Association Study of Transcription Factors Regulating Insulin Secretion and Action in Type 2 Diabetes in Chinese

HO Sin Ka Janice

A Thesis Submitted in Partial Fulfilment

of the Requirements for the Degree of

Master of Philosophy

in

Medical Sciences

Supervisor: Dr Maggie CY Ng

© The Chinese University of Hong Kong

May 2008

The Chinese University of Hong Kong holds the copyright of this thesis. Any person(s) intending to use a part or whole of the materials in the thesis in a proposed publication must seek copyright release from the Dean of the Graduate School.

香港中文大學
大學圖書館系統
20 JAN 2010
UNIVERSITY LIBRARY SYSTEM
THE CHINESE UNIVERSITY OF HONG KONG

Abstract for the thesis entitled

Association Study of Transcription Factors Regulating

Insulin Secretion and Action in Type 2 Diabetes in Chinese

Submitted by Ho Sin Ka Janice

For the degree of Master of Philosophy in Medical Sciences

At the Chinese University of Hong Kong in May 2008

Type 2 diabetes (T2D) is a highly prevalent complex disease characterized by defective insulin secretion in the pancreatic β cells and insulin resistance in the liver, muscle and adipose tissues, but their pathogenic mechanisms are not fully understood. This study hypothesized that common polymorphisms in the genes encoding nutrients-, glucose-, and insulin-responsive transcriptions factors and their downstream targets may be implicated in these pathways and influence susceptibility to T2D.

Eight candidate genes from the insulin secretion pathway (*HNF4A, HNF1A, PDX1, PBX1, NEUROD1, GCK, KCNJ11* and *ABCC8*) and six candidate genes from the insulin action pathway (*PPARG, PPARA, PPARGC1A, ADIPOQ, LPL* and *LIPC*) were selected respectively. A two-stage case-control association study in 1462 T2D cases and 600 controls from the Hong Kong Chinese population was conducted. The first stage investigated 152 SNPs in 467 T2D cases with early onset familial diabetes and 290 controls. The second stage followed up 22 nominally significant SNPs in 994 random T2D cases and 310 controls. The association between genetic polymorphisms

and T2D, T2D-subtypes and metabolic trait, as well as gene-gene interactions and additive effects were explored.

The key findings of this study included 1) polymorphisms in *HNF4A, ADIPOQ, LPL* and *PPARA* were associated with T2D susceptibility; 2) different forms of diabetes were attributed by genes from different pathways. Genes implicated primarily in insulin secretion (e.g. *HNF4A*) lead to early onset T2D, whereas genes implicated in insulin action pathway (e.g *ADIPOQ*) conferred risk for T2D later in life with coexistence of metabolic syndrome; 3) these genes likely regulate both glucose and lipid metabolisms, and their joint effects had substantial stronger implications for T2D development.

中文摘要

2 型糖尿病(T2D)是一種非常普遍的複雜性疾病，具有胰腺β細胞胰島素分泌不足及肝臟、肌肉及脂肪組織胰島素抵抗綜合症的特性，但其發病機制尚未完全清楚。本研究假設胰島素反應性轉錄因子基因及其下游靶基因中的常見多態性與 2 型糖尿病致病通路相關，因而影響其發病易感性。

本研究應用兩階段病例對照的研究方法分析了胰島素分泌通路中的 8 個候選基因（包括肝細胞核因子 4A，肝細胞核因子 1A，胰腺及十二指腸同源框 1，前 ß 淋巴細胞白血病轉錄因子，神經分化因子 1，葡萄糖激酶，內向整流鉀通道 和磺脲類藥物受體 1)及胰島素作用通路中的 6 個候選基因（包括過氧化物酶體增殖物激活受體 G，過氧化物酶體增殖物激活受體 A，過氧化物增殖激活受體協同激活子 1A，脂聯素，脂蛋白脂酶，和肝脂酶）。以 1462 名 2 型糖尿病患者及 600 名健康對照者（皆為香港中國人）為研究對象，進行基因檢測及分析。第一階段在 467 名早發家族性 2 型糖尿病患者及 290 名健康對照者中調查了 152 個基因多態性, 而其中 22 個重點多態性則在第二階段（994 名隨機 2 型糖尿病患者及 310 名健康對照者）中加以調查。本研究着重探討基因多態性與 2 型糖尿病，糖尿

病亞型(T2D-subtype)及其代謝特徵的相關性， 同時也探索了基因與基因的相互及相加作用對糖尿病及其亞型的影響。

本研究的主要發現如下：

1. 肝細胞核因子 4A、脂聯素、脂蛋白脂酶及過氧化物酶體增殖物激活受體 A 基因多態性與 2 型糖尿病易感性相關。
2. 不同通路的基因與不同類型的 2 型糖尿病亞型相關。胰島素分泌通路中的基因（如肝細胞核因子 4A）主要與早發 2 型糖尿病相關；而胰島素作用通路中的基因（如脂聯素）則增加伴隨代謝綜合症的遲發 2 型糖尿病的發病危險性。
3. 這些基因可能通過對葡萄糖、酯脂代謝及其聯合作用的調控，在 2 型糖尿病的發生發展中發揮了重要作用。

Publication

Poster presentation

J. Ho, S. Germer, M. Ng, R. Ma, W. So, M. Martin, J. Chan, Associations of HNF4α Variants with Type 2 Diabetes in Hong Kong Chinese [abstract program number 2379]. Presented at the annual meeting of The American Society of Human Genetics, October 24, 2007, San Diego, California. Available from http://www.ashg.org/genetics/ashg07s/index.shtml

Grant Support

This study was supported by the Hong Kong Innovation and Technology Support Fund (ITS/33/00) and the Research Grants Council Central Allocation Scheme (CUHK 1/04C).

Author's Contribution

All study subjects were recruited and phenotyped by the staff at Prince of Wales Hospital Diabetes and Endocrine Center. The phenotype data were maintained by Mr. Lunan Chow, Ms. Claudia Tam, Dr. Xilin Yang and Mr. Kevin Yu in the Prince of Wales Hospital Diabetes Registry. Biochemical measurements were performed by Mr. Stanley Ho, Mr. Alex Ng, Ms. Patty Tse and Mr. Vincent Lam at the Department of Medicine and Therapeutics, and laboratory at the Department of Chemical Pathology. Janice Ho extracted 120 DNA samples from blood, and the rest were extracted by fellow endocrine team members. With advices from Dr. Maggie Ng, Dr. Mitchell Martin and Dr. Soren Germer, Janice Ho selected all candidate genes and tag SNPs. Two SNPs in stage 1 were genotyped by Janice Ho, and the rest were genotyped by the staff at Roche Pharmaceutical, Hong Kong University, or McGill University and Genome Quebec Innovation Center. Janice Ho performed Y chromosome assay on stage 1 samples as a sample quality control. Janice Ho performed all data quality control and statistical analyses. Janice Ho wrote the thesis, and Dr. Maggie Ng edited it.

Dedication

In memory of my grandma Yeung Suet Ying,

I hope I have made you proud.

To my beloved parents,

without their love, support

and continuous supply of coffee,

this thesis could not have been completed.

Acknowledgement

I am thankful to my supervisor, Dr. Maggie Ng, for her guidance throughout the study. Learning from her has been a life transforming experience, and I appreciate her many advices. I am fortunate to be her student.

I am deeply grateful to Dr. Juliana Chan for initiating my study, and for sharing her vision on diabetes. She has inspired my passion for research. I thank Dr. Wing Yee So and Dr. Ronald Ma for their help on the clinical aspect of this study.

I appreciate the expertise that Dr. Mitchell Martin and Dr. Soren Germer at Roche Pharmaceutical have provided on this project. It was a pleasure to collaborate with them. I thank Mr. Vincent Lam for his help in sample preparation.

I am most grateful to the Prince of Wales Hospital (PWH) Diabetes and Endocrine Centre staff for recruiting and phenotyping patients, to Mr. Lunan Chow, Mr. Kevin Yu, and Dr. Xilin Yang for managing the PWH Diabetes Registry, to Ms. Ruchi Choski and Ms. Polliana Macedo at Roche Pharmaceutical for genotyping SNPs, and to all study subjects for participating in this study. This project would not have been possible without their many contributions.

I thank Ms. Janice Mong, Mr. Alex Ng and Ms. Claudia Tam for their generous help throughout my research. Their companionship lightened those difficult days. I thank Ms. Emily Poon, Ms. Cindy Tong, Ms. Patty Tse and Dr. Ying Wang for their invaluable assistances in the laboratory.

To see a world in a grain of genes

And a disease in a wild allele

Hold significance in the palm of one hand

And replication in another

— William Blake, *Auguries of innocence*, modified by Janice Ho

Table of Contents

CHAPTER 1. Introduction

1.1. Epidemiology of Type 2 Diabetes 1

1.2. Risk factors contributing to Type 2 Diabetes.................... 3

1.2.1. Environmental and physiological factors 3

1.2.2. Genetic factors 3

1.3. Disruption of energy homeostasis in the pathogenesis of type 2 diabetes.......6

1.3.1. Clinical spectrum of diabetes 6

1.3.2. Insulin as a key regulator of energy homeostasis....................7

1.3.3. Insulin secretion and glucose metabolism....................8

1.3.4. Insulin action and lipid metabolism....................9

1.3.5. Lipotoxicity and glucotoxicity....................12

1.3.6. Role of transcription factors as metabolic switch....................13

1.4. Candidate genes implicated in type 2 diabetes susceptibility....................15

1.4.1. Candidate genes involved in insulin secretion pathway....................15

1.4.1.1. *HNF4A*....................15

1.4.1.2. *HNF1A*....................16

1.4.1.3. *PDX1/PBX1*....................17

1.4.1.4. *NEUROD1*....................17

1.4.1.5. *GCK*....................17

1.4.1.6. *KCNJ11/ABCC8*....................18

1.4.2 Candidate genes involved in insulin action pathway....................19

1.4.2.1. *PPARG*....................19

1.4.2.2.*PPARA*....................20

1.4.2.3. *PPARGC1A*....................20

1.4.2.4. *ADIPOQ*....................21

1.4.2.5. *LPL*....................21

1.4.2.6. *LIPC*....................22

1.5. Hypothesis and objectives of the study 23

CHAPTER 2. Materials and methods

2.1. Study design....................25

2.1.1. Two-stage candidate gene association design....................25

2.1.2. Power calculation....27
2.2. Study cohort....29
2.2.1. Subject recruitment....29
2.2.2. Clinical and biochemical measurements....30
2.2.3. Clinical definitions....31
2.3. Genetic study....32
2.3.1. Candidate gene selection....32
2.3.2. SNP selection....32
2.3.3. DNA sample preparation....35
2.3.4. Genotyping methods....36
2.3.4.1. Allele specific Tm shift assay....36
2.3.4.2. Mass spectrometry assay....40
2.4. Data quality control....42
2.4.1. Stage 1....42
2.4.2. Stage 2....42
2.5. Statistical analysis....45
2.5.1. Stage 1 analysis....45
2.5.2. Stage 2 analysis....45
2.5.3. Stage 1 and 2 combined analysis....46

CHAPTER 3. Results
3.1. Clinical characteristics of subjects in stages 1 and 2 studies....48
3.2. Case-control associations in stage 1....51
3.2.1. Association with T2D....51
3.2.2. Association with T2D subset by metabolic syndrome....54
3.3. Case-control associations in stage 2....60
3.3.1. SNP selection for genotyping....60
3.3.2. Association with T2D....63
3.3.3. Association with T2D subset by metabolic syndrome....64
3.4. Case-control associations in combined stages 1 and 2....66
3.4.1. Association with T2D....66
3.4.2. Association with T2D subset by metabolic syndrome....70
3.4.3. Association with T2D subset by age at diagnosis....74
3.4.4. Association with T2D subset by gender....76

3.4.5. Genetic epistasis for T2D association.. 79
3.5. Metabolic traits associations in control subjects in combined stages 1 and 2 studies 83

CHAPTER 4. Discussion....86
4.1. Role of insulin secretion genes in type 2 diabetes 87
4.2. Role of insulin action genes in type 2 diabetes 92
4.3. Combined genetic effects on risk for type 2 diabetes 97
4.4. Summary 98
4.5. Limitation of this study and future direction 101

REFERENCES....104

APPENDICES....119
Appendix 1: Gene structure and linkage disequilibrium of genotyped SNPs of candidate genes....119
Appendix 2: Information of SNPs genotyped in stage 1....130
Appendix 3: T2D association results (additive model) of 152 SNPs for stage 1 case-control samples....137
Appendix 4: T2D association results (additive model) of 152 SNPs for stage 1 case-control samples subset by metabolic syndrome status in cases....144
Appendix 5: T2D association results (additive model) of 22 SNPs for stage 2 case-control samples....151
Appendix 6: T2D association results (additive model) of 22 SNPs for stage 2 case-control samples subset by metabolic syndrome status in cases....153

List of Tables

Table 2.1: D' and r^2 color scheme in Haploview....................................34

Table 2.2: Summary of SNP selection and genotyping....................................43

Table 3.1: Clinical characteristics of subjects in stages 1 and 2 studies....................49

Table 3.2: Significant T2D association results of SNPs in stage 1 case-control samples....................................53

Table 3.3: Significant T2D association results of SNPs in stage 1 case-control samples subset by metabolic syndrome status in cases....................................58

Table 3.4 Summary of 22 SNPs successfully genotyped in stage 2....................61

Table 3.5: T2D association results of SNPs in the stage 1+2 case-control samples....68

Table 3.6: Haplotype association of significant genes in stage 1+2 case-control samples....................................69

Table 3.7: T2D association results of SNPs in the stage 1+2 case-control samples subset by metabolic syndrome status in cases....................................72

Table 3.8: T2D association results of SNPs in the stage 1+2 case-control samples subset by age at diagnosis in cases....................................75

Table 3.9: T2D association results of SNPs in the stage 1+2 case-control samples subset by gender....................................77

Table 3.10: Dosage effect of risk alleles of five significant SNPs in stage 1+ 2 case-control association....................................82

Table 3.11: Significant association of candidate genes with metabolic traits in 600 normal controls....................................85

List of Figures

Figure 1.1: Estimated increasing prevalence of diabetes from 2003 to 2025............2

Figure 1.2: Two pathways implicated in the pathogenesis of T2D........................5

Figure 1.3: Hyperbolic relation between β-cell function and insulin sensitivity.........7

Figure 1.4: Glucose stimulated insulin secretion in pancreatic β cells...................9

Figure 1.5: Insulin action on target tissues..10

Figure 1.6: FFA action on target tissues...11

Figure 1.7: Effects of glucotoxicity and lipotoxicity on insulin secretion and action pathways..13

Figure 2.1: Power calculation under a allelic model at 5% type 1 error rate for total sample sizes..28

Figure 2.2: Protocol for allele specific T_m shift assay.....................................39

Figure 2.3: Sequenome i-PLEX genotyping protocol.......................................41

Figure 3.1: Graphical summary of association results for T2D for stage 1 SNPs in case-control samples 52

Figure 3.2: Graphical summaries of association results for stage 1 SNPs 57

Figure 3.3: Graphical summary of association results for T2D for stage 2 SNPs in case-control samples..63

Figure 3.4: Graphical summaries of association results for stage 2 SNPs 65

Figure 3.5: Graphical summary of association results for metabolic traits in stage 1 and 2 control samples 84

List of Abbreviations

AAD	Age of diagnosis
BMI	Body mass index
CHB	Han Chinese population in HapMap
ChIP	Chromatin immunoprecipitation
CI	Confidence interval
DBP	Diastolic blood pressure
DF	Degree of freedom
FCHL	Familial combined hyperlipidemia
FFA	Free fatty acid
FINS	Fasting plasma insulin
FPG	Fasting plasma glucose
FPLD	Family partial lipodystrophy
GLU30	Glucose at OGTT 30 mins
GSIS	Glucose stimulated insulin secretion
GWAS	Genome-wide association studies
HDL	High-density lipoprotein cholesterol
HOMA-IR	Homeostasis model assessment for insulin resistance
HOMA-β	Homeostasis model assessment for β cell function
HWE	Hardy Weinberg equilibrium
ID	Insulin secretory deficiency
IDI	Insulin disposition index
IDL	Intermediate-density lipoprotein
IGT	Impaired glucose tolerance
INS30	Insulin at OGTT 30 mins
IR	Insulin resistance
ISI	Insulin sensitivity index

LD	Linkage disequilibrium
LDL	Low-density lipoprotein cholesterol
MAF	Minor allele frequency
MALDI-TOF MS	Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry
MetS	Metabolic syndrome
MODY	Maturity-onset diabetes of the young
NCBI	The National Center for Biotechnology Information
NCEP III	National Cholesterol Education Program Adult Treatment Panel III
NGT	Normal glucose tolerance
OD	Optical density
OGTT	Oral glucose tolerance test
OR	Odds ratio
PCR	Polymerase chain reaction
PNDM	Permanent neonatal diabetes
PPRE	Peroxisome proliferator response element
PWH	Prince of Wales Hospital
QC	Quality control
SBP	Systolic blood pressure
SD	Standard deviation
SNP	Single nucleotide polymorphism
T1D	Type 1 diabetes
T2D	Type 2 diabetes
TF	Transcription factor
TG	Triglyceride
Tm	Melting temperature
VLDL	Very-low density lipoprotein

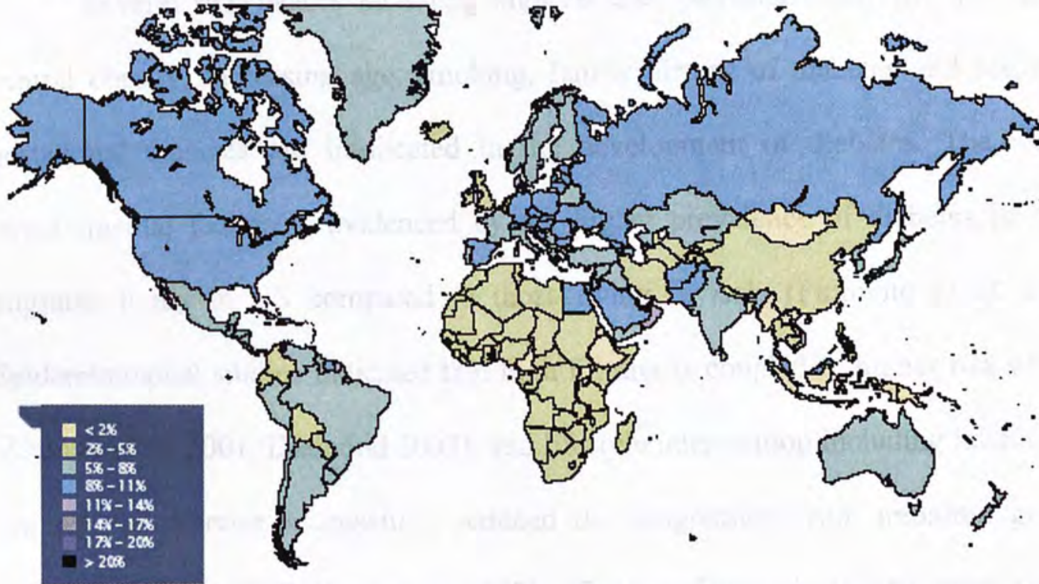
WHO	World Health Organization
WHR	Waist-to-hip ratio

CHAPTER 1. Introduction

1.1. Epidemiology of Type 2 Diabetes

Diabetes is one of the most prevalent complex diseases in the world. Over 150 million individuals are affected globally, and another 150 million are estimated to remain undiagnosed (Zimmet *et al.* 2001; Diamond 2003). The number of diabetic patients is projected to increase to 366 million in 2030 (Wild *et al.* 2004). While type 1 diabetes (T1D) is relatively rare, type 2 diabetes (T2D) is the predominant form accounting for more than 90% of all forms of diabetes. T2D represents a cluster of metabolic diseases characterized by elevated glucose level, and its development is subjected to genetic and environmental factors. Epidemiological studies revealed that T2D is more prevalent in populations that adopt an affluent lifestyle such as Mauritius, Singapore and Taiwan (13%, 12% and 9%) compared to less affluent areas such as rural China (< 5%) (Zimmet *et al.* 2001; Diamond 2003). Faced with the increasingly westernized lifestyle in developing countries such as China and India, the World Health Organization (WHO) predicted that one-third of the world's diabetics will be Asians by year 2030 (Wild *et al.* 2004). Poorly controlled diabetes is associated with various complications such as retinopathy, nephropathy, neuropathy, cardiovascular diseases and stroke (Meetoo *et al.* 2007), and treatments of these diseases pose major therapeutic challenges. In addition to deteriorating the quality of life in the diabetic patients, these diseases also strain the public health care system by accounting for 15% of annual health care costs in the US (Zimmet *et al.* 2001; Diamond 2003). In light of the enormous medical, social and economic burden that T2D imposes on the world especially in Asia, a better understanding of its pathogenesis is important for controlling this disease.

Map 1.1
Prevalence estimates of diabetes, 2003



Map 1.2
Prevalence estimates of diabetes, 2025

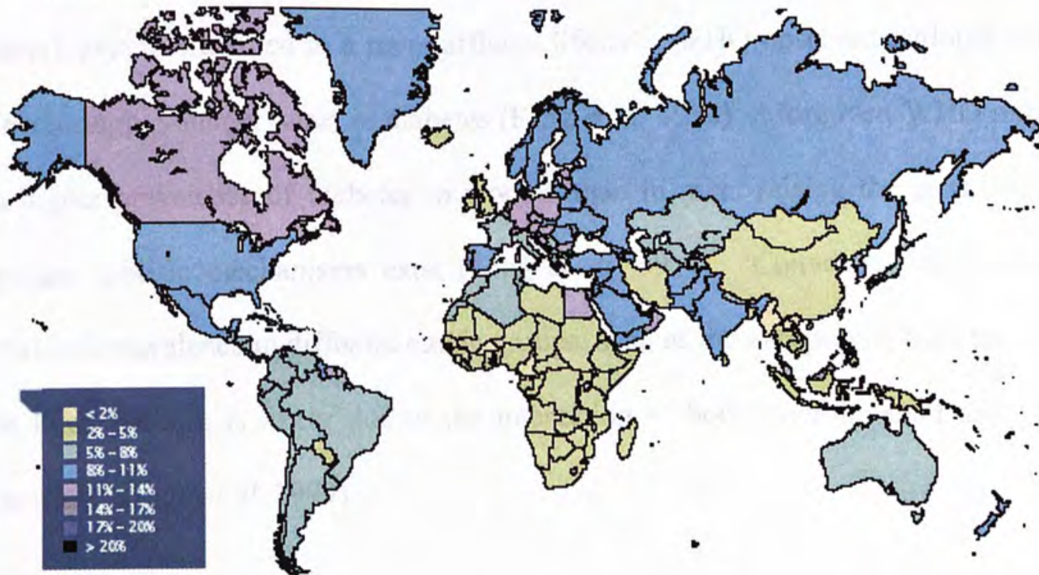


Figure 1.1: Estimated increasing prevalence of diabetes from 2003 to 2025. Adopted from International Diabetes Federation, 2003 (International Diabetes Federation 2003)

1.2. Risk factors contributing to Type 2 Diabetes

1.2.1. Environmental and physiological factors

Several risk factors including high fat diet, physical inactivity, general and central obesity, increasing age, smoking, family history of diabetes and history of gestational diabetes are implicated in the development of diabetes. The role of environmental factors is evidenced by the higher prevalence of diabetes in Asian migrants living in US compared to those living in Asia (Fujimoto *et al.* 1994). Epidemiological studies indicated that high fat diet is coupled to higher risk of T2D (Zimmet *et al.* 2001; Diamond 2003), and lifestyle intervention including healthy diet and regular exercise successfully reduced the progression from impaired glucose tolerance (IGT) to T2D (Sato *et al.* 2007). The beneficial role of long term exercise has also been shown to ameliorate T2D (Sato *et al.* 2007). With the socio-economic development which led to a more affluent lifestyle, WHO observed a global trend of increasingly younger onset of diabetes (King *et al.* 1998). Moreover, WHO indicated a higher prevalence of diabetes in women than in men, raising the possibility that gender-specific mechanisms exist (King *et al.* 1998). Conversely, differences in diabetes prevalence in different ethnic groups, such as the exceptional high prevalence in Pima Indians, is likely due to the interaction of both environmental and genetic factors (Savage *et al.* 1979).

1.2.2. Genetic factors

The genetic component of T2D is evidenced by family studies. Higher concordance rate of diabetes was observed in monozygotic twins (41% to 55%) compared to dizygotic twins (10% to 15%) (Harvald *et al.* 1963; Newman *et al.* 1987).

Offspring with both parents affected also had higher prevalence of diabetes (60% at 60 years old) compared to those with only one parent affected (38% at 80 years old) (Pierce *et al.* 1995; The DECODE study group 1999). Heritability studies also demonstrated high heritability (h^2 = 0.45-0.63) in diabetes related traits including anthropometric indices, blood pressure, lipids, insulin resistance and β-cell function (Li *et al.* 2006).

Depending on the penetrance of the at risk genetic variants, diabetes is divided into two forms: monogenic or polygenic. Monogenic diabetes is rare, usually caused by single gene mutations with high penetrance. They are characterized by severe defects in either insulin secretion (e.g. maturity-onset diabetes of the young (MODY), maternally inherited diabetes with deafness, permanent and transient neonatal diabetes) or insulin action (e.g. insulin receptor gene mutation and familial lipodystrophy) (Owen *et al.* 2007). In contrast, the more common form of T2D is likely caused by many common and/or rare genetic variants with modest effects, and additional gene-gene and gene-environment interactions are often required for the manifestation of the disease. Since the presentation of elevated blood glucose resulted from the convergence and interaction of diverse primary disturbances (McCarthy *et al.* 2002), numerous genetic loci in multiple pathways may be involved. Moreover, these genetic variants may exert different effects on different forms of diabetes (e.g. diabetes with or without other metabolic risks) or in different populations, and interactions with environmental factors may also modify their effects. Thus, T2D may be considered as a lifestyle disorder with higher prevalence in populations with greater genetic susceptibility (Zimmet *et al.* 2001; Diamond 2003). This complex interplay among genetic and environmental factors poses difficulty in studying the genetics of T2D. Fortunately, studies in monogenic diabetes had unmasked some

critical genes implicated in the impairment of insulin secretion (e.g. *GCK, HNF4A, HNF1A, TCF2, PDX1, NEUROD1, INS, KCNJ11, ABCC8, PLAGL1 (ZAC)/HYMA1, CEL* and mitochondria A3243G) and insulin sensitivity (e.g. *INSR, LMNA, LMNB2, PPARG, AKT2, AGPAT2* and *BSCL2*) in the pathogenesis of diabetes (Owen *et al.* 2007). This leads to the hypothesis that at least two distinct groups of genes from the two respective pathways may be implicated in the development of the more common form of T2D.

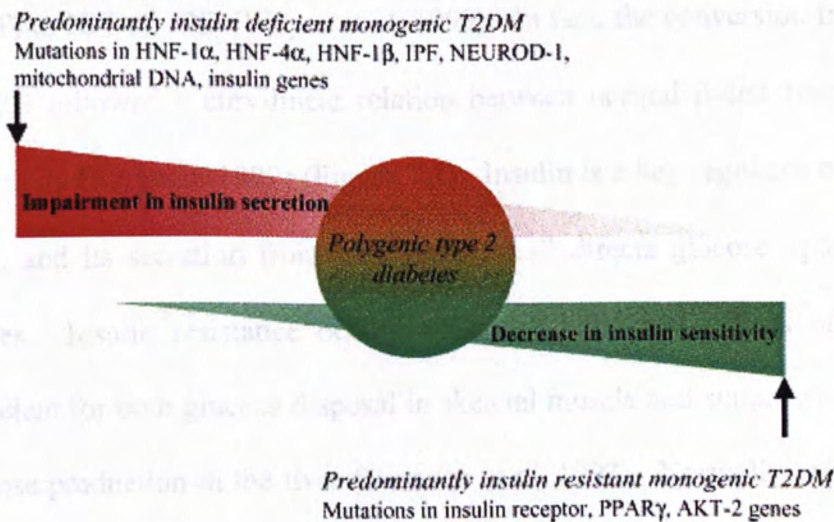


Figure 1.2: Two pathways implicated in the pathogenesis of T2D. Adopted from Malecki, 2005 (Malecki 2005)

1.3. Disruption of energy homeostasis in the pathogenesis of type 2 diabetes

1.3.1. Clinical spectrum of type 2 diabetes

Unlike those of monogenic diabetes, patients with T2D often display defects in both insulin secretion and sensitivity at varying degrees. The large number of possible interactions between the environmental and genetic factors underlies this wide pathophysiological spectrum. Insulin resistance (IR) often precedes IGT and overt T2D by many years (Kahn 1994). However, only in the presence of pancreatic β-cell dysfunction will IR triggers the progression from normal glucose tolerance (NGT) to IGT to T2D (Weyer *et al.* 1999). In fact, the conversion from NGT to IGT to T2D followed a curvilinear relation between normal β-cell function and insulin sensitivity (Bergman 1989) (Figure 1.3). Insulin is a key regulator of plasma glucose level, and its secretion from pancreatic β-cell directs glucose uptake in peripheral tissues. Insulin resistance occurs when the biological effects of insulin are not sufficient for both glucose disposal in skeletal muscle and suppression of endogenous glucose production in the liver (Dinneen *et al.* 1992). Normally, when IR is present, β cells can compensate for the decreased insulin sensitivity by hyper-secreting insulin to maintain NGT status, even though they have moved up along the glucose intolerance hyperbola. When individuals fail to compensate for IR by hyper-secretion of insulin, insulin secretory deficiency (ID) occurs and they become IGT and later T2D.

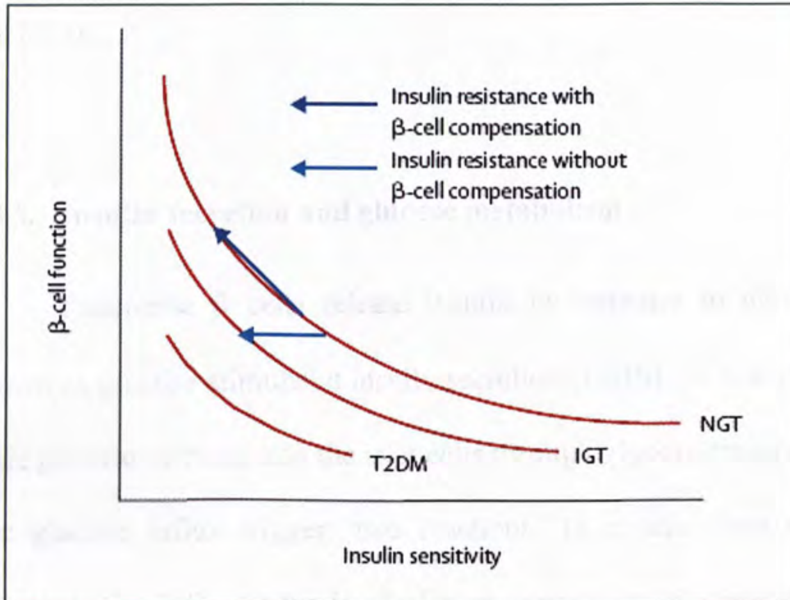


Figure 1.3: Hyperbolic relation between β-cell function and insulin sensitivity. Adopted from Stumvoll *et al.* (Stumvoll *et al.* 2005)

1.3.2. Insulin as a key regulator of energy homeostasis

Throughout evolutionary history, humans (and mammals in general) frequently faced shortage of food. Thus, survival depends on a species' efficiency on glucose and fat utilization. Fat contains twice the energy per unit mass of carbohydrates, but requires more energy to be metabolized (Rosen *et al.* 2006). Plasma glucose level is maintained between 3.9 and 6.1 mmol/l for normal function. When energy is abundant (i.e. during feeding state), insulin is secreted to promote glucose and free fatty acid (FFA) uptake in target tissues. It also prevents liver and adipose tissues from releasing stored glucose and FFA. Thus insulin maintains energy homeostasis by removing excess fuels from blood and storing them appropriately. When energy is deficient (i.e. during fasting state), insulin is not

secreted, and liver and adipose tissues maintain homeostasis by circulating glucose and FFAs.

1.3.3. Insulin secretion and glucose metabolism

Pancreatic β cells release insulin in response to elevated plasma glucose, known as glucose stimulated insulin secretion (GSIS). When plasma glucose level is high, glucose diffuses into the islet cells through glucose transporter type 2 (GLUT2). The glucose influx triggers two reactions: 1) insulin gene expression; 2) insulin secretion (Fig 1.4). At the level of gene expression, glucose stimulates transcription factors (TFs) such as hepatic nuclear factor 4α (HNF4A), neurogenic differentiation 1 (NEUROD1), and the heterodimers of pre-B-cell leukemia homeobox 1 and pancreatic and duodenal homeobox 1 (PBX1/PDX1). These TFs interact to upregulate insulin transcription (Wang *et al.* 2005). At the level of insulin secretion, glucose stimulates HNF4A, which in turn activates hepatic nuclear factor 1α (HNF1A) and promotes GLUT2 transcription to uptake more glucose. Glucokinase (GCK) controls the first rate-limiting step of glycolysis by activate the phosphorylation of glucose into glucose 6-phosphate, followed by mitochondrial oxidative phosphorylation to generate ATP. The increased ATP/ADP ratio leads to closure of the ATP-sensitive potassium channels (KCNJ11/ABCC8), followed by depolarization of plasma membrane and influx of calcium. This calcium influx then leads to exocytosis of insulin from their stored vesicles.

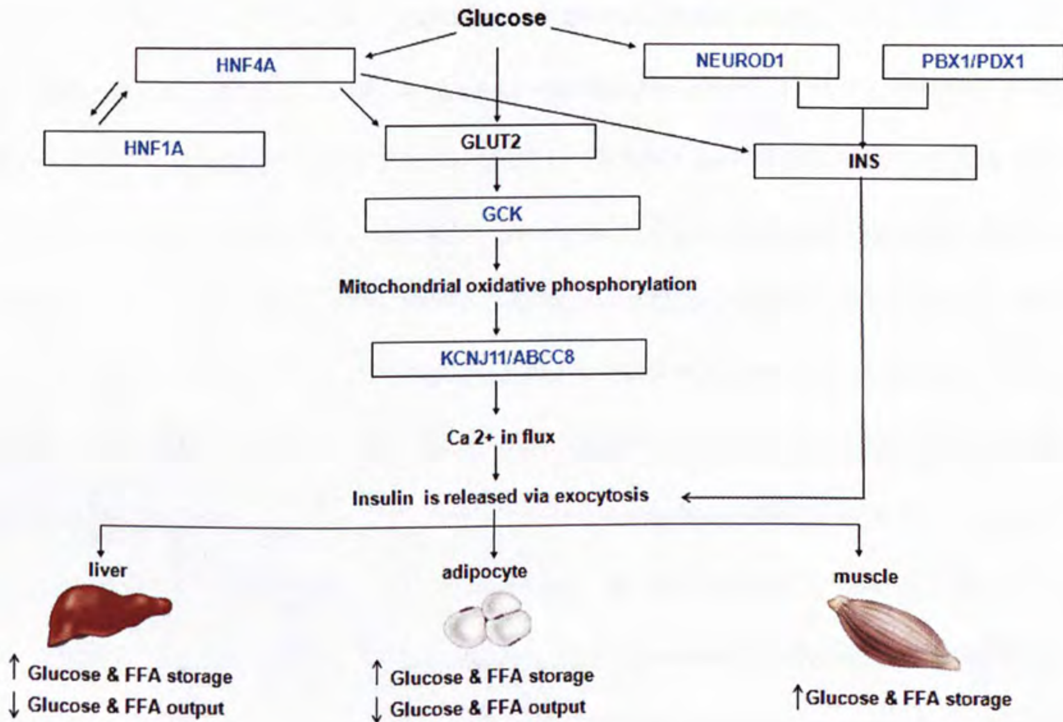


Figure 1.4: Glucose stimulated insulin secretion in pancreatic β cells. Blue colored genes were genotyped.

1.3.4. Insulin action and lipid metabolism

Insulin exerts two major effects in its target tissues: 1) promotes glucose and FFA uptake in the liver, muscle, and adipose tissues; 2) prevents glucose and FFA output by the liver and adipose tissues (Fig 1.5). The efficiency of insulin action is dependent on the tissue's insulin sensitivity. For fuel uptake, insulin binds to receptors on the cell surface, activating master transcription factors such as peroxisome proliferator-activated receptor γ subtype (PPARG) in adipose tissue, α subtype (PPARA) in liver, and PPARG coactivator 1α (PPARGC1A) in muscle (Kota *et al.* 2005; Handschin *et al.* 2006). These TFs in turn activate downstream target genes to achieve a coordinated response in glucose and lipid metabolisms. For

example, they up regulate the expression of glucose transporter type 4 (GLUT4) to promote glucose uptake in their respective tissues (Hamm *et al.* 1999; Michael *et al.* 2001; Kota *et al.* 2005; Armoni *et al.* 2007). PPARG and PPARA also up regulate the expression of lipases such as lipoprotein lipase (LPL) and hepatic lipase (LIPC) in adipose tissue and liver respectively, so that ingested triglyceride (Wittrup *et al.*) can be released as FFA into the blood stream and oxidized by other cells (Schoonjans *et al.* 1996). In addition, PPARG and PPARA up regulate lipogenic gene expression such as fatty acid synthase (FAS) and acetyl-CoA carboxylase (ACC) to convert excess glucose and FFA into triglycerides for storage in the adipose tissue and the liver (Motojima *et al.* 1998). Insulin also induces the expression of adiponectin (ADIPOQ) to interact with these TFs to promote glucose and FFA oxidation in muscle (Tomas *et al.* 2002; Yamauchi *et al.* 2002). Conversely, insulin inhibits the break down of stored glycogen and TG, forcing the body to metabolize the new substrates.

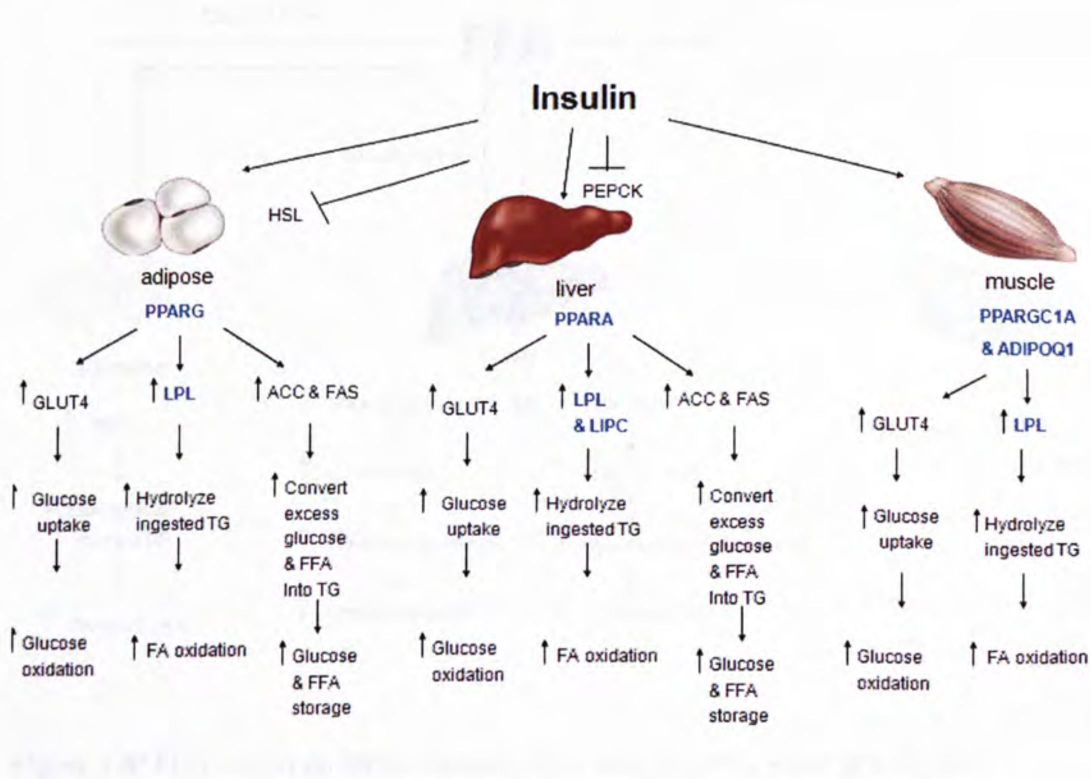


Figure 1.5: Insulin action on target tissues. Blue colored genes were genotyped.

During energy deficient state (in which insulin is absent), liver and adipose tissues maintain energy homeostasis by releasing glucose and FFA. In the liver, HNF4A and PPARGC1A cooperate to activate gluconeogenic genes such as phosphoenolpyruvate carboxykinase (PEPCK) and glucose-6-phosphatase (G6P) to generate glucose from carbon precursors (Herzig *et al.* 2001; Yoon *et al.* 2001). In the adipose tissue, hormone sensitive lipase (HSL) releases free FA (FFA) by hydrolyzing stored TG. The high energy FFA can act as a substrate as well as a signaling molecule, and promotes glycogen and TG degradation which are repressed by insulin. FFA can down regulate glucose and fatty acid uptake in muscle (Roden *et al.* 1996; Roden *et al.* 2000). FFA can also up-regulate PPARA and its downstream lipases (LIPC and LPL) in liver to generate more FFA rapidly (Jump *et al.* 2005). Thus, FFA has opposite effects to insulin and integrate the regulation of both lipid and glucose metabolisms.

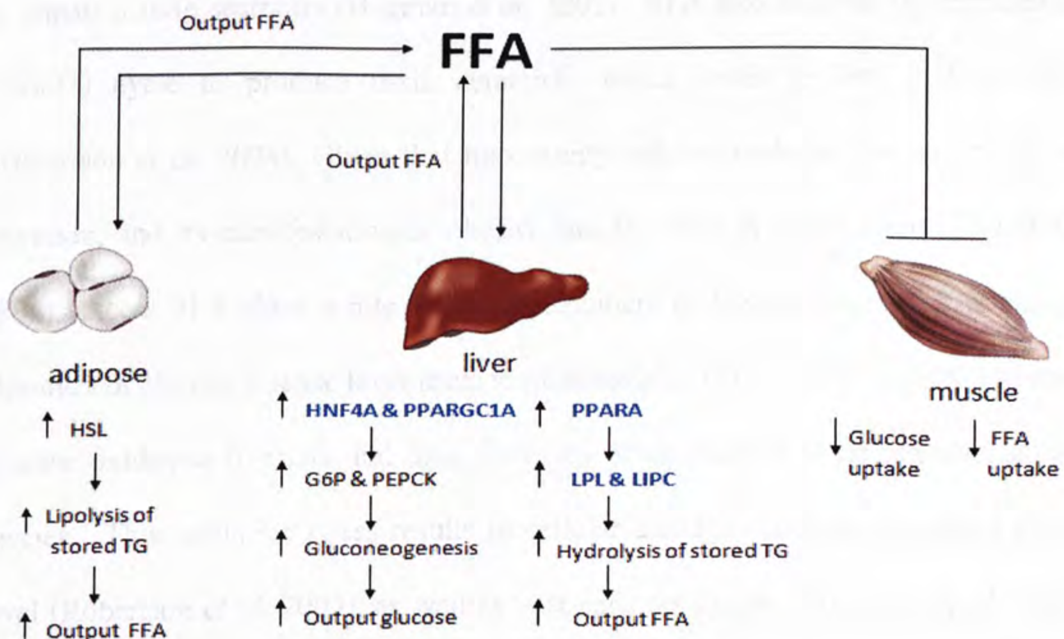


Figure 1.6: FFA action on target tissues. Blue colored genes were genotyped.

1.3.5. Lipotoxicity and glucotoxicity

The presence of impaired insulin secretion and action leads to excessive accumulation of glucose and FFA in the circulation. While acute increase of glucose and FFA enhances GSIS, chronic increase inhibits GSIS (Leahy *et al.* 1988; Leahy *et al.* 1988; Boden *et al.* 1995; Boden 1997). Given that β cell dysfunction triggers the conversion from NGT to IGT to T2D (Weyer *et al.* 1999), lipotoxicity and glucotoxicity are implicated in the pathogenesis of T2D. The chronic elevation of plasma FFA level leads to a phenomenon known as lipotoxicity (Sivitz 2001). In insulin target tissues such as the muscle and adipose tissues, FFA lowers GLUT 4 expression and therefore glucose uptake (Leahy *et al.* 1988; Leahy *et al.* 1988; Boden *et al.* 1995; Boden 1997). In the liver, FFA promotes glucose output during feeding state (Leahy *et al.* 1988; Leahy *et al.* 1988; Boden *et al.* 1995; Boden 1997). In islets, FFA down-regulates insulin transcription-activating TFs such as FOXO1 and PDX1 to impair insulin secretion (Hagman *et al.* 2005). FFA also disrupts the intracellular FFA/TG cycle to produce toxic ceramide, which leads to beta cell apoptosis (Robertson *et al.* 2004). Given that lipotoxicity reduces both insulin sensitivity and secretion, and its manifestation as obesity and IR often precedes overt T2D (Kahn 1994), excess FFA plays a role in the development of hyperglycemia. The chronic elevation of plasma glucose level leads to glucotoxicity (Sivitz 2001), which increases glucose oxidation in islets and thus increases production of toxic reactive oxygen species. This oxidative stress results in cellular damages such as decreased PDX1 level (Robertson *et al.* 2003), as well as islet cells apoptosis (Stumvoll *et al.* 2005). Furthermore, excess glucose impairs insulin secretion by up-regulating insulin transcription-repressing TFs such as CEBPB (El-Assaad *et al.* 2003). The adverse effects of glucotoxicity highlight the importance of

maintaining plasma glucose level within a narrow range. Taken together, lipotoxicity and glucotoxicity can aggravate each other through their effects on insulin secretion and action, forming a vicious cycle that results in β cell failure and eventually development of T2D (Fig 1.7).

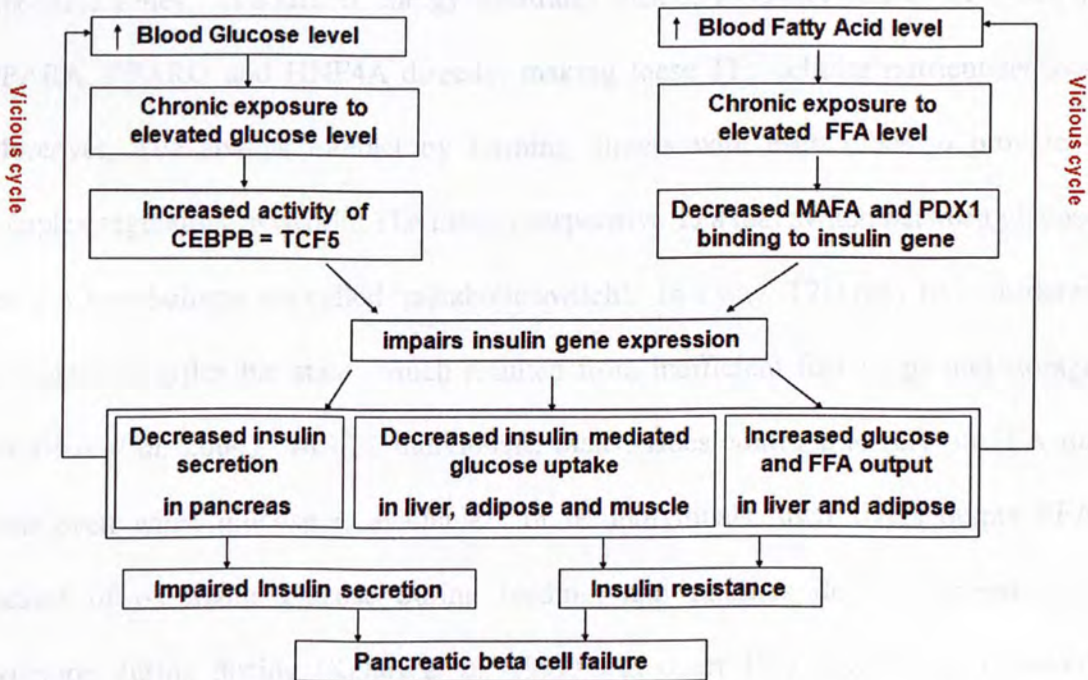


Figure 1.7: Effects of glucotoxicity and lipotoxicity on insulin secretion and action pathways.

1.3.6. Role of transcription factors as metabolic switch

Faced with the dilemma of irregular food supply and continuous energy demand, humans have evolved to use insulin to direct efficient usage of energy. During feeding state, insulin directs most tissues to use glucose as the primary fuel. This preference for glucose ensures that all ingested glucose is either burned or stored.

During fasting state, the absence of insulin signals most tissues to use FFA as the primary fuel and leave glucose for usage by the brain. Thus, insulin plays an important role in inducing or inhibiting a cascade of TFs, which in turn activate subsets of different genes, to modulate the switch between glucose and FA metabolisms for maintaining energy homeostasis. For example, insulin stimulates PPARA, PPARG and PPARGC1A, which up-regulate a set of glycolytic and lipogenic genes. In addition, energy substrates such as FFA can bind to TFs such as PPARA, PPARG and HNF4A directly, making these TFs cellular nutrient-sensors. Moreover, TFs always interact by forming dimers with each other to provide a complex regulatory network. The insulin-responsive TFs that switch between glucose and FA metabolisms are called 'metabolic switch'. In a way, T2D may be considered a 'metabolic inflexible state' which resulted from inefficient fuel usage and storage (Storlien *et al.* 2004). In IGT individuals, their tissues continue to rely on FFA for fuels even when glucose is available. In IR individuals, their livers output FFA instead of oxidizing glucose during feeding, and muscles do not increase FA oxidation during fasting (Kelley *et al.* 1999). In overt T2D individuals, impaired insulin secretion and action result in defective fuel uptake and usage. Given the critical role of TFs in maintaining energy homeostasis, and the link between T2D and metabolic inflexibility, this study hypothesized that genetic variants in the insulin-responsive transcription factors and their downstream targets may be implicated in the pathogenesis of T2D.

1.4. Candidate genes implicated in type 2 diabetes susceptibility

As discussed in section 1.3, several TFs and their downstream targets involved in the insulin secretion and action pathways are implicated in T2D development. In this study, 14 well-studied candidate genes, including eight TFs and six of their target genes, were examined for association with T2D in a Hong Kong Chinese population. For the insulin secretion pathway, genes implicated in MODY and neonatal diabetes were selected. MODY is characterized by autosomal dominant inheritance and early onset of diabetes before the age of 25. Neonatal diabetes refers to the development of diabetes within the first six months of life. Both types of diabetes are caused by single gene mutations leading to primarily insulin deficient phenotype. For the insulin action pathway, genes implicated in the glucose and lipid metabolisms in the liver, muscle and in particular, the adipose tissue, were selected. Recent studies showed that adipose tissue plays a key role in regulating energy homeostasis via controlling fat mass and releasing cytokines (Rosen *et al.* 2006). In addition, many of these gene products are drug targets including GCK, KCNJ11/ABCC8 and PPARs. Thus it is plausible that common variants in the respective genes may be implicated in the common form of T2D. Details of the candidate genes and their association with T2D in previous studies are described below in sections 1.4.1 and 1.4.2.

1.4.1. Candidate genes involved in insulin secretion pathway

1.4.1.1. *HNF4A*

Hepatic nuclear factor 4α gene (*HNF4A*) encodes a TF that belongs to the steroid/thyroid hormone receptor super family (Olefsky 2001). It is expressed in the liver, kidney, intestine, pancreatic islets and insulinoma cells (Odom *et al.* 2004).

Molecular studies showed that HNF4A exerts its effects through inducing other TFs (for example, HNF1A and FOXO1A), which in turn induce the transcription of downstream target genes (Thomas *et al.* 2001). Thus it is considered a master regulator in the pancreatic and hepatic transcription programs. During feeding state, it activates pancreatic gene expressions to promote GSIS. During fasting state, it enhances hepatic glucose and FFA output. *HNF4A* was the first gene that was discovered to cause MODY (MODY1). Heterozygous mutation carriers demonstrated defective GSIS as well as lower TG, apoAII, apoCIII and lipoproteins levels (Malecki 2005). Two SNPs (rs1884614 and rs2144908) located in the 5' promoter region have been associated with T2D in Finnish and Ashkenazi populations (Love-Gregory *et al.* 2004; Silander *et al.* 2004). Other SNPs have also been associated with higher lipid levels and metabolic syndrome (MetS) risk in Finnish families with familial combined hyperlipidemia (Weissglas-Volkov *et al.* 2006).

1.4.1.2. *HNF1A*

Hepatic nuclear factor 1α (*HNF1A*) is also known as transcription factor 1 (*TCF1*). It is expressed in the liver, kidney, intestine and pancreatic islets. It encodes a TF that binds with HNF4A to regulate pancreatic and hepatic transcription programs. A chromatin immunoprecipitation (ChIP) study revealed that it binds to at least 222 target genes in the liver and 106 target genes in the islets, with 30% of these target genes being common to both tissues (Odom *et al.* 2004). Its functions are similar to those of HNF4A. In addition, carriers of this gene mutation (MODY3) have defective GSIS as well as lower apoM level (Malecki 2005). A L27I variant in *HNF1A* has been associated with insulin secretion and insulin sensitivity in Caucasians (Chiu *et al.* 2000; Chiu *et al.* 2003).

1.4.1.3. *PDX1/PBX1*

Pancreatic and duodenal homeobox 1 (*PDX1*), also known as insulin promoter factor 1 (*IPF1*), is expressed in duodenum and pancreatic islets. Pre B-cell leukemia transcription factor (*PBX1*) is widely expressed in all tissues, except in cells of the B and T lineage. The two TFs form dimers to induce insulin gene transcription and regulate normal pancreatic function (Kim *et al.* 2002). Heterozygous mutations in *PDX1* leads to MODY (MODY4) while homozygous mutations leads to neonatal diabetes and pancreatic agenesis (Stoffers *et al.* 1997). Common variants in *PDX1* and *PBX1* have also been associated with T2D in African American and Caucasian populations (Karim *et al.* 2005; Wang *et al.* 2005).

1.4.1.4. *NEUROD1*

Neurogenic differentiation factor 1 (*NEUROD1*) is also known as beta cell E-Box transactivator 2 (*BETA2*). It is expressed in the brain, intestine, kidney and pancreatic islets. This TF binds to the E-box sequence at the insulin promoter region and regulates its expression (Naya *et al.* 1997). *In vivo* study of mice with homozygous gene disruption demonstrated insufficient insulin expression, leading to abnormal pancreatic islet morphogenesis and overt T2D (Naya *et al.* 1997). As the MODY6 gene, mutation carriers showed low serum insulin and C peptide levels (Malecki *et al.* 1999). An A45T variant has been associated with early onset T2D in Chinese (Liu *et al.* 2006), as well as type 1 diabetes in Japanese and Caucasian populations (Iwata *et al.* 1999; Malecki *et al.* 2003).

1.4.1.5. *GCK*

Glucokinase (*GCK*), also called hexokinase 4 (*HK4*), is expressed primarily in the liver and pancreatic islets. This enzyme catalyzes a rate limiting step of phosphorylation of glucose into glucose-6-phosphate, thus acts as a glucose sensor in β cells to commit extracellular glucose into the glycolytic pathway and subsequent GSIS (Matschinsky *et al.* 1993). As the MODY2 gene, carriers with *GCK* mutations demonstrated chronic hyperglycemia due to reduced glucose sensitivity, and those with more severe mutations are inflicted with permanent neonatal diabetes (PNDM) (Velho *et al.* 1992; Njolstad *et al.* 2001). A common variant A-30G within the β cell promoter region has been associated with reduced β cell function in a Japanese American population (Stone *et al.* 1996).

1.4.1.6. *KCNJ11/ABCC8*

The ATP-sensitive potassium (K_{ATP}) channel in the pancreatic β cells is formed by two subunits. The potassium inwardly-rectifying channel J11 gene (*KCNJ11*) encodes Kir6.2 which forms the pore-forming subunit, whereas the ATP-binding cassette C8 gene (*ABCC8*) encodes the sulfonylurea receptor regulatory subunit (SUR1). The K_{ATP} channel is expressed in the brain, heart, skeletal muscle, intestine, kidney and pancreatic islets. During feeding state, K_{ATP} channel plays a key role in insulin exocytosis by its closure incited by increased intracellular ATP/ADP ratios during glucose oxidation (Inagaki *et al.* 1995). The hypoglycemic drug, sulfonylurea, improves insulin secretion through activating the closure of the K_{ATP} channel. *KCNJ11/ABCC8* activating mutations cause PNDM (Gloyn *et al.* 2004) while inactivating mutations lead to hyperinsulinemia (Nestorowicz *et al.* 1997). A

E23K variant (rs5219) in *KCNJ11* was reproducibly associated with T2D in both candidate gene and genome-wide association studies (GWAS), with the rare KK genotype conferring a 14% increased risk for T2D in Europeans (Florez *et al.* 2004; Scott *et al.* 2007; The Wellcome Trust Case Control Consortium 2007; Zeggini *et al.* 2007).

1.4.2. Candidate genes involved in the insulin action pathway

1.4.2.1. *PPARG*

Peroxisome proliferator-activated receptor γ gene (*PPARG*) encodes a TF that belongs to the steroid/thyroid hormone receptor super family (Olefsky 2001). It is expressed in the brain, liver, large intestine, skeletal muscle, macrophages, adipocytes and pancreatic islets (Elbrecht *et al.* 1996; Fajas *et al.* 1997; Mukherjee *et al.* 1997). PPARG is known as a key 'fat regulator' (Picard *et al.* 2004) by binding to the peroxisome proliferator response element (PPRE) of target genes to regulate adipose transcription program (Kota *et al.* 2005). PPARG is sensitive to both insulin and FFA and is the target for a class of insulin sensitizer known as thiazolidinediones (TZDs). During feeding state, it induces lipogenic and glycolytic genes to promote fuel uptake and storage in the adipocytes. During fasting state, it enhances hepatic glucose and FA output. Carriers with loss-of-function mutations are inflicted with familial partial lipodystrophy type 3 (FPLD3) and are characterized by dominantly inherited IR, T2D, hypertension, and partial lipodystrophy (Barroso *et al.* 1999). A P12A variant (rs1801282) was associated with T2D in candidate gene studies and was further confirmed in recent GWAS, with the common CC genotype conferring a 14% increased risk for T2D in Europeans (Altshuler *et al.* 2000; Scott *et al.* 2007; The Wellcome Trust Case Control Consortium 2007; Zeggini *et al.* 2007).

1.4.2.2. *PPARA*

Peroxisome proliferator-activated receptor α gene (*PPARA*) encodes a TF that belongs to the steroid/thyroid hormone receptor super family (Olefsky 2001). It is expressed in the liver, kidney, heart, skeletal muscle and macrophages (Braissant *et al.* 1996; Auboeuf *et al.* 1997; Chinetti *et al.* 1998). PPARA is sensitive to both insulin and FFA (Djouadi *et al.* 1998). It modulates insulin action in the liver by binding to the PPRE of the respective target genes (Kota *et al.* 2005). During feeding state, it induces lipogenic and glycolytic genes to promote fuel uptake and storage in the liver. During fasting state, it enhances hepatic glucose and FFA output (Patsouris *et al.* 2004). It is the target of fenofibrate (a class of fibrates) action which controls TG, total cholesterol and apoB levels (Tai *et al.* 2002; Kota *et al.* 2005). A haplotype has been associated with earlier age of onset in T2D (Flavell *et al.* 2005), while a V227A polymorphism has been associated with serum lipid level (Yamakawa-Kobayashi *et al.* 2002).

1.4.2.3. *PPARGC1A*

Peroxisome proliferator-activated receptor γ coactivator 1α gene (*PPARGC1A*) encodes a transcriptional coactivator that mediates the tissue-specific effects of other TFs. It is expressed in the liver, kidney, heart, skeletal muscle, brain, pancreas and brown adipose tissue (Esterbauer *et al.* 1999; Wu *et al.* 1999; Knutti *et al.* 2000). PPARGC1A does not directly bind to DNA sequences, but facilitates selective binding between TFs such as HNF4A, PPARG and PPARA and their target genes so that only the correct program is induced for each tissue (Handschin *et al.* 2006). During feeding state, PPARGC1A binds to a TF (MEF2C) to promote FA and glucose oxidation in muscle (Michael *et al.* 2001). During fasting state, PPARGC1A binds to

HNF4A to promote glucose output in the liver (Herzig *et al.* 2001; Yoon *et al.* 2001). In subjects with IGT or T2D, lower level of *PPARGC1A* is associated with reduced oxidative phosphorylation in the muscle (Mootha *et al.* 2003; Patti *et al.* 2003; Ling *et al.* 2004). A S482G polymorphism has been associated with T2D in Caucasian and Korean populations (Kunej *et al.* 2004; Kim *et al.* 2005).

1.4.2.4. *ADIPOQ*

Adipocyte, C1q and collagen domain-containing protein (*ADIPOQ*), also called adipose most abundant gene transcript (*APM1*), is exclusively expressed in the adipose tissue (Scherer *et al.* 1995). It encodes the protein adiponectin that acts as a 'starvation signal' to promote fuel oxidation in the peripheral tissues (Powell 2007). Adiponectin also acts as an 'insulin-sensitizer', and it amplifies insulin's action by stimulating the same target kinase (AMPK) (Berg *et al.* 2001; Tomas *et al.* 2002; Yamauchi *et al.* 2002). During feeding state, adiponectin demotes gluconeogenesis in the liver, and cooperates with PPARGC1A to promote FA and glucose oxidation in the muscle (Berg *et al.* 2001; Tomas *et al.* 2002; Yamauchi *et al.* 2002). Studies in Pima Indians suggested that a low adiponectin level was correlated with IR and T2D, as well as an independent predictor for prospective development of diabetes (Lindsay *et al.* 2002; Stefan *et al.* 2002; Williams *et al.* 2004). A T45G polymorphism (rs2241766) has been associated with T2D in both Asian and Caucasian populations (Menzaghi *et al.* 2002; Li *et al.* 2007).

1.4.2.5. *LPL*

Lipoprotein lipase gene (*LPL*) encodes an enzyme that serves as a physiological FFA provider by mediating the rate-limiting step of lipolysis of TG into FFA, as well as facilitating lipoprotein uptake (Pappan *et al.* 2005). It is expressed in the heart, muscle, liver, adipose tissue, and pancreatic islets. During feeding state, it hydrolyzes ingested TG to promote efficient FFA uptake in adipose tissue. During fasting state, it hydrolyzes stored TG to promote FA oxidation in the liver and muscle. Carriers with severe LPL mutations demonstrated hypertriglyceridemia (Wilson *et al.* 1983), while carriers with milder mutations are inflicted with familial combined hyperlipidemia (FCHL) characterized by high TG and low HDL levels (Reymer *et al.* 1995). A T495G variant (rs320) has also been associated with TG and HDL levels in an early-onset T2D Chinese population (Ma *et al.* 2003).

1.4.2.6. *LIPC*

Hepatic lipase gene (*LIPC*) is exclusively expressed in the liver. It shares high sequence homology and possesses similar function as *LPL* (Cai *et al.* 1989). LIPC metabolizes very-low-density lipoprotein (VLDL) into intermediate- (IDL), low- (LDL) and high- (HDL) density lipoproteins, respectively (Zambon *et al.* 1998). Furthermore, it controls plasma HDL concentration by converting HDL_2 into HDL_3 (Santamarina-Fojo *et al.* 2004). It hydrolyzes ingested TG to promote hepatic FFA uptake during feeding state, and hydrolyzes stored TG into FFA during fasting state. Carriers with LIPC mutations demonstrated hypercholesterolemia and hypertriglyceridemia. A C514T variant (rs1800588) has been associated with decreased insulin sensitivity and increased TG level in a non-diabetic Japanese population (Yabu *et al.* 2005).

1.5. Hypothesis and objectives of the study

T2D is a highly prevalent complex disease, but its pathophysiology remains obscure. Faced with the rising epidemic of T2D in China, and the ethnic differences in environmental factors (e.g. lifestyle), risk factor profiles (body composition and insulin secretion/resistance patterns) and genetic backgrounds (linkage disequilibrium pattern and risk allele frequencies), a better understanding of its pathogenesis in Chinese is merited.

To unravel the underlying genetic factors, several methods such as linkage in affected families and association in case control samples using either candidate gene or genome-wide approaches have been adopted. While a genome-wide approach allows identification of novel genes, a pathway-based candidate gene study allows more in-depth examination of each gene region and the analysis of gene-gene and gene-environmental interaction with less compromise on study power.

With the availability of the international HapMap project, systematic survey and selection of a catalog of representative "tag" SNPs for each gene becomes feasible. This study capitalized on this development and conducted a pathway-based two-stage case-control association in a cohort of 1462 T2D cases and 600 controls from the Hong Kong Chinese population. Candidate genes which are TFs and/or common targets of these TFs from both insulin secretion and action pathways were selected, given their implications in the development of T2D.

The three objectives of this study are as followed:

1. To systematically examine common genetic polymorphisms in candidate genes with a known role in insulin secretion or action pathways, and assess their associations with T2D susceptibility using a two-stage approach.

2. To determine whether these genes, which have been previously studied predominantly in Caucasian populations, have a major role in determining type 2 diabetes susceptibility in the Hong Kong Chinese.
3. To assess whether these genes have a larger impact in the more homogenous subgroups of subjects with and without type 2 diabetes.
4. To detect possible gene-gene interactions.

CHAPTER 2. Materials and Methods

2.1. Study design

2.1.1. Two-stage candidate gene association design

T2D is a complex disease that is often described as a "geneticist's nightmare" (Neel 1976). To unravel the genetics of T2D, several methods are currently available. While linkage studies in families successfully identified genes for Mendelian diseases with high penetrance such as MODY, linkage is sub-optimal for detecting disease loci with modest effects for common diseases. Thus far, only *WFS1*, *CAPN10* and *TCF7L2* were found to affect T2D using this method (Strom *et al.* 1998; Horikawa *et al.* 2000; Grant *et al.* 2006). On the other hand, with the availability of the human genome sequence, resequencing is ideal for detecting both rare and common causal variants in common diseases. However, this method is expensive and laborious and thus prohibits examination of large number of subjects and large genomic regions, despite the recent advances in this technology and the lowered cost. Alternatively, the availability of SNPs and linkage disequilibrium (LD) information through the HapMap project makes case-control association studies to be one of the most popular designs due to its relative low cost and ease in genotyping, making large scale studies possible. Under the "common disease – common variant" hypothesis, common diseases are likely caused by multiple common variants (allele frequency $\geq$ 5%) with modest effects, in addition to their interactions with other genes and environmental factors. This hypothesis is validated by recent genome-wide association studies (GWAS) which found at least 14 susceptibility genes for T2D. These include *SLC30A8, HHEX, CDKAL1, CDKN2A/B, IGF2BP2, FTO, PPARG, KCNJ11, JAZF1, CDC123-CAMK1D, TSPAN8-LGR5, THADA, ADAMTS9* and *NOTCH2* (Saxena *et al.*

2007; Scott *et al.* 2007; Sladek *et al.* 2007; Steinthorsdottir *et al.* 2007; Zeggini *et al.* 2007). Interestingly, many of these genes encode TFs that play important roles in the pancreatic β-cell insulin secretion. Due to the large number of SNPs being tested in GWAS, thousands of samples are required to justify multiple comparisons while maintaining study power, which is beyond the scope of this project.

Instead, we chose a case-control association design for selected candidate genes that encode TFs and their downstream targets from two well studied pathways on insulin secretion and action. Several steps were taken to improve the chance of detecting association. Firstly, SNPs with previously reported T2D-associations in other studies (literature significant SNPs) and tag SNPs were selected to validate known candidates as well as search for novel SNPs that are relevant to our Chinese population. Secondly, a two-stage design is used. In the stage 1 study, patients with early onset of diabetes and positive family history were selected due to their likely greater genetic loading for T2D. In the stage 2 study, a larger number of patients consisting of both early- and late-onset T2D were selected to improve power. We then used replication and overall significance in the combined samples to declare significance instead of stringent correction of multiple comparisons due to limited sample size. Thirdly, the availability of detailed clinical phenotypes allows subgroup and metabolic trait analyses which help to dissect the respective pathway/phenotype related to the risk alleles, which otherwise may be masked by the genetic and phenotypic heterogeneous nature of T2D. In the combined samples, patients were subdivided according to age at diagnosis to examine for possible sharing of risk alleles across groups. Similarly, MetS often coexists in patients with insulin resistance and different genes may be responsible for different form of diabetes. A slightly higher female predominance of diabetes as well as sexual differences in phenotypes

(e.g. lipids, body composition) and other risk factors (e.g. smoking) also suggest possible gene-gender interaction for T2D. For these reasons, subgroup analyses were performed in three conditions including 1) family history; 2) MetS; and 3) gender. Lastly, given that the genetic effect of a single variant is likely to be modest, additional interaction or joint effect with other genetic variants and environmental factors will increase the overall effect on T2D susceptibility.

2.1.2. Power calculation

The power to detect an association depends on sample size, effect size, risk allele frequency and type I error rate (Fig. 2.1). Using a nominal type I error rate of 5%, the 800 case-control samples in stage 1 have 80% power to detect odds ratios (ORs) between 1.3 and 1.8 at risk allele frequency ≥ 0.05. The respective ORs range between 1.25 and 1.6 for the 1300 samples in stage 2, and range between 1.2 and 1.5 in the 2100 combined samples. The latter OR range is similar to those found in recent GWAS for T2D (OR = 1.1 to 1.5). The power calculation was performed using PS calculator (Dupont *et al.* 1990).

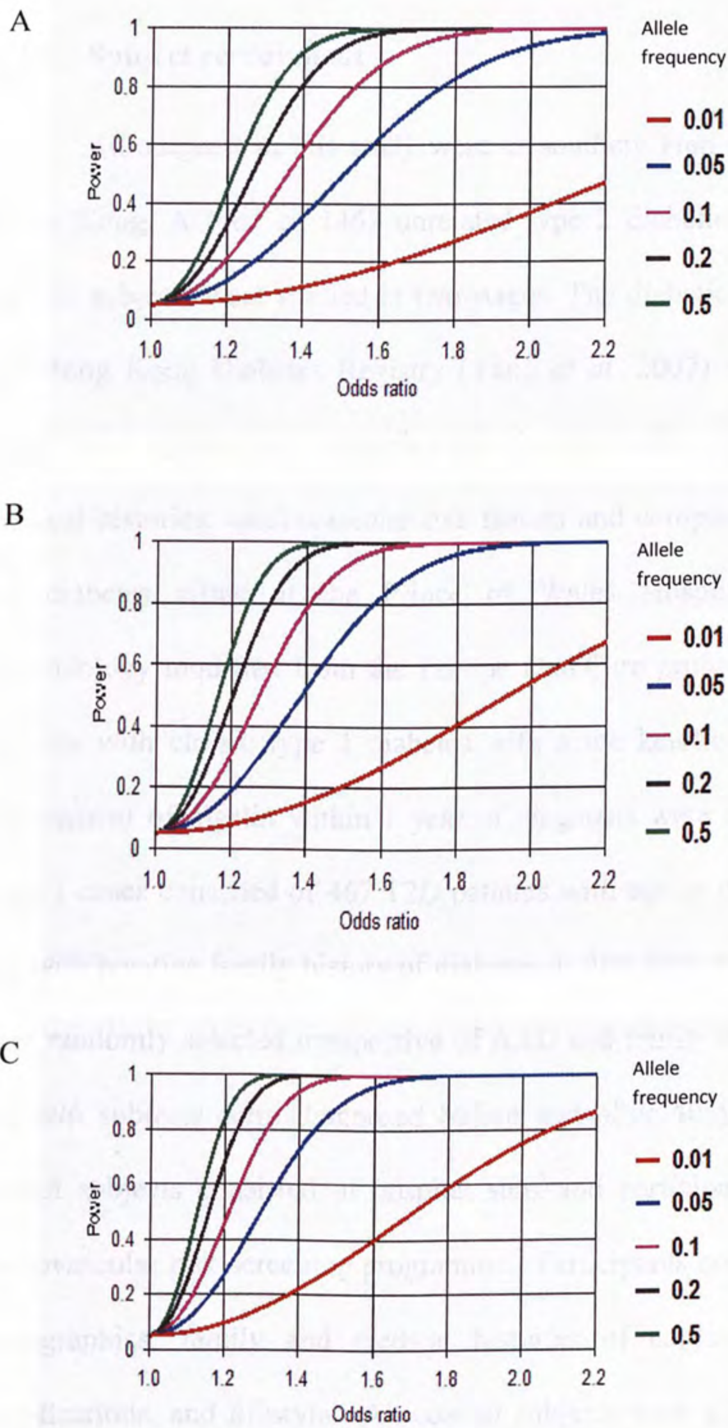


Figure 2.1: Power calculation under a allelic model at 5% type 1 error rate for total sample sizes used in a) stage 1; b) stage 2; and c) combined stages 1 and 2

2.2. Study cohort

2.2.1. Subject recruitment

All subjects in this study were of southern Han Chinese ancestry residing in Hong Kong. A total of 1461 unrelated type 2 diabetic patients and 600 unrelated control subjects were studied in two stages. The diabetic patients were selected from the Hong Kong Diabetes Registry (Yang *et al.* 2007) recruited between 1995 and 2002. This registry has comprehensive documentation of demographics, family and medical histories, cardiovascular risk factors and complications of patients attending the diabetes clinic at the Prince of Wales Hospital (PWH) using standard methodology modified from the Europe DiabCare protocol (Piwernetz *et al.* 1993). Patients with classic type 1 diabetes with acute ketotic presentation or continuous requirement of insulin within 1 year of diagnosis were excluded in this study. The stage 1 cases consisted of 467 T2D patients with age of diagnosis (AAD) $\leq$ 40 years and with positive family history of diabetes in first degree relatives. The stage 2 cases were randomly selected irrespective of AAD and family history. Amongst these, 298 and 696 subjects were diagnosed before and after 40 years old respectively. The control subjects consisted of hospital staff and participants of a community-based cardiovascular risk screening programme. Participants completed a questionnaire on demographics, family and medical histories of cardiovascular risk factors and complications, and lifestyle. All control subjects were glucose tolerant with fasting plasma glucose (FPG) <6.1 mmol/l. A subset of 435 subjects also underwent 75g oral glucose tolerance test (OGTT) and their 2-hour glucose level were < 7.8 mmol/l. Stage 1 consisted of 290 controls and stage 2 consisted of 310 controls. Informed consent was obtained for each participating subject. This study was approved by the Clinical Research Ethics Committee of the Chinese University of Hong Kong.

2.2.2. Clinical and biochemical measurements

Patients attending the PWH diabetes clinic or control subjects attending the Diabetes and Endocrine Centre underwent comprehensive assessment after an 8-hour overnight fast. Body weight, height, waist (the minimum circumference between the umbilicus and xiphoid process) and hip (the maximum circumference around the buttocks and symphysis pubis) circumferences were measured in subjects wearing light clothing and no shoes. Body mass index (BMI) was calculated as weight (kg) divided by height (m) squared. Waist-to-hip ratio (WHR) was calculated as waist circumference (cm) divided by hip circumference (cm). Systolic blood pressure (SBP) and diastolic blood pressure (DBP) were measured in triplicates separated by one minute interval after at least 5 mins in the sitting position, using a Dinamap automated sphygmomanometer (Critikon Inc, CA, USA). Fasting blood samples were taken for routine laboratory assays including measurement of FPG, lipids (total cholesterol, high-density lipoprotein cholesterol (HDL) and triglycerides (Wittrup *et al.*)), liver and renal functions. Low-density lipoprotein cholesterol (LDL) was calculated by the Friedewald's equation for TG < 4.5 mmol/l (Friedewald, Levy et al. 1972), where LDL (mmol/l) = TC – HDL – (TG/2.2). 10 ml of whole blood was also collected for DNA assay.

A 75 g OGTT was also performed in 435 control subjects. Three fasting blood samples collected at 5 minute intervals were assessed for mean FPG and insulin (FINS) levels. Blood samples were also collected at 15, 30, 60 and 120 mins during the OGTT for measurement of plasma glucose and insulin. Indices for insulin sensitivity (HOMA-IR, ISI) and insulin secretion capacity (HOMA-β, IDI) were calculated as follows (Matthews *et al.* 1985; Matsuda *et al.* 1999):

1) Homeostasis model assessment for insulin resistance:

$$\text{HOMA-IR} = \text{FINS (mU/l)} \times \text{FPG (mmol/l)}/22.5$$

2) Homeostasis model assessment for β cell function:

$$\text{HOMA-}\beta = \text{FINS} \times 20/(\text{FPG}-3.5)$$

3) Insulin sensitivity index (ISI) from OGTT:

$$\text{ISI} = 10000 \,/\, (\text{FPG (mmol/l)} \times \text{FINS (mU/l)} \times \text{MGLUOGTT (mmol/l)} \times \text{MINSOGTT (mU/l)})$$, where

MGLUOGTT = average of OGTT glucose at 0, 30, 60, 90 and 120 mins;

MINSOGTT = average of OGTT insulin at 0, 30, 60, 90 and 120 mins

4) Insulin disposition index from OGTT:

$$\text{IDI} = (\text{INS30} - \text{FINS})\text{ (mU/l)} / (\text{GLU30} - \text{FPG})\text{ (mmol/l)} \times \text{ISI}/100$$, where

INS30 and GLU30 represent insulin and glucose at OGTT 30 mins

Plasma glucose (hexokinase method), TC (enzymatic method), TG (enzymatic method without glycerol blanking) and HDL (direct method using PEG-modified enzymes and dextran sulfate) were measured on a Roche Modular Analytics system (Roche Diagnostics GmbH, Mannheim, Germany) using standard reagent kits supplied by the manufacturer of the analyzer. The precision was within the manufacturer's specifications. Insulin were measured with an enzyme linked immunosorbent assay (DAKOCytomation, Cambridgeshire, UK).

2.2.3. Clinical definitions

Diabetes was defined according to the World Health Organization (WHO) 1998 criteria (Alberti *et al.* 1998). Subjects with FPG $\geq$ 7 mmol/l and/or 2-hour glucose $\geq$ 11.1 mmol/l during OGTT were considered diabetic. Metabolic syndrome (MetS) was defined according to the National Cholesterol Education Program Adult Treatment Panel III (NCEP III) guidelines (NCEPIII 2001). Subjects who have at

least three of the following five risk factors were classified as having metabolic syndrome:

1) Hyperglycemia with known diabetes or FPG ≥ 6.1 mmol/l;

2) Hypertension defined as SBP ≥130 mmHg and/or DBP ≥ 85 mmHg or taking antihypertensive medication;

3) Hypertriglyceridemia defined as TG ≥ 1.7 mmol/l;

4) Low HDL-cholesterol defined as HDL <1.0 mmol/l in men or <1.3 mmol/l in women;

5) Central obesity defined as waist circumference >90 cm in men or >80 cm in women. The definition of central obesity was modified for Asian populations (World Health Organization 2000)

2.3. Genetic study

2.3.1. Candidate gene selection

According to the hypothesis, eight genes from the insulin secretion pathway (*HNF4A, HNF1A, PDX1/PBX1, NEUROD1, GCK* and *KCNJ11/ABCC8*) and six genes from the insulin action pathway (*PPARG, PPARA, PPARGC1A, ADIPOQ, LPL* and *LIPC*) were investigated.

2.3.2. SNP selection

Single nucleotide polymorphism (SNP), defined as a nucleotide change at a single base, is the most common form of genetic variation in the human genome. With the genotyping of more than 1 million SNPs in populations of Europeans, Asians (Chinese and Japanese) and Africans, the international haplotype map (HapMap) project demonstrated that most SNPs are not independently associated

with each other. The non-random association between SNPs is described as linkage disequilibrium (LD), and can be quantified as D' or r^2. While D' reflects the recombination events that shaped the haplotyep block structure during evolution (Balding 2006), r^2 measures the frequency of the alleles of two SNPs cosegregate together on the same haplotype. Within a haplotype block, SNPs in high r^2 are redundant in information and thus a subset of "tag" SNPs that can uniquely identify all the haplotypes is sufficient for association study.

For tag SNP selection in this study, firstly the Han Chinese (CHB) genotype data of each gene, including its 2kb upstream and downstream flanking regions, were downloaded from HapMap (http://www.hapmap.org) using the NCBI build 35 assembly. The large intronic regions in four genes (*HNF4A, PBX1, PPARGC1A* and *LIPC*) were excluded to lower genotyping cost as they are less likely to harbor functional variants.

Secondly, the HapMap data were uploaded into the Haploview program (v.3.32) (Barrett *et al.* 2005). Common SNPs with minor allele frequency (MAF) ≥ 5% were used for tag SNP selection. Using a pair-wise tagging approach, only one tag SNP from a group of correlated SNPs ($r^2 \geq 0.8$) was chosen (de Bakker *et al.* 2005). The association between SNPs for each gene was visualized as D' and r^2 in various shades of red and black, respectively (Table 2.1 and Appendix 1).

Table 2.1: A) D' color scheme; B) r^2 color scheme in Haploview

A)

	D' < 1	D' = 1
LOD < 2	white	blue
LOD ≥ 2	shades of pink/red	bright red

B)

$r^2 = 0$	$0 < r^2 < 1$	$r^2 = 1$
white	shades of grey	black

Thirdly, during tag SNP selection, coding SNPs and previously-associated SNPs were preferably selected or force included. Coding SNPs are more likely to exert a functional effect. In addition, SNPs that showed evidence for T2D associations in previous studies warrant further investigation in our Hong Kong Chinese cohort.

Fourthly, SNPs that may fail in assay design or genotyping were preferably excluded as tag SNPs. An in-house bioinformatic program was developed to analyze the 50bp DNA amplicon surrounding each test SNP. SNPs that showed one or more of the following conditions may be excluded: 1) amplicon with 4 or more repeated nucleotides that may lead to non-specific primer binding; b) < 40% or > 60% GC content which may cause sub-optimal primer binding and dissociation; and c) presence of a secondary SNP within 21 bp of the target SNP that may alter primer binding efficiency to amplicons with different alleles.

Details of the gene and SNP selection are shown in Table 2.2.

2.3.3. DNA sample preparation

10 ml of stored peripheral venous blood was de-thawed and then transferred to a 50 ml falcon tube. 15 ml of 1x lysis buffer (10mM $KHCO_3$, 155mM NH_4Cl, 0.1 mM EDTA) was added. The mixture was iced for 15 mins, then centrifuged at 3000 rpm at 4°C for 15 mins. After decanting the supernatant, the cell pellet was re-suspended in 15 ml of lysis buffer, iced, centrifuged, and decanted as before. 3 ml of TE buffer, 600 μl of 10% SDS and 50 μl of proteinase K solution were then added to the cell pellet. After vigorous shaking, the sample was incubated at 65°C overnight.

The digested sample was cooled for 15 mins before 1 ml of 6 M NaCl was added and mixed. After centrifugation at 3000 rpm at 4°C for 15 mins, the supernatant was decanted to a 15 ml falcon tube. 1 ml of TE equilibrated phenol and 1 ml of chloroform was gently mixed in. Following centrifugation, the upper aqueous layer was aspirated to a new tube, and 2 ml chloroform was gently mixed in. The centrifugation and aspiration were repeated. Two volumes of ice-cold absolute ethanol were added and gently inverted to precipitate the genomic DNA. The mixture was iced for 15 mins, centrifuged, and the supernatant was discarded. The precipitated DNA pellet was transferred to a 1.5 ml eppendorf tube and washed with 1 ml 70% ethanol. Centrifugation was repeated, the alcohol was drained off, and the DNA pellet was left to air dry for 30 mins. Depending on the size of the pellet, 200-500 μl of TE buffer was added to dissolve the DNA. The quality and quantity of DNA was checked by optical density (OD) measurement in a spectrometer. One unit of OD260 is equivalent to 50μg/ml of DNA. A ratio of 1.8-2.0 for OD260/280 absorbance was considered as good quality.

2.3.4. Genotyping methods

In this project, two genotyping platforms were employed. Of the 198 selected SNPs, 135 SNPs in 11 genes were genotyped using the allele specific melting temperature (T_m) shift assay, while 63 SNPs in 3 genes were genotyped using the Sequenom i-PLEX gold assay. As the study of the 11 genes was in collaboration with Roche Pharmaceuticals, the 135 SNPs were genotyped by a technician and me at Roche Pharmaceuticals in New Jersey, USA. Of the remaining 63 SNPs, 28 SNPs in *PBX1* were genotyped by the Genome Research Center in Hong Kong University, and 35 SNPs in *PPARA* and *PPARGC1A* were genotyped by the McGill University and Genome Québec Innovation Centre in Quebec, Canada.

2.3.4.1. Allele specific T_m shift assay

This singleplex technology makes use of allele-specific polymerase chain reaction (PCR), and relies on the differential melting temperatures between amplified products to discriminate between genotypes (Germer *et al.* 1999). Each allele-specific primer will only bind to the template containing the respective allele for the DNA polymerase to amplify. By designing two allele-specific primers with unequal length and GC contents, their respective amplified products will have different T_m. The longer the primer length, the more unbalance the GC content, the higher T_m, which leads to a higher temperature melting curve.

Figure 2.2 demonstrated the steps of this assay. Three primers (two allele-specific forward primers and a common reverse primer) were designed for each SNP. An in-house algorithm developed by Roche Pharmaceuticals called T_m Calculator was used to calculate the T_m, GC content, and primer length for each SNP (Figure

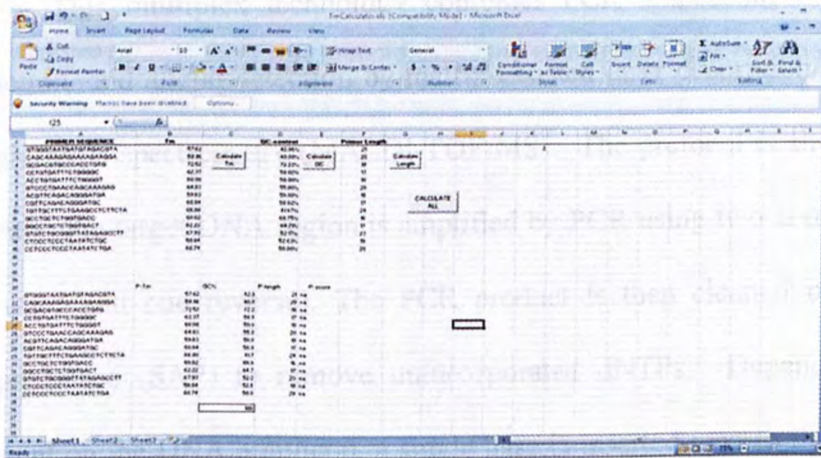
2.2, Panel A). The allele-specific primer that ended with a lower T_m base (A or T) would receive a short GC-rich tail of 6 bp, while the other allele-specific primer that ended with a higher T_m base (G or C) would receive a long GC-rich tail of 14 bp (Wang *et al.* 2005) (Figure 2.2, Panel B). The GC-rich tails maximized the T_m difference between the alleles to generate two non-overlapping melting peaks. The common reverse primer was placed less than 20 bp downstream from the SNP to achieve a good balance between amplification efficiency and allele specificity.

For the PCR amplification, a master mix consisted of 10mM Tris-HCl, 40 mM KCl, 2mM $MgCl_2$, 50 μM each of dATP, dCTP, dGTP, and dTTP, 0.2x SYBR Green, 0.01% Tween 20, 4% DMSO, 2% glycerol, 1.8 U of Stoffel Gold polymerase, and sterile water was prepared. Then, 9 ng of genomic DNA and 0.2 μM each of the three SNP-specific primers were added to make up a total reaction volume of 15 μl. The PCR was then carried out on an ABI 9700 dual block 384-well thermal cyclers (Applied Biosystems, Foster City, CA) at the following condition: 95°C for 12 mins to heat-activate the Stoffel Gold polymerase, followed by 40 cycles of denaturation at 95°C for 20s, annealing at 58°C for 40s and extension at 72°C for 20s.

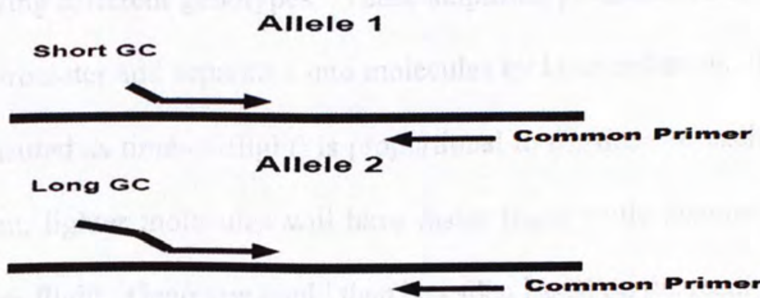
Following PCR amplification, the samples were transferred to an ABI PRISM 7900 HT real-time thermal cycler (Applied Biosystems, Foster City, CA) for melting curve analysis. The resulting data were imported to an in-house bioinformatic program called GCSNPClust for genotype calling. The rate of product dissociation (plotted on the y-axis) was compared to the melting temperature (plotted on the x-axis) to give the T_m profile. Samples homozygous for allele bound by shorter primer showed lower T_m curve whereas samples homozygous for allele bound by longer primer showed higher T_m curve. Heterozygous samples showed both T_m curves but with lower peak height (Figure 2.2, Panel C). Based on the melting curve patterns,

the software identified the amplified allele and clustered the samples by their genotypes. A tight cluster indicated a reliable genotype calling.

A)



B)



C)

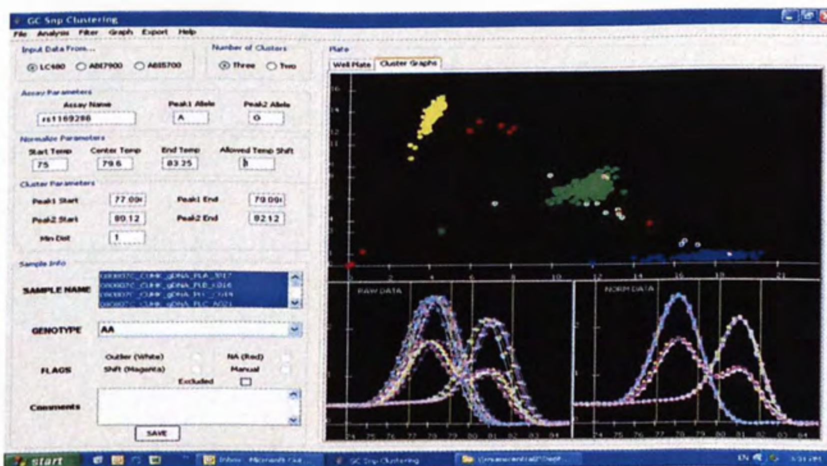


Figure 2.2: Protocol for allele specific T_m shift assay. A) T_m calculator; B) allele specific primer design; C) GCSNPClust for genotype calling

2.3.4.2. Mass spectrometry assay

This multiplex technology combines PCR amplification, primer extension reaction, and allelic-detection by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS). The protocol is shown in Figure 2.3. Firstly, the target DNA region is amplified by PCR using two common primers (one forward and one reverse). The PCR product is then cleaned by shrimp alkaline phosphatase (SAP) to remove unincorporated dNTPs. Depending on the allele present on the DNA amplicon, a single base is inserted next to the extension primer and amplified. This leads to a difference in mass between the amplified products carrying different genotypes. These amplified products are then placed in the mass spectrometer and separated into molecules by laser radiation. The rate of separation (measured as time-of-flight) is proportional to the mass of each nucleotide. For this reason, lighter molecules will have faster flight while heavier molecules will have slower flight. Genotype could then be called based on the resulting mass spectrum.

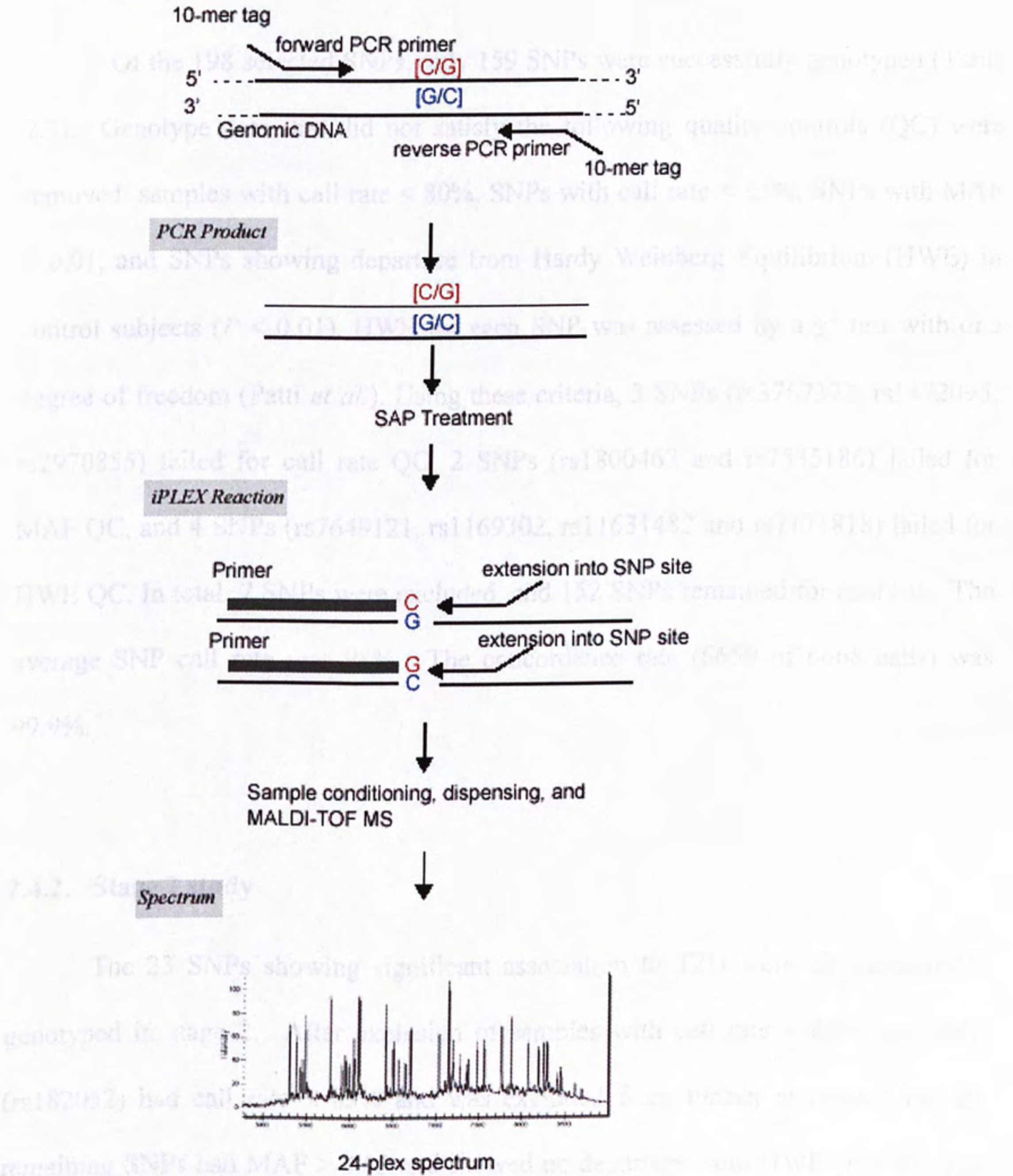


Figure 2.3: Sequenome i PLEX genotyping protocol

2.4. Data quality control

2.4.1. Stage 1 study

Of the 198 selected SNPs, only 159 SNPs were successfully genotyped (Table 2.2). Genotype data that did not satisfy the following quality controls (QC) were removed: samples with call rate < 80%, SNPs with call rate < 85%, SNPs with MAF ≤ 0.01, and SNPs showing departure from Hardy Weinberg Equilibrium (HWE) in control subjects ($P < 0.01$). HWE for each SNP was assessed by a χ^2 test with one degree of freedom (Patti *et al.*). Using these criteria, 3 SNPs (rs3767372, rs1472095, rs2970855) failed for call rate QC, 2 SNPs (rs1800467 and rs7535186) failed for MAF QC, and 4 SNPs (rs7649121, rs1169302, rs11631482 and rs7171818) failed for HWE QC. In total, 7 SNPs were excluded, and 152 SNPs remained for analysis. The average SNP call rate was 98%. The concordance rate (6659 of 6668 calls) was 99.9%.

2.4.2. Stage 2 study

The 23 SNPs showing significant association to T2D were all successfully genotyped in stage 2. After exclusion of samples with call rate < 80%, one SNP (rs182052) had call rate < 85% and was excluded from further analysis. The 22 remaining SNPs had MAF > 0.05 and showed no departure from HWE ($P > 0.01$) in both cases and controls. The average SNP call rate was 97%. The concordance rate (557 of 558 calls) was 99.8%.

Table 2.2: Summary of SNP selection and genotyping

Pathway	Gene	Selected region (B35)	Hapmap SNPs[a]	Selected SNPs	Failed-assay SNPs	Genotyped SNPs	Failed-QC SNPs	Analyzed SNPs	Hapmap SNPs[b]	db SNPs[c]	% of Hapmap SNPs captured[d]
Insulin secretion	HNF4A	chr20:42,403,212-42,423,128	55	25	6	19	0	**19**	18	1	80%
		chr20:42,457,463-42490894									86%
Insulin secretion	HNF1A	chr12:119,877,268-119,905,031	29	12	3	9	1	**8**	8		86%
Insulin secretion	IPF1	chr13:27,390,176-27,399,393	2	2	1	1	0	**1**	1		50%
Insulin secretion	PBX1	chr1:161255602-161264206	69	28	8	20	2	**18**	16	2	80%
		chr1:161493633-161554494									53%
Insulin secretion	NEUROD1	chr2:182,364,700-182,372,759	4	2	1	1	0	**1**	1		50%
Insulin secretion	GCK	chr7:43,955,109-44,004,277	42	14	2	12	0	**12**	10	2	67%
Insulin secretion	ABCC8/ KCNJ11	Chr11:17,306,970-17378462	43	16	6	10	1	**9**	5	4	49%
Insulin action	PPARG	chr3:12,302,358-12,452,839	91	16	3	13	0	**13**	12	1	88%
Insulin action	PPARA	Chr22: 44,865,017-44,962,171	26	12	0	12	0	**12**	11	1	77%
Insulin action	PPARGC1A	Chr4: 23,467,913-23513192	54	23	0	23	0	**23**	22	1	100%
		Chr4: 23557168-23,569,968									79%
Insulin action	ADIPOQ	chr3:188,041,164-188,060,951	19	10	2	8	1	**7**	6	1	74%
Insulin action	LPL	chr8:19,839,057-19,871,048	42	12	3	9	0	**9**	8	1	52%
Insulin action	LIPC	chr15:56509466-56513466	68	26	4	22	2	**20**	20	0	62%
		chr15:56619285-56650363									82%
Total SNPs			544	198	39	159	7	**152**	138	14	

a: Only Hapmap SNPs with MAF ≥ 5% in the selected region are included

b: Number of Hapmap SNPs within analyzed SNPs

c: Number of dbSNPs within analyzed SNPs

d: % of region coverage (at max. $r^2 \geq 0.8$) in CHB+JPT. Calculated with Tagger.

2.5. Statistical analysis

All statistical analyses were performed with PLINK (v.0.99, available at http://pngu.mgh.harvard.edu/~purcell/plink) (Purcell *et al.* 2007), Haploview (v.3.32) (Barrett *et al.* 2005) or SPSS 14.0 for Windows (SPSS Inc., Chicago, IL, USA). A nominal P value < 0.05 (2-tailed) was considered to be statistically significant.

For descriptive statistics, continuous data were expressed as mean ± standard deviation (SD) or geometric mean (95% confidence interval [CI]), and were logarithm transformed if necessary. Categorical data were compared using a χ^2 test.

2.5.1. Stage 1 analysis

For disease association analyses, genotype frequencies between T2D cases and controls were compared using logistic regression under an additive model. Odds ratio (OR) with 95% confidence intervals (CIs) were presented with respect to the minor allele in cases. Logistic regression was favored over allelic association for its robustness against departures from HWE. We hypothesize that candidate genes from different pathways may exert different genetic risks in patients with different phenotypes. Thus we divided the T2D cases into two subgroups according to the presence or absence of MetS, and then compared to all controls to examine for possible modifier effect of MetS on genetic association with T2D.

2.5.2. Stage 2 analysis

22 SNPs that showed significant association with T2D and/or MetS subgroups were further examined in an independent case-control sample in stage 2 using the same statistics as in stage 1.

2.5.3. Stage1 and 2 combined analysis

In the combined samples, the increased sample size allowed us to examine the phenotype-genotype association in more details without compromising the power. For the 22 tested SNPs, the best fit genetic models were tested by comparing the additive, dominant and recessive models using logistic regression. Haplotypes with frequencies > 5% were tested for T2D association using haplotype specific test implemented in Haploview. To examine the possible confounding or modifying effects of MetS, age at diagnosis and gender, subgroup analyses were tested by logistic regression under an additive model. Cases with or without MetS were compared to all controls respectively. Similarly, cases with early-onset (age at diagnosis ≤ 40 years) or late onset T2D were compared to all controls respectively. For gender effect, cases and controls were divided by gender and compared to their corresponding counterparts.

We further tested for possible interaction and joint effects amongst 5 independent SNPs that showed consistent associations in T2D, MetS subgroup and/or age at diagnosis subgroup analyses. Each of the possible two SNPs interaction was tested one at a time to a logistic model containing the 5 SNPs coded under the additive model. In addition, by assuming similar effect size, the joint effects of the 5 SNPs, with or without division into insulin secretion or action pathways, were assessed for association with T2D. Each case or control subject was counted for the number of risk alleles carried, and then assigned to a low risk, medium risk, moderately high risk, or high risk allele group using an arbitrary allele number cutoff. Due to the modifying effects of MetS and age at diagnosis, the cases were further divided into 4 subgroups: 1) MetS negative and early onset T2D; 2) MetS negative and late onset T2D; 3) MetS positive and early onset T2D; 4) MetS positive and late onset T2D. The cases were then compared to all controls to examine the interacting effect of risk allele and phenotypes on T2D association.

We also examined for possible associations between SNPs and metabolic traits in the 600 control subjects. Continuous data were logarithm transformed and winsorized if necessary. Distributions of metabolic traits were compared for different genotypes under an additive model by linear regression with adjustment for age and gender.

CHAPTER 3. Results

3.1. Clinical characteristics of subjects in stages 1 and 2 studies

The clinical characteristics of the study subjects in stages 1 and 2 were summarized in Table 3.1. In total, 1461 T2D cases and 600 controls were studied. The cases had worse metabolic profiles compared to controls, as evidenced by their higher prevalence of MetS (56.2% vs. 4.2% in cases vs. controls) and other MetS risk factors (59.4% vs. 27.6% for hypertension, 43.4% vs. 14.9% for low HDL level, 35.2% vs. 14.0% for hypertriglyceridemia and 48.2% vs. 14.8 for central obesity).

In this staged study, 467 cases and 290 controls were studied in stage 1, while 994 cases and 310 controls were studied in stage 2. As expected from the study design, the patients in stage 1 were younger (39.2 ± 8.5 vs. 54.8 ± 12.8 yrs) and had earlier age at diagnosis (31.3 ± 6.0 vs. 49.6 ± 12.3 yrs) than those in stage 2. However, the stage 1 cases also had better lipid profiles, blood pressure and lower MetS prevalence (52.1% vs. 58.1%) than the stage 2 cases. Amongst the controls, the stage 1 subjects had lower anthropometric measures, lipid profiles and blood pressure compared to the stage 2 subjects, possibly due to their younger age (39.4 ± 10.6 vs. 42.9 ± 10 yrs) and lower degree of family history of diabetes (6.6% vs. 31%).

Table 3.1: Clinical characteristics of subjects in stages 1 and 2 studies

	Stage 1		Stage 2		Stage 1+2		Stage 1 vs. 2	Stage 1 vs. 2	Stage 1+2
	Case (N = 467)	Control (N = 290)	Case (N = 994)	Control (N = 310)	Case (N=1461)	Control (N = 600)	Case *P*	Control *P*	Case vs. Control *P*
Male (%)	36.6	36.9	42.2	53.2	40.5	45.3	**0.046**	**< 0.001**	**< 0.043**
Family history of diabetes (%)	100.0	6.6	39.9	31.0	59.1	19.2	**< 0.001**	**< 0.001**	**< 0.001**
Age (year)	39.2 ± 8.5	39.4 ± 10.6	54.8 ± 12.8	42.9 ± 10.0	49.8 ± 13.7	41.2 ± 10.4	**< 0.001**	**< 0.001**	**< 0.001**
Age at diagnosis (year)	31.3 ± 6.0	-	49.6 ± 12.3	-	43.8 ± 13.7	-	**< 0.001**		
BMI (kg/m^2)	25.7 ± 4.6	22.3 ± 3.0	24.8 ± 3.8	23.5 ± 3.5	25.1 ± 4.1	22.9 ± 3.3	**< 0.001**	**< 0.001**	**< 0.001**
Waist circumference (cm)	84.1 ± 11.5	74.5 ± 9.1	84.8 ± 9.8	78.8 ± 9.4	84.6 ± 10.4	76.8 ± 9.5	0.299	**< 0.001**	**< 0.001**
SBP (mmHg)	125 ± 18	112 ± 13	137 ± 22	118 ± 18	133 ± 22	115 ± 16	**< 0.001**	**< 0.001**	**< 0.001**
DBP (mmHg)	75 ± 10	70 ± 10	82 ± 11	75 ± 12	79 ± 11	72 ± 11	**< 0.001**	**< 0.001**	**< 0.001**
FPG (mmol/l)	8.8 ± 3.4	4.8 ± 0.4	9.1 ± 3.6	4.9 ± 0.4	9.0 ± 3.6	4.8 ± 0.4	0.056	**0.006**	**< 0.001**
TC (mmol/l)	5.28 ± 1.2	4.8 ± 0.8	5.5 ± 1.2	5.2 ± 1.0	5.4 ± 1.2	5.0 ± 0.9	**< 0.001**	**< 0.001**	**< 0.001**
TG (mmol/l) [A]	1.40 (1.31-1.49)	0.81 (0.77-0.86)	1.45 (1.38-1.52)	1.05 (0.98-1.12)	1.44 (1.37-1.52)	0.98 (0.91-1.05)	0.348	**< 0.001**	**< 0.001**
HDL (mmol/l)	1.3 ± 0.4	1.6 ± 0.4	1.3 ± 0.4	1.5 ± 0.4	1.3 ± 0.4	1.5 ± 0.4	0.904	0.762	**< 0.001**
LDL (mmol/l)	3.2 ± 0.9	2.9 ± 0.7	3.5 ± 1.0	3.1 ± 0.9	3.4 ± 1	3.0 ± 0.8	**< 0.001**	**< 0.001**	**< 0.001**
MetS (%)	52.1	2.1	58.1	6.1	56.2	4.2	**0.031**	**0.013**	**< 0.001**
Hypertension (%)	41.4	19.4	67.9	33.5	59.4	27.6	**< 0.001**	**0.001**	**< 0.001**

Low HDL (%)	46.4	13.4	41.9	16.3	43.4	14.9	0.106	0.33	**< 0.001**
High TG (%)	34	6.9	35.8	20.8	35.2	14	0.487	**< 0.001**	**< 0.001**
Central obesity (%)	47.6	11.6	48.4	17.7	48.2	14.8	0.780	**0.047**	**< 0.001**
Anti-hypertensive drug (%)	17.1	-	23.0	-	21.1	-	1.000		
Lipid lowering drug (%)	7.7	-	5.0	-	5.9	-	0.100		
Oral hypoglycemic agent (OHA) (%)	52.7	-	51.6	-	52.0	-	0.317		
Insulin treatment (%)	21.4	-	10.2	-	13.8	-	**< 0.001**		

Data are expressed as mean +/- SD or [a] geometric mean (95% CIs)

3.2. Case-control associations in stage 1

3.2.1. Association with T2D

A total of 152 SNPs were successfully genotyped with good quality in stage 1 samples consisting of 467 early onset T2D cases with positive family history and 290 controls. Appendix 2 showed the basic information and HWE statistics of all SNPs. To investigate the association between these 152 SNPs and T2D, their genotype frequencies were compared between cases and controls under the additive model. The complete association results were shown in Figure 1 and Appendix 3. Table 1 summarized the 12 SNPs (7.9%) in 8 genes that demonstrated nominal significant ($P < 0.05$) association with T2D.

Amongst the 12 SNPs, *ADIPOQ, KCNJ11* and *HNF4A* showed multiple associated SNPs ($0.006 < P < 0.025$). In *ADIPOQ*, rs2241766 coding for a synonymous change (G15G) had an odds ratio (OR) of 0.76 ($P = 0.011$) for the protective G allele. Another SNP (rs1063539) in high LD ($r^2 = 0.83$) had a similar OR of 0.74 ($P = 0.006$). In *KCNJ11,* two significant SNPs (rs5219 and rs5215) had r^2 of 0.9 and with similar ORs of 1.28 - 1.29 ($0.024 < P < 0.025$). In addition, three significant SNPs in *HNF4A* were in high LD ($r^2 \geq 0.69$) and were all located in the P2 promoter region. Their effect sizes were similar with ORs ranged from 1.27 to 1.33 ($0.007 < P < 0.021$) for the risk C allele for rs4812828, T allele for rs1884614 and A allele for rs2144908. The high r^2 shared by the significant SNPs in these three genes suggested that only one signal is present for each gene. For the remaining genes showing significant association to T2D, rs881740 of *PPARA* showed the largest effect (OR = 1.59, $P = 0.01$), followed by rs1472095 of *PPARGC1A* (OR = 1.52), rs13239289 of *GCK* (OR = 1.48), rs2242062 of *LIPC* (OR = 1.28) and rs1169286 of *HNF1A* (OR = 1.23).

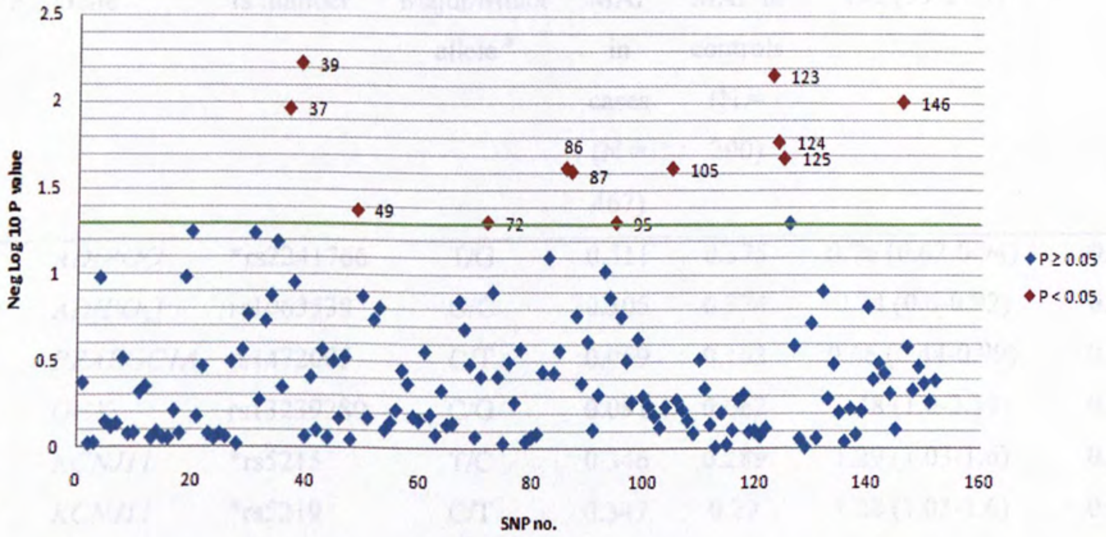


Figure 3.1: Graphical summary of association results for T2D for stage 1 SNPs in case control samples (shown as negative $\log_{10} P$ value). The green line indicates the threshold of $P = 0.05$. The SNP number corresponds to the SNP order shown in Appendix 2.

Table 3.2: Significant T2D association results of SNPs in stage 1 case-control samples

SNP no.	Gene	rs number	Major/Minor allele [a]	MAF in cases (N = 467)	MAF in controls (N = 290)	OR (95% CI)	*P*
37	*ADIPOQ*	*rs2241766	T/G	0.311	0.375	0.76 (0.62-0.94)	**0.011**
39	*ADIPOQ*	rs1063539	G/C	0.305	0.374	0.74 (0.6-0.92)	**0.006**
49	*PPARGC1A*	rs1472095	C/T	0.069	0.103	0.66 (0.44-0.99)	**0.042**
72	*GCK*	rs13239289	C/G	0.091	0.062	1.48 (1.0-2.19)	**0.049**
86	*KCNJ11*	*rs5215	T/C	0.346	0.289	1.29 (1.03-1.6)	**0.024**
87	*KCNJ11*	*rs5219	C/T	0.347	0.29	1.28 (1.03-1.6)	**0.025**
95	*HNF1A*	rs1169286	A/G	0.461	0.409	1.23 (1-1.51)	**0.049**
105	*LIPC*	rs2242062	A/G	0.375	0.317	1.28 (1.03-1.59)	**0.024**
123	*HNF4A*	rs4812828	C/T	0.454	0.526	0.75 (0.61-0.92)	**0.007**
124	*HNF4A*	*rs1884614	C/T	0.463	0.399	1.28 (1.05-1.57)	**0.017**
125	*HNF4A*	*rs2144908	G/A	0.469	0.406	1.27 (1.04-1.56)	**0.021**
146	*PPARA*	rs881740	A/G	0.135	0.089	1.59 (1.12-2.25)	**0.010**

Significant associations at $P < 0.05$ are bolded.

* literature significant SNPs.

[a] minor alleles were defined according to the frequencies in cases shown in Appendix 2.

3.2.2. Association with T2D subset by metabolic syndrome

Metabolic syndrome (MetS) often coexists in T2D patients with insulin resistance, while T2D patients with primarily insulin deficiency are less likely to have multiple metabolic risk factors. To investigate the possible interacting effect of MetS and SNPs on T2D association, subset analyses which compared 223 MetS negative cases and 244 MetS positive cases to 290 controls respectively were performed. MetS was defined according to the NCEP criteria with central obesity definition modified for Asians (Section 2.2.3). The complete association results were shown in Figure 3.2 and Appendix 3.

A total of 10 SNPs (6.6%) in 5 genes and 12 SNPs (7.9%) in 7 genes were significantly associated with MetS negative T2D and MetS positive T2D, respectively. Apart from the 10 (83%) T2D-associated SNPs (Table 3.2) that remained significant in the subset analyses, 10 additional SNPs demonstrated previously undetected associations. It is of interest to note that 4 genes (*ADIPOQ, HNF4A*, *KCNJ11* and *LPL*) had multiple association signals in the same MetS subgroup. Moreover, most SNPs were significant in only one subgroup except for rs291 of *LPL* and rs4812828 of *HNF4A*.

In *ADIPOQ*, five SNPs in two LD blocks were significantly associated with MetS positive T2D. In LD block 2 (Appendix 1: Figure A1.3), two closely associated SNPs (rs2241766 and rs1063539) that showed previous association in overall T2D, as well as an additional SNP rs6773957 ($r^2 = 0.3$ with rs2241766 and rs1063539) were significant (OR = 1.32 - 1.43, $0.005 < P < 0.027$). In addition, two SNPs in LD block 1 (rs266729 and rs182052, $r^2 = 0.5$) showed independent association to MetS positive T2D (OR = 1.35 - 1.36, $0.018 < P < 0.036$).

In *HNF4A,* the three significant SNPs for overall T2D (rs4812828, rs1884614 and rs2144908) located in the P2 promoter regions of LD block 1 were more significant in the

MetS negative T2D subset (OR = 1.32 - 1.35, $0.016 < P < 0.027$) than in the MetS positive subset (OR = 1.22 - 1.32, $0.021 < P < 0.089$). In addition, an independent SNP located in LD block 2 (Appendix 1: Figure A1.10), rs2071199, was also associated with MetS negative T2D subset with a similar effect size (OR = 1.43, $P = 0.031$).

In *KCNJ11*, the two associated SNPs (rs5215 and rs5219) are in strong LD ($r^2 = 1$), with stronger association in the MetS positive T2D subgroup (OR = 1.34 – 1.35, $0.023 < P < 0.026$) than the MetS negative T2D subgroup (OR = 1.23 – 1.24, $P > 0.05$).

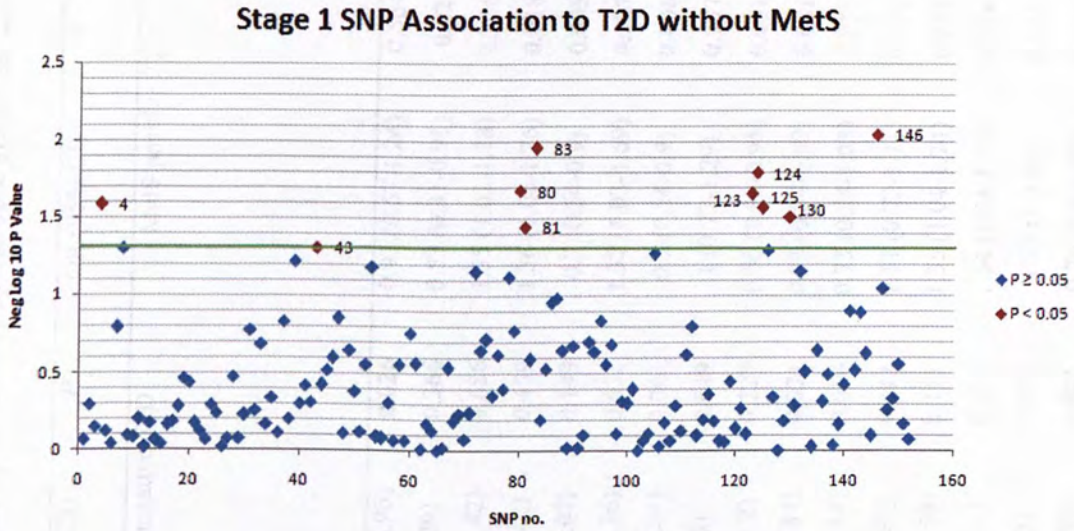
In *LPL*, none of the SNPs showed association with overall T2D but three SNPs were significantly associated with MetS negative T2D. This finding is due to the paradoxical opposite associations in the MetS subset analyses. While the C allele of rs291, G allele of rs320 and T allele of rs4921684 increased risk for MetS negative T2D (OR = 1.37 – 1.70, $0.011 < P < 0.036$), the former two risk alleles conferred trend of decreased risk for MetS positive T2D (OR = 1.37 – 1.70, $0.011 < P < 0.036$).

For the other genes with single association signals, SNPs in *PBX1* and *PPARA* showed suggestive association to MetS negative T2D, whereas SNPs in *PPARG* and *LIPC* showed suggestive association to MetS positive T2D. Two SNPs in *PPARGC1A* showed trend of associations in different MetS subgroups, making the result inconclusive.

In summary, additional loci and independent signals were observed in MetS subset analyses that were not detected in the overall case-control comparison. This is in part due to a stronger genetic effect when phenotypic homogeneity of sample was increased. In particular, the clustering of association signals in *ADIPOQ* in MetS positive T2D and *HNF4A* in MetS negative T2D were consistent with their expected roles in insulin action and insulin secretion, respectively. The association of *PPARG* and *LIPC* with MetS positive T2D, and *PBX1* with MetS negative T2D were also consistent with

their roles in the respective pathways. However, the association results for *PPARGC1A*, *KCNJ11* and *PPARA* did not entirely fit the pathways discussed in the hypothesis and warrant further investigation.

A)



B)

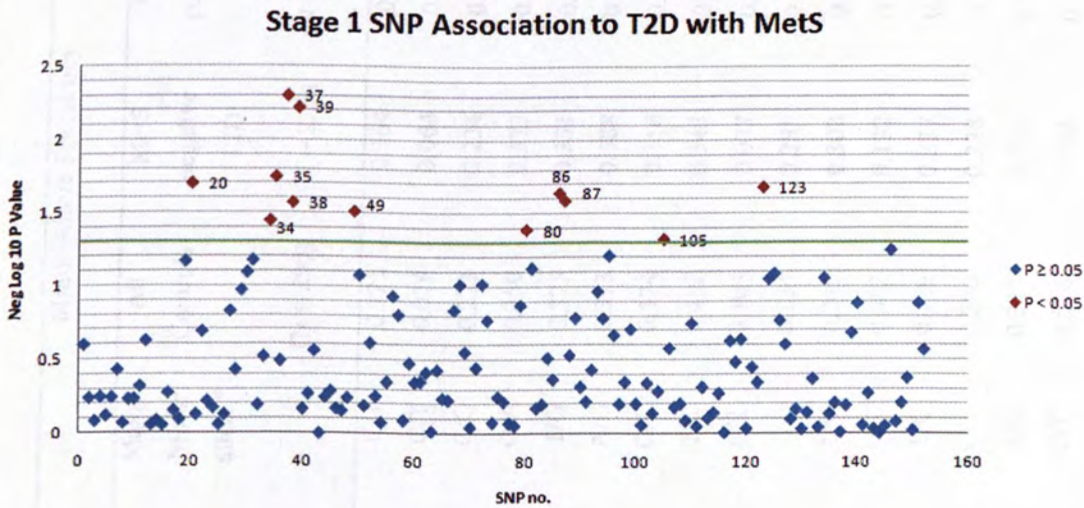


Figure 3.2: Graphical summaries of association results for stage 1 SNPs (shown as negative $\log_{10}$ P value) in A) T2D without metabolic syndrome versus all controls; B) T2D with metabolic syndrome versus all controls. The green line indicates the threshold of $P = 0.05$. The SNP number corresponds to the SNP order shown in Appendix 2.

Table 3.3: Significant T2D association results of SNPs in stage 1 case-control samples subset by metabolic syndrome status in cases

				Minor Allele Frequency			OR (95% CI)	*P*	OR (95% CI)	*P*
SNP no.	Gene	rs number	Major/ Minor allele [a]	All Controls (N = 290)	MetS negative T2D (N = 223)	MetS positive T2D (N = 244)	MetS negative T2D		MetS positive T2D	
4	*PBX1*	rs3767374	C/T	0.221	0.164	0.206	0.68 (0.48-0.96)	**0.026**	0.91 (0.67-1.24)	0.564
20	*PPARG*	rs2972164	C/T	0.079	0.064	0.045	0.8 (0.5-1.29)	0.366	0.53 (0.31-0.91)	**0.02**
34	*ADIPOQ*	*rs266729	C/G	0.223	0.234	0.279	1.06 (0.79-1.42)	0.686	1.35 (1.02-1.78)	**0.036**
35	*ADIPOQ*	rs182052	G/A	0.350	0.373	0.421	1.1 (0.85-1.42)	0.459	1.36 (1.06-1.76)	**0.018**
37	*ADIPOQ*	*rs2241766	T/G	0.375	0.331	0.293	0.83 (0.65-1.07)	0.149	0.7 (0.54-0.9)	**0.005**
38	*ADIPOQ*	rs6773957	A/G	0.373	0.388	0.439	1.06 (0.83-1.36)	0.632	1.32 (1.03-1.69)	**0.027**
39	*ADIPOQ*	rs1063539	G/C	0.374	0.316	0.294	0.79 (0.61-1.01)	0.061	0.7 (0.54-0.9)	**0.006**
43	*PPARGC1A*	rs12650562	T/C	0.482	0.543	0.482	1.3 (1-1.68)	**0.049**	1 (0.77-1.29)	0.997
49	*PPARGC1A*	rs1472095	C/T	0.103	0.077	0.061	0.75 (0.47-1.2)	0.227	0.56 (0.33-0.95)	**0.031**
80	*LPL*	rs291	T/C	0.234	0.299	0.185	1.39 (1.05-1.84)	**0.021**	0.73 (0.54-0.99)	**0.041**
81	*LPL*	*rs320	T/G	0.245	0.308	0.198	1.37 (1.02-1.84)	**0.036**	0.75 (0.54-1.03)	0.075
83	*LPL*	rs4921684	C/T	0.082	0.132	0.090	1.7 (1.13-2.55)	**0.011**	1.11 (0.72-1.71)	0.632
86	*KCNJ11*	*rs5215	T/C	0.289	0.337	0.355	1.23 (0.95-1.59)	0.111	1.35 (1.04-1.74)	**0.023**
87	*KCNJ11*	*rs5219	C/T	0.290	0.338	0.354	1.24 (0.96-1.6)	0.105	1.34 (1.04-1.73)	**0.026**
105	*LIPC*	rs2242062	A/G	0.317	0.375	0.375	1.28 (1-1.65)	0.054	1.29 (1-1.66)	**0.048**
123	*HNF4A*	rs4812828	C/T	0.526	0.454	0.454	0.75 (0.58-0.96)	**0.022**	0.76 (0.6-0.96)	**0.021**

124	*HNF4A*	*rs1884614	C/T	0.399	0.474	0.452	1.35 (1.06-1.73)	**0.016**	1.22 (0.97-1.55)	0.089
125	*HNF4A*	*rs2144908	G/A	0.406	0.476	0.462	1.32 (1.03-1.69)	**0.027**	1.23 (0.97-1.55)	0.082
130	*HNF4A*	rs2071199	T/C	0.146	0.197	0.148	1.43 (1.03-1.96)	**0.031**	1.02 (0.73-1.41)	0.929
146	*PPARA*	rs881740	A/G	0.090	0.143	0.127	1.71 (1.14-2.55)	**0.009**	1.48 (0.99-2.22)	0.057

Significant associations at $P < 0.05$ are bolded.

* literature significant SNPs.

[a] minor alleles were defined according to the frequencies in cases shown in Appendix 2.

3.3. Case-control associations in stage 2

3.3.1. SNP selection for genotyping

From the stage 1 association analyses on 1) all T2D cases vs. controls, 2) MetS negative T2D cases vs. controls, and 3) MetS positive T2D cases vs. controls, 22 SNPs in 12 genes, except those in *PDX1* and *NEUROD1*, showed nominal significance in one or more analyses. These SNPs were prioritized for genotyping in an independent stage 2 samples consisted of 994 random T2D patients and 310 control subjects. Due to the strong LD between rs5215 and rs5219 of *KCNJ11* ($r^2 = 0.9$), only rs5219 was genotyped in stage 2. In addition, two literature significant SNPs (rs745975 of *HNF4A* and rs4148643 of *ABCC8*) were also included to refine the association signals within the large LD blocks in the respective genes. After removal of rs182052 in *ADIPOQ* that failed for call rate, a total of 22 SNPs were successfully genotyped and analyzed in stage 2 (Table 3.4).

Table 3.4: Summary of 22 SNPs successfully genotyped in stage 2

				Minor allele frequency				OR (95%CI)	*P*	OR (95%CI)	*P*	OR (95%CI)	*P*	
SNP no	*Geme*	rs number	Major/ Minor allele [a]	All controls (N=290)	All cases (N=467)	MetS negative T2D (N=223)	MetS positive T2D (N=244)	All cases		MetS negative T2D		MetS positive T2D		
4	*PBX1*	rs3767374	C/T	0.221	0.186	0.164	0.206	0.8 (0.61-1.05)	0.106	0.68 (0.48-0.96)	**0.026**	0.91 (0.67-1.24)	0.564	√
20	*PPARG*	rs2972164	C/T	0.079	0.055	0.064	0.045	0.67 (0.44-1.01)	0.057	0.8 (0.5-1.29)	0.366	0.53 (0.31-0.91)	**0.020**	√
34	*ADIPOQ*	*rs266729	C/G	0.223	0.257	0.234	0.279	1.2 (0.94-1.54)	0.134	1.06 (0.79-1.42)	0.686	1.35 (1.02-1.78)	**0.036**	√
37	*ADIPOQ*	*rs2241766	T/G	0.375	0.311	0.331	0.293	0.76 (0.62-0.94)	**0.011**	0.83 (0.65-1.07)	0.149	0.7 (0.54-0.9)	**0.005**	√
38	*ADIPOQ*	rs6773957	A/G	0.373	0.415	0.388	0.439	1.18 (0.96-1.46)	0.112	1.06 (0.83-1.36)	0.632	1.32 (1.03-1.69)	**0.027**	√
39	*ADIPOQ*	rs1063539	G/C	0.374	0.305	0.316	0.294	0.74 (0.6-0.92)	**0.006**	0.79 (0.61-1.01)	0.061	0.7 (0.54-0.9)	**0.006**	√
43	*PPARGC1A*	rs12650562	T/C	0.482	0.512	0.543	0.482	1.13 (0.91-1.4)	0.272	1.3 (1-1.68)	**0.049**	1 (0.77-1.29)	0.997	√
49	*PPARGC1A*	rs1472095	C/T	0.103	0.069	0.077	0.061	0.66 (0.44-0.99)	**0.042**	0.75 (0.47-1.2)	0.227	0.56 (0.33-0.95)	**0.031**	√
72	*GCK*	rs13239289	C/G	0.062	0.091	0.093	0.089	1.48 (1-2.19)	**0.049**	1.51 (0.97-2.35)	0.071	1.47 (0.93-2.31)	0.096	√
80	*LPL*	rs291	T/C	0.234	0.239	0.299	0.185	1.02 (0.8-1.3)	0.855	1.39 (1.05-1.84)	**0.021**	0.73 (0.54-0.99)	**0.041**	√
81	*LPL*	*rs320	T/G	0.245	0.251	0.308	0.198	1.03 (0.8-1.33)	0.825	1.37 (1.02-1.84)	**0.036**	0.75 (0.54-1.03)	0.075	√
83	*LPL*	rs4921684	C/T	0.082	0.11	0.132	0.09	1.38 (0.96-1.98)	0.079	1.7 (1.13-2.55)	**0.011**	1.11 (0.72-1.71)	0.632	√
87	*KCNJ11*	*rs5219	T/C	0.289	0.346	0.337	0.355	1.29 (1.03-1.6)	**0.024**	1.23 (0.95-1.59)	0.111	1.35 (1.04-1.74)	0.023	√
88	*ABCC8*	*rs4148643	G/A	0.081	0.062	0.061	0.063	0.76 (0.52-1.13)	0.172	0.75 (0.47-1.2)	0.227	0.79 (0.5-1.23)	0.296	√
95	*HNF1A*	rs1169286	A/G	0.409	0.461	0.455	0.466	1.23 (1-1.51)	**0.049**	1.2 (0.94-1.52)	0.146	1.26 (0.99-1.6)	0.062	√
105	*LIPC*	rs2242062	A/G	0.317	0.375	0.375	0.375	1.28 (1.03-1.59)	**0.024**	1.28 (1-1.65)	0.054	1.29 (1-1.66)	**0.048**	√
123	*HNF4A*	rs4812828	C/T	0.526	0.454	0.454	0.454	0.75 (0.61-0.92)	**0.007**	0.75 (0.58-0.96)	**0.022**	0.76 (0.6-0.96)	**0.021**	√
124	*HNF4A*	*rs1884614	C/T	0.399	0.463	0.474	0.452	1.28 (1.05-1.57)	**0.017**	1.35 (1.06-1.73)	**0.016**	1.22 (0.97-1.55)	0.089	√

125	*HNF4A*	*rs2144908	G/A	0.406	0.469	0.476	0.462	1.27 (1.04-1.56)	**0.021**	1.32 (1.03-1.69)	**0.027**	1.23 (0.97-1.55)	0.082	√
130	*HNF4A*	rs2071199	T/C	0.146	0.172	0.197	0.148	1.21 (0.91-1.6)	0.190	1.43 (1.03-1.96)	**0.031**	1.02 (0.73-1.41)	0.929	√
132	*HNF4A*	*rs745975	G/A	0.2	0.169	0.156	0.181	0.81 (0.62-1.06)	0.124	0.74 (0.54-1.02)	0.070	0.88 (0.65-1.2)	0.428	√
146	*PPARA*	rs881740	A/G	0.09	0.135	0.143	0.127	1.59 (1.12-2.25)	**0.010**	1.71 (1.14-2.55)	**0.009**	1.48 (0.99-2.22)	0.057	√

Significant associations at $P < 0.05$ are bolded.

* literature significant SNPs.

[a] minor alleles were defined according to the frequencies in cases shown in Appendix 2.

√ indicates SNPs successfully genotyped in stage 2.

3.3.2. Association with T2D

Figure 3.3 and Appendix 5 summarized the stage 2 association results of 22 SNPs in 994 cases and 310 controls. None of the significant SNPs in stage 1remained significant in stage 2. Despite the rare SNP rs2972164 of *PPARG* demonstrated a trend of association (P = 0.067), the direction of association was opposite to that in stage 1 [OR (95 CI) 0.67 (0.44 1.01) in stage 1 vs. 1.5 (0.97-2.31) in stage 2].

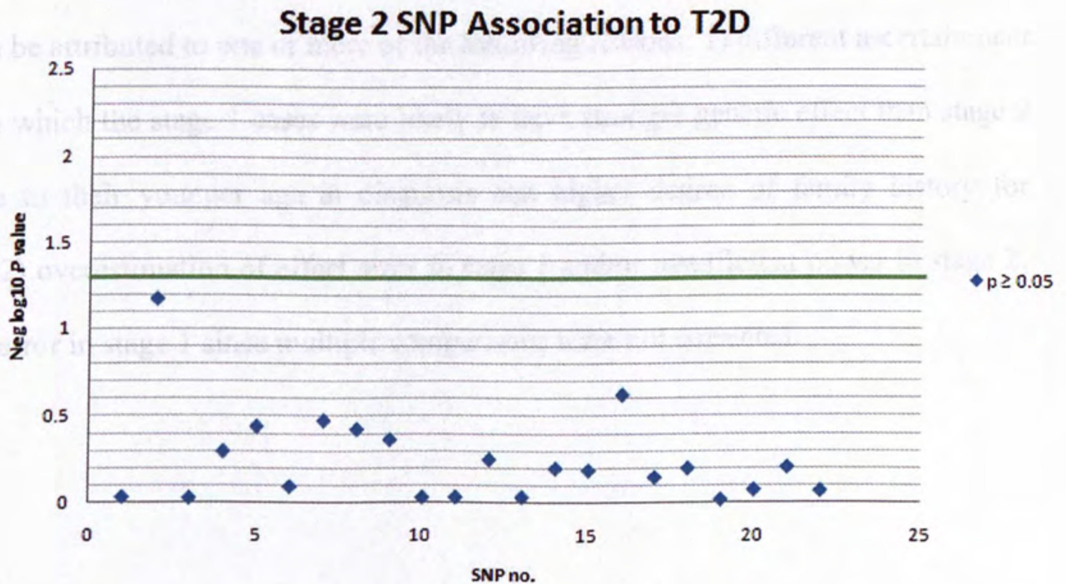


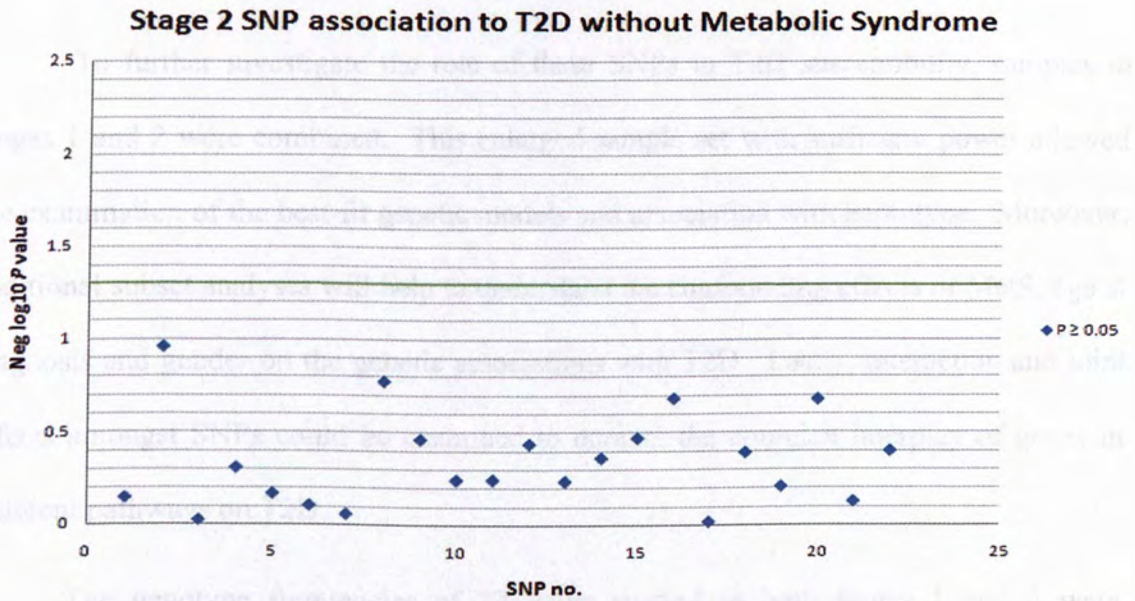
Figure 3.3: Graphical summary of association results for T2D for stage 2 SNPs in case-control samples (shown as negative $\log_{10}$ P value). The green line indicates the threshold of P = 0.05. The SNP number corresponds to the SNP order shown in Appendix 5.

3.3.3. Association with T2D subset by metabolic syndrome

Subsequent subset analyses compared the genotype frequencies of 417 MetS negative cases and 578 MetS positive cases to 310 controls respectively. Figure 3.4 and Appendix 6 summarized the association results of two subset analyses in stage 2. Similar to the findings in the overall case-control comparison, none of the 22 SNPs remained significant in the subset analyses. However, the minor G allele of rs6773957 of *ADIPOQ* showed trend of increased association with MetS positive T2D, similar to that found in stage 1 [OR (95%CI) = 1.32 (1.03-1.69) in stage 1 vs. 1.20 (0.98-1.46) in stage 2].

The non-replication of association in overall and subset analyses between stages 1 and 2 can be attributed to one or more of the following reasons: 1) different ascertainment criteria in which the stage 1 cases were likely to have stronger genetic effect than stage 2 cases due to their younger age at diagnosis and higher degree of family history for diabetes; 2) overestimation of effect sizes in stage 1 and/or insufficient power in stage 2; 3) type 1 error in stage 1 since multiple comparisons were not corrected.

A)



B)

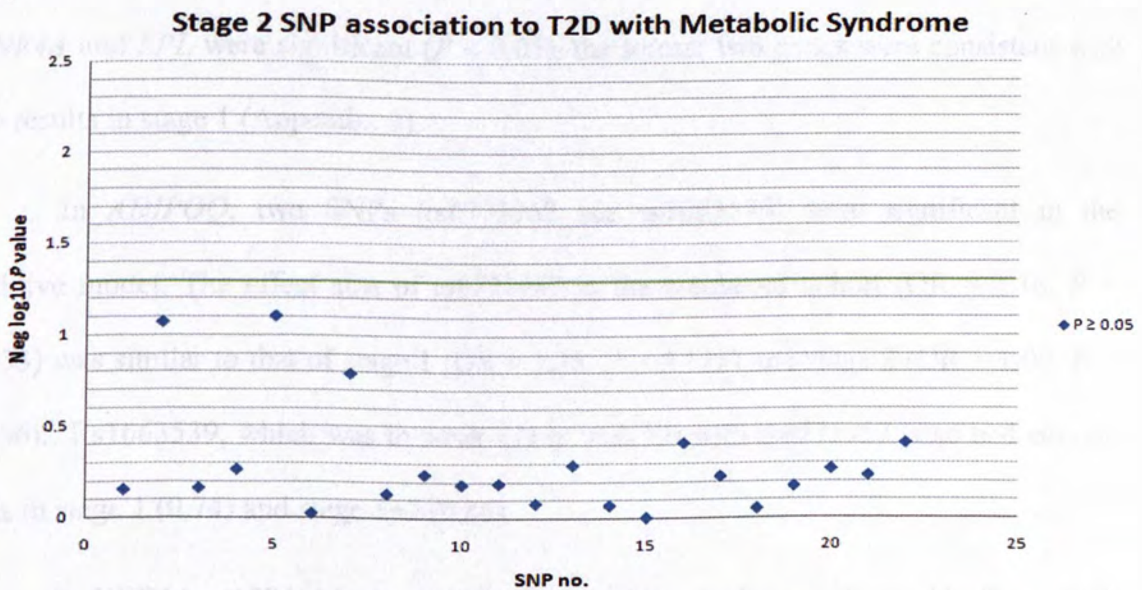


Figure 3.4: Graphical summaries of association results for stage 2 SNPs (shown as negative $\log_{10}$ P value) in A) T2D without metabolic syndrome versus all controls; B) T2D with metabolic syndrome versus all controls. The green line indicates the threshold of $P = 0.05$. The SNP number corresponds to the SNP order shown in Appendix 6.

3.4. Case-control associations in combined stages 1 and 2

3.4.1. Association with T2D

To further investigate the role of these SNPs in T2D susceptibility, samples in stages 1 and 2 were combined. This enlarged sample set with sufficient power allowed the examination of the best-fit genetic models and association with haplotype. Moreover, additional subset analyses will help to understand the confounding effects of MetS, age at diagnosis and gender on the genetic associations with T2D. Lastly, interaction and joint effects amongst SNPs could be examined to unravel the complex interplay of genes in different pathways on T2D.

The genotype frequencies of 22 SNPs studied in both stages 1 and 2 were compared between 1462 cases and 600 controls (Table 3.5). In total, 4 SNPs in *ADIPOQ, HNF4A* and *LPL* were significant ($P < 0.05$), the former two genes were consistent with the results in stage 1 (Appendix 3).

In *ADIPOQ*, two SNPs (rs6773957 and rs1063539) were significant in the additive model. The effect size of rs6773957 in the combined cohort (OR = 1.16, P = 0.033) was similar to that of stage 1 (OR = 1.18, P = 0.112) and stage 2 (OR = 1.09, P = 0.366). Rs1063539, which was in weak LD (r^2 = 0.29) with rs6773957, also had similar ORs in stage 1 (0.74) and stage 1+2 (0.86).

In *HNF4A*, rs1884614 was significant in both the additive (OR = 1.15, P = 0.049) and dominant (OR = 1.24, P = 0.038) models, similar to that found in stage 1 (OR = 1.28, P = 0.017). Furthermore, an independent SNP (rs745975) also showed a trend of association with T2D (OR = 0.84, P = 0.075).

In *LPL*, rs320 was associated with T2D only in the recessive model (OR = 1.63, P = 0.044) with the rare GG genotype (0.061 in cases vs. 0.040 in controls) being at risk. For SNPs in other genes, most of the effects were lost in the combined samples.

Table 3.5: T2D association results of SNPs in the stage 1+2 case-control samples

SNP no.	Gene	rs number	Major/ Minor allele	Genotype frequency[a] Cases (N = 1462)			Genotype frequency[a] Controls (N = 600)			Additive		Recessive		Dominant	
				11	12	22	11	12	22	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P
1	*PBX1*	rs3767374	C/T	0.046	0.32	0.634	0.042	0.351	0.607	0.94 (0.78-1.12)	0.481	1.11 (0.66-1.87)	0.702	0.89 (0.72-1.11)	0.315
2	*PPARG*	rs2972164	C/T	0.006	0.109	0.885	0.002	0.12	0.878	0.99 (0.75-1.32)	0.963	3.26 (0.41-26.16)	0.265	0.96 (0.71-1.29)	0.77
3	*ADIPOQ*	*rs266729	C/G	0.066	0.379	0.555	0.053	0.367	0.579	1.09 (0.93-1.28)	0.265	1.19 (0.79-1.8)	0.413	1.1 (0.91-1.34)	0.329
4	*ADIPOQ*	*rs2241766	T/G	0.097	0.408	0.495	0.122	0.413	0.464	0.88 (0.76-1.02)	0.082	0.77 (0.57-1.06)	0.107	0.87 (0.72-1.07)	0.181
5	*ADIPOQ*	rs6773957	A/G	0.188	0.495	0.317	0.156	0.488	0.356	1.16 (1.01-1.34)	**0.033**	1.26 (0.97-1.63)	0.083	1.2 (0.98-1.46)	0.08
6	*ADIPOQ*	rs1063539	G/C	0.099	0.391	0.51	0.119	0.418	0.463	0.86 (0.75-0.99)	**0.037**	0.81 (0.6-1.09)	0.163	0.83 (0.68-1)	0.052
7	*PPARGC1A*	rs12650562	T/C	0.253	0.505	0.242	0.212	0.535	0.252	1.11 (0.97-1.28)	0.144	1.26 (1-1.59)	0.055	1.05 (0.84-1.32)	0.644
8	*PPARGC1A*	rs1472095	C/T	0.006	0.142	0.852	0.009	0.154	0.838	0.89 (0.69-1.15)	0.381	0.68 (0.22-2.08)	0.494	0.9 (0.68-1.18)	0.436
9	*GCK*	rs13239289	C/G	0.009	0.135	0.856	0.007	0.129	0.864	1.09 (0.84-1.41)	0.504	1.31 (0.42-4.02)	0.641	1.09 (0.83-1.44)	0.545
10	*LPL*	rs291	T/C	0.058	0.331	0.61	0.04	0.374	0.587	0.99 (0.84-1.16)	0.886	1.54 (0.96-2.47)	0.071	0.9 (0.74-1.1)	0.310
11	*LPL*	*rs320	T/G	0.061	0.327	0.611	0.04	0.381	0.579	0.98 (0.83-1.15)	0.761	1.63 (1.01-2.64)	**0.044**	0.87 (0.72-1.07)	0.185
12	*LPL*	rs4921684	C/T	0.01	0.179	0.811	0.004	0.161	0.835	1.21 (0.95-1.56)	0.128	2.99 (0.68-13.18)	0.149	1.19 (0.91-1.55)	0.203
13	*KCNJ11*	*rs5219	C/T	0.117	0.425	0.458	0.104	0.409	0.487	1.11 (0.96-1.28)	0.151	1.17 (0.86-1.6)	0.322	1.14 (0.94-1.38)	0.189
14	*ABCC8*	*rs4148643	G/A	0.002	0.122	0.876	0.008	0.134	0.857	0.82 (0.62-1.1)	0.185	0.25 (0.06-1.11)	0.069	0.85 (0.63-1.15)	0.303
15	*HNF1A*	rs1169286	A/G	0.193	0.509	0.297	0.182	0.477	0.341	1.12 (0.98-1.29)	0.107	1.08 (0.85-1.39)	0.527	1.22 (0.99-1.49)	0.058
16	*LIPC*	rs2242062	A/G	0.122	0.459	0.419	0.119	0.447	0.434	1.05 (0.91-1.21)	0.534	1.06 (0.79-1.42)	0.702	1.06 (0.87-1.29)	0.553
17	*HNF4A*	rs4812828	C/T	0.229	0.496	0.274	0.258	0.488	0.253	0.9 (0.79-1.03)	0.139	0.85 (0.68-1.06)	0.157	0.89 (0.72-1.11)	0.310
18	*HNF4A*	*rs1884614	C/T	0.206	0.487	0.307	0.185	0.46	0.355	1.15 (1-1.31)	**0.049**	1.14 (0.9-1.46)	0.276	1.24 (1.01-1.52)	**0.038**
19	*HNF4A*	*rs2144908	G/A	0.219	0.475	0.306	0.201	0.455	0.344	1.11 (0.97-1.27)	0.120	1.11 (0.88-1.41)	0.371	1.18 (0.97-1.45)	0.105
20	*HNF4A*	rs2071199	T/C	0.026	0.258	0.715	0.026	0.247	0.727	1.06 (0.87-1.28)	0.570	1.01 (0.55-1.85)	0.980	1.07 (0.87-1.33)	0.521
21	*HNF4A*	*rs745975	G/A	0.025	0.279	0.696	0.038	0.303	0.659	0.84 (0.69-1.02)	0.075	0.65 (0.37-1.16)	0.149	0.84 (0.68-1.05)	0.129
22	*PPARA*	rs881740	A/G	0.02	0.215	0.765	0.014	0.187	0.799	1.21 (0.98-1.5)	0.081	1.46 (0.66-3.22)	0.352	1.22 (0.96-1.55)	0.097

Significant associations at $P < 0.05$ are bolded.
* literature significant SNPs.
[a] allele 1 refers to the minor allele defined according to the frequency in cases shown in Appendix 2.

Haplotype analysis was also performed to examine the association signals in *ADIPOQ, HNF4A* and *LPL*. The frequency of the CGAC haplotype of *ADIPOQ* was significantly lower in the cases than in the controls (P = 0.019) (Table 3.6). The protection effect of this haplotype was mainly driven by the protective G allele of rs2241766 and C allele of rs1063539, which had r^2 of 0.83. Similarly, the common at risk CTATG haplotype of *HNF4A* (P = 0.049) was mainly driven by the at risk T allele of rs1884614 and A allele of rs2144908 which were in strong LD ($r^2 = 0.94$). No association was observed for *LPL* haplotypes. It is likely that single association signals in these genes contributed to T2D susceptibility.

Table 3.6: Haplotype association of significant genes in stage 1+2 case-control samples

Gene	SNPs	Haplotype	Frequency		*P* value
			Case (N = 1461)	Control (N = 600)	
ADIPOQ	*rs266729,	GTGG	0.163	0.151	0.339
	*rs2241766,	CTGG	0.283	0.261	0.165
	rs6773957,	GTAG	0.087	0.081	0.563
	rs1063539	CGAC	0.280	0.318	**0.019**
		CTAG	0.187	0.190	0.863
LPL	rs291,	CGT	0.096	0.083	0.200
	*rs320,	CGC	0.127	0.142	0.204
	rs4921684	TTC	0.777	0.775	0.889
HNF4A	rs4812828,	CTATA	0.074	0.078	0.685
	*rs1884614,	TCGTA	0.089	0.109	0.064
	*rs2144908,	TCGCG	0.101	0.102	0.876
	rs2071199,	CTATG	0.354	0.32	**0.049**
	*rs745975	TCGTG	0.322	0.333	0.506
		CCGTG	0.060	0.058	0.761

Significant associations at $P < 0.05$ are bolded.

* literature significant SNPs.

3.4.2. Association with T2D subset by metabolic syndrome

Given the demonstrated value of combined analysis and increased phenotypic homogeneity, additional subgroup analyses involving 1) MetS; 2) age at diagnosis; and 3) gender were examined in the combined stages 1 and 2 samples in the following sections.

From Table 3.7, 822 MetS negative cases and 640 MetS positive cases were compared to 600 controls respectively. Three SNPs (rs6773957 and rs1063539 in *ADIPOQ* and rs881740 in *PPARA*) were significant in the MetS positive subgroup comparison, while two SNPs (rs1884614 in *HNF4A*) and (rs4921684 in *LPL*) were significant in the MetS negative subgroup comparison.

In the comparison of MetS positive cases and all controls, the at-risk G allele of rs6773957 in *ADIPOQ* were consistently associated with T2D in stage 1 (OR = 1.32, P = 0.027), stage 2 (OR = 1.20, P = 0.079), and stage 1+2 (OR = 1.29, P = 0.001). Rs1063539 displayed a similar trend with its protective C allele. Even though the associations of rs266729 (P = 0.11) and rs2241766 (P = 0.07) did not achieve statistical significance, their at-risk alleles consistently trended towards association with the MetS positive T2D phenotype. In addition, rs881740 of *PPARA*, a gene upstream of ADIPOQ in the insulin action pathway, was also associated with MetS positive T2D with a modest OR of 1.29 (P = 0.036). However, this SNP showed stronger association with MetS negative T2D in stage 1, in contrast to the combined analyses.

In the MetS negative subgroup analysis, the at-risk T allele of rs1884614 in *HNF4A* was consistent throughout the three stages, with ORs of 1.35 in stage 1, 1.09 in stage 2 and 1.20 in stage 1+2. A similar trend of association was also observed for

MetS negative T2D for an adjacent SNP (rs2144908) in strong LD ($r^2 = 0.94$). For *LPL*, the at-risk T allele of rs4921684 also showed consistent trend of association across all three stages.

Table 3.7: T2D association results of SNPs in the stage 1+2 case-control samples subset by metabolic syndrome status in cases

				Minor Allele Frequency			OR (95% CI)	*P*	OR (95% CI)	*P*
SNP no.	Gene	rs number	Major/ Minor allele	All Controls (N = 600)	MetS negative T2D (N = 822)	MetS positive T2D (N = 640)	MetS negative T2D		MetS positive T2D	
1	*PBX1*	rs3767374	C/T	0.217	0.19	0.218	0.84 (0.68-1.05)	0.129	1.01 (0.83-1.22)	0.958
2	*PPARG*	rs2972164	C/T	0.062	0.064	0.057	1.04 (0.75-1.44)	0.821	0.92 (0.67-1.27)	0.620
3	*ADIPOQ*	*rs266729	C/G	0.237	0.245	0.264	1.04 (0.86-1.26)	0.670	1.15 (0.97-1.38)	0.110
4	*ADIPOQ*	*rs2241766	T/G	0.329	0.31	0.294	0.92 (0.77-1.1)	0.352	0.86 (0.72-1.01)	0.070
5	*ADIPOQ*	rs6773957	A/G	0.4	0.404	0.461	1.02 (0.86-1.2)	0.853	1.29 (1.1-1.5)	**0.001**
6	*ADIPOQ*	rs1063539	G/C	0.328	0.301	0.289	0.89 (0.75-1.05)	0.168	0.84 (0.72-0.99)	**0.030**
7	*PPARGC1A*	rs12650562	T/C	0.48	0.504	0.506	1.1 (0.94-1.3)	0.235	1.12 (0.96-1.3)	0.167
8	*PPARGC1A*	rs1472095	C/T	0.086	0.086	0.071	1 (0.75-1.34)	0.983	0.81 (0.6-1.08)	0.147
9	*GCK*	rs13239289	C/G	0.071	0.076	0.077	1.06 (0.79-1.43)	0.691	1.09 (0.82-1.44)	0.570
10	*LPL*	rs291	T/C	0.227	0.254	0.201	1.16 (0.96-1.39)	0.128	0.85 (0.71-1.03)	0.100
11	*LPL*	*rs320	T/G	0.23	0.253	0.204	1.13 (0.93-1.36)	0.218	0.85 (0.7-1.03)	0.100
12	*LPL*	rs4921684	C/T	0.084	0.112	0.09	1.37 (1.03-1.82)	**0.028**	1.08 (0.81-1.43)	0.620
13	*KCNJ11*	*rs5219	C/T	0.309	0.336	0.324	1.13 (0.95-1.33)	0.162	1.07 (0.91-1.26)	0.390
14	*ABCC8*	*rs4148643	G/A	0.075	0.06	0.066	0.78 (0.56-1.09)	0.143	0.86 (0.63-1.18)	0.357
15	*HNF1A*	rs1169286	A/G	0.421	0.456	0.442	1.15 (0.98-1.35)	0.085	1.09 (0.94-1.27)	0.260
16	*LIPC*	rs2242062	A/G	0.343	0.347	0.356	1.02 (0.86-1.2)	0.837	1.06 (0.9-1.24)	0.480

17	*HNF4A*	rs4812828	C/T	0.503	0.472	0.482	0.88 (0.75-1.04)	0.132	0.92 (0.79-1.07)	0.290
18	*HNF4A*	*rs1884614	C/T	0.415	0.461	0.441	1.2 (1.02-1.41)	**0.025**	1.11 (0.95-1.28)	0.180
19	*HNF4A*	*rs2144908	G/A	0.428	0.467	0.447	1.17 (0.99-1.36)	0.059	1.07 (0.93-1.24)	0.340
20	*HNF4A*	rs2071199	T/C	0.149	0.152	0.158	1.02 (0.82-1.27)	0.870	1.07 (0.87-1.32)	0.520
21	*HNF4A*	*rs745975	G/A	0.19	0.161	0.167	0.82 (0.66-1.03)	0.087	0.85 (0.69-1.06)	0.143
22	*PPARA*	rs881740	A/G	0.107	0.119	0.135	1.12 (0.87-1.44)	0.388	1.28 (1.02-1.62)	**0.036**

Significant associations at $P < 0.05$ are bolded.

* literature significant SNPs.

[a] minor alleles were defined according to the frequencies in cases shown in Appendix 2.

3.4.3. Association with T2D subset by age at diagnosis

Familial diabetes usually occurs at a young age due to increased genetic loading. To investigate for a possible modifying effect of age of onset on diabetes, subset analyses comparing 759 early onset cases (age at diagnosis ≤ 40 years) and 689 late onset cases to 600 controls were performed respectively (Table 3.8).

Three SNPs (rs2241766, rs6773957 and rs1063539) in *ADIPOQ* were significantly associated with late onset T2D (OR = 1.20 – 1.32, $0.001 < P < 0.046$). On the other hand, SNPs from *LPL* (rs4921684) and *PPARA* (rs881740) were associated with early onset T2D (OR = 1.29 - 1.35, $0.032 < P < 0.034$). Moreover, two SNPs in *HNF4A* (rs1884614 and rs745975) also showed a trend for association with early onset T2D, despite the former SNP may also be associated with late onset T2D ($P < 0.1$).

Interestingly, the associated SNPs in *ADIPOQ* for late onset T2D coincided with that of MetS positive T2D, whereas the associated SNPs in *HNF4A* and *LPL* for early onset T2D coincided with that of MetS negative T2D, suggesting a possible correlation between the two phenotypes.

Table 3.8: T2D association results of SNPs in the stage 1+2 case-control samples subset by age at diagnosis in cases

				Minor Allele Frequency			OR (95% CI)	*P*	OR (95% CI)	*P*
SNP no.	Gene	Rs number	Major/Minor allele	All Controls (N = 600)	Early Onset T2D (N = 759)	Late Onset T2D (N = 689)	Early Onset T2D		Late Onset T2D	
1	*PBX1*	rs3767374	C/T	0.217	0.199	0.214	0.89 (0.73-1.1)	0.289	0.98 (0.8-1.2)	0.854
2	*PPARG*	rs2972164	C/T	0.062	0.057	0.063	0.92 (0.66-1.27)	0.601	1.03 (0.75-1.43)	0.854
3	*APM1*	*rs266729	C/G	0.237	0.257	0.257	1.11 (0.93-1.33)	0.244	1.12 (0.93-1.34)	0.24
4	*APM1*	rs2241766	T/G	0.329	0.311	0.289	0.92 (0.78-1.09)	0.337	0.83 (0.7-1)	**0.046**
5	*APM1*	rs6773957	A/G	0.4	0.411	0.465	1.05 (0.9-1.23)	0.553	1.32 (1.12-1.55)	**0.001**
6	*APM1*	rs1063539	G/C	0.328	0.308	0.276	0.92 (0.78-1.08)	0.292	0.79 (0.67-0.94)	**0.006**
7	*PPARGC1A*	rs12650562	T/C	0.48	0.507	0.502	1.12 (0.95-1.3)	0.172	1.1 (0.93-1.29)	0.271
8	*PPARGC1A*	rs1472095	C/T	0.086	0.071	0.083	0.81 (0.61-1.1)	0.175	0.96 (0.72-1.28)	0.789
9	*GCK*	rs13239289	C/G	0.071	0.079	0.075	1.11 (0.83-1.47)	0.486	1.05 (0.78-1.42)	0.734
10	*LPL*	rs291	T/C	0.227	0.231	0.216	1.02 (0.85-1.23)	0.798	0.94 (0.78-1.14)	0.532
11	*LPL*	*rs320	T/G	0.23	0.236	0.214	1.03 (0.86-1.25)	0.727	0.91 (0.75-1.1)	0.32
12	*LPL*	rs4921684	C/T	0.084	0.11	0.086	1.35 (1.03-1.77)	**0.032**	1.03 (0.76-1.39)	0.849
13	*KCNJ11*	*rs5219	C/T	0.309	0.334	0.326	1.12 (0.95-1.31)	0.186	1.08 (0.91-1.28)	0.367
14	*ABCC8*	*rs4148643	G/A	0.075	0.062	0.063	0.81 (0.59-1.12)	0.207	0.83 (0.6-1.15)	0.253
15	*HNF1A*	rs1169286	A/G	0.421	0.454	0.442	1.14 (0.98-1.33)	0.096	1.09 (0.93-1.28)	0.284
16	*LIPC*	rs2242062	A/G	0.343	0.357	0.344	1.06 (0.9-1.24)	0.468	1 (0.85-1.19)	0.971
17	*HNF4A*	rs4812828	C/T	0.503	0.478	0.479	0.91 (0.78-1.06)	0.209	0.91 (0.78-1.07)	0.239
18	*HNF4A*	*rs1884614	C/T	0.415	0.45	0.449	1.15 (0.98-1.33)	0.079	1.15 (0.98-1.34)	0.088
19	*HNF4A*	*rs2144908	G/A	0.428	0.459	0.453	1.12 (0.97-1.3)	0.13	1.1 (0.94-1.29)	0.227
20	*HNF4A*	rs2071199	T/C	0.149	0.157	0.155	1.06 (0.86-1.31)	0.576	1.04 (0.84-1.29)	0.728
21	*HNF4A*	*rs745975	G/A	0.19	0.16	0.168	0.81 (0.65-1.01)	0.058	0.86 (0.7-1.07)	0.179
22	*PPARA*	rs881740	A/G	0.107	0.136	0.118	1.29 (1.02-1.64)	**0.034**	1.1 (0.86-1.41)	0.43

Significant associations at $P < 0.05$ are bolded.

* literature significant SNPs.

[a] minor alleles were defined according to the frequencies in cases shown in Appendix 2.

3.4.4. Association with T2D subset by gender

In view of a slightly higher female predominance of diabetes as well as sexual differences in phenotypes (e.g. lipids, body composition) and other environmental risk factors, subset analyses for T2D association were performed in men and women separately (Table 3.9). By comparing the genotype frequencies of 592 male cases with 272 male controls, rs4921684 of *LPL* and rs4148643 of *ABCC8* demonstrated nominal significant association with T2D (OR = 1.49 - 1.51, $0.028 < P < 0.044$). On the other hand, rs6773957 of *ADIPOQ* showed significant association (OR = 1.23, $P = 0.03$) in the comparison of 870 female cases and 328 female controls.

Table 3.9: T2D association results of SNPs in the stage 1+2 case-control samples subset by gender

				Minor Allele Frequency				OR (95% CI)	*P*	OR (95% CI)	*P*
SNP no.	Gene	rs number	Major/ Minor allele	Male Controls (N = 272)	Male T2D (N = 592)	Female Controls (N = 328)	Female T2D (N = 870)	Male Case-Control		Female Case-Control	
1	*PBX1*	rs3767374	C/T	0.218	0.2	0.216	0.211	0.89 (0.68-1.18)	0.432	0.97 (0.76-1.23)	0.787
2	*PPARG*	rs2972164	C/T	0.06	0.072	0.063	0.052	1.21 (0.79-1.86)	0.373	0.83 (0.57-1.2)	0.320
3	*ADIPOQ*	*rs266729	C/G	0.252	0.245	0.225	0.263	0.96 (0.75-1.23)	0.746	1.22 (0.98-1.5)	0.069
4	*ADIPOQ*	*rs2241766	T/G	0.325	0.295	0.332	0.306	0.87 (0.69-1.1)	0.237	0.89 (0.73-1.08)	0.241
5	*ADIPOQ*	rs6773957	A/G	0.417	0.437	0.386	0.435	1.08 (0.88-1.33)	0.447	1.23 (1.02-1.49)	**0.030**
6	*ADIPOQ*	rs1063539	G/C	0.321	0.298	0.333	0.292	0.9 (0.73-1.12)	0.348	0.83 (0.69-1.01)	0.059
7	*PPARGC1A*	rs12650562	T/C	0.485	0.511	0.476	0.502	1.12 (0.9-1.38)	0.315	1.11 (0.92-1.33)	0.272
8	*PPARGC1A*	rs1472095	C/T	0.086	0.08	0.086	0.075	0.92 (0.63-1.36)	0.687	0.87 (0.62-1.22)	0.430
9	*GCK*	rs13239289	C/G	0.072	0.07	0.071	0.081	0.96 (0.65-1.43)	0.846	1.16 (0.82-1.63)	0.406
10	*LPL*	rs291	T/C	0.242	0.235	0.214	0.217	0.96 (0.76-1.22)	0.743	1.02 (0.82-1.27)	0.880
11	*LPL*	*rs320	T/G	0.245	0.234	0.218	0.219	0.94 (0.74-1.21)	0.643	1.00 (0.80-1.26)	0.972
12	*LPL*	rs4921684	C/T	0.083	0.121	0.086	0.085	1.51 (1.05-2.18)	**0.028**	1.00 (0.71-1.41)	0.990
13	*KCNJ11*	*rs5219	C/T	0.31	0.327	0.307	0.331	1.08 (0.86-1.35)	0.502	1.11 (0.92-1.34)	0.290

14	*ABCC8*	*rs4148643	G/A	0.1	0.068	0.057	0.06	0.67 (0.45-0.99)	**0.044**	1.05 (0.69-1.59)	0.830
15	*HNF1A*	rs1169286	A/G	0.404	0.448	0.434	0.448	1.2 (0.97-1.48)	0.093	1.06 (0.88-1.27)	0.546
16	*LIPC*	rs2242062	A/G	0.351	0.359	0.336	0.347	1.04 (0.83-1.29)	0.743	1.05 (0.87-1.27)	0.627
17	*HNF4A*	rs4812828	C/T	0.498	0.477	0.506	0.478	0.92 (0.75-1.13)	0.413	0.89 (0.75-1.07)	0.226
18	*HNF4A*	*rs1884614	C/T	0.404	0.448	0.425	0.451	1.19 (0.97-1.47)	0.093	1.11 (0.93-1.33)	0.262
19	*HNF4A*	*rs2144908	G/A	0.418	0.455	0.437	0.457	1.15 (0.94-1.41)	0.166	1.08 (0.90-1.29)	0.408
20	*HNF4A*	rs2071199	T/C	0.142	0.148	0.156	0.161	1.05 (0.78-1.41)	0.745	1.04 (0.81-1.33)	0.767
21	*HNF4A*	*rs745975	G/A	0.203	0.164	0.179	0.165	0.77 (0.58-1.03)	0.074	0.90 (0.70-1.17)	0.439
22	*PPARA*	rs881740	A/G	0.116	0.128	0.101	0.128	1.12 (0.81-1.54)	0.488	1.29 (0.97-1.73)	0.083

Significant associations at $P < 0.05$ are bolded.

* literature significant SNPs.

[a] minor alleles were defined according to the frequencies in cases shown in Appendix 2.

3.4.5. Genetic epistasis for T2D association

We observed consistent significant association of 5 independent SNPs (rs6773957 and rs1063539 of *ADIPOQ*, rs1884614 of *HNF4A*, rs4921684 of *LPL*, and rs881740 of *PPARA*) in one or more of the stage 1 and 2 analyses for overall T2D and subsets by MetS and age at diagnosis (Tables 3.5, 3.7 and 3.8). In order to test for possible interaction effect amongst these five SNPs on T2D association, a logistic regression model including the main effects of the five SNPs was built under an additive model. Then each of the possible two SNPs interaction was entered in a logistic regression model to test for significance of interaction, but none of the interactions were significant ($P > 0.05$).

In view of the heterogeneous nature of T2D and the different contribution of genes in specific pathways on different subtypes of diabetes, we hypothesized that 1) combination of genetic effects from genes in the same pathway may have stronger effect than that of a single gene in predisposition to T2D related to that pathway; 2) the coexistence of at risk genes from multiple pathways (insulin secretion and action in this study) may further increase the genetic predisposition to T2D.

To test these hypotheses, the five SNPs were divided into two groups according to their association with MetS. As MetS often coexists with insulin resistance, three SNPs in *ADIPOQ* and *PPARA* that showed significant association with MetS positive T2D were considered related to the insulin action pathway. Conversely, two SNPs in *HNF4A* and *LPL* that showed significant association with MetS negative T2D were considered related to the insulin secretion pathway (Table 3.7). For each and the combined pathways, the number of at risk alleles carried by each subject was counted. For the 5 T2D-associated

SNPs, each subject was classified into a low risk (0-2 risk alleles), medium risk (3 risk alleles), moderately high risk (4 risk alleles), or high risk allele (5-8 risk alleles) group according to the number of risk alleles carried. The 3 insulin action SNPs were similarly classified (0-1 risk alleles for low risk, 2 risk alleles for medium risk, 3 risk alleles for moderately high risk, and 4-6 risk alleles for high risk). The 2 insulin secretion SNPs had fewer risk alleles, so carriers with 0, 1, or 2-4 risk alleles were assigned to the low, medium, and high risk groups respectively.

In view of the tendency of clustering of SNP association in MetS positive and late-onset T2D, as well as MetS negative and early-onset T2D, 1182 diabetic subjects with complete genotypes for the five SNPs were classified into four groups to increase phenotypic homogeneity: 1) 331 MetS negative and early-onset T2D; 2) 188 MetS negative and late-onset T2D; 3) 315 MetS positive and early-onset T2D; 4) 348 MetS positive and late-onset T2D. Each of the diabetic subgroup was then compared with all controls according to the pathways involved.

The results of the joint effects of the 5 SNPs for T2D association in different conditions were summarized in Table 3.10. For the three insulin action SNPs in *ADIPOQ* and *PPARA*, stronger association with risk for MetS positive late-onset T2D was observed in carriers with increasing number of risk alleles compared to low risk carriers (OR = 1.44 - 2.24, *P* for trend = 0.00008). Similarly, the two insulin secretion SNPs in *HNF4A* and *LPL* were strongly associated with MetS negative early onset T2D in a dosage-dependent manner (OR = 1.41 - 1.71, *P* for trend = 0.005). When combining SNPs from insulin secretion and action pathways, the association with MetS positive late-onset T2D remained similar (OR = 1.44 - 2.24 for insulin action SNPs vs. OR = 1.66 -

2.08 for five associated SNPs), but the trend became insignificant for the MetS negative early-onset T2D (OR = 1.41 - 1.71 for insulin secretion SNPs vs. OR = 1.22 - 1.62 for all five associated SNPs). When we combined all cases and compared them to controls, the effect sizes of the insulin action and insulin secretion SNPs were similar (OR = 1.33 – 1.49 for insulin action SNPs and OR = 1.26 – 1.43 for insulin secretion SNPs), and combining all five SNPs led to a stronger trend of association (OR = 1.43 – 1.65). The increased effect size by combining risk alleles and dividing patients into subgroups confirmed previous subset results and suggested the importance of dosage and phenotypic homogeneity in association study for T2D.

Table 3.10: Dosage effect of risk alleles of five significant SNPs in stage 1+ 2 case-control association

		Frequency						OR	*P*	OR	*P*	OR	*P*	OR	*P*	OR	*P*
	Allele risk group	All Controls (N = 507)	All Cases (N =1182)	MetS negative early-onset T2D (N = 331)	MetS negative late-onset T2D (N = 188)	MetS positive early-onset T2D (N = 315)	MetS positive late-onset T2D (N = 348)	All cases vs. all controls		MetS negative early-onset T2D vs. all controls		MetS negative late-onset T2D vs. all controls		MetS positive early-onset T2D vs. all controls		MetS positive late-onset T2D vs. all controls	
Five associated SNPs [a]	Low risk (0-2 risk alleles)	0.300	0.218	0.242	0.207	0.241	0.181	1.00	**0.002**	1.00	0.161	1.00	0.114	1.00	0.090	1.00	**0.0002**
	Medium risk (3 risk alleles)	0.229	0.238	0.227	0.261	0.244	0.230	1.43 (1.06-1.92)		1.23 (0.83-1.83)		1.65 (1.01-2.67)		1.33 (0.89-1.98)		1.66 (1.11-2.51)	
	Moderately high risk (4 risk alleles)	0.213	0.256	0.278	0.271	0.216	0.264	1.65 (1.23-2.23)		1.62 (1.10- 2.39)		1.84 (1.13-2.99)		1.26 (0.84-1.90)		2.06 (1.37-3.08)	
	High risk (5-8 risk alleles)	0.258	0.288	0.254	0.261	0.298	0.325	1.53 (1.15-2.03)		1.22 (0.83-1.79)		1.46 (0.90-2.36)		1.44 (0.98-2.10)		2.08 (1.41-3.06)	
Insulin action SNPs [b]	Low risk (0-1 risk allele)	0.250	0.196	0.230	0.181	0.213	0.158	1.00	**0.020**	1.00	0.699	1.00	0.240	1.00	0.352	1.00	**0.00008**
	Medium risk (2 risk alleles)	0.272	0.283	0.287	0.303	0.308	0.247	1.33 (0.99-1.78)		1.15 (0.78-1.69)		1.54 (0.95-2.51)		1.33 (0.90-1.98)		1.44 (0.95-2.18)	
	Moderately high risk (3 risk alleles)	0.268	0.277	0.272	0.309	0.238	0.299	1.32 (0.98-1.77)		1.11 (0.75-1.63)		1.59 (0.98-2.59)		1.05 (0.69-1.57)		1.77 (1.18-2.65)	
	High risk (4-6 risk alleles)	0.209	0.244	0.211	0.207	0.241	0.296	1.49 (1.09-2.03)		1.10 (0.73-1.67)		1.37 (0.81-2.33)		1.36 (0.90-2.06)		2.24 (1.48-3.4)	
Insulin secretion SNPs [c]	Low risk (0 risk allele)	0.316	0.258	0.233	0.250	0.279	0.267	1.00	**0.011**	1.00	**0.005**	1.00	0.207	1.00	0.121	1.00	0.117
	Medium risk (1 risk allele)	0.428	0.442	0.444	0.479	0.416	0.443	1.26 (0.98-1.62)		1.41 (1.00-1.98)		1.41 (0.94-2.12)		1.1 (0.78-1.54)		1.22 (0.88-1.70)	
	High risk (2-4 risk alleles)	0.256	0.300	0.323	0.271	0.305	0.290	1.43 (1.09-1.89)		1.71 (1.18-2.48)		1.34 (0.84-2.11)		1.34 (0.93-1.94)		1.34 (0.93-1.92)	

Significant associations at $P < 0.05$ are bolded.

[a] The five associated SNPs consisted of rs6773957 and rs1063539 of *ADIPOQ*, rs1884614 of *HNF4A*, rs4921684 of *LPL*, and rs881740 of *PPARA*.

[b] SNPs involved in insulin action are rs6773957 and rs1063539 of *ADIPOQ*, and rs881740 of *PPARA*.

[c] SNPs involved in insulin secretion are rs1884614 of *HNF4A* and rs4921684 of *LPL*.

3.5. Metabolic trait associations in control subjects in combined stages 1 and 2 studies

To investigate the effect of the 22 SNPs on metabolic risk factors, linear regression was performed on 9 metabolic traits associated with: 1) lipid (HDL and TG); 2) blood pressure (SBP and DBP); 3) obesity (WHR); 4) insulin sensitivity (HOMA-IR and ISI); and 5) insulin secretion (HOMA-β and IDI) in 600 control subjects (Figure 3.5). Table 3.11 summarized the significant association of SNPs with quantitative traits. Among the nine traits tested, six genes (*LPL, HNF4A, PPARA, PPARGC1A, ADIPOQ* and *LIPC*) demonstrated significant associations to five traits related to lipids, insulin sensitivity and insulin secretion ($P < 0.05$).

For *LPL*, the major T alleles of two SNPs (rs320 and rs291) in strong LD ($r^2 = 0.95$) were associated with decreased HDL level ($0.002 < P < 0.003$), insulin sensitivity (ISI) ($P = 0.026$), and insulin secretion (IDI) ($0.022 < P < 0.044$). While the association of this lipid gene with HDL level is expected, the association with insulin parameters is interesting. It is of note that five SNPs in other lipid genes (*PPARA, PPARGC1A, ADIPOQ* and *LIPC*) were also associated with insulin secretion including IDI and HOMA-β. Similarly, the major G allele of rs745975 in *HNF4A* was associated with both increased TG level and decreased insulin secretion (IDI). The association of multiple genes in regulation of both lipid levels and insulin secretion suggests that glucose and fatty acid metabolisms are interlinked and may be regulated by a common set of candidate genes.

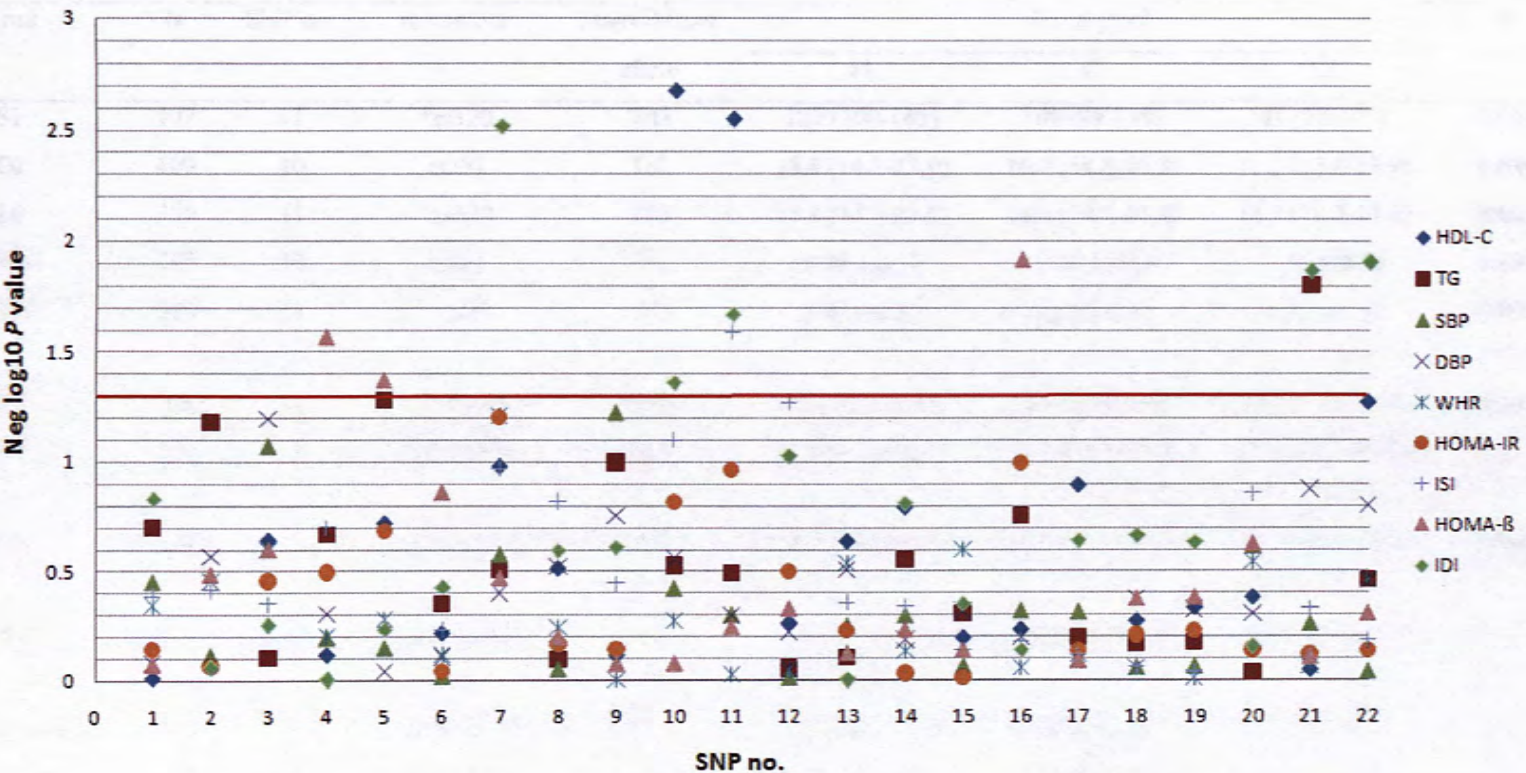


Figure 3.5: Graphical summary of association results for metabolic traits in stage 1 and 2 control samples (shown as negative $\log_{10} P$ value). The red line indicates the threshold of $P = 0.05$. The SNP number corresponds to the SNP order shown in Appendix 6.

Table 3.11: Significant association of candidate genes with metabolic traits in 600 normal controls

Gene	Trait	N	SNP no.	rs number	Major/Minor allele	Genotype [a] 11	12	22	P
LPL	ISI	397	11	*rs320	T/G	127 (100-162)	108 (99-119)	99 (92-107)	**0.026**
	IDI	409	10	rs291	T/C	18.4 (14.1-23.9)	16.7 (14.8-18.8)	15.1 (13.6-16.9)	**0.044**
	IDI	389	11	*rs320	T/G	18.4 (14.1-23.9)	16.6 (14.7-18.8)	14.7 (13.2-16.4)	**0.022**
	HDL	573	10	rs291	T/C	1.70 ± 0.32	1.58 ± 0.42	1.51 ± 0.40	**0.002**
	HDL	550	11	*rs320	T/G	1.70 ± 0.33	1.57 ± 0.42	1.51 ± 0.39	**0.003**
HNF4A	TG	470	21	*rs745975	G/A	0.71 (0.53-0.95)	0.86 (0.79-0.93)	0.93 (0.87-0.99)	**0.016**
	IDI	304	21	*rs745975	G/A	30.0 (17.9-50.3)	18.7 (15.0-23.2)	15.6 (14.1-17.2)	**0.014**
PPARGC1A	IDI	416	7	rs12650562	T/C	12.7 (10.8-15.0)	16.0 (14.4-17.8)	18.1 (15.7-20.9)	**0.003**
PPARA	IDI	417	22	rs881740	A/G	14.8 (6.7-32.4)	13.5 (11.0-16.6)	16.5 (15.2-17.8)	**0.013**
ADIPOQ	HOMA-β	526	4	*rs2241766	T/G	122 (102-147)	105 (95-116)	95 (87-105)	**0.027**
	HOMA-β	559	5	rs6773957	A/G	94 (79-112)	98 (89-107)	112 (101-123)	**0.042**
LIPC	HOMA-β	554	16	rs2242062	A/G	84 (68-103)	102 (93-111)	109 (98-120)	**0.012**

Data are expressed as mean +/- SD or geometric mean (95% CI).
Linear regression was performed for the association of SNPs under an additive genetic model with adjustment for age and gender.
Significant associations at $P < 0.05$ are bolded.
* literature significant SNPs.
[a] allele 1 refers to the minor allele defined according to the frequency in cases shown in Appendix 2.

CHAPTER 4. Discussion

Type 2 diabetes (T2D) is a complex disease with a substantial genetic component. However, its pathophysiological mechanism is not fully understood. While the genome-wide approach is a hypothesis-generating tool to discover novel genes, a candidate gene association study is a hypothesis-testing tool to discern the genetic basis of complex diseases based on known molecular pathways.

In this study, 14 candidate genes belonging to the insulin secretion or insulin action pathways were examined for association with type 2 diabetes in a Hong Kong Chinese population. The study was conducted in two stages. In stage 1, 152 SNPs were genotyped in 467 cases and 290 controls. A subset of 22 significant SNPs in 12 genes were then genotyped in an independent cohort of 994 cases and 310 controls in stage 2 to balance genotyping cost and efficiency. As stated by Skol *et al.* (Skol *et al.* 2006), combining the data from stages 1 and 2 will increase the power to detect an association. Thus we combined data from the two stages and further investigated the modifying effects of metabolic syndrome (MetS), age at diagnosis and gender, as well as gene-gene epistasis and joint effect on the susceptibility to T2D. While the combined sample size was sufficiently powered for detecting SNPs with a moderate effect size when using a nominal significance *P* value level of 0.05 (Section 2.1.2), the power after multiple comparisons (e.g. Bonferroni correction) was greatly diminished given the large number of SNPs tested. To minimize type 1 error while maintaining good study power, we declared significance by reproducible findings in different analyses.

Among the 14 investigated genes, four (*HNF4A, PPARA, LPL* and *ADIPOQ*) demonstrated consistent strong associations with subtypes of T2D. While HNF4A is

a transcription factor (Noland *et al.*) regulating glucose stimulated insulin secretion (GSIS), PPARA, LPL and ADIPOQ are linked to adipocytes and play roles in lipid metabolism and insulin action pathway. Interestingly, these four genes seem to have dual regulatory roles in maintaining glucose and lipid metabolisms and thus enable cross-talk between the insulin secretion and action pathways. The key findings of this study are summarized in the following sections.

4.1. Role of insulin secretion genes in type 2 diabetes

Insulin is a key hormone for maintaining glucose homeostasis, thus defective insulin secretion is an important pathological mechanism for manifestation of T2D. Thus far, 6 genes have been identified to cause maturity-onset diabetes of the young (Yang *et al.*), a monogenic form that accounts for 2-5% of diabetes (Fajans *et al.* 2001). Remarkably, all these genes are primarily involved with the secretion (or the regulation) of insulin in the pancreatic β cells. This study examined the common polymorphisms of five of these genes (*HNF4A, HNF1A, GCK*, *PBX1/PDX1* and *NEUROD1*), in addition to *ABCC8* and *KCNJ11* which are implicated in neonatal diabetes, for association with the more common form of T2D.

HNF4A

Hepatic nuclear factor 4α (HNF4α) primarily regulates gene expressions in the pancreatic β cells and the liver. In the β cells, it is the master regulator that initiates GSIS. In the liver, it responds to insulin stimulation by forming a regulatory network with other TFs (such as HNF1α, PPARGC1α, HNF1β and FOXA2) to control hepatic gene expressions, leading to inhibition of hepatic glucose output and increased glucose uptake.

In the combined samples consisting of 1462 cases and 600 controls, the T allele of rs1884614 was significantly associated with increased risk for overall T2D (OR = 1.15), and with stronger effect in the diabetic subgroup without MetS (OR = 1.20) (Tables 3.5 and 3.7). Two SNPs (rs4812828 and rs2144908) with $r^2 \geq 0.71$ to rs1884614 also showed trend of association (*P* < 0.2) with both overall T2D and MetS negative T2D. In addition, the common G allele of rs745975 ($r^2 = 0$ to the three SNPs above) also conferred trend of increased risk to overall (OR = 1.19), MetS negative (OR = 1.22) and early-onset T2D (OR = 1.23) (*P* < 0.1), further supporting its effect on T2D susceptibility. Interestingly, the G allele of rs745975 also demonstrated lower insulin secretion capacity (IDI) and higher TG level in healthy controls.

Rs1884614, rs4812828 and rs2144908 are located at the P2 promoter region of *HNF4A*. Consistent with our results, the T allele of rs1884614 and other surrogate markers or haplotypes have been repeatedly shown to increase risk for T2D (OR = 1.14 to 1.3) in multiple populations including Japanese, Europeans, Amish and Ashkenazi Jewish (Damcott *et al.* 2004; Hara *et al.* 2006; Johansson *et al.* 2007; Lehman *et al.* 2007), although negative findings were also reported (Vaxillaire *et al.* 2005; Winckler *et al.* 2005). Moreover, rs2144908 was associated with an earlier age of diagnosis of diabetes in Mexican American (Lehman *et al.* 2007). Despite a similar effect size observed across different ethnic groups, the T allele frequency of rs1884614 in Asians (0.40 – 0.45, this study and Hara, Horikoshi *et al.* 2006) is different from that of Caucasians (0.2 - 0.27) and Pima Indians (0.83), suggesting stronger contribution of this gene region on T2D susceptibility in native Americans and Asians than in Europeans.

Another potential T2D-associated SNP, rs745975, is located 2 bp downstream from a consensus 3' splice site in intron 2. Contrary to a Japanese study with the A allele being at risk (Yokoi *et al.* 2006), the G allele showed trend of increased risk for T2D in this study. Nevertheless, a haplotype containing the G allele of rs745975 is associated with elevated TG level in Mexican and Finnish samples (Weissglas-Volkov *et al.* 2006), which is consistent with our findings. Given that the Mexican and Finnish samples consisted of individuals with familial combined hyperlipidemia (FCHL) or premature history of coronary heart disease, their early onset of disease and dyslipidemic family history corroborates with our observed association with early onset diabetes. The G allele was also associated with lower IDI level in the current study, implying that the risk-conferring effect may be mediated through impaired insulin secretion.

Functionally, the P2 promoter region of HNF4α is important for its specific regulation of gene expression in pancreatic islets through binding by other TFs including HNF1α, HNF1β and PDX1 (Thomas *et al.* 2001). SNPs flanking the promoter region may also affect splicing (Weissglas-Volkov *et al.* 2006). Defective function or expression of HNF4α will likely affect insulin secretion from the pancreatic β cells, as supported by the association of rs1884614 and rs745975 with development of MetS negative T2D and decreased IDI level, respectively. The association between rs745975 and TG level was supported by the recent observation that TG in the islets can induce GSIS. Taken together, we observed association of *HNF4A* polymorphisms with risk to develop early onset T2D in the absence of MetS. This is consistent with the *a priori* hypothesis and the literature, providing empirical evidence that link HNF4α-induced insulin secretory defect to T2D.

KCNJ11/ABCC8

Two adjacent genes on chromosome 11 code for the islet ATP-sensitive potassium channel (Kir 6.2 subunit encoded by *KCNJ11*, and sulfonylurea receptor subunit encoded by *ABCC8*). The T allele of rs5219 (E23K) in *KCNJ11* has been widely replicated to confer risk for developing T2D in European populations (Nielsen *et al.* 2003; van Dam *et al.* 2005; Saxena *et al.* 2007; The Wellcome Trust Case Control Consortium 2007; Zeggini *et al.* 2007) with an OR of ~1.14. The T allele is also associated with decreased serum insulin level during OGTT (Nielsen *et al.* 2003), suggesting that this effect is mediated through impairment of GSIS. Although this SNP did not reach statistical significance in this study, similar ORs of 1.07 – 1.13 were observed in the overall T2D and MetS subset analyses of combined stages 1 and 2 samples. The non-replication of rs5219 with T2D in Hong Kong Chinese may be due to insufficient power and thus a further large scale study in Chinese is warranted.

Rs4148643 (also called rs1799859, R1273R and K649K) of *ABCC8* is located in the same LD block (D' ≥ 0.8) but has low r^2 (0.02) with rs5219. The A allele was reported to be associated with a 2 fold increased risk for T2D in 490 Finnish with impaired glucose tolerance (Laukkanen *et al.* 2004). However, the G allele was at-risk for T2D in men in this study (OR = 0.67, P = 0.044). The inconsistency of direction of association and low minor allele frequency in Chinese (0.075 in controls) suggests that this SNP is less likely a genuinely associated SNP, although ethnic differences in the susceptibility allele could not be excluded.

Other insulin secretion genes

Although 6 genes belonging to the insulin secretion pathway demonstrated some evidence of association with T2D in stage 1, SNPs in 5 genes (*PBX1, GCK,*

KCNJ11, *ABCC8* and *HNF1A*) did not maintain their significance in the combined samples including 1462 cases and 600 controls for both overall T2D and subset analyses. In the combined stage 1+2, their ORs were modest (1.09 - 1.12 in the additive model, $P > 0.05$), with conflicting direction of association as compared to stage 1 and stage 2 respectively. This could be due to insufficient study power and/or type I error.

Among these SNPs, rs1169286 of *HNF1A* may warrant further investigation. The G allele showed some trend for increased risk for overall T2D in the combined samples (OR = 1.12 and 1.22 for additive and dominant models, respectively, $P \leq 0.1$). Similar to the pattern of association for rs1884614 of *HNF4A*, the G allele of rs1169286 also showed trend for association with MetS negative T2D (OR = 1.15, $P = 0.085$) and early-onset T2D (OR = 1.14, $P = 0.10$) (Tables 3.7 and 3.8). Given that HNF1α and HNF4α physically and functionally interact to regulate the expression of downstream genes, it is reasonable to speculate that their genetic variants would modulate similar disease phenotypes.

The consistent association of polymorphisms at *HNF4A* with T2D susceptibility suggests that TFs that act as master regulators of genes for insulin secretion play a key role in T2D development. This observation is supported by the recent GWAS, which identified several T2D genes including *TCF7L2* and *HHEX*, which are TFs regulating insulin secretion (Grant *et al.* 2006; Saxena *et al.* 2007; Scott *et al.* 2007; Sladek *et al.* 2007; Zeggini *et al.* 2007).

4.2. Role of insulin action genes in type 2 diabetes

Under normal circumstances, insulin activates tissue-specific protein cascades, inducing liver to convert excess glucose into glycogen and triglyceride (Wittrup *et al.*), as well as adipose tissue and muscle to store free fatty acids (FFA) as TG. However, the presence of insulin resistance leads to decreased insulin response of these target tissues, so that muscle and adipose reduce glucose and FFA uptake, while the liver outputs glucose instead of storing it. These effects disrupt the normal fat partition in the body and leads to elevated plasma glucose. In addition, while elevated glucose induces insulin secretion which controls FFA oxidation in the peripheral tissues, FFA also feeds back to the pancreas to release insulin to control plasma glucose level. Thus the insulin secretion and insulin action pathways interact closely to maintain energy homeostasis. Genetic defects that disrupt this regulatory network may lead to pathogenic consequences such as development of MetS and T2D.

Among others, 6 genes (*PPARG, PPARA, PPARGC1A, ADIPOQ, LPL* and *LIPC*) have been implicated in the control of lipid metabolism. PPARγ regulates FFA uptake in adipose, while PPARα regulates FFA oxidation in the muscle and the liver. In the liver, lipases such as LPL and LIPC are responsible for removing TG from plasma lipids, and releasing FFA that can be easily absorbed by the cells. In the muscle, PPARGC1α and ADIPOQ promote FFA oxidation to provide energy for activity.

PPARA

PPARA encodes the TF peroxisome proliferator-activated receptor-α, a key regulator of ADIPOQ and LPL activities that are involved in the FFA oxidation in the liver and muscle. Similar to PPARγ, PPARα is a drug target for lowering TG and

cholesterol in hyperlipidemic patients, as well as improving insulin sensitivity and GSIS in diabetic patients.

In the combined stages 1 and 2 analyses, the G allele of rs881740 was modestly associated with MetS positive T2D (OR = 1.28, $P = 0.036$), as well as early-onset T2D (OR = 1.29, $P = 0.034$) (Tables 3.7 and 3.8). Rs881740 is a novel T2D-associated SNP discovered through the tagging approach. The association of this SNP with MetS in the diabetic patients is consistent with the regulatory role of PPARα in both glucose and lipid metabolisms, which are often disrupted in patients with MetS. In animal studies, PPARα-null mice demonstrated decreased hepatic glucose production during fasting as compared to the wildtype mice (Patsouris *et al.* 2004). In addition, these null mice also accumulated lipids in the heart and liver when FFA influx was inhibited (Djouadi *et al.* 1998). Our finding of an association with early-onset T2D was also supported by a recent study which found association of a *PPARA* haplotype with 3.75-fold increased risk of developing early onset T2D before the age of 45 years old (Flavell *et al.* 2005).

ADIPOQ

ADIPOQ encodes the hormone adiponectin which is exclusively expressed in the adipocytes that regulates energy homeostasis. As a major insulin-sensitizing cytokine, adiponectin stimulates fat storage in the adipose tissue, inhibits glucose output in the liver and provides TG as an energy fuel in the muscle.

In the combined stages 1 and 2 analyses, two independent SNPs of *ADIPOQ* (rs6773957 and rs1063539) were significantly associated with T2D, especially the forms with MetS positive (OR = 1.19-1.29) and late-onset T2D (OR = 1.30-1.33), as well as for diabetes in women (OR = 1.2-1.23). In addition, the T allele of rs2241766

(r^2 = 0.83 with rs1063539) also showed a trend for increased risk for T2D in the subgroup analyses. The moderately increased ORs in the subgroups, as compared to the whole group, indicate that increased phenotypic (and thereby genetic) homogeneity may help to clarify the specific effects of the SNPs. Interestingly, the T2D risk alleles of rs2241766 and rs6773957 were also associated with lower HOMA-β index, suggesting that adiponectin may also play a role in maintaining homeostasis in β-cell insulin secretion.

Rs6773957, rs1063539 and rs2241766 are all located in LD block 2. Rs2241766 is a synonymous coding SNP located in exon 2, whereas rs6773957 and rs1063539 are located in the 3' untranslated region near exon 3. SNPs in this LD block have been previously reported to be associated with T2D (Hara *et al.* 2002; Menzaghi *et al.* 2002; Yang *et al.* 2007), although negative findings were also reported (Gibson *et al.* 2004; Gu *et al.* 2004; Lee *et al.* 2005). Consistent with our findings, the T allele of rs2241766 confered risk for MetS related phenotypes (especially obesity) in an Italian population (Menzaghi *et al.* 2002), and was further associated with insulin resistance and cardiovascular diseases in a subsequent meta-analysis in Caucasians (Menzaghi *et al.* 2007). In a healthy Austrian population, this allele is also associated with lower serum adiponectin level (Mackevics *et al.* 2006), a predisposition to T2D development. However, a meta-analysis involving 2379 Chinese subjects suggested that the G allele was at risk for T2D (P = 0.05) (Li *et al.* 2007).

In this study, it is unclear why SNPs in the LD block 2 were associated with HOMA-β rather than with HOMA-IR which is closely linked to MetS. One possible mechanism is by indirect modulation of pancreatic β-cell function through regulation of FFA metabolism. On the other hand, the female-specific risk for T2D may be

partially explained by the known sexual variation in plasma adiponectin level (Kadowaki *et al.* 2006). In a study involving 1727 Caucasians, 18% of the variance of adiponectin concentration was explained by gender alone (Heid *et al.* 2006). A Taiwanese study showed that girls have higher plasma adiponectin level than boys, and that adiponectin level was inversely correlated with plasma insulin level (Hung *et al.* 2006). Stumvoll *et al.* demonstrated that genetic association of *ADIPOQ* with insulin sensitivity was only found in the subgroup without first degree family history of diabetes (Stumvoll *et al.* 2002). Despite some conflicting results in ours and others' studies, which may be due to difference in ascertainment, ethnicity and/or type 1 error, the data collectively suggest that polymorphisms in *ADIPOQ* may aggravate insulin resistance and/or impair GSIS in the presence of other risk factors such as aging and obesity.

LPL

LPL encodes the enzyme lipoprotein lipase that controls the rate-limiting step in the lipolysis of TG. In the adipose tissue, insulin stimulates LPL to convert plasma lipoproteins into FFA for cellular nutrient uptake, whereas in the muscle and liver, LPL releases FFA as an energy fuel when insulin is absent. It has also been suggested that LPL inversely controls insulin secretion through the long-chain acyl-CoA (Lc-CoA) and the GPR40 signaling pathway in the pancreatic β cells, leading to enhanced insulin secretion in response to insulin resistance. Thus, LPL may play multiple roles in both glucose and fatty acid metabolisms.

In the combined stages 1 and 2 analyses, the T allele of rs4921684 in *LPL* demonstrated significant association with MetS negative T2D (OR = 1.37), early-onset T2D (OR = 1.35), and increased T2D risk in men (OR = 1.51). Given the known function of LPL in the regulation of lipid levels, its associations with MetS

negative and early-onset T2D seemed to be contradicting as both phenotypes may be more affiliated with defective insulin secretion (as observed in rs1884614 of *HNF4A*). Given that lipid traits are regulated by hormones such as insulin, glucagon and growth hormone that are subject to gender-specific variations, lipid-related genes may also exert gender effect. An earlier study has reported an association of Ser447X (in low r2 with rs4921684) with T2D in women (Larson *et al.* 1999). The at-risk effect of rs4921684 in men in this study is unclear, but it could be linked to indirect effects such as stress and other environmental factors.

Apart from rs4921684, two correlated SNPs (rs320 and rs291) demonstrated nominal association with overall T2D (OR = 1.54 – 1.63, $P < 0.1$) under a recessive model. Interestingly, these two SNPs were also associated with lipids (HDL), insulin sensitivity (ISI) and insulin secretion (IDI) in the control subjects. However, the T2D conferring G allele was associated with better metabolic profiles. Rs320 (also known as *Hind*III and T495G) is an intronic SNP that disrupts a *Hind*III restriction enzyme recognition site by replacing thymine (T) with guanine (G). It has been repeatedly associated with hypertension (Chen *et al.* 2005; Yang *et al.* 2005), atherosclerosis (McGladdery *et al.* 2001) (Gotoda 1999) and cardiovascular diseases (Elosua *et al.* 2006) in several Caucasian and Chinese cohorts. However, few studies on T2D have been investigated, and the results are conflicting (Wang *et al.* 1996; Larson *et al.* 1999). In a case-control hypertension study in Chinese (Ma *et al.* 2003), the G allele (H- allele) was associated with elevated HDL, consistent with the current findings.

Taken together, the results of rs320 and rs4921684 suggest a role of LPL in the regulation of both glucose and lipid metabolisms. In a recent GWAS for blood lipids in 6650 healthy subjects, *LPL* was the only loci confirmed to be associated with TG and HDL levels (Kathiresan *et al.* 2007). Decreasing LPL's activity in mice has

led to elevated TG and reduced HDL. A clinical trial in Japanese found that patients with TG genotype (and not GG genotype) for rs320 had significantly reduced area under the curve of plasma glucose when treated with Simvastatin (Onai *et al.* 1999), which suggests an effect on glucose metabolism. Indeed, LPL is regulated by glucose-induced insulin signaling cascade in the adipose tissues, while LPL releases FFA to inversely regulate insulin secretion in the islets. In animal study, acute exposure of FFA to rat islets increases GSIS, while chronic exposure insensitizes islet's response to glucose (Sako *et al.* 1990), perhaps due to lipotoxicity-induced β cell apoptosis (Lee *et al.* 1994).

Other insulin action genes

Although 6 genes belonging to the insulin action pathway reached statistical significance for association with T2D in stage 1, 4 SNPs in 3 genes (*PPARG, PPARGC1A* and *LIPC*) did not maintain their significance when 1462 cases were compared to 600 controls in the combined stages 1 and 2 ($P > 0.05$). These 4 SNPs had not been previously reported to be associated with T2D. It is possible that the apparent associations of rs12650562 (in *PPARGC1A*) to IDI, and rs2242062 (in *LIPC*) to HOMA-β may reflect statistical fluctuations. A Pro12Ala (rs1801282) polymorphism in *PPARG* has been widely replicated to be associated with T2D in Caucasians (Saxena *et al.* 2007). However, this SNP was not significant in the present stage 1 study and failed for genotyping in stage 2. A previous study in 2730 Chinese also observed inconsistent results (Tai *et al.* 2004).

4.3. Combined genetic effects on risk for type 2 diabetes

Despite the absence of gene-gene interaction, a joint or additive effect was observed for the five associated SNPs (rs6773957 and rs1063539 of *ADIPOQ*, rs1884614 of *HNF4A*, rs4921684 of *LPL*, and rs881740 of *PPARA*). Due to the predominant effect of *HNF4A* and *LPL* on MetS positive and early-onset T2D, and the preferential effect of *ADIPOQ* and *PPARA* on MetS negative and late-onset T2D (Table 3.7), these two sets of genes were examined for association with different subtypes of T2D separately and jointly.

The results in Table 3.10 demonstrated that insulin secretion SNPs (*HNF4A* and *LPL*) were only associated with MetS negative early-onset T2D (OR = 1.41 - 1.71, *P* for trend = 0.005) whereas insulin action SNPs (*ADIPOQ* and *PPARA*) were only associated with MetS positive late-onset T2D (OR = 1.44 - 2.24, *P* for trend = 0.00008). The stronger effect size achieved by: 1) combining SNPs; 2) increasing number of risk alleles; and 3) the same direction of association in the joint SNP analyses (as compared to individual SNPs) confirm that these SNPs influence the susceptibility of different forms of T2D, in agreement with the heterogeneous nature of the disease. Moreover, in view that all five insulin secretion and action risk SNPs demonstrated increased effect size (OR = 1.43 – 1.65) as compared to SNPs in individual pathways (OR = 1.26 – 1.49) for overall T2D, possible detrimental synergistic effect of multiple risk alleles in multiple pathways exists for T2D development. These joint genetic effects were also observed in a recent Chinese study on seven diabetes genes discovered through GWAS (Ng *et al.* 2008).

4.4. Summary

T2D is an emerging world epidemic, and it has particularly strong impact in China given its recent industrialization and the adaptation to an unhealthy lifestyle. Based on previous literature, this study hypothesizes that defects in transcription factors and their downstream target genes in the insulin secretion and action pathways may confer increased susceptibility for T2D. To investigate this hypothesis, 8 genes (*HNF4A, HNF1A, KCNJ11, ABCC8, GCK, NEUROD1* and *PDX1*) were selected from the insulin secretion pathway and 6 genes (*PPARA, PPARG, PPARGC1A, ADIPOQ, LPL* and *LIPC*) were selected from the insulin action pathway. Literature significant SNPs and tagSNPs that captured most of the common variations in each gene were genotyped in a subset of samples in stage 1, and interesting signals were followed up in a larger stage 2 cohort. The overall results suggest that these 4 genes play a subtle role in development of T2D and regulation of its related metabolic traits.

The patterns of genetic associations

In the current study, genetic variants in *HNF4A, PPARA, ADIPOQ* and *LPL* were associated with T2D in the Hong Kong Chinese population using multiple analytical methods. From the current knowledge, HNF4α belongs to the insulin secretion pathway, while PPARα, ADIPOQ and LPL belong to the insulin action pathway. As expected, genes in different pathways display different patterns of association, such as the association of *HNF4A* with MetS negative and early-onset T2D, and the association of *ADIPOQ* with MetS positive late-onset T2D. Conversely, the lack of association with metabolic traits in *PPARA* and the association of MetS negative early-onset T2D in *LPL* prompt additional investigations for their role in T2D development.

The role of transcription factors and adipocyte-secreted proteins in T2D

The important roles of TFs and adipocyte-related proteins on T2D are highlighted in this study. TFs regulate the expression of a large set of target genes, and they act as genetic switches between metabolic pathways to coordinate appropriate changes in response to environmental stimulus. As hypothesized, two TFs (*HNF4A* and *PPARA*) were associated with early-onset T2D. HNF4α is a master regulator of the signaling cascade that induces GSIS. Other TFs such as TCF7L2 and HHEX are also implicated for T2D through action on GSIS (Grant *et al.* 2006; Sladek *et al.* 2007). It is unclear why other TFs were not associated with T2D in the current cohort, but larger studies are required to confirm or refute these negative findings.

Adipose tissue is increasingly recognized for its role in controlling lipid and glucose metabolisms throughout the body (Rosen *et al.* 2006). Two of its secreted proteins, ADIPOQ and LPL, were associated with T2D in the present study. When fat cells are unsaturated, they release starvation signals (such as ADIPOQ) to promote glucose uptake in insulin target tissues. When fat cells are saturated, they release satisfaction signals to inhibit glucose uptake, and induce LPL to convert TG into FFA. These secreted proteins thus maintain the energy balance between glucose and FFA in the adipose tissue. The key role of adipocyte proteins in T2D is supported by the association of other adipocytokines (such as leptin and TNFα) and T2D. The presence of severe insulin resistance and glucose intolerance in adiponectin knock-out mice (Matsuzawa 2005) also support our observations.

An integral pathway for T2D

Extensive studies have implicated the role of insulin secretion genes in glucose metabolism, and insulin action genes in lipid metabolism. In the current

study, however, the four T2D susceptibility genes were linked to both lipid and glucose metabolic traits irrespective of their pathways. For example, *HNF4A* was associated with TG level while lipogenic genes such as *ADIPOQ* and *LPL* were associated with insulin secretion index in control subjects. This paradoxical picture suggests that these candidate genes may control glucose and lipid metabolisms simultaneously in a complex regulatory network.

Key findings of this study

There are three key findings in the present study. Firstly, genetic variants in *HNF4A, PPARA, LPL* and *ADIPOQ* were associated with T2D in our Hong Kong Chinese population, suggesting the role of GSIS-related TFs and adipocyte-related proteins in the development of T2D. Secondly, the four associated genes were likely involved in the regulation of both glucose and lipid metabolisms, and through their dual regulatory roles achieved cross-talk between the insulin secretion and action pathways. Thirdly, joint effects of genes from one or multiple pathways greatly improved the predictive power for T2D, and supported the usefulness of an integrated approach in the genetic studies of complex diseases such as T2D.

4.5. Limitation of this study and future direction

The effect of power

This study was designed to have sufficient power (> 80%) to capture common SNP association at a nominal type I error rate of 5% with odds ratios (ORs) ranging from 1.3 to 1.8 in the stage 1 samples, and ORs of 1.2 to 1.5 in the combined stages 1 and 2 samples. Given that most T2D genes discovered in recent GWAS have very modest effect sizes (OR < 1.3) except for *TCF7L2* (Zeggini *et al.* 2007; Zeggini *et al.*

2008), our study may not have sufficient power to detect associations for rare SNPs and SNPS with smaller effects. Due to limited power, multiple comparison correction was not performed. By accepting a larger type 1 error rate, the data were interpreted with caution and guided by consistency across different analyses.

The effect of genetic and phenotypic heterogeneity

Inconsistent or non-replicated findings may be attributable to several reasons including ethnic differences in risk markers, alleles and/or frequencies, different study designs and over-estimation of effect size by first report (winner's curse). One method to overcome winner's curse is to increase sample size in the follow up studies. In the current study, we did not observe replication of stage 1 results using the stage 2 samples, but some associations were observed in the combined analyses of stages 1 and 2, in particular during the subset analyses. This can be partly explained by the sampling differences and phenotypic heterogeneity between stage 1 and stage 2. Unlike stage 1 cases with all early-onset diabetes, stage 2 cases were more heterogeneous with both early- and late-onset diabetes and with weaker family history of diabetes. This difference in genetic loading makes direct comparison of the two stages results difficult. On the other hand, combining samples in the two stages conferred greater power and supported subset analyses based on phenotypes. Indeed, the increased phenotypic (and thus genetic) homogeneity in the sub-phenotypic analyses revealed associations unobserved in the overall samples comparison. Together, these results highlighted the importance of using a homogeneous cohort to explore small genetic effects in complex diseases.

Future direction

This study contributes to the current literature by identifying *HNF4A, PPARA, ADIPOQ*, and *LPL* as the T2D candidate genes in Hong Kong Chinese. The importance of transcription factors and adipocyte-related proteins in T2D development was highlighted. Further studies are warrant to understand the role of these genes. Firstly, these results should be validated by association studies in larger cohorts of Chinese or Asians. Prospective cohort studies will be particularly useful for the predictive role of these genes on T2D. Secondly, gene-gene interactions should be explored with other genes in the insulin secretion and action pathways. Thirdly, functional studies such as reporter assays and gene knock-out studies would further elucidate the physiological mechanism of these SNPs or genes.

REFERENCES

Alberti, K. G. and P. Z. Zimmet (1998). "Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation." Diabet Med **15**(7): 539-53.

Altshuler, D., J. N. Hirschhorn, M. Klannemark, C. M. Lindgren, M. C. Vohl, J. Nemesh, C. R. Lane, S. F. Schaffner, S. Bolk, C. Brewer, T. Tuomi, D. Gaudet, T. J. Hudson, M. Daly, L. Groop and E. S. Lander (2000). "The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes." Nat Genet **26**(1): 76-80.

Armoni, M., C. Harel and E. Karnieli (2007). "Transcriptional regulation of the GLUT4 gene: from PPAR-gamma and FOXO1 to FFA and inflammation." Trends Endocrinol Metab **18**(3): 100-7.

Auboeuf, D., J. Rieusset, L. Fajas, P. Vallier, V. Frering, J. P. Riou, B. Staels, J. Auwerx, M. Laville and H. Vidal (1997). "Tissue distribution and quantification of the expression of mRNAs of peroxisome proliferator-activated receptors and liver X receptor-alpha in humans: no alteration in adipose tissue of obese and NIDDM patients." Diabetes **46**(8): 1319-27.

Balding, D. J. (2006). "A tutorial on statistical methods for population association studies." Nat Rev Genet **7**(10): 781-91.

Barrett, J. C., B. Fry, J. Maller and M. J. Daly (2005). "Haploview: analysis and visualization of LD and haplotype maps." Bioinformatics **21**(2): 263-5.

Barroso, I., M. Gurnell, V. E. Crowley, M. Agostini, J. W. Schwabe, M. A. Soos, G. L. Maslen, T. D. Williams, H. Lewis, A. J. Schafer, V. K. Chatterjee and S. O'Rahilly (1999). "Dominant negative mutations in human PPARgamma associated with severe insulin resistance, diabetes mellitus and hypertension." Nature **402**(6764): 880-3.

Berg, A. H., T. P. Combs, X. Du, M. Brownlee and P. E. Scherer (2001). "The adipocyte-secreted protein Acrp30 enhances hepatic insulin action." Nat Med **7**(8): 947-53.

Bergman, R. N. (1989). "Lilly lecture 1989. Toward physiological understanding of glucose tolerance. Minimal-model approach." Diabetes **38**(12): 1512-27.

Boden, G. (1997). "Role of fatty acids in the pathogenesis of insulin resistance and NIDDM." Diabetes **46**(1): 3-10.

Boden, G., X. Chen, J. Rosner and M. Barton (1995). "Effects of a 48-h fat infusion on insulin secretion and glucose utilization." Diabetes **44**(10): 1239-42.

Braissant, O., F. Foufelle, C. Scotto, M. Dauca and W. Wahli (1996). "Differential expression of peroxisome proliferator-activated receptors (PPARs): tissue distribution of PPAR-alpha, -beta, and -gamma in the adult rat." Endocrinology **137**(1): 354-66.

Cai, S. J., D. M. Wong, S. H. Chen and L. Chan (1989). "Structure of the human hepatic triglyceride lipase gene." Biochemistry **28**(23): 8966-71.

Chinetti, G., S. Griglio, M. Antonucci, I. P. Torra, P. Delerive, Z. Majd, J. C. Fruchart, J. Chapman, J. Najib and B. Staels (1998). "Activation of proliferator-activated receptors alpha and gamma induces apoptosis of human monocyte-derived macrophages." J Biol Chem **273**(40): 25573-80.

Chiu, K. C., L. M. Chuang, A. Chu and M. Wang (2003). "Transcription factor 1 and beta-cell function in glucose-tolerant subjects." Diabet Med **20**(3): 225-30.

Chiu, K. C., L. M. Chuang, J. M. Ryu, G. P. Tsai and M. F. Saad (2000). "The I27L amino acid polymorphism of hepatic nuclear factor-1alpha is associated with insulin resistance." J Clin Endocrinol Metab **85**(6): 2178-83.

Damcott, C. M., N. Hoppman, S. H. Ott, L. J. Reinhart, J. Wang, T. I. Pollin, J. R. O'Connell, B. D. Mitchell and A. R. Shuldiner (2004). "Polymorphisms in both promoters of hepatocyte nuclear factor 4-alpha are associated with type 2 diabetes in the Amish." Diabetes **53**(12): 3337-41.

de Bakker, P. I., R. Yelensky, I. Pe'er, S. B. Gabriel, M. J. Daly and D. Altshuler (2005). "Efficiency and power in genetic association studies." Nat Genet **37**(11): 1217-23.

Diamond, J. (2003). "The double puzzle of diabetes." Nature **423**(6940): 599-602.

Dinneen, S., J. Gerich and R. Rizza (1992). "Carbohydrate metabolism in non-insulin-dependent diabetes mellitus." N Engl J Med **327**(10): 707-13.

Djouadi, F., C. J. Weinheimer, J. E. Saffitz, C. Pitchford, J. Bastin, F. J. Gonzalez and D. P. Kelly (1998). "A gender-related defect in lipid metabolism and glucose homeostasis in peroxisome proliferator- activated receptor alpha- deficient mice." J Clin Invest **102**(6): 1083-91.

Dupont, W. D. and W. D. Plummer, Jr. (1990). "Power and sample size calculations. A review and computer program." Control Clin Trials **11**(2): 116-28.

El-Assaad, W., J. Buteau, M. L. Peyot, C. Nolan, R. Roduit, S. Hardy, E. Joly, G. Dbaibo, L. Rosenberg and M. Prentki (2003). "Saturated fatty acids synergize with elevated glucose to cause pancreatic beta-cell death." Endocrinology **144**(9): 4154-63.

Elbrecht, A., Y. Chen, C. A. Cullinan, N. Hayes, M. Leibowitz, D. E. Moller and J. Berger (1996). "Molecular cloning, expression and characterization of human peroxisome proliferator activated receptors gamma 1 and gamma 2." Biochem Biophys Res Commun **224**(2): 431-7.

Elosua, R., L. A. Cupples, C. S. Fox, J. F. Polak, R. A. D'Agostino, Sr., P. A. Wolf, C. J. O'Donnell and J. M. Ordovas (2006). "Association between well-characterized lipoprotein-related genetic variants and carotid intimal medial thickness and stenosis: The Framingham Heart Study." Atherosclerosis **189**(1): 222-8.

Esterbauer, H., H. Oberkofler, F. Krempler and W. Patsch (1999). "Human peroxisome proliferator activated receptor gamma coactivator 1 (PPARGC1) gene: cDNA sequence, genomic organization, chromosomal localization, and tissue expression." Genomics **62**(1): 98-102.

Fajans, S. S., G. I. Bell and K. S. Polonsky (2001). "Molecular mechanisms and clinical pathophysiology of maturity-onset diabetes of the young." N Engl J Med **345**(13): 971-80.

Fajas, L., D. Auboeuf, E. Raspe, K. Schoonjans, A. M. Lefebvre, R. Saladin, J. Najib, M. Laville, J. C. Fruchart, S. Deeb, A. Vidal-Puig, J. Flier, M. R. Briggs, B. Staels, H. Vidal and J. Auwerx (1997). "The organization, promoter analysis, and expression of the human PPARgamma gene." J Biol Chem **272**(30): 18779-89.

Flavell, D. M., H. Ireland, J. W. Stephens, E. Hawe, J. Acharya, H. Mather, S. J. Hurel and S. E. Humphries (2005). "Peroxisome proliferator-activated receptor alpha gene variation influences age of onset and progression of type 2 diabetes." Diabetes **54**(2): 582-6.

Florez, J. C., N. Burtt, P. I. de Bakker, P. Almgren, T. Tuomi, J. Holmkvist, D. Gaudet, T. J. Hudson, S. F. Schaffner, M. J. Daly, J. N. Hirschhorn, L. Groop and D. Altshuler (2004). "Haplotype structure and genotype-phenotype correlations of the sulfonylurea receptor and the islet ATP-sensitive potassium channel gene region." Diabetes **53**(5): 1360-8.

Fujimoto, W. Y., R. W. Bergstrom, E. J. Boyko, J. L. Kinyoun, D. L. Leonetti, L. L. Newell-Morris, L. R. Robinson, W. P. Shuman, W. C. Stolov, C. H. Tsunehara and et al. (1994). "Diabetes and diabetes risk factors in second- and third-generation Japanese Americans in Seattle, Washington." Diabetes Res Clin Pract **24 Suppl**: S43-52.

Germer, S. and R. Higuchi (1999). "Single-tube genotyping without oligonucleotide probes." Genome Res **9**(1): 72-8.

Gloyn, A. L., E. R. Pearson, J. F. Antcliff, P. Proks, G. J. Bruining, A. S. Slingerland, N. Howard, S. Srinivasan, J. M. Silva, J. Molnes, E. L. Edghill, T. M. Frayling, I. K. Temple, D. Mackay, J. P. Shield, Z. Sumnik, A. van Rhijn, J. K. Wales, P. Clark, S. Gorman, J. Aisenberg, S. Ellard, P. R. Njolstad, F. M. Ashcroft and A. T. Hattersley (2004). "Activating mutations in the gene encoding the ATP-sensitive potassium-channel subunit Kir6.2 and permanent neonatal diabetes." N Engl J Med **350**(18): 1838-49.

Gotoda, T. (1999). "[Lipoprotein lipase and atherosclerosis]." Tanpakushitsu Kakusan Koso **44**(8 Suppl): 1294-301.

Grant, S. F., G. Thorleifsson, I. Reynisdottir, R. Benediktsson, A. Manolescu, J. Sainz, A. Helgason, H. Stefansson, V. Emilsson, A. Helgadottir, U. Styrkarsdottir, K. P. Magnusson, G. B. Walters, E. Palsdottir, T. Jonsdottir, T. Gudmundsdottir, A. Gylfason, J. Saemundsdottir, R. L. Wilensky, M. P. Reilly, D. J. Rader, Y. Bagger, C. Christiansen, V. Gudnason, G. Sigurdsson, U. Thorsteinsdottir, J. R. Gulcher, A.

Kong and K. Stefansson (2006). "Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes." Nat Genet **38**(3): 320-3.

Hagman, D. K., L. B. Hays, S. D. Parazzoli and V. Poitout (2005). "Palmitate inhibits insulin gene expression by altering PDX-1 nuclear localization and reducing MafA expression in isolated rat islets of Langerhans." J Biol Chem **280**(37): 32413-8.

Hamm, J. K., A. K. el Jack, P. F. Pilch and S. R. Farmer (1999). "Role of PPAR gamma in regulating adipocyte differentiation and insulin-responsive glucose uptake." Ann N Y Acad Sci **892**: 134-45.

Handschin, C. and B. M. Spiegelman (2006). "Peroxisome proliferator-activated receptor gamma coactivator 1 coactivators, energy homeostasis, and metabolism." Endocr Rev **27**(7): 728-35.

Hara, K., M. Horikoshi, H. Kitazato, C. Ito, M. Noda, J. Ohashi, P. Froguel, K. Tokunaga, K. Tobe, R. Nagai and T. Kadowaki (2006). "Hepatocyte nuclear factor-4alpha P2 promoter haplotypes are associated with type 2 diabetes in the Japanese population." Diabetes **55**(5): 1260-4.

Heid, I. M., S. A. Wagner, H. Gohlke, B. Iglseder, J. C. Mueller, P. Cip, G. Ladurner, R. Reiter, A. Stadlmayr, V. Mackevics, T. Illig, F. Kronenberg and B. Paulweber (2006). "Genetic architecture of the APM1 gene and its influence on adiponectin plasma levels and parameters of the metabolic syndrome in 1,727 healthy Caucasians." Diabetes **55**(2): 375-84.

Herzig, S., F. Long, U. S. Jhala, S. Hedrick, R. Quinn, A. Bauer, D. Rudolph, G. Schutz, C. Yoon, P. Puigserver, B. Spiegelman and M. Montminy (2001). "CREB regulates hepatic gluconeogenesis through the coactivator PGC-1." Nature **413**(6852): 179-83.

Horikawa, Y., N. Oda, N. J. Cox, X. Li, M. Orho-Melander, M. Hara, Y. Hinokio, T. H. Lindner, H. Mashima, P. E. Schwarz, L. del Bosque-Plata, Y. Horikawa, Y. Oda, I. Yoshiuchi, S. Colilla, K. S. Polonsky, S. Wei, P. Concannon, N. Iwasaki, J. Schulze, L. J. Baier, C. Bogardus, L. Groop, E. Boerwinkle, C. L. Hanis and G. I. Bell (2000). "Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus." Nat Genet **26**(2): 163-75.

Hung, Y. J., N. F. Chu, S. C. Wang, C. H. Hsieh, C. T. He, C. H. Lee and S. C. Fan (2006). "Correlation of plasma leptin and adiponectin with insulin sensitivity and beta-cell function in children - the Taipei Children Heart Study." Int J Clin Pract **60**(12): 1582-7.

Iwata, I., S. Nagafuchi, H. Nakashima, S. Kondo, T. Koga, Y. Yokogawa, T. Akashi, T. Shibuya, Y. Umeno, T. Okeda, S. Shibata, S. Kono, M. Yasunami, H. Ohkubo and Y. Niho (1999). "Association of polymorphism in the NeuroD/BETA2 gene with type 1 diabetes in the Japanese." Diabetes **48**(2): 416-9.

Johansson, S., H. Raeder, S. A. Eide, K. Midthjell, K. Hveem, O. Sovik, A. Molven and P. R. Njolstad (2007). "Studies in 3,523 Norwegians and meta-analysis in 11,571

subjects indicate that variants in the hepatocyte nuclear factor 4 alpha (HNF4A) P2 region are associated with type 2 diabetes in Scandinavians." Diabetes **56**(12): 3112-7.
Jump, D. B., D. Botolin, Y. Wang, J. Xu, B. Christian and O. Demeure (2005). "Fatty acid regulation of hepatic gene transcription." J Nutr **135**(11): 2503-6.

Kadowaki, T., T. Yamauchi, N. Kubota, K. Hara, K. Ueki and K. Tobe (2006). "Adiponectin and adiponectin receptors in insulin resistance, diabetes, and the metabolic syndrome." J Clin Invest **116**(7): 1784-92.

Kahn, C. R. (1994). "Banting Lecture. Insulin action, diabetogenes, and the cause of type II diabetes." Diabetes **43**(8): 1066-84.

Karim, M. A., X. Wang, T. C. Hale and S. C. Elbein (2005). "Insulin Promoter Factor 1 variation is associated with type 2 diabetes in African Americans." BMC Med Genet **6**: 37.

Kathiresan, S., A. K. Manning, S. Demissie, R. B. D'Agostino, A. Surti, C. Guiducci, L. Gianniny, N. P. Burtt, O. Melander, M. Orho-Melander, D. K. Arnett, G. M. Peloso, J. M. Ordovas and L. A. Cupples (2007). "A genome-wide association study for blood lipid phenotypes in the Framingham Heart Study." BMC Med Genet **8 Suppl 1**: S17.

Kelley, D. E., B. Goodpaster, R. R. Wing and J. A. Simoneau (1999). "Skeletal muscle fatty acid metabolism in association with insulin resistance, obesity, and weight loss." Am J Physiol **277**(6 Pt 1): E1130-41.

Kim, J. H., H. D. Shin, B. L. Park, Y. M. Cho, S. Y. Kim, H. K. Lee and K. S. Park (2005). "Peroxisome proliferator-activated receptor gamma coactivator 1 alpha promoter polymorphisms are associated with early-onset type 2 diabetes mellitus in the Korean population." Diabetologia **48**(7): 1323-30.

Kim, S. K., L. Selleri, J. S. Lee, A. Y. Zhang, X. Gu, Y. Jacobs and M. L. Cleary (2002). "Pbx1 inactivation disrupts pancreas development and in Ipf1-deficient mice promotes diabetes mellitus." Nat Genet **30**(4): 430-5.

King, H., R. E. Aubert and W. H. Herman (1998). "Global burden of diabetes, 1995-2025: prevalence, numerical estimates, and projections." Diabetes Care **21**(9): 1414-31.

Knutti, D., A. Kaul and A. Kralli (2000). "A tissue-specific coactivator of steroid receptors, identified in a functional genetic screen." Mol Cell Biol **20**(7): 2411-22.

Kota, B. P., T. H. Huang and B. D. Roufogalis (2005). "An overview on biological mechanisms of PPARs." Pharmacol Res **51**(2): 85-94.

Kunej, T., M. Globocnik Petrovic, P. Dovc, B. Peterlin and D. Petrovic (2004). "A Gly482Ser polymorphism of the peroxisome proliferator-activated receptor-gamma coactivator-1 (PGC-1) gene is associated with type 2 diabetes in Caucasians." Folia Biol (Praha) **50**(5): 157-8.

Larson, I., M. M. Hoffmann, J. M. Ordovas, E. J. Schaefer, W. Marz and J. Kreuzer (1999). "The lipoprotein lipase HindIII polymorphism: association with total cholesterol and LDL-cholesterol, but not with HDL and triglycerides in 342 females." Clin Chem **45**(7): 963-8.

Laukkanen, O., J. Pihlajamaki, J. Lindstrom, J. Eriksson, T. T. Valle, H. Hamalainen, P. Ilanne-Parikka, S. Keinanen-Kiukaanniemi, J. Tuomilehto, M. Uusitupa and M. Laakso (2004). "Polymorphisms of the SUR1 (ABCC8) and Kir6.2 (KCNJ11) genes predict the conversion from impaired glucose tolerance to type 2 diabetes. The Finnish Diabetes Prevention Study." J Clin Endocrinol Metab **89**(12): 6286-90.

Leahy, J. L., S. Bonner-Weir and G. C. Weir (1988). "Minimal chronic hyperglycemia is a critical determinant of impaired insulin secretion after an incomplete pancreatectomy." J Clin Invest **81**(5): 1407-14.

Leahy, J. L. and G. C. Weir (1988). "Evolution of abnormal insulin secretory responses during 48-h in vivo hyperglycemia." Diabetes **37**(2): 217-22.

Lee, Y., H. Hirose, M. Ohneda, J. H. Johnson, J. D. McGarry and R. H. Unger (1994). "Beta-cell lipotoxicity in the pathogenesis of non-insulin-dependent diabetes mellitus of obese rats: impairment in adipocyte-beta-cell relationships." Proc Natl Acad Sci U S A **91**(23): 10878-82.

Lehman, D. M., D. K. Richardson, C. P. Jenkinson, K. J. Hunt, T. D. Dyer, R. J. Leach, R. Arya, H. E. Abboud, J. Blangero, R. Duggirala and M. P. Stern (2007). "P2 promoter variants of the hepatocyte nuclear factor 4alpha gene are associated with type 2 diabetes in Mexican Americans." Diabetes **56**(2): 513-7.

Li, J. K., M. C. Ng, W. Y. So, C. K. Chiu, R. Ozaki, P. C. Tong, C. S. Cockram and J. C. Chan (2006). "Phenotypic and genetic clustering of diabetes and metabolic syndrome in Chinese families with type 2 diabetes mellitus." Diabetes Metab Res Rev **22**(1): 46-52.

Li, S., L. Li, K. Li, X. Qi, D. Hoekema, H. Liu and G. Yang (2007). "Association of adipose most abundant transcript 1 gene (apM1) with type 2 diabetes mellitus in a Chinese population: a meta-analysis of case-control studies." Clin Endocrinol (Oxf).

Lindsay, R. S., T. Funahashi, R. L. Hanson, Y. Matsuzawa, S. Tanaka, P. A. Tataranni, W. C. Knowler and J. Krakoff (2002). "Adiponectin and development of type 2 diabetes in the Pima Indian population." Lancet **360**(9326): 57-8.

Liu, L., W. Jia, T. Zheng, M. Li, H. Lu and K. Xiang (2006). "Ala45Thr variation in neuroD1 gene is associated with early-onset type 2 diabetes with or without diabetic pedigree in Chinese." Mol Cell Biochem **290**(1-2): 199-204.

Love-Gregory, L. D., J. Wasson, J. Ma, C. H. Jin, B. Glaser, B. K. Suarez and M. A. Permutt (2004). "A common polymorphism in the upstream promoter region of the hepatocyte nuclear factor-4 alpha gene on chromosome 20q is associated with type 2 diabetes and appears to contribute to the evidence for linkage in an ashkenazi jewish population." Diabetes **53**(4): 1134-40.

Ma, Y. Q., G. N. Thomas, M. C. Ng, J. A. Critchley, J. C. Chan and B. Tomlinson (2003). "The lipoprotein lipase gene HindIII polymorphism is associated with lipid levels in early-onset type 2 diabetic patients." Metabolism **52**(3): 338-43.

Mackevics, V., I. M. Heid, S. A. Wagner, P. Cip, H. Doppelmayr, A. Lejnieks, H. Gohlke, G. Ladurner, T. Illig, B. Iglseder, F. Kronenberg and B. Paulweber (2006). "The adiponectin gene is associated with adiponectin levels but not with characteristics of the insulin resistance syndrome in healthy Caucasians." Eur J Hum Genet **14**(3): 349-56.

Malecki, M. T. (2005). "Genetics of type 2 diabetes mellitus." Diabetes Res Clin Pract **68 Suppl1**: S10-21.

Malecki, M. T., U. S. Jhala, A. Antonellis, L. Fields, A. Doria, T. Orban, M. Saad, J. H. Warram, M. Montminy and A. S. Krolewski (1999). "Mutations in NEUROD1 are associated with the development of type 2 diabetes mellitus." Nat Genet **23**(3): 323-8.

Malecki, M. T., T. Klupa, D. K. Moczulski and J. J. Rogus (2003). "The Ala45Thr polymorphism of BETA2/NeuroD1 gene and susceptibility to type 1 diabetes mellitus in caucasians." Exp Clin Endocrinol Diabetes **111**(5): 251-4.

Matschinsky, F., Y. Liang, P. Kesavan, L. Wang, P. Froguel, G. Velho, D. Cohen, M. A. Permutt, Y. Tanizawa, T. L. Jetton and et al. (1993). "Glucokinase as pancreatic beta cell glucose sensor and diabetes gene." J Clin Invest **92**(5): 2092-8.

Matsuzawa, Y. (2005). "Adipocytokines and metabolic syndrome." Semin Vasc Med **5**(1): 34-9.

McCarthy, M. I. and P. Froguel (2002). "Genetic approaches to the molecular understanding of type 2 diabetes." Am J Physiol Endocrinol Metab **283**(2): E217-25.

McGladdery, S. H., S. N. Pimstone, S. M. Clee, J. F. Bowden, M. R. Hayden and J. J. Frohlich (2001). "Common mutations in the lipoprotein lipase gene (LPL): effects on HDL-cholesterol levels in a Chinese Canadian population." Atherosclerosis **156**(2): 401-7.

Meetoo, D., P. McGovern and R. Safadi (2007). "An epidemiological overview of diabetes across the world." Br J Nurs **16**(16): 1002-7.

Menzaghi, C., T. Ercolino, R. Di Paola, A. H. Berg, J. H. Warram, P. E. Scherer, V. Trischitta and A. Doria (2002). "A haplotype at the adiponectin locus is associated with obesity and other features of the insulin resistance syndrome." Diabetes **51**(7): 2306-12.

Menzaghi, C., V. Trischitta and A. Doria (2007). "Genetic influences of adiponectin on insulin resistance, type 2 diabetes, and cardiovascular disease." Diabetes **56**(5): 1198-209.

Michael, L. F., Z. Wu, R. B. Cheatham, P. Puigserver, G. Adelmant, J. J. Lehman, D. P. Kelly and B. M. Spiegelman (2001). "Restoration of insulin-sensitive glucose

transporter (GLUT4) gene expression in muscle cells by the transcriptional coactivator PGC-1." Proc Natl Acad Sci U S A **98**(7): 3820-5.

Motojima, K., P. Passilly, J. M. Peters, F. J. Gonzalez and N. Latruffe (1998). "Expression of putative fatty acid transporter genes are regulated by peroxisome proliferator-activated receptor alpha and gamma activators in a tissue- and inducer-specific manner." J Biol Chem **273**(27): 16710-4.

Mukherjee, R., L. Jow, G. E. Croston and J. R. Paterniti, Jr. (1997). "Identification, characterization, and tissue distribution of human peroxisome proliferator-activated receptor (PPAR) isoforms PPARgamma2 versus PPARgamma1 and activation with retinoid X receptor agonists and antagonists." J Biol Chem **272**(12): 8071-6.

Naya, F. J., H. P. Huang, Y. Qiu, H. Mutoh, F. J. DeMayo, A. B. Leiter and M. J. Tsai (1997). "Diabetes, defective pancreatic morphogenesis, and abnormal enteroendocrine differentiation in BETA2/neuroD-deficient mice." Genes Dev **11**(18): 2323-34.

NCEPIII (2001). "Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III)." Jama **285**(19): 2486-97.

Neel, J. V. (1976). The Genetics of Diabetes Mellitus. W.Creutzfeldt, J.Kobberling, J.Neel and Eds, Springer-Verlag, Berlin; New York: pp. 1-11.

Ng, M. C., K. S. Park, B. Oh, C. H. Tam, Y. M. Cho, H. D. Shin, V. K. Lam, R. C. Ma, W. Y. So, Y. S. Cho, H. L. Kim, H. K. Lee, J. C. Chan and N. H. Cho (2008). "Implication of Genetic Variants near TCF7L2, SLC30A8, HHEX, CDKAL1, CDKN2A/B, IGF2BP2 and FTO in Type 2 Diabetes and Obesity in 6719 Asians." Diabetes.

Nielsen, E. M., L. Hansen, B. Carstensen, S. M. Echwald, T. Drivsholm, C. Glumer, B. Thorsteinsson, K. Borch-Johnsen, T. Hansen and O. Pedersen (2003). "The E23K variant of Kir6.2 associates with impaired post-OGTT serum insulin response and increased risk of type 2 diabetes." Diabetes **52**(2): 573-7.

Njolstad, P. R., O. Sovik, A. Cuesta-Munoz, L. Bjorkhaug, O. Massa, F. Barbetti, D. E. Undlien, C. Shiota, M. A. Magnuson, A. Molven, F. M. Matschinsky and G. I. Bell (2001). "Neonatal diabetes mellitus due to complete glucokinase deficiency." N Engl J Med **344**(21): 1588-92.

Noland, R. C., T. L. Woodlief, B. R. Whitfield, S. M. Manning, J. R. Evans, R. W. Dudek, R. M. Lust and R. N. Cortright (2007). "Peroxisomal-mitochondrial oxidation in a rodent model of obesity-associated insulin resistance." Am J Physiol Endocrinol Metab **293**(4): E986-E1001.

Odom, D. T., N. Zizlsperger, D. B. Gordon, G. W. Bell, N. J. Rinaldi, H. L. Murray, T. L. Volkert, J. Schreiber, P. A. Rolfe, D. K. Gifford, E. Fraenkel, G. I. Bell and R. A. Young (2004). "Control of pancreas and liver gene expression by HNF transcription factors." Science **303**(5662): 1378-81.

Olefsky, J. M. (2001). "Nuclear receptor minireview series." J Biol Chem **276**(40): 36863-4.

Onai, T., S. Okada, K. Oshima and M. Masatomo (1999). "Effects of Sinvastatin on Glucose Metabolism in Non-Insulin Dependent Diabetic Patients. The Influence of LPL Gene Hind III Polymorphism." Journal of the Japan Diabetic Society **42**(8): 673-678.

Owen, K. R. and M. I. McCarthy (2007). "Genetics of type 2 diabetes." Curr Opin Genet Dev **17**(3): 239-44.

Pappan, K. L., Z. Pan, G. Kwon, C. A. Marshall, T. Coleman, I. J. Goldberg, M. L. McDaniel and C. F. Semenkovich (2005). "Pancreatic beta-cell lipoprotein lipase independently regulates islet glucose metabolism and normal insulin secretion." J Biol Chem **280**(10): 9023-9.

Patsouris, D., S. Mandard, P. J. Voshol, P. Escher, N. S. Tan, L. M. Havekes, W. Koenig, W. Marz, S. Tafuri, W. Wahli, M. Muller and S. Kersten (2004). "PPARalpha governs glycerol metabolism." J Clin Invest **114**(1): 94-103.

Patti, M. E., A. J. Butte, S. Crunkhorn, K. Cusi, R. Berria, S. Kashyap, Y. Miyazaki, I. Kohane, M. Costello, R. Saccone, E. J. Landaker, A. B. Goldfine, E. Mun, R. DeFronzo, J. Finlayson, C. R. Kahn and L. J. Mandarino (2003). "Coordinated reduction of genes of oxidative metabolism in humans with insulin resistance and diabetes: Potential role of PGC1 and NRF1." Proc Natl Acad Sci U S A **100**(14): 8466-71.

Picard, F., M. Kurtev, N. Chung, A. Topark-Ngarm, T. Senawong, R. Machado De Oliveira, M. Leid, M. W. McBurney and L. Guarente (2004). "Sirtl promotes fat mobilization in white adipocytes by repressing PPAR-gamma." Nature **429**(6993): 771-6.

Pierce, M., H. Keen and C. Bradley (1995). "Risk of diabetes in offspring of parents with non-insulin-dependent diabetes." Diabet Med **12**(1): 6-13.

Piwernetz, K., P. D. Home, O. Snorgaard, M. Antsiferov, K. Staehr-Johansen and M. Krans (1993). "Monitoring the targets of the St Vincent Declaration and the implementation of quality management in diabetes care: the DIABCARE initiative. The DIABCARE Monitoring Group of the St Vincent Declaration Steering Committee." Diabet Med **10**(4): 371-7.

Powell, K. (2007). "Obesity: the two faces of fat." Nature **447**(7144): 525-7.

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. de Bakker, M. J. Daly and P. C. Sham (2007). "PLINK: a tool set for whole-genome association and population-based linkage analyses." Am J Hum Genet **81**(3): 559-75.

Reymer, P. W., B. E. Groenemeyer, E. Gagne, L. Miao, E. E. Appelman, J. C. Seidel, D. Kromhout, S. M. Bijvoet, K. van de Oever, T. Bruin and et al. (1995). "A

frequently occurring mutation in the lipoprotein lipase gene (Asn291Ser) contributes to the expression of familial combined hyperlipidemia." Hum Mol Genet **4**(9): 1543-9.
Robertson, R. P., J. Harmon, P. O. Tran and V. Poitout (2004). "Beta-cell glucose toxicity, lipotoxicity, and chronic oxidative stress in type 2 diabetes." Diabetes **53 Suppl 1**: S119-24.

Robertson, R. P., J. Harmon, P. O. Tran, Y. Tanaka and H. Takahashi (2003). "Glucose toxicity in beta-cells: type 2 diabetes, good radicals gone bad, and the glutathione connection." Diabetes **52**(3): 581-7.

Roden, M., T. B. Price, G. Perseghin, K. F. Petersen, D. L. Rothman, G. W. Cline and G. I. Shulman (1996). "Mechanism of free fatty acid-induced insulin resistance in humans." J Clin Invest **97**(12): 2859-65.

Roden, M., H. Stingl, V. Chandramouli, W. C. Schumann, A. Hofer, B. R. Landau, P. Nowotny, W. Waldhausl and G. I. Shulman (2000). "Effects of free fatty acid elevation on postabsorptive endogenous glucose production and gluconeogenesis in humans." Diabetes **49**(5): 701-7.

Rosen, E. D. and B. M. Spiegelman (2006). "Adipocytes as regulators of energy balance and glucose homeostasis." Nature **444**(7121): 847-53.

Sako, Y. and V. E. Grill (1990). "A 48-hour lipid infusion in the rat time-dependently inhibits glucose-induced insulin secretion and B cell oxidation through a process likely coupled to fatty acid oxidation." Endocrinology **127**(4): 1580-9.

Santamarina-Fojo, S., H. Gonzalez-Navarro, L. Freeman, E. Wagner and Z. Nong (2004). "Hepatic lipase, lipoprotein metabolism, and atherogenesis." Arterioscler Thromb Vasc Biol **24**(10): 1750-4.

Sato, Y., M. Nagasaki, M. Kubota, T. Uno and N. Nakai (2007). "Clinical aspects of physical exercise for diabetes/metabolic syndrome." Diabetes Res Clin Pract **77 Suppl 1**: S87-91.

Savage, P. J., P. H. Bennett, R. G. Senter and M. Miller (1979). "High prevalence of diabetes in young Pima Indians: evidence of phenotypic variation in a genetically isolated population." Diabetes **28**(10): 937-42.

Saxena, R., B. F. Voight, V. Lyssenko, N. P. Burtt, P. I. de Bakker, H. Chen, J. J. Roix, S. Kathiresan, J. N. Hirschhorn, M. J. Daly, T. E. Hughes, L. Groop, D. Altshuler, P. Almgren, J. C. Florez, J. Meyer, K. Ardlie, K. Bengtsson Bostrom, B. Isomaa, G. Lettre, U. Lindblad, H. N. Lyon, O. Melander, C. Newton-Cheh, P. Nilsson, M. Orho-Melander, L. Rastam, E. K. Speliotes, M. R. Taskinen, T. Tuomi, C. Guiducci, A. Berglund, J. Carlson, L. Gianniny, R. Hackett, L. Hall, J. Holmkvist, E. Laurila, M. Sjogren, M. Sterner, A. Surti, M. Svensson, M. Svensson, R. Tewhey, B. Blumenstiel, M. Parkin, M. Defelice, R. Barry, W. Brodeur, J. Camarata, N. Chia, M. Fava, J. Gibbons, B. Handsaker, C. Healy, K. Nguyen, C. Gates, C. Sougnez, D. Gage, M. Nizzari, S. B. Gabriel, G. W. Chirn, Q. Ma, H. Parikh, D. Richardson, D. Ricke and S. Purcell (2007). "Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels." Science **316**(5829): 1331-6.

Scherer, P. E., S. Williams, M. Fogliano, G. Baldini and H. F. Lodish (1995). "A novel serum protein similar to C1q, produced exclusively in adipocytes." J Biol Chem **270**(45): 26746-9.

Schoonjans, K., J. Peinado-Onsurbe, A. M. Lefebvre, R. A. Heyman, M. Briggs, S. Deeb, B. Staels and J. Auwerx (1996). "PPARalpha and PPARgamma activators direct a distinct tissue-specific transcriptional response via a PPRE in the lipoprotein lipase gene." Embo J **15**(19): 5336-48.

Scott, L. J., K. L. Mohlke, L. L. Bonnycastle, C. J. Willer, Y. Li, W. L. Duren, M. R. Erdos, H. M. Stringham, P. S. Chines, A. U. Jackson, L. Prokunina-Olsson, C. J. Ding, A. J. Swift, N. Narisu, T. Hu, R. Pruim, R. Xiao, X. Y. Li, K. N. Conneely, N. L. Riebow, A. G. Sprau, M. Tong, P. P. White, K. N. Hetrick, M. W. Barnhart, C. W. Bark, J. L. Goldstein, L. Watkins, F. Xiang, J. Saramies, T. A. Buchanan, R. M. Watanabe, T. T. Valle, L. Kinnunen, G. R. Abecasis, E. W. Pugh, K. F. Doheny, R. N. Bergman, J. Tuomilehto, F. S. Collins and M. Boehnke (2007). "A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants." Science **316**(5829): 1341-5.

Silander, K., K. L. Mohlke, L. J. Scott, E. C. Peck, P. Hollstein, A. D. Skol, A. U. Jackson, P. Deloukas, S. Hunt, G. Stavrides, P. S. Chines, M. R. Erdos, N. Narisu, K. N. Conneely, C. Li, T. E. Fingerlin, S. K. Dhanjal, T. T. Valle, R. N. Bergman, J. Tuomilehto, R. M. Watanabe, M. Boehnke and F. S. Collins (2004). "Genetic variation near the hepatocyte nuclear factor-4 alpha gene predicts susceptibility to type 2 diabetes." Diabetes **53**(4): 1141-9.

Sivitz, W. I. (2001). "Lipotoxicity and glucotoxicity in type 2 diabetes. Effects on development and progression." Postgrad Med **109**(4): 55-9, 63-4.

Skol, A. D., L. J. Scott, G. R. Abecasis and M. Boehnke (2006). "Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies." Nat Genet **38**(2): 209-13.

Sladek, R., G. Rocheleau, J. Rung, C. Dina, L. Shen, D. Serre, P. Boutin, D. Vincent, A. Belisle, S. Hadjadj, B. Balkau, B. Heude, G. Charpentier, T. J. Hudson, A. Montpetit, A. V. Pshezhetsky, M. Prentki, B. I. Posner, D. J. Balding, D. Meyre, C. Polychronakos and P. Froguel (2007). "A genome-wide association study identifies novel risk loci for type 2 diabetes." Nature **445**(7130): 881-5.

Stefan, N., J. C. Bunt, A. D. Salbe, T. Funahashi, Y. Matsuzawa and P. A. Tataranni (2002). "Plasma adiponectin concentrations in children: relationships with obesity and insulinemia." J Clin Endocrinol Metab **87**(10): 4652-6.

Steinthorsdottir, V., G. Thorleifsson, I. Reynisdottir, R. Benediktsson, T. Jonsdottir, G. B. Walters, U. Styrkarsdottir, S. Gretarsdottir, V. Emilsson, S. Ghosh, A. Baker, S. Snorradottir, H. Bjarnason, M. C. Ng, T. Hansen, Y. Bagger, R. L. Wilensky, M. P. Reilly, A. Adeyemo, Y. Chen, J. Zhou, V. Gudnason, G. Chen, H. Huang, K. Lashley, A. Doumatey, W. Y. So, R. C. Ma, G. Andersen, K. Borch-Johnsen, T. Jorgensen, J. V. van Vliet-Ostaptchouk, M. H. Hofker, C. Wijmenga, C. Christiansen, D. J. Rader, C. Rotimi, M. Gurney, J. C. Chan, O. Pedersen, G. Sigurdsson, J. R. Gulcher, U.

Thorsteinsdottir, A. Kong and K. Stefansson (2007). "A variant in CDKAL1 influences insulin response and risk of type 2 diabetes." Nat Genet **39**(6): 770-5.

Stoffers, D. A., J. Ferrer, W. L. Clarke and J. F. Habener (1997). "Early-onset type-II diabetes mellitus (MODY4) linked to IPF1." Nat Genet **17**(2): 138-9.

Storlien, L., N. D. Oakes and D. E. Kelley (2004). "Metabolic flexibility." Proc Nutr Soc **63**(2): 363-8.

Strom, T. M., K. Hortnagel, S. Hofmann, F. Gekeler, C. Scharfe, W. Rabl, K. D. Gerbitz and T. Meitinger (1998). "Diabetes insipidus, diabetes mellitus, optic atrophy and deafness (DIDMOAD) caused by mutations in a novel gene (wolframin) coding for a predicted transmembrane protein." Hum Mol Genet **7**(13): 2021-8.

Stumvoll, M., B. J. Goldstein and T. W. van Haeften (2005). "Type 2 diabetes: principles of pathogenesis and therapy." Lancet **365**(9467): 1333-46.

Tai, E. S., D. Corella, M. Deurenberg-Yap, X. Adiconis, S. K. Chew, C. E. Tan and J. M. Ordovas (2004). "Differential effects of the C1431T and Pro12Ala PPARgamma gene variants on plasma lipids and diabetes risk in an Asian population." J Lipid Res **45**(4): 674-85.

Tai, E. S., S. Demissie, L. A. Cupples, D. Corella, P. W. Wilson, E. J. Schaefer and J. M. Ordovas (2002). "Association between the PPARA L162V polymorphism and plasma lipid levels: the Framingham Offspring Study." Arterioscler Thromb Vasc Biol **22**(5): 805-10.

The DECODE study group (1999). "Glucose tolerance and mortality: comparison of WHO and American Diabetes Association diagnostic criteria. The DECODE study group. European Diabetes Epidemiology Group. Diabetes Epidemiology: Collaborative analysis Of Diagnostic criteria in Europe." Lancet **354**(9179): 617-21.

The Wellcome Trust Case Control Consortium (2007). "Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls." Nature **447**(7145): 661-78.

Thomas, H., K. Jaschkowitz, M. Bulman, T. M. Frayling, S. M. Mitchell, S. Roosen, A. Lingott-Frieg, C. J. Tack, S. Ellard, G. U. Ryffel and A. T. Hattersley (2001). "A distant upstream promoter of the HNF-4alpha gene connects the transcription factors involved in maturity-onset diabetes of the young." Hum Mol Genet **10**(19): 2089-97.

Tomas, E., T. S. Tsao, A. K. Saha, H. E. Murrey, C. Zhang Cc, S. I. Itani, H. F. Lodish and N. B. Ruderman (2002). "Enhanced muscle fat oxidation and glucose transport by ACRP30 globular domain: acetyl-CoA carboxylase inhibition and AMP-activated protein kinase activation." Proc Natl Acad Sci U S A **99**(25): 16309-13.

Vaxillaire, M., C. Dina, S. Lobbens, A. Dechaume, V. Vasseur-Delannoy, N. Helbecque, G. Charpentier and P. Froguel (2005). "Effect of common polymorphisms in the HNF4alpha promoter on susceptibility to type 2 diabetes in the French Caucasian population." Diabetologia **48**(3): 440-4.

Velho, G., P. Froguel, K. Clement, M. E. Pueyo, B. Rakotoambinina, H. Zouali, P. Passa, D. Cohen and J. J. Robert (1992). "Primary pancreatic beta-cell secretory defect caused by mutations in glucokinase gene in kindreds of maturity onset diabetes of the young." Lancet **340**(8817): 444-8.

Wang, H., W. Chu, X. Wang, Z. Zhang and S. C. Elbein (2005). "Evaluation of sequence variants in the pre-B cell leukemia transcription factor 1 gene: a positional and functional candidate for type 2 diabetes and impaired insulin secretion." Mol Genet Metab **86**(3): 384-91.

Wang, J., K. Chuang, M. Ahluwalia, S. Patel, N. Umblas, D. Mirel, R. Higuchi and S. Germer (2005). "High-throughput SNP genotyping by single-tube PCR with Tm-shift primers." Biotechniques **39**(6): 885-93.

Wang, X. L., R. M. McCredie and D. E. Wilcken (1996). "Common DNA polymorphisms at the lipoprotein lipase gene. Association with severity of coronary artery disease and diabetes." Circulation **93**(7): 1339-45.

Weissglas-Volkov, D., A. Huertas-Vazquez, E. Suviolahti, J. Lee, C. Plaisier, S. Canizales-Quinteros, T. Tusie-Luna, C. Aguilar-Salinas, M. R. Taskinen and P. Pajukanta (2006). "Common hepatic nuclear factor-4alpha variants are associated with high serum lipid levels and the metabolic syndrome." Diabetes **55**(7): 1970-7.

Weyer, C., C. Bogardus, D. M. Mott and R. E. Pratley (1999). "The natural history of insulin secretory dysfunction and insulin resistance in the pathogenesis of type 2 diabetes mellitus." J Clin Invest **104**(6): 787-94.

Wild, S., G. Roglic, A. Green, R. Sicree and H. King (2004). "Global prevalence of diabetes: estimates for the year 2000 and projections for 2030." Diabetes Care **27**(5): 1047-53.

Williams, M. A., C. Qiu, M. Muy-Rivera, S. Vadachkoria, T. Song and D. A. Luthy (2004). "Plasma adiponectin concentrations in early pregnancy and subsequent risk of gestational diabetes mellitus." J Clin Endocrinol Metab **89**(5): 2306-11.

Wilson, D. E., C. Q. Edwards and I. F. Chan (1983). "Phenotypic heterogeneity in the extended pedigree of a proband with lipoprotein lipase deficiency." Metabolism **32**(12): 1107-14.

Winckler, W., R. R. Graham, P. I. de Bakker, M. Sun, P. Almgren, T. Tuomi, D. Gaudet, T. J. Hudson, K. G. Ardlie, M. J. Daly, J. N. Hirschhorn, L. Groop and D. Altshuler (2005). "Association testing of variants in the hepatocyte nuclear factor 4alpha gene with risk of type 2 diabetes in 7,883 people." Diabetes **54**(3): 886-92.

Wittrup, H. H., B. G. Nordestgaard, R. Steffensen, G. Jensen and A. Tybjaerg-Hansen (2002). "Effect of gender on phenotypic expression of the S447X mutation in LPL: the Copenhagen City Heart Study." Atherosclerosis **165**(1): 119-26.

World Health Organization (2000). "International Obesity Task Force: The Asia-Pacific perspective: redefining obesity and its treatment."

Wu, Z., P. Puigserver, U. Andersson, C. Zhang, G. Adelmant, V. Mootha, A. Troy, S. Cinti, B. Lowell, R. C. Scarpulla and B. M. Spiegelman (1999). "Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC-1." Cell **98**(1): 115-24.

Yabu, Y., K. Noma, K. Nakatani, J. Nishioka, M. Suematsu, A. Katsuki, Y. Hori, Y. Yano, Y. Sumida, H. Wada and T. Nobori (2005). "C-514T polymorphism in hepatic lipase gene promoter is associated with elevated triglyceride levels and decreasing insulin sensitivity in nondiabetic Japanese subjects." Int J Mol Med **16**(3): 421-5.

Yamakawa-Kobayashi, K., H. Ishiguro, T. Arinami, R. Miyazaki and H. Hamaguchi (2002). "A Val227Ala polymorphism in the peroxisome proliferator activated receptor alpha (PPARalpha) gene is associated with variations in serum lipid levels." J Med Genet **39**(3): 189-91.

Yamauchi, T., J. Kamon, Y. Minokoshi, Y. Ito, H. Waki, S. Uchida, S. Yamashita, M. Noda, S. Kita, K. Ueki, K. Eto, Y. Akanuma, P. Froguel, F. Foufelle, P. Ferre, D. Carling, S. Kimura, R. Nagai, B. B. Kahn and T. Kadowaki (2002). "Adiponectin stimulates glucose utilization and fatty-acid oxidation by activating AMP-activated protein kinase." Nat Med **8**(11): 1288-95.

Yang, Q., T. E. Graham, N. Mody, F. Preitner, O. D. Peroni, J. M. Zabolotny, K. Kotani, L. Quadro and B. B. Kahn (2005). "Serum retinol binding protein 4 contributes to insulin resistance in obesity and type 2 diabetes." Nature **436**(7049): 356-62.

Yang, X., W. Y. So, A. P. Kong, C. S. Ho, C. W. Lam, R. J. Stevens, R. R. Lyu, D. D. Yin, C. S. Cockram, P. C. Tong, V. Wong and J. C. Chan (2007). "Development and validation of stroke risk equation for Hong Kong Chinese patients with type 2 diabetes: the Hong Kong Diabetes Registry." Diabetes Care **30**(1): 65-70.

Yokoi, N., M. Kanamori, Y. Horikawa, J. Takeda, T. Sanke, H. Furuta, K. Nanjo, H. Mori, M. Kasuga, K. Hara, T. Kadowaki, Y. Tanizawa, Y. Oka, Y. Iwami, H. Ohgawara, Y. Yamada, Y. Seino, H. Yano, N. J. Cox and S. Seino (2006). "Association studies of variants in the genes involved in pancreatic beta-cell function in type 2 diabetes in Japanese subjects." Diabetes **55**(8): 2379-86.

Yoon, J. C., P. Puigserver, G. Chen, J. Donovan, Z. Wu, J. Rhee, G. Adelmant, J. Stafford, C. R. Kahn, D. K. Granner, C. B. Newgard and B. M. Spiegelman (2001). "Control of hepatic gluconeogenesis through the transcriptional coactivator PGC-1." Nature **413**(6852): 131-8.

Zambon, A., S. S. Deeb, J. E. Hokanson, B. G. Brown and J. D. Brunzell (1998). "Common variants in the promoter of the hepatic lipase gene are associated with lower levels of hepatic lipase activity, buoyant LDL, and higher HDL2 cholesterol." Arterioscler Thromb Vasc Biol **18**(11): 1723-9.

Zeggini, E., L. J. Scott, R. Saxena, B. F. Voight, J. L. Marchini, T. Hu, P. I. de Bakker, G. R. Abecasis, P. Almgren, G. Andersen, K. Ardlie, K. B. Bostrom, R. N. Bergman, L. L. Bonnycastle, K. Borch-Johnsen, N. P. Burtt, H. Chen, P. S. Chines, M. J. Daly,

P. Deodhar, C. J. Ding, A. S. Doney, W. L. Duren, K. S. Elliott, M. R. Erdos, T. M. Frayling, R. M. Freathy, L. Gianniny, H. Grallert, N. Grarup, C. J. Groves, C. Guiducci, T. Hansen, C. Herder, G. A. Hitman, T. E. Hughes, B. Isomaa, A. U. Jackson, T. Jorgensen, A. Kong, K. Kubalanza, F. G. Kuruvilla, J. Kuusisto, C. Langenberg, H. Lango, T. Lauritzen, Y. Li, C. M. Lindgren, V. Lyssenko, A. F. Marvelle, C. Meisinger, K. Midthjell, K. L. Mohlke, M. A. Morken, A. D. Morris, N. Narisu, P. Nilsson, K. R. Owen, C. N. Palmer, F. Payne, J. R. Perry, E. Pettersen, C. Platou, I. Prokopenko, L. Qi, L. Qin, N. W. Rayner, M. Rees, J. J. Roix, A. Sandbaek, B. Shields, M. Sjogren, V. Steinthorsdottir, H. M. Stringham, A. J. Swift, G. Thorleifsson, U. Thorsteinsdottir, N. J. Timpson, T. Tuomi, J. Tuomilehto, M. Walker, R. M. Watanabe, M. N. Weedon, C. J. Willer, T. Illig, K. Hveem, F. B. Hu, M. Laakso, K. Stefansson, O. Pedersen, N. J. Wareham, I. Barroso, A. T. Hattersley, F. S. Collins, L. Groop, M. I. McCarthy, M. Boehnke and D. Altshuler (2008). "Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes." Nat Genet **40**(5): 638-45.

Zeggini, E., M. N. Weedon, C. M. Lindgren, T. M. Frayling, K. S. Elliott, H. Lango, N. J. Timpson, J. R. Perry, N. W. Rayner, R. M. Freathy, J. C. Barrett, B. Shields, A. P. Morris, S. Ellard, C. J. Groves, L. W. Harries, J. L. Marchini, K. R. Owen, B. Knight, L. R. Cardon, M. Walker, G. A. Hitman, A. D. Morris, A. S. Doney, M. I. McCarthy and A. T. Hattersley (2007). "Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes." Science **316**(5829): 1336-41.

Zimmet, P., K. G. Alberti and J. Shaw (2001). "Global and societal implications of the diabetes epidemic." Nature **414**(6865): 782-7.

APPENDICES

Appendix 1: Gene structure and linkage disequilibrium of genotyped SNPs of candidate genes

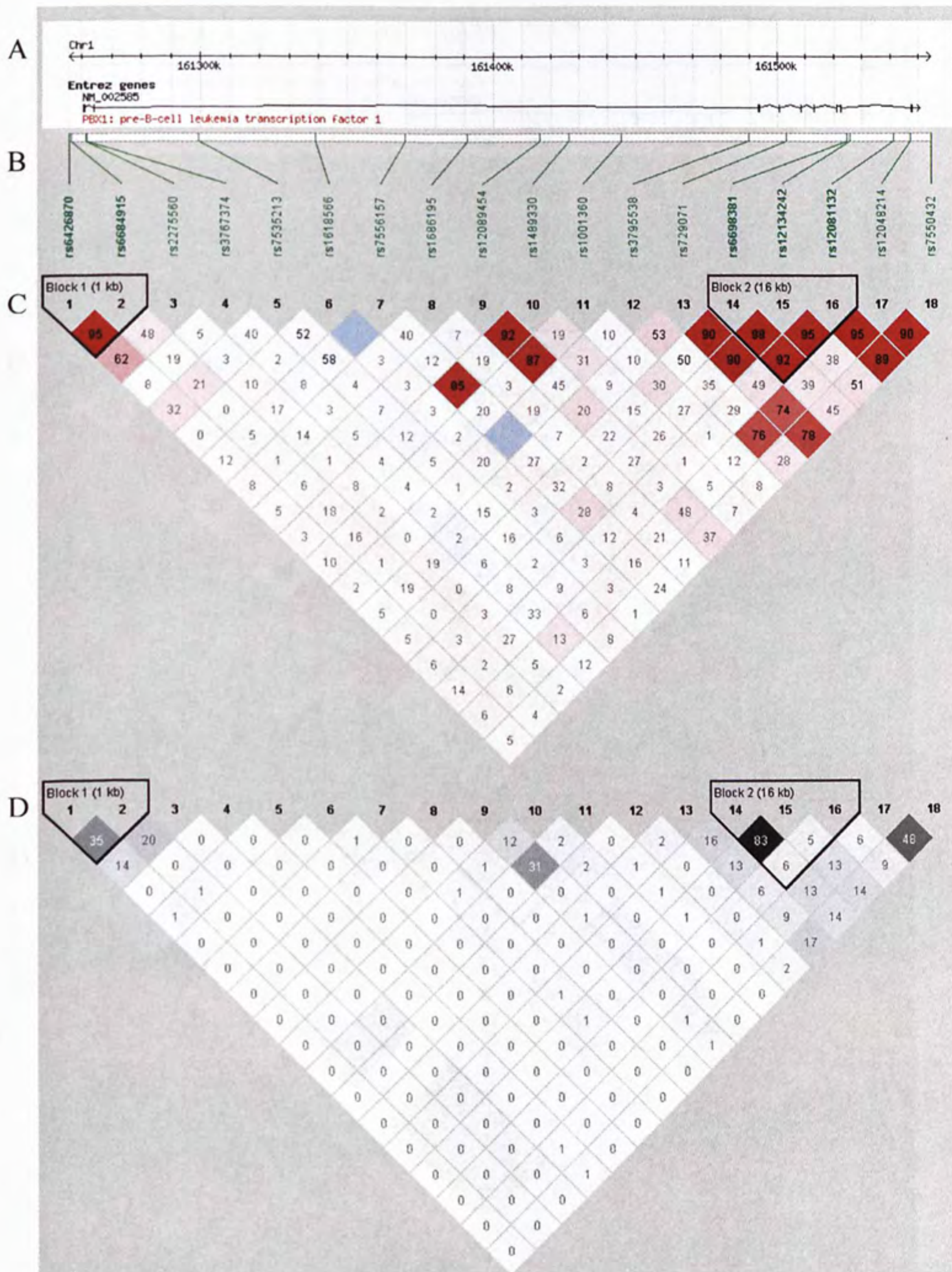


Figure A1.1: *PBX1*. A) Gene structure; B) Analyzed SNPs; C) D' plot; D) r^2 plot.

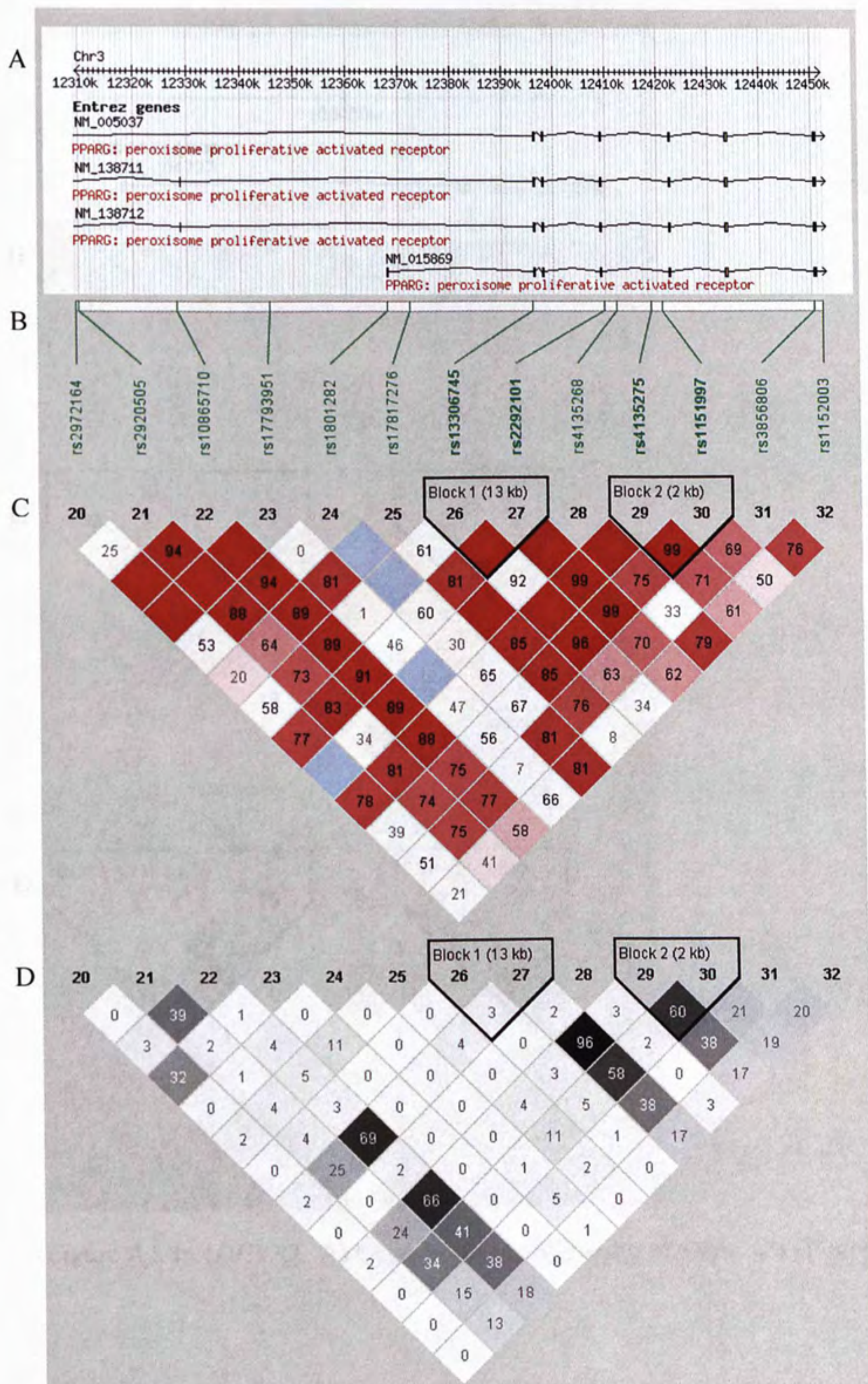


Figure A1.2: *PPARG*. A) Gene structure; B) Analyzed SNPs; C) D' plot; D) r^2 plot.

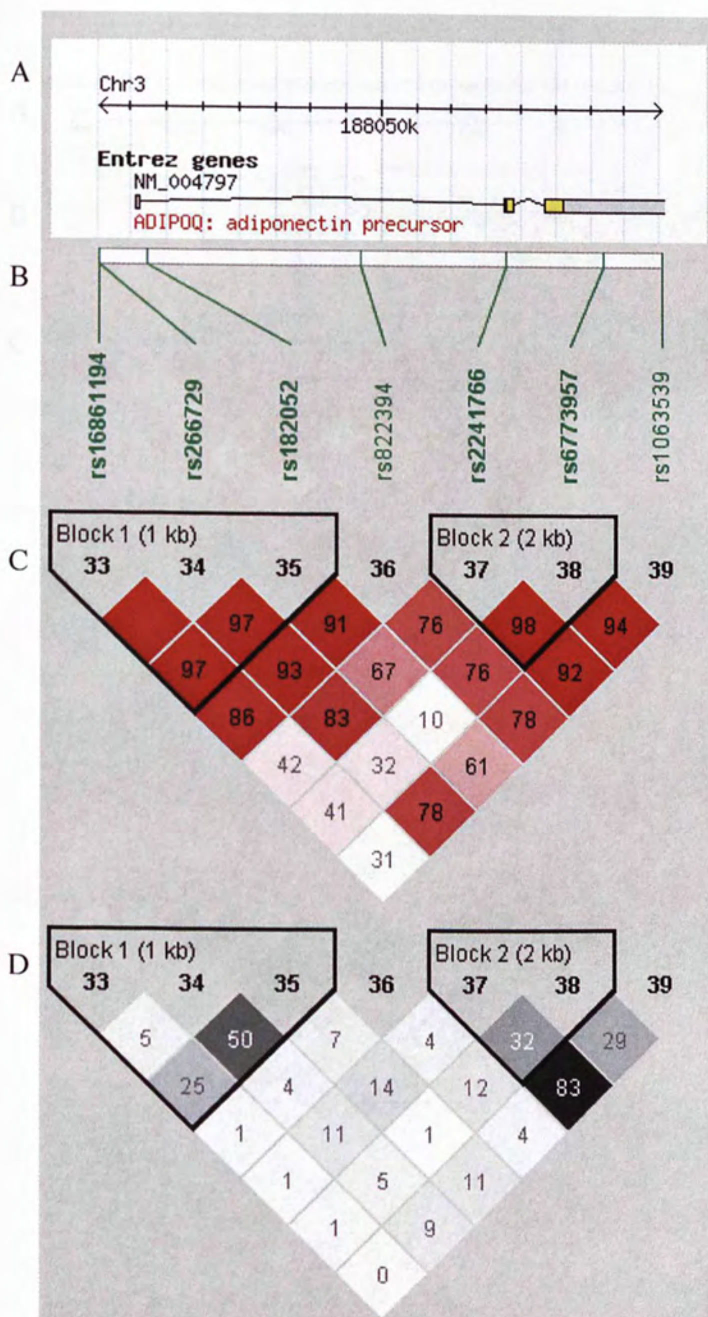


Figure A1.3: *ADIPOQ*. A) Gene structure; B) Analyzed SNPs; C) D' plot; D) r^2 plot.

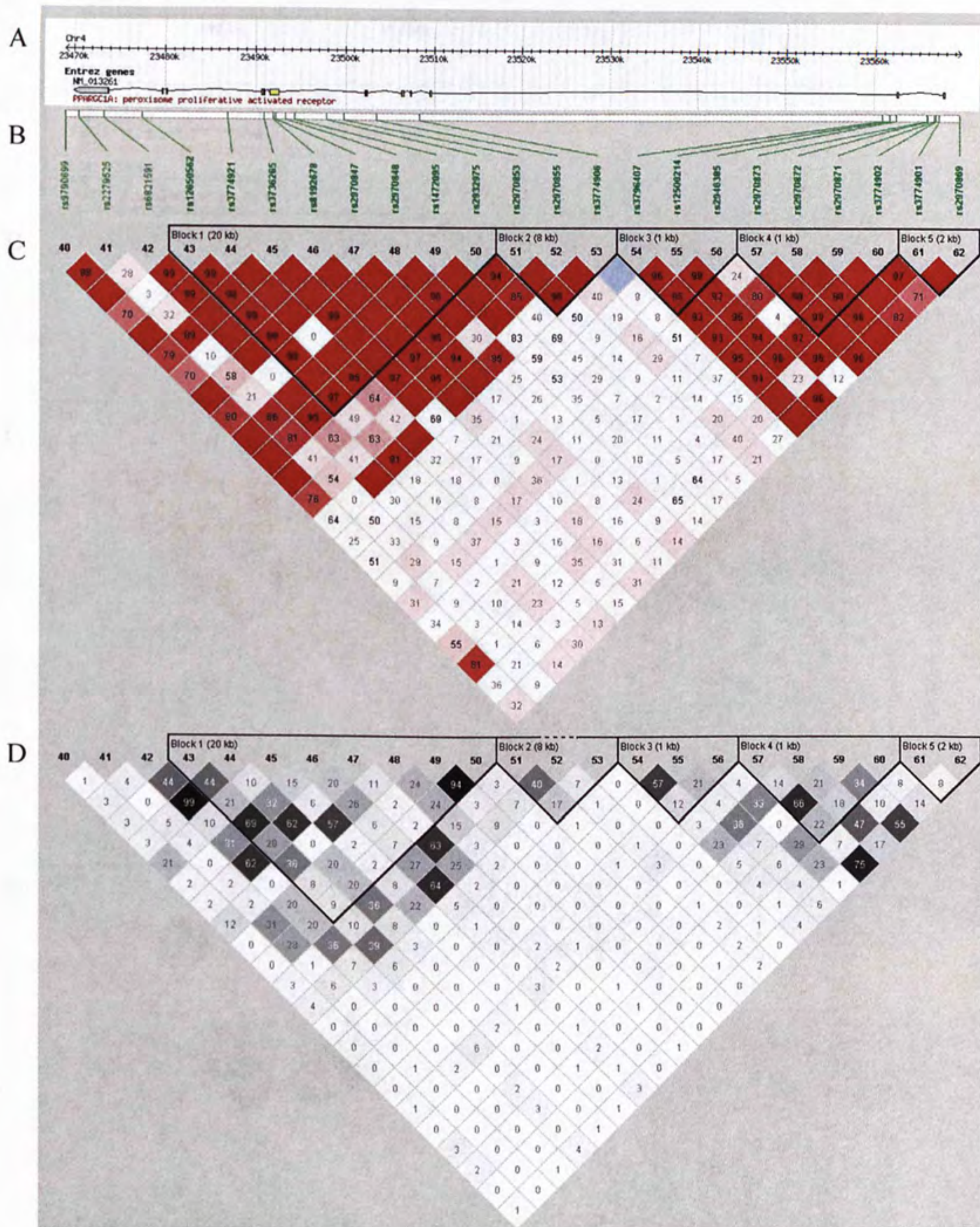


Figure A1.4: *PPARGC1A*. A) Gene structure; B) Analyzed SNPs; C) D' plot; D) r^2 plot.

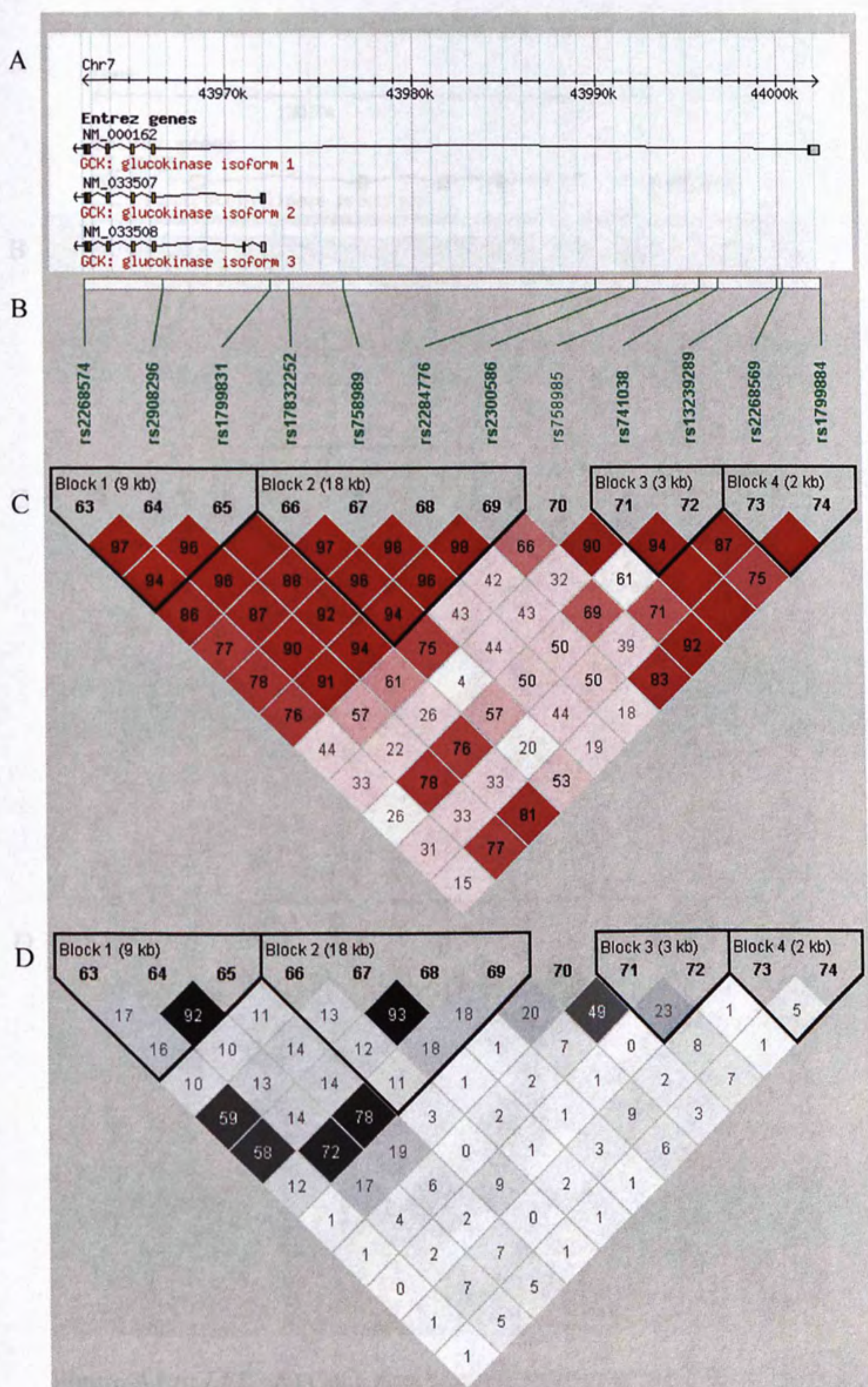


Figure A1.5: *GCK*. A) Gene structure; B) Analyzed SNPs; C) D' plot; D) r^2 plot.

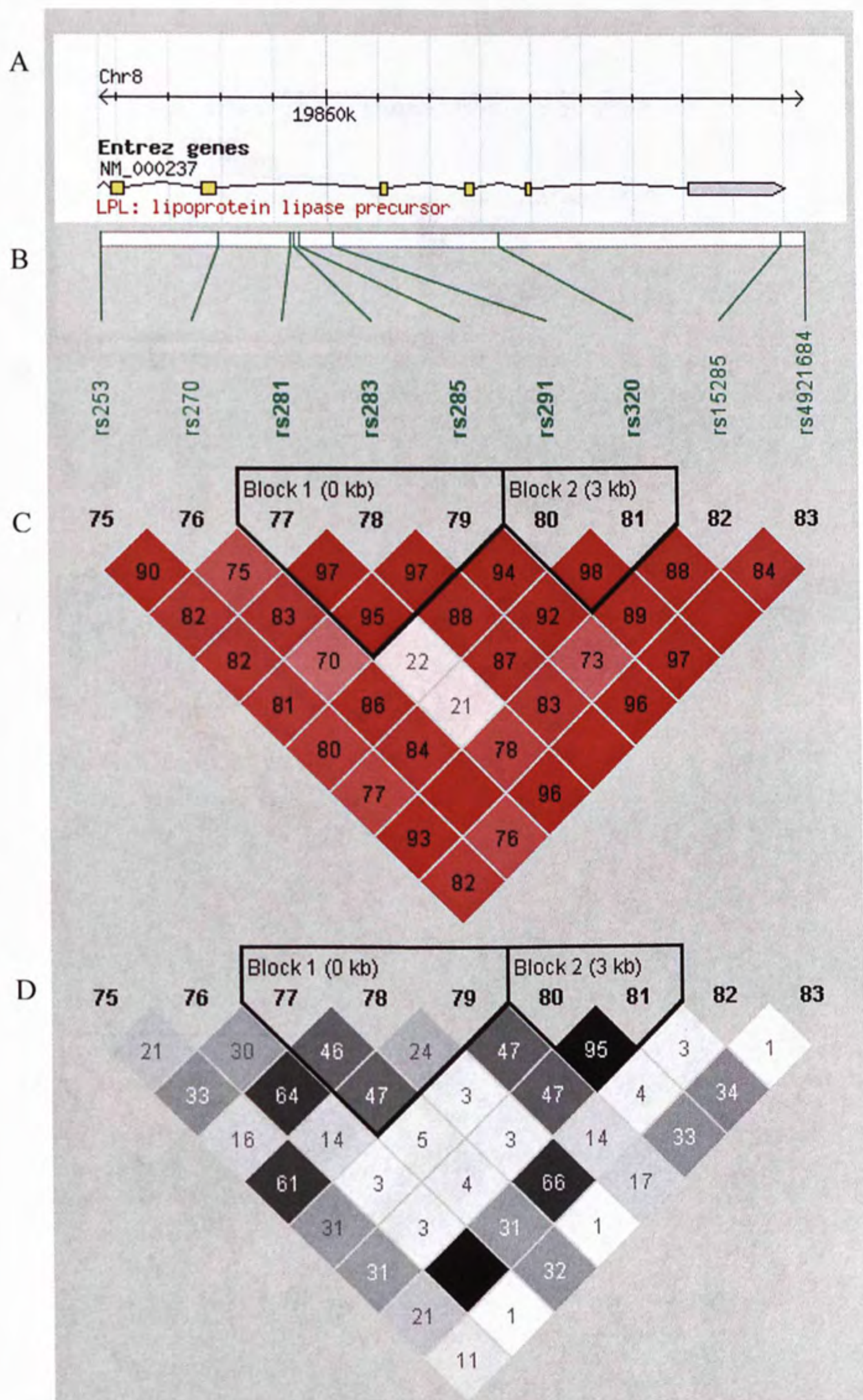


Figure A1.6: *LPL*. A) Gene structure; B) Analyzed SNPs; C) D' plot; D) r^2 plot.

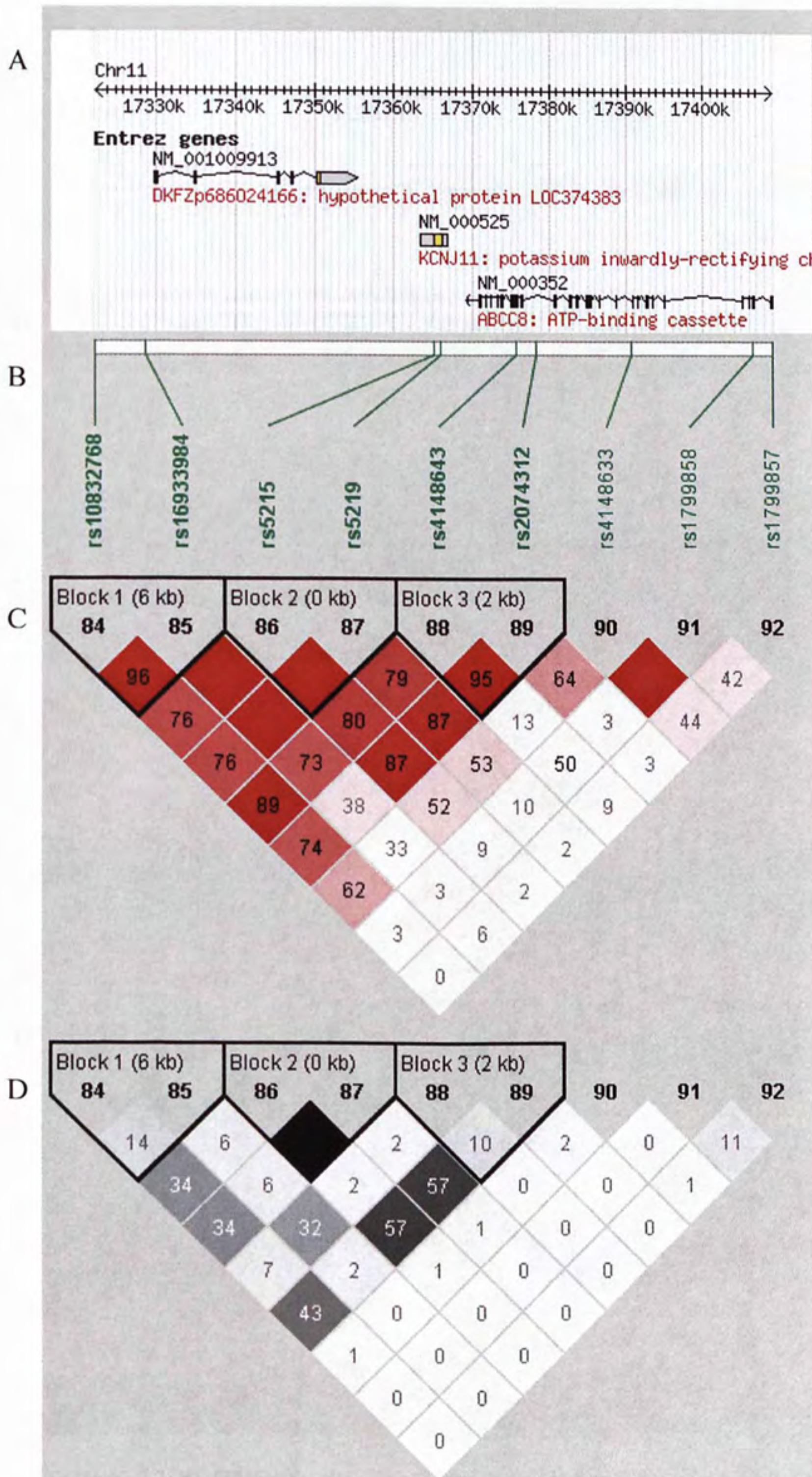


Figure A1.7: *KCNJ11/ABCC8*. A) Gene structure; B) Analyzed SNPs; C) D' plot; D) r^2 plot.

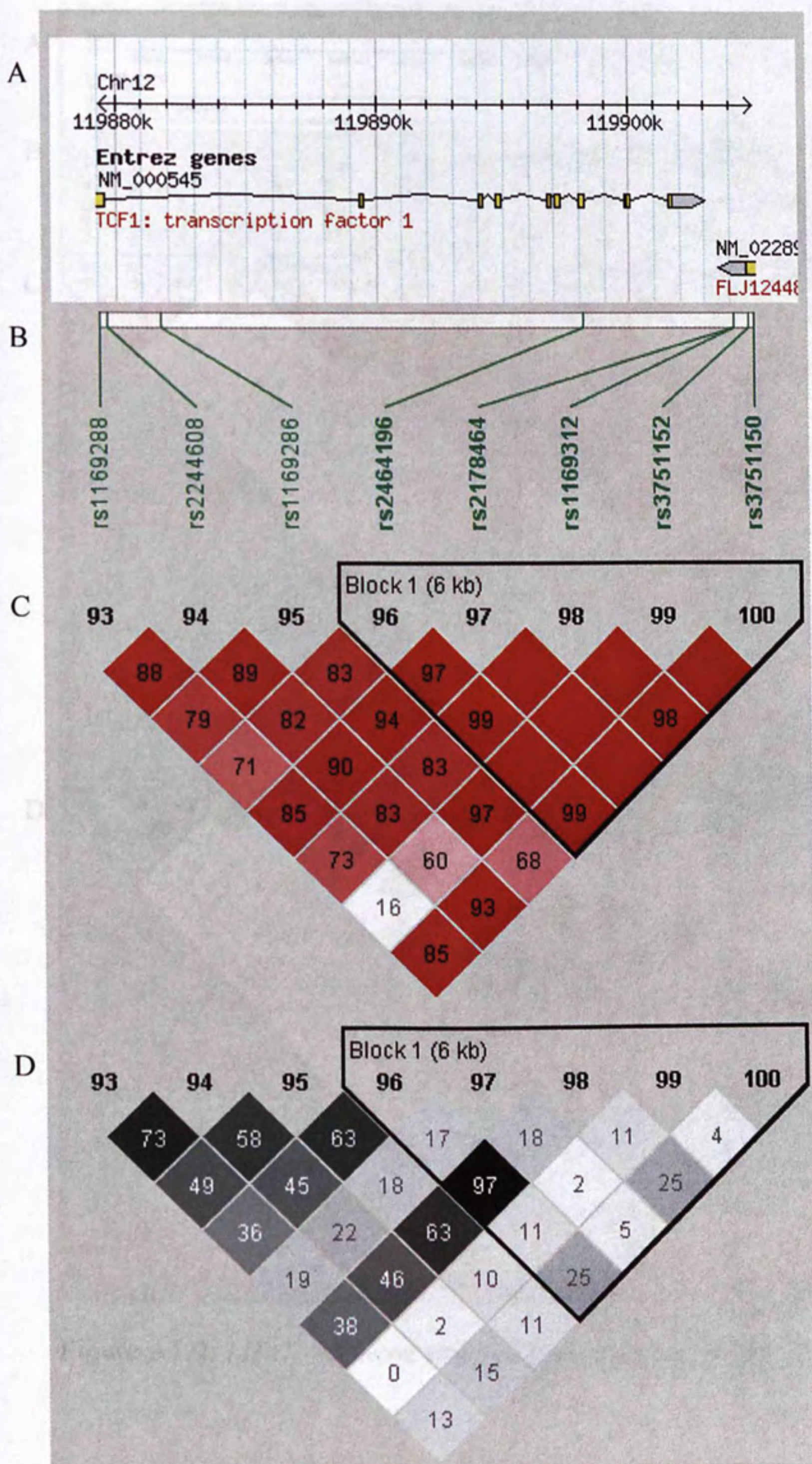


Figure A1.8: *HNF1A*. A) Gene structure; B) Analyzed SNPs; C) D' plot; D) r^2 plot.

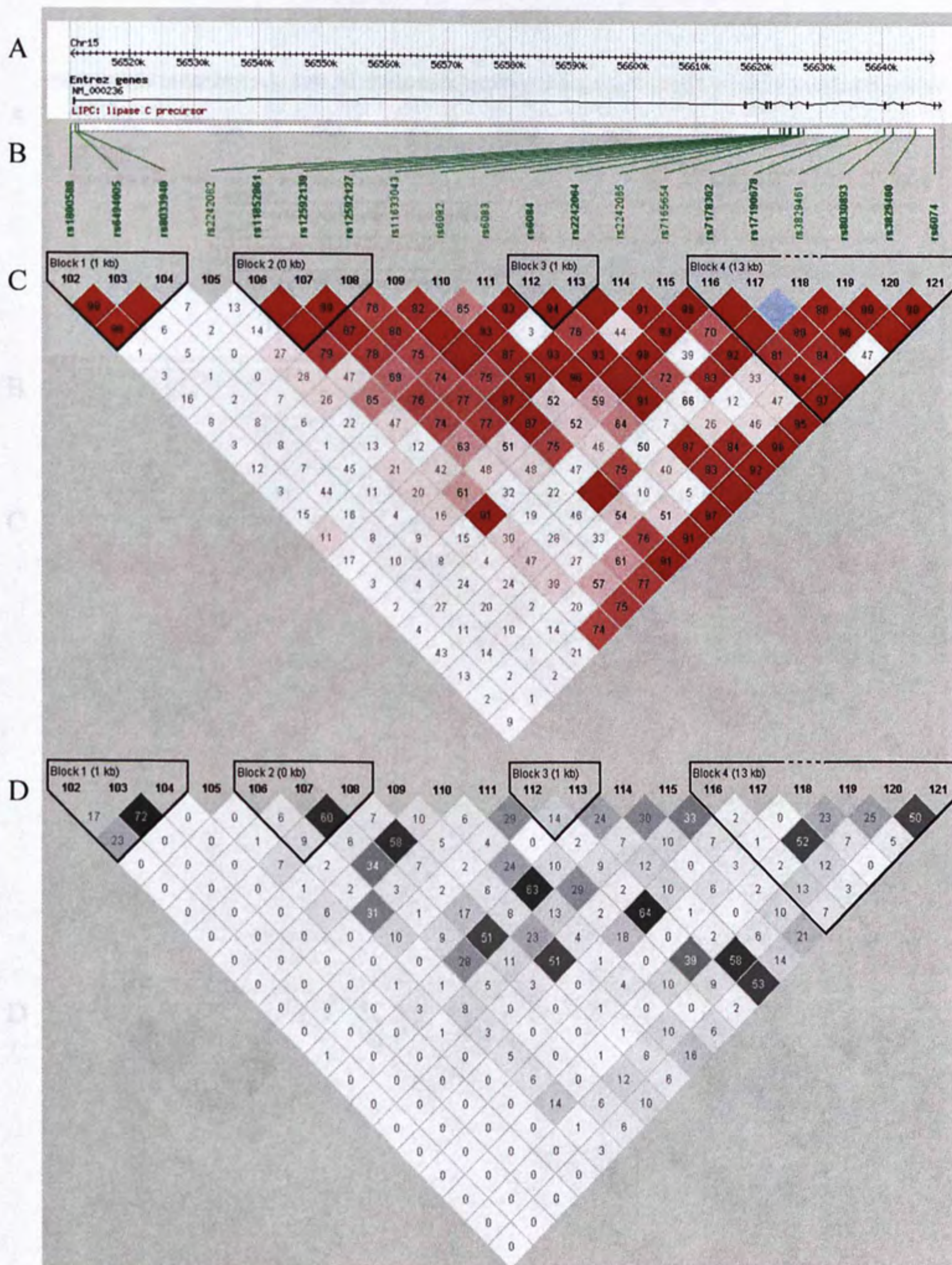


Figure A1.9: *LIPC*. A) Gene structure; B) Analyzed SNPs; C) D' plot; D) r^2 plot.

A

B

C

D

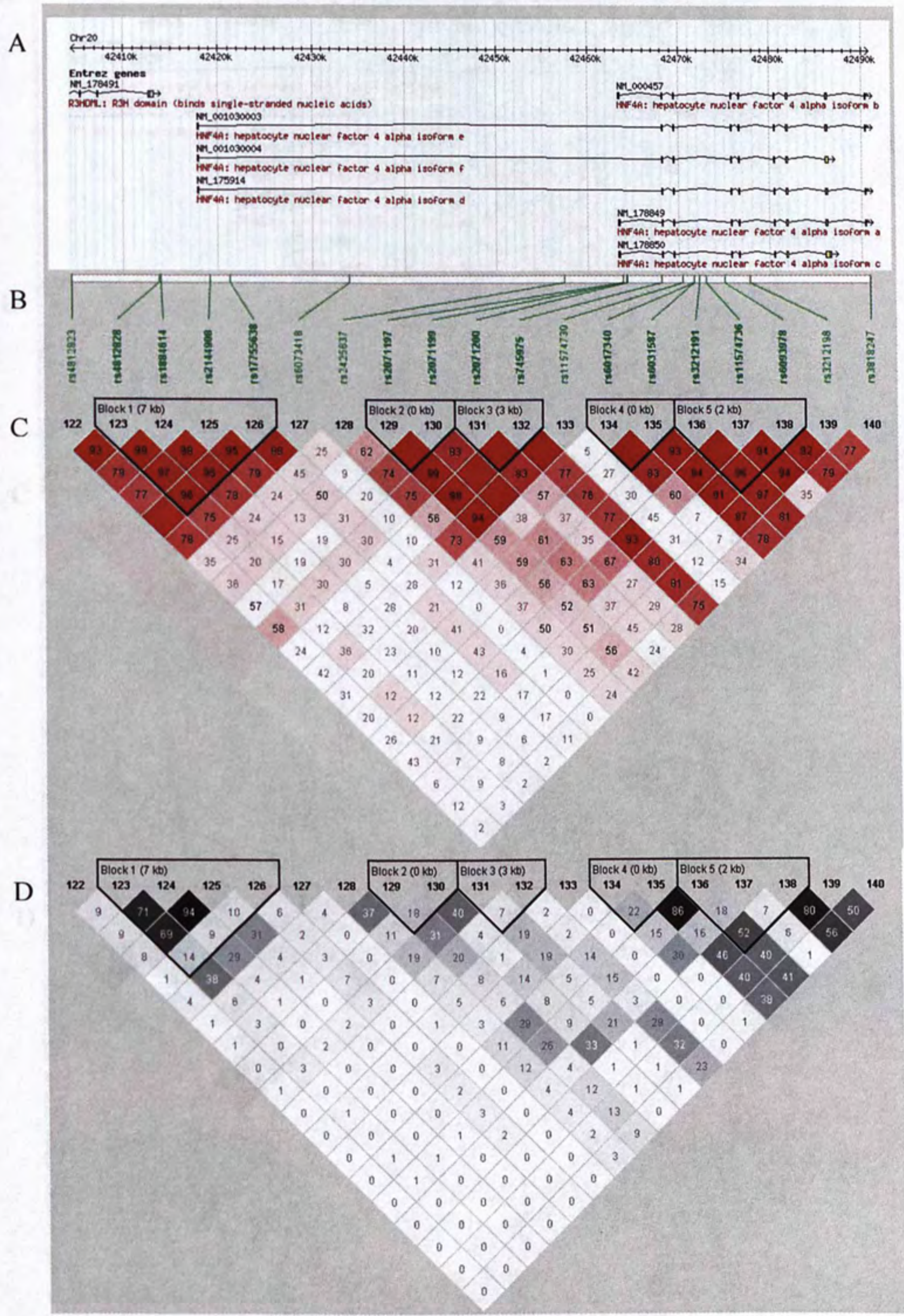


Figure A1.10: *HNF4A*. A) Gene structure; B) Analyzed SNPs; C) D' plot; D) r^2 plot.

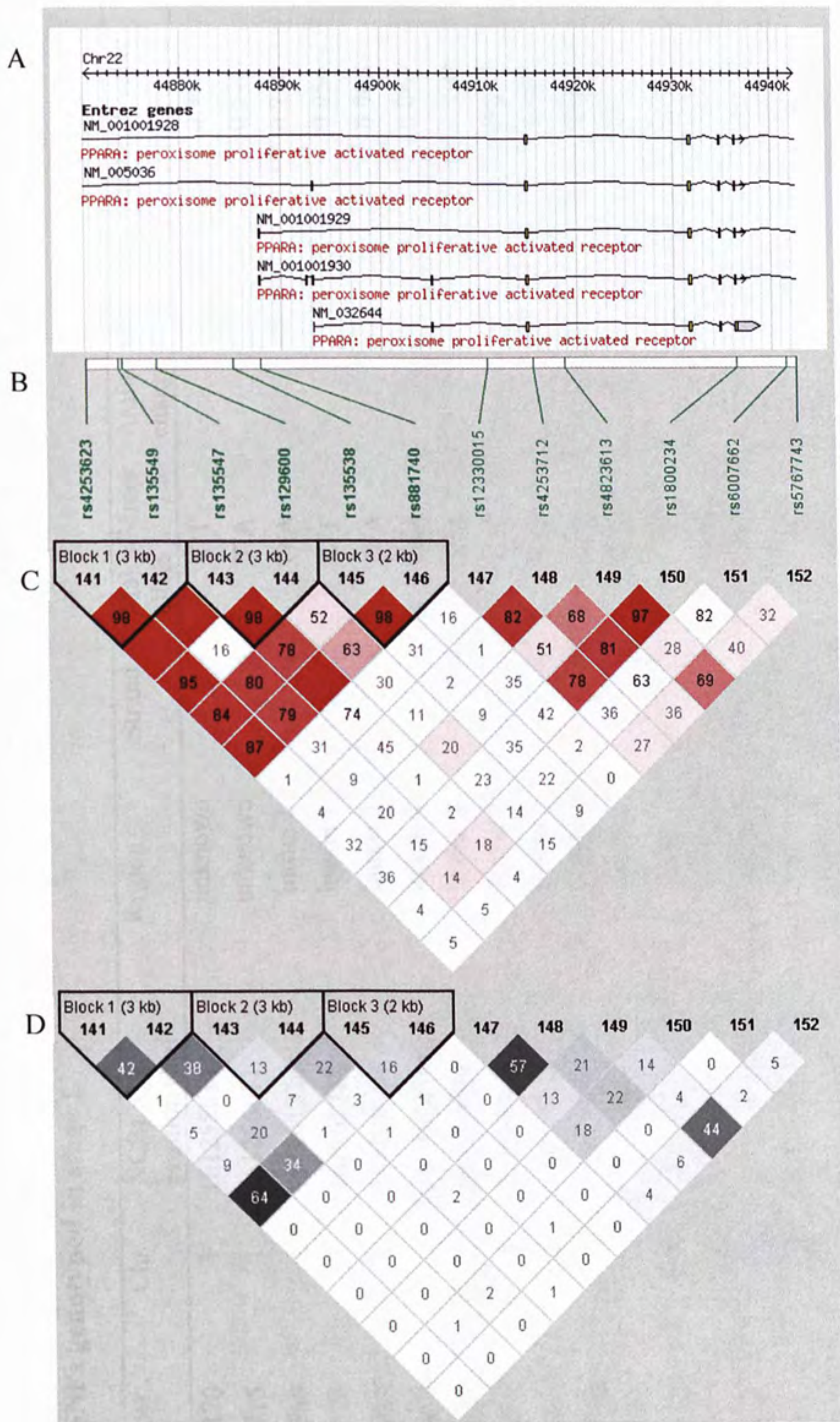


Figure A1.11: *PPARA*. A) Gene structure; B) Analyzed SNPs; C) D' plot; D) r^2 plot.

Appendix 2: Information of SNPs genotyped in stage 1

SNP no.	Gene	rs number	Chr	NCBI b35 position (bp)	Region	Strand	Major/Minor allele	Amino acid change	HWE *P* value
1	*PBX1*	rs6426870	1	161255611	unknown	+	C/T		0.961
2	*PBX1*	rs6684915	1	161256611	unknown	+	T/A		0.575
3	*PBX1*	rs2275560	1	161261324	intron 1	+	G/A		0.230
4	*PBX1*	rs3767374	1	161261788	intron 1	-	C/T		0.954
5	*PBX1*	rs7535213	1	161299384	intron 2	+	G/A		0.908
6	*PBX1*	*rs1618566	1	161339004	intron 2	+	G/A		0.043
7	*PBX1*	rs7556157	1	161369628	intron 2	+	G/A		1.000
8	*PBX1*	rs1686195	1	161391248	intron 2	-	C/T		0.285
9	*PBX1*	rs12089454	1	161416516	intron 2	+	C/T		0.451
10	*PBX1*	rs1489330	1	161426748	intron 2	+	G/T		0.603
11	*PBX1*	rs1001360	1	161444976	intron 2	+	A/G		0.509
12	*PBX1*	rs3795538	1	161489811	intron 2	+	T/C		0.042
13	*PBX1*	rs729071	1	161503359	intron 4	-	A/C		0.466
14	*PBX1*	rs6698381	1	161524836	intron 8	+	T/C		0.846
15	*PBX1*	rs12134242	1	161525911	intron 8	+	T/A		0.629
16	*PBX1*	rs12081132	1	161541345	intron 8	+	A/G		0.036
17	*PBX1*	rs12048214	1	161547125	intron 8	+	C/T		0.704
18	*PBX1*	rs7550432	1	161554494	unknown	+	T/C		1.000
19	*NEUROD1*	*rs1801262	2	182368961	exon1	-	G/A	A45T	0.789
20	*PPARG*	rs2972164	3	12309416	intron A1	+	C/T		1.000
21	*PPARG*	rs2920505	3	12310081	intron A1	-	C/T		0.376

22	*PPARG*	rs10865710	3	12328198	intron A1	+	C/G		0.622
23	*PPARG*	rs17793951	3	12345737	intron A2	+	A/G		1.000
24	*PPARG*	*rs1801282	3	12368125	exon B	+	C/G	P12A	1.000
25	*PPARG*	rs17817276	3	12372392	intron B	+	A/G		1.000
26	*PPARG*	rs13306745	3	12396056	intron B	+	G/T		0.074
27	*PPARG*	rs2292101	3	12409901	intron 3	+	C/T		0.364
28	*PPARG*	rs4135268	3	12412237	intron 3	+	C/G		1.000
29	*PPARG*	rs4135275	3	12418844	intron 3	+	C/G		0.455
30	*PPARG*	rs1151997	3	12421004	intron 3	-	C/T		0.382
31	*PPARG*	rs3856806	3	12450557	exon 6	+	C/T	H447H	0.779
32	*PPARG*	rs1152003	3	12452055	unknown	+	G/C		0.394
33	*ADIPOQ*	*rs16861194	3	188042127	locus	+	A/G		0.867
34	*ADIPOQ*	*rs266729	3	188042176	locus	+	C/G		0.895
35	*ADIPOQ*	rs182052	3	188043484	intron 1	+	G/A		0.924
36	*ADIPOQ*	rs822394	3	188049430	intron 1	+	C/A		1.000
37	*ADIPOQ*	*rs2241766	3	188053594	exon 2	+	T/G	G15G	0.177
38	*ADIPOQ*	rs6773957	3	188056407	exon 3	+	A/G		0.482
39	*ADIPOQ*	rs1063539	3	188058094	exon 3	+	G/C		0.179
40	*PPARGC1A*	rs9790699	4	23469037	unknown	+	C/T		1.000
41	*PPARGC1A*	rs2279525	4	23470521	untranslated	+	T/C		0.320
42	*PPARGC1A*	rs6821591	4	23473269	exon 1	+	T/C		0.102
43	*PPARGC1A*	rs12650562	4	23477456	intron 1	+	T/C		0.534
44	*PPARGC1A*	rs3774921	4	23486916	intron 3	-	A/G		0.115
45	*PPARGC1A*	rs3736265	4	23490976	exon 5	+	G/A	M612T	0.457
46	*PPARGC1A*	*rs8192678	4	23491931	exon 6	-	G/A	S482G	0.800

47	*PPARGC1A*	*rs2970847	4	23492193	exon 6	+	C/T	T394T	0.600
48	*PPARGC1A*	rs2970848	4	23493296	intron 6	+	A/G		0.491
49	*PPARGC1A*	rs1472095	4	23494346	intron 6	+	C/T		0.204
50	*PPARGC1A*	rs2932975	4	23497878	intron 6	+	G/A		0.266
51	*PPARGC1A*	rs2970853	4	23499788	intron 6	+	G/A		1.000
52	*PPARGC1A*	rs2970855	4	23503514	intron 8	+	A/T		0.638
53	*PPARGC1A*	rs3774906	4	23508353	intron 10	-	C/A		0.282
54	*PPARGC1A*	rs3796407	4	23560980	intron 11	+	G/A		0.495
55	*PPARGC1A*	rs12500214	4	23561837	intron 11	+	G/A		0.019
56	*PPARGC1A*	rs2946385	4	23562592	intron 11	+	G/T		0.732
57	*PPARGC1A*	rs2970873	4	23566050	intron 12	-	C/T		0.472
58	*PPARGC1A*	rs2970872	4	23566082	intron 12	-	T/C		0.503
59	*PPARGC1A*	rs2970871	4	23566851	intron 12	-	G/A		1.000
60	*PPARGC1A*	rs3774902	4	23567051	intron 12	-	C/T		0.650
61	*PPARGC1A*	rs3774901	4	23567363	intron 12	-	G/A		0.892
62	*PPARGC1A*	rs2970869	4	23569638	locus	-	G/A		0.274
63	*GCK*	rs2268574	7	43962561	intron 4	-	C/T		0.290
64	*GCK*	rs2908296	7	43966702	intron 9	-	C/A		0.469
65	*GCK*	*rs1799831	7	43972382	intron 9	-	G/A		0.775
66	*GCK*	rs17832252	7	43973339	intron 9	+	G/T		1.000
67	*GCK*	rs758989	7	43976246	intron 9	-	A/G		0.084
68	*GCK*	rs2284776	7	43989994	intron 9	-	A/G		0.171
69	*GCK*	rs2300586	7	43992096	intron 9	-	C/T		0.405
70	*GCK*	rs758985	7	43995711	intron 9	-	C/T		0.170
71	*GCK*	rs741038	7	43996719	intron 9	-	T/C		0.194

72	*GCK*	rs13239289	7	43999928	intron 9	+	C/G		0.096
73	*GCK*	rs2268569	7	44000260	intron 9	-	G/A		0.513
74	*GCK*	*rs1799884	7	44002308	promoter	-	G/A		0.819
75	*LPL*	rs253	8	19855697	intron 4	+	C/T		0.690
76	*LPL*	rs270	8	19857956	intron 6	+	C/A		0.172
77	*LPL*	rs281	8	19859303	intron 6	-	T/A		0.599
78	*LPL*	rs283	8	19859378	intron 6	-	G/A		0.777
79	*LPL*	rs285	8	19859469	intron 6	+	C/T		1.000
80	*LPL*	rs291	8	19860132	intron 6	+	T/C		0.900
81	*LPL*	*rs320	8	19863357	intron 8	+	T/G	T495G	0.937
82	*LPL*	rs15285	8	19868947	exon 10	-	G/A		0.861
83	*LPL*	rs4921684	8	19869408	locus	+	C/T		0.967
84	*KCNJ11*	rs10832768	11	17322464	unknown	+	G/T		0.098
85	*KCNJ11*	rs16933984	11	17328924	locus	+	T/G		0.745
86	*KCNJ11*	*rs5215	11	17365206	exon 1	+	T/C	I337V	0.184
87	*KCNJ11*	*rs5219	11	17366148	exon 1	+	C/T	E23K	0.210
88	*ABCC8*	*rs4148643	11	17375855	exon 9	-	G/A	L829L	0.302
89	*ABCC8*	rs2074312	11	17378366	intron 10	-	T/C		0.195
90	*ABCC8*	rs4148633	11	17390860	exon 19	-	C/T	L829L	1.000
91	*ABCC8*	*rs1799858	11	17406505	exon 25	-	G/A	K649K	0.716
92	*ABCC8*	rs1799857	11	17409068	exon 27	-	C/T	H562H	0.851
93	*HNF1A*	*rs1169288	12	119879370	exon 1	-	T/G	L27I	0.052
94	*HNF1A*	rs2244608	12	119879708	intron 1	-	T/C		0.436
95	*HNF1A*	rs1169286	12	119881776	intron 1	-	A/G		0.517
96	*HNF1A*	rs2464196	12	119898147	exon 7	-	C/T	N487S	0.719

97	*HNF1A*	rs2178464	12	119904160	untranslated	+	T/C		0.924
98	*HNF1A*	rs1169312	12	119904181	untranslated	+	G/T		0.719
99	*HNF1A*	rs3751152	12	119904784	unknown	-	C/T	N227N	0.116
100	*HNF1A*	rs3751150	12	119904934	unknown	-	C/T	I117I	1.000
101	*IPF1*	rs7982864	13	27398496	unknown	+	C/T		0.756
102	*LIPC*	*rs1800588	15	56510967	locus	+	C/T		0.214
103	*LIPC*	rs6494005	15	56511816	intron 1	+	A/G		0.042
104	*LIPC*	rs8033940	15	56512134	intron 1	+	A/G		0.012
105	*LIPC*	rs2242062	15	56621683	intron 3	+	A/G		0.344
106	*LIPC*	rs11852861	15	56623600	intron 4	+	C/T		0.691
107	*LIPC*	rs12592139	15	56624230	intron 4	+	G/A		0.593
108	*LIPC*	rs12592127	15	56624381	intron 4	+	A/G		0.405
109	*LIPC*	rs11633043	15	56625014	intron 4	+	G/A		0.931
110	*LIPC*	rs6082	15	56625249	exon 5	+	A/G	G197G	1.000
111	*LIPC*	rs6083	15	56625302	exon 5	+	G/A	S215N	1.000
112	*LIPC*	rs6084	15	56625330	exon 5	+	C/G	T224T	0.565
113	*LIPC*	rs2242064	15	56626564	intron 5	+	T/G		0.180
114	*LIPC*	rs2242065	15	56626590	intron 5	+	C/T		1.000
115	*LIPC*	rs7165654	15	56627331	intron 5	+	A/G		0.095
116	*LIPC*	rs7178362	15	56634555	intron 6	+	T/C		0.808
117	*LIPC*	rs17190678	15	56634664	intron 6	+	C/G		0.073
118	*LIPC*	rs3829461	15	56640401	exon 7	-	C/T	T336T	0.029
119	*LIPC*	rs8030893	15	56641709	intron 7	+	G/C		0.076
120	*LIPC*	rs3829460	15	56645237	intron 8	-	A/T		0.596
121	*LIPC*	rs6074	15	56648255	exon 9	+	C/A	T479T	0.478

122	*HNF4A*	rs4812823	20	42404514	locus	+	T/C	0.291
123	*HNF4A*	rs4812828	20	42413734	locus	+	C/T	0.789
124	*HNF4A*	*rs1884614	20	42413933	promoter	+	C/T	0.110
125	*HNF4A*	*rs2144908	20	42419131	intron 1	+	G/A	0.091
126	*HNF4A*	rs17755638	20	42421255	intron 1	+	T/C	0.222
127	*HNF4A*	*rs6073418	20	42434004	intron 1	+	C/T	0.047
128	*HNF4A*	rs2425637	20	42457463	intron 1	+	T/G	0.600
129	*HNF4A*	rs2071197	20	42463849	intron 1	+	G/A	0.346
130	*HNF4A*	rs2071199	20	42464280	intron 1	+	T/C	0.579
131	*HNF4A*	rs2071200	20	42464370	intron 1	+	C/T	0.753
132	*HNF4A*	rs745975	20	42468107	intron 1	-	G/A	0.503
133	*HNF4A*	rs11574730	20	42470427	intron 3	+	G/A	0.94
134	*HNF4A*	rs6017340	20	42471622	intron 3	+	C/T	0.811
135	*HNF4A*	rs6031587	20	42471663	intron 3	+	T/C	0.764
136	*HNF4A*	rs3212191	20	42472196	intron 3	+	C/T	0.888
137	*HNF4A*	rs11574736	20	42472951	intron 3	-	C/G	0.873
138	*HNF4A*	rs6093978	20	42475005	intron 3	+	C/T	1.000
139	*HNF4A*	rs3212198	20	42477776	intron 5	+	C/T	0.977
140	*HNF4A*	*rs3818247	20	42490894	intron 9	+	T/G	0.917
141	*PPARA*	rs4253623	22	44870625	intron 2	+	A/G	0.781
142	*PPARA*	rs135549	22	44873827	intron 2	-	A/G	0.198
143	*PPARA*	rs135547	22	44874169	intron 2	-	G/C	0.778
144	*PPARA*	rs129600	22	44877680	intron 2	-	T/C	0.042
145	*PPARA*	rs135538	22	44885147	intron 2	-	G/C	0.900
146	*PPARA*	rs881740	22	44887907	intron 2	+	A/G	1.000

147	*PPARA*	rs12330015	22	44910797	intron 4	+	A/G		1.000
148	*PPARA*	rs4253712	22	44915554	intron 5	+	A/G		0.599
149	*PPARA*	rs4823613	22	44918826	intron 5	+	A/G		0.828
150	*PPARA*	*rs1800234	22	44936399	exon 8	+	T/C	V227A	0.912
151	*PPARA*	rs6007662	22	44941564	intron 8	+	A/G		0.756
152	*PPARA*	rs5767743	22	44942513	intron 8	+	T/C		0.979

Chromosome position, allele nomenclature and strand are indicated using NCBI dbSNP build 35.

The major/minor alleles were defined according to the frequencies in cases.

* literature significant SNPs.

Appendix 3: T2D association results (additive model) of 152 SNPs for stage 1 case-control samples

SNP no.	Gene	rs number	Major/Minor allele	MAF in case [a] (N = 467)	MAF in control [a] (N = 290)	OR (95% CI)	*P*
1	*PBX1*	rs6426870	C/T	0.346	0.325	1.1 (0.87-1.38)	0.427
2	*PBX1*	rs6684915	T/A	0.165	0.164	1.01 (0.75-1.35)	0.955
3	*PBX1*	rs2275560	G/A	0.162	0.163	0.99 (0.73-1.34)	0.944
4	*PBX1*	rs3767374	C/T	0.186	0.221	0.8 (0.61-1.05)	0.106
5	*PBX1*	rs7535213	G/A	0.081	0.075	1.08 (0.73-1.59)	0.717
6	*PBX1*	*rs1618566	G/A	0.262	0.255	1.04 (0.81-1.34)	0.765
7	*PBX1*	rs7556157	G/A	0.034	0.03	1.11 (0.61-2.04)	0.732
8	*PBX1*	rs1686195	C/T	0.364	0.337	1.12 (0.9-1.4)	0.309
9	*PBX1*	rs12089454	C/T	0.234	0.239	0.97 (0.76-1.25)	0.838
10	*PBX1*	rs1489330	G/T	0.331	0.325	1.03 (0.82-1.29)	0.829
11	*PBX1*	rs1001360	A/G	0.432	0.451	0.93 (0.75-1.15)	0.483
12	*PBX1*	rs3795538	T/C	0.111	0.097	1.14 (0.81-1.62)	0.446
13	*PBX1*	rs729071	A/C	0.453	0.457	0.98 (0.79-1.22)	0.883
14	*PBX1*	rs6698381	T/C	0.194	0.2	0.97 (0.74-1.27)	0.815
15	*PBX1*	rs12134242	T/A	0.164	0.161	1.02 (0.77-1.36)	0.878
16	*PBX1*	rs12081132	A/G	0.246	0.249	0.98 (0.76-1.27)	0.884
17	*PBX1*	rs12048214	C/T	0.167	0.178	0.93 (0.71-1.23)	0.62
18	*PBX1*	rs7550432	T/C	0.249	0.254	0.97 (0.76-1.24)	0.833
19	*NEUROD1*	*rs1801262	G/A	0.09	0.066	1.38 (0.93-2.04)	0.105

20	*PPARG*	rs2972164	C/T	0.055	0.079	0.67 (0.44-1.01)	0.057
21	*PPARG*	rs2920505	C/T	0.425	0.436	0.96 (0.78-1.17)	0.668
22	*PPARG*	rs10865710	C/G	0.374	0.35	1.11 (0.9-1.37)	0.343
23	*PPARG*	rs17793951	A/G	0.021	0.022	0.93 (0.45-1.89)	0.837
24	*PPARG*	*rs1801282	C/G	0.027	0.026	1.05 (0.55-2.01)	0.887
25	*PPARG*	rs17817276	A/G	0.114	0.117	0.97 (0.7-1.33)	0.836
26	*PPARG*	rs13306745	G/T	0.064	0.067	0.97 (0.65-1.44)	0.868
27	*PPARG*	rs2292101	C/T	0.332	0.308	1.12 (0.9-1.39)	0.321
28	*PPARG*	rs4135268	C/G	0.057	0.057	1.01 (0.65-1.58)	0.956
29	*PPARG*	rs4135275	C/G	0.339	0.311	1.13 (0.91-1.4)	0.274
30	*PPARG*	rs1151997	C/T	0.464	0.428	1.15 (0.94-1.42)	0.172
31	*PPARG*	rs3856806	C/T	0.286	0.242	1.25 (0.99-1.59)	0.058
32	*PPARG*	rs1152003	G/C	0.478	0.495	0.94 (0.77-1.15)	0.536
33	*ADIPOQ*	*rs16861194	A/G	0.15	0.126	1.23 (0.9-1.67)	0.187
34	*ADIPOQ*	*rs266729	C/G	0.257	0.223	1.2 (0.94-1.54)	0.134
35	*ADIPOQ*	rs182052	G/A	0.397	0.35	1.23 (0.99-1.52)	0.065
36	*ADIPOQ*	rs822394	C/A	0.127	0.114	1.13 (0.82-1.55)	0.45
37	*ADIPOQ*	*rs2241766	T/G	0.311	0.375	0.76 (0.62-0.94)	**0.011**
38	*ADIPOQ*	rs6773957	A/G	0.415	0.373	1.18 (0.96-1.46)	0.112
39	*ADIPOQ*	rs1063539	G/C	0.305	0.374	0.74 (0.6-0.92)	**0.006**
40	*PPARGC1A*	rs9790699	C/T	0.071	0.068	1.04 (0.68-1.58)	0.863
41	*PPARGC1A*	rs2279525	T/C	0.184	0.203	0.89 (0.69-1.16)	0.391
42	*PPARGC1A*	rs6821591	T/C	0.312	0.319	0.97 (0.77-1.23)	0.794
43	*PPARGC1A*	rs12650562	T/C	0.512	0.482	1.13 (0.91-1.4)	0.272
44	*PPARGC1A*	rs3774921	A/G	0.32	0.316	1.02 (0.8-1.29)	0.875

45	*PPARGC1A*	rs3736265	G/A	0.188	0.167	1.15 (0.87-1.51)	0.319
46	*PPARGC1A*	*rs8192678	G/A	0.41	0.421	0.96 (0.77-1.19)	0.677
47	*PPARGC1A*	*rs2970847	C/T	0.232	0.209	1.15 (0.88-1.49)	0.3
48	*PPARGC1A*	rs2970848	A/G	0.274	0.277	0.98 (0.78-1.24)	0.895
49	*PPARGC1A*	rs1472095	C/T	0.069	0.103	0.66 (0.44-0.99)	**0.042**
50	*PPARGC1A*	rs2932975	G/A	0.083	0.107	0.76 (0.54-1.09)	0.134
51	*PPARGC1A*	rs2970853	G/A	0.296	0.285	1.05 (0.83-1.33)	0.67
52	*PPARGC1A*	rs2970855	A/T	0.479	0.518	0.86 (0.68-1.08)	0.183
53	*PPARGC1A*	rs3774906	C/A	0.062	0.081	0.74 (0.48-1.12)	0.156
54	*PPARGC1A*	rs3796407	G/A	0.085	0.081	1.06 (0.71-1.56)	0.786
55	*PPARGC1A*	rs12500214	G/A	0.117	0.11	1.06 (0.77-1.47)	0.712
56	*PPARGC1A*	rs2946385	G/T	0.39	0.404	0.94 (0.75-1.18)	0.603
57	*PPARGC1A*	rs2970873	C/T	0.291	0.314	0.9 (0.72-1.13)	0.357
58	*PPARGC1A*	rs2970872	T/C	0.259	0.24	1.1 (0.87-1.41)	0.427
59	*PPARGC1A*	rs2970871	G/A	0.383	0.395	0.95 (0.77-1.18)	0.661
60	*PPARGC1A*	rs3774902	C/T	0.351	0.361	0.96 (0.77-1.19)	0.696
61	*PPARGC1A*	rs3774901	G/A	0.15	0.129	1.19 (0.87-1.62)	0.278
62	*PPARGC1A*	rs2970869	G/A	0.335	0.347	0.95 (0.76-1.18)	0.643
63	*GCK*	rs2268574	C/T	0.317	0.322	0.98 (0.79-1.22)	0.838
64	*GCK*	rs2908296	C/A	0.273	0.291	0.92 (0.73-1.15)	0.466
65	*GCK*	*rs1799831	G/A	0.278	0.286	0.96 (0.76-1.21)	0.73
66	*GCK*	rs17832252	G/T	0.229	0.237	0.96 (0.75-1.22)	0.717
67	*GCK*	rs758989	A/G	0.324	0.287	1.18 (0.95-1.47)	0.143
68	*GCK*	rs2284776	A/G	0.316	0.285	1.15 (0.92-1.44)	0.206
69	*GCK*	rs2300586	C/T	0.294	0.317	0.89 (0.71-1.13)	0.33

70	*GCK*	rs758985	C/T	0.168	0.164	1.03 (0.78-1.35)	0.854
71	*GCK*	rs741038	T/C	0.25	0.23	1.11 (0.87-1.42)	0.386
72	*GCK*	rs13239289	C/G	0.091	0.062	1.48 (1-2.19)	**0.049**
73	*GCK*	rs2268569	G/A	0.196	0.229	0.82 (0.64-1.06)	0.126
74	*GCK*	*rs1799884	G/A	0.185	0.167	1.13 (0.86-1.48)	0.386
75	*LPL*	rs253	C/T	0.392	0.389	1.01 (0.82-1.25)	0.931
76	*LPL*	rs270	C/A	0.134	0.153	0.86 (0.65-1.15)	0.321
77	*LPL*	rs281	T/A	0.243	0.229	1.08 (0.85-1.37)	0.545
78	*LPL*	rs283	G/A	0.125	0.145	0.85 (0.63-1.14)	0.268
79	*LPL*	rs285	C/T	0.364	0.367	0.99 (0.8-1.23)	0.912
80	*LPL*	rs291	T/C	0.239	0.234	1.02 (0.8-1.3)	0.855
81	*LPL*	*rs320	T/G	0.251	0.245	1.03 (0.8-1.33)	0.825
82	*LPL*	rs15285	G/A	0.13	0.146	0.87 (0.65-1.17)	0.363
83	*LPL*	rs4921684	C/T	0.11	0.082	1.38 (0.96-1.98)	0.079
84	*KCNJ11*	rs10832768	G/T	0.459	0.435	1.1 (0.9-1.34)	0.369
85	*KCNJ11*	rs16933984	T/G	0.104	0.122	0.83 (0.6-1.16)	0.273
86	*KCNJ11*	*rs5215	T/C	0.346	0.289	1.29 (1.03-1.6)	**0.024**
87	*KCNJ11*	*rs5219	C/T	0.347	0.29	1.28 (1.03-1.6)	**0.025**
88	*ABCC8*	*rs4148643	G/A	0.062	0.081	0.76 (0.52-1.13)	0.172
89	*ABCC8*	rs2074312	T/C	0.4	0.378	1.09 (0.89-1.34)	0.42
90	*ABCC8*	rs4148633	C/T	0.027	0.037	0.71 (0.4-1.26)	0.242
91	*ABCC8*	*rs1799858	G/A	0.182	0.188	0.96 (0.74-1.25)	0.782
92	*ABCC8*	rs1799857	C/T	0.256	0.272	0.92 (0.73-1.16)	0.49
93	*HNF1A*	*rs1169288	T/G	0.4	0.356	1.19 (0.97-1.47)	0.095
94	*HNF1A*	rs2244608	T/C	0.384	0.346	1.18 (0.95-1.45)	0.135

95	*HNF1A*	rs1169286	A/G	0.461	0.409	1.23 (1-1.51)	**0.049**
96	*HNF1A*	rs2464196	C/T	0.48	0.444	1.16 (0.94-1.43)	0.174
97	*HNF1A*	rs2178464	T/C	0.135	0.153	0.86 (0.64-1.15)	0.314
98	*HNF1A*	rs1169312	G/T	0.473	0.457	1.07 (0.87-1.32)	0.539
99	*HNF1A*	rs3751152	C/T	0.112	0.132	0.82 (0.59-1.14)	0.234
100	*HNF1A*	rs3751150	C/T	0.226	0.241	0.92 (0.72-1.17)	0.499
101	*IPF1*	rs7982864	C/T	0.177	0.188	0.93 (0.71-1.21)	0.567
102	*LIPC*	*rs1800588	C/T	0.379	0.367	1.05 (0.85-1.29)	0.65
103	*LIPC*	rs6494005	A/G	0.234	0.227	1.04 (0.82-1.31)	0.756
104	*LIPC*	rs8033940	A/G	0.293	0.279	1.07 (0.85-1.33)	0.578
105	*LIPC*	rs2242062	A/G	0.375	0.317	1.28 (1.03-1.59)	**0.024**
106	*LIPC*	rs11852861	C/T	0.153	0.141	1.1 (0.82-1.46)	0.528
107	*LIPC*	rs12592139	G/A	0.259	0.247	1.07 (0.84-1.36)	0.6
108	*LIPC*	rs12592127	A/G	0.363	0.353	1.05 (0.84-1.3)	0.69
109	*LIPC*	rs11633043	G/A	0.19	0.195	0.97 (0.74-1.26)	0.816
110	*LIPC*	rs6082	A/G	0.396	0.37	1.12 (0.9-1.38)	0.309
111	*LIPC*	rs6083	G/A	0.181	0.197	0.9 (0.69-1.18)	0.452
112	*LIPC*	rs6084	C/G	0.074	0.079	0.93 (0.64-1.37)	0.727
113	*LIPC*	rs2242064	T/G	0.339	0.339	1 (0.8-1.25)	0.98
114	*LIPC*	rs2242065	C/T	0.332	0.319	1.06 (0.85-1.32)	0.625
115	*LIPC*	rs7165654	A/G	0.428	0.43	0.99 (0.81-1.22)	0.944
116	*LIPC*	rs7178362	T/C	0.203	0.209	0.96 (0.75-1.25)	0.782
117	*LIPC*	rs17190678	C/G	0.099	0.089	1.12 (0.79-1.58)	0.52
118	*LIPC*	rs3829461	C/T	0.048	0.04	1.19 (0.73-1.94)	0.486
119	*LIPC*	rs8030893	G/C	0.139	0.134	1.04 (0.78-1.39)	0.798

120	*LIPC*	rs3829460	A/T	0.375	0.382	0.97 (0.78-1.21)	0.79
121	*LIPC*	rs6074	C/A	0.239	0.243	0.98 (0.76-1.25)	0.841
122	*HNF4A*	rs4812823	T/C	0.105	0.101	1.06 (0.74-1.5)	0.763
123	*HNF4A*	rs4812828	C/T	0.454	0.526	0.75 (0.61-0.92)	**0.007**
124	*HNF4A*	*rs1884614	C/T	0.463	0.399	1.28 (1.05-1.57)	**0.017**
125	*HNF4A*	*rs2144908	G/A	0.469	0.406	1.27 (1.04-1.56)	**0.021**
126	*HNF4A*	rs17755638	T/C	0.107	0.14	0.73 (0.53-1)	0.05
127	*HNF4A*	*rs6073418	C/T	0.37	0.4	0.89 (0.72-1.09)	0.256
128	*HNF4A*	rs2425637	T/G	0.496	0.5	0.98 (0.8-1.21)	0.876
129	*HNF4A*	rs2071197	G/A	0.491	0.492	1 (0.81-1.22)	0.977
130	*HNF4A*	rs2071199	T/C	0.172	0.146	1.21 (0.91-1.6)	0.19
131	*HNF4A*	rs2071200	C/T	0.252	0.248	1.02 (0.8-1.29)	0.868
132	*HNF4A*	rs745975	G/A	0.169	0.2	0.81 (0.62-1.06)	0.124
133	*HNF4A*	rs11574730	G/A	0.081	0.09	0.89 (0.62-1.28)	0.515
134	*HNF4A*	rs6017340	C/T	0.157	0.176	0.87 (0.66-1.15)	0.326
135	*HNF4A*	rs6031587	T/C	0.467	0.454	1.05 (0.86-1.3)	0.62
136	*HNF4A*	rs3212191	C/T	0.463	0.46	1.01 (0.82-1.25)	0.912
137	*HNF4A*	rs11574736	C/G	0.143	0.133	1.09 (0.81-1.46)	0.587
138	*HNF4A*	rs6093978	C/T	0.324	0.329	0.98 (0.78-1.22)	0.83
139	*HNF4A*	rs3212198	C/T	0.307	0.32	0.94 (0.76-1.18)	0.605
140	*HNF4A*	*rs3818247	T/G	0.362	0.327	1.17 (0.94-1.46)	0.153
141	*PPARA*	rs4253623	A/G	0.096	0.082	1.19 (0.79-1.78)	0.399
142	*PPARA*	rs135549	A/G	0.204	0.183	1.15 (0.87-1.52)	0.322
143	*PPARA*	rs135547	G/C	0.081	0.095	0.84 (0.58-1.22)	0.369
144	*PPARA*	rs129600	T/C	0.389	0.409	0.93 (0.75-1.14)	0.481

145	*PPARA*	rs135538	G/C	0.439	0.432	1.03 (0.83-1.28)	0.775
146	*PPARA*	rs881740	A/G	0.135	0.089	1.59 (1.12-2.25)	**0.01**
147	*PPARA*	rs12330015	A/G	0.125	0.147	0.83 (0.59-1.15)	0.26
148	*PPARA*	rs4253712	A/G	0.119	0.132	0.89 (0.64-1.22)	0.462
149	*PPARA*	rs4823613	A/G	0.231	0.253	0.89 (0.69-1.13)	0.337
150	*PPARA*	*rs1800234	T/C	0.042	0.051	0.82 (0.5-1.34)	0.421
151	*PPARA*	rs6007662	A/G	0.131	0.143	0.91 (0.67-1.24)	0.536
152	*PPARA*	rs5767743	T/C	0.219	0.238	0.9 (0.7-1.16)	0.405

* literature significant SNPs.

Significant associations at $P < 0.05$ are bolded.

[a] Minor alleles were defined according to the frequencies in cases shown in Appendix 2.

The odds ratios (ORs) were calculated with reference to the minor allele.

Appendix 4: T2D association results (additive model) of 152 SNPs for stage 1 case-control samples subset by metabolic syndrome status in cases

				Minor Allele Frequency			OR (95% CI)	*P*	OR (95% CI)	*P*
SNP no.	Gene	rs number	Major/Minor allele [a]	All Controls (N = 290)	MetS negative T2D (N = 223)	MetS positive T2D (N = 244)	MetS negative T2D		MetS positive T2D	
1	*PBX1*	rs6426870	C/T	0.325	0.331	0.361	1.03 (0.77-1.36)	0.862	1.17 (0.89-1.53)	0.253
2	*PBX1*	rs6684915	T/A	0.164	0.18	0.151	1.12 (0.79-1.59)	0.511	0.9 (0.63-1.29)	0.574
3	*PBX1*	rs2275560	G/A	0.163	0.154	0.168	0.93 (0.65-1.35)	0.713	1.04 (0.73-1.48)	0.824
4	*PBX1*	rs3767374	C/T	0.221	0.164	0.206	0.68 (0.48-0.96)	**0.026**	0.91 (0.67-1.24)	0.564
5	*PBX1*	rs7535213	G/A	0.075	0.08	0.081	1.07 (0.67-1.72)	0.765	1.07 (0.69-1.68)	0.756
6	*PBX1*	*rs1618566	G/A	0.255	0.253	0.271	0.98 (0.72-1.34)	0.921	1.09 (0.81-1.45)	0.569
7	*PBX1*	rs7556157	G/A	0.03	0.047	0.021	1.62 (0.82-3.2)	0.164	0.69 (0.31-1.55)	0.372
8	*PBX1*	rs1686195	C/T	0.337	0.4	0.331	1.3 (1-1.7)	0.050	0.98 (0.75-1.27)	0.849
9	*PBX1*	rs12089454	C/T	0.239	0.246	0.224	1.04 (0.77-1.4)	0.814	0.92 (0.69-1.23)	0.584
10	*PBX1*	rs1489330	G/T	0.325	0.319	0.342	0.97 (0.73-1.29)	0.827	1.08 (0.83-1.4)	0.581
11	*PBX1*	rs1001360	A/G	0.451	0.436	0.429	0.94 (0.73-1.21)	0.639	0.91 (0.71-1.17)	0.478
12	*PBX1*	rs3795538	T/C	0.097	0.098	0.123	1.01 (0.66-1.54)	0.969	1.26 (0.86-1.86)	0.237
13	*PBX1*	rs729071	A/C	0.457	0.444	0.462	0.95 (0.73-1.23)	0.67	1.02 (0.79-1.31)	0.875
14	*PBX1*	rs6698381	T/C	0.2	0.195	0.194	0.97 (0.7-1.35)	0.872	0.96 (0.7-1.33)	0.817
15	*PBX1*	rs12134242	T/A	0.161	0.163	0.164	1.02 (0.73-1.43)	0.915	1.03 (0.73-1.43)	0.879
16	*PBX1*	rs12081132	A/G	0.249	0.26	0.233	1.06 (0.78-1.44)	0.692	0.91 (0.67-1.23)	0.531
17	*PBX1*	rs12048214	C/T	0.178	0.167	0.168	0.93 (0.67-1.29)	0.656	0.94 (0.68-1.29)	0.698
18	*PBX1*	rs7550432	T/C	0.254	0.237	0.262	0.91 (0.68-1.22)	0.526	1.04 (0.78-1.38)	0.799

19	*NEUROD1*	*rs1801262	G/A	0.066	0.082	0.097	1.25 (0.79-1.97)	0.345	1.5 (0.97-2.33)	0.068
20	*PPARG*	rs2972164	C/T	0.079	0.064	0.045	0.8 (0.5-1.29)	0.366	0.53 (0.31-0.91)	**0.02**
21	*PPARG*	rs2920505	C/T	0.436	0.424	0.426	0.95 (0.74-1.22)	0.68	0.96 (0.76-1.22)	0.741
22	*PPARG*	rs10865710	C/G	0.35	0.359	0.388	1.04 (0.8-1.34)	0.765	1.17 (0.92-1.49)	0.205
23	*PPARG*	rs17793951	A/G	0.022	0.024	0.018	1.08 (0.47-2.45)	0.862	0.79 (0.33-1.89)	0.602
24	*PPARG*	*rs1801282	C/G	0.026	0.032	0.022	1.28 (0.61-2.68)	0.509	0.84 (0.38-1.86)	0.667
25	*PPARG*	rs17817276	A/G	0.117	0.107	0.121	0.89 (0.6-1.33)	0.581	1.03 (0.71-1.49)	0.87
26	*PPARG*	rs13306745	G/T	0.067	0.067	0.062	1.01 (0.63-1.62)	0.959	0.92 (0.57-1.48)	0.741
27	*PPARG*	rs2292101	C/T	0.308	0.313	0.35	1.03 (0.79-1.34)	0.842	1.2 (0.94-1.54)	0.149
28	*PPARG*	rs4135268	C/G	0.057	0.071	0.044	1.28 (0.77-2.13)	0.335	0.78 (0.45-1.35)	0.372
29	*PPARG*	rs4135275	C/G	0.311	0.317	0.358	1.03 (0.79-1.33)	0.846	1.23 (0.96-1.57)	0.107
30	*PPARG*	rs1151997	C/T	0.428	0.444	0.482	1.07 (0.84-1.37)	0.598	1.24 (0.97-1.57)	0.081
31	*PPARG*	rs3856806	C/T	0.242	0.279	0.292	1.22 (0.92-1.6)	0.168	1.28 (0.98-1.67)	0.067
32	*PPARG*	rs1152003	G/C	0.495	0.476	0.48	0.93 (0.73-1.19)	0.553	0.95 (0.75-1.19)	0.639
33	*ADIPOQ*	*rs16861194	A/G	0.126	0.153	0.147	1.26 (0.88-1.81)	0.209	1.2 (0.85-1.7)	0.303
34	*ADIPOQ*	*rs266729	C/G	0.223	0.234	0.279	1.06 (0.79-1.42)	0.686	1.35 (1.02-1.78)	**0.036**
35	*ADIPOQ*	rs182052	G/A	0.35	0.373	0.421	1.1 (0.85-1.42)	0.459	1.36 (1.06-1.76)	**0.018**
36	*ADIPOQ*	rs822394	C/A	0.114	0.12	0.134	1.06 (0.73-1.53)	0.773	1.2 (0.83-1.73)	0.323
37	*ADIPOQ*	*rs2241766	T/G	0.375	0.331	0.293	0.83 (0.65-1.07)	0.149	0.7 (0.54-0.9)	**0.005**
38	*ADIPOQ*	rs6773957	A/G	0.373	0.388	0.439	1.06 (0.83-1.36)	0.632	1.32 (1.03-1.69)	**0.027**
39	*ADIPOQ*	rs1063539	G/C	0.374	0.316	0.294	0.79 (0.61-1.01)	0.061	0.7 (0.54-0.9)	**0.006**
40	*PPARGC1A*	rs9790699	C/T	0.069	0.08	0.062	1.18 (0.73-1.91)	0.501	0.9 (0.54-1.49)	0.68
41	*PPARGC1A*	rs2279525	T/C	0.204	0.181	0.188	0.87 (0.64-1.19)	0.382	0.91 (0.66-1.24)	0.537
42	*PPARGC1A*	rs6821591	T/C	0.318	0.338	0.288	1.1 (0.83-1.46)	0.488	0.85 (0.64-1.14)	0.276
43	*PPARGC1A*	rs12650562	T/C	0.482	0.543	0.482	1.3 (1-1.68)	**0.049**	1 (0.77-1.29)	0.997
44	*PPARGC1A*	rs3774921	A/G	0.316	0.341	0.3	1.14 (0.86-1.5)	0.377	0.92 (0.69-1.22)	0.566
45	*PPARGC1A*	rs3736265	G/A	0.167	0.193	0.183	1.19 (0.86-1.64)	0.305	1.11 (0.81-1.54)	0.515

46	*PPARGC1A*	*rs8192678	G/A	0.421	0.385	0.433	0.86 (0.66-1.12)	0.253	1.06 (0.82-1.36)	0.675
47	*PPARGC1A*	*rs2970847	C/T	0.208	0.247	0.218	1.26 (0.93-1.71)	0.141	1.06 (0.78-1.45)	0.7
48	*PPARGC1A*	rs2970848	A/G	0.278	0.286	0.262	1.04 (0.79-1.37)	0.777	0.92 (0.7-1.22)	0.571
49	*PPARGC1A*	rs1472095	C/T	0.103	0.077	0.061	0.75 (0.47-1.2)	0.227	0.56 (0.33-0.95)	**0.031**
50	*PPARGC1A*	rs2932975	G/A	0.108	0.092	0.076	0.84 (0.56-1.27)	0.412	0.68 (0.44-1.05)	0.084
51	*PPARGC1A*	rs2970853	G/A	0.285	0.293	0.298	1.05 (0.79-1.38)	0.756	1.07 (0.81-1.4)	0.646
52	*PPARGC1A*	rs2970855	A/T	0.518	0.48	0.477	0.86 (0.65-1.13)	0.278	0.85 (0.65-1.12)	0.245
53	*PPARGC1A*	rs3774906	C/A	0.081	0.052	0.071	0.61 (0.35-1.03)	0.066	0.87 (0.53-1.4)	0.559
54	*PPARGC1A*	rs3796407	G/A	0.081	0.085	0.084	1.06 (0.66-1.68)	0.821	1.05 (0.66-1.66)	0.839
55	*PPARGC1A*	rs12500214	G/A	0.11	0.106	0.127	0.96 (0.65-1.43)	0.837	1.16 (0.8-1.68)	0.447
56	*PPARGC1A*	rs2946385	G/T	0.404	0.428	0.353	1.1 (0.84-1.44)	0.487	0.8 (0.61-1.06)	0.117
57	*PPARGC1A*	rs2970873	C/T	0.315	0.31	0.273	0.98 (0.74-1.28)	0.857	0.82 (0.63-1.08)	0.157
58	*PPARGC1A*	rs2970872	T/C	0.241	0.271	0.248	1.17 (0.88-1.57)	0.28	1.04 (0.78-1.38)	0.81
59	*PPARGC1A*	rs2970871	G/A	0.396	0.401	0.366	1.02 (0.79-1.33)	0.864	0.88 (0.69-1.14)	0.335
60	*PPARGC1A*	rs3774902	C/T	0.359	0.318	0.382	0.83 (0.63-1.09)	0.177	1.1 (0.85-1.43)	0.45
61	*PPARGC1A*	rs3774901	G/A	0.13	0.153	0.146	1.23 (0.85-1.8)	0.275	1.15 (0.8-1.66)	0.445
62	*PPARGC1A*	rs2970869	G/A	0.348	0.349	0.321	1.01 (0.77-1.31)	0.972	0.89 (0.69-1.15)	0.385
63	*GCK*	rs2268574	C/T	0.322	0.31	0.323	0.94 (0.72-1.23)	0.667	1.01 (0.78-1.29)	0.96
64	*GCK*	rs2908296	C/A	0.291	0.281	0.266	0.96 (0.73-1.25)	0.743	0.89 (0.68-1.15)	0.372
65	*GCK*	*rs1799831	G/A	0.286	0.286	0.271	1 (0.76-1.32)	0.993	0.93 (0.71-1.21)	0.576
66	*GCK*	rs17832252	G/T	0.237	0.236	0.223	0.99 (0.74-1.33)	0.957	0.93 (0.7-1.22)	0.59
67	*GCK*	rs758989	A/G	0.287	0.317	0.33	1.15 (0.88-1.51)	0.293	1.2 (0.94-1.54)	0.145
68	*GCK*	rs2284776	A/G	0.285	0.298	0.333	1.07 (0.81-1.4)	0.648	1.24 (0.96-1.59)	0.097
69	*GCK*	rs2300586	C/T	0.317	0.301	0.286	0.93 (0.7-1.22)	0.583	0.86 (0.66-1.13)	0.281
70	*GCK*	rs758985	C/T	0.164	0.169	0.167	1.03 (0.75-1.43)	0.85	1.02 (0.75-1.39)	0.899
71	*GCK*	rs741038	T/C	0.23	0.245	0.255	1.09 (0.81-1.45)	0.572	1.14 (0.86-1.52)	0.359
72	*GCK*	rs13239289	C/G	0.062	0.093	0.089	1.51 (0.97-2.35)	0.071	1.47 (0.93-2.31)	0.096

73	*GCK*	rs2268569	G/A	0.229	0.197	0.194	0.83 (0.62-1.12)	0.225	0.82 (0.61-1.09)	0.17
74	*GCK*	*rs1799884	G/A	0.167	0.198	0.172	1.24 (0.9-1.7)	0.193	1.04 (0.75-1.43)	0.835
75	*LPL*	rs253	C/T	0.389	0.413	0.372	1.1 (0.86-1.41)	0.443	0.93 (0.73-1.19)	0.563
76	*LPL*	rs270	C/A	0.153	0.126	0.141	0.81 (0.57-1.15)	0.241	0.92 (0.66-1.28)	0.605
77	*LPL*	rs281	T/A	0.229	0.252	0.234	1.13 (0.85-1.49)	0.399	1.03 (0.78-1.37)	0.844
78	*LPL*	rs283	G/A	0.145	0.107	0.142	0.72 (0.5-1.04)	0.076	0.97 (0.69-1.36)	0.877
79	*LPL*	rs285	C/T	0.367	0.41	0.323	1.19 (0.93-1.54)	0.169	0.82 (0.64-1.06)	0.134
80	*LPL*	rs291	T/C	0.234	0.299	0.185	1.39 (1.05-1.84)	**0.021**	0.73 (0.54-0.99)	**0.041**
81	*LPL*	*rs320	T/G	0.245	0.308	0.198	1.37 (1.02-1.84)	**0.036**	0.75 (0.54-1.03)	0.075
82	*LPL*	rs15285	G/A	0.146	0.122	0.137	0.81 (0.57-1.17)	0.259	0.93 (0.66-1.31)	0.669
83	*LPL*	rs4921684	C/T	0.082	0.132	0.09	1.7 (1.13-2.55)	**0.011**	1.11 (0.72-1.71)	0.632
84	*KCNJ11*	rs10832768	G/T	0.435	0.45	0.467	1.06 (0.84-1.34)	0.629	1.13 (0.89-1.42)	0.309
85	*KCNJ11*	rs16933984	T/G	0.122	0.101	0.107	0.81 (0.55-1.2)	0.296	0.86 (0.59-1.25)	0.43
86	*KCNJ11*	*rs5215	T/C	0.289	0.337	0.355	1.23 (0.95-1.59)	0.111	1.35 (1.04-1.74)	**0.023**
87	*KCNJ11*	*rs5219	C/T	0.29	0.338	0.354	1.24 (0.96-1.6)	0.105	1.34 (1.04-1.73)	**0.026**
88	*ABCC8*	*rs4148643	G/A	0.081	0.061	0.063	0.75 (0.47-1.2)	0.227	0.79 (0.5-1.23)	0.296
89	*ABCC8*	rs2074312	T/C	0.378	0.377	0.421	0.99 (0.78-1.26)	0.955	1.19 (0.93-1.51)	0.164
90	*ABCC8*	rs4148633	C/T	0.037	0.024	0.03	0.63 (0.3-1.3)	0.212	0.79 (0.41-1.53)	0.484
91	*ABCC8*	*rs1799858	G/A	0.188	0.189	0.175	1.01 (0.74-1.37)	0.957	0.92 (0.68-1.26)	0.607
92	*ABCC8*	rs1799857	C/T	0.272	0.265	0.248	0.96 (0.73-1.27)	0.798	0.89 (0.68-1.16)	0.372
93	*HNF1A*	*rs1169288	T/G	0.356	0.395	0.404	1.18 (0.92-1.51)	0.2	1.22 (0.96-1.55)	0.108
94	*HNF1A*	rs2244608	T/C	0.346	0.382	0.387	1.17 (0.91-1.5)	0.231	1.19 (0.93-1.53)	0.165
95	*HNF1A*	rs1169286	A/G	0.409	0.455	0.466	1.2 (0.94-1.52)	0.146	1.26 (0.99-1.6)	0.062
96	*HNF1A*	rs2464196	C/T	0.444	0.478	0.482	1.15 (0.89-1.48)	0.277	1.17 (0.91-1.49)	0.215
97	*HNF1A*	rs2178464	T/C	0.153	0.126	0.143	0.79 (0.55-1.14)	0.207	0.92 (0.66-1.29)	0.631
98	*HNF1A*	rs1169312	G/T	0.457	0.466	0.48	1.04 (0.81-1.34)	0.78	1.1 (0.86-1.41)	0.449
99	*HNF1A*	rs3751152	C/T	0.132	0.117	0.107	0.87 (0.59-1.28)	0.483	0.78 (0.53-1.14)	0.196

100	*HNF1A*	rs3751150	C/T	0.241	0.223	0.228	0.9 (0.67-1.21)	0.494	0.93 (0.7-1.24)	0.627
101	*IPF1*	rs7982864	C/T	0.188	0.168	0.185	0.87 (0.63-1.2)	0.402	0.98 (0.72-1.32)	0.872
102	*LIPC*	*rs1800588	C/T	0.367	0.367	0.389	1 (0.78-1.28)	0.996	1.1 (0.86-1.39)	0.457
103	*LIPC*	rs6494005	A/G	0.227	0.232	0.236	1.03 (0.77-1.36)	0.86	1.05 (0.8-1.38)	0.726
104	*LIPC*	rs8033940	A/G	0.279	0.288	0.298	1.04 (0.8-1.35)	0.775	1.09 (0.84-1.4)	0.52
105	*LIPC*	rs2242062	A/G	0.317	0.375	0.375	1.28 (1-1.65)	0.054	1.29 (1-1.66)	**0.048**
106	*LIPC*	rs11852861	C/T	0.141	0.14	0.165	0.99 (0.69-1.4)	0.938	1.21 (0.87-1.68)	0.265
107	*LIPC*	rs12592139	G/A	0.247	0.26	0.259	1.07 (0.81-1.41)	0.653	1.06 (0.81-1.4)	0.663
108	*LIPC*	rs12592127	A/G	0.353	0.358	0.367	1.02 (0.79-1.32)	0.866	1.06 (0.83-1.36)	0.627
109	*LIPC*	rs11633043	G/A	0.195	0.179	0.2	0.9 (0.65-1.23)	0.508	1.04 (0.76-1.42)	0.814
110	*LIPC*	rs6082	A/G	0.37	0.38	0.411	1.04 (0.81-1.34)	0.75	1.18 (0.93-1.5)	0.179
111	*LIPC*	rs6083	G/A	0.197	0.168	0.194	0.83 (0.6-1.14)	0.242	0.98 (0.72-1.33)	0.891
112	*LIPC*	rs6084	C/G	0.079	0.056	0.091	0.71 (0.44-1.14)	0.158	1.16 (0.76-1.77)	0.487
113	*LIPC*	rs2242064	T/G	0.339	0.347	0.331	1.04 (0.79-1.35)	0.8	0.97 (0.75-1.25)	0.785
114	*LIPC*	rs2242065	C/T	0.319	0.333	0.33	1.07 (0.82-1.39)	0.633	1.05 (0.82-1.34)	0.722
115	*LIPC*	rs7165654	A/G	0.43	0.404	0.449	0.91 (0.71-1.15)	0.43	1.08 (0.85-1.36)	0.536
116	*LIPC*	rs7178362	T/C	0.209	0.197	0.208	0.93 (0.69-1.26)	0.645	1 (0.74-1.35)	0.981
117	*LIPC*	rs17190678	C/G	0.089	0.086	0.111	0.97 (0.64-1.46)	0.874	1.26 (0.86-1.85)	0.236
118	*LIPC*	rs3829461	C/T	0.04	0.042	0.054	1.05 (0.58-1.89)	0.876	1.3 (0.76-2.23)	0.33
119	*LIPC*	rs8030893	G/C	0.134	0.115	0.161	0.84 (0.59-1.21)	0.359	1.22 (0.88-1.68)	0.23
120	*LIPC*	rs3829460	A/T	0.382	0.371	0.379	0.95 (0.74-1.23)	0.72	0.99 (0.77-1.26)	0.92
121	*LIPC*	rs6074	C/A	0.243	0.26	0.219	1.1 (0.82-1.46)	0.531	0.88 (0.66-1.16)	0.358
122	*HNF4A*	rs4812823	T/C	0.101	0.096	0.115	0.94 (0.62-1.43)	0.784	1.17 (0.78-1.74)	0.453
123	*HNF4A*	rs4812828	C/T	0.526	0.454	0.454	0.75 (0.58-0.96)	**0.022**	0.76 (0.6-0.96)	**0.021**
124	*HNF4A*	*rs1884614	C/T	0.399	0.474	0.452	1.35 (1.06-1.73)	**0.016**	1.22 (0.97-1.55)	0.089
125	*HNF4A*	*rs2144908	G/A	0.406	0.476	0.462	1.32 (1.03-1.69)	**0.027**	1.23 (0.97-1.55)	0.082
126	*HNF4A*	rs17755638	T/C	0.14	0.102	0.113	0.67 (0.45-1)	0.051	0.77 (0.53-1.12)	0.17

127	*HNF4A*	*rs6073418	C/T	0.4	0.377	0.365	0.91 (0.71-1.16)	0.449	0.87 (0.69-1.1)	0.249
128	*HNF4A*	rs2425637	T/G	0.5	0.5	0.492	1 (0.78-1.28)	1	0.97 (0.76-1.23)	0.795
129	*HNF4A*	rs2071197	G/A	0.492	0.476	0.504	0.95 (0.75-1.19)	0.641	1.05 (0.83-1.33)	0.683
130	*HNF4A*	rs2071199	T/C	0.146	0.197	0.148	1.43 (1.03-1.96)	**0.031**	1.02 (0.73-1.41)	0.929
131	*HNF4A*	rs2071200	C/T	0.248	0.266	0.239	1.09 (0.83-1.44)	0.518	0.95 (0.72-1.26)	0.72
132	*HNF4A*	rs745975	G/A	0.2	0.156	0.181	0.74 (0.54-1.02)	0.07	0.88 (0.65-1.2)	0.428
133	*HNF4A*	rs11574730	G/A	0.09	0.073	0.088	0.79 (0.51-1.24)	0.312	0.98 (0.64-1.48)	0.905
134	*HNF4A*	rs6017340	C/T	0.176	0.178	0.138	1.01 (0.74-1.39)	0.946	0.75 (0.54-1.04)	0.088
135	*HNF4A*	rs6031587	T/C	0.454	0.493	0.444	1.16 (0.91-1.47)	0.228	0.96 (0.75-1.22)	0.735
136	*HNF4A*	rs3212191	C/T	0.46	0.483	0.445	1.09 (0.86-1.39)	0.486	0.94 (0.74-1.2)	0.62
137	*HNF4A*	rs11574736	C/G	0.133	0.155	0.132	1.19 (0.84-1.69)	0.322	0.99 (0.7-1.4)	0.974
138	*HNF4A*	rs6093978	C/T	0.329	0.333	0.316	1.01 (0.79-1.31)	0.912	0.94 (0.73-1.22)	0.643
139	*HNF4A*	rs3212198	C/T	0.32	0.332	0.285	1.06 (0.82-1.37)	0.679	0.85 (0.65-1.1)	0.207
140	*HNF4A*	*rs3818247	T/G	0.327	0.353	0.371	1.12 (0.87-1.45)	0.374	1.21 (0.94-1.56)	0.13
141	*PPARA*	rs4253623	A/G	0.082	0.113	0.079	1.42 (0.91-2.24)	0.126	0.96 (0.58-1.59)	0.874
142	*PPARA*	rs135549	A/G	0.184	0.209	0.199	1.19 (0.86-1.64)	0.307	1.11 (0.8-1.54)	0.534
143	*PPARA*	rs135547	G/C	0.095	0.068	0.093	0.69 (0.43-1.11)	0.128	0.98 (0.64-1.49)	0.925
144	*PPARA*	rs129600	T/C	0.409	0.37	0.408	0.86 (0.67-1.1)	0.238	1 (0.78-1.27)	0.991
145	*PPARA*	rs135538	G/C	0.433	0.441	0.438	1.03 (0.8-1.33)	0.803	1.02 (0.79-1.31)	0.889
146	*PPARA*	rs881740	A/G	0.09	0.143	0.127	1.71 (1.14-2.55)	**0.009**	1.48 (0.99-2.22)	0.057
147	*PPARA*	rs12330015	A/G	0.147	0.107	0.142	0.7 (0.46-1.06)	0.09	0.96 (0.65-1.41)	0.831
148	*PPARA*	rs4253712	A/G	0.13	0.117	0.12	0.89 (0.61-1.3)	0.547	0.91 (0.62-1.33)	0.62
149	*PPARA*	rs4823613	A/G	0.252	0.231	0.23	0.9 (0.67-1.2)	0.463	0.89 (0.67-1.19)	0.423
150	*PPARA*	*rs1800234	T/C	0.05	0.035	0.049	0.71 (0.38-1.33)	0.28	0.98 (0.56-1.72)	0.947
151	*PPARA*	rs6007662	A/G	0.143	0.153	0.111	1.08 (0.76-1.54)	0.675	0.75 (0.51-1.09)	0.13
152	*PPARA*	rs5767743	T/C	0.237	0.231	0.208	0.97 (0.72-1.3)	0.839	0.85 (0.63-1.14)	0.271

* literature significant SNPs.

Significant associations at $P < 0.05$ are bolded.

[a] Minor alleles were defined according to the frequencies in cases shown in Appendix 2.

The odds ratios (ORs) were calculated with reference to the minor allele.

Appendix 5: T2D association results (additive model) of 22 SNPs for stage 2 case-control samples

SNP no.	Gene	rs number	Major/Minor allele [a]	MAF in cases (N = 994)	MAF in controls (N = 310)	OR (95% CI)	*P*
1	*PBX1*	rs3767374	C/T	0.213	0.212	1.01 (0.77-1.33)	0.943
2	*PPARG*	rs2972164	C/T	0.064	0.044	1.5 (0.97-2.31)	0.067
3	*ADIPOQ*	*rs266729	C/G	0.253	0.252	1.01 (0.81-1.24)	0.956
4	*ADIPOQ*	*rs2241766	T/G	0.294	0.279	1.08 (0.86-1.34)	0.51
5	*ADIPOQ*	rs6773957	A/G	0.447	0.426	1.09 (0.9-1.31)	0.366
6	*ADIPOQ*	rs1063539	G/C	0.288	0.283	1.02 (0.84-1.25)	0.819
7	*PPARGC1A*	rs12650562	T/C	0.501	0.479	1.1 (0.91-1.32)	0.338
8	*PPARGC1A*	rs1472095	C/T	0.081	0.07	1.18 (0.82-1.69)	0.38
9	*GCK*	rs13239289	C/G	0.07	0.079	0.87 (0.62-1.23)	0.438
10	*LPL*	rs291	T/C	0.217	0.219	0.99 (0.79-1.24)	0.949
11	*LPL*	*rs320	T/G	0.215	0.216	0.99 (0.79-1.24)	0.944
12	*LPL*	rs4921684	C/T	0.094	0.086	1.11 (0.78-1.57)	0.569
13	*KCNJ11*	*rs5219	C/T	0.324	0.326	0.99 (0.81-1.21)	0.942
14	*ABCC8*	*rs4148643	G/A	0.063	0.07	0.9 (0.57-1.41)	0.639
15	*HNF1A*	rs1169286	A/G	0.441	0.431	1.04 (0.87-1.26)	0.658
16	*LIPC*	rs2242062	A/G	0.343	0.37	0.89 (0.74-1.08)	0.24
17	*HNF4A*	rs4812828	C/T	0.487	0.479	1.03 (0.86-1.24)	0.722
18	*HNF4A*	*rs1884614	C/T	0.445	0.434	1.05 (0.87-1.26)	0.64

19	*HNF4A*	*rs2144908	G/A	0.452	0.452	1 (0.83-1.2)	0.978
20	*HNF4A*	rs2071199	T/C	0.147	0.151	0.98 (0.75-1.27)	0.855
21	*HNF4A*	*rs745975	G/A	0.165	0.175	0.93 (0.7-1.25)	0.632
22	*PPARA*	rs881740	A/G	0.124	0.122	1.02 (0.78-1.34)	0.866

* literature significant SNPs.

[a] Minor alleles were defined according to the frequencies in cases shown in Appendix 2.

Appendix 6: T2D association results (additive model) of 22 SNPs for stage 2 case-control samples subset by metabolic syndrome status in cases

				Minor Allele Frequency			OR (95% CI)	*P*	OR (95% CI)	*P*
SNP no.	Gene	rs number	Major/Minor allele [a]	All Controls (N = 310)	MetS negative T2D (N = 417)	MetS positive T2D (N = 578)	MetS negative T2D		MetS positive T2D	
1	*PBX1*	rs3767374	C/T	0.212	0.202	0.221	0.94 (0.7-1.28)	0.712	1.06 (0.79-1.41)	0.706
2	*PPARG*	rs2972164	C/T	0.044	0.063	0.064	1.48 (0.91-2.4)	0.111	1.5 (0.94-2.37)	0.085
3	*ADIPOQ*	*rs266729	C/G	0.249	0.248	0.258	0.99 (0.78-1.27)	0.946	1.05 (0.83-1.32)	0.693
4	*ADIPOQ*	*rs2241766	T/G	0.278	0.296	0.293	1.09 (0.85-1.41)	0.494	1.08 (0.85-1.36)	0.549
5	*ADIPOQ*	rs6773957	A/G	0.426	0.415	0.47	0.96 (0.77-1.19)	0.678	1.2 (0.98-1.46)	0.079
6	*ADIPOQ*	rs1063539	G/C	0.284	0.29	0.285	1.03 (0.82-1.3)	0.785	1.01 (0.81-1.25)	0.952
7	*PPARGC1A*	rs12650562	T/C	0.479	0.483	0.513	1.02 (0.82-1.26)	0.874	1.15 (0.94-1.41)	0.163
8	*PPARGC1A*	rs1472095	C/T	0.07	0.09	0.074	1.32 (0.89-1.98)	0.17	1.07 (0.72-1.58)	0.743
9	*GCK*	rs13239289	C/G	0.08	0.067	0.072	0.82 (0.55-1.24)	0.347	0.9 (0.62-1.31)	0.581
10	*LPL*	rs291	T/C	0.218	0.231	0.209	1.08 (0.83-1.39)	0.571	0.95 (0.74-1.21)	0.659
11	*LPL*	*rs320	T/G	0.215	0.228	0.206	1.08 (0.83-1.39)	0.569	0.95 (0.74-1.21)	0.651
12	*LPL*	rs4921684	C/T	0.086	0.101	0.089	1.2 (0.81-1.78)	0.366	1.04 (0.71-1.52)	0.836
13	*KCNJ11*	*rs5219	C/T	0.327	0.341	0.311	1.07 (0.85-1.34)	0.58	0.93 (0.75-1.16)	0.516

14	*ABCC8*	*rs4148643	G/A	0.07	0.058	0.067	0.82 (0.5-1.35)	0.436	0.96 (0.59-1.55)	0.856
15	*HNF1A*	rs1169286	A/G	0.43	0.456	0.43	1.11 (0.9-1.37)	0.341	1 (0.82-1.22)	0.998
16	*LIPC*	rs2242062	A/G	0.369	0.337	0.348	0.87 (0.69-1.08)	0.209	0.91 (0.74-1.12)	0.391
17	*HNF4A*	rs4812828	C/T	0.479	0.48	0.492	1.01 (0.81-1.24)	0.958	1.06 (0.87-1.29)	0.597
18	*HNF4A*	*rs1884614	C/T	0.434	0.456	0.437	1.09 (0.88-1.35)	0.408	1.01 (0.83-1.24)	0.896
19	*HNF4A*	*rs2144908	G/A	0.452	0.465	0.442	1.06 (0.85-1.3)	0.62	0.96 (0.79-1.17)	0.679
20	*HNF4A*	rs2071199	T/C	0.151	0.127	0.162	0.82 (0.6-1.12)	0.212	1.09 (0.83-1.44)	0.546
21	*HNF4A*	*rs745975	G/A	0.175	0.167	0.163	0.95 (0.68-1.31)	0.746	0.92 (0.67-1.25)	0.593
22	*PPARA*	rs881740	A/G	0.122	0.107	0.137	0.87 (0.63-1.2)	0.401	1.13 (0.85-1.51)	0.394

* literature significant SNPs.

[a] Minor alleles were defined according to the frequencies in cases shown in Appendix 2.

It is the time I have spent on my rose that makes my rose so important.

—Antoine de Saint- Exupéry, *Le Petit Prince*

CUHK Libraries
004561311