LUND UNIVERSITY

DNA methylation patterns in human adipose tissue in relation to diet and type 2 diabetes

Perfilyev, Alexander

2018

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Perfilyev, A. (2018). *DNA methylation patterns in human adipose tissue in relation to diet and type 2 diabetes*. [Doctoral Thesis (compilation), Department of Clinical Sciences, Malmö]. Lund University, Faculty of Medicine.

Total number of authors:
1

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Download date: 07. Oct. 2026

DNA methylation patterns in human adipose tissue in relation to diet and type 2 diabetes

ALEXANDER PERFILYEV
DEPARTMENT OF CLINICAL SCIENCES | LUND UNIVERSITY

DNA methylation patterns in human adipose tissue in relation to diet and type 2 diabetes

Alexander Perfilyev

LUND
UNIVERSITY

DOCTORAL DISSERTATION
by due permission of the Faculty of Medicine, Lund University, Sweden.
To be defended at Aula Auditorium, CRC Malmö,
on November 20th, 2018 at 13:00.

Faculty opponent
Professor Andrew P. Feinberg,
Baltimore, Maryland, U.S.A

Organization LUND UNIVERSITY	Document name Doctoral Dissertation
	Date of issue November 20th, 2018
Author(s) Alexander Perfilyev	Sponsoring organization

Title and subtitle DNA methylation patterns in human adipose tissue in relation to diet and type 2 diabetes

Abstract

Type 2 diabetes (T2D) is a common metabolic disease and its prevalence is increasing worldwide. Adipose tissue plays an important role in metabolic processes. Environmental factors may affect metabolic phenotypes and epigenetics may mediate this influence. We used Illumina 450k microarrays to study correlations between epigenetic patterns in human adipose tissue and environmental factors such as high-fat diets, impaired intrauterine environment and low birth weight as well as with type 2 diabetes.

In study 1 we investigated whether the DNA methylation pattern in human adipose tissue is affected differently by a 7-week exposure to a diet high in either polyunsaturated fatty acids or saturated fatty acids. We found differences in methylation related to general fat overfeeding as well as between the two dietary groups.

In study 2 we studied the effect of 5 days of overfeeding with a high fat diet on DNA methylation in adipose tissue compared to a control diet. The participants were either born with a normal or low birth weight and it allowed also to compare the effects of the diets on these two groups separately. There were differences in methylation between subjects born with a low compared with normal birth weight.

In study 3 we used adipose tissue of a cohort of 14 monozygotic twin pairs discordant for type 2 diabetes in order to eliminate most of known confounding factors in comparing DNA methylation patterns as well as a case-control cohort for T2D. We found numerous differences in methylation in adipose tissue from subjects with T2D compared with controls, while less differences were found in the discordant twins.

In study 4 we compared DNA methylation in adipose-derived stem and differentiated cells in subjects born with a low or normal birth weight. We found differences in gene expression in adipose-derived stem cells between low birth weight and normal birth weight groups but did not identify differences in methylation patterns possibly due to lack of statistical power.

Overall, these studies contribute to a better understanding of the influence of environmental factors and type 2 diabetes on DNA methylation in human adipose tissue.

Key words DNA methylation, type 2 diabetes, Illumina 450k, adipose tissue		
Classification system and/or index terms (if any)		
Supplementary bibliographical information		Language English
ISSN and key title 1652-8220		ISBN 978-91-7619-704-2
Recipient's notes	Number of pages	Price
	Security classification	

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature Date 2018-10-15

DNA methylation patterns in human adipose tissue in relation to diet and type 2 diabetes

Alexander Perfilyev

LUND
UNIVERSITY

Cover picture by Vesta Shevchenko. Edited by Liza Shkirando.

Copyright (p 1-47) Alexander Perfilyev

Paper 1 © 2017 American Society of Nutrition

Paper 2 © Springer-Verlag Berlin Heidelberg 2016

Paper 3 © 2014 by the American Diabetes Association

Paper 4 © Springer-Verlag Berlin Heidelberg 2016

Faculty of Medicine
Department of Clinical Sciences

ISBN 978-91-7619-704-2
ISSN 1652-8220

Printed in Sweden by Media-Tryck, Lund University
Lund 2018

MADE IN SWEDEN

Media-Tryck is an environmentally certified and ISO 14001 certified provider of printed material. Read more about our environmental work at www.mediatryck.lu.se

To Vesta

Table of contents

List of publications 8
Publications included in the thesis 8
Publications not included in the thesis 9
Abbreviations 11

Introduction 13
Type 2 diabetes 13
Risk factors for type 2 diabetes 13
Obesity and adipose tissue 14
Food fat composition 15
Epigenetics and DNA methylation 15
Bioinformatics of DNA methylation 16

Aims 21
Study 1 21
Study 2 21
Study 3 21
Study 4 21

Materials and methods 23
Ethics statement 23
Study participants 23
Study 1 23
Study 2 24
Study 3 25
Study 4 27
DNA methylation 27
DNA extraction and bisulfite conversion 27
Illumina 450k BeadChip array 28
Gene expression 28
Affymetrix GeneChip Human Gene ST Array chip 28
Illumina HumanHT-12 v4 Expression BeadChip 28
Statistical and bioinformatic analysis 29
Statistical and analysis tools 29

Results **31**
Study 1 31
Study 2 32
Study 3 33
Study 4 34
Discussion **35**
Summary and general conclusion **39**
Acknowledgements **41**
References **43**

List of publications

Publications included in the thesis

1. **Perfilyev A**, Dahlman I, Gillberg L, Rosqvist F, Iggman D, Volkov P, Nilsson E, Risérus U, Ling C. *Impact of polyunsaturated and saturated fat overfeeding on the DNA-methylation pattern in human adipose tissue: a randomized controlled trial.* Am J Clin Nutr. 2017 Apr;105(4):991-1000.
2. Gillberg L, **Perfilyev A**, Brøns C, Thomasen M, Grunnet LG, Volkov P, Rosqvist F, Iggman D, Dahlman I, Risérus U, Rönn T, Nilsson E, Vaag A, Ling C. *Adipose tissue transcriptomics and epigenomics in low birthweight men and controls: role of high-fat overfeeding.* Diabetologia. 2016 Apr;59(4):799-812.
3. Nilsson E, Jansson PA, **Perfilyev A**, Volkov P, Pedersen M, Svensson MK, Poulsen P, Ribel-Madsen R, Pedersen NL, Almgren P, Fadista J, Rönn T, Klarlund Pedersen B, Scheele C, Vaag A, Ling C. *Altered DNA methylation and differential expression of genes influencing metabolism and inflammation in adipose tissue from subjects with type 2 diabetes.* Diabetes. 2014 Sep;63(9):2962-76
4. Broholm C, Olsson AH, **Perfilyev A**, Hansen NS, Schrölkamp M, Strasko KS, Scheele C, Ribel-Madsen R, Mortensen B, Jørgensen SW, Ling C, Vaag A. *Epigenetic programming of adipose-derived stem cells in low birthweight individuals.* Diabetologia. 2016 Dec;59(12):2664-2673.

Publications not included in the thesis

1. Rönn T, Volkov P, Gillberg L, Kokosar M, **Perfilyev A**, Jacobsen AL, Jørgensen SW, Brøns C, Jansson PA, Eriksson KF, Pedersen O, Hansen T, Groop L, Stener-Victorin E, Vaag A, Nilsson E, Ling C. *Impact of age, BMI and HbA1c levels on the genome-wide DNA methylation and mRNA expression patterns in human adipose tissue and identification of epigenetic biomarkers in blood.* Hum Mol Genet. 2015 Jul 1;24(13):3792-813.

2. Agardh E, Lundstig A, **Perfilyev A**, Volkov P, Freiburghaus T, Lindholm E, Rönn T, Agardh CD, Ling C. *Genome-wide analysis of DNA methylation in subjects with type 1 diabetes identifies epigenetic modifications associated with proliferative diabetic retinopathy.* BMC Med. 2015 Aug 6;13:182.

3. Nilsson E, Matte A, **Perfilyev A**, de Mello VD, Käkelä P, Pihlajamäki J, Ling C. *Epigenetic Alterations in Human Liver From Subjects With Type 2 Diabetes in Parallel With Reduced Folate Levels.* J Clin Endocrinol Metab. 2015 Nov;100(11):E1491-501.

4. Kokosar M, Benrick A, **Perfilyev A**, Fornes R, Nilsson E, Maliqueo M, Behre CJ, Sazonova A, Ohlsson C, Ling C, Stener-Victorin E. *Epigenetic and Transcriptional Alterations in Human Adipose Tissue of Polycystic Ovary Syndrome.* Sci Rep. 2016 Mar 15;6:22883.

5. Dayeh T, Tuomi T, Almgren P, **Perfilyev A**, Jansson PA, de Mello VD, Pihlajamäki J, Vaag A, Groop L, Nilsson E, Ling C. *DNA methylation of loci within ABCG1 and PHOSPHO1 in blood DNA is associated with future type 2 diabetes risk.* Epigenetics. 2016 Jul 2;11(7):482-8.

6. Broholm C, Olsson AH, **Perfilyev A**, Hansen NS, Schrölkamp M, Strasko KS, Scheele C, Ribel-Madsen R, Mortensen B, Jørgensen SW, Ling C, Vaag A. *Epigenetic programming of adipose-derived stem cells in low birthweight individuals.* Diabetologia. 2016 Dec;59(12):2664-2673.

7. Broholm C, Olsson AH, **Perfilyev A**, Gillberg L, Hansen NS, Ali A, Mortensen B, Ling C, Vaag A. *Human adipogenesis is associated with genome-wide DNA methylation and gene-expression changes.* Epigenomics. 2016 Dec;8(12):1601-1617.

8. Bysani M, **Perfilyev A**, de Mello VD, Rönn T, Nilsson E, Pihlajamäki J, Ling C. *Epigenetic alterations in blood mirror age-associated DNA methylation and gene expression changes in human liver.* Epigenomics. 2017 Feb;9(2):105-122.

9. Davegårdh C, Broholm C, **Perfilyev A**, Henriksen T, García-Calzón S, Peijs L, Hansen NS, Volkov P, Kjøbsted R, Wojtaszewski JF, Pedersen M, Pedersen BK, Ballak DB, Dinarello CA, Heinhuis B, Joosten LA, Nilsson E, Vaag A, Scheele C, Ling C. *Abnormal epigenetic changes during differentiation of human skeletal muscle stem cells from obese subjects.* BMC Med. 2017 Feb 22;15(1):39.

10. de Mello VD, Matte A, **Perfilyev A**, Männistö V, Rönn T, Nilsson E, Käkelä P, Ling C, Pihlajamäki J. *Human liver epigenetic alterations in non-alcoholic steatohepatitis are related to insulin action.* Epigenetics. 2017 Apr 3;12(4):287-295.

11. García-Calzón S, **Perfilyev A**, Männistö V, de Mello VD, Nilsson E, Pihlajamäki J, Ling C. *Diabetes medication associates with DNA methylation of metformin transporter genes in the human liver.* Clin Epigenetics. 2017 Sep 21;9:102.

12. Kokosar M, Benrick A, **Perfilyev A**, Nilsson E, Källman T, Ohlsson C, Ling C, Stener-Victorin E. *A Single Bout of Electroacupuncture Remodels Epigenetic and Transcriptional Changes in Adipose Tissue in Polycystic Ovary Syndrome.* Sci Rep. 2018 Jan 30;8(1):1878.

13. García-Calzón S, **Perfilyev A**, Mello VD, Pihlajamäki J, Ling C. *Sex differences in the methylome and transcriptome of the human liver and circulating HDL-cholesterol levels*. J Clin Endocrinol Metab. 2018 May 28.

Abbreviations

ADSC	Adipose-derived stem cells
BMI	Body mass index
DMR	Differentially methylated region
DZ	Dizygotic
FDR	False discovery rate
FWER	Familywise error rate
HFO	High fat overfeeding
IDF	International Diabetes Federation
KEGG	Kyoto Encyclopedia of Genes and Genomes
LBW	Low birth weight
MZ	Monozygotic
NBW	Normal birth weight
NGT	Normal glucose tolerance
PCR	Polymerase chain reaction
PUFA	Polyunsaturated fatty acids
RefSeq	The Reference Sequence database
SFA	Saturated fatty acids
SVA	Surrogate variable analysis
T2D	Type 2 diabetes
UCSC	University of California, Santa Cruz
WGBS	Whole-genome bisulfite sequencing
WHO	World Health Organization
WHR	Waist to hip ratio

Introduction

Type 2 diabetes

Diabetes mellitus (DM) is a disease characterized by chronically elevated blood glucose levels. International Diabetes Federation (IDF) estimates that over 425 million people worldwide have diabetes (http://www.diabetesatlas.org/key-messages.html). According to IDF only half of these cases are diagnosed. Diabetes is associated with a number of complications such as cardiovascular disease, blindness, kidney failure, and lower limb amputation, some of which are associated with increased mortality (1).

According to the World Health Organization (WHO), diabetes may be diagnosed using one of the following measurements: by fasting plasma glucose levels ≥ 7 mmol/l (126 mg/dl), by a plasma glucose concentration ≥ 11.1 mmol/l (200 mg/dl) 2 h after 75g oral glucose load or hemoglobin A1c (HbA1c) of above 6.5% (48 mmol/mol). HbA1c is a glycosylated form of hemoglobin that represents the average blood glucose levels during the last eight to twelve weeks (2).

There are two major types of diabetes, Type 1 Diabetes (T1D) and Type 2 Diabetes (T2D). In T1D autoimmune mechanisms destroy pancreatic beta-cells leading to absolute insulin deficiency. It is often diagnosed in children and adults younger than 35 years (3). T2D is a multifactorial, polygenic disease characterized by a combination of insulin resistance and dysregulated insulin secretion. Insulin resistant people often firstly hypersecrete insulin in order to compensate for the increased insulin demand. As the disease progresses, the beta-cells fail to compensate and T2D develops. Other types of diabetes include the maturity onset diabetes of the young (MODY), latent autoimmune diabetes in adults (LADA), gestational diabetes and other subtypes (3).

Risk factors for type 2 diabetes

T2D is often linked to higher age and increased body mass index (BMI) (3). As many as 80% of people with T2D are overweight or obese (4). Other risk factors include low physical activity, energy rich diets, smoking, ethnicity, family history

of T2D, history of gestational diabetes mellitus, impaired intrauterine environment and different drug treatments (5–10).

An impaired intrauterine environment can be reflected by a low birth weight (LBW), which has been linked to certain phenotypic differences in adulthood compared to individuals born with a normal birth weight (NBW). These differences include reduced insulin secretion (11), higher abdominal fat mass (12) and reduced lean body mass (13). Together these differences may lead to an increased risk of developing T2D (14).

The strong genetic contribution to T2D was established through twin and family-based studies (15,16). Twin studies indicate higher concordance of T2D in monozygotic twins (~60-70%) compared to dizygotic twins (~20-40%) (9,17). Over 120 loci associated with susceptibility to T2D were identified by genome wide association studies (GWAS) in the past decade (18). However only a small proportion of heritability of T2D is explained by the identified single nucleotide polymorphisms (SNPs) (3). Epigenetics is one of the possible explanations for the missing heritability.

Obesity and adipose tissue

The prevalence of obesity is increasing worldwide and obesity is now considered to be global epidemic. According to WHO, a person with BMI in the range between 25 and 30 is considered overweight. A person with BMI over 30 is considered obese. Overweight and obesity are some of the strongest risk factors for T2D development.

The adipose tissue is one of the main organs regulating the whole-body energy homeostasis by controlling lipid storage and release. It consists of adipocytes that store energy in the form of triacylglycerol (TAG), endothelial cells, fibroblasts and macrophages. During large weight gain, adipose tissue expansion may result in dysfunctional adipocytes with impaired lipid metabolism, lipid uptake into other tissues, immune cell infiltration and subsequent low-grade inflammation. As a consequence, local and peripheral insulin resistance, hyperglycemia and T2D may develop (19).

Several studies indicate that insulin resistance may originate in the adipose tissue and this may later induce insulin resistance in skeletal muscle and liver before the development of obesity and T2D (20,21). Excessive energy intake has negative effects on the whole-body insulin sensitivity independently from body weight changes. It was shown in both healthy lean (22–24) and obese diabetic individuals (25). Additionally, short- and long-term overfeeding has been shown to have effects on gene expression in adipose tissue (26–29).

Food fat composition

Dietary fatty acids can be divided into 4 categories: saturated (SFA), monounsaturated, polyunsaturated (PUFA) and trans fats. The physiological effects of each fatty acid type have been subject of numerous published studies. For instance, a diet high in vegetable n-6 PUFAs decreased abdominal fat content and peripheral insulin resistance compared with a diet high in SFAs (30). SFAs have been shown to correlate positively (31,32) and PUFAs have been shown to correlate negatively (32) with the liver fat accumulation. At the same time excessive liver fat has been strongly associated with metabolic disorders and T2D (33,34). Some studies indicate that replacing SFAs with PUFAs may reduce risk of a coronary artery disease (CAD) events (35,36) as well as risk of developing diabetes (37).

In two of the studies presented in this thesis we investigated the effects of general high-fat overfeeding and overfeeding with a specific fatty acid type on adipose tissue transcriptomics and epigenomics.

In study 2, a cohort of young men was exposed to both a control and a high-fat overfeeding (HFO) diet separated by a 6-8-week wash-out period. The control diet was calculated based on the individual energy requirement with 35% of the energy coming from fat, 50% from carbohydrates, and 15% from protein. The HFO diet, on the other hand, was calculated with 50% total energy more than the control diet and 60% of the energy came from fat, 32.5% from carbohydrates, and 7.5% from protein. The fat proportion of the diets was 1/3 monounsaturated, 1/3 polyunsaturated and 1/3 saturated fatty acids. The control diet was 3 days long and had average calorie intake of 2,818 ± 239 Cal. The HFO diet was 5 days long and had average calorie intake of 4,228 ± 334 Cal (24).

In study 1, a cohort of healthy young individuals was exposed to a 7-week randomized dietary intervention. The participants maintained their habitual diet and were given muffins containing either PUFA or SFA in a randomized fashion thus making the diet high in fat and calories. The amount was individually adjusted to achieve a 3% weight gain during this period. The muffins provided 51% of energy from fat, 5% from protein and the remaining 44% from carbohydrates. PUFAs came from sunflower oil and SFAs came from refined palm oil (38).

Epigenetics and DNA methylation

An epigenetic trait is a stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence (39). Epigenetic changes play an important role in a number of biological processes such as cell differentiation, development, regulation of cell specific gene expression, imprinting, inactivation of the X-chromosomes in females and genomic stability (40).

DNA methylation is an epigenetic modification that occurs when a methyl group is added to a cytosine base of a DNA molecule (41). In differentiated mammalian cells DNA methylation mainly occurs in so-called CpG sites, sequence of a cytosine followed by a guanine. DNA methyl transferases (DNMTs) regulate DNA methylation. DNMT1 is mainly responsible for DNA methylation maintenance during replication and *de novo* methylation is regulated by DNMT3A and 3B (42,43). Demethylation can be either passive, when DNA methylation marks are not reproduced during replication, or active when the methyl group is removed by the enzymes involved in DNA repair. Alternatively, active demethylation can occur with TET (ten-eleven translocation) enzymes converting 5-methylcytosine to 5-hydroxymethylcytosine and subsequently to formylcytosine and carboxylcytosine (44,45).

The human genome contains approximately 29 million CpG sites (46) and 70% to 80% of those are methylated (47). Their distribution across the genome is not even and there are regions where CpG sites occur at a higher frequency. These regions are called CpG islands and usually they tend to have lower methylation levels and often are found in promoter regions. One of the definitions of a CpG island states that it is a genomic region longer than 500bp with C and G content not less than 55% and the observed/expected ratio of CpG sites of 0.65 or more (48). About 70% of human genes have CpG islands in their promoter regions and higher DNA methylation of these CpG islands has been correlated with lower gene transcription (40,47).

Bioinformatics of DNA methylation

Bisulfite treatment of DNA is a common technique used to determine its pattern of methylation. The treatment converts cytosine to uracil by deamination while leaving 5-methylcytosine intact. After polymerase chain reaction (PCR), uracil is converted to thymine (49) (**Figure 1**). This process allows to discriminate between methylated and unmethylated cytosines using microarray or sequencing technologies.

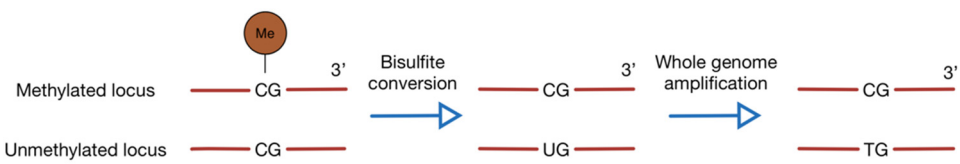

Figure 1.
Bisulfite treatment of DNA.

In the studies included in this thesis we used a microarray solution from Illumina (San Diego, CA, USA), Illumina Infinium HumanMethylation450 BeadChip (Illumina 450k array) (50) that was the most comprehensive method available at the time. Illumina 450k array is a 12-sample microarray chip that includes 482,421 oligonucleotide probes designed to measure DNA methylation of CpG sites across the genome, 3,091 probes measure methylation in non-CpG locations and 65 control SNPs. The array covers 21,231 out of 21,474 UCSC RefSeq genes and 96% of UCSC CpG islands (50). DNA methylation is traditionally measured as a proportion between the number of cells methylated at a certain genomic position and the total number of analyzed cells. This proportion is referred to as beta-value and can also be represented as percentage. While very intuitive and useful this measure has certain issues. Du et al. (51) have shown that it exhibits heteroscedastic properties in hyper- and hypomethylated regions of the overall distribution (**Figure 2**). In other words, variance depends on mean value and, in this case, it was smaller when the mean value was closer to 0 or 1. This represents a problem since many statistical analyses such as linear models or ANOVA, assume homoscedasticity of the analyzed data (52). Alternatively, DNA methylation can be measured as M-value that is defined as the log2 ratio of the intensities of methylated probe versus unmethylated probe: $log_2\left(\frac{M}{U}\right)$. The relationship between beta-values and M-values is a logistic function: $M = log_2\left(\frac{beta}{1-beta}\right)$. According to Du et al. (51) this transformation effectively resolves the heteroscedasticity problem (**Figure 3**).

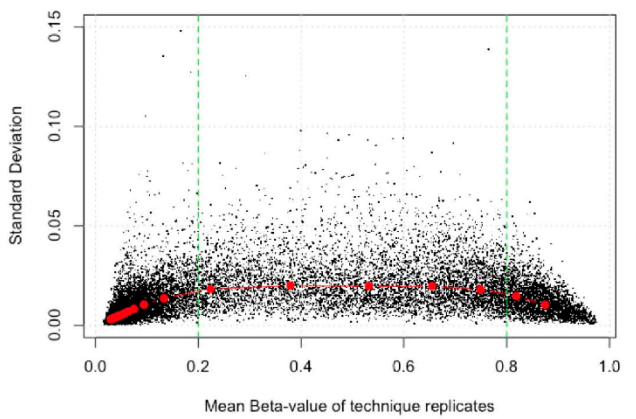

Figure 2.
The mean and standard deviation relations of technical replicates. Beta-values. As published in (51).

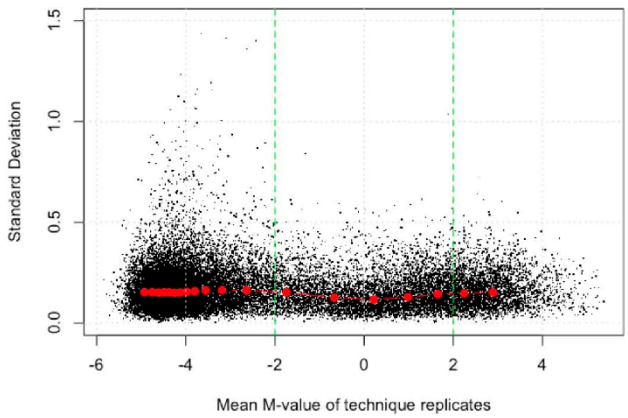

Figure 3.
The mean and standard deviation relations of technical replicates. M-values. As published in (51).

Another challenge associated with analyzing Illumina 450k array data is that the array utilizes two types of probes with two different assay chemistries (50). Infinium type I probes cover 135,501 sites and use two bead types for methylated and unmethylated states of a target site. Infinium type II probes cover 350,076 sites and have one bead type corresponding to both DNA methylation states based on single base extension (50). It was shown that the distribution of DNA methylation values differs between probe types and that type II probes typically have a much lower dynamic range compared to type I probes (53). It makes it impossible to compare type I and type II probes which leads to several problems. First, it introduces a type I enrichment bias meaning it is more likely to detect differentially methylated type I probes than type II probes due to the higher dynamic range of type I probes. Then it makes it impossible to use methods looking at differentially methylated regions. Beta-mixture quantile dilation (BMIQ) normalization method was proposed to correct for the probes bias (54). The method involves, as described in the paper, "application of a three-state beta-mixture model to assign probes to methylation states, subsequent transformation of probabilities into quantiles and finally a methylation-dependent dilation transformation to preserve the monotonicity and continuity of the data".

Batch effects can be defined as unwanted non-biological differences between groups of analyzed samples. The differences can be caused by many factors including samples being processed by different people at a different time, using a different batch of a reagent or samples being placed on different microarray chips. Batch effects make samples from different batches not directly comparable. In our studies we have been using ComBat procedure (55) as implemented in the R package sva (56). It uses parametric and non-parametric empirical Bayes

frameworks and has been shown to work well with small batch sizes (<25 samples) (55).

It has been shown that as many as ~8.5% of Infinium Type I probes and ~5.1% of Infinium Type II probes may potentially match multiple genomic locations (57). About 14% of the probes target a CpG site with a known SNP (58) meaning that differences in measured values may be due to the removal of a CpG site and not to actual changes in methylation.

Aims

Study 1

To determine whether overfeeding with a diet high in PUFA or SFA affected DNA methylation in adipose tissue differently in a randomized trial.

Study 2

To investigate the effects of high-fat overfeeding on mRNA expression and DNA methylation in adipose tissue and to compare these effects between individuals who were born with normal (NBW) or low birth weight (LBW).

Study 3

By using a cohort of monozygotic twins discordant for T2D we aimed at eliminating the genetic component in comparison of the genome-wide DNA methylation and expression patterns in adipose tissue between diabetic and non-diabetic individuals.

Study 4

Our aim was to investigate whether phenotypic differences between individuals born with LBW or NBW can be explained by different RNA expression and DNA methylation patterns in both adipose-derived stem cells and mature adipocytes.

Materials and methods

Ethics statement

All studies were approved by the local ethics committees and written informed consent was obtained from all participants.

Study participants

Study 1

31 healthy normal-weight people participated in this study (Lipogain). They were aged between 20 and 38 years, had BMI between 18 and 27 kg/m^2 and they exercised less than 3 hours/week. They had no history of diabetes or liver diseases. Participants were instructed to keep their usual diet and exercise levels during the study. Adipose tissue biopsies were extracted before and after the dietary intervention. Clinical characteristics of the participants of Lipogain cohort before and after overfeeding can be found in Table 1.

Table 1. Clinical characteristics of the Lipogain cohort (n=31) at baseline and after 7 weeks overfeeding of either SFA or PUFA

Data are presented as means ± SDs. *P<0.05 baseline compared to overfeeding using Wilcoxon's paired test.

	SFA group		PUFA group	
	Baseline	**Overfeeding**	**Baseline**	**Overfeeding**
n (male/female)	17 (13/4)		14 (9/5)	
Age (years)	26.94 ± 4.68		27 ± 4.01	
Weight (kg)	68.35 ± 7.33*	70 ± 7.26*	64.12 ± 6.43*	65.76 ± 7.02*
Weight change (%)	2.44 ± 1.26		2.49 ± 1.53	
BMI (kg/m^2)	21.64 ± 2.41*	22.15 ± 2.34*	20.22 ± 1.42*	20.7 ± 1.35*
Glucose (mmol/l)	4.71 ± 0.42	4.77 ± 0.42	4.65 ± 0.36	4.57 ± 0.28
Insulin (mU/ l)	5.49 ± 2.7	6.52 ± 2.43	5.45 ± 1.84*	6.33 ± 2.33*
Cholesterol (mmol/l)	4.35 ± 0.86	4.24 ± 0.69	4.21 ± 0.91	4.39 ± 0.87
HDL (mmol/l)	1.37 ± 0.32*	1.47 ± 0.33*	1.45 ± 0.43	1.54 ± 0.53
LDL (mmol/l)	2.41 ± 0.89	2.25 ± 0.69	2.32 ± 0.52	2.36 ± 0.47
LDL/HDL ratio	1.87 ± 1.02*	1.63 ± 0.68*	1.69 ± 0.46	1.64 ± 0.5
Triglycerides (mmol/l)	0.71 ± 0.26	0.81 ± 0.4	0.61 ± 0.21	0.56 ± 0.19

Study 2

40 healthy young men were recruited from the Danish National Birth Registry. 16 of them had a LBW (<10th percentile) and 24 had a NBW (between 50th and 90th percentile). Adipose tissue biopsies were taken after both control and a five-day high fat overfeeding (HFO) diet. Clinical characteristics can be found in Table 2.

Table 2. Clinical characteristics of the 40 young men in the HFO cohort

Data are presented as means ± SDs. *P<0.01 control diet compared to HFO using Wilcoxon's paired test.

	Control diet	**High-fat overfeeding**
Weight (kg)	78.3 ± 9.3	78.5 ± 9.8
BMI (kg/m^2)	24.0 ± 3.0	24.1 ± 3.1
Body fat (%)	19.1 ± 7.4	19.0 ± 7.3
WHR	0.88 ± 0.05	0.89 ± 0.05
HbA1c (%)	5.2 ± 0.3	5.2 ± 0.2
HbA1c (mmol/mol)	33	33
Fasting plasma glucose (mmol/l)	4.8 ± 0.5*	5.1 ± 0.4*

Study 3

14 monozygotic (MZ) twin pairs discordant for T2D were recruited from Scandinavian twin registries. Additionally, 120 unrelated individuals with normal glucose tolerance (NGT) or T2D were part of the Case-Control cohort 1 and 56 unrelated individuals with NGT or T2D were part of the Case-Control cohort 2. They were matched for age and sex. 10 MZ and 10 same-sex dizygotic (DZ) twin pairs (all with NGT) were used for heritability estimates. Clinical characteristics of all these cohorts can be found in Table 3.

Table 3. Clinical characteristics of study subjects included in the discordant twin cohort, the two case-control cohorts, and the twin cohort for heritability estimates
Data are presented as means ± SDs. *P<0.05 T2D compared to non-diabetic subjects.

	Discordant twins		Case-control cohort 1		Case-control cohort 2		Twin cohort for heritability estimates	
	Non-diabetic	T2D	NGT	T2D	NGT	T2D	MZ	DZ
n (male/female)	14 (9/5)	14 (9/5)	70 (32/38)	50 (26/24)	28 (15/13)	28 (15/13)	20 (10/10)	20 (10/10)
Age (years)	67.6 ± 7.7	67.6 ± 7.7	53.4 ± 7.4	55.5 ± 6.3	74.3 ± 4.3	74.5± 4.2	71.0 ± 5.4	70.6 ± 5.9
BMI (kg/m^2)	29.8 ± 6.8*	32.0 ± 7.1*	29.4 ± 5.9	29.6 ± 5.8	27.0 ± 3.6	27.4 ± 3.6	26.7 ± 2.1	26.5 ± 2.9
Fat%	30.5 ± 8.8*	33.6 ± 9.4*						
Fasting plasma glucose (mmol/l)	6.0 ± 0.5*	9.3 ± 2.6*	5.6 ± 0.5*	9.3 ± 3.0*	5.5 ± 0.5*	7.1 ± 1.9*	5.4 ± 0.3	5.4 ± 0.4
2-hour glucose (mmol/l)	8.3 ± 1.8*	16.1 ± 5.2*	6.0 ± 1.3*	16.5 ± 5.0*	6.6 ± 0.7*	15.1 ± 5.0*	6.1 ± 1.2	5.9 ± 1.0
HbA1c (%)	5.9 ± 0.4*	7.5 ± 1.8*	5.5 ± 0.3*	6.8 ± 1.1*	5.7 ± 0.3*	6.6 ± 1.3*	5.6 ± 0.19	5.7 ± 0.23
HbA1c (mmol/mol)	41.0 ± 4.4*	58.0 ± 19.7*	37 ± 3.3*	51 ± 12.0*	39.0 ± 3.3*	49.0 ± 14.0*	38.0 ± 2.1	39.0 ± 2.5

Study 4

26 healthy young men were recruited through The Danish National Birth Register with 13 of them being born with a LBW and 13 with NBW. They had no family history of diabetes, BMI < 30 kg/m^2 and their physical activity levels were below 10h/week. Abdominal subcutaneous adipose biopsies were obtained from the individuals after an overnight fast. Adipose precursor cells were isolated from the subcutaneous biopsies. Clinical characteristics can be found in Table 4.

Table 4. Clinical characteristics of adipose-derived stem cells (ADSC) donors
Data are presented as means ± SDs. *P<0.05 LBW compared to NBW.

	LBW (n = 13)	NBW (n = 13)
Birthweight (kg)	2.7 ± 0.1*	3.7 ± 0.2*
Age (years)	22.4 ± 1.7	23.2 ± 1.6
Weight (kg)	73.9 ± 7.3*	80.8 ± 5.4*
Height (cm)	178.0 ± 4.1	182.5 ± 6.9
BMI (kg/m^2)	23.3 ± 2.2	24.3 ± 2.0
WHR	0.90 ± 0.1	0.86 ± 0.1
Total fat mass (kg)	13.2 ± 4.0	13.1 ± 2.8
Trunk fat mass (%)	16.6 ± 4.7	15.4 ± 3.5
Lean body mass (kg)	58.5 ± 4.5*	64.4 ± 4.4*
Fasting insulin (pmol/l)	28.8 ± 17.4	37.3 ± 20.5
Fasting glucose (mmol/l)	4.9 ± 0.4	5.1 ± 0.2
HbA1c (%)	5.2 ± 0.3	5.1 ± 0.2
HbA1c (mmol/mol)	32.8 ± 3.1	32.2 ± 2.5

DNA methylation

DNA extraction and bisulfite conversion

Genomic DNA was extracted from adipose tissue specimens using either QiAamp DNA Mini kit or DNeasy Blood & Tissue kit (Qiagen, Hilden, Germany). DNA (500 ng) was treated with bisulfite using the EZ DNA Methylation kit (Zymo Research, Orange, CA).

Illumina 450k BeadChip array

DNA methylation was analyzed using Infinium HumanMethylation450 BeadChip array (Illumina, San Diego, CA, USA). With some variation the following procedure was used to analyze the data in these studies. Raw methylation data was exported from GenomeStudio software (Illumina) as intensity values that correspond to the number of methylated and unmethylated cells. The data was later processed using R (59) and Lumi (51,60–62) and Methylumi (63) packages. Intensity values were used to calculate M-values (logit function of percentage of methylation) (51) that were used for further analysis. Probes that targeted SNPs (58), non CpG-sites and cross-reactive probes (57) were removed. Low quality probes that are characterized by a detection p-value (provided by GenomeStudio) higher than 0.01 were also removed. Average signal of 614 negative control probes was subtracted from the signal of all remaining probes to remove the background noise. The data was quantile normalized to correct for the intra-array variation. Beta-mixture quantile (BMIQ) normalization was used to correct for probe design bias in Illumina 450k array (54). The data was batch corrected using ComBat (55).

Gene expression

Affymetrix GeneChip Human Gene ST Array chip

In the studies 1-3 mRNA expression was analyzed using Affymetrix GeneChip Human Gene ST arrays (Affymetrix, Santa Clara, CA, USA) following the Affymetrix protocol. R software (59) and the Oligo package (64) were used to process the data. The data was quantile normalized and then batch corrected using ComBat (55).

Illumina HumanHT-12 v4 Expression BeadChip

In the study 4 Illumina HumanHT-12 v4 Expression BeadChip arrays (Illumina, San Diego, CA, USA) were used to analyze mRNA expression. The data was processed using R (59) and Lumi package (51,60–62). The data was background corrected and quantile normalized. ComBat (55) was used for batch correction.

Statistical and bioinformatic analysis

Statistical and analysis tools

All statistical analyses were done using R software (59). Groups were compared using either Mann-Whitney U test or Wilcoxon signed-rank test using base R functionality. Linear regression models were used to study associations between DNA methylation or mRNA expression and multiple variables such as age, BMI and others.

Benjamini-Hochberg FDR (65) was used to correct for multiple testing.

Results

Study 1

7 weeks of excessive SFA or PUFA intake affected DNA methylation in adipose tissue differently.

We started with studying the effects of fat overfeeding on DNA methylation in adipose tissue without separating the two diets (n=31). We identified 1444 genes that significantly ($q < 0.05$) changed (>1%) mean DNA methylation of the CpG sites annotated to them after 7 weeks overfeeding. Only 2 of these genes showed a decrease, while the majority of genes showed increased methylation in response to fat overfeeding.

We then analyzed individual CpG sites and found that 4933 of them changed in response to the overfeeding diet with 4795 of them increasing methylation. We tested whether the genes annotated to those 4795 sites were related to any biological pathways using WebGestalt tool kit (http://bioinfo.vanderbilt.edu/webgestalt). The genes that were annotated to these sites were significantly enriched in 6 KEGG pathways including pathways in cancer, cell cycle, and protein processing in the endoplasmic reticulum.

Analysis of gene expression (n=31) showed that 1117 transcripts significantly changed after the 7-week diet ($q < 0.05$). We performed pathway analysis of 776 upregulated and 341 downregulated transcripts separately. Genes that are involved in carbohydrate metabolism (i.e., citrate cycle, pyruvate metabolism, starch and sucrose metabolism, and glycolysis and gluconeogenesis), lipid metabolism (i.e., biosynthesis of unsaturated fatty acids (FAs), glycerolipid metabolism, and FA metabolism), and oxidative phosphorylation were among the most significant KEGG pathways of genes that were upregulated by overfeeding.

We then compared the SFA and PUFA groups after 7 weeks of the dietary intervention. We found 22 genes that differed > 1% between the groups in mean DNA methylation of the CpG sites annotated to them ($q < 0.05$). 4 of them were higher in the SFA group and 18 were higher in the PUFA group. Genes that differed the most in the mean DNA methylation degree between the groups were *CYP2E1*, *C6orf138* and *DEFB115*. We also identified 4875 individual CpG sites that differed in methylation between the two diet groups ($q < 0.05$).

Then we separated the groups (SFA with n=17 and PUFA with n=14) and compared DNA methylation and gene expression before vs after the intervention. We found that SFA overfeeding only increased the mean methylation of 125 genes by >1% ($q < 0.05$), while PUFA overfeeding changed the mean methylation of 1797 genes by >1% (1795 genes increased, and 2 genes decreased) ($q < 0.05$) including *FTO*, *IL6* and *POMC*. Mean methylation of 38 genes was affected by both diets.

We identified 22 genes which expression was altered by the SFA diet ($q < 0.05$) but there were no significant changes found in response to the PUFA overfeeding.

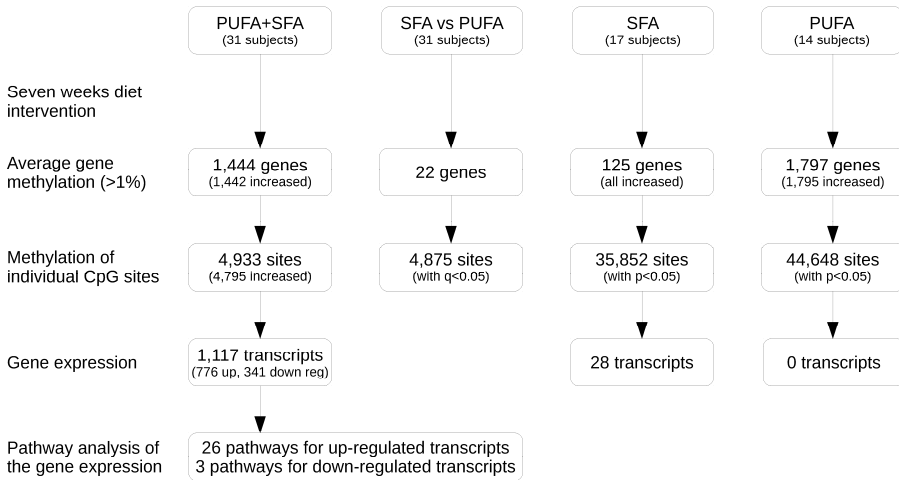

Figure 4.
Study design and workflow of the analysis of gene expression and DNA methylation data in study 1. Numbers represent results with $q < 0.05$ unless otherwise specified. CpG, Cytosine-phosphate-guanine; down reg, downregulated; up, upregulated.

Study 2

Young men with LBW were found to have epigenetic variation in adipose tissue that may influence insulin resistance and risk of T2D. 5 days of high-fat overfeeding influenced gene expression and, to some extent, DNA methylation in adipose tissue.

We compared adipose tissue from LBW and NBW men after the control diet and after the high-fat overfeeding (HFO) diet. We found 13 CpG sites that were differentially methylated between LBW and NBW after the control diet and 40 CpG sites after HFO. These sites were annotated to genes that have previously been associated with T2D including *ACAT1*, *CPLX1*, *FADS2*, *GPRC5B*, *HCCA2* and *IGF2R*. 5 of the 13 CpG sites that were different in LBW vs NBW men after the control diet were also differentially methylated after HFO ($p < 0.05$). Also 9 of the

40 CpG sites that were significantly different between birthweight groups during HFO were also differentially methylated during the control diet. No difference was found in the same comparisons for gene expression.

We next combined LBW and NBW groups to investigate the effects of fat overfeeding. 3276 genes were found to be differentially expressed ($q < 0.05$) in adipose tissue between the control and high fat diets. 59% of significantly different transcripts were upregulated and 41% were downregulated. Genes with the most pronounced increase included *ELOVL6*, *FADS2*, *NNAT* and *DGAT2* and genes with the most pronounced decrease included *SLC27A2* and *CIDEA*.

We used WebGestalt to test if the genes with altered expression in response to HFO could be linked to any biological pathways. Significantly upregulated transcripts were mapped to 40 KEGG pathways that predominantly related to metabolism. Oxidative phosphorylation, citrate cycle and pyruvate metabolism were the most significant KEGG pathways among transcripts that were upregulated by HFO. Many genes involved in the fatty acid metabolism, glycolysis/gluconeogenesis and biosynthesis of unsaturated fatty acids pathways were also upregulated by HFO. TGF-beta signaling was the only significantly enriched KEGG pathway among transcripts that were downregulated by HFO. Additionally, we found that many genes in the insulin signaling pathway were either upregulated (27 genes) or downregulated (16 genes).

Study 3

Gene expression patterns in adipose tissue varied between diabetic and non-diabetic monozygotic twins. DNA methylation differences were modest between MZ twins discordant for T2D, but they were pronounced in adipose tissue of T2D case-control cohort of unrelated subjects.

We compared gene expression in diabetic and NGT twins in the Discordant Twin Cohort and found that 197 genes were differentially expressed ($q < 0.15$). 116 of these were upregulated and 81 genes were downregulated in diabetic twins. The expression differences of these genes ranged from 5% to 43%.

We selected six genes to be replicated in an independent cohort (case-control cohort 1). The genes were selected based on their known functions related to fat metabolism (*ELOVL6* and *FADS1*), glucose metabolism (*GYS2*) and inflammation (*SPP1* [*OPN*], *CCL18*, and *IL1RN*). All six genes were replicated using Real-Time PCR and showed different expression between unrelated patients with diabetes and non-diabetic subjects. We further looked at the relation between the expression of these genes and obesity in the NGT part of the cohort. We found that the expression of *ELOVL6* and *GYS2* was significantly lower and the expression of *SPP1* and

IL1RN was significantly higher in obese versus non-obese NGT individuals. The expression levels of the 6 genes correlated significantly with BMI and/or glucose levels in the NGT individuals supporting that their expression is regulated already before the disease develops.

We did not find any CpG site to be significantly differentially methylated between diabetic and NGT twins in the Discordant Twin Cohort after controlling for multiple testing (FDR). However, when we analyzed a bigger T2D case-control cohort of unrelated individuals matched for age and sex (case-control cohort 2), 15,627 CpG sites annotated to 7,046 unique genes were differentially methylated between the diabetic vs non-diabetic subjects with $q < 0.15$. These data suggest that genetic components in the twins may control the methylation pattern. Furthermore, 1,410 CpG sites were also differentially methylated in the same direction in the discordant twins based on $P < 0.05$ as the ones found in the case-control cohort 2.

Study 4

DNA methylation and RNA expression patterns were different between adipose-derived stem cells from individuals with LBW and NBW.

506 genes were differentially expressed ($q<0.05$) in ADSC between LBW and NBW groups. 281 of those were upregulated and 225 were downregulated in LBW. These include *CCNT2* downregulated in LBW and *STAT2* upregulated in LBW. However, no genes were found to be differentially expressed between the birth weight groups after the differentiation.

Principal component analysis suggested a correlation between global DNA methylation levels and birth weight in ADSCs but not in differentiated adipocytes. No individual CpG sites were identified as differentially methylated when comparing LBW and NBW in ADSCs or after differentiation.

Discussion

Several technologies are available for studying DNA methylation. Methods such as pyrosequencing are used for analyzing DNA methylation of specific genes or regions. Microarrays such as Illumina 450k or more recent Illumina MethylationEPIC 850k allow interrogating hundreds of thousands CpG sites at the same time. They cover 2-3% of all CpG sites in the human genome but these sites are strategically chosen to cover a large number of CpG islands, known gene bodies and promoter regions. Whole-genome bisulfite sequencing (WGBS) is the technology that allows studying of DNA methylation across all the genome.

Illumina 450k was chosen for the studies in this thesis as the most comprehensive and cost-effective solution available at the time. There were certain challenges associated with analysis of the Illumina 450k data some of which are specific for this chip and some are applied to analysis of microarray data in general.

In the analyses of microarray data, it is possible to characterize the variation in the data as either related to biological differences or as technical variation. While it is desirable to detect these biological differences, technical variation may make it harder to do. Technical variation may be due to differences in sample preparation, production of the arrays and so on. It can be differences between samples or groups of samples (in this case referred to as batch effect).

Various normalization methods were proposed to deal with technical variation on the individual sample level (66). Quantile normalization remains one of widely used such methods (60). In short, the method consists in sorting the data for each sample from highest to lowest value, then substituting the values with a mean value across all the samples for that particular rank. So that the highest value for all the samples becomes the mean of the highest values, then the same for the second highest and so on. It effectively makes the distribution of probe intensities for each array identical.

Quantile normalization aims at making possible direct comparisons between the samples, but it does not take into consideration batch effects. This may or may not represent a problem depending on type of analysis that needs to be done. In case of microarray data one of the biggest sources of batch effect is placement at different microarray chips. In case of pairwise comparisons it is possible to place the samples from the same pairs at the same chip in which case batch effect does not affect the

outcome of the statistical analysis. However, statistical analyses such as linear models require to take batch effects into consideration. First and a very important step is to assure even distribution of the samples across the batches based on the variables of interest in order to avoid correlation between batch effects and a variable of interest (67,68). Several methods exist for batch correction such as surrogate variable analysis (*sva*) (69,70) and ComBat (55). *sva* is based on identifying and estimating surrogate variables for unknown sources of variation while ComBat directly removes known batch effects using an empirical Bayes approach. sva may be preferable in case there is a number of potential unknown confounders while ComBat may be more appropriate in case there are known batch variables and the studied biological groups are known to be heterogeneous (56).

There were certain challenges associated specifically with the analysis of Illumina 450k data.

Soon after the array was available it became apparent that due to different chemistry the distributions of DNA methylation values produced by type I and type II probes of the array were very different and higher dynamic range of type I probes created a type I enrichment bias. This eventually led to developing of within-sample normalization methods such as BMIQ normalization designed to adjust the statistical distribution of type II probes into a statistical distribution of type I probes.

A significant number of probes was found to be cross-reactive meaning that the measured values could not be unequivocally matched to a specific genomic location. Some probes were targeting CpG sites with known SNPs. All these probes have to be excluded from the analyses since the measurements can be due to the factors unrelated to the variables of interest.

In statistics type I error is defined as the rejection of a true null hypothesis, when type II error is defined as not rejecting a false null hypothesis. In other words, type I error can be labeled as “false positive”, while type II error as “false negative”. Finding a cut-off to define statistical significance is a trade-off between type I and type II errors. Study of genomics data often means that thousands of tests are done simultaneously. In case of one test a result is said to be statistically significant if it is very unlikely to have occurred given the null hypothesis. But increasing the number of tests done simultaneously, drives up the probability of getting a significant result just by chance. Several methods were proposed to address this issue.

Familywise error rate (FWER) is the probability of making at least one Type I error. Bonferroni correction is one of procedures implementing it and it is widely used to correct for multiple testing. However, by controlling the probability of at least one Type I error Bonferroni correction can be conservative, can be prone to producing “false negatives” thus reducing the statistical power (71).

With advancement of high-throughput technologies alternative methods for correction for multiple testing were proposed. For technological and financial reasons, it is common that thousands of variables (e.g. genes or CpG sites) are tested with relatively small samples sizes. In this case the number of type II errors produced by FWER became less acceptable. Alternative methods such as false discovery rate (FDR) instead of controlling the probability of at least one Type I error control the expected proportion of "false positives" (65). One of the assumptions of FDR is about p-values being independent or weakly dependent. It has been shown (72) that in human genome exist differentially methylated regions (DMR) meaning that that assumption may not always be correct when studying DNA methylation in humans. As described by J. Storey in his article (73) taking into consideration possible dependencies between p-values may be an area of improvement for the correction for multiple testing in the future.

As discussed in (74) epigenetics might in part mediate gene–environment interactions which play role in development of common and complex diseases. In particular it has been hypothesized that DNA methylation might be one of the mechanisms that link risk factors for T2D such as diet and impaired intrauterine environment to cellular functions in tissues such as adipose tissue. In the studies, included in this thesis, we looked at the relation between genome-wide methylation patterns in adipose tissue and these risk factors. Studies 1 and 2 provide insights into the impact of high-fat diets on the epigenome, while study 4 looks into correlations between impaired intrauterine environment and epigenetic differences in stem cells derived from adipose tissue of grown adults. Study 3 aimed at eliminating as many confounding factors (such as genetics, age, sex) as possible by studying epigenome in monozygotic twins discordant for T2D. It is possible that the lack of success in trying to identify significant differences in methylation between the twins after FDR is due to the small sample size as we analyzed only 14 twin pairs across approximately 450,000 CpG sites. About a hundred of CpG sites had the lowest possible p-value given the sample size and were not significant after FDR. Another reason could be that genetics is a strong component that control DNA methylation more than disease (75,76). It is also possible that the sample size in the stem cell study was too small and the study underpowered for detection of altered methylation. Future genome-wide epigenetic studies should hence aim to be performed in large human cohorts in order to increase the statistical power. However, one cannot rule out that there are no differences in methylation between the two birthweight groups, without increasing the sample size.

Summary and general conclusion

Type 2 diabetes is a common metabolic disease and its prevalence is increasing worldwide. Adipose tissue plays an important role in metabolic processes. Environmental factors may affect metabolic phenotypes and epigenetics may mediate this influence. We used Illumina 450k microarrays to study correlations between epigenetic patterns in human adipose tissue and environmental factors such as high-fat diets, impaired intrauterine environment and low birth weight as well as with type 2 diabetes.

In study 1 we investigated whether the DNA methylation pattern in human adipose tissue is affected differently by a 7-week exposure to a diet high in either polyunsaturated fatty acids or saturated fatty acids. We found differences in methylation related to general fat overfeeding as well as between the two dietary groups.

In study 2 we studied the effect of 5 days of overfeeding with a high fat diet on DNA methylation in adipose tissue compared to a control diet. The participants were either born with a normal or low birth weight and it allowed also to compare the effects of the diets on these two groups separately. There were differences in methylation between subjects born with a low compared with normal birth weight.

In study 3 we used adipose tissue of a cohort of 14 monozygotic twin pairs discordant for type 2 diabetes in order to eliminate most of known confounding factors in comparing DNA methylation patterns as well as a case-control cohort for T2D. We found numerous differences in methylation in adipose tissue from subjects with T2D compared with controls, while less differences were found in the discordant twins.

In study 4 we compared DNA methylation in adipose-derived stem and differentiated cells in subjects born with a low or normal birth weight. We found differences in gene expression in adipose-derived stem cells between LBW and NBW groups but did not identify differences in methylation patterns possibly due to lack of statistical power.

Overall, these studies contribute to a better understanding of the influence of environmental factors and type 2 diabetes on DNA methylation in human adipose tissue.

Acknowledgements

My dear colleagues, my friends, thank you to all of you for making these last five years truly an amazing experience. You made it all possible, inspired me and helped me along the way. Looking back at me at the beginning of this journey I can hardly recognize myself. These were years of huge personal transformations and growth, years of learning and realizing how much there is to learn. And each one of you made an impact and I can only say again thank you!

References

1. International Diabetes Federation. IDF diabetes atlas 2017 [Internet]. [cited 2018 Sep 24]. Available from: http://www.diabetesatlas.org/across-the-globe.html
2. Meetoo D, McGovern P, Safadi R. An epidemiological overview of diabetes across the world. Br J Nurs Mark Allen Publ. 2007 Sep 13;16(16):1002–7.
3. Prasad RB, Groop L. Genetics of type 2 diabetes-pitfalls and possibilities. Genes. 2015 Mar 12;6(1):87–123.
4. Daousi C, Casson IF, Gill GV, MacFarlane IA, Wilding JPH, Pinkney JH. Prevalence of obesity in type 2 diabetes in secondary care: association with cardiovascular risk factors. Postgrad Med J. 2006 Apr;82(966):280–4.
5. Mykkänen L, Kuusisto J, Pyörälä K, Laakso M. Cardiovascular disease risk factors as predictors of type 2 (non-insulin-dependent) diabetes mellitus in elderly subjects. Diabetologia. 1993 Jun;36(6):553–9.
6. Lyssenko V, Jonsson A, Almgren P, Pulizzi N, Isomaa B, Tuomi T, et al. Clinical risk factors, DNA variants, and the development of type 2 diabetes. N Engl J Med. 2008 Nov 20;359(21):2220–32.
7. Hu FB. Globalization of diabetes: the role of diet, lifestyle, and genes. Diabetes Care. 2011 Jun;34(6):1249–57.
8. Noble D, Mathur R, Dent T, Meads C, Greenhalgh T. Risk models and scores for type 2 diabetes: systematic review. BMJ. 2011 Nov 28;343:d7163.
9. Lyssenko V, Laakso M. Genetic screening for the risk of type 2 diabetes: worthless or valuable? Diabetes Care. 2013 Aug;36 Suppl 2:S120-126.
10. Kahn SE, Cooper ME, Del Prato S. Pathophysiology and treatment of type 2 diabetes: perspectives on the past, present, and future. Lancet Lond Engl. 2014 Mar 22;383(9922):1068–83.
11. Jensen CB, Storgaard H, Dela F, Holst JJ, Madsbad S, Vaag AA. Early differential defects of insulin secretion and action in 19-year-old caucasian men who had low birth weight. Diabetes. 2002 Apr;51(4):1271–80.
12. Rasmussen EL, Malis C, Jensen CB, Jensen J-EB, Storgaard H, Poulsen P, et al. Altered fat tissue distribution in young adult men who had low birth weight. Diabetes Care. 2005 Jan;28(1):151–3.
13. Gale CR, Martyn CN, Kellingray S, Eastell R, Cooper C. Intrauterine programming of adult body composition. J Clin Endocrinol Metab. 2001 Jan;86(1):267–72.
14. Whincup PH, Kaye SJ, Owen CG, Huxley R, Cook DG, Anazawa S, et al. Birth weight and risk of type 2 diabetes: a systematic review. JAMA. 2008 Dec 24;300(24):2886–97.

15. Hemminki K, Li X, Sundquist K, Sundquist J. Familial risks for type 2 diabetes in Sweden. Diabetes Care. 2010 Feb;33(2):293–7.
16. Medici F, Hawa M, Ianari A, Pyke DA, Leslie RD. Concordance rate for type II diabetes mellitus in monozygotic twins: actuarial analysis. Diabetologia. 1999 Feb;42(2):146–50.
17. Wheeler E, Barroso I. Genome-wide association studies and type 2 diabetes. Brief Funct Genomics. 2011 Mar;10(2):52–60.
18. Scott RA, Scott LJ, Mägi R, Marullo L, Gaulton KJ, Kaakinen M, et al. An Expanded Genome-Wide Association Study of Type 2 Diabetes in Europeans. Diabetes. 2017;66(11):2888–902.
19. Cusi K. The role of adipose tissue and lipotoxicity in the pathogenesis of type 2 diabetes. Curr Diab Rep. 2010 Aug;10(4):306–15.
20. Ouchi N, Parker JL, Lugus JJ, Walsh K. Adipokines in inflammation and metabolic disease. Nat Rev Immunol. 2011 Feb;11(2):85–97.
21. Smith U. Impaired ('diabetic') insulin signaling and action occur in fat cells long before glucose intolerance--is insulin resistance initiated in the adipose tissue? Int J Obes Relat Metab Disord J Int Assoc Study Obes. 2002 Jul;26(7):897–904.
22. Brøns C, Jacobsen S, Hiscock N, White A, Nilsson E, Dunger D, et al. Effects of high-fat overfeeding on mitochondrial function, glucose and fat metabolism, and adipokine levels in low-birth-weight subjects. Am J Physiol Endocrinol Metab. 2012 Jan 1;302(1):E43-51.
23. Cornier M-A, Bergman BC, Bessesen DH. The effects of short-term overfeeding on insulin action in lean and reduced-obese individuals. Metabolism. 2006 Sep;55(9):1207–14.
24. Brøns C, Jensen CB, Storgaard H, Hiscock NJ, White A, Appel JS, et al. Impact of short-term high-fat feeding on glucose and insulin metabolism in young healthy men. J Physiol. 2009 May 15;587(Pt 10):2387–97.
25. Kelley DE, Wing R, Buonocore C, Sturis J, Polonsky K, Fitzsimmons M. Relative effects of calorie restriction and weight loss in noninsulin-dependent diabetes mellitus. J Clin Endocrinol Metab. 1993 Nov;77(5):1287–93.
26. Alligier M, Meugnier E, Debard C, Lambert-Porcheron S, Chanseaume E, Sothier M, et al. Subcutaneous adipose tissue remodeling during the initial phase of weight gain induced by overfeeding in humans. J Clin Endocrinol Metab. 2012 Feb;97(2):E183-192.
27. Shea J, French CR, Bishop J, Martin G, Roebothan B, Pace D, et al. Changes in the transcriptome of abdominal subcutaneous adipose tissue in response to short-term overfeeding in lean and obese men. Am J Clin Nutr. 2009 Jan;89(1):407–15.
28. Rosqvist F, Iggman D, Kullberg J, Cedernaes J, Johansson H-E, Larsson A, et al. Overfeeding polyunsaturated and saturated fat causes distinct effects on liver and visceral fat accumulation in humans. Diabetes. 2014 Jul;63(7):2356–68.
29. Franck N, Gummesson A, Jernås M, Glad C, Svensson P-A, Guillot G, et al. Identification of adipocyte genes regulated by caloric intake. J Clin Endocrinol Metab. 2011 Feb;96(2):E413-418.

30. Summers LKM, Fielding BA, Bradshaw HA, Ilic V, Beysen C, Clark ML, et al. Substituting dietary saturated fat with polyunsaturated fat changes abdominal fat distribution and improves insulin sensitivity. Diabetologia. 2002 Mar;45(3):369–77.
31. Tiikkainen M, Bergholm R, Vehkavaara S, Rissanen A, Häkkinen A-M, Tamminen M, et al. Effects of identical weight loss on body composition and features of insulin resistance in obese women with high and low liver fat content. Diabetes. 2003 Mar;52(3):701–7.
32. Petersson H, Arnlöv J, Zethelius B, Risérus U. Serum fatty acid composition and insulin resistance are independently associated with liver fat markers in elderly men. Diabetes Res Clin Pract. 2010 Mar;87(3):379–84.
33. Yki-Järvinen H. Liver fat in the pathogenesis of insulin resistance and type 2 diabetes. Dig Dis Basel Switz. 2010;28(1):203–9.
34. Fabbrini E, Magkos F, Mohammed BS, Pietka T, Abumrad NA, Patterson BW, et al. Intrahepatic fat, not visceral fat, is linked with metabolic complications of obesity. Proc Natl Acad Sci U S A. 2009 Sep 8;106(36):15430–5.
35. Mozaffarian D, Micha R, Wallace S. Effects on coronary heart disease of increasing polyunsaturated fat in place of saturated fat: a systematic review and meta-analysis of randomized controlled trials. PLoS Med. 2010 Mar 23;7(3):e1000252.
36. Astrup A, Dyerberg J, Elwood P, Hermansen K, Hu FB, Jakobsen MU, et al. The role of reducing intakes of saturated fat in the prevention of cardiovascular disease: where does the evidence stand in 2010? Am J Clin Nutr. 2011 Apr;93(4):684–8.
37. Risérus U, Willett WC, Hu FB. Dietary fats and prevention of type 2 diabetes. Prog Lipid Res. 2009 Jan;48(1):44–51.
38. Iggman D, Rosqvist F, Larsson A, Arnlöv J, Beckman L, Rudling M, et al. Role of dietary fats in modulating cardiometabolic risk during moderate weight gain: a randomized double-blind overfeeding trial (LIPOGAIN study). J Am Heart Assoc. 2014 Oct 15;3(5):e001095.
39. Berger SL, Kouzarides T, Shiekhattar R, Shilatifard A. An operational definition of epigenetics. Genes Dev. 2009 Apr 1;23(7):781–3.
40. Deaton AM, Bird A. CpG islands and the regulation of transcription. Genes Dev. 2011 May 15;25(10):1010–22.
41. Laird PW. Principles and challenges of genomewide DNA methylation analysis. Nat Rev Genet. 2010 Mar;11(3):191–203.
42. Subramaniam D, Thombre R, Dhar A, Anant S. DNA methyltransferases: a novel target for prevention and therapy. Front Oncol. 2014;4:80.
43. Xu F, Mao C, Ding Y, Rui C, Wu L, Shi A, et al. Molecular and enzymatic profiles of mammalian DNA methyltransferases: structures and targets for drugs. Curr Med Chem. 2010;17(33):4052–71.
44. Han J-A, An J, Ko M. Functions of TET Proteins in Hematopoietic Transformation. Mol Cells. 2015 Nov;38(11):925–35.
45. Rawłuszko-Wieczorek AA, Siera A, Jagodziński PP. TET proteins in cancer: Current "state of the art." Crit Rev Oncol Hematol. 2015 Dec;96(3):425–36.

46. Fouse SD, Nagarajan RO, Costello JF. Genome-scale DNA methylation analysis. Epigenomics. 2010 Feb;2(1):105–17.
47. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet. 2012 May 29;13(7):484–92.
48. Takai D, Jones PA. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proc Natl Acad Sci U S A. 2002 Mar 19;99(6):3740–5.
49. Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, et al. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc Natl Acad Sci U S A. 1992 Mar 1;89(5):1827–31.
50. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, et al. High density DNA methylation array with single CpG site resolution. Genomics. 2011 Oct;98(4):288–95.
51. Du P, Zhang X, Huang C-C, Jafari N, Kibbe WA, Hou L, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinformatics. 2010 Nov 30;11:587.
52. Durbin BP, Hardin JS, Hawkins DM, Rocke DM. A variance-stabilizing transformation for gene-expression microarray data. Bioinforma Oxf Engl. 2002;18 Suppl 1:S105-110.
53. Dedeurwaerder S, Defrance M, Calonne E, Denis H, Sotiriou C, Fuks F. Evaluation of the Infinium Methylation 450K technology. Epigenomics. 2011 Dec;3(6):771–84.
54. Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinforma Oxf Engl. 2013 Jan 15;29(2):189–96.
55. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostat Oxf Engl. 2007 Jan;8(1):118–27.
56. Leek J, Johnson W, Parker H, Fertig E, Jaffe A, Storey J, et al. sva: Surrogate Variable Analysis. 2018.
57. Chen Y, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. Epigenetics. 2013 Feb;8(2):203–9.
58. Price ME, Cotton AM, Lam LL, Farré P, Emberly E, Brown CJ, et al. Additional annotation enhances potential for biologically-relevant analysis of the Illumina Infinium HumanMethylation450 BeadChip array. Epigenetics Chromatin. 2013 Mar 3;6(1):4.
59. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2018. Available from: https://www.R-project.org/
60. Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. Bioinforma Oxf Engl. 2008 Jul 1;24(13):1547–8.
61. Lin SM, Du P, Huber W, Kibbe WA. Model-based variance-stabilizing transformation for Illumina microarray data. Nucleic Acids Res. 2008 Feb;36(2):e11.

62. Du P, Kibbe WA, Lin SM. nuID: a universal naming scheme of oligonucleotides for illumina, affymetrix, and other microarrays. Biol Direct. 2007 May 31;2:16.
63. Davis S, Du P, Bilke S, Triche, Jr T, Bootwalla M. methylumi: Handle Illumina methylation data. 2017.
64. Carvalho BS, Irizarry RA. A framework for oligonucleotide microarray preprocessing. Bioinforma Oxf Engl. 2010 Oct 1;26(19):2363–7.
65. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B Methodol. 1995;57(1):289–300.
66. Bolstad BM, Irizarry RA, Åstrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics. 2003 Jan 22;19(2):185–93.
67. Leek JT, Scharpf RB, Bravo HC, Simcha D, Langmead B, Johnson WE, et al. Tackling the widespread and critical impact of batch effects in high-throughput data. Nat Rev Genet. 2010;11(10):733–9.
68. Buhule OD, Minster RL, Hawley NL, Medvedovic M, Sun G, Viali S, et al. Stratified randomization controls better for batch effects in 450K methylation analysis: a cautionary tale. Front Genet. 2014;5:354.
69. Leek JT, Storey JD. Capturing heterogeneity in gene expression studies by surrogate variable analysis. PLoS Genet. 2007 Sep;3(9):1724–35.
70. Leek JT, Storey JD. A general framework for multiple testing dependence. Proc Natl Acad Sci U S A. 2008 Dec 2;105(48):18718–23.
71. Nakagawa S. A farewell to Bonferroni: the problems of low statistical power and publication bias. Behav Ecol. 2004 Nov 1;15(6):1044–5.
72. Ziller MJ, Gu H, Müller F, Donaghey J, Tsai LT-Y, Kohlbacher O, et al. Charting a dynamic DNA methylation landscape of the human genome. Nature. 2013 Aug 22;500(7463):477–81.
73. Storey JD. False Discovery Rates. 2011;7.
74. Feinberg AP. The Key Role of Epigenetics in Human Disease Prevention and Mitigation. N Engl J Med. 2018 Apr 5;378(14):1323–34.
75. Liu Y, Aryee MJ, Padyukov L, Fallin MD, Hesselberg E, Runarsson A, et al. Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. Nat Biotechnol. 2013 Feb;31(2):142–7.
76. Volkov P, Olsson AH, Gillberg L, Jørgensen SW, Brøns C, Eriksson K-F, et al. A Genome-Wide mQTL Analysis in Human Adipose Tissue Identifies Genetic Variants Associated with DNA Methylation, Gene Expression and Metabolic Traits. PloS One. 2016;11(6):e0157776.

Printed by Media-Tryck, Lund 2018 NORDIC SWAN ECOLABEL 3041 0903

LUND UNIVERSITY

FACULTY OF MEDICINE

Department of Clinical Sciences

Lund University, Faculty of Medicine
Doctoral Dissertation Series 2018:136
ISBN 978-91-7619-704-2
ISSN 1652-8220