

**A GENETIC INVESTIGATION OF THE ECTONUCLEOTIDE  
PYROPHOSPHATASE/PHOSPHODIESTERASE 1 (*ENPP1*) VARIANTS WITH  
DIABETES AND GLYCEMIA TRAITS IN  
AFRO-CARIBBEAN MEN**

by

Tennille Shintana Leak

B.S. in Biology, High Point University, 1999

Ph.D. in Molecular Genetics and Genomics, Wake Forest University School of Medicine, 2008

Submitted to the Graduate Faculty of

Epidemiology

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2010

UNIVERSITY OF PITTSBURGH

Graduate School of Public Health

This thesis was presented

by

**Tennille S. Leak**

It was defended on

November 11, 2010

and approved by

**Thesis Chair:**

Joseph M. Zmuda, Ph.D.  
Associate Professor, Departments Epidemiology/Human Genetics  
Graduate School of Public Health  
University of Pittsburgh

**Committee Member:**

Anne B. Newman, M.D., M.P.H.  
Chair, Department of Epidemiology  
Professor of Epidemiology and Medicine  
Graduate School of Public Health  
University of Pittsburgh

**Committee Member:**

Candace M. Kammerer, Ph.D.  
Associate Professor, Department Human Genetics  
Graduate School of Public Health  
University of Pittsburgh

Copyright © by Tennille S. Leak

2010

**A GENETIC INVESTIGATION OF THE ECTONUCLEOTIDE PYROPHOSPHATASE/  
PHOSPHODIESTERASE 1 (*ENPP1*) VARIANTS WITH DIABETES AND GLYCEMIA  
TRAITS IN AFRO-CARIBBEAN MEN**

Tennille S. Leak, M.S.

University of Pittsburgh, 2010

Ectonucleotide pyrophosphatase/phosphodiesterase 1, which downregulates insulin signaling by inhibiting insulin-receptor tyrosine kinase activity, is encoded by the *ENPP1* gene. Common variants in *ENPP1* have been associated with body mass index (BMI), diabetes and glycemia related traits in populations of European Ancestry, but data in African ancestry populations are sparse. Our objective was to evaluate common *ENPP1* variants for association with diabetes and glycemia related traits in a high risk Afro-Caribbean population. Thirty-four single nucleotide polymorphisms (SNPs) based on pair-wise tagging ( $r^2 \geq 0.80$ ;  $MAF \geq 0.05$ ) were successfully genotyped in 380 cases and 1,455 controls without diabetes. Associations with BMI, fasting glucose, fasting insulin and HOMA-IR were also analyzed in non-diabetic controls. The most interesting association was observed between rs1044498 K121 allele and lower BMI (age-adjusted  $P = 0.018$ ). Nominal associations were observed with *ENPP1* SNPs and fasting glucose (age- and BMI-adjusted  $P=0.001$  to  $0.020$ ). Also, six SNPs showed nominal evidence for association ( $P \leq 0.05$ ) with diabetes in one or more genotypic model. The most significant associations were observed with SNPs in intron 11 (rs17060836; OR=1.32 [1.04-1.67]; dominant  $P = 0.019$ ), two SNPs in intron 1 (rs703184, rs7749493; OR= 0.78 to 1.39) and three SNPs in the 3' untranslated region (rs7754561, rs7769993 and rs9373000; OR = 0.69-1.38). In this population of Afro-Caribbean men, the *ENPP1*-rs1044498 the K121 allele and intronic variants may

modulate BMI and glucose. Also, variants in the 3'UTR confer an increased risk of developing diabetes confirming and extending reports in European and African Americans. After accounting for multiple testing, we conclude that *ENPP1* is not a major contributor to diabetes related traits; nevertheless, our results reveal that variants in the *ENPP1* gene may modulate BMI and maintain glucose homeostasis in this population of Afro-Caribbean men. These studies are of public health relevance or importance because they contribute epidemiologic information to the genetic etiology of type 2 diabetes in men of African ancestry.

## TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	X
1.0 INTRODUCTION.....	1
1.1 DIABETES.....	1
1.2 GENETICS OF TYPE 2 DIABETES.....	2
1.3 IDENTIFYING GENES FOR COMPLEX DISEASES.....	3
1.4 ECTONUCLEOTIDE PYROPHOSPHATASE/PHOSPHODIESTERASE 1.....	3
1.5 SINGLE NUCLEOTIDE POLYMORPHISMS.....	4
1.6 INTERPRETATION OF ASSOCIATION RESULTS.....	5
1.7 LINKAGE DISEQUILIBRIUM AND HAPLOTYPES.....	5
1.8 ADMIXTURE.....	6
2.0 GOALS AND SPECIFIC AIMS .....	7
2.1 SPECIFIC AIM 1: TO COMPREHENSIVELY INVESTIGATE <i>ENPP1</i> VARIANTS FOR ASSOCIATION WITH BODY MASS INDEX, FASTING GLUCOSE, FASTING INSULIN AND INSULIN RESISTANCE IN NON-DM CONTROLS.....	7
2.2 SPECIFIC AIM 2: EVALUATE K121Q VARIANT AND COMMON <i>ENPP1</i> VARIANTS FOR ASSOCIATION WITH T2D.....	7
3.0 RESEARCH DESIGN AND METHODS.....	8
3.1 STUDY POPULATION.....	8
3.2 BLOOD SAMPLE COLLECTION AND GLUCOSE MEASUREMENT.....	8
3.3 SNP SELECTION .....	9
3.4 DNA EXTRACTION AND GENOTYPING.....	10
3.5 STATISTICAL METHODS.....	11

3.5.1 Correction for Multiple Testing.....	12
4.0 RESULTS.....	13
4.1 POPULATION CHARACTERISTICS.....	13
4.2 SNP ASSOCIATION ANALYSES.....	13
4.2.1 Hardy-Weinberg Equilibrium and Linkage Disequilibrium.....	13
4.2.2 Association Analyses with Analyses with BMI in Controls.....	15
4.2.3 Association Analyses with Analyses with Glucose in Controls.....	15
4.2.4 Association Analyses with Insulin Levels and Insulin Resistance Estimated by HOMA in Controls.....	15
4.2.5 Association Analyses with Type 2 Diabetes.....	16
5.0 DISCUSSION.....	17
6.0 CONCLUSIONS AND PUBLIC HEALTH RELEVANCE.....	21
APPENDIX: TABLES.....	25
BIBLIOGRAPHY.....	40

## LIST OF TABLES

Table 1. Characteristic of the Afro Caribbean Men.....	30
Table 2. Nominally Significant ( $P<0.05$ ) Single-SNP Genotypic Tests of Association with BMI, Fasting Glucose, Insulin Resistance using HOMA-IR and Fasting Insulin in Non-Diabetic Controls.....	31
Table 3. Nominally Significant ( $P<0.05$ ) Single-SNP Genotypic Tests of Association with Type 2 Diabetes.....	32
Table 4. Single SNP Association Results for BMI ( $\text{kg}/\text{m}^2$ ) .....	33
Table 5. Single SNP Association Results for Fasting Glucose ( $\text{md}/\text{dl}$ ) .....	35
Table 6. Single SNP Association Results for Fasting Insulin ( $\mu\text{U}/\text{ml}$ ).....	37
Table 7. Single SNP Association Results for Insulin Resistance Estimated by HOMA.....	39
Table 8. Single SNP Genotypic Association Tests with T2D.....	41
Table 9. Genotype Frequencies in Afro-Caribbean Cases and Controls.....	43



## LIST OF FIGURES

Figure1. Linkage Disequilibrium structure of <i>ENPP1</i> in Afro-Caribbean Controls.....	14
Figure 2. Linkage Disequilibrium structure of <i>ENPP1</i> in Afro-Caribbean Cases.....	14
Figure 3. Progression and Direction for Genomic Studies and their Implications.....	24

## ACKNOWLEDGEMENTS

*‘When you are driven by PURPOSE and it’s stronger than PASSION, there is always a PRICE’*

I would like to thank and acknowledge my advisor Dr. Zmuda, for his patient supervision, dedication and support during this work. Thank you for being a superb advisor. I thank each of my committee members: Dr. Newman and Dr. Kammerer for their additional support and guidance during my post-graduate training. Additionally, I acknowledge the faculty and staff of the Epidemiology program. I would like to thank all of my family and friends for their support and prayers throughout my post-graduate school career. I extend a very special thank you to my church family especially my spiritual parents Bishop and Mother Frederick D. Patterson. Thank you both for the hugs and words of encouragement. I would also like to extend the biggest thanks and admiration to my parents, Naja Leak and Lance Addison, my siblings Torre, Shatara and aunties girls Chloe Sky and Gabrielle London, I love you all dearly. I am deeply appreciative for my King in Waiting, Stanley A. Johnson, Jr.; you have been an incredible source of encouragement, love and support throughout this entire process. I love you! Last but not least, I dedicate this thesis to my grandparents Evangelist Marie and Mr. James Oscar Leak, thank you for the love and the determination to press on. Grandpa Leak although you are not here physically I know you are with me spiritually, therefore I dedicate this thesis to you. Thank you for always being there.

## **1.0 INTRODUCTION**

### **1.1 DIABETES**

Diabetes is a disease that is characterized by high blood glucose levels due to the body's inability to produce insulin, and/or the body's inability to respond properly to insulin. Clinically, diabetes is diagnosed if an individual has a fasting blood glucose level of 126 mg/dl or higher using the Fasting Plasma Glucose Test (FPGT), or a two-hour blood glucose level greater than or equal to 200 mg/dl using the Oral Glucose Tolerance Test (OGTT) <sup>1</sup>. Diabetes is a growing epidemic in the United States and worldwide. It is currently estimated that 20.8 million people, or 7% of the total US population, have diabetes, of which 6.2 million (nearly one-third) are undiagnosed (2007, Kaiser Family Foundation). In 2005 there were 1.5 million new cases of diabetes in Americans aged 20 years and older <sup>2</sup>. Diabetes is the sixth leading cause of death in the United States and has a major impact on both the economy and health <sup>3</sup>. In 2002, diabetes costs in the United States totaled \$132 billion in both direct and indirect costs <sup>4</sup>. Diabetes and complication due to diabetes can greatly affect quality of life. Diabetes can reduce one's life expectancy 5-10 years, increases the risk for stroke 2-4 times, and is the leading cause of kidney failure, non-traumatic lower-limb amputations, and blindness <sup>2</sup>.

There are multiple types of diabetes including type 1, type 2, and gestational diabetes. The most prevalent form of diabetes, type 2 diabetes mellitus (T2D), accounts for 90-95% of all cases of diabetes. T2D is a complex disease with both genetic and environmental influences. Risk factors for developing T2D include older age, obesity, physical inactivity, and a family history of diabetes <sup>4</sup>. Additionally, certain ethnicities such as African Americans, Hispanic/Latino Americans, and American Indians, have a higher risk for developing T2D. In the United States,

3.2 million African Americans aged 20 years and older, or 13.3% of the nation's African American population, have diabetes (2007, Kaiser Family Foundation). On average, an African American is nearly two times more likely to develop diabetes as compared to a European American of similar age <sup>5</sup>. African Americans are also at increased risk for complications associated with diabetes, including amputation rates (1.2-1.5 times higher) <sup>6</sup>, and kidney disease, including end-stage renal disease (2.6-5.6 times greater) <sup>7</sup>, as compared to European Americans. Mortality rates, due to diabetes, are also higher for African American as compared to European Americans. In 2003, diabetes accounted for 49.2 deaths per 100,000 population in African Americans as compared to 23.0 deaths per 100,000 population in Caucasians <sup>8</sup>.

## **1.2 GENETICS OF TYPE 2 DIABETES**

Genetic epidemiological studies support evidence that several variables contribute to T2D risk, including multiple genetic loci with small effect, environmental factors, and genetic heterogeneity <sup>9</sup>. Evidence of a genetic contribution to T2D susceptibility has been established through twin <sup>10</sup>, familial aggregation <sup>11</sup>, and segregation studies <sup>12</sup>. Barnett *et al*, 1981 showed that 48 of 53 middle aged or elderly monozygotic twins were concordant for T2D (>90% concordant rate). Of the five discordant twin pairs, all of the unaffected twins were beginning to show metabolic abnormalities (i.e. reduced insulin response to glucose) at the time of the report <sup>13</sup>. Goodman and Chung, 1975 calculated that the heritability ( $h^2$ ) for diabetes in individuals aged 40 years and older (assuming a multifactorial mode of inheritance) was 0.66 <sup>12</sup>. This estimate was supported by  $h^2$  values of 0.70 for T2D by Rich (1990), who also reported risk ratios of 12.3, 3.5, and 1.9 for monozygotic twins, first degree relatives, and second degree

relatives of probands with T2D; respectively <sup>9</sup>. Similarly, familial aggregation studies reveal that relatives of African American probands with T2D had a 2.95 fold increased prevalence of diabetes (95% CI = 1.55-6.62) as compared to relatives of unaffected individuals <sup>11</sup>.

### **1.3 IDENTIFYING GENES FOR COMPLEX DISEASES**

Until the recent use of Genome Wide Association Studies (GWAS), the two commonly used methods for identifying genes for complex diseases were the candidate gene approach and the positional cloning approach. The candidate gene approach evaluates variants in genes of known function and relevance to a particular phenotype. For example, logical candidate genes for T2D susceptibility have included genes related to insulin (e.g. the genes encoding the insulin protein or the insulin receptor protein) or genes involved in glucose homeostasis (e.g. glucose transporters or enzymes such as glucokinase). In contrast, the positional cloning approach evaluates genes based on their physical location in the genome alone, independent of their function. To date, the candidate gene approach has identified several genes that are involved in the etiology of T2D susceptibility: *INSR* (insulin receptor) <sup>14</sup>, *IRS1* (insulin receptor substrate 1) <sup>15</sup>, *PPARG* (peroxisome proliferator-activated receptor gamma) <sup>16</sup>, *PTPNI* (protein tyrosine phosphatase 1) <sup>17</sup> and *ENPP1* (ectonucleotide pyrophosphatase/phosphodiesterase family member 1) <sup>10</sup>.

#### **1.4 ECTONUCLEOTIDE PYROPHOSPHATASE/PHOSPHODIESTERASE 1**

Ectonucleotide pyrophosphatase/phosphodiesterase 1 (ENPP1) downregulates insulin signaling by inhibiting insulin-receptor tyrosine kinase activity <sup>18</sup>. In addition, ENPP1 has been shown to be involved in regulation of adipocyte maturation and highly regulated during adipogenesis thus

suggesting a role of insulin resistance in the absence of obesity<sup>19</sup>. The gene that codes *ENPPI* spans over 83 kb and is located on chromosome 6q22-23. Previous studies of variant rs1044498, located in exon 4 which causes an amino acid change from lysine to glutamine at codon 121 (K121Q)<sup>20</sup>, has shown positive associations with diabetes<sup>21,22 23-26</sup>, insulin resistance<sup>27</sup>, glucose homeostasis<sup>28</sup>, metabolic syndrome<sup>29</sup>, obesity<sup>30-32</sup> and diabetic nephropathy<sup>33-35</sup>. However, negative association results question the reproducibility of these positive findings<sup>36-38</sup>. In support of a potential role of *ENPPI* and diabetes risk, functional studies have demonstrated that mutations in the gene result in overexpression of the glycoprotein leading to decreased insulin receptor tyrosine kinase activity<sup>39-41</sup>.

### **1.5 SINGLE NUCLEOTIDE POLYMORPHISMS**

Association studies do not require the analysis of large pedigrees and are generally conducted in affected and unaffected (case-control) sample populations. These studies compare the prevalence of genetic markers, usually single nucleotide polymorphisms (SNPs), to determine if there is a statistically significant difference in the number of observed alleles for a marker in cases (affected individuals) as compared to controls (unaffected individuals). SNPs are mostly commonly used in genetic association studies because they occur more frequently in the human genome as compared to other polymorphisms such as microsatellites and insertions/deletions. Although humans are roughly 99.9% identical, SNPs with a minor allele frequency greater than 10% are present approximately every 600 base pair in the human genome<sup>42</sup>. Due to the collaborative efforts in human genetics research, databases such as NCBI's dbSNP and the International HapMap project, provide data for more than 10 millions SNPs.

The goal of the international HapMap project is to provide a catalog of the common genetic variants in humans. In 2001, the HapMap project released its phase I data for over 1 million SNPs <sup>43</sup>, with hopes of genotyping and releasing data for 4.6 million additional SNPs in phase II of the project. The HapMap project provides genotypic data for four ethnically diverse populations. These populations include 90 Yoruban individuals (thirty trios) from Ibadan, Nigeria (YRI), 45 unrelated Japanese individuals from Tokyo, Japan, 45 unrelated Han Chinese individuals from Beijing, China, and 90 (thirty trios) United States residents from Utah (CEPH), with northern and western European ancestry. The availability of SNP data from databases such as the International HapMap project continue to greatly advance SNP selection strategies for association studies in populations of all ethnicities.

### **1.6 INTERPRETATION OF ASSOCIATION RESULTS**

Significant results from an association study may be interpreted several ways. One possibility is that the positive association is a false positive. False positives may arise due to poor genotyping quality, population substructure, or admixture. A second alternative possibility may be that the association observed is a true association for which the actual causal variant for the disease is the variant being analyzed. This could be a result of an effective SNP selection process or due to chance alone. A third possibility may be that the associated SNP represents true association, although the genotyped SNP is in linkage disequilibrium (LD) with the actual causal variant.

### **1.7 LINKAGE DISEQUILIBRIUM AND HAPLOTYPES**

Linkage Disequilibrium (LD) is defined as the non-random assortment between alleles at two or more loci. In other words, markers are considered to be in LD when an allele at one locus is

found together with an allele at a second locus more often than if they were segregating independently. Generally LD will be greater if the loci are in close proximity to one another, although LD may be observed for SNPs separated by large distances. The extent of LD varies across the genome as well as from one ethnic group to another, as described by Gabriel *et al*, 2002, where they noted that the mean block size in Yoruban Nigerians and African American samples was 11 kb, compared with 22 kb observed in European and Asian samples. LD block size varied from <1 to 94 kb in the Yoruban Nigerians and African Americans and <1 to 173 kb in the Europeans and Asians <sup>44</sup>.

## **1.8 ADMIXTURE**

Association studies of unrelated individuals are widely used in order to identify susceptibility loci; even though it is not usually addressed, population substructure can produce false association results. Admixture is defined as the intermixing of once isolated populations <sup>45</sup>. In the U.S., African Americans and Hispanic Americans are the two largest admixed populations. Genetic studies in admixed populations such as Hispanic populations and African American are presented with the challenge of addressing the contributions due to admixture. In order to estimate the genetic contribution of ancestral populations to admixed individuals or populations one can use genetic markers known as ancestry informative markers (AIMs) <sup>46</sup>. AIMs have also been used to determine whether a given trait is associated with ancestral proportions. Addressing population substructure is necessary for tests of genetic association in order to account for spurious associations.



## **2.0 GOALS AND SPECIFIC AIMS**

Although diabetes is a growing epidemic in the African American community, there are few research efforts/projects that address the genetic contributions to the disease in populations with very little admixture. The overall goal of this project explores possible genetic mechanisms of the *ENPP1* gene with diabetes and glycemic related traits in an Afro Caribbean population from Tobago. The working hypothesis of this study was that the candidate gene approach would successfully identify variants in *ENPP1* that contribute to these diseases with complex etiologies. Elucidation of genetic components of diabetes is anticipated to drive efforts toward a mechanistic understanding of the disease and its complications.

### **2.1 SPECIFIC AIM 1: TO COMPREHENSIVELY INVESTIGATE *ENPP1* VARIANTS FOR ASSOCIATION WITH BODY MASS INDEX, FASTING GLUCOSE, FASTING INSULIN AND INSULIN RESISTANCE IN NON-DM CONTROLS.**

We aim to replicate positive association results in a subset of Afro-Caribbean controls from the Tobago Health Study Cohort, consisting of 1,454 healthy individuals.

### **2.2 SPECIFIC AIM 2: EVALUATE K121Q VARIANT AND COMMON *ENPP1* VARIANTS FOR ASSOCIATION WITH T2D.**

We will use existing data. SNPs will be genotyped in 384 diabetic (cases) individuals and 1454 healthy (controls) individuals African origin men aged 40+ in the Tobago Health Study Cohort and analyzed for single SNP genotypic association. The most strongly disease-associated variant(s) will be identified.

### **3.0 RESEARCH DESIGN AND METHODS**

#### **3.1 STUDY POPULATION**

Three-thousand two-hundred men aged  $\geq 40$  years were recruited on the Caribbean island of Tobago between 1998 and 2003 for a population-based prostate specific antigen screening study among previously unscreened men<sup>47</sup>. The population is predominantly of West African origin. Ancestry informative molecular markers have estimated 94% African, 4.6% European and 1.4% Native American ancestry in this population<sup>48</sup>. In brief, recruitment was accomplished by word of mouth, hospital flyers and radio broadcasting. To be eligible, men had to be 40 years and older, ambulatory and not terminally ill. Approximately 60% of all age-eligible men on the island participated and participation was similar across the island Parishes. Questionnaires were administered to obtain information on demographic characteristics, occupation, medical history, and lifestyle related factors. All men who participated in the baseline exam were invited to participate in a follow-up clinic exam between 2004 and 2007 and 2,031 men in the cohort (70% of survivors) and 451 new participants completed the visit. Written informed consent was obtained using forms approved by the Institutional Review Boards of the University of Pittsburgh and the Tobago Ministry of Health.

#### **3.2 BLOOD SAMPLE COLLECTION AND GLUCOSE MEASUREMENT**

Blood samples were obtained by venipuncture in the morning after a 12-h fast and stored at  $-80^{\circ}\text{C}$  until assays were completed. Serum glucose was measured by using an enzymatic procedure<sup>49</sup>. An RIA procedure developed by Linco Research, Inc. was used to measure insulin. To quantify insulin resistance we used the homeostatic model assessment (HOMA) equation (fasting glucose \* fasting insulin/22.5)<sup>50</sup>.

### 3.3 SNP SELECTION

We used a two-stage strategy to prioritize SNPs for genotyping. First, we employed a tagging SNP approach to capture common genetic variation across the *ENPP1* gene region. We initially identified 149 SNPs across the *ENPP1* gene region (including 10kb downstream and 10kb upstream of the transcript) using publicly available SNP data from Phase II of the International HapMap project (<http://www.hapmap.org>) that were obtained using samples from the Yoruban (YRI) population in Ibadan, Nigeria. A subset of informative SNPs was then selected from this larger reference SNP panel using a pair-wise correlation method with  $r^2 \geq 0.80$  and minor allele frequency (MAF)  $\geq 0.05$  using the program HClust<sup>51</sup>. Using this approach, we identified 52 SNPs (36 singletons and 16 tagging SNPs) from the 91 SNPs in the reference SNP panel that had a MAF  $\geq 0.05$ . We then used the Function Analysis and Selection Tool for Single Nucleotide Polymorphisms (FASTSNP) (<http://fastsnp.ibms.sinica.edu.tw>)<sup>52</sup> and the EIDorado tool (version 4.5) from the Genomatix software package (Genomatix Suite release 3.4; <http://www.genomatix.de>) to prioritize the 36 singleton SNPs for genotyping based on their predicted functional effects. FASTSNP identifies SNPs that may alter: 1) the amino acid sequence of the encoded protein to one with different structural characteristics or a premature termination of an amino-acid sequence; 2) an exonic splicing enhancer/silencer binding site in a coding sequence that may affect splicing regulation; 3) a consensus splicing site sequence; 4) a putative binding site for a transcription factor in the promoter or an intronic enhancer region; or 5) a 3' untranslated region motif likely to be involved in post-transcriptional regulation. The EIDorado tool was used to identify SNPs that may create or abolish a putative transcription

factor binding site in the promoter/regulatory region of genes. Singleton SNPs that were not predicted to be potentially functional were not genotyped. This approach yielded a total of 38 SNPs (4 SNPs in promoter region, 1 intergenic, 2 coding, 2 intron boundary, 26 non-coding and 3 in the 3'UTR). Average SNP density was 1 SNP every 3.3 kb, with the largest gap 19.6 kb and the smallest 55 bp.

### **3.4 DNA EXTRACTION AND GENOTYPING**

Genomic DNA was extracted from 5 ml frozen blood clots (baseline sample) or whole blood (follow-up) using the Qiagen DNA Blood Kit (Qiagen, Inc., Valencia, CA). One hundred nineteen bi-allelic Admixture Informative Markers (AIMs), were selected to maximize European and African allele frequency differences<sup>53</sup>. PCR primers were purchased from Invitrogen (Carlsbad, CA) and genotyping was performed using Sequenom MassARRAY iPLEX Gold technology (Sequenom, Inc; San Diego, CA). The average genotyping completeness rate was 96.3%. The average genotyping consensus rate among ~7% blind replicate samples was 98.4%.

Genotyping of *ENPP1* SNPs was performed using pre-designed TaqMan SNP genotyping assays (Applied Biosystems). Genotyping was completed according to the manufacture's protocol on a 7900HT Fast Real-Time PCR system (Foster City, CA). The reactions were cycled with standard TaqMan conditions (50°C for 2 min hold, 95°C for 10 min hold, 95°C for 15 sec and 60°C for 1 min for 40 cycles and then cool down to 4°C). The genotypes were called with the Applied Biosystems SDS 2.2.2 software package. Final genotyping results were evaluated manually by

two trained technicians who were blinded to case-control status. Four SNPs (rs9493105, rs7773477, rs9493116 and rs9402345) were dropped from further analysis due to low genotyping success rates. Our final marker set consisted of 34 SNPs. The average genotyping completeness rate was 95.9%. The average genotyping consensus rate among 5% blind replicate samples was 99.4%. Using the program Tagger, we estimate that the 34 genotyped SNPs captured 77% of the common variation in the *ENPP1* gene region<sup>54</sup>.

### 3.5 STATISTICAL METHODS

Characteristics of cases and controls were tested for significant differences using unpaired Student's t-tests. Hardy-Weinberg equilibrium (HWE) was assessed using the  $\chi^2$  goodness-of-fit statistic ( $P \leq 0.001$ ). Haplotype block structure was established using Haploview 4.2<sup>55</sup>, using the block definition from Gabriel *et al.*<sup>56</sup>. Haplotype frequencies were estimated using the EM algorithm implemented in the software program Dandelion 3.0 using a permutation test of the multinomial likelihood ratio statistic using 1000 permutations<sup>57</sup>.

To model population substructure and obtain a multidimensional 'ancestry' score for each individual, principal component analyses which is especially sensitive to subtle population structure was used<sup>58-61</sup>. Eigenvectors was developed from these SNPs and plotted against each other to evaluate substructure and outliers. Our experience with black populations after removal of 'outlier' indicates that 2-3 PCs are sufficient, especially to distinguish between ethnic groups. The present analysis is limited only to men at the follow-up visit with the comprehensive data on

fasting glucose and insulin and who were not outliers based on 130 African Ancestry informative molecular markers (N=120). Among these men, 380 had diabetes defined as a fasting serum glucose concentration  $\geq 126$  mg/dL or current use of antidiabetic medication and 1,455 Afro-Caribbean controls without a current diagnosis of diabetes were used for the current analysis.

Tests of association under the three *a priori* genetic models (additive, dominant, and recessive) are reported. Due to a lack of validity of the large sample  $\chi^2$  test statistic, only the dominant model was considered for SNPs with ten or fewer individuals that were homozygous for the minor allele. Quantitative traits (insulin, and homeostatic model assessment of insulin resistance (HOMA-IR)) association analysis for every SNP in the control population was performed using a series of analysis of variance (ANOVA) tests implemented in QSNPGWA<sup>62</sup>.

Unadjusted odds ratios (ORs) and 95% confidence intervals (CI) in the case-control sample were computed using SNPGWA to test all SNPs for genotypic association with diabetes<sup>63</sup>.

SNPs that showed nominal evidence for association were further adjusted for age- and body mass index (BMI). A *P* value of  $<0.05$  was considered significant. All adjusted multivariable linear regression and unconditional logistic regression analyses were performed using Stata 10 (College Station, TX).

### **3.5.1 Correction for multiple testing**

To correct for multiple tests at the gene level the conservative Bonferonni method was used, with a *P*-value  $<0.0004$  considered significant evidence for association when assuming independence based on linkage disequilibrium (LD).

## 4.0 RESULTS

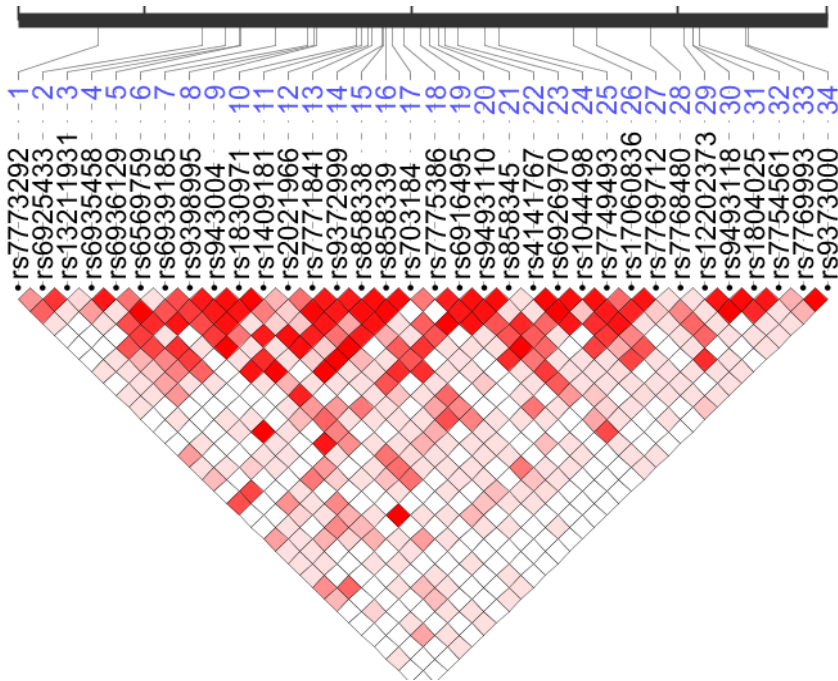
### 4.1 POPULATION CHARACTERISTICS

Baseline characteristics of the Afro-Caribbean cases and controls are summarized in Table 1. The mean age of examination for the non-diabetic controls was 5 years younger than the mean age of those with diabetes ( $P < 0.0001$ ). Additionally, the diabetes cases were more insulin resistant and hyperglycemic ( $166.6 \pm 67.4$  vs.  $91.7 \pm 11.6$ ) compared to non diabetic controls ( $P < 0.0001$ ).

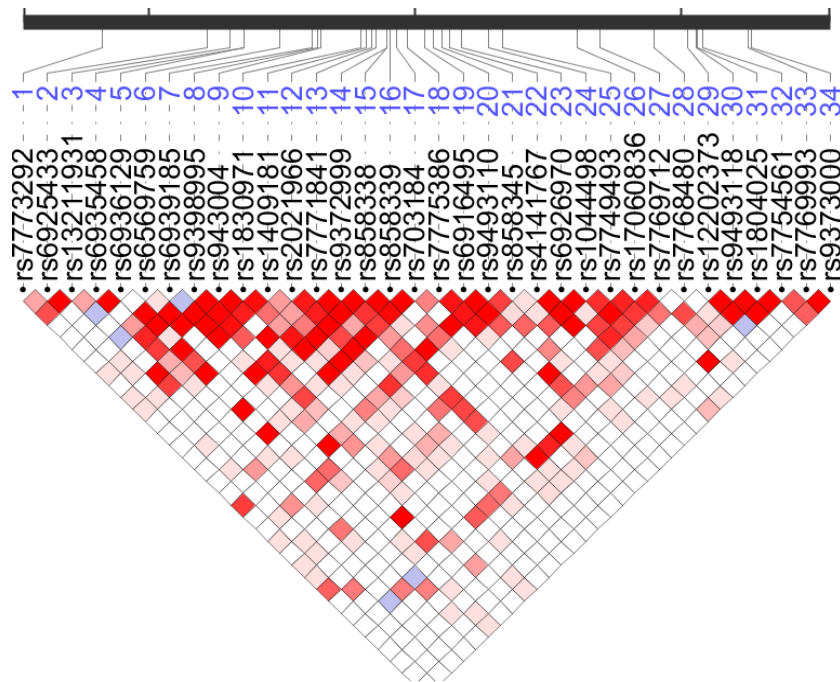
### 4.2 SNP ASSOCIATION ANALYSES

#### 4.2.1 Hardy-Weinberg Equilibrium and linkage disequilibrium

The genotypic distribution of all *ENPP1* variants in cases and controls was in Hardy-Weinberg equilibrium (all  $P > 0.001$ ). The linkage disequilibrium (LD) structure present in controls (Figure 1) and cases (Figure 2) was compared by evaluating the total number of blocks and average block size (kb) between cases and controls. Using the Confidence Interval method<sup>56</sup> implemented in Haploview (with default settings), there were 9 blocks in cases (average block length 2.9 kb) and 9 LD blocks in controls (average block size of 2.8kb). Figures 1 and 2; respectively.



**Figure 1. Linkage Disequilibrium Structure of *ENPPI* in Afro-Caribbean Controls (n=1,455).** LD determined using Confidence Interval method<sup>56</sup> implemented in Haploview 4.2 (with default settings), with blocks extended if pairwise  $D' > 0.80$ . Red squares indicate high pairwise LD ( $D' = 1.00$ ), gradually coloring down to white squares of low pairwise LD.



**Figure 2. Linkage Disequilibrium Structure of *ENPPI* in Afro-Caribbean Cases (n=380).** LD determined using Confidence Interval method<sup>56</sup> implemented in Haploview 4.2 (with default settings), with blocks extended if pairwise  $D' > 0.80$ . Red squares indicate high pairwise LD ( $D' = 1.00$ ), gradually coloring down to white squares of low pairwise LD.



#### **4.2.2 Association Analyses with Analyses with BMI in Controls**

One intron 1 SNP rs703184 were nominally associated with higher BMI (Table 2). SNP rs703184 was nominally associated with higher BMI ( $30.6 \pm 10.5$  kg/m<sup>2</sup> for “CC” genotype compared with  $27.1 \pm 4.5$  kg/m<sup>2</sup> for “GG”/“GC”; age-adjusted  $P=0.006$ ). In contrast, rs1044498 (minor “A” allele corresponding to the K121 allele) was nominally associated with lower BMI ( $25.9 \pm 3.9$  kg/m<sup>2</sup> for “AA” genotype compared with  $27.2 \pm 4.6$  kg/m<sup>2</sup> for “CC”/“CA”; age-adjusted  $P=0.018$ ). Detailed association results for BMI are presented in Table 4.

#### **4.2.3 Association Analyses with Analyses with Glucose in Controls**

Six SNPs showed nominal evidence for association with fasting glucose in one or more genotypic model (Table 2). Four intronic SNPs (rs858345, rs6569759, rs1830971, rs1409181, rs17060836) and rs7768480, located in intron 20–exon 21 boundary (age-and BMI adjusted  $P$  values ranging from 0.001 to 0.020). The most significant association was observed for rs6569759, located in intron 1 being associated in all three genotypic models (age-and BMI adjusted  $P = 0.003, 0.001, 0.009$  for the dominant, additive, and recessive models. Further results for glucose are presented in Table 5.

#### **4.2.4 Association Analyses with Insulin Levels and Insulin Resistance Estimated by HOMA in Controls**

Association analyses revealed one SNP nominally associated with insulin resistance estimated by HOMA-IR and two SNPs associated with fasting insulin levels (Table 2). The homozygote recessive genotype (“TT”) for SNP rs858345, located in intron 1, was nominally associated with decreased insulin resistance (unadjusted  $P=0.043$ ). The most significant association was between

insulin and SNPs rs6935458 and rs9372999, located in the promoter region and intron 1, respectively. As shown in Table 2, rs9372999 was nominally associated with increased fasting insulin ( $13.3 \pm 8.4$   $\mu\text{U/ml}$  for recessive “AA” genotype compared with  $12.2 \pm 6.9$   $\mu\text{U/ml}$ ) for “CC”/“CA”;  $P=0.044$ ). In contrast, SNP rs6935458 was nominally associated with lower fasting insulin (dominant  $P=0.033$ ;  $11.6 \pm 5.9$   $\mu\text{U/ml}$  for the “AA” genotype vs.  $12.2 \pm 6.4$   $\mu\text{U/ml}$  for the “AG”/“GG” genotypes). Association results for insulin levels and insulin resistance are presented in Tables 6 and 7; respectively.

#### **4.2.5 Association Analyses with Type 2 Diabetes**

Genotypic associations ( $P \leq 0.05$ ) were observed with SNPs rs703184 (dominant  $P=0.022$ , adjusted  $P_a = 0.068$ ); rs7749493 (dominant  $P=0.037$ ,  $P_a = 0.025$ ; additive  $P=0.019$ ,  $P_a = 0.014$ ), rs17060836 (dominant  $P=0.020$ ;  $P_a = 0.019$ ), rs7754561 (dominant  $P=0.020$ ,  $P_a = 0.034$ ); rs776993 (dominant  $P=0.025$ ;  $P_a = 0.032$ ) and rs9373000 (dominant  $P=0.008$ ,  $P_a = 0.012$ ; additive  $P=0.018$ ,  $P_a=0.023$ ). Single-SNP genotypic association results are presented in Table 8. Genotype frequencies and counts for each SNP are shown in Table 9.

## 5.0 DISCUSSION

We examined the association of common variants in *ENPP1* with BMI, glycemia traits and diabetes prevalence in an Afro-Caribbean population at high risk of diabetes<sup>64</sup>. We found that the widely studied rs1044498 minor K121 allele was associated with lower BMI among non-diabetic controls. In addition, we found that variants in intron 1 and the 3'UTR were associated with diabetes independent of age and BMI. Our findings confirm and extend previous reports of associations between common variants in *ENPP1* with obesity and diabetes.

Two reports to date have shown that the Q121 genotype is associated with lower BMI in non-diabetic populations of European and African American populations<sup>26,65</sup> whereas others have reported that the Q121 genotype is associated with higher BMI<sup>66,67</sup>. Given the discordant genotypic association between rs1044498 (K121Q) and BMI our association of K121 in our study population with lower BMI might be attributable to differences in genetic backgrounds between these populations.

The allele frequency of the minor K121 allele of rs1044498 in Afro-Caribbean controls (0.14) closely resembles the frequencies of the HapMap YRI (0.08) and previously reported African-American control and case subjects<sup>23,25,26</sup>. However, the frequency in the Afro-Caribbean Dominican Republic population (0.46) is dramatically higher which may reflect a higher degree of European admixture in this cohort than in the HapMap Nigerian population or our Afro-Caribbean population of Tobago<sup>24</sup>.

Nominal associations with intron 1 SNPs (rs6569759, rs1830971, rs1409181, rs858345) were observed with fasting glucose (Table 2). Jenkinson, et al.<sup>68</sup>, identified four SNPs (rs1800949, rs6918013, rs7756163, rs7754609) spanning intron 1 associated with fasting glucose in nondiabetic Mexican Americans from the San Antonio Family Diabetes Study (SAFDS) with P-values ranging from 0.0047-0.048. Bochenski<sup>69</sup> et al., revealed that rs997509, also located in intron 1 is not associated with diabetes in nonobese individuals from Poland, although nonsignificant ( $P > 0.40$ ) there is a 32% decreased risk of diabetes. However in obese subjects, rs997509 was found to be associated with a 4.7 fold increased risk of diabetes<sup>69</sup>. These observations suggest that polymorphisms in this region of ENPP1 may be heterogeneous and further exploration of the region is required.

Modest association with diabetes was detected with two intron 1 variants (rs703184 and rs7749493), an intron 11 variant (rs17060836) and three variants located in the 3' untranslated region (UTR) (rs7754561, rs7769993 and rs9373000). Men having at least one copy of the major allele for rs7749493 and rs7754561 had a 21 to 31% decreased risk of diabetes. In contrast, the major alleles for SNPs rs703184, rs17060836, rs9373000 and rs7769993 were associated with an increased risk of diabetes with ORs between 1.33-1.39. After adjusting for age and BMI, associations with all SNPs except rs703184 remained nominally significant. Additionally, haplotype analysis showed nominal evidence of association with 2 three-marker haplotypes (Supplementary Table 7). The three-marker haplotypes are located in the promoter region and the 3'UTR, with *P* values from 0.034 to 0.039; respectively.

In a recent report by Keene et al.<sup>25</sup> rs7754586 located approximately 48bp – 8.9 kb away from rs7754561, rs7769993 and rs9373000 was associated with increased diabetes risk (OR ranging from 1.34-1.50) in 577 African American cases with diabetes-end stage renal disease and 596 African American controls. Functional studies of a cluster of SNPs (rs1044548, rs11964389 and rs1044558) which make up a 3-SNP haplotype located in the 3'UTR performed by Frittitta et al.<sup>41</sup> demonstrated that chinese hamster ovary (CHO) cells transfected with the 3-SNP haplotype revealed overexpression of ENPP1 protein due to increased mRNA stability. We did not directly investigate these markers in our Afro-Caribbean population due to our SNP selection process. However, markers rs7754561, rs7769993 and rs9373000 fall within one 15 kb block of high LD consisting of 37 HapMap SNPs spanning the 3' UTR.

The replicated K121Q (rs1044498) SNP in European populations<sup>67,70-72</sup> was analyzed in our case-control population, however it was not associated with diabetes (*P* values ranging from 0.079 to 0.792). Although nonsignificant, the trend for rs1044498 risk is in the direction and the OR close to reported values in African Americans as reported by Keene et al.<sup>23,25,26</sup>.

ENPP1 is known to have potential pleiotropic effects on obesity and diabetes<sup>73</sup> thus; we performed our analyses with and without adjustments for BMI. We found that adjustments for BMI reduced the strength of the observed association between rs858345 with HOMA-IR and rs6935458 and rs9372999 with insulin. However, associations between several *ENPP1* SNPs with diabetes remained nominally significant after adjusting for age-and BMI. This observation suggests that at least some of the potential effect of *ENPP1* variants on diabetes susceptibility

may be mediated through mechanisms independent of obesity. Indeed, there is evidence that ENPP1 has a direct role in insulin signaling and insulin sensitivity <sup>74</sup>. We corrected our associations for multiple comparisons using the conservative Bonferroni method assuming independence based on LD. Although our study is the largest to date on ENPP1 and obesity and diabetes related measures in African ancestry individuals, none of the nominally significant associations surpassed the conservative Bonferonni threshold ( $P < 0.0004$ ). Larger studies of *ENPP1* in African ancestry individuals will be needed to replicate and extend our findings in this population group.

Overall, this study supports an association of the widely studied and replicated ENPP1-rs1044498 variant with lower BMI <sup>65</sup>. In addition, our analysis identified and confirmed associations with common variants in the 3'UTR with diabetes <sup>25</sup>. Denser SNP genotyping and direct sequencing along with functional studies investigating the 3'UTR of the *ENPP1* gene merits further examination to investigate the biological relevance of variations in this region.

## 6.0 CONCLUSIONS AND PUBLIC HEALTH RELEVANCE

The *ENPP1* gene was analyzed for association with T2D-ESRD in this study. To comprehensively evaluate variants in *ENPP1* for association with T2D-ESRD, 34 SNPs located in the coding and flanking regions of *ENPP1* were successfully genotyped in 380 Afro-Caribbean men with diabetes and 1,455 men without diabetes. Positive association was detected with eight SNPs, one located in exon 4, two in intron 1, one in intron 11 variant and three variants located in the 3'UTR. Seven of the associated SNPs had not been reported in diabetes and diabetes related association studies of *ENPP1*.

The close proximity of the two most significantly associated SNPs in the 3'UTR are in high LD ( $r^2 = 0.98$ ) and the prior identification of variants in 3'UTR provide evidence that variants in the distal region of *ENPP1* are important in relation to diabetes susceptibility in African ancestry population.

The contributions of associated SNPs have not been fully examined. Due to the frequency and overall odds ratio, it appears that these SNPs would only modestly account for the disparity of T2D. These findings, coupled with the linkage data in African Americans<sup>75</sup>, insulin levels, insulin resistance, and leptin concentrations<sup>69 76</sup>, support a possible pleiotropic effect for variants in *ENPP1*.

Additionally, addressing the confounding affects of admixture in genetic association studies of is very important. Through the use of ancestry informative markers, this study has presented the

first exploratory analysis of the impact of admixture in studies of Afro-Caribbean. This supports the use of ancestral proportions as a means of reducing spurious associations due to admixture. Several studies have tested for association with markers for ancestry and ancestral estimates<sup>77-80</sup>, however it appears that only one study (outside of our laboratory) has performed admixture adjustments for genetic association studies<sup>81</sup>.

Over the last few years and during the span of this project, several advances and findings have been reported in the field of diabetes genetics. The recent use of Genome Wide Association Studies (GWAS) has overshadowed the traditional approach of linkage analysis for identifying genes. Although both approaches have their pros and cons, there is a growing trend towards GWAS in case-control populations as opposed to linkage analysis in families. The combined efforts of both approaches has led to the identification of several significant associations, however the reproducibility of the most of those associations are questionable. To date, there are 8 or so genes that are widely accepted in the diabetes genetics research community as “T2D genes.” These genes include peroxisome proliferator-activated receptor  $\gamma$  (*PPARG*)<sup>16</sup>, ATP-sensitive inwardly-rectifying potassium channel subunit Kir6.2 (*KCNJ11*)<sup>82</sup>, transcription factor 7-like 2 (*TCF7L2*)<sup>83</sup>, solute carrier family 30 member 8 (*SLC30A8*)<sup>84-87</sup>, CDK5 regulatory subunit associated protein (*CDKAL1*)<sup>84,86-89</sup>, insulin-like growth factor 2 mRNA binding protein (*IGF2BP2*)<sup>84,86,87,89</sup>, and the region encoding cyclin-dependent kinase inhibitor 2A and cyclin-dependent kinase inhibitor 2B (*CDKN2A/CDKNA2B*)<sup>84,87,89</sup>. Outside of *PPARG*, *KCNJ11*, and *TCF7L2*, the majority of these findings (five) were identified using GWAS, many of which have been confirmed across several studies.

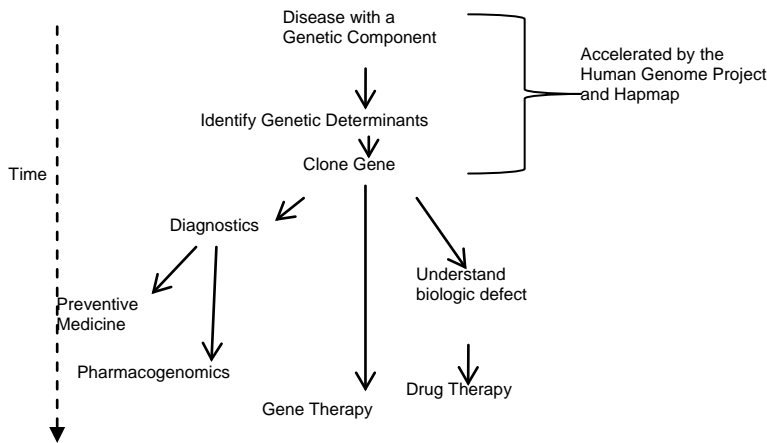


These findings suggest that positive associations observed in Europeans do not necessarily hold true in populations of African descent and variants unique to populations of African descent probably exist. Although *TCF7L2* associations have been replicated across numerous populations, the mechanism behind the association is not fully understood.

In conclusion, with the increasing number of genes identified that contribute to T2D, replication studies across several populations and ethnicities are necessary to confirm findings. Due to the increased genetic diversity (i.e. more variants and less LD) for African derived populations as compared to populations of European descent, it is possible that some of the significant associations observed in European populations may not hold true in African ancestry genetic association studies. Current GWAS in Europeans capture approximately 78% of the common SNPs, defined as having a  $MAF \geq 5\%$  in HapMap CEU, but the same SNP set would capture less variation in African populations. Until funding (and adequate SNP sets) for large scale GWAS as well as sequencing efforts in African populations are available, positional cloning approaches, candidate gene studies, and mapping by admixture linkage disequilibrium (MALD) provide the most productive avenue for identifying genes that contribute to diabetes and/or related traits. These studies, however, must incorporate a well characterized, adequately powered sample size sufficient to detect true associations, thorough SNP coverage across genes, and take into account any influences due to admixture. In my opinion, the successes reported from GWAS will ultimately trigger a paradigm shift in the field away from a focus on identifying genes, and instead towards determining the mechanisms and functions of variants in those genes.

We anticipate with the increasing number of genetic information obtained via the use of dense genetic information this will lead to the need for larger and more in-depth study of this topic.

The information obtained will aid to advance the care of populations of African ancestry affected and those at risk of T2D. As public health has shown mortality and morbidity of T2D is higher in populations of African descent. T2D is a major health issue whose risk factors include older age, obesity, ethnicity and family history. Gene identification remains a daunting task while the next challenge is the identification of genes and genomic regions responsible for gene regulation. Knowledge along with each gene function and gene regulation is expected to be the cornerstone for the development of new drugs and improved treatment. As genetic information becomes more involved in medical practice, (i.e personalized medicine) knowing ones genetic makeup can aid in therapeutic strategies, disease prevention, etc. Thus the identification of genes will significantly benefit public awareness and public health (Figure 3).



**Figure 3. Progression and Direction for Genomic Studies and their Implications**

## APPENDIX: TABLES

**Table 1. Characteristic of the Afro Caribbean Men**

<b>Trait</b>	<b>T2D Cases (n=380)</b>	<b>Controls (n=1454)</b>
Age at exam (years±SD)	62.8±10.2	57.9±10.1‡
BMI (kg/m <sup>2</sup> )	28.4±5.6	27.1±4.6‡
Insulin (iU/ml)	13.2±8.4	12.0±6.2†
Glucose (mg/dl)	166.6±67.4	91.7±11.6‡
HOMA-IR	5.3±3.6	2.8±1.5‡

Mean differs significantly between cases and controls P-value ‡<0.001 and † <0.05.

**Table 2. Nominally Significant (P<0.05) Single-SNP Genotypic Tests of Association with BMI, Fasting Glucose, Insulin Resistance using HOMA-IR and Fasting Insulin in Non-Diabetic Controls.**

Phenotype	Marker	Location	Major <sup>+</sup> / Minor alleles	MAF Controls	*1/2 Mean ± SD	1/1 Mean±SD	1/2 Mean ±SD	Additive P-value	Dominant P-value	Recessive P-value
BMI	rs703184	Intron	G\C	0.09	26.9±4.4 (221)	27.2±4.5 (1150)	30.6±10.5 (13)	0.511	0.982	<b>0.007</b> ( <b>0.006</b> ) <sup>a</sup>
BMI	rs1044498	Exon	C\A	0.14	27.0±4.6 (302)	27.2±4.5 (1036)	25.4±3.8 (38)	<b>0.040</b> (0.066) <sup>a</sup>	0.147	<b>0.013</b> ( <b>0.018</b> ) <sup>a</sup>
Glucose	rs6569759	Intron 1	G\A	0.24	90.9±11.8 (485)	92.5±11.7 (809)	94.8±9.9 (88)	<b>0.0005</b> ( <b>0.001</b> )	<b>0.002</b> ( <b>0.003</b> )	<b>0.010</b> ( <b>0.009</b> )
Glucose	rs858345	Intron	C\T	0.36	91.8±11.8 (591)	92.1±11.6 (562)	89.3±11.4 (179)	0.085	0.611	<b>0.006</b> ( <b>0.012</b> )
Glucose	rs17060836	Intron	T\C	0.23	92.4±12.1 (464)	90.3±11.1 (814)	92.1±11.6 (76)	<b>0.029</b> ( <b>0.018</b> )	<b>0.004</b> ( <b>0.002</b> )	0.742
Glucose	rs7768480	Intron (boundary)	G\A	0.46	91.4±11.9 (670)	91.3±11.7 (418)	93.0±11.3 (301)	0.094	0.551	<b>0.025</b> ( <b>0.019</b> )
Glucose	rs1830971	Intron	A\G	0.26	91.8±11.5 (526)	91.9±12.0 (747)	88.9±10.9 (100)	0.161	0.601	<b>0.016</b> ( <b>0.005</b> )
Glucose	rs1409181	Intron	C\G	0.31	91.9±11.5 (568)	91.9±11.9 (662)	89.5±10.9 (141)	0.121	0.499	<b>0.022</b> ( <b>0.020</b> )
Insulin	rs6935458	Promoter	A\G	0.43	12.1±6.4 (639)	11.6±5.9 (468)	12.6±6.6 (268)	0.063	<b>0.033</b> (0.108) <sup>b</sup>	0.102
Insulin	rs9372999	Intron	C\A	0.26	11.9±6.6 (490)	11.9±5.8 (743)	13.3±8.4 (106)	0.497	0.169	<b>0.044</b> (0.264) <sup>b</sup>
HOMA-IR	rs858345	Intron	C\T	0.36	2.8±1.6 (591)	2.8±1.6 (562)	2.5±1.2 (179)	0.350	0.937	<b>0.043</b> (0.205) <sup>b</sup>

Test models refer to the minor allele. \*1= Major Allele; 2= Minor Allele in Controls. **Bold:** P-values <0.05. Unadjusted *P* values without parentheses and *P* values with parentheses adjusted for (a) Age; (b) Age- and BMI

**Table 3 Nominally Significant (P<0.05) Single-SNP Genotypic Tests of Association with Type 2 Diabetes.**

Locations	Marker	Major <sup>†</sup> / Minor alleles	Dominant* P-value	OR (95% CI)	Additive P-value	OR (95% CI)	Recessive P-value	OR (95% CI)
Intron	rs703184	G\C	<b>0.022</b> (0.068)	1.39 (1.05-1.85)	-	-	-	-
Intron	rs7749493	G\T	<b>0.037</b> ( <b>0.025</b> )	0.78 (0.61-0.99)	<b>0.019</b> ( <b>0.014</b> )	0.79 (0.65-0.96)	0.095	0.64 (0.38-1.08)
Intron	rs17060836	T\C	<b>0.020</b> ( <b>0.019</b> )	1.32 (1.04-1.67)	0.092	1.18 (0.97-1.43)	0.562	0.85 (0.50-1.46)
3'UTR	rs7754561	G\A	<b>0.020</b> ( <b>0.034</b> )	0.70 (0.52-0.95)	-	-	-	-
3'UTR	rs7769993	A\G	<b>0.025</b> ( <b>0.032</b> )	1.33 (1.03-1.71)	0.249	1.10 (0.93-1.30)	0.533	0.91 (0.67-1.23)
3'UTR	rs9373000	A\G	<b>0.008</b> ( <b>0.013</b> )	1.38 (1.08-1.76)	<b>0.018</b> ( <b>0.023</b> )	1.22 (1.03-1.44)	0.339	1.17 (0.85-1.62)

\*Only the dominant model was considered where the minor allele homozygote count for either cases or controls was <10. Test models refer to the minor allele. <sup>†</sup>Major allele is defined as most common allele in controls. <sup>‡</sup>Dominant model. **Bold:** P-values <0.05. Unadjusted *P* values without parentheses and *P* values with parentheses adjusted for age- and BMI.

**Table 4. Single SNP Association Results for BMI (kg/m<sup>2</sup>)**

Marker	Position	Location	Major†/ Minor	MAF	*1/2	1/1	2/2	Additive	Dominant	Recessive
			alleles	Controls	Mean±SD	Mean ± SD	Mean ±SD	P-value	P-value	P-value
rs7773292	132141454		C\T	0.22	27.3±4.0 (430)	27.1±4.9 (814)	27.2±4.4 (79)	0.687	0.635	0.958
rs6925433	132161059	Promoter	A\G	0.19	27.4±4.4 (406)	27.1±4.7 (885)	26.5±3.9 (48)	0.870	0.541	0.286
rs13211931	132165417	Promoter	G\T	0.09	27.1±4.7 (207)	27.2±4.6 (1150)	27.4±3.3 (17)	0.813	0.747	0.842
rs6935458	132168013	Promoter	A\G	0.43	26.9±4.8 (639)	27.1±4.4 (468)	27.6±4.5 (268)	0.233	0.769	0.071
rs6936129	132168133	Promoter	G\C	0.09	27.3±4.8 (223)	27.2±4.6 (1104)	26.8±3.3 (16)	0.735	0.641	0.743
rs6569759	132174809	Intron	G\A	0.24	27.3±4.9 (485)	27.2±4.4 (809)	26.6±4.5 (88)	0.618	0.957	0.253
rs6939185	132180880	Intron	G\A	0.09	27.2±5.1 (228)	27.1±4.5 (1161)	26.6±5.7 (15)	0.988	0.885	0.636
rs9398995	132181896	Intron	T\C	0.22	27.0±4.4 (443)	27.3±4.7 (849)	26.8±4.3 (75)	0.268	0.305	0.474
rs943004	132182569	Intron	C\T	0.22	27.0±4.7 (483)	27.1±4.6 (832)	27.3±4.1 (62)	0.874	0.757	0.775
rs1830971	132190046	Intron	A\G	0.26	27.0±4.5 (526)	27.2±4.6 (747)	27.5±5.0 (100)	0.810	0.876	0.379
rs1409181	132190993	Intron	C\G	0.31	27.1±4.4 (568)	27.2±4.6 (662)	27.1±4.8 (141)	0.649	0.649	0.806
rs2021966	132192132	Intron	A\G	0.30	27.1±4.4 (575)	27.2±4.7 (677)	27.2±4.8 (123)	0.945	0.824	0.815
rs7771841	132192798	Intron	G\A	0.22	27.2±4.4 (457)	27.1±4.5 (846)	27.7±4.5 (68)	0.309	0.441	0.305
rs9372999	132194845	Intron	C\A	0.26	27.2±4.3 (490)	27.1±4.7 (743)	27.0±4.5 (106)	0.997	0.882	0.779
rs858338	132194900	Intron	G\T	0.12	26.9±4.5 (302)	27.2±4.5 (1070)	29.3±9.0 (19)	0.873	0.693	0.039
rs858339	132195590	Intron	T\A	0.32	26.9±4.3 (601)	27.3±4.5 (654)	27.5±5.5 (147)	0.492	0.145	0.379
rs703184	132196688	Intron	G\C	0.09	26.9±4.4 (221)	27.2±4.5 (1150)	30.6±10.5 (13)	0.511	0.982	<b>0.007</b>
rs7775386	132198842	Intron	T\C	0.49	27.2±4.5 (650)	27.2±4.6 (373)	27.1±4.7 (338)	0.722	0.753	0.788
rs6916495	132201958	Intron	C\T	0.11	26.7±4.6 (261)	27.3±4.6 (1092)	27.6±4.3 (23)	0.221	0.128	0.642
rs9493110	132203671	Intron	C\T	0.24	27.2±4.5 (505)	27.1±4.7 (782)	27.4±4.6 (78)	0.457	0.487	0.651
rs858345	132205310	Intron	C\T	0.36	27.3±4.6 (591)	27.1±4.5 (562)	26.7±4.2 (179)	0.696	0.660	0.154
rs4141767	132206700	Intron	A\G	0.41	27.2±4.7 (654)	27.0±4.5 (485)	27.2±4.1 (226)	0.593	0.455	0.968
rs6926970	132208983	Intron	A\C	0.21	27.2±4.7 (429)	27.2±4.6 (868)	26.4±4.5 (80)	0.454	0.810	0.154
rs1044498	132214061	Exon	C\A	0.14	27.0±4.6 (302)	27.2±4.5 (1036)	25.4±3.8 (38)	<b>0.040</b>	0.147	<b>0.013</b>
rs7749493	132216752	Intron	G\T	0.24	27.3±4.8 (476)	27.1±4.4 (804)	26.8±4.9 (98)	0.970	0.642	0.422

rs17060836	132230641	Intron	T\C	0.23	27.3±4.8 (464)	26.9±4.4 (814)	27.5±4.6 (76)	0.114	0.127	0.383
rs7769712	132235088	Intron	A\C	0.31	27.3±4.3 (560)	27.2±4.8 (632)	26.7±4.5 (137)	0.549	0.980	0.205
		Intron								
rs7768480	132245125	(boundary)	G\A	0.46	27.2±4.8 (670)	27.2±4.6 (418)	27.0±4.0 (301)	0.728	0.975	0.570
rs12202373	132251394	Intron	T\C	0.09	27.7±4.9 (215)	27.1±4.5 (1139)	26.8±4.4 (12)	0.140	0.095	0.761
rs9493118	132253058	Intron	C\T	0.31	27.2±4.4 (538)	27.1±4.4 (622)	27.9±6.1 (138)	0.104	0.310	0.060
rs1804025	132253227	Exon	A\G	0.23	27.0±4.7 (459)	27.1±4.3 (842)	27.7±6.2 (83)	0.582	0.936	0.215
rs7754561	132254387	3' UTR	G\A	0.13	26.9±3.9 (292)	27.2±4.6 (1019)	27.8±5.0 (30)	0.680	0.455	0.425
rs7769993	132262945	3' UTR	A\G	0.42	27.2±4.8 (641)	27.1±4.3 (489)	27.1±4.7 (257)	0.965	0.757	0.765
rs9373000	132263399	3' UTR	A\G	0.36	27.1±4.8 (620)	27.2±4.3 (575)	27.0±4.6 (185)	0.711	0.794	0.712

\* 1= Major Allele; 2= Minor Allele in Controls. Test models refer to the minor allele. **Bold:** P-values <0.05

**Table 5. Single SNP Association Results for Fasting Glucose (mg/dl)**

Marker	Position	Location	Major†/ Minor	MAF	*1/2	1/1	2/2	Additive	Dominant	Recessive
			alleles	Controls	Mean±SD	Mean ± SD	Mean ±SD	P-value	P-value	P-value
rs7773292	132141454		C\T	0.22	91.4±11.5 (430)	92.2±11.7 (814)	90.3±12.8 (79)	0.861	0.472	0.309
rs6925433	132161059	Promoter	A\G	0.19	91.3±11.4 (406)	92.5±12.1 (885)	91.7±14.1 (48)	0.200	0.132	0.991
rs13211931	132165417	Promoter	G\T	0.09	91.5±11.8 (207)	92.4±10.9 (1150)	91.0±9.6 (17)	0.476	0.386	0.811
rs6935458	132168013	Promoter	A\G	0.43	92.2±11.1 (639)	91.4±12.4 (468)	91.5±10.9 (268)	0.374	0.271	0.768
rs6936129	132168133	Promoter	G\C	0.09	91.9±11.5 (223)	91.3±12.5 (1104)	89.0±7.8 (16)	0.247	0.319	0.325
rs6569759	132174809	Intron	G\A	0.24	90.9±11.8 (485)	92.5±11.7 (809)	94.8±9.9 (88)	<b>0.0005</b>	<b>0.002</b>	<b>0.010</b>
rs6939185	132180880	Intron	G\A	0.09	91.3±11.8 (228)	92.6±11.1 (1161)	95.5±7.2 (15)	0.052	0.076	0.193
rs9398995	132181896	Intron	T\C	0.22	92.1±11.6 (443)	91.4±11.8 (849)	89.7±11.4 (75)	0.064	0.125	0.116
rs943004	132182569	Intron	C\T	0.22	92.0±11.6 (483)	91.2±11.7 (832)	92.3±12.5 (62)	0.407	0.251	0.698
rs1830971	132190046	Intron	A\G	0.26	91.8±11.5 (526)	91.9±12.0 (747)	88.9±10.9 (100)	0.161	0.601	<b>0.016</b>
rs1409181	132190993	Intron	C\G	0.31	91.9±11.5 (568)	91.9±11.9 (662)	89.5±10.9 (141)	0.121	0.499	<b>0.022</b>
rs2021966	132192132	Intron	A\G	0.30	92.3±11.9 (575)	91.4±11.2 (677)	90.4±11.8 (123)	0.057	0.094	0.164
rs7771841	132192798	Intron	G\A	0.22	91.3±11.4 (457)	92.4±12.1 (846)	91.4±12.1 (68)	0.216	0.116	0.872
rs9372999	132194845	Intron	C\A	0.26	91.6±11.8 (490)	91.9±11.7 (743)	90.4±11.6 (106)	0.699	0.914	0.264
rs858338	132194900	Intron	G\T	0.12	91.4±11.5 (302)	92.5±11.9 (1070)	92.2±10.7 (19)	0.175	0.152	0.852
rs858339	132195590	Intron	T\A	0.32	91.1±11.2 (601)	92.4±12.2 (654)	91.5±11.3 (147)	0.200	0.069	0.848
rs703184	132196688	Intron	G\C	0.09	91.6±11.6 (221)	92.1±11.8 (1150)	94.2±11.3 (13)	0.386	0.461	0.429
rs7775386	132198842	Intron	T\C	0.49	91.9±11.8 (650)	91.8±11.6 (373)	91.1±11.5 (338)	0.312	0.529	0.298
rs6916495	132201958	Intron	C\T	0.11	91.7±11.6 (261)	91.5±12.2 (1092)	94.1±7.7 (23)	0.793	0.995	0.337
rs9493110	132203671	Intron	C\T	0.24	91.5±11.8 (505)	92.3±11.4 (782)	89.9±12.9 (78)	0.941	0.558	0.153
rs858345	132205310	Intron	C\T	0.36	91.8±11.8 (591)	92.1±11.6 (562)	89.3±11.4 (179)	0.085	0.611	<b>0.006</b>
rs4141767	132206700	Intron	A\G	0.41	92.1±11.5 (654)	91.6±12.1 (485)	91.6±11.1 (226)	0.545	0.463	0.853
rs6926970	132208983	Intron	A\C	0.21	91.5±11.7 (429)	91.8±11.6 (868)	93.0±12.1 (80)	0.321	0.462	0.303
rs1044498	132214061	Exon	C\A	0.14	91.7±11.8 (302)	90.9±11.1 (1036)	93.4±12.1 (38)	0.721	0.431	0.338
rs7749493	132216752	Intron	G\T	0.24	91.7±11.6 (476)	91.6±11.8 (804)	91.7±11.8 (98)	0.907	0.884	0.994
rs17060836	132230641	Intron	T\C	0.23	92.4±12.1 (464)	90.3±11.1 (814)	92.1±11.6 (76)	<b>0.029</b>	<b>0.004</b>	0.742



rs7769712	132235088	Intron	A\C	0.31	91.9±11.4 (560)	91.5±11.9 (632)	92.0±11.6 (137)	0.822	0.634	0.769
		Intron								
rs7768480	132245125	(boundary)	G\A	0.46	91.4±11.9 (670)	91.3±11.7 (418)	93.0±11.3 (301)	0.094	0.551	<b>0.025</b>
rs12202373	132251394	Intron	T\C	0.09	91.8±11.6 (215)	91.2±12.0 (1139)	88.5±9.67 (12)	0.296	0.373	0.340
rs9493118	132253058	Intron	C\T	0.31	91.9±11.5 (538)	91.7±11.5 (622)	91.6±12.1 (138)	0.786	0.787	0.879
rs1804025	132253227	Exon	A\G	0.23	91.5±11.6 (459)	91.8±11.8 (842)	91.9±11.4 (83)	0.665	0.675	0.805
rs7754561	132254387	3' UTR	G\A	0.13	91.5±11.6 (292)	92.3±12.1 (1019)	91.8±7.6 (30)	0.381	0.329	0.947
rs7769993	132262945	3' UTR	A\G	0.42	91.6±12.1 (641)	91.7±11.3 (489)	91.9±11.5 (257)	0.785	0.874	0.759
rs9373000	132263399	3' UTR	A\G	0.36	91.7±12.1 (620)	91.8±11.4 (575)	91.4±11.5 (185)	0.809	0.928	0.720

\* 1= Major Allele; 2= Minor Allele in Controls. Test models refer to the minor allele. **Bold:** P-values <0.05

**Table 6. Single SNP Association Results for Fasting Insulin ( $\mu\text{U/ml}$ )**

Marker	Position	Location	Major†/ Minor	MAF	*1/2	1/1	2/2	Additive	Dominant	Recessive
			alleles	Controls	Mean±SD	Mean ± SD	Mean ±SD	P-value	P-value	P-value
rs7773292	132141454		C\T	0.22	12.5±6.6 (430)	11.8±6.1 (814)	12.2±6.7 (79)	0.118	0.194	0.894
rs6925433	132161059	Promoter	A\G	0.19	12.5±6.2 (406)	11.8±6.2 (885)	11.9±7.4 (48)	0.116	0.195	0.892
rs13211931	132165417	Promoter	G\T	0.09	11.6±5.8 (207)	12.1±6.3 (1150)	13.0±4.1 (17)	0.394	0.562	0.506
rs6935458	132168013	Promoter	A\G	0.43	12.1±6.4 (639)	11.6±5.9 (468)	12.6±6.6 (268)	0.063	<b>0.033</b>	0.102
rs6936129	132168133	Promoter	G\C	0.09	12.5±6.6 (223)	11.9±6.1 (1104)	12.5±5.6 (16)	0.210	0.225	0.761
rs6569759	132174809	Intron	G\A	0.24	12.3±6.8 (485)	11.9±5.9 (809)	11.7±5.3 (88)	0.505	0.753	0.586
rs6939185	132180880	Intron	G\A	0.09	11.9±5.2 (228)	12.0±6.4 (1161)	9.9±3.8 (15)	0.612	0.431	0.195
rs9398995	132181896	Intron	T\C	0.22	11.7±5.8 (443)	12.1±6.3 (849)	12.4±7.1 (75)	0.441	0.672	0.595
rs943004	132182569	Intron	C\T	0.22	12.0±6.3 (483)	11.9±6.1 (832)	13.4±8.2 (62)	0.358	0.160	0.078
rs1830971	132190046	Intron	A\G	0.26	11.7±5.8 (526)	12.2±6.6 (747)	12.2±6.5 (100)	0.201	0.364	0.797
rs1409181	132190993	Intron	C\G	0.31	11.9±6.4 (568)	12.2±6.3 (662)	11.9±6.1 (141)	0.410	0.438	0.736
rs2021966	132192132	Intron	A\G	0.30	11.8±5.9 (575)	12.1±6.6 (677)	11.9±6.3 (123)	0.414	0.505	0.932
rs7771841	132192798	Intron	G\A	0.22	12.0±5.9 (457)	12.1±6.5 (846)	11.2±4.9 (68)	0.631	0.444	0.321
rs9372999	132194845	Intron	C\A	0.26	11.9±6.6 (490)	11.9±5.8 (743)	13.3±8.4 (106)	0.497	0.169	<b>0.044</b>
rs858338	132194900	Intron	G\T	0.12	12.4±5.7 (302)	11.9±6.3 (1070)	11.5±5.1 (19)	0.311	0.399	0.734
rs858339	132195590	Intron	T\A	0.32	11.9±5.8 (601)	12.1±6.8 (654)	11.9±5.5 (147)	0.587	0.662	0.953
rs703184	132196688	Intron	G\C	0.09	12.2±5.8 (221)	11.9±6.2 (1150)	12.1±5.4 (13)	0.484	0.504	0.929
rs7775386	132198842	Intron	T\C	0.49	11.9±6.4 (650)	11.9±6.3 (373)	12.0±5.9 (338)	0.840	0.819	0.862
rs6916495	132201958	Intron	C\T	0.11	11.7±5.8 (261)	12.1±6.3 (1092)	13.7±6.9 (23)	0.591	0.909	0.194
rs9493110	132203671	Intron	C\T	0.24	12.2±6.8 (505)	11.9±5.8 (782)	12.5±6.9 (78)	0.412	0.363	0.537
rs858345	132205310	Intron	C\T	0.36	12.3±6.2 (591)	12.0±6.7 (562)	11.3±4.8 (179)	0.985	0.419	0.095
rs4141767	132206700	Intron	A\G	0.41	11.7±5.9 (654)	12.2±6.9 (485)	12.3±5.2 (226)	0.326	0.809	0.411
rs6926970	132208983	Intron	A\C	0.21	12.2±6.5 (429)	11.9±6.2 (868)	11.8±5.8 (80)	0.499	0.663	0.785
rs1044498	132214061	Exon	C\A	0.14	12.2±6.4 (302)	11.9±6.2 (1036)	10.5±4.3 (38)	0.961	0.663	0.147

rs7749493	132216752	Intron	G\T	0.24	12.2±6.5 (476)	11.9±6.2 (804)	11.1±4.5 (98)	0.882	0.598	0.119
rs17060836	132230641	Intron	T\C	0.23	11.9±6.2 (464)	11.9±6.4 (814)	12.2±5.6 (76)	0.878	0.976	0.687
rs7769712	132235088	Intron	A\C	0.31	11.9±5.9 (560)	12.1±6.5 (632)	12.1±5.7 (137)	0.612	0.720	0.959
rs7768480	132245125	Intron (boundary)	G\A	0.46	11.9±6.3 (670)	12.2±6.4 (418)	11.9±6.0 (301)	0.422	0.539	0.861
rs12202373	132251394	Intron	T\C	0.09	11.5±5.8 (215)	12.1±6.2 (1139)	13.6±8.6 (12)	0.318	0.482	0.356
rs9493118	132253058	Intron	C\T	0.31	12.1±6.4 (538)	11.9±6.1 (622)	11.8±6.2 (138)	0.826	0.969	0.659
rs1804025	132253227	Exon	A\G	0.23	11.9±6.3 (459)	12.0±6.2 (842)	12.3±6.9 (83)	0.972	0.900	0.696
rs7754561	132254387	3' UTR	G\A	0.13	12.2±5.9 (292)	11.9±6.3 (1019)	11.9±4.8 (30)	0.591	0.644	0.981
rs7769993	132262945	3' UTR	A\G	0.42	12.0±6.5 (641)	12.0±6.1 (489)	11.8±6.1 (257)	0.842	0.729	0.695
rs9373000	132263399	3' UTR	A\G	0.36	12.0±6.5 (620)	12.0±6.1 (575)	11.8±5.9 (185)	0.919	0.783	0.683

\* 1= Major Allele; 2= Minor Allele in Controls. Test models refer to the minor allele. **Bold:** P-values <0.05

**Table 7. Single SNP Association Results for Insulin Resistance Estimated by HOMA**

Marker	Position	Location	Major†/ Minor	MAF	*1/2	1/1	2/2	Additive	Dominant	Recessive
			alleles	Controls	Mean±SD	Mean ± SD	Mean ±SD	P-value	P-value	P-value
rs7773292	132141454		C\T	0.22	2.9±1.6 (430)	2.7±1.5 (814)	2.7±1.4 (79)	0.303	0.144	0.730
rs6925433	132161059	Promoter	A\G	0.19	2.9±1.5 (406)	2.7±1.6 (885)	2.7±1.4 (48)	0.275	0.153	0.704
rs13211931	132165417	Promoter	G\T	0.09	2.7±1.5 (207)	2.8±1.6 (1150)	2.9±1.0 (17)	0.681	0.538	0.607
rs6935458	132168013	Promoter	A\G	0.43	2.8±1.6 (639)	2.7±1.5 (468)	2.9±1.6 (268)	0.073	0.136	0.143
rs6936129	132168133	Promoter	G\C	0.09	2.8±1.6 (223)	2.7±1.5 (1104)	2.8±1.3 (16)	0.462	0.421	0.982
rs6569759	132174809	Intron	G\A	0.24	2.8±1.7 (485)	2.7±1.5 (809)	2.8±1.4 (88)	0.402	0.292	0.994
rs6939185	132180880	Intron	G\A	0.09	2.8±1.3 (228)	2.7±1.6 (1161)	2.3±0.9 (15)	0.769	0.963	0.318
rs9398995	132181896	Intron	T\C	0.22	2.7±1.5 (443)	2.8±1.6 (849)	2.8±1.7 (75)	0.529	0.402	0.889
rs943004	132182569	Intron	C\T	0.22	2.7±1.5 (483)	2.7±1.5 (832)	3.1±1.9 (62)	0.309	0.629	0.087
rs1830971	132190046	Intron	A\G	0.26	2.7±1.5 (526)	2.8±1.6 (747)	2.7±1.5 (100)	0.303	0.274	0.697
rs1409181	132190993	Intron	C\G	0.31	2.7±1.6 (568)	2.8±1.6 (662)	2.6±1.4 (141)	0.263	0.370	0.331
rs2021966	132192132	Intron	A\G	0.30	2.7±1.5 (575)	2.8±1.6 (677)	2.7±1.5 (123)	0.380	0.354	0.710
rs7771841	132192798	Intron	G\A	0.22	2.8±1.5 (457)	2.8±1.6 (846)	2.5±1.2 (68)	0.561	0.859	0.242
rs9372999	132194845	Intron	C\A	0.26	2.8±1.6 (490)	2.7±1.5 (743)	2.9±1.9 (106)	0.340	0.618	0.179
rs858338	132194900	Intron	G\T	0.12	2.9±1.5 (302)	2.7±1.6 (1070)	2.6±1.2 (19)	0.313	0.227	0.696
rs858339	132195590	Intron	T\A	0.32	2.8±1.5 (601)	2.8±1.7 (654)	2.7±1.4 (147)	0.842	0.906	0.811
rs703184	132196688	Intron	G\C	0.09	2.8±1.5 (221)	2.7±1.5 (1150)	2.8±1.3 (13)	0.410	0.388	0.911
rs7775386	132198842	Intron	T\C	0.49	2.8±1.6 (650)	2.7±1.6 (373)	2.7±1.4 (338)	1.00	0.839	0.834
rs6916495	132201958	Intron	C\T	0.11	2.7±1.5 (261)	2.8±1.6 (1092)	3.2±1.7 (23)	0.983	0.653	0.177
rs9493110	132203671	Intron	C\T	0.24	2.8±1.7 (505)	2.7±1.5 (782)	2.8±1.5 (78)	0.534	0.456	0.975
rs858345	132205310	Intron	C\T	0.36	2.8±1.6 (591)	2.8±1.6 (562)	2.5±1.2 (179)	0.350	0.937	<b>0.043</b>
rs4141767	132206700	Intron	A\G	0.41	2.7±1.5 (654)	2.8±1.7 (485)	2.8±1.4 (226)	0.802	0.340	0.443
rs6926970	132208983	Intron	A\C	0.21	2.8±1.6 (429)	2.7±1.5 (868)	2.7±1.5 (80)	0.513	0.425	0.969
rs1044498	132214061	Exon	C\A	0.14	2.8±1.5 (302)	2.7±1.5 (1036)	2.5±1.2 (38)	0.584	0.835	0.261

rs7749493	132216752	Intron	G\T	0.24	2.8±1.6 (476)	2.7±1.5 (804)	2.5±1.1 (98)	0.548	0.963	0.123
rs17060836	132230641	Intron	T\C	0.23	2.7±1.5 (464)	2.8±1.6 (814)	2.8±1.5 (76)	0.729	0.495	0.584
rs7769712	132235088	Intron	A\C	0.31	2.7±1.5 (560)	2.8±1.6 (632)	2.8±1.5 (137)	0.695	0.526	0.852
rs7768480	132245125	Intron (boundary)	G\A	0.46	2.7±1.6 (670)	2.8±1.6 (418)	2.8±1.5 (301)	0.731	0.489	0.866
rs12202373	132251394	Intron	T\C	0.09	2.6±1.5 (215)	2.8±1.5 (1139)	3.0±1.9 (12)	0.501	0.383	0.575
rs9493118	132253058	Intron	C\T	0.31	2.8±1.6 (538)	2.7±1.5 (622)	2.7±1.5 (138)	0.894	0.896	0.616
rs1804025	132253227	Exon	A\G	0.23	2.7±1.6 (459)	2.7±1.5 (842)	2.8±1.7 (83)	0.814	0.903	0.728
rs7754561	132254387	3' UTR	G\A	0.13	2.8±1.5 (292)	2.7±1.5 (1019)	2.7±1.1 (30)	0.513	0.420	0.868
rs7769993	132262945	3' UTR	A\G	0.42	2.7±1.6 (641)	2.8±1.5 (489)	2.7±1.5 (257)	0.824	0.858	0.850
rs9373000	132263399	3' UTR	A\G	0.36	2.8±1.6 (620)	2.8±1.5 (575)	2.7±1.4 (185)	0.766	0.897	0.680

\* 1= Major Allele; 2= Minor Allele in Controls. Test models refer to the minor allele. **Bold:** P-values <0.05

**Table 8. Single-SNP Genotypic Tests of Association with T2D**

Locations	Marker	Major†/Minor alleles	Dominant* P-value	OR (95% CI)	Additive P-value	OR (95% CI)	Recessive P-value	OR (95% CI)
	rs7773292	C\T	0.623	1.06 (0.83-1.35)	0.995	1.00 (0.82-1.22)	0.311	0.75 (0.43-1.31)
Promoter	rs6925433	A\G	0.179	1.18 (0.93-1.51)	0.206	1.14 (0.93-1.40)	0.734	1.11 (0.61-2.04)
Promoter	rs13211931	G\T	0.270	1.18 (0.88-1.60)	-	-	-	-
Promoter	rs6935458	A\G	0.139	0.84 (0.66-1.06)	0.310	0.92 (0.78-1.08)	0.945	0.99 (0.74-1.32)
Promoter	rs6936129	G\C	0.110	0.77 (0.55-1.06)	-	-	-	-
Intron	rs6569759	G\A	0.622	0.94 (0.74-1.19)	0.906	1.01 (0.84-1.22)	0.209	1.32 (0.86-2.03)
Intron	rs6939185	G\A	0.482	1.11 (0.83-1.49)	-	-	-	-
Intron	rs9398995	T\C	0.627	0.94 (0.74-1.20)	0.596	0.95 (0.78-1.15)	0.725	0.91 (0.54-1.54)
Intron	rs943004	C\T	0.182	0.85 (0.67-1.08)	0.251	0.89 (0.72-1.09)	0.940	0.98 (0.56-1.72)
Intron	rs1830971	A\G	0.848	0.98 (0.78-1.23)	0.907	0.99 (0.82-1.19)	0.932	1.02 (0.66-1.59)
Intron	rs1409181	C\G	0.534	0.93 (0.74-1.17)	0.335	0.92 (0.77-1.09)	0.276	0.80 (0.53-1.20)
Intron	rs2021966	A\G	0.590	0.94 (0.74-1.18)	0.560	0.95 (0.79-1.14)	0.705	0.92 (0.61-1.40)
Intron	rs7771841	G\A	0.440	0.91 (0.71-1.16)	0.302	0.90 (0.73-1.10)	0.285	0.72 (0.39-1.32)
Intron	rs9372999	C\A	0.287	1.14 (0.90-1.44)	0.509	1.06 (0.88-1.28)	0.686	0.91 (0.58-1.43)
Intron	rs858338	G\T	0.101	1.24 (0.96-1.62)	-	-	-	-
Intron	rs858339	T\A	0.473	1.09 (0.86-1.37)	0.543	1.05 (0.89-1.25)	0.880	1.03 (0.71-1.49)
Intron	rs703184	G\C	<b>0.023</b>	1.39 (1.05-1.85)	-	-	-	-
Intron	rs7775386	T\C	0.924	1.01 (0.78-1.31)	0.473	0.94 (0.80-1.11)	0.193	0.83 (0.63-1.10)
Intron	rs6916495	C\T	0.193	1.20 (0.91-1.58)	-	-	-	-
Intron	rs9493110	C\T	0.527	1.08 (0.85-1.36)	0.553	1.06 (0.88-1.28)	0.843	1.05 (0.65-1.71)
Intron	rs858345	C\T	0.633	0.94 (0.74-1.20)	0.531	0.95 (0.80-1.12)	0.568	0.90 (0.63-1.29)
Intron	rs4141767	A\G	0.681	1.05 (0.83-1.34)	0.587	1.05 (0.89-1.24)	0.626	1.08 (0.80-1.46)
Intron	rs6926970	A\C	0.052	0.78 (0.61-1.00)	0.058	0.82 (0.67-1.01)	0.400	0.79 (0.46-1.36)
Exon	rs1044498	C\A	0.079	0.78 (0.58-1.03)	-	-	-	-

Intron	rs7749493	G\T	<b>0.037</b>	0.78 (0.61-0.99)	<b>0.020</b>	0.79 (0.65-0.96)	0.095	0.64 (0.38-1.08)
Intron	rs17060836	T\C	<b>0.021</b>	1.32 (1.04-1.67)	0.092	1.18 (0.97-1.43)	0.562	0.85 (0.50-1.46)
Intron	rs7769712	A\C	0.567	1.07 (0.85-1.36)	0.822	1.02 (0.85-1.22)	0.650	0.91 (0.61-1.36)
Intron (boundary)	rs7768480	G\A	0.762	0.96 (0.75-1.23)	0.559	0.95 (0.81-1.12)	0.498	0.91 (0.68-1.20)
Intron	rs12202373	T\C	0.259	0.83 (0.60-1.15)	-	-	-	-
Intron	rs9493118	C\T	0.220	1.16 (0.91-1.48)	0.269	1.10 (0.93-1.32)	0.681	1.08 (0.74-1.58)
Exon	rs1804025	A\G	0.231	1.15 (0.91-1.46)	0.368	1.09 (0.90-1.32)	0.867	0.96 (0.59-1.57)
3'UTR	rs7754561	G\A	<b>0.020</b>	0.70 (0.52-0.95)	-	-	-	-
3'UTR	rs7769993	A\G	<b>0.026</b>	1.33 (1.03-1.71)	0.249	1.10 (0.93-1.30)	0.533	0.91 (0.67-1.23)
3'UTR	rs9373000	A\G	<b>0.008</b>	1.38 (1.08-1.76)	<b>0.019</b>	1.22 (1.03-1.44)	0.339	1.17 (0.85-1.62)

\*Only the dominant model was considered where the minor allele homozygote count for either cases or controls was <10. Test models refer to the minor allele. †Major allele is defined as most common allele in controls. ‡Dominant model. **Bold:** P-values <0.05.

**Table 9. Genotype Frequencies in Afro-Caribbean Case and Control Groups**

Marker	Marker Position*	Marker Location	Major†/ Minor alleles	T2D Cases Frequency (n)			Controls Frequency (n)		
				1/1	1/2	2/2	1/1	1/2	2/2
rs7773292	132141454		C\T	0.598 (207)	0.355 (123)	0.046 (16)	0.615 (814)	0.325 (430)	0.059 (79)
rs6925433	132161059	Promoter	A\G	0.620 (219)	0.339 (120)	0.039 (14)	0.661 (885)	0.303 (406)	0.036 (48)
rs13211931	132165417	Promoter	G\T	0.813 (295)	0.179 (65)	0.008 (3)	0.837 (1150)	0.151 (207)	0.012 (17)
rs6935458	132168013	Promoter	A\G	0.383 (140)	0.426 (156)	0.191 (70)	0.340 (468)	0.465 (639)	0.195 (268)
rs6936129	132168133	Promoter	G\C	0.859 (307)	0.132 (47)	0.008 (3)	0.822 (1104)	0.166 (223)	0.012 (16)
rs6569759	132174809	Intron	G\A	0.598 (216)	0.319 (115)	0.083 (30)	0.585 (809)	0.351 (485)	0.064 (88)
rs6939185	132180880	Intron	G\A	0.809 (298)	0.177 (65)	0.014 (5)	0.827 (1161)	0.162 (228)	0.011 (15)
rs9398995	132181896	Intron	T\C	0.635 (228)	0.315 (113)	0.050 (18)	0.621 (849)	0.324 (443)	0.055 (75)
rs943004	132182569	Intron	C\T	0.644 (230)	0.310 (111)	0.045 (16)	0.604 (832)	0.351 (483)	0.045 (62)
rs1830971	132190046	Intron	A\G	0.549 (200)	0.376 (137)	0.074 (27)	0.544 (747)	0.383 (526)	0.073 (100)
rs1409181	132190993	Intron	C\G	0.501 (180)	0.415 (149)	0.084 (30)	0.483 (662)	0.414 (568)	0.103 (141)
rs2021966	132192132	Intron	A\G	0.509 (184)	0.407 (147)	0.083 (30)	0.492 (677)	0.418 (575)	0.089 (123)
rs7771841	132192798	Intron	G\A	0.641 (230)	0.323 (116)	0.036 (13)	0.617 (846)	0.333 (457)	0.049 (68)
rs9372999	132194845	Intron	C\A	0.522 (178)	0.405 (138)	0.073 (25)	0.554 (743)	0.366 (490)	0.079 (106)
rs858338	132194900	Intron	G\T	0.727(264)	0.259 (94)	0.014 (5)	0.769 (1070)	0.217 (302)	0.014 (19)
rs858339	132195590	Intron	T\A	0.446 (166)	0.446 (166)	0.108 (40)	0.466 (654)	0.429 (601)	0.105 (147)
rs703184	132196688	Intron	G\C	0.779 (285)	0.216 (79)	0.005 (2)	0.831 (1150)	0.159 (221)	0.009 (13)
rs7775386	132198842	Intron	T\C	0.272 (98)	0.514 (185)	0.214 (77)	0.274 (373)	0.478 (650)	0.248 (338)
rs6916495	132201958	Intron	C\T	0.762 (279)	0.227 (83)	0.011 (4)	0.794 (1092)	0.189 (261)	0.017 (23)
rs9493110	132203671	Intron	C\T	0.554 (201)	0.386 (140)	0.061 (22)	0.573 (782)	0.369 (505)	0.057 (78)
rs858345	132205310	Intron	C\T	0.437 (153)	0.443 (155)	0.120 (42)	0.422 (562)	0.444 (591)	0.134 (179)
rs4141767	132206700	Intron	A\G	0.344 (125)	0.482 (175)	0.174 (63)	0.355 (485)	0.479 (654)	0.166 (226)
rs6926970	132208983	Intron	A\C	0.684 (249)	0.269 (98)	0.047 (17)	0.630 (868)	0.312 (429)	0.058 (80)
rs1044498	132214061	Exon	C\A	0.797 (286)	0.178 (64)	0.025 (9)	0.753 (1036)	0.219 (302)	0.028 (38)



rs7749493	132216752	Intron	G\T	0.643 (234)	0.310 (113)	0.047 (17)	0.583 (804)	0.345 (476)	0.071(98)
rs17060836	132230641	Intron	T\C	0.531 (187)	0.420 (148)	0.048 (17)	0.601 (814)	0.343 (464)	0.056 (76)
rs7769712	132235088	Intron	A\C	0.459 (160)	0.448 (156)	0.091 (32)	0.476 (632)	0.421 (560)	0.103 (137)
rs7768480	132245125	Intron (boundary)	G\A	0.309 (114)	0.492 (181)	0.198 (73)	0.301 (418)	0.482 (670)	0.217 (301)
rs12202373	132251394	Intron	T\C	0.858 (309)	0.122 (44)	0.019 (7)	0.833 (1139)	0.157 (215)	0.009 (12)
rs9493118	132253058	Intron	C\T	0.442 (150)	0.442 (150)	0.115 (39)	0.479 (622)	0.414 (538)	0.106 (138)
rs1804025	132253227	Exon	A\G	0.575 (207)	0.367 (132)	0.058 (21)	0.608 (842)	0.332 (459)	0.059 (83)
rs7754561	132254387	3' UTR	G\A	0.819 (280)	0.175 (60)	0.006 (2)	0.759 (1019)	0.218 (292)	0.022 (30)
rs7769993	132262945	3' UTR	A\G	0.288 (104)	0.540 (195)	0.172 (62)	0.353 (489)	0.462 (641)	0.185 (257)
rs9373000	132263399	3' UTR	A\G	0.339 (123)	0.507 (184)	0.154 (56)	0.417 (575)	0.4491 (620)	0.134 (185)

\*NCBI Build 36.1 (March 2006). †Major allele is defined as most common allele in control.

## BIBLIOGRAPHY

1. Diagnosis and classification of diabetes mellitus. *Diabetes Care* **30 Suppl 1**, S42-7 (2007).
2. National Diabetes Statistics fact sheet: general information and national estimates on diabetes in the United States. (National Institute of Diabetes and Digestive and Kidney Diseases, 2005).
3. Minino, A.M., Heron, M.P. & Smith, B.L. Deaths: Preliminary Data for 2004. *National Vital Statistics Reports* **54**, 1-50 (2006).
4. American. National Diabetes Facts Sheet. (American Diabetes Association, 2005).
5. Brancati, F.L., Kao, W.H., Folsom, A.R., Watson, R.L. & Szklo, M. Incident type 2 diabetes mellitus in African American and white adults: the Atherosclerosis Risk in Communities Study. *Jama* **283**, 2253-9 (2000).
6. Collins, T.C., Johnson, M., Henderson, W., Khuri, S.F. & Daley, J. Lower extremity nontraumatic amputation among veterans with peripheral arterial disease: is race an independent factor? *Med Care* **40**, I106-16 (2002).
7. Smith, S.R., Svetkey, L.P. & Dennis, V.W. Racial differences in the incidence and progression of renal diseases. *Kidney Int* **40**, 815-22 (1991).
8. State Health Facts Online. (Kaiser Family Foundation, 2007).
9. Rich, S.S. Mapping genes in diabetes. Genetic epidemiological perspective. *Diabetes* **39**, 1315-9 (1990).
10. Hawkes, C.H. Twin studies in diabetes mellitus. *Diabet Med* **14**, 347-52 (1997).
11. Rotimi, C., Cooper, R., Cao, G., Sundarum, C. & McGee, D. Familial aggregation of cardiovascular diseases in African-American pedigrees. *Genet Epidemiol* **11**, 397-407 (1994).
12. Goodman, M.J. & Chung, C.S. Diabetes mellitus: discrimination between single locus and multifactorial models of inheritance. *Clin Genet* **8**, 66-74 (1975).
13. Barnett, A.H., Eff, C., Leslie, R.D. & Pyke, D.A. Diabetes in identical twins. A study of 200 pairs. *Diabetologia* **20**, 87-93 (1981).
14. Haft, C.R. & Taylor, S.I. Deletion of 343 amino acids from the carboxyl terminus of the beta-subunit of the insulin receptor inhibits insulin signaling. *Biochemistry* **33**, 9143-51 (1994).
15. Almind, K. et al. Aminoacid polymorphisms of insulin receptor substrate-1 in non-insulin-dependent diabetes mellitus. *Lancet* **342**, 828-32 (1993).
16. Altshuler, D. et al. The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat Genet* **26**, 76-80 (2000).
17. Bento, J.L. et al. Association of protein tyrosine phosphatase 1B gene polymorphisms with type 2 diabetes. *Diabetes* **53**, 3007-12 (2004).
18. Maddux, B.A. et al. Membrane glycoprotein PC-1 and insulin resistance in non-insulin-dependent diabetes mellitus. *Nature* **373**, 448-51 (1995).
19. Liang, J., Fu, M., Ciociola, E., Chandalia, M. & Abate, N. Role of ENPP1 on adipocyte maturation. *PLoS One* **2**, e882 (2007).
20. Pizzuti, A. et al. A polymorphism (K121Q) of the human glycoprotein PC-1 gene coding region is strongly associated with insulin resistance. *Diabetes* **48**, 1881-4 (1999).
21. Meyre D, F.P. ENPP1, the first example of common genetic link between childhood and adult obesity and type 2 diabetes. *Med Sci (Paris)*. **30**, 233-7 (2006 Mar).

22. Bochenski J, P.G., Wanic K, Malecki M, Sieradzki J, Warram JH, Krolewski AS. New polymorphism of ENPP1 (PC-1) is associated with increased risk of type 2 diabetes among obese individuals. *Diabetes* **55**, 2626-30 (2006 Sep).
23. Chandalia, M., Grundy, S.M., Adams-Huet, B. & Abate, N. Ethnic differences in the frequency of ENPP1/PC1 121Q genetic variant in the Dallas Heart Study cohort. *J Diabetes Complications* **21**, 143-8 (2007).
24. Hamaguchi, K. et al. The PC-1 Q121 allele is exceptionally prevalent in the Dominican Republic and is associated with type 2 diabetes. *J Clin Endocrinol Metab* **89**, 1359-64 (2004).
25. Keene, K.L. et al. Association of the distal region of the ectonucleotide pyrophosphatase/phosphodiesterase 1 gene with type 2 diabetes in an African-American population enriched for nephropathy. *Diabetes* **57**, 1057-62 (2008).
26. Matsuoka, N. et al. Association of K121Q polymorphism in ENPP1 (PC-1) with BMI in Caucasian and African-American adults. *Int J Obes (Lond)* **30**, 233-7 (2006).
27. Grarup N, U.S., Ek J, Albrechtsen A, Glumer C, Borch-Johnsen K, Jorgensen T, Hansen T, Pedersen O. Studies of the relationship between the ENPP1 K121Q polymorphism and type 2 diabetes, insulin resistance and obesity in 7,333 Danish white subjects. *Diabetologia* **49**, 2097-104 (2006 Sep).
28. Baratta, R. et al. Role of the ENPP1 K121Q polymorphism in glucose homeostasis. *Diabetes* **57**, 3360-4 (2008).
29. Tasic I, M.M., Sunder-Plassmann R, Lazarevic G, Tasic NM, Stefanovic V. The association of PC-1 (ENPP1) K121Q polymorphism with metabolic syndrome in patients with coronary heart disease. *Clin Chim Acta.* **377**, 237-42 (2007 Feb).
30. Bottcher Y, K.A., Reinehr T, Enigk B, Kiess W, Stumvoll M, Kovacs P. ENPP1 variants and haplotypes predispose to early onset obesity and impaired glucose and insulin metabolism in German obese children. *J Clin Endocrinol Metab* **91**, 4948-52 (2006 Dec).
31. Matsuoka N, P.A., Tiwari HK, Allison DB, Johnson SB, Gregersen PK, Leibel RL, Chung WK. Association of K121Q polymorphism in ENPP1 (PC-1) with BMI in Caucasian and African-American adults. *Int J Obes (Lond).* **30**, 233-7 (2006 Feb).
32. Meyre D, B.-N.N., Tounian A, Samson C, Lecoecur C, Vatin V, Ghousaini M, Wachter C, Hercberg S, Charpentier G, Patsch W, Pattou F, Charles MA, Tounian P, Clement K, Jouret B, Weill J, Maddux BA, Goldfine ID, Walley A, Boutin P, Dina C, Froguel P. Variants of ENPP1 are associated with childhood and adult obesity and increase the risk of glucose intolerance and type 2 diabetes. *Nat Genet* **37**, 863-7 (2005 Aug).
33. Eller, P. et al. Impact of ENPP1 genotype on arterial calcification in patients with end-stage renal failure. *Nephrol Dial Transplant* **23**, 321-7 (2008).
34. Wu, L.S. et al. Association and interaction analyses of genetic variants in ADIPOQ, ENPP1, GHSR, PPARgamma and TCF7L2 genes for diabetic nephropathy in a Taiwanese population with type 2 diabetes. *Nephrol Dial Transplant* **24**, 3360-6 (2009).
35. Canani, L.H. et al. Polymorphism in ecto-nucleotide pyrophosphatase/phosphodiesterase 1 gene (ENPP1/PC-1) and early development of advanced diabetic nephropathy in type 1 diabetes. *Diabetes* **51**, 1188-93 (2002).
36. Keshavarz P, I.H., Sakamoto Y, Kunika K, Tanahashi T, Nakamura N, Yoshikawa T, Yasui N, Shiota H, Itakura M. No evidence for association of the ENPP1 (PC-1) K121Q variant with risk of type 2 diabetes in a Japanese population. *J Hum Genet* **51**, 559-66 (2006).

37. Lyon HN, F.J., Bersaglieri T, Saxena R, Winckler W, Almgren P, Lindblad U, Tuomi T, Gaudet D, Zhu X, Cooper R, Ardlie KG, Daly MJ, Altshuler D, Groop L, Hirschhorn JN. Common variants in the ENPP1 gene are not reproducibly associated with diabetes or obesity. *Diabetes* **55**, 3180-4 (2006 Nov).
38. Weedon MN, S.B., Hitman G, Walker M, McCarthy MI, Hattersley AT, Frayling TM. No evidence of association of ENPP1 variants with type 2 diabetes or obesity in a study of 8,089 U.K. Caucasians. *Diabetes* **55**, 3175-9 (2006 Nov).
39. Kumakura, S., Maddux, B.A. & Sung, C.K. Overexpression of membrane glycoprotein PC-1 can influence insulin action at a post-receptor site. *J Cell Biochem* **68**, 366-77 (1998).
40. Goldfine, I.D. et al. Membrane glycoprotein PC-1 and insulin resistance. *Mol Cell Biochem* **182**, 177-84 (1998).
41. Frittitta, L. et al. A cluster of three single nucleotide polymorphisms in the 3'-untranslated region of human glycoprotein PC-1 gene stabilizes PC-1 mRNA and is associated with increased PC-1 protein content and insulin resistance-related abnormalities. *Diabetes* **50**, 1952-5 (2001).
42. Kruglyak, L. & Nickerson, D.A. Variation is the spice of life. *Nat Genet* **27**, 234-6 (2001).
43. Thorisson, G.A., Smith, A.V., Krishnan, L. & Stein, L.D. The International HapMap Project Web site. *Genome Res* **15**, 1592-3 (2005).
44. Gabriel SB, S.S., Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M, Liu-Cordero SN, Rotimi C, Adeyemo A, Cooper R, Ward R, Lander ES, Daly MJ, Altshuler D. The structure of haplotype blocks in the human genome. *Science* **296**, 2225-9 (2002 Jun).
45. Tang, H., Peng, J., Wang, P. & Risch, N.J. Estimation of individual admixture: analytical and study design considerations. *Genet Epidemiol* **28**, 289-301 (2005).
46. Pfaff, C.L., Barnholtz-Sloan, J., Wagner, J.K. & Long, J.C. Information on ancestry from genetic markers. *Genet Epidemiol* **26**, 305-15 (2004).
47. Bunker, C.H. et al. High prevalence of screening-detected prostate cancer among Afro-Caribbeans: the Tobago Prostate Cancer Survey. *Cancer Epidemiol Biomarkers Prev* **11**, 726-9 (2002).
48. Miljkovic-Gacic, I., Ferrell, R.E., Patrick, A.L., Kammerer, C.M. & Bunker, C.H. Estimates of African, European and Native American ancestry in Afro-Caribbean men on the island of Tobago. *Hum Hered* **60**, 129-33 (2005).
49. Bondar, R.J. & Mead, D.C. Evaluation of glucose-6-phosphate dehydrogenase from *Leuconostoc mesenteroides* in the hexokinase method for determining glucose in serum. *Clin Chem* **20**, 586-90 (1974).
50. Matthews, D.R. et al. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* **28**, 412-9 (1985).
51. Roeder, K., Bacanu, S.A., Sonpar, V., Zhang, X. & Devlin, B. Analysis of single-locus tests to detect gene/disease associations. *Genet Epidemiol* **28**, 207-19 (2005).
52. Yuan, H.Y. et al. FASTSNP: an always up-to-date and extendable service for SNP function analysis and prioritization. *Nucleic Acids Res* **34**, W635-41 (2006).
53. Kosoy, R. et al. Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America. *Hum Mutat* **30**, 69-78 (2009).

54. de Bakker, P.I. et al. Transferability of tag SNPs in genetic association studies in multiple populations. *Nat Genet* **38**, 1298-303 (2006).
55. Barrett, J.C., Fry, B., Maller, J. & Daly, M.J. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263-5 (2005).
56. Gabriel, S.B. et al. The structure of haplotype blocks in the human genome. *Science* **296**, 2225-9 (2002).
57. Green, L.E., Lange, E.M. & Langefeld, C.D. Power comparisons of phase-known versus phase-unknown haplotype analyses for case-control designs. *Am. J. Hum. Genet* **69**, 1948a (2001).
58. Li, Q. & Yu, K. Improved correction for population stratification in genome-wide association studies by identifying hidden population structures. *Genet Epidemiol* **32**, 215-26 (2008).
59. Patterson, N., Price, A.L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet* **2**, e190 (2006).
60. Price, A.L. et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38**, 904-9 (2006).
61. Roeder, K. & Luca, D. Searching for disease susceptibility variants in structured populations. *Genomics* **93**, 1-4 (2009).
62. Lewis, J.P. et al. Association analysis in african americans of European-derived type 2 diabetes single nucleotide polymorphisms from whole-genome association studies. *Diabetes* **57**, 2220-5 (2008).
63. Leak, T.S. et al. Variants in intron 13 of the ELMO1 gene are associated with diabetic nephropathy in African Americans. *Ann Hum Genet* **73**, 152-9 (2009).
64. Miljkovic-Gacic, I. et al. Adipose tissue infiltration in skeletal muscle: age patterns and association with diabetes among men of African ancestry. *Am J Clin Nutr* **87**, 1590-5 (2008).
65. Prudente, S. et al. The Q121/Q121 genotype of ENPP1/PC-1 is associated with lower BMI in non-diabetic whites. *Obesity (Silver Spring)* **15**, 1-4 (2007).
66. Barroso, I. et al. Candidate gene association study in type 2 diabetes indicates a role for genes involved in beta-cell function as well as insulin action. *PLoS Biol* **1**, E20 (2003).
67. Meyre, D. et al. Variants of ENPP1 are associated with childhood and adult obesity and increase the risk of glucose intolerance and type 2 diabetes. *Nat Genet* **37**, 863-7 (2005).
68. Jenkinson, C.P. et al. Association of genetic variation in ENPP1 with obesity-related phenotypes. *Obesity (Silver Spring)* **16**, 1708-13 (2008).
69. Bochenski, J. et al. New polymorphism of ENPP1 (PC-1) is associated with increased risk of type 2 diabetes among obese individuals. *Diabetes* **55**, 2626-30 (2006).
70. Abate, N. et al. ENPP1/PC-1 K121Q polymorphism and genetic susceptibility to type 2 diabetes. *Diabetes* **54**, 1207-13 (2005).
71. Grarup, N. et al. Studies of the relationship between the ENPP1 K121Q polymorphism and type 2 diabetes, insulin resistance and obesity in 7,333 Danish white subjects. *Diabetologia* **49**, 2097-104 (2006).
72. McAteer, J.B. et al. The ENPP1 K121Q polymorphism is associated with type 2 diabetes in European populations: evidence from an updated meta-analysis in 42,042 subjects. *Diabetes* **57**, 1125-30 (2008).

73. Prudente, S. & Trischitta, V. Editorial: The pleiotropic effect of the ENPP1 (PC-1) gene on insulin resistance, obesity, and type 2 diabetes. *J Clin Endocrinol Metab* **91**, 4767-8 (2006).
74. Zhou, H.H. et al. Suppression of PC-1/ENPP-1 expression improves insulin sensitivity in vitro and in vivo. *Eur J Pharmacol* **616**, 346-52 (2009).
75. Leak, T.S. et al. Evaluation of a SNP map of 6q24-27 confirms diabetic nephropathy loci and identifies novel associations in type 2 diabetes patients with nephropathy from an African-American population. *Hum Genet* **124**, 63-71 (2008).
76. Duggirala, R. et al. A major locus for fasting insulin concentrations and insulin resistance on chromosome 6q with strong pleiotropic effects on obesity-related phenotypes in nondiabetic Mexican Americans. *Am J Hum Genet* **68**, 1149-64 (2001).
77. Shaffer, J.R. et al. Genetic markers for ancestry are correlated with body composition traits in older African Americans. *Osteoporos Int* **18**, 733-41 (2007).
78. Wassel Fyr, C.L. et al. Genetic admixture, adipocytokines, and adiposity in Black Americans: the Health, Aging, and Body Composition study. *Hum Genet* **121**, 615-24 (2007).
79. Peralta, C.A. et al. African ancestry, socioeconomic status, and kidney function in elderly African Americans: a genetic admixture analysis. *J Am Soc Nephrol* **17**, 3491-6 (2006).
80. Fernandez, J.R. et al. Association of African genetic admixture with resting metabolic rate and obesity among women. *Obes Res* **11**, 904-11 (2003).
81. Zuo, L., Kranzler, H.R., Luo, X., Covault, J. & Gelernter, J. CNR1 Variation Modulates Risk for Drug and Alcohol Dependence. *Biol Psychiatry* (2007).
82. Gloyn, A.L. et al. Large-scale association studies of variants in genes encoding the pancreatic beta-cell KATP channel subunits Kir6.2 (KCNJ11) and SUR1 (ABCC8) confirm that the KCNJ11 E23K variant is associated with type 2 diabetes. *Diabetes* **52**, 568-72 (2003).
83. Grant, S.F. et al. Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet* **38**, 320-3 (2006).
84. Scott, L.J. et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* **316**, 1341-5 (2007).
85. Sladek, R. et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* **445**, 881-5 (2007).
86. Zeggini, E. et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* **316**, 1336-41 (2007).
87. Saxena, R. et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* **316**, 1331-6 (2007).
88. Steinthorsdottir, V. et al. A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. *Nat Genet* **39**, 770-5 (2007).
89. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661-78 (2007).