietary pattern. In addition, the Western and Eastern dietary patterns each explained less than 5% of the between-person variation in dietary intake 136.

7.3.3 Genetic Analyses

The Sequenom MassARRAY platform used in the current thesis has several features that are important for accurate custom SNP genotyping. It has modest multiplexing and minimal assay setup steps due to unmodified oligonucleotide primers. It utilizes a homogeneous reaction format with a single extension primer to generate allele-specific products with distinct masses, multiplexed polymerase chain reactions, a single termination mix and universal reaction conditions for all SNPs 218. This technique was the standard approach for SNP analysis when the genetic material for the parent study was collected between 2006-2010. However, over the past decade, there have been several additional SNPs identified as well as repeat patterns (copies) of risk allele presence have been discovered in CVD cases 219, that have not been analyzed in the current thesis. Although most 9p21 SNPs have been shown to be in linkage disequilibrium, and these results were in agreement with the study population of this thesis, only five SNPs were available for analysis, hence limiting the findings only to the five SNPs studied here.
7.3.4 Proteomics Analyses

Concentrations of common plasma proteins were measured by using a LC/MRM-MS without enrichment or affinity depletion as described elsewhere \(^9^1\). The targeted mass spectrometry, with a stable-isotope labeled internal standard peptide (SIS peptide) corresponding to each targeted peptide is the gold standard abundant proteome quantification \(^9^1,2^0^7,2^0^8\). The concentrations for most of the 54 proteins included in the original assay were within the range of reported values, \(^9^6\). These methods have been validated in an intra-lab investigation \(^1^8^7\). However, the proteomics assay used was based on the measurement of peptides that are surrogates of specific proteins, assuming 100% digestion efficiency. Incomplete digestion might have resulted in some inaccuracies, however, low intra- and inter-assay variability for the 54 proteins used indicate high reproducibility and precise comparison between samples.

7.3.5 Study Design and Population

The study design was cross-sectional hence no cause-effect relationships could be concluded from the associations observed. The TNHS population was self-selected university students form one campus of the University of Toronto, hence it might not be representative of all young adults in Toronto, Canada or elsewhere. In addition to socioeconomic status and education of the population, an unequal number of each ethnocultural groups was presented in the study which might have affected the results. Furthermore, ethnicity was self-identified by participants and then grouped by study administrators into four general groups. This approach might have improperly grouped ethnocultural groups together compared to more objective approaches that use DNA-linked ancestry and ethnicity categorization \(^2^2^0\).
7.4 Future Research

The findings of this research indicate that 9p21 genotype is associated with insulin markers as well as some plasma proteins, and that these associations might vary according to sex, ethnicity, HC use and dietary patterns in a group of young adults. These novel findings, along with the lack of associations with the many of the traditional CVD markers, may suggest unique and alternative mechanisms by which 9p21 genotype is associated with CVD events that have been reported in many other studies globally.

Further research could build on the current findings to discover new mechanisms and biomarkers of CVD that might aid in better prediction and treatment of CVD compared to those currently in clinical use. Finding additional pathways in which 9p21 interacts with CVD development early in life would be an important discovery to aid in personalized medicine and nutritional interventions that would be most effective for carriers of 9p21 risk alleles. This type of targeted approach would likely yield better results in preventing CVD events if combined with current approaches, given the high population attribution of this genotype for CVD development.
7.5 Implications

The most important implication of this study is how 9p21 genotype is associated with insulin markers at a young age, and that this association is modifiable by a Prudent Dietary pattern only in those carrying the highest genetic risk. Genetic predisposition to CVD is often associated with earlier onset of CVD events and the findings of this research support this in the most robust and common genotype associated with CVD. This is especially important in the absence of other CVD markers because this group would likely be missed in clinical risk screening for CVD. As demonstrated by past studies, and confirmed in the current thesis, the gene-diet interaction might indicate specific pathways that are modulated by dietary components present in fresh fruits and vegetables, suggesting personalized interventions that might be most effective for high risk allele carriers of 9p21.

In addition, the association of 9p21 genotype with plasma proteins is a novel finding that could be built on in order to discover clinically relevant biomarkers of CVD as they relate to this genotype. Another important implication of the study is the differences observed in specific plasma protein associations between the ethnicities presented in this study. This along with the sex and HC use differences seen in insulin markers suggest important of considering these variables in clinical and epidemiological findings.
7.6 Thesis Conclusion

In summary, this thesis examined the association between 9p21 genotype and early and novel biomarkers of CVD, while assessing the potential effects of HC use, sex, and ethnocultural variations on these relationships. We identified sex, and HC use codified the association between 9p21 genotype and fasting insulin, regardless of ethnicity. In addition, habitual consumption of a high Prudent dietary pattern appeared to be attenuate an association between homozygous risk alleles of 9p21 and risk of glycemic dysregulation. We also identified four plasma proteins to be associated with 9p21 genotype, while ethnocultural variation affected these relationships, sex, and HC use did not meaningfully interact with these relationships.
References


43. Janus A, Szahidewicz-Krupska E, Mazur G, Doroszko A. Insulin resistance and endothelial dysfunction constitute a common therapeutic target in cardiometabolic


68. Manolio TA, Savage PJ, Burke GL, et al. Association of fasting insulin with blood


75. Wang F, Han L, Hu D. Fasting insulin, insulin resistance and risk of hypertension in the


doi:10.3945/ajcn.111.022657.


111. Broadbent HM, Peden JF, Lorkowski S, et al. Susceptibility to coronary artery disease and


132. Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. *Curr Opin*


188. Levy AP, Hochberg I, Jablonski K, et al. Haptoglobin phenotype is an independent risk


doi:10.1074/jbc.M702163200.


