From Department of Molecular Medicine and Surgery
Karolinska Institutet, Stockholm, Sweden

SEARCHING FOR NEW BIOMARKERS OF DIABETIC NEPHROPATHY IN MALAYSIAN SUBJECTS WITH TYPE 2 DIABETES

Norhashimah Abu Seman

Stockholm 2015

Published by Karolinska Institutet.
Printed by Universitets service US-AB.

ISBN 978-91-7549-989-5

Searching for New Biomarkers of Diabetic Nephropathy in Malaysian Subjects with Type 2 Diabetes

THESIS FOR DOCTORAL DEGREE (Ph.D.)

AKADEMISK AVHANDLING
som för avläggande av medicine doktorsexamen vid Karolinska Institutet offentligen försvaras i Rehab salen Norrbacka/Rehabsalen (76pl), Karolinska Universitetssjukhuset, Solna, Fredagen den 12 juni 2015, kl 13:00

Av

Norhashimah Abu Seman

Principal Supervisor:
Associate Professor Harvest F. Gu
Department of Molecular Medicine and Surgery
Karolinska Institutet
Stockholm, Sweden

Co-supervisor(s):
Professor Kerstin Brismar
Department of Molecular Medicine and Surgery
Karolinska Institutet
Stockholm, Sweden

Dr Wan Nazaimoon Wan Mohamud
Institute for Medical Research
Diabetes and Endocrine Unit
Cardiovascular Diabetes and Nutrition
Research Centre
Kuala Lumpur, Malaysia

Professor Claes-Göran Östenson
Department of Molecular Medicine and Surgery
Karolinska Institutet
Stockholm, Sweden

Opponent:
Professor Alexander Maxwell
Nephrology Unit
Department of Medical Science
Queen's University
Belfast, UK

Examination Board:
Professor Anita Aperia
Department of Women's and Children's Health
Karolinska Institutet
Stockholm, Sweden

Associate Professor Leonid Padyukov
Department of Medicine
Karolinska Institutet
Stockholm, Sweden

Associate Professor Maria Svensson
Department of Medicine
Sahlgrenska Akademin
Gothenburg University
Gothenburg, Sweden

For my family with love

ABSTRACT

According to the latest National Health and Morbidity Survey reports in year 2011, the prevalence of type 2 diabetes (T2D) in Malaysia has risen to 20.8%, much faster than expected. Of concern, 57% of T2D patients in Malaysia had diabetic nephropathy (DN). T2D and DN are complex diseases, in which genetic and environmental factors have been reported to play an important role. In this thesis, we collected a total of 1142 blood samples from a cohort of Malaysian subjects, including normal glucose tolerance (NGT) subjects and T2D patients with or without DN and conducted genetic, epigenetic and protein analyses. The objectives of this thesis were to improve our knowledge for better understanding of the pathogenesis of T2D and DN, and to search for new biomarker of the disease.

We investigated two members of solute carrier (SLC) family. SLC12A3 is a sodium/chloride transporter in kidneys. In **Study I**, we found that Arg913Gln polymorphism in the *SLC12A3* gene was associated with T2D and DN in a Malaysian cohort. The allele Arg913 conferred the susceptibility risk to the diseases. We also found that slc12a3 led to structural abnormality of kidney pronephric distal duct at 1-cell stage by knockdown of zebrafish ortholog. Slc12a3 mRNA and protein expression levels were upregulated in kidneys of *db/db* mice from 6, 12, and 26 weeks at the age. This study provided evidence suggesting that SLC12A3 may be a new biomarker for DN. To evaluate this biomarker, analyzing plasma or serum SLC12A3 levels with advanced protein analysis technique and prospective study in the patients with T2D and DN have been taken into our future research plan. SLC30A8 is a zinc efflux transporter and is highly expressed in the pancreas, particularly in alpha, beta and PP cells of the islets of Langerhans. In **Study II,** we replicated the genetic association of *SLC30A8* polymorphisms with T2D and further analyzed DNA methylation of this gene. Results demonstrated that increased DNA methylation levels of this gene were associated with T2D but not with DN. This study thus provided evidence that *SLC30A8* has epigenetic effects in T2D.

Recent research has been implicated that the inflammation may be a key pathophysiological mechanism in DN. Intercellular adhesion molecule 1 (ICAM-1) is an acute phase marker of inflammation. In **Study III,** we found that the *ICAM1* K469E(A/G) polymorphism was significantly associated with DN. The plasma ICAM-1 levels were increased in T2D patients with DN compared with the patients without DN. Among T2D patients with DN, the carriers with heterozygous genotype had higher plasma ICAM-1 levels than both homozygous carriers. Therefore, analyzing ICAM-1 protein according to the *ICAM1* K469E(A/G) polymorphism genotypes may be useful for predicting susceptibility to DN. Pentraxin 3 (PTX3) is an acute-phase glycoprotein and a soluble receptor acting as an opsonin. This protein is expressed in vascular endothelial cells and macrophages and its level may reflect more directly the inflammatory status of the vasculature. In **Study IV**, we investigated the plasma PTX3 levels and found that decreased plasma PTX3 levels were associated with T2D and DN in Malay men.

Taking together, we have analyzed four candidate genes of *SLC12A3*, *SLC30A8*, *ICAM1* and *PTX3* in Malaysian subjects with T2D and DN. Results may provide useful information for better understanding of the pathogenesis of the diseases. We shall add our effort to analyze SLC12A3 protein variation in diabetes patients and to further evaluate this molecule as a new biomarker for DN.

LIST OF SCIENTIFIC PAPERS

I. **Abu Seman N**, He B, Ojala JR, Wan Mohamud WN, Ostenson CG, Brismar K, Gu HF. Genetic and Biological Effects of Sodium-Chloride Cotransporter (SLC12A3) in Diabetic Nephropathy. Am J Nephrol. 2014 40(5):408-416.

II. **Abu Seman N**, Wan Mohamud WN, Östenson CG, Brismar K, Gu HF. Increased DNA Methylation of the SLC30A8 Gene Promoter is Associated with Type 2 Diabetes in a Malay Population. Clinical Epigenetics. 2015 7(1):30.

III. **Abu Seman N**, Anderstam B, Wan Mohamud WN, Östenson CG, Brismar K and Gu HF. Genetic, Epigenetic and Protein Analyses of Intercellular Adhesion Molecule 1 in Malaysian Subjects with Type 2 Diabetes and Diabetic Nephropathy. Submitted manuscript.

IV. **Abu Seman N,** Witasp A, Wan Mohamud WN, Anderstam B, Brismar K, Stenvinkel P, Gu HF. Evaluation of the Association of Plasma Pentraxin 3 Levels with Type 2 Diabetes and Diabetic Nephropathy in a Malay Population. J Diabetes Res. 2013:298019.

Other publications not included in the thesis:

I. Shen C, Sharm M, Reid DC, Celver J, **Abu Seman N**, Chen J, Vasan SK, Wang H, Gu T, Liu Y, Wan Mohamud WN, Shen H, Brismar K, Fairbrother WG, Kovoor A, Gu HF. A Polymorphic Microdeletion in the RGS9 Gene Suppresses PTB Binding and Associates with Obesity. J Diabetes Metab. 2014 5:437.

II. Gu HF, Zheng X, **Abu Seman N**, Gu T, Botusan IR, Sunkari VG, Lokman EF, Brismar K, Catrina SB. Impact of the Hypoxia-Inducible Factor-1 α (HIF-1α) Pro582Ser Polymorphism on Diabetes Nephropathy. Diabetes Care. 2013 36(2):415-21.

III. Seed Ahmed M, Kovoor A, Nordman S, **Abu Seman N**, Gu T, Efendic S, Brismar K, Ostenson CG, Gu HF. Increased Expression of Adenylyl Cyclase 3 in Pancreatic Islets and Central Nervous System of Diabetic Goto-Kakizaki Rats: A Possible Regulatory Role in Glucose Homeostasis. Islets. 2012 4(5):343-8.

IV. Gu T, Horová E, Möllsten A, **Abu Seman N**, Falhammar H, Prázný M, Brismar K, Gu HF. IGF2BP2 and IGF2 Genetic Effects in Diabetes and Diabetic Nephropathy. J Diabetes Complications. 2012 26(5):393-8.

CONTENTS

LIST OF ABBREVIATIONS

CNV	Copy number variations
CKD	Chronic kidney disease
CVD	Chronic cardiovascular disease
DN	Diabetic nephropathy
DNA	Deoxyribonucleic acid
DNMT	DNA-methytransferase
ESRD	End stage renal disease
EWAS	Epigenome wide association study
GWAS	Genome-wide association study
HbA1c	Glycated hemoglobin
ICAM-1	Intercellular adhesion molecule 1
IGT	Impaired fasting glucose
IGF	Impaired glucose tolerance
IGFBP1	Insulin-like growth factor binding protein 1
IR	Insulin resistance
KCNJ11	Potassium inwardly-rectifying channel, subfamily J, member 11
LINE	Long interspersed repeat element
MDRD	Modification of diet in renal disease
miRNA	MicroRNA
mRNA	Messenger ribonucleic acid
NGT	Normal glucose tolerance
NHMS	National Health Morbidity Survey
OGTT	Oral glucose tolerance test
PCR	Polymerase chain reaction
PLA	Proximity ligation assay
PPARG	Peroxisome proliferator-activated receptor gamma
PTX3	Pentraxin 3
RCA	Rolling circle amplification
RNA	Ribonucleic acid
rRNA	Ribosomal ribonucleic acid
SINE	Small interspersed repeat element
SLC12A3	Solute carrier family 12 (sodium/chloride transporters) member 3
SLC30A8	Solute Carrier family 30 member 8
SAT	Subcutaneous abdominal adipose tissue
SNP	Single nucleotide polymorphism
TCF7L2	Transcription factor 7-like 2
tRNA	Transfer ribonucleic acid
VAT	Visceral adipose tissue

1 BACKGROUND

1.1 DIABETES AND DIABETIC NEPHROPATHY

1.1.1 Type 2 Diabetes Mellitus in Malaysia

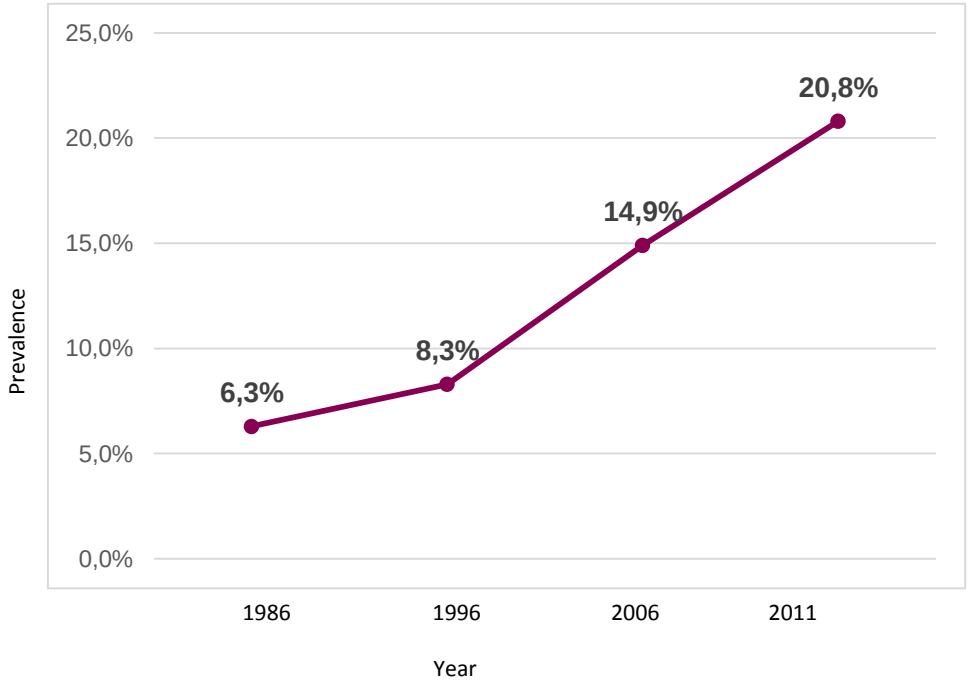
Type 2 diabetes (T2D) is a complex disorder caused by the combined effects of genetic inheritance and environmental factors. T2D accounts for most cases of diabetes in Malaysia. It is characterized by chronic hyperglycaemia due to impaired insulin secretion by pancreatic beta cells as well as insulin resistance in the target tissues such as liver and skeletal muscle. Individuals can experience different signs and symptoms of diabetes. The common symptoms of diabetes include frequent urination, excessive thirst, weight loss, tiredness and tingling sensation or numbness in the hands or feet [1].

Diabetes is primarily diagnosed based on plasma glucose levels at fasting condition and/or 2 hours after taking a 75g oral glucose tolerant test (OGTT). The individual is considered to have diabetes if fasting plasma glucose is ≥7.0 mmol/L or 2h-OGTT postload glucose is ≥11.1 mmol/L [2]. Since year 2009, glycated haemoglobin (HbA1c) has been added into diabetes diagnostic criteria [3]. HbA1c is an indicator of average blood glucose level for approximately the past 120 days and individual with HbA1c ≥6.5% (48 mmol/mol) is having diabetes. Pre-diabetes is a condition when individual can be said to have either impaired fasting glucose (IFG) and/or impaired glucose tolerance (IGT). Individual with IFG presents with fasting plasma glucose in the range between 6.1 to 6.9 mmol/L according to WHO definition [2] or between 5.6 to 6.9 mmol/L according to ADA [4] while one is said to have impaired glucose tolerance (IGT) when 2h-OGTT postload glucose ranged between 7.8-11.0 mmol/L (WHO 2006) or HbA1c between 5.7-6.4% [5]. Pre-diabetes means that a person is at increased risk for developing T2D within a decade unless they adopt a healthier lifestyle that includes weight loss and more physical activity.

According to the International Diabetes Federation (IDF) Diabetes Atlas, in 2013, 382 million people worldwide had diabetes. This number is expected to rise to 592 million by 2035 [6]. The data were represented 130 countries including Malaysia. T2D prevalence rate in Malaysia has risen much faster than expected. The first National Health and Morbidity Survey (NHMS)

conducted among adults aged above 30 years old in year 1986 reported a prevalence of T2D of 6.3%. The prevalence had risen to 8.3% in the second NHMS 10 years later [7]. According to the latest NHMS reports in year 2011 [8], the prevalence of T2D in Malaysia was 20.8% compared to 14.9% in year 2006 (Figure 1).

Figure 1 Prevalence of diabetes mellitus in Malaysia (>30 years age group) Data from the National Health and Morbidity Surveys



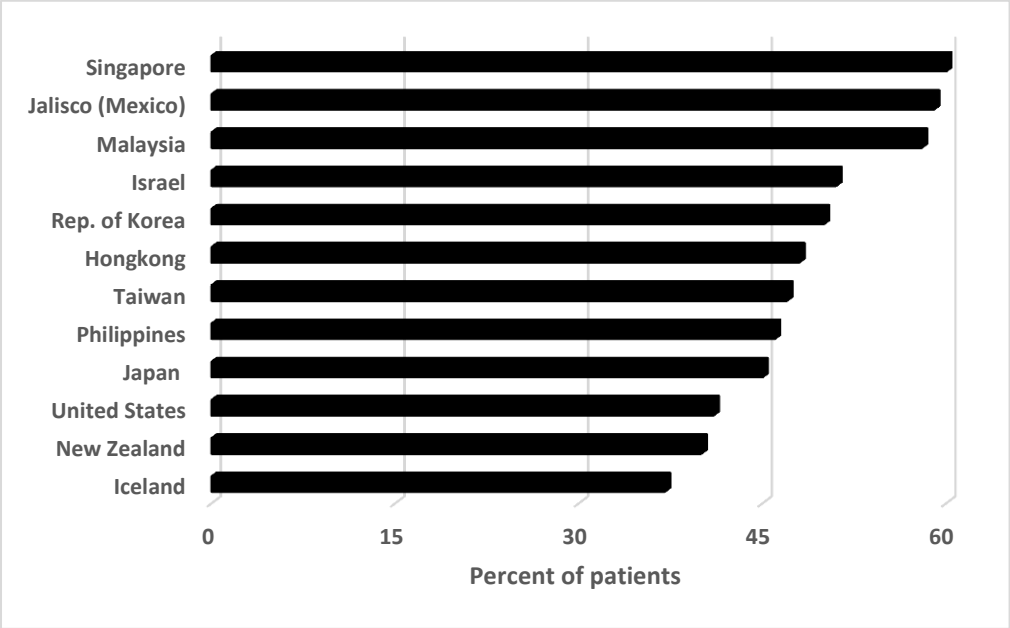
1.1.2 Diabetic Nephropathy in Malaysia

Diabetic nephropathy (DN) is a syndrome characterized by abnormal urine albumin excretion (>30 µg/mg creatinine or >30 mg/24 hr or >20 µg/min) or diabetic glomerular lesions with loss of glomerular filtration rate (GFR) with and without albuminuria. The disease is the most common cause of end-stage renal disease (ESRD) and contributes to 57% of patients with T2D in Malaysia. Although T2D represents a preventable and treatable cause of ESRD, the number of ESRD cases caused by T2D has increased and accounts for more than 50% of incident dialysis patients [9],[10]. The increasing of DN patients in Malaysia are parallel with high prevalence of diabetes, hypertension, the poor glycemic and blood pressure control [11]. The public burden

from diabetes and DN in this nation is massive.

Also in the United States, DN is the most common cause of ESRD, accounting for nearly 50% of the cases. The economic burden of ESRD has increased, reaching a current estimate of USD 17.9 billion, an increase of 7.2% from 1998. Data from United States Renal Data Systems (USRDS) indicate that the incidence of ESRD attributed to diabetes mellitus has increased more than 10 fold over the past two decades [12].

Figure 2 Percentage of incident patients with ESRD due to diabetes, 2011 according to US Renal Data Systems *Figure 12.4 (Volume 2)*



DN is associated with increase morbidity and mortality in cardiovascular disease, already in the early phase of the disease, long before ESRD. While the pathophysiology of DN is incompletely understood, some factors contributing to the disease have been suggested and identified. Non-modifiable risk factors such as ethnicity, family history and elevated GFR early after diagnosis of T2D have been proven in many studies. Risk factors for developing DN that can be modified are glycemia, hypertension, smoking status, obesity and dyslipidemia. Research has shown that if the patients take good care of the modified risk factors such as glycemic controls, the progress of

developing DN could be delayed.

1.2 BIOMARKERS IN DIABETIC NEPHROPATHY

1.2.1 Biomarkers

Biological markers or biomarkers are any biological features of a biological state. It can be any naturally molecule or gene associated with particular pathological or physiological features of the disease. Biomarkers can be measured or evaluated as an indicator of biological processes. In 1998, the National Institutes of Health Biomarkers Definitions Working Group defined a biomarker as "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention." [13]. A joint venture on chemical safety, the International Programme on Chemical Safety, led by the WHO in coordination with the United Nations and the International Labor Organization, has defined a biomarker as "any substance, structure, or process that can be measured in the body or its products and influence or predict the incidence of outcome or disease" [14].

1.2.2 Current biomarkers of diabetic nephropathy

DN tends to be a progressive disease that often leads to EASD. The clinical problem is that once the disease has become overt, a kidney damage has already occurred, and the opportunity for intervention is limited. The current biomarker used to detect DN is microalbuminuria, the presence of albumin in the urine. Microalbuminuria has been used for screening DN patients by albumin-to-creatinine ration (ACR) in a random spot collection, preferably in the first morning urin or by measurement of albumin in a timed urine collection (24 hours or less). Usually, the screening is performed in type 1 diabetic (T1D) patients with diabetes duration more than 5 years and in all T2D patients. This biomarker has been used clinically as it is less expensive, easy to perform and has been shown to be a good predictor of progressive DN. However, this biomarker is susceptible to false-negative and -positive determinations as a result of variation in urine concentration due to hydration and other factors such as diet or urinary tract infection.

With the presence of microalbuminuria, National Kidney Foundation [15] classification has suggested that the levels of glomerular filtration rate (GFR) may be used to stage chronic kidney disease (CKD). The GFR can be easily estimated using the modification of diet in renal disease (MDRD) equation by calculating it from the results of blood creatinine test, age, gender and race. The stage of CKD can be divided as Table 1. The eGFR is a good test to detect kidney disease. However, it may not be accurate for people younger than 18, pregnant women, very overweight and very muscular peoples.

Table 1. Stages of chronic kidney disease.

Stage	Description	(GFR)
1	Kidney damage with normal kidney function	90 or above
2	Kidney damage with mild loss of kidney function	89 to 60
3a	Mild to moderate loss of kidney function	59 to 44
3b	Moderate to severe loss of kidney function	44 to 30
4	Severe loss of kidney function	29 to 15
5	Kidney failure	Less than 15

1.2.2.1 Serum creatinine

Serum creatinine is excreted unchanged by the kidneys and become one of the indicators for kidney health. Creatinine, majorly being produced by the muscles, is removed primarily by glomerular filtration but also by proximal tubular secretion. When the filtration in kidney is deficient, creatinine levels in the blood will be risen. Serum creatinine levels are also used to calculate the estimated GFR (eGFR). However, serum creatinine is unable to detect the kidney injury at the early stage as it is observed only with marked damage to functioning nephrons. Moreover, serum creatinine levels are influenced by the body muscles, make the creatinine levels different between men and women and false low in old people with decreasing muscle mass. Cystatin C is now recommended in the estimate of GFR. Golden standard is iohexol clearance. Furthermore, increased dietary intake of creatine or protein could increase daily creatinine excretion [16].

1.2.3 Potential new biomarkers of diabetic nephropathy

Despite the current markers that have been used in the clinic, the development of more effective treatment strategies for DN could assist successful therapeutic interventions at the early stage. Sensitive and specific biomarkers need to be carefully considered and their usefulness validated. The ideal renal biomarker should be easy to measure and noninvasive; it should be accurate and highly reproducible; and it should demonstrate high sensitivity, high specificity, be cost-effective in its ability to predict the presence of disease, prognosis of the disorder and progression of the condition [17].

Proteomic approaches have been proposed to offer potential tools for the identification of new biomarkers for the kidney disease. Characterization and validation of these proteomic signatures may represent an important step to the noninvasive early diagnosis of kidney diseases. Biomarkers that have been identified include specific collagen fragments, β2-microglobulin, ubiquitin, proinflammatory cytokines, RBP4, transthyretin, apolipoprotein A1, apolipoprotein C1, and cystatin C [18], [19].

1.3 GENETIC STUDIES OF TYPE 2 DIABETES AND DIABETIC NEPHROPATHY

T2D and DN are complex diseases, which involve interaction between genetic and environmental factors. Many studies have shown that genetic susceptibility plays an important role in the pathogenesis of T2D and DN. Identification of the susceptibility genes and evaluation of their genetic effects will provide useful information for better understanding the pathogenesis and for further developing novel therapeutic approaches for the diseases.

1.3.1 Human genome

The human genome, which is the complete set of genetic information for humans, was completely sequenced in year 2012 by Human Genome Project. The total length of human genome is approximately 3 billion base pairs. Human body cells have 23 paired chromosomes, of which 2 of 46 chromosomes are sex chromosomes x and y, which determine an individual's gender. Human genome content is divided into coding and non-coding DNA sequences. Those sequences that can be transcribed into mRNA and translated into protein were defined as coding DNA, which accounts a small fraction of the genome (<2%). Most of the human genome content

are made up of non-coding DNA that are not used to encode protein. Numerous classes of noncoding DNA have been identified, including genes for noncoding RNA (e.g. tRNA and rRNA), pseudogenes, introns, untranslated regions of mRNA, regulatory DNA sequences, repetitive DNA sequences, and sequences related to mobile genetic elements as well as noncoding RNA (e.g. tRNA, rRNA), and untranslated components of protein-coding genes (e.g. introns, and 5' and 3' untranslated regions of mRNA). Although a major part of the human genome has been sequenced, the functions of the non-coding DNA are still not fully understood compared to the coding DNA.

1.3.2 Single nucleotide polymorphisms

A single nucleotide polymorphism (SNP) is a variation in a single base pair in a DNA sequence. They are the most common source of genetic variation in the human genome. Each SNP represents a difference of nucleotide. They are region, single-copy and repetitive sequences. They are highly abundant and widely distributed in the human genome [20]. SNPs have become standard genetic markers used to identify associated alleles in different diseases. SNPs are distributed across the whole genome including both coding and non-coding DNAs. Non-coding consists of tandem repeats and short and long interspersed elements (SINEs and LINEs). SNPs within the coding regions will not necessarily change the amino acid sequence of the protein that is produced. A SNP which does not change the polypeptide sequence is called synonymous (sometimes also called a silent mutation), and is called non-synonymous if a different polypeptide sequence is produced.

1.4 GENETIC STUDY APPROACHES

In the recent years, different genetic approaches and strategies have been effectively applied to identify risk loci in diabetes and DN. There are three approaches including candidate gene association study, family-based association study and genome-wide association study. Each of these approaches has its advantages, motivations and limitations and will be briefly described.

1.4.1 Candidate gene population association study

The candidate gene association study is the most commonly used approach in genetic study of diabetes and DN, and often based upon the design of population genetic association study (cases

and controls). A case-control study approach compares the frequency of SNP alleles in two well defined groups of individuals; cases that have been diagnosed with the disease under study and controls who are known to be non-disease are selected from the population. The candidate genes usually are selected from biological pathways that harbor other previously associated risk loci. In present, many genes have been predicted to be associated with diabetic nephropathy using this approaches such as *SLC12A3* [21] and *APOL1* [22].

1.4.2 Family-based association study

Family based approach can be used to avoid the ethnic variation. The variation of the gene among family members or sibling can be studied if we match each case with the unaffected family members in the family. This approach provides the best matching on ethnicity. However, there is a limitation to use this approach in DN. Compared to other disease like cancer, DN often occurs late in life after 20-30 years of duration; parents are often not available for genetic study. Because of the difficulty in sample collection, this approach with family-based design is rarely used in genetic study of DN.

1.4.3 Genome wide association study

Genome wide association (GWAS) study is a hypothesis free approach to investigate the association between common genetic markers and disease. In the recent years, GWAS has been successfully facilitated rapid progress in genetic study of T2D. Most of the genetic variants identified by GWAS are common variants with modest effects on T2D. These common genetic variants for T2D loci are shared among different ethnic groups.

1.5 GENETIC STUDIES OF TYPE 2 DIABETES

GWASs have identified about 70 loci that associated with T2D. Most of the identified variants are located in the non-coding regions of the genome, make it difficult to link disease-associated-variants with the pathogenesis of T2D. Based on the result from twin studies, heritability of T2D is estimated to range between 30% to 70% depending on the age of diabetes onset and the glycemic status of cases [23]. Therefore, although GWASs have identified over 70 loci for T2D, the collective effect of those loci accounts only a small part of the heritability of T2D. Therefore,

it is common when genetic studies were conducted in a small samples sizes, detection of variants's effects becomes difficult.

In 2006, the strongest association of common intronic variants was identified as the genetic risk factor for T2D within the transcription factor 7-like 2 (*TCF7L2*) gene and has been replicated in many ethnicities [24]. A bundle of replication studies have approved that the *TCF7L2* variants have a substantially stronger effect on T2D risk than those in *PPARG* and *KCNJ11*, with a per-allele odds ratio of ~1.4 [25]-[28]. *TCF7L2* gene encodes a transcription factor that is active in the Wnt-signalling pathway and that had no 'track-record' as a candidate for T2D. Zhou et al. 2014 reported that compelling evidence for strong involvement of *TCF7L2* into pancreatic beta cell function including insulin production and processing [29]. However, some reports also suggested the possibility that the beta cell dysfunction related to *TCF7L2* was indirect and was instead the consequence of disruptions in liver, brain, or gut [30],[31]. Saxena et al. 2010 reported that in human, the T2D risk allele in *TCF7L2* is markedly associated with a pancreatic phenotype evoking a primary islet dysfunction, and not with insulin resistance or with liver abnormalities as reported previously [32].

Besides the common variants that have been discussed, rare or low frequency variants have been identified. Recently, Majithia et al. 2014 have demonstrated that rare loss-of-function variants in the nuclear receptor encoded by *PPARG* increased T2D risk [33]. The authors identified a total of 52 rare non-synonymous *PPARG* variants that were not associated with T2D, even after variant stratification according to the *in silico* prediction of their putative functional effect. Another GWAS-identified T2D gene found to harbor rare mutations associated with T2D is *SLC30A8*. The meta-analysis of all these studies showed a significant protection of rare *SLC30A8* variants against T2D with an OR of 0.34 [34].

1.6 GENETIC STUDIES OF DIABETIC NEPHROPATHY

Not every T2D patients develop DN (only 30-40%), suggesting that genetic factors play a role. Various risk factors like, diabetes duration, hyperglycemia, increased blood pressure, obesity, hyperlipidemia and genetic alterations may influence an individual to develop DN in the future. Recently, accumulative findings have been proven that apart from the above risk factors, there is

a strong association between an individual's genetic make-ups in his predisposition to DN. The first genome wide study using linkage approach has used multigenerational families' members or compared pairs of siblings to explore DN. This approach is suitable for study of small to moderate effect size as it is challenging to collect the subjects from a multiple generation because of the disease always appears at the late age. By using this approach, a few possible candidate genes/variants have been recognized. Furthermore, diabetic siblings of patients with ESRD due to diabetes are known to be at 5-fold higher risk of ESRD compared with those without family history [35], [36]. Other study on diabetic siblings of ESRD patients due to diabetes have a higher frequency of albuminuria (46%), suboptimal BP control (65%), suboptimal glycemic control (HbA1c >7.0%: 43%), smoking (26%), and failure to receive RAAS-modifying agents (42%) [37].

Although several candidate genes have been identified, replication of the genetic association in the different ethnic populations and functional analyses are tough. This is because many of the candidate genes reported earlier were based on the limited sample sizes. A large samples size is required to study the association between the studied genes and DN.

1.7 EPIGENETIC STUDIES OF TYPE 2 DIABETES AND DIABETIC NEPHROPATHY

1.7.1 Basics of epigenetics

Genetic factors alone unable to explain the involvement of additional factors that contribute to the disease development. Recently, epigenetic effects have been introduced to explain the causal link between genetics and environmental exposures. It refers to heritable changes in gene expression that does not change the DNA sequence in the genome. The study of epigenetics focuses on the mechanisms by which the environment interacts with the genotype to produce a variety of phenotypes by either modified the chromatin structure or control the mRNA translation [38], [39]. At least three systems including DNA methylation, histone modification and non-coding RNA (ncRNA)-associated gene silencing are currently considered to initiate and sustain epigenetic changes.

The mechanisms on how epigenetic effects could regulate the gene expression are still not fully understood. However, increased methylation levels of DNA is generally associated with silencing of a gene whereas demethylation of DNA is associated transcriptional activation thus increased the gene expression. These mechanisms can be active during intrauterine and early postnatal development, as well as throughout adult life [40]. Epigenetic regulation could have both short-term (non-heritable) and long-term (heritable) effects. The short-term epigenetic mechanism enables cells to quickly respond to changing environmental factors [41]. Meanwhile, the long-term epigenetic effects are formed in response to long-acting environmental stimuli and can be transferred as memory to offspring cells [42].

1.7.2 DNA methylation alteration

DNA methylation occurs at the cytosine bases of eukaryotic DNA, where cytosine bases are converted to 5-methylcytosine by DNA methyltransferase (DNMT) enzymes by adding a methyl (CH_3) group. DNA methylation is conducted by the methyltransferase (DNMTs) family encompassing DNMT1, DNMT3A and DNTM3B [43]. In mammalian cells, DNMT1 is required to maintain methylation during replication, whereas DNMT3A and DNTM3B provide de novo methylation patterns [44].

It is the best-studied epigenetic modification and governs transcriptional regulation and silencing [45]. In mammals, DNA methylation is found distributed in definite CpG sequences throughout the entire genome, with the exception of CpG islands, or certain stretches (approximately 1 kilobase in length) where high CpG contents are found. About 75% of human gene promoters are associated with CpG islands [45], [46] which are clusters of 500 bp to 2 kb length with a comparatively high frequency of CpG dinucleotides.

1.7.3 DNA methylation analyses in type 2 diabetes and diabetic nephropathy

Several studies have been performed in the specific tissues of pancreatic islets. Epigenome wide association study (EWAS) in T2D has been performed in human pancreatic islets by Dayeh et al. 2013 [47]. They found low degrees of methylation in the genomic regions that located close to the transcription start site. Meanwhile higher degree of methylation was observed in the regions further away from the transcription start site such as the gene body, 3'-UTR and intergenic

regions. They also identified 1,649 CpG sites and 853 genes, including *TCF7L2*, *FTO* and *KCNQ1*, with differential DNA methylation levels in T2D islets after correction for multiple testing. Previously, Volkmar et al. 2012 uncovered 276 CpG loci affiliated to the promoters of 254 genes displaying significant differential DNA methylation in diabetic islets [48].

Using bisulfite pyrosequencing, our research group has conducted DNA methylation analyses in the *IGFBP1* gene. DNA methylation levels of the *IGFBP1* gene were found to be increased in T2D patients compared to NGT controls. IGFBP-1 serum circulating levels were found to lower in T2D patients [49].

A EWAS performed for CKD using blood DNA samples demonstrated that the differential methylations were associated with CKD in the cutlike homeobox 1, *ELMO1*, FK506-binding protein 5, protein tyrosine phosphatase receptor type N polypeptide 2, and *PRKAG2* genes [50]. Studies of DNA methylation profiles in genomic DNA of diabetic patients with or without DN also revealed differential methylation changes in several genes, including *UNC13B*, which has been suggested to mediate apoptosis in glomerular cells as a result of hyperglycaemia [51]. Furthermore, global DNA hypermethylation of peripheral blood leukocytes from patients with CKD found to be associated with inflammation and increased mortality [52].

1.8 ETHNIC AND GENDER DIFFERENCES

1.8.1 Ethnic difference

Many studies have shown that different ethnic groups have different risks to develop T2D. In Malaysia, Malay, Chinese and Indians live together in one country. The prevalence of known T2D: Indians had the highest prevalence of 19.9% followed by Malay 11.9% and Chinese 11.4% [7]. Studies in young children showed that African Americans, Latino and Native Americans children had the highest rates of T2D compared to other ethnicities [53], [54]. Similar to African Americans, Native Americans adults have robust insulin in response to glucose, but produce lower insulin sensitivity [55]. These make them to have a higher risk for developing T2D.

In USA, the risk for developing DN is much higher in African-American, Native American, and Hispanic compared with white population. The cumulative incidence of ESRD in Pima Indians

was 61%, at 15 years after the onset of proteinuria compared to 17% in Caucasian with T2D [12].

Obesity is one of the risk factor for developing T2D. Increased subcutaneous abdominal adipose (SAT) and visceral adipose tissues (VAT) have been shown to have positive associations with fasting insulin levels and markers of insulin resistance [56]. Unfortunately, these do not explain why African Americans are more resistant to T2D compared with Caucasians and Latinos despite having the lowest levels of VAT [57]. Compared to Caucasian, African Americans and Latino adults, Asians have the largest accumulation of VAT and deep SAT with increasing adiposity, which is the risk factor for T2D, regardless the lowest BMI [58]. Cumulative studies have reported that for the same BMI, age and gender, Asians had a higher body fat percentage compared to Caucasians [59]-[61]. These may result in Asians to have a higher predisposition to insulin resistance at lower degree of obesity compared to Caucasians.

1.8.2 Gender difference

Gender has been identified as one of the risk factors for T2D and DN. Good glycemic control is one of the key to delay the development of T2D complications. Hormones for regulating glycemic control are affected by gender and body type. Sex hormones regulate not only sex characteristics and fertility, but also metabolism and adipose tissue [62]. In men, low testosterone levels have been reported to be associated with abdominal obesity and insulin resistance [63] which are independent risk factors for developing T2D [64]. Compared to men, women with increased androgen levels have higher insulin resistance [65] and increased risk of T2D [66] cause by reduced the glucose uptake [67] and increased lipolysis [62]. Low circulating levels of IGFBP-1 have been shown to be associated with insulin resistance, T2D and metabolic syndrome.

In Malaysia, based on NHMS III report and several other studies [52], [9], [10] a higher prevalence of T2D and DN was found in males. DNA polymorphism in the genes of sex-determining region Y-box 2, angiotensin II type 1, type 2 receptors also have reported to be associated with DN with gender-specific effects [68]-[70]. Nearly all degrees of nephropathy are

more common in men than women in many populations. One study found albuminuria as stronger risk factor for CVD among men with T2D compared with women [71].

2 AIMS

2.1 GENERAL HYPOTHESIS

We hypothesized that genetic polymorphisms, epigenetic alterations as well as protein variations in the candidate genes (*SLC12A3, SLC30A8. ICAM-1* and *PTX3*) may be associated with T2D and DN.

2.2 AIMS

The overall aim of this thesis is to search for new biomarkers for DN in Malaysian subjects with T2D.

The specific aims for each study included in this thesis are as follows:

i. To determine the genetic effects of solute carrier family 12 (sodium/chloride transporters) member 3 (*SLC12A3*) in Malay subjects with T2D and DN.

ii. To investigate the epigenetic effects of solute carrier family 30 member 8 (*SLC30A8)* in Malay subjects with T2D.

iii. To analyse the genetic polymorphism and protein of intercellular adhesion molecule 1 (*ICAM-1*) in T2D and DN among the Malaysian population.

iv. To evaluate the association of plasma pentraxin 3 (PTX3) levels in Malay subjects with T2D and DN.

3 MATERIALS AND METHODS

3.1 SUBJECTS

3.1.1 Human subjects- Collection of Malaysian cohort

Malaysia is located in South East Asia and consists of two geographical regions divided by the South China Sea, with multi-ethnic populations. With the population 30 million, Malaysia consists 67.4% of bumiputera (Malay 63.1% and aborigines 4.3%), 24.6% Chinese, 7.3% Indians and 0.7% others [72].

The ethnic distribution of our studied subjects are 67.6% Malay, 15.3% Indians, 14.8% Chinese and 2.3% Indigenous Sabahans and Sarawakians. In this thesis, a total of 1142 (551 males/591 females) blood samples were collected from Malay subjects with normal-glucose tolerance (NGT), patients with T2D with and without DN from collaborating centres all over Peninsular Malaysia.

Diagnosis of T2D was done based on World Health Organization (WHO) criteria [2], while diagnosis of DN was based on urine albumin-to-creatinine ratio (ACR) suggested by ADA [4]. All subjects gave informed written consent prior to the study, answered a set of questionnaires and underwent clinical and physical examinations. Subjects fasted for 10-12 hours prior to study visit. For those subjects who claimed not to have diabetes at the start of study, the oral glucose tolerance test was done. Blood and urine samples were collected from all subjects and stored in -70ºC or -20ºC freezers until used.

3.1.2 Animal models

Choosing the correct and suitable animal model is important, in order to understand the molecular basis of various factors that involved during the development of T2D and DN. In this thesis, we used *db/db* mice and zebrafish as a model for diabetic nephropathy. All animal experiments were approved by the local ethics committees.

3.1.2.1 *db/db* mice

The *db/db* mouse was identified initially in 1996 in Jackson Labs. It was an obese mouse that was hyperphagic soon on weaning [73]. The diabetic gene (*db*) is transmitted as an autosomal

recessive trait. The *db* gene is encodes for a G-to-T point mutations of the leptin receptor, leading to abnormal splicing and defective signaling of the adipocyte-derived hormone leptin [74]. Hyperinsulinemia is shown after 10 days of age and blood glucose levels are elevated at 1 month old of age [75]. The phenotypes of *db/db* mice are similar to T2D in humans such as obesity, insulin resistance and diabetes. These mice have been used as DN model since it develops progressive renal histologic changes and functional derangements. The *db/db* mice showed renal hypertrophy, glomerular enlargement, albuminuria and mesangial matrix expansion, which are the similar characterizations of DN in human [76].

3.1.2.2 Zebrafish

Zebrafish has been used as a model organism for several pathophysiological conditions which are related to human diseases. A part from that, zebrafish has also been used for drug screening, in tumor biology, systems biology and in infection research [77]. Recently, zebrafish genome has been sequenced which made zebrafish to be increasingly used as model human disease particularly in metabolic diseases. Zebrafish's unique characteristics such as short generation intervals, transparent embryos offers unique imaging opportunities. For genetic study, zebrafish genome can be manipulated using morpholinos, mutant or transgenic fish lines, which made zebrafish becomes one of the most important models in developmental biology.

In kidney disease research, the simple nature of the zebrafish pronephric kidney makes it as a suitable system to study the early developmental events compared to other complex kidneys. The molecular and segmental organization between human and zebrafish nephrons also was found to be similar.

3.2 METHODS

3.2.1 TaqMan allelic discrimination

TaqMan allelic discrimination is an optimized method to genotype SNP. The 5' nuclease assay is used for amplifying and detecting specific SNP alleles in purified genomic DNA samples. Two TaqMan minor groove binder (MGB) probes are used to target two different alleles. Each probe has oligonucleotide with a 5'-reporter dye and a 3'-quencher dye. Exonuclease cleavage activity

of an allele-specific 5'-reporter dye during extension produced higher fluorescence intensity. The intensity of the fluorescence later is determined by laser detection.

Genomic DNA extraction was isolated from fresh peripheral blood samples using a DNeasy Blood and Tissue Extraction Kit according to the manufacturer's protocol (Qiagen, Hilden, Germany) and quantified using a spectrophotometry (Biophotometer Plus, Eppendorf, Germany)**.** All SNP genotyping assays were designed and purchased from Life Technologies (Grand Island, USA). Genotyping experiments in the present study were performed in ABI 7300 sequence detector with a Taqman allelic discrimination protocol (Applied Biosystems, Foster City, USA). DNA samples were distributed randomly across plates with cases and controls for genotyping quality control. All PCR reactions were run in 20 µl volumes using 10-20 ng genomic DNA. Millipore water was used as negative controls (blanks) on each plate.

3.2.2 Bisulfite pyrosequencing

Bisulfite pyrosequencing is becoming a common technique to measure DNA methylation levels right down to the single base. Pyrosequencing methylation analysis of CpG sites is a sensitive and accurate protocol [78], [79]. This method is suitable for DNA methylation analysis of single gene loci and relatively cost- and time-effective.

First, DNA sample was treated with sodium bisulfite using EpiTect Bisulfite kit (Qiagen). At this step, un-methylated cytosine was converted to uracil, whereas the methylated cytosine was remains unchanged. This generates a DNA strand that is differentiable upon following sequencing.

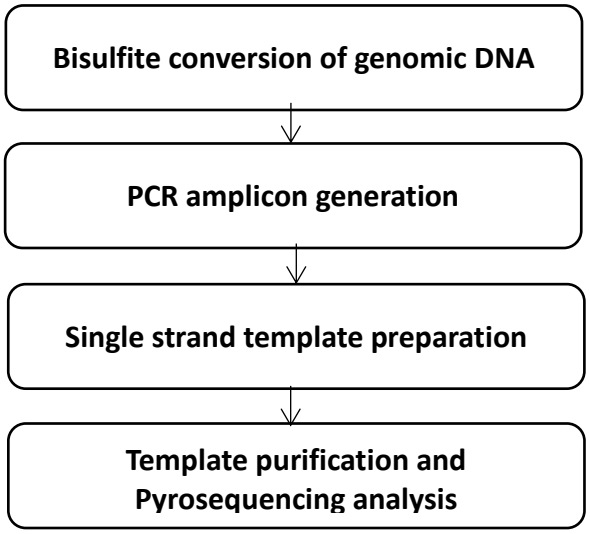
Bisulfite converted DNA then was purified. The target DNA was amplified using PyroMark CpG assay purchased from Qiagen and PyroMark Gold Q96 Reagents kits (Qiagen) in a PyroMark Q96 system (Biotage AB, Uppsala Sweden). The size of the amplicon is between 150-250 bp.

One of the primer (forward or reverse) is tagged at the 5' end with a biotin label. During pre-pyrosequencing preparation steps, streptavidin-coated sepharose beads are added to the PCR product to capture the biotin labeled product. The product was then released into the mixture of

annealing buffer with pyrosequencing primer for sequencing step. The DNA is denatured to produce single stranded molecule for sequencing.

Methylation levels of these CpG sites were detected by using the PyroMark Q96 ID Pyrosequencing System (Biotage, Uppsala, Sweden). PyroQ-CpG software (Biotage) was used for methylation data analysis. The unmethylated and unconverted DNA samples from Qiagen were used for control of conversion efficiency in bisulfite treatment and accuracy in methylation analyses.

Figure 3 The procedure of Bisulfite pyrosequencing DNA methylation



3.2.3 Real time RT-PCR

To exact total mRNAs, homogenization of kidney tissues was prepared with a Mini Beadbeater according to the protocol that we developed in our laboratory. Total mRNAs from kidney tissues in each animal were extracted by using an RNeasy Mini Kit (Qiagen). The integrity and concentration of mRNA samples were assessed by 1.2% agarose gel electrophoresis and determined spectrophotometrically in using NanoDrop ND2000 (Thermo, Uckfield, UK). Based upon the templates of mRNA samples, cDNAs were synthesized with QuantiTect Reverse Transcription Kit (Qiagen). TaqMan real time RT-PCR experiments were performed with a

standard protocol using the ABI7300 real-time PCR system (Applied Biosystems, Foster City, USA). The experiments were replicated on two or three occasions.

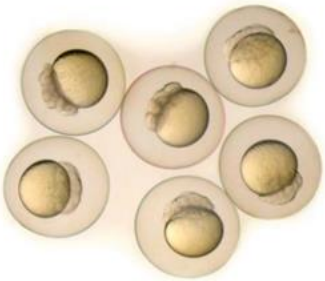
3.2.4 Immunostaining

To detect slc12a3 protein expression levels in kidneys of *db/db* and control mice, we used an immunostaining protocol. Freshly collected tissues/kidneys were embedded into OCT and stored at -80^{o}C until sectioning. One hour prior to sectioning the tissue blocks were brought to -20^{o}C and cut into (5-15 µm) cryosections placed on a frosted glass. After drying the sections for 30 min at room temperature they were fixed with ice-cold aceton for 20 min. Aceton was evaporated and the sections were hydrated for 10 min with PBS after which they were blocked with 5 % goat serum in PBS for 30 min at room temperature. The sections were stained with rabbit anti-SLC12A3 Abcam (1:200); polyclonal antibodies for 1 h at +37°C in the blocking buffer. Goat anti-rabbit Alexa Fluor 488 -conjugates (Life technologies, Foster City, USA) were used as secondary antibodies (1:1000), and incubated with the sections together with a nuclear DAPI (1:2000) staining for 30 min at 37°C. The sections were mounted with (Dako) fluorescent mounting medium and analyzed with conventional fluorescence microscopy (Leica DM RB).

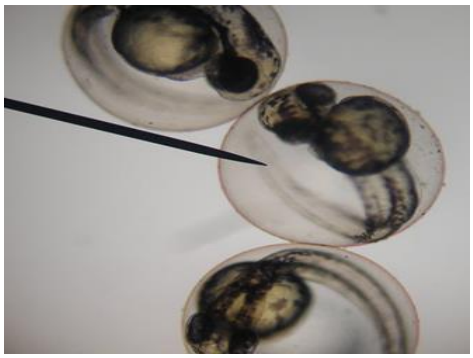
3.2.5 Analysis of *slc12a3* using zebrafish system

Zebrafish model was used to study role of *slc12a3* in the development of kidneys. Experiment was done in the core facility at Karolinska Institutet. The *slc12a3* gene in zebrafish was knocked down using morpholino antisense oligonucleotide (MO) from Gene Tools (Philomath Oregon USA). The MO targeting the exon 3 of the *slc12a3* gene. The MO with 350 µm was injected into the yolk of one- to two-cell embryos from Tg (cdh17:mCherry) line. The injected zebrafish was analysed using a fluorescence microscopy (Leica TCS SP8, Germany).

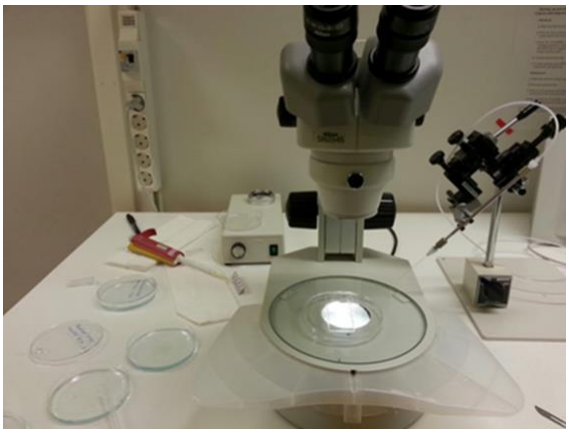
Figure 4 The procedures of slc12a3 analyses with zebrafish system



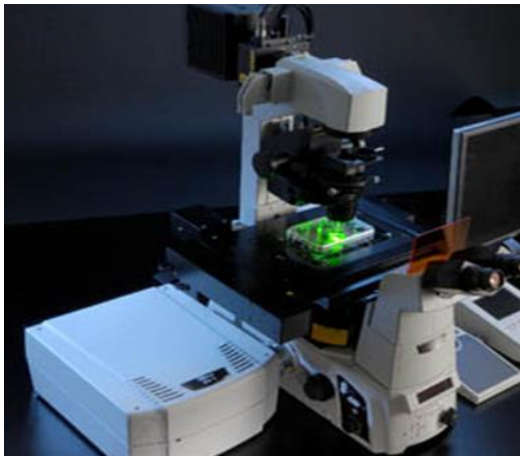
1. Zebrafish embryos were maintain in the KI core facility



3. Day 2: Injection of 350 μm MO into zebrafish's yolk followed by 24 hour incubation



2. Day 1: Selection of zebrafish embryos



4. Day 4: Imaging and analysing using fluorescence microscopy

3.2.6 ELISA

Plasma PTX3 and ICAM-1 concentrations were determined using a commercial enzyme-linked immunosorbent assay kit (R&D Systems Inc., Minneapolis, USA). Experiments were carried out according to the manufacturer's instructions.

Briefly, 20 µl of standard and plasma samples were assayed duplicate in the micro-titer plate wells coated with a specific monoclonal antibody followed by incubation at room temperature for 2 hour. The wells were then washed four times with a buffered surfactant solution. Anti-

protein of interest polyclonal antibody conjugated to alkaline phosphatase was added to each well and incubated for two hours at room temperature. After washing step, 200 µl of substrate solution was added to each well followed by incubation for 30 minutes at room temperature. The solution of 2N sulfuric acid was added to each well to stop the reaction. Absorbance was measured at 450 nm with corrections set at 540 nm using micro-plate reader. The values of plasma protein levels were extrapolated from a curve drawn using a standard.

3.3 BIOINFORMATICS

In medical research, bioinformatics is used to the management and analysis of biological data. Transcription Element Search System (TESS) and Protein Structure Prediction Server (PS 2) are automated homology modelling servers that can be used for conducting analyses of protein structures [80]. The methods were used in both template selection and target- template alignment [81]. A final three-dimensional structure of protein is built using the modeling package. In this thesis, these two programs were used to predict the possibly functional differences of SLC12A3 protein caused by the Arg913Gln polymorphism. The SLC12A3 protein structures were analysed by adding amino acid sequences with either Arg or Gln at the position of 913 into the programs. Further analyses of clinical parameters including eGFR and serum creatinine according to the genotypes of SLC12A3 Arg913Gln polymorphism in NGT, T2D with and without DN were performed.

3.4 STATISTICS

Allele frequency and genotype distribution of the studied SNPs were tested for Hardy-Weinberg equilibrium (HWE). For differences between NGT subjects and T2D patients, two models were tested comparing either allele frequencies in 2x2 contingency tables (dominant) or genotypes in 3X2 contingency tables (additive). Cochran Armitage test was used for detection of trends. Odd ratios (OD) and 95% confidence intervals (CI) were calculated to test the relative risk for association. Tests for association among genotypes and quantitative traits were performed by using ANOVA and Tukey's HSD (Honestly Significant Difference) post-hoc test. Data were given as the means±SD. All p-values were two-tailed and the values less than 0.05 were considered significant. Statistical calculations were performed by using PASW Statistic Base 22.2 (SPSS Inc, Chicago, Illinois, USA).

4 RESULTS

4.1 STUDY I

Genetic and Biological Effects of Sodium-Chloride Co-transporter (SLC12A3) in Diabetic Nephropathy

Association of SLC12A3 Arg913Gln polymorphism with diabetic nephropathy

The association of SLC12A3 Arg913Gln polymorphism with T2D and DN was studied in the Malaysian population. The frequency of 913Gln allele in SNP rs11643718 was 12.3% in the Malaysian population. This polymorphism was found to be associated with reduced risk in T2D (P = 0.028, OR = 0.772, 95% CI = 0.612–0.973) and DN (P = 0.038, OR = 0.547, 95% CI = 0.308–0.973). Furthermore, meta-analysis of the present data (Malaysian) and previous genetic studies in Japanese, Koreans and Americans Caucasians populations showed an association with reduced risk for DN in T2D (Z-value = –1.992, P = 0.046, OR = 0.792, 95% CI = 0.629–0.996).

Clinical Parameters in the patients according to the genotypes of SLC12A3 Arg913Gln polymorphism

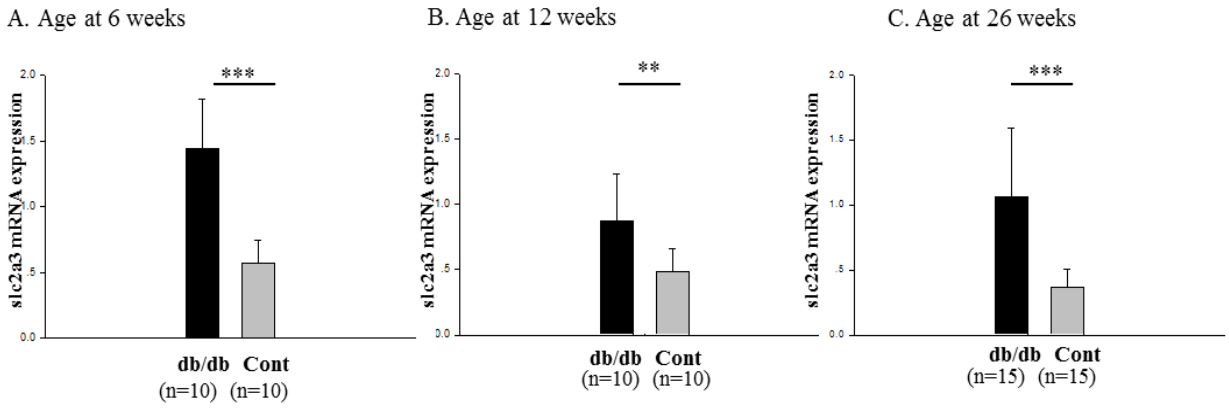
SLC12A3 Arg913Gln polymorphism is a non-synonymous variant where the amino-acid change from Arginine to Glutamine. First, we analyzed the changes of SLC12A3 protein structures with bioinformatics tool [80], [81]. The images implicated that the protein structure of SLC12A3 was altered when the mutant allele 913Gln substituted the wild allele Arg913 in amino acid sequences, suggesting that this polymorphism might have a functional relevance. We further analyzed clinical parameters according to the genotypes of *SLC12A3* Arg913Gln polymorphism. Data showed that among T2D patients, the carriers with Gln913Gln genotype had relatively low serum creatinine and high eGFR levels compared with the patients carrying the Arg913Arg genotype. However, the differences were not statistically significant mainly due to high standard deviations.

Up-regulation of the slc12a3 gene expression in kidneys of db/db mice

We used *db/db* mice as an animal model of DN to explore the functional role of *SLC12A3* in DN by study the expression at both mRNA and protein levels in *db/db* mice at age 6, 12 and 26

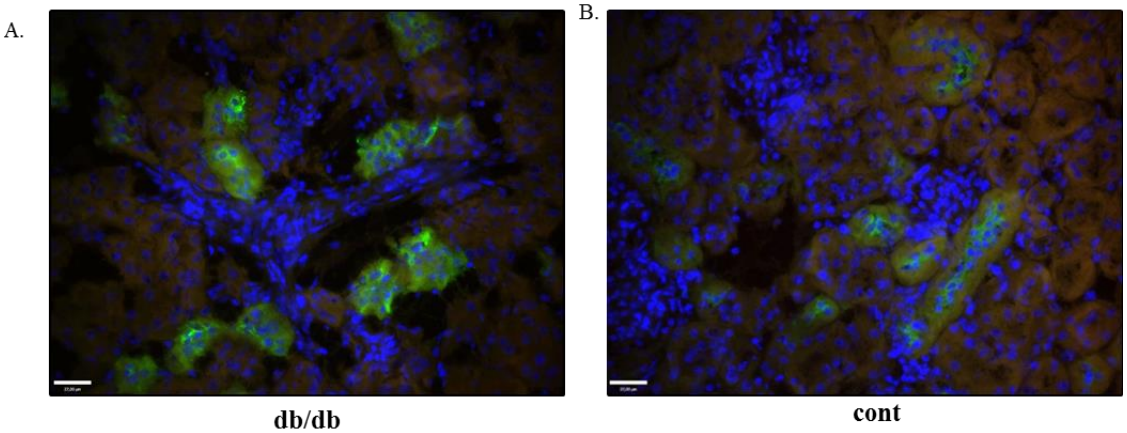
weeks. The *slc12a3* gene in kidneys of *db/db* mice at the ages of 6, 12 and 26 weeks was found significantly over-expressed at mRNA levels compared with the control mice at the same ages (Figure 5a, b and c). Figure 6a and b demonstrated that slc12a3 protein with the stained antibody was distributed in kidney distal convoluted tubule of *db/db* and control mice at the age of 6-weeks. The signal intensity in kidneys of *db/db* mice was significantly higher than that in the control mice.

Figure 5 The slc12a3 gene expression levels were higher in the kidney of db/db compared to control mice at age 6, 14 and 26 weeks.



Data are presented as means±SD. *P*-values are <0.001*** or <0.01**.

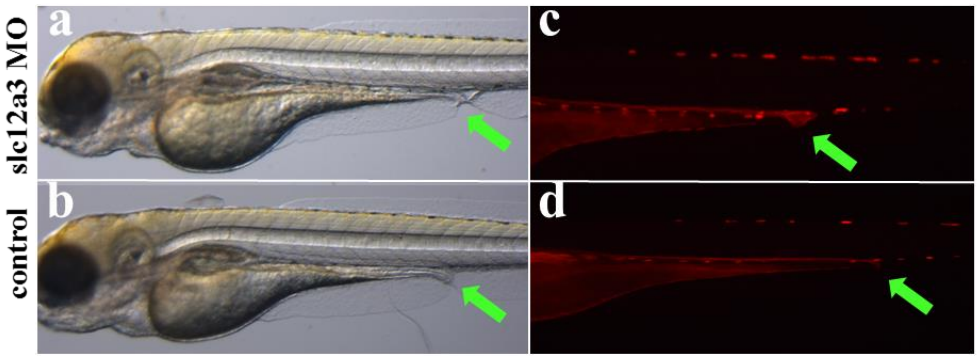
Figure 6 Higher intensities of slc12a3 protein was found in the distal convoluted tubule kidney of db/db compared to control mice at the age of 6 weeks



Role of slc12a3 in zebrafish pronephric duct epithelium

Because the *slc12a3* gene expression at mRNA and protein levels in the kidneys of *db/db* mice at the age of 6-weeks was found to be significantly increased compared with the controls, we were interested to explore whether *slc12a3* plays a role in kidney tubular epithelium. The zebrafish *slc12a3* gene is conserved with 62% of amino acid identity compared with the human. We applied a specific MO-mediated antisense knockdown approach in zebrafish and found that knockdown of zebrafish *slc12a3* did not lead to global alteration of embryonic development compared to the wild-type embryos. Under fluorescence microscopic analysis, however, pronephric duct epithelial structure defined by red signal (mCherry) in the cloacal portion was significantly altered at 4 dpf (Figure 7). The penetrance of this abnormal morphology was 35%. This implicated the importance of *slc12a3* in zebrafish pronephric distal duct, particularly in the cloacal development.

Figure 7 Zebrafish embryos were injected with 350μM slc12a3 morpholino (MO) at the 1-cell stage.



4.2 STUDY II

Increased DNA methylation of the *SLC30A8* gene promoter is associated with type 2 diabetes in a Malay population

Association of SLC30A8 genetic polymorphisms with type 2 diabetes and diabetic nephropathy

In the study II, we first analyzed the association of rs11558471 (A/G) and rs13266634 (C/T) with T2D patients with and without DN. We found that the A allele frequencies of all T2D patients compared to NGT subjects were 0.552 vs. 0.620, *P* = 0.002, OR = 1.334, 95%CI = 1.110-1.602. The results indicated that the A allele of rs11558471 (A/G) was strongly associated with T2D. Moderate association was found when we compared DN subjects with T2D patients without DN 0.593 vs. 0.671, P=0.041, OR=1.399, 95%CI = 1.013-1.932). The association of SNP rs13266634 (C/T) with T2D and DN was not significant (*P* = 0.053, OR = 1.200, 95%CI=0.997-1.443; and *P* = 0.098, OR = 1.313, 95%CI = 0.950-1.815).

Association of SLC30A8 DNA methylation with type 2 diabetes and diabetic nephropathy

In the DNA methylation analyses, only the age-matched NGT subjects and T2D patients were included in order to avoid error caused by ages. We found that the average DNA methylation levels of all 6 CpG sites that located in the *SLC30A8* promoter were high (~78.5%). DNA methylation levels at 5 CpG sites of the gene (except CpG2) in T2D patients were found to be higher than those in NGT subjects, respectively (CpG1 83.9 vs. 81.9%, *P* = 0.031; 82.1% vs. 84.8% CpG3 vs. *P* = 0.003; CpG4 69.6 vs. 66.3%, *P* = 0.001; CpG5 86.2 vs. 83.7%, *P* = 0.004; and CpG6 79.8 vs. 78.1%, *P* = 0.001). Combining all 6 CpG sites together, total mean values of *SLC30A8* DNA methylation levels were significantly increased in T2D patients compared with NGT subjects (79.9%, 95% CI = 79.2-80.5% vs. 77.1%, 95% CI=75.4-78.6%, *P* = 0.002). No significant difference was found when we compared the *SLC30A8* DNA methylation levels between T2D patient without and with DN.

4.3 STUDY III

Genetic, Epigenetic and Protein Analyses of Intercellular Adhesion Molecule 1 in Type 2 Diabetes and Diabetic Nephropathy among a Malay Population

Genetic association of the ICAM1 K469E(A/G) polymorphism with type 2 diabetes and diabetic nephropathy in the Malay population

We found that the *ICAM1* K469E (A/G) polymorphism was associated with T2D (*P*=0.038, OR=1.190 95% CI 1.009-1.404) and DN (P=0.039, OR=1.278 95% CI 1.012-1.614) in the Malay population when Chinese subjects were excluded for the analyses. This polymorphism showed a high heterozygous index in the Malay population but not in Chinese subjects and found to be significantly associated with T2D (*P*=3.0x10-5, OR=2.808 95% CI=1.703-4.630) and DN (*P*=1.7x10-6, OR=2.909, 95% CI=1.857-4.556) in the Malay population.

Plasma ICAM-1 concentrations in Malay subjects with normal glucose tolerance, and with type 2 diabetes without and with diabetic nephropathy

The plasma ICAM-1 levels were significantly increased from NGT (206.9±113.1 ng/ml) to T2D without DN (303.5±113.4) (*P*=0.001). T2D patients with DN had a higher plasma ICAM-1 levels (352.6±156.7) (P<0.001) compared with T2D without DN. All T2D patients without and with DN had higher plasma ICAM-1 levels compared with NGT subjects (P <0.001 both). However, no statistical significance was found between T2D with and without DN. The plasma ICAM-1 levels were found to be elevated from NGT to T2D without and with DN in subjects with BMI <23 kg/m^2. No significant difference of plasma ICAM-1 levels between T2D without and with DN was found (*P*=0.368).

Plasma ICAM-1 concentrations in Malay subjects with normal glucose tolerance, and with type 2 diabetes without and with diabetic nephropathy according to the genotypes of the ICAM1 K469E(A/G) polymorphism

The NGT subjects carrying K469(A/A) genotype were found to have higher plasma ICAM-1 levels compared with the subjects carrying with K469E(A/G) (*P*=0.009) and 469E(G/G) (*P*=0.012) genotypes, respectively. However, there was no significant difference of plasma ICAM-1 levels among the patients without and with DN according to the genotypes of *ICAM1* K469E(A/G) polymorphism.

Detection of the ICAM1 DNA methylation levels in Malay subjects with normal glucose tolerance and, with type 2 diabetes without and with diabetic nephropathy

We found that in all Malay subjects, the average DNA methylation levels of the *ICAM1* gene including 7 CpG sites were low about 3.5%. The DNA methylation levels among these 7 CpG sites were varied between 0.9% and 8.4%. No significant difference was found among subjects with NGT (3.3%), T2D without (3.3%) and with DN (3.7) (*P*=0.398).

4.4 STUDY IV

Evaluation of the association of plasma pentraxin 3 levels with type 2 diabetes and diabetic nephropathy in a Malay population

In this study, we measured plasma PTX3 levels in normal glucose tolerance (NGT) subjects, and T2D patients with and without DN. Plasma PTX3 levels were found to differ significantly between males and females. In males subjects, plasma PTX3 levels were found to be decreased gradually from NGT subjects to T2D patients to DN patients (3.98 vs 2.62 vs 1.63 ng/mL, P = 0.008). No significant difference was found in female subjects.

Furthermore, we analyzed plasma PTX3 levels according to body mass index (BMI). We found inverse correlation between plasma PTX3 levels and BMI only in male subjects with NGT (r = -0.390, P = 0.012) but not in females. The correlation between PTX3 levels and BMI was not found in all male and female T2D patients with or without DN. However, in males with overweight, we found that plasma PTX3 levels were lower in DN patients compared to T2D patient without DN and NGT subjects (1.42 vs 2.60 vs 3.68 ng/mL, P = 0.044).

5 DISCUSSION AND FUTURE PERSPECTIVE

In this thesis, we have performed the genetic and epigenetic studies together with protein analysis of *SLC12A3*, *SLC30A8*, *ICAM1* and *PTX3* in Malaysians subjects with T2D and DN. The results from our studies and further investigations are discussed and summarized as below.

5.1 GENETIC AND BIOLOGICAL EFFECTS OF *SLC12A3* IN DIABETIC NEPHROPATHY

Cumulative epidemiological findings have shown that genetic susceptibility plays an important role in the pathogenesis of DN. GWAS identified *SLC12A3* gene as the new candidates for DN [82]. The Arg913Gln polymorphism in this gene has been reported to be associated with DN in T2D in Japanese, Koreans and American Caucasians populations but with inconsistent conclusions [83], [21], [84]. We thus conducted a genetic study of the *SLC12A3* gene in the Malaysian population (Study I). In 2003, Tanada et al. published the first genetic study of *SLC12A3* and demonstrated that the *SLC12A3* Arg913Gln polymorphism was associated with DN by a protective effect in a Japanese population [83]. Later, Ng et al. found the negative association of the *SLC12A3* genetic polymorphisms with DN in a Caucasian population and thereby explained two possibilities [84]. First, there was a publication bias caused by "winner's curse" in the report of Tanada et al. because there was no direct biological evidence in the literature. Second, the genetic effect of *SLC12A3* may be population specific. In our study, we found that the *SLC12A3* Arg913Gln polymorphism was associated with DN in the Malaysian population with a protective effect. We support the report from Tanada et al. and disagree with the so called "winner's curse". In the recent years, several GWAS of DN have been performed in the Caucasian populations. The results from these GWAS were inconsistent, and no association of *SLC12A3* with DN was reported. Most likely, *SLC12A3* genetic effects are population specific to Asian but not to the Caucasian populations. In addition, Kim et al. reported that the *SLC12A3* Arg913Gln polymorphism was significantly associated with ESRD in a Korean population. But the effect of the 913Gln allele was increased risk for DN [21]. The conclusion from Kim's study is opposite to the data from Tanada's. Kim et al. did not sufficiently explain how to understand the contradiction although they stated that the use of ESRD patients as cases might have led to

the strong survival bias. Whatever the selection that have caused the bias, it may not be sufficient to explain the opposite genetic effects of SLC12A3 in DN. Thus it is still difficult to understand the data from Kim's report.

There was no biological study of *SLC12A3* in DN except our study. In our study, we found that the *SLC12A3* Arg913Gln polymorphism was associated with reduced risk on T2D and DN in the Malaysian population. We also found that patients carrying the Gln913Gln genotype had a relatively low serum creatinine and higher eGFR levels compared with the Arg913Arg genotype carriers. Furthermore, our study in kidneys of the *db/db* mice demonstrated that slc12a3 protein was distributed in the kidney distal convoluted tubule of *db/db* mice at the age of 6-weeks with higher signal intensity compared to control mice. Using zebrafish model to study the role of *slc12a3* in the kidney development, we found that *slc12a3* may affect the pronephric distal duct structure when the *slc12a3* gene activity was knocked down.

The data from previous and our studies suggested that *SLC12A3* is a susceptibility gene to DN in Malaysian population. This molecule may be a new biomarker for DN. Further investigation of plasma SLC12A3 analysis and prospective study in the Malaysian cohort with T2D and DN has been taken into our consideration. At present, however, there is no assay and/or protocol available for plasma or serum SLC12A3 analysis.

5.1.1 Proximity ligation assay and method

Solute carrier (SLC) family is large and there are many homologous members. This may be the main reason why it is difficult to analyze SLC12A3 in plasma and/or serum samples. In recent years, a new technique named as proximity ligation assay (PLA) has been developed to analyze the proteins and to validate biomarkers in plasma and/or serum samples. The attraction of plasma or serum biomarkers lies in their potential to reveal disease processes throughout the body and to guide selection of therapy and follow-up using minimally invasive blood sampling. We shall develop a single PLA assay to analyze plasma SLC12A3 levels in the Malaysian subjects with T2D and DN.

PLA was first described by Fredriksson *et al.* is an immunoassay for detection of protein molecules via DNA ligation and amplification, offering high specificity and sensitivity [85]. In PLA, pairs of affinity probes directed against the same target molecule are modified by attaching short single-stranded DNA molecules, creating so-called PLA probes. Upon proximal binding of a pair of PLA probes to a target molecule, the DNA strands are brought in close proximity and allowed to hybridize to a connector oligonucleotide. The DNA strands can then be joined by enzymatic ligation, forming a reporter DNA molecule. This new DNA sequence can be quantified by sensitive and specific nucleic acid detection techniques, such as quantitative real time PCR (q-PCR). The first form of PLA was a homogeneous-phase assay where the antigen was recognized by DNA aptamers in solution before ligation and amplification with real time detection. The assay has also been performed on solid supports by immobilizing antibodies directly on the walls of PCR tubes [85] or by immobilizing biotinylated antibodies on the surface of streptavidin-coated tubes [86]. The PLA technique has been implemented for a wide variety of applications to visualize proteins *in situ* [87], to reveal infectious agents [86] and protein-DNA interactions [88], and for biomarker detection in both singleplex [89], [90] and multiplex [91], [92].

The PLA approach is a great method to detect proteins present at low concentrations in highly complex mixture of protein. DNA ligation products that can be identified by real-time PCR make it available to measure many different proteins in many samples in a single run [93]. This technique has been applied to measure numerous potential diagnostic targets, such as Aβ protofibrils, suspected to contribute to Alzheimer's disease, in brain homogenates from mice transgenic for a human allele predisposing to Aβ aggregation [94], growth differentiation factor 15 (GDF15) in plasma samples from colorectal cancer patients [95], parvovirus and intracellular bacteria in infected samples [86].

5.1.2 Prospective study of SLC12A3 in diabetic nephropathy

To evaluate *SLC12A3* as a new biomarker in DN, it is necessary to perform a prospective study in the Malaysian population. The prospective study on the effect of *SLC12A3* gene is important to further elucidate its efficiency as a new biomarker for T2D and DN. At the baseline, subjects

will be grouped based on their *SLC12A3* Arg913Gln polymorphism genotype. Over the certain time, subjects will be followed up to see their risk of developing T2D and DN. Furthermore, PLA method could be used to measure the plasma levels of SLC12A3 protein in the study subjects. This type of study allows us to understand whether duration is an independent factor for plasma SLC12A3. Combining all data from the genotyping together with the protein levels, *SLC12A3* may be evaluated as a new biomarker for early diagnosis of DN.

5.2 SLC30A8 DNA METHYLATION CHANGES IN TYPE 2 DIABETES

T2D is a complex metabolic disorder influenced by genetic and environmental factors. In recent years, GWASs have identified a number of confirmed genetic susceptibility variants including *SLC30A8* for T2D. However, GWAS findings can only explain ~10% of the overall heritable risk of T2D, which challenges our expectations to translate genetic information into clinical practice [96]-[98]. One of the reasons causing the missing information on heritability could be that epigenetic factors are involved in the complex interplay between genes and environment. The knowledge regarding epigenetic factors associated with T2D is still limited. Therefore, epigenetic studies may provide further information for better understanding of the pathogenesis of T2D [99]-[101].

In study II, we demonstrated that *SLC30A8* DNA methylation levels were increased in T2D patients compared to NGT subjects. The 6 CpG sites are located in the promoter region of the *SLC30A8* gene. A study on palmitate-treated human islets using genome-wide DNA methylation showed differential *SLC30A8* gene expression and an increased global DNA methylation [102]. Another study using human pancreatic islets suggests that interactions between genetic and epigenetic factors may affect T2D as several risk SNPs that influence islet function such as *SLC30A8* by associated with alternative splicing events in the human islets [103]. In our study, we had no tissue sample of pancreatic islets available for analysis, which was a limitation. However, epigenetic study with blood samples is clinical accessible. A recent report has indicated that both approaches of whole-blood DNA methylation profiling and adipose tissue specific methylation analysis for study of epigenetic changes related to BMI and suggested that

the analysis of blood DNA methylation is worthwhile and can reflect changes in relevant tissues for a phenotype [104].

5.3 *ICAM1* K469E (A/G) POLYMORPHISM GENOTYPE DISTRIBUTION IN THE MALAY POPULATION

In study III, we investigated genetic association of the *ICAM1* K469E(A/G) polymorphism with T2D and DN in the Malay subjects. The data indicated that the major allele A of this polymorphism conferred the risk susceptibility to diabetes and DN. A recent meta-analysis based upon our and other studies has confirmed that the *ICAM1* K469E(A/G) polymorphism affects individual susceptibility to diabetes and diabetic microvascular complications [105]. Particularly, a high heterozygous index of the *ICAM1* K469E(A/G) polymorphism is presented in the Malaysian, Swedish [106] but not in Chinese populations [107].

ICAM-1 protein acts as ligands and the primary receptors for ICAM-1 are integrins, which mediate cell-cell interactions and allow signal transduction. Unlike most integrin-binding proteins, ICAM-1 does not contain an RGD (Arg-Gly-Asp) motif to promote integrin binding, but targets to leukocyte adhesion protein 1 (LFA-1) and Mac-1 [108]. The *ICAM1* K469E(A/G) polymorphism resides in the 5th Ig-like domain of ICAM-1 protein. This domain is essential for dimerization, surface presentation and solubilization of the protein and subsequently plays a crucial role in the activity of ICAM-1 protein in the interaction with LFA-1 and the adhesion of B cells [109]. The *ICAM1* polymorphism K469E(A/G) is non-synonymous and results in a change of the amino acid sequence (from glutamic acid to lysine). Therefore, it is necessary to analyze the circulating ICAM-1 levels according to the genotypes in order to better understand the biological effect of the *ICAM1* K469E(A/G) polymorphism. Analyzing the ratio of the two forms of ICAM-1 protein according to the ICAM1 K469E(A/G) polymorphism genotypes may be useful for predicting susceptibility to diabetes and DN.

5.4 PLASMA PTX3 LEVELS IN TYPE 2 DIABETES AND DIABETIC NEPHROPATHY

In study IV, we reported the gender difference of plasma PTX3 levels in a Malay cohort. Plasma PTX3 levels were found to be decreased in males with T2D and DN. An inverse correlation between PTX3 and body mass index was found in male subjects with NGT. Previously, a group in Japan reported that plasma PTX3 levels are different between males and females in a healthy Japanese population [110]. Conversely, plasma PTX 3 levels were found to be elevated in CKD patients [111]. They found that in CKD patients, plasma PTX3 was a significant predictor of *PTX3* mRNA independent of age, sex and diabetes. They also reported no significant correlations between BMI and plasma PTX3 in both patients and control group. The same observation was found in the study conducted among CKD patients [112]. Therefore, we suggested that PTX3 may have different effects in DN and CKD.

5.5 TISSUE SPECIFIC DNA METHYLATION ANALYSIS

DNA methylation has been extensively studied in medical research and has become a potential biomarker due to its specificity and stability in human samples. They can reflect past environmental exposures, predict disease onset or course, or determine a patient's response to therapy [113]. The analysis of DNA methylation biomarkers could be done in many types of tissue samples including cell-based samples such as blood and tumor cell material and cell-free DNA samples such as plasma. The homogeneity at cellular stage within a tissue samples is the most desired characteristic for a DNA methylation analysis. However, the samples such as blood or even blood fractions such as mononuclear cells which are clinical accessible exhibit cellular heterogeneity [114]-[116]. Fortunately, the variation in DNA methylation that occurs from the different cell-type can be assessed using either differential cell counts [117] or statistical adjustment in post hoc regression models [118]-[120].

In cancer research, DNA hypermethylation analysis of the SEPT9 promoter has been used in the clinic. The detection of DNA methylation levels in the v2 region of the SEPT9 promoter in the blood is sampled in the clinic and followed by analyzed in the laboratory [121]. Association study between DNA methylation and BMI has been performed recently by Dick and his colleagues. They reported the association of BMI and specific HIF3A methylation sites was the

same both in blood and adipose tissues [104]. Those reports support the possibility DNA methylation blood-based analysis to be used as a biomarker for prediction of the disease.

In this thesis, we analyzed the blood DNA methylation levels of *SLC30A8* and *ICAM1* in T2D and DN subjects. Although *SLC30A8* is mainly expressed in the pancreatic islets, based on the reports from other groups indicated that DNA methylation analyses using whole blood is still relevant [121], [104]. However, it will be interesting to measure the *SLC30A8* DNA methylation levels both in the whole blood and pancreatic islets from the same subject. This will help us to further understand the DNA methylation role and the pathogenesis of the disease.

In T2D, the previous reports have showed strong evidence that environmental exposure in early life would affect the progress of the disease later [122]. This opens opportunity for discovering DNA methylation as biomarkers to estimate and predict the risk of developing T2D.

6 CONCLUSIONS

Overall, we have analyzed four candidate genes of *SLC12A3, SLC30A8, ICAM1* and *PTX3* in the Malaysian subjects with T2D and DN. Results may provide useful information for better understanding of the pathogenesis of the diseases.

Study I provided the first biological and further genetic evidence that *SLC12A3* has genetic susceptibility in the development of DN, while the minor 913Gln allele in this gene confers a protective effect in the disease. We shall add our effort to analyze SLC12A3 protein variation in diabetes patients and evaluate this molecule as a new biomarker in DN.

Study II demonstrated that the average DNA methylation levels of the *SLC30A8* gene in the Malay population were at the high levels (~81.4%) and provided the first evidence that increased DNA methylation of the *SLC30A8* gene promoter is associated with T2D but not DN.

Study III showed that the *ICAM1* K469E(A/G) polymorphism in parallel with increased plasma ICAM-1 levels was associated with DN in the Malaysian population and suggested that analyzing ICAM-1 protein according to the *ICAM1* K469E(A/G) polymorphism genotypes may be useful for predicting susceptibility to DN.

Study IV indicated that plasma PTX3 variation had gender difference in the Malay population. The decreased plasma PTX3 levels were associated with T2D and DN in Malays men.

7 ACKNOWLEDGEMENTS

Firstly, I would like to say ‘Alhamdulillah’.

This work was performed at the Department of Molecular Medicine and Surgery (MMK) Karolinska Institutet, Stockholm, Sweden, with financial support from the Stig and Gunborg Westmans Foundation and the Family Erling-Persson Foundation. I am very grateful to the Ministry of Science Technology and Innovative, Malaysia for providing me the scholarship during my Ph.D. study in Sweden. I also would like to thank the Ministry of Health Malaysia and the Institute for Medical Research (IMR) for allowing me to further my education in Sweden. I wish to thank all Malaysian subjects who have volunteered to participate in this Ph.D. study.

In particular, I would like to express my gratitude to the following:

First of all, I would like to thank my main supervisor, **Associate Professor Harvest F. Gu** for making the thesis possible. You have guided me a lot from the day I came to Stockholm until the day I left. Thank you for your patience, guidance, encouragement, and for always open you door for scientific questions. Your knowledge in genetics and molecular biology are impressive, and thank you for sharing it with me. I feel very lucky to be chosen as your student and it has been a real pleasure to work with and learn from you. I will always be grateful!

My co-supervisor **Professor Kerstin Brismar**, thank you for your continuous support and encouragement. Thank you for sharing your knowledge of clinical research. I really adore your energy, positivity and kindness. I wish to be like you when I get older.

Dr. Wan Nazaimoon, my Malaysian supervisor, for the opportunity and guidance from the beginning. Thank you for helping me with the samples collection back home in Malaysia when I was struggle to start my Ph.D program. Thanks for your passion, support and effort during all these years!

Professor Claes-Göran Östenson, my co-supervisor. Thank you for kind comments and discussion on my Ph.D. study.

Professor Halimah, as my mentor. Thank you for being my mentor.

Professor Peter Stenvinkel, for your idea and nice collaboration in the PTX3 study.

Sergiu-Bogdan Catarina, for your comments on my study during our group meetings.

My all co-authors in our studies, thank you for the nice collaboration. Particularly, **Juha Ojala** for the great assistance with the immunostaining, **Bing He** for introducing and teaching me hand by hand with the zebrafish experiments. **Anna, Monica, Ann-Christin Bragfors-Helin, Björn** for the nice environment in the Huddinge lab. It was my pleasure to work with all of you.

To all my friends in M1:03: **Agneta Hilding, Anneli Björklund, Carole Muller, Elizabeth Noren-Krog, Faradianna Lokman, Galyna Bryzgalova, Julien Pelletier, Kamal Yassin, Mohammed Seed Ahmed, Neil Portwood, Saad Alqahtani, Senthil Vasan, Silvia Zambrana, Tianwei Gu, Tina Wirström and Zuheng Ma**, thank you for the friendship and nice working environment, for the fika. Do contact me if you come to Malaysia.

To all the colleagues at L1: 01: **Christina Bark, Elisabete Forsberg, Ileana Botusan, Ishrath Ansurudeen, Ismael Valladolid, Jacob Grunler, Jing Wang, Marianna del sole, Michael Tekle, Noah Moruzzi, Teresa Daraio, Vivek Sunkari** and **Xiaowei Zheng** for the scientific discussion in group meeting every Monday and Friday, for great time we had together.

I would like to thank the administrative staff at MMK: **Katarina Breitholtz, Britt-Marie Witasp, Kerstin Florell** and **Ann-Brit Wikström** for kind and professional help, **Jan-Erik Karre** and **Thomas Westerberg** for professional IT support.

To all my research colleagues in Diabetes and Endocrine Unit, IMR, Malaysia, who were involved in the samples collection. Thank you for your support.

To all **Malaysian students** in Stockhom and Uppsala, all **Malaysian families** in Stockholm, Sweden. Thank you for the great time and memories. Thank you for organizing such memorable activities and making me feel like at home surrounding with wonderful people like all of you.

To my dear **families in Malaysia**, thank you for the unlimited support, encouragement and support during my study.

My beloved husband, **Kamarul**. Thank you Abang for your understanding and love throughout these years. My precious **Khayr Sufyan and Khayra Sumayyah**, may both of you grow up and be a good person. Always show respect to other people and believe in yourself. Ibu will always pray for your success.

8 REFERENCES

1. www.idf.org/signs-and-symptoms-diabetes

2. WHO. Definition, Diagnosis and Classification of Diabetes Mellitus and its Complications. Part 1: Diagnosis and Classification of Diabetes Mellitus. WHO/NCD/NCS/99.2. Geneva: World Health Organization, 1999.

3. International A, Committee E, Diabe- A, Federation ID, Committee IE, and International T. International Expert Committee report on the role of the A1C assay in the diagnosis of diabetes. Diabetes Care 2009; 32(7)1327–34.

4. Molitch ME, DeFronzo RA, Franzetal MJ. Nephropathy in diabetes. Diabetes Care 2004; (2): 27S79–S83.

5. I. Classification, Standards of medical care in diabetes--2014, Diabetes Care vol. 37 Suppl 1, no. October 2013, pp. S14–80, Jan. 2014.

6. Guariguata L, Whiting DR, Hambleton I, Beagley J, Linnenkamp U, Shaw JE. Global estimates of diabetes prevalence for 2013 and projections for 2035. Diabetes Res Clin Pract 2014; 103(2):137-49.

7. Letchuman GR, Wan Nazaimoon WM, Wan Mohamad WB, Chandran LR, Tee GH, Jamaiyah H, Isa MR, Zanariah H, Fatanah I, Ahmad Faudzi Y. Prevalence of diabetes in the Malaysian National Health Morbidity Survey III 2006. Med J Malaysia 2010; 65:180– 186.

8. Mohamud WN, Ismail AA, Sharifuddin A, Ismail IS, Musa KI, Kadir KA, Kamaruddin NA, Yaacob NA, Mustafa N, Ali O, Harnida S, Bebakar WM. Prevalence of metabolic syndrome and its risk factors in adult Malaysians: results of a nationwide survey. Diabetes Res Clin Pract 2011; 91(2):239-45.

9. Lim YN, Lim TO, Lee DG, Wong HS, Ong LM, Shaariah W, Rozina G, Morad Z. A report of the Malaysian dialysis registry of the National Renal Registry, Malaysia. Med J Malaysia 2008; 63 Suppl C:5-8.

10. Shaza AM, Rozina G, Izham MI, Azhar SS. Dialysis for end stage renal disease: a descriptive study in Penang Hospital. Med J Malaysia 2005; 60(3):320-7.

11. Hooi LS, Wong HS, Morad Z. Prevention of renal failure: the Malaysian experience. Kidney Int Suppl 2005; (94):S70-4.

12. USRDS. USRDS 2005 Annual Data Report: Atlas of end Stage renal Disease in the United States. Bethesd, MD: National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Disease. 2005.

13. Strimbu K, Tavel JA. What are biomarkers? Curr Opin HIV AIDS 2010; 5(6):463-6.

14. WHO International Programme on Chemical Safety Biomarkers in Risk Assessment: Validity and Validation. 2001.

15. https://www.kidney.org/

16. Taylor, E. Howard. Clinical Chemistry. New York: John Wiley and Sons 1989; 4:58–62.

17. Jha JC, Jandeleit-Dahm KA, Cooper ME. New insights into the use of biomarkers of diabetic nephropathy. Adv Chronic Kidney Dis 2014; 21(3):318-26.

18. Moresco RN, Sangoi MB, De Carvalho JA, Tatsch E, Bochi GV. Diabetic nephropathy: traditional to proteomic markers. Clin Chim Acta 2013; 421:17–30.

19. Starkey JM, Tilton RG. Proteomics and systems biology for understanding diabetic nephropathy. J Cardiovasc Transl Res 2012; 5 (4):479–490.

20. Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, Sherry S, Mullikin JC, Mortimore BJ, Willey DL et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 2001; 409(6822):928-933.

21. Kim JH, Shin HD, Park BL, Moon MK, Cho YM, Hwang YH, Oh KW, Kim SY, Lee HK, Ahn C, Park KS. SLC12A3 (solute carrier family 12 member [sodium/chloride] 3) polymorphisms are associated with end-stage renal disease in diabetic nephropathy. Diabetes 2006; 55(3):843-8.

22. Freedman BI, Langefeld CD, Murea M, Ma L, Otvos JD, Turner J,Antinozzi PA,Divers J, Hicks PJ, Bowden DW, Rocco MV, Parks JS. Apolipoprotein L1 nephropathy risk variants associate with HDL subfraction concentration in African Americans. Nephrol Dial Transplant 2011; 26(11):3805-10.

23. Almgren P, Lehtovirta M, Isomaa B, Sarelin L, Taskinen MR, Lyssenko V, Tuomi T, Groop L. Botnia Study Group Heritability and familiality of type 2 diabetes and related quantitative traits in the Botnia Study Diabetologia 2011; 54:2811–2819.

24. Cauchi S, et al. TCF7L2 is reproducibly associated with type 2 diabetes in various ethnic groups: a global meta-analysis. J. Mol. Med 2007; 85:777–782.

25. Sladek R, et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 2007; 445:881–885.

26. Saxena R, et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science 2007; 316:1331–1336.

27. Scott LJ, et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 2007; 316:1341–1345.

28. V. Steinthorsdottir, et al. A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. Nat. Genet 2007; 39:770–775.

29. Zhou Y, Park S.-Y, Su J, Bailey K, Ottosson-Laakso E, Shcherbina L, Oskolkov N, Zhang E, Thevenin T, Fadista J, et al. TCF7L2 is a master regulator of insulin production and processing. Hum. Mol. Genet 2014; 23:6419–6431.

30. Boj SF, van Es, J.H, Huch M, Li VSW, Jose´, A, Hatzis P, Mokry M, Haegebarth A, van den Born M, Chambon P, et al. Diabetes risk gene and Wnt effector Tcf7l2/TCF4 controls hepatic response to perinatal and adult metabolic demand. Cell 2012; 151:1595–1607.

31. Shao W, Wang D, Chiang YT, Ip W, Zhu L, Xu F, Columbus J, Belsham DD, Irwin DM, Zhang H, et al. The Wnt signaling pathway effector TCF7L2 controls gut and brain proglucagon gene expression and glucose homeostasis. Diabetes 2013; 62:789–800.

32. Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, Roix JJ, Kathiresan S, Hirschhorn JN, Daly MJ, et al. Diabetes Genetics Initiative of Broad Institute of Harvard and MIT, Lund University, and Novartis Institutes of Bio Medical Research. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science 2007; 316:1331–1336.

33. Majithia AR, Flannick J, Shahinian P, Guo M, Bray MA, Fontanillas P, GabrielSB; GoT2D Consortium; NHGRI JHS/FHS Allelic Spectrum Project; SIGMA T2D Consortium; T2D-GENES Consortium, Rosen ED, Altshuler D. Rare variants in PPARG with decreased activity in adipocyte differentiation are associated with increased risk of type 2 diabetes. Proc Natl Acad Sci U S A 2014; 111(36):13127-32.

34. Flannick J et al. Loss-of-function mutations in SLC30A8 protect against type 2 diabetes. Nat Genet 2014; 46(4):357-63.

35. Freedman BI, Bostrom M, Daeihagh P, Bowden DW. Genetic factors in diabetic nephropathy. Clin J Am Soc Nephrol 2007 Nov; 2(6):1306-16.

36. Thomas MC, Groop PH, Tryggvason K. Towards understanding the inherited susceptibility for nephropathy in diabetes. Curr Opin Nephrol Hypertens 2012; 21(2):195-202.

37. Bleyer AJ, Sedor JR, Freedman BI, O'Brien A, Russell GB, Graley J, Schelling JR. Risk factors for development and progression of diabetic kidney disease and treatment patterns among diabetic siblings of patients with diabetic kidney disease. Am J Kidney Dis 2008; 51(1):29-37.

38. Meagher RB, Müssar KJ. The influence of DNA sequence on epigenome induced pathologies. Epigenetics Chromatin 2012; 5:11.

39. Fraga MF, Ballestar E, M.F. Paz, S. Ropero, F. Setien, M.L. Ballestar, et al. Epigenetic differences arise during the lifetime of monozygotic twins. Proc Natl Acad Sci USA 2005; 102:10604–10609.

40. Drong AW, Lindgren CM, McCarthy MI. The genetic and epigenetic basis of type 2 diabetes and obesity. Clinical Pharmacology and Therapeutics 2012; 92: 707–715.

41. Karaguzel G, Ozer S, Akcurin S, Turkkahraman D, Bircan I. Type 1 diabetes-related epidemiological, clinical and laboratory findings. An evaluation with special regard to autoimmunity in children. Saudi Med J 2007; 28: 584–589.

42. Fetita LS, Sobngwi E, Serradas P, Calvo F, Gautier JF. Consequences of fetal exposure to maternal diabetes in offspring. J Clin Endocrinol Metab 2006; 91:3718–3724.

43. Jin B, Robertson KD. DNA methyltransferases, DNA damage repair, and cancer. Adv Exp Med Biol 2013; 754:3–29.

44. Ooi SKT, Bestor TH. The colorful history of active DNA demethylation. Cell 2008; 133:1145–1148.

45. Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet 2008; 9(6):465-76.

46. Jones PA, Baylin SB. The epigenomics of cancer. Cell 2007; 128(4):683-92.

47. Dayeh TA, Olsson AH, Volkov P, Almgren P, Rönn T, Ling C. Identification of CpG-SNPs associated with type 2 diabetes and differential DNA methylation inhuman pancreatic islets. Diabetologia 2013; 56(5):1036-46.

48. Volkmar M, Dedeurwaerder S, Cunha DA, Ndlovu MN, Defrance M, Deplus R, Calonne E, Volkmar U, Igoillo-Esteve M, Naamane N, Del Guerra S, Masini M, Bugliani M Marchetti P, Cnop M, Eizirik DL, Fuks F. DNA methylation profiling identifies epigenetic dysregulation in pancreatic islets from type 2 diabetic patients. EMBO J 2012; 31(6):1405-26.

49. Gu T, Gu HF, Hilding A, Sjöholm LK, Ostenson CG, Ekström TJ, Brismar K. Increased DNA methylation levels of the insulin-like growth factor binding protein 1 gene are associated with type 2 diabetes in Swedish men. Clin Epigenetics 2013; 5(1):21.

50. Smyth LJ, McKay GJ, Maxwell AP, McKnight AJ. DNA hypermethylation and DNA hypomethylation is present at different loci in chronic kidney disease. Epigenetics 2014; 9:366–376.

51. Bell, CG. et al. Genome-wide DNA methylation analysis for diabetic nephropathy in type 1 diabetes mellitus. BMC Med. Genomics 2010; 3:33.

52. Stenvinkel P, Karimi M, Johansson S et al. Impact of inflammation on epigenetic DNA methylation - a novel risk factor for cardiovascular disease? J Intern Med 2007; 261:488–499.

53. Dabelea D. The predisposition to obesity and diabetes in offspring of diabetic mothers. Diabetes Care 2007; 30(2):S169-74.

54. Dabelea D. et al. Diabetes in Navajo youth: prevalence, incidence, and clinical characteristics: the SEARCH for Diabetes in Youth Study. Diabetes Care 2009; 32(2):S141–7.

55. Stefan N. et al. Exaggerated insulin secretion in Pima Indians and African Americans but higher insulin resistance in Pima Indians compared with African Americans and Caucasians. Diabet Med 2004; 21:1090–5.

56. Indulekha K, Anjana RM, Surendar J, Mohan V. Association of visceral and subcutaneous fat with glucose intolerance, insulin resistance, adipocytokines and inflammatory markers in Asian Indians (CURES-113). Clin Biochem 2011; 44:281–7.

57. Goran MI, Bergman RN, Cruz ML, Watanabe R. Insulin resistance and associated compensatory responses in African American and Hispanic children. Diabetes Care 2002; 25:2184–90.

58. Nazare J-A, Smith JD, Borel A-L, et al. Ethnic influences on the relations between abdominal subcutaneous and visceral adiposity, liver fat, and cardio metabolic risk profile: the International Study of Prediction of Intra-Abdominal Adiposity and its Relationship with Cardio metabolic Risk/Intra-Abdominal Adiposity. Am J Clin Nutr 2012; 96:714-26.

59. Deurenberg P, Deurenberg-Yap M, Foo LF, Schmidt G, Wang J. Differences inbody composition between Singapore Chinese, Beijing Chinese and Dutch children. Eur J Clin Nutr 2003; 57:405–9.

60. Chung S, Song MY, Shin HD, et al. Korean and Caucasian overweight premenopausal women have different relationship of body mass index to percent body fat with age. J Appl Physiol 2005; 99:103–7.

61. Gurrici S, Hartriyanti Y, Hautvast JG, Deurenberg P. Differences in the relationship between body fat and body mass index between two different Indonesian ethnic groups: the effect of body build. Eur J Clin Nutr 1999; 53:468–72.

62. Federman DD. The biology of human sex differences. N Engl J Med 2006; 354(14):1507–1514.

63. Seidell JC, Bjorntorp P, Sjostrom L, Kvist H, Sannerstedt R. Visceral fat accumulation in men is positively associated with insulin, glucose, and C-peptide levels, but negatively with testosterone levels. Metabolism 1990; 39(9):897–901.

64. Tibblin G, Adlerberth A, Lindstedt G, Bjorntorp P. The pituitary-gonadal axis and health in elderly men: a study of men born in 1913. Diabetes 1996; 45(11):1605–1609.

65. Polderman KH, Gooren LJ, Asscheman H, Bakker A, Heine RJ. Induction of insulin resistance by androgens and estrogens. J Clin Endocrinol Metab 1994; 79(1):265–271.

66. Ding EL, Song Y, Malik VS, Liu S. Sex differences of endogenous sex hormones and risk of type 2 diabetes: a systematic review and meta-analysis. JAMA 2006; 295(11):1288–1299.

67. Rincon J, Holmang A, Wahlstrom EO, et al. Mechanisms behind insulin resistance in rat skeletal muscle after oophorectomy and additional testosterone treatment. Diabetes 1996; 45(5):615–621.

68. Pettersson-Fernholm K, Frojdo S, Fagerudd J, Tomas MC, Forsblom C, Wessman M, Groop PH, for the FinnDiane study group. The AT2 gene may have a gender-specific effect on kidney function and pulse pressure in type 1 diabetic patients. Kid Int 2006; 69:1880-4.

69. Gu HF, Alvarsson A, Efendic S, Brismar K. SOX2 Has Gender-Specific Genetic Effects on Diabetic Nephropathy in Samples From Patients With Type 1 Diabetes Mellitus in the GoKinD Study. Gender Medicine 2009; 6:555-564.

70. Retnakaran R, Cull CA, Thorne KI, Adler AI, Holman RR; UKPDS Study Group. Risk factors for renal dysfunction in type 2 diabetes: U.K. Prospective Diabetes Study 74. Diabetes 2006; 55(6):1832-9.

71. Nakhjavani M, Morteza A, Jenab Y, et al. Gender difference in albuminuria and ischemic heart disease in type 2 diabetes. Clin Med Res 2012; 10(2):51–56.

72. http://www.statistics.gov.my

73. Hummel KP, Dickie MM, Coleman DL. Diabetes, a new mutation in the mouse. Science 1966; 153(3740):1127-8.

74. Chen H, Charlat O, Tartaglia LA, Woolf EA, Weng X, Ellis SJ, Lakey ND, Culpepper J, Moore KJ, Breitbart RE, Duyk GM, Tepper RI, Morgenstern JP. Evidence that the diabetes gene encodes the leptin receptor: Identification of a mutation in the leptin receptor gene in db/db mice. Cell 1996; 84:491-495.

75. Lee GH, Proenca R, Montez JM, Carroll KM, Darvishzadah JG, Lee GI, and Freidman JM. Abnormal splicing of the leptin receptor in diabetic mice. Nature 1996; 379:632-635.

76. Sharma K, McCue P, Dunn SR. Diabetic kidney disease in the db/db mouse. Am J Physiol Renal Physiol 2003; 284:F1138-1144.

77. Lieschke GJ, Currie PD. Animal models of human disease: zebrafish swim into view. Nat Rev Genet 2007; 8(5):353-67.

78. Colella S, Shen L, Baggerly KA, Issa JP, Krahe R. Sensitive and quantitative universal pyrosequencing methylation analysis of CpG sites. Biotechniques 2003; 35(1):146–50.

79. Tost J, Dunker J, Gut IG. Analysis and quantification of multiple methylation variable positions in CpG islands by pyrosequencing. Biotechniques 2003; 35(1):152–6.

80. Wallace AC, Borkakoti N, Thornton JM. TESS: a geometric hashing algorithm for deriving 3D coordinate templates for searching structural databases. Application to enzyme active sites. Protein Sci 1997; 6:2308–2323.

81. Chen CC, Hwang JK, Yang JM. (PS) 2: protein structure prediction server. Nucleic Acids Res 2006; 34:W152–W157.

82. Maeda S. Genome-wide search for susceptibility gene to diabetic nephropathy by gene-based SNP. Diabetes Res Clin Pract 2004; 66(Suppl 1):S45–S47.

83. Tanaka N, Babazono T, Saito S, Sekine A, Tsunoda T, Haneda M, Tanaka Y, Fujioka T, Kaku K, Kawamori R, Kikkawa R, Iwamoto Y, Nakamura Y, Maeda S. Association of solute carrier family 12 (sodium/chloride) member 3 with diabetic nephropathy, identified by genome-wide analyses of single nucleotide polymorphisms. Diabetes 2003; 52(11):2848-53.

84. Ng DP, Nurbaya S, Choo S, Koh D, Chia KS, Krolewski AS. Genetic variation at the SLC12A3 locus is unlikely to explain risk for advanced diabetic nephropathy in Caucasians with type 2 diabetes. Nephrol Dial Transplant 2008; 23(7):2260-4.

85. Fredriksson S, Gullberg M, Jarvius J, Olsson C, Pietras K, Gústafsdóttir SM, Ostman A, Landegren U. Protein detection using proximity-dependent DNA ligation assays. Nat. Biotechnol 2002; 20:473– 477.

86. Gustafsdottir SM, Nordengrahn A, Fredriksson S, Wallgren P, Rivera E, Schallmeiner E, Merza M, Landegren U. Detection of individual microbial pathogens by proximity ligation. Clin Chem 2006; 52(6):1152-60.

87. Söderberg O, Gullberg M, Jarvius M, Ridderstråle K, Leuchowius KJ, Jarvius J, Wester K, Hydbring P, Bahram F, Larsson LG, Landegren U. Direct observation of individual endogenous protein complexes in situ by proximity ligation. Nat. Methods 2006; 3:995–1000.

88. Gustafsdottir SM, Schlingemann J, Rada-Iglesias A, Schallmeiner E, Kamali-Moghaddam M, Wadelius C, Landegren U. In vitro analysis of DNA-protein interactions by proximity ligation. Proc. Natl. Acad. Sci. U.S.A 2007; 104:3067– 3072.

89. Zhu L, Koistinen H, Landegren U, Stenman UH. Proximity ligation measurement of the complex between prostate specific antigen and alpha1-protease inhibitor. Clin. Chem 2009; 55:1665– 1671.

90. Gullberg M, Gústafsdóttir SM, Schallmeiner E, Jarvius J, Bjarnegård M, Betsholtz C, Landegren U, Fredriksson S. Cytokine detection by antibody-based proximity ligation. Proc. Natl. Acad. Sci. U.S.A 2004; 101:8420– 8424.

91. Fredriksson S, Dixon W, Ji H., Koong AC, Mindrinos M, Davis RW. Multiplexed protein detection by proximity ligation for cancer biomarker validation. Nat. Methods 2007; 4:327– 329.

92. Fredriksson S, Horecka J, Brustugun OT, Schlingemann J, Koong AC, Tibshirani R, Davis RW. Multiplexed proximity ligation assays to profile putative plasma biomarkers relevant to pancreatic and ovarian cancer. Clin. Chem 2008; 54:582– 589.

93. Blokzijl A, Nong R, Darmanis S, Hertz E, Landegren U, Kamali-Moghaddam M. Protein biomarker validation via proximity ligation assays. Biochim Biophys Acta 2014; 1844(5):933-9.

94. Kamali-Moghaddam M, Pettersson FE, Wu D, Englund H, Darmanis S, Lord A, Tavoosidana G, Sehlin D, Gustafsdottir S, Nilsson LN, Lannfelt L, Landegren U. Sensitive detection of Aβ protofibrils by proximity ligation--relevance for Alzheimer's disease. BMC Neurosci 2010; 11:124.

95. Wallin U, Glimelius B, Jirström K, Darmanis S, Nong RY, Pontén F, Johansson C, Påhlman L, Birgisson H. Growth differentiation factor 15: a prognostic marker for recurrence in colorectal cancer. Br J Cancer 2011; 104(10):1619-27.

96. Billings LK, Florez JC. The genetics of type 2 diabetes: what have we learned from GWAS? Ann N Y Acad Sci 2010; 1212:59-77.

97. Ahlqvist E, Ahluwalia TS, Groop L. Genetics of type 2 diabetes. Clin Chem 2011; 57(2):241-54.

98. Imamura M, Maeda S. Genetics of type 2 diabetes: the GWAS era and future perspectives. Endocr J 2011; 58(9):723-39.

99. Ling C, Groop L. Epigenetics: a molecular link between environmental factors and type 2 diabetes. Diabetes 2009; 58(12):2718-25.

100. Drong AW, Lindgren CM, McCarthy MI. The genetic and epigenetic basis of type 2 diabetes and obesity. Clin Pharmacol Ther 2012; 92(6):707-15.

101. Kirchner H, Osler ME, Krook A, Zierath JR. Epigenetic flexibility in metabolic regulation: disease cause and prevention? Trends Cell Biol 2013; 23(5):203-9.

102. Hall E, Volkov P, Dayeh T, Bacos K, Rönn T, Nitert MD, Ling C. Effects of palmitate on genome-wide mRNA expression and DNA methylation patterns in human pancreatic islets. BMC Med 2014; 12:103.

103. Dayeh TA, Olsson AH, Volkov P, Almgren P, Rönn T, Ling C. Identification of CpG-SNPs associated with type 2 diabetes and differential DNA methylation in human pancreatic islets. Diabetologia 2013; 56(5):1036-46.

104. Dick KJ, Nelson CP, Tsaprouni L, Sandling JK, Aïssi D, Wahl S, Meduri E,Morange PE, Gagnon F, Grallert H, Waldenberger M, Peters A, Erdmann J, Hengstenberg C, Cambien F, Goodall AH, Ouwehand WH, Schunkert H, Thompson JR, Spector TD, Gieger C, Trégouët DA, Deloukas P, Samani NJ. DNA methylation and body-mass index: a genome-wide analysis. Lancet 2014; S0140-6736(13):62674-4.

105. Su X, Chen X, Liu L, Chang X, Yu X, Sun K. Intracellular adhesion molecule-1 K469E gene polymorphism and risk of diabetic microvascular complications: a meta-analysis. PLoS One 2013; 8(7):e69940.

106. Ma J, Zhang D, Brismar K, Efendic S, Gu HF. Evaluation of the association between the common E469K polymorphism in the ICAM-1 gene and diabetic nephropathy among type 1 diabetic patients in GoKinD population. BMC Med Genet 2008; 9:47.

107. He Q, Lin X, Wang F, Xu J, Ren Z, Chen W, Xing X. Associations of a polymorphism in the intercellular adhesion molecule-1 (ICAM1) gene and ICAM1 serum levels with migraine in a Chinese Han population. J Neurol Sci 2014; 345(1-2):148-53.

108. Miller J, Knorr R, Ferrone M, Houdei R, Carron CP, Dustin ML. Intercellular adhesion molecule-1 dimerization and its consequences for adhesion mediated by lymphocyte function associated-1. J Exp Med 1995; 182(5):1231-41.

109. Brown A, Turner L, Christoffersen S, Andrews KA, Szestak T, Zhao Y, Larsen S, Craig AG, Higgins MK. Molecular architecture of a complex between an adhesion protein from the malaria parasite and intracellular adhesion molecule 1. J Biol Chem 2013; 288(8):5992-6003.

110. Yamasaki K, Kurimura M, Kasai T, Sagara M, Kodama T, Inoue K. Determination of physiological plasma pentraxin 3 (PTX3) levels in healthy populations. Clin Chem Lab Med 2009; 47(4):471-7.

111. Witasp A, Rydén M, Carrero JJ, Qureshi AR, Nordfors L, Näslund E, Hammarqvist F, Arefin S, Kublickiene K, Stenvinkel P. Elevated circulating levels and tissue expression of pentraxin 3 in uremia: a reflection of endothelial dysfunction. PLoS One 2013; 8(5).

112. Tong M, Carrero JJ, Qureshi AR, Anderstam B, Heimbürger O, Bárány P, Axelsson J, Alvestrand A, Stenvinkel P, Lindholm B, Suliman ME. Plasma pentraxin 3 in patients with chronic kidney disease: associations with renal function, protein-energy wasting, cardiovascular disease, and mortality. Clin J Am Soc Nephrol 2007; 2(5):889-97

113. Mikeska T, Craig JM. DNA methylation biomarkers: cancer and beyond. Genes (Basel) 2014; 5(3):821-64.

114. Adalsteinsson BT, Gudnason H, Aspelund T, Harris TB, Launer LJ, Eiriksdottir G, Smith AV, Gudnason V. Heterogeneity in white blood cells has potential to confound DNA methylation measurements. PLoS One 2012; 7:e46705.

115. Loh M, Liem N, Lim PL, Vaithilingam A, Cheng CL, Salto-Tellez M, Yong WP, Soong R. Impact of sample heterogeneity on methylation analysis. Diagn. Mol. Pathol 2010; 19:243–247.

116. Montano CM, Irizarry RA, Kaufmann WE, Talbot K, Gur RE, Feinberg AP, Taub MA. Measuring cell-type specific differential methylation in human brain tissue. Genome Biol 2013; 14:R94.

117. Moverare-Skrtic S, Mellstrom D, Vandenput L, Ehrich M, Ohlsson C. Peripheral blood leukocyte distribution and body mass index are associated with the methylation pattern of the androgen receptor promoter. Endocrine 2009; 35:204–210.

118. Zou J, Lippert C, Heckerman D, Aryee M, Listgarten J. Epigenome-wide association studies without the need for cell-type composition. Nat. Methods 2014; 11:309–311.

119. Accomando WP, Wiencke JK, Houseman EA, Nelson HH, Kelsey KT. Quantitative reconstruction of leukocyte subsets using DNA methylation. Genome Biol 2014; 15:R50.

120.Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, Wiencke JK, Kelsey KT. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinform 2012; 13:86.

121.deVos T. et al. Circulating methylated SEPT9 DNA in plasma is a biomarker for colorectal cancer. Clin Chem 2009; 55(7):1337–46.

122. Schultz NS, Broholm C, Gillberg L, Mortensen B, Jørgensen SW, Schultz HS, Scheele C, Wojtaszewski JF, Pedersen BK, Vaag A. Impaired leptin gene expression and release in cultured pre adipocytes isolated from individuals born with low birth weight. Diabetes 2014; 63(1):111-21.