Clinical Biochemistry xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Clinical Biochemistry

journal homepage: www.elsevier.com/locate/clinbiochem

# Differential DNA methylation of microRNAs within promoters, intergenic and intragenic regions of type 2 diabetic, pre-diabetic and non-diabetic individuals

Carmen Pheiffer [a,*,1], Rajiv T. Erasmus [b,c], Andre P. Kengne [c,d], Tandi E. Matsha [e,**,1]

[a] Biomedical Research and Innovation Platform, South African Medical Research Council, Cape Town, South Africa
[b] Division of Chemical Pathology, Faculty of Medicine and Health Sciences, National Health Laboratory Service (NHLS) and University of Stellenbosch, Cape Town, South Africa
[c] Non-Communicable Diseases Research Unit, South African Medical Research Council, Cape Town, South Africa
[d] Department of Medicine, University of Cape Town, Cape Town, South Africa
[e] Department of Biomedical Sciences, Faculty of Health and Wellness Sciences, Cape Peninsula University of Technology, Cape Town, South Africa

ARTICLE INFO

*Article history:*
Received 6 May 2015
Received in revised form 24 November 2015
Accepted 30 November 2015
Available online xxxx

*Keywords:*
DNA methylation
MicroRNAs
Type 2 diabetes
Pre-diabetes
Promoters
Intragenic regions
Intragenic regions

ABSTRACT

**Objective:** Accumulating evidence supports the role of epigenetic modifications, and in particular DNA methylation and non-coding RNAs in the pathophysiology of type 2 diabetes. Alterations in methylation patterns within promoter regions are linked with aberrant transcription and pathological gene expression; however the role of methylation within non-promoter regions is not yet fully elucidated.

**Design and methods:** We performed whole genome methylated DNA immunoprecipitation sequencing (MeDIP-Seq) in peripheral-blood-derived DNA from age–gender–body mass index (BMI)–ethnicity matched type 2 diabetic, pre-diabetic and non-diabetic individuals.

**Results:** The density of methylation normalized to the average length of the promoter, intergenic and intragenic regions and to CpG count was 3.17, 9.80 and 0.09 for the promoter, intergenic and intragenic regions, respectively. Methylation within these regions varied according to glucose tolerance status and was associated with hypermethylation rather than hypomethylation. MicroRNA–DNA methylation peaks accounted for 4.8% of the total number of peaks detected. Differential DNA methylation of these microRNA peaks was observed during dysglycemia, with the promoter, intergenic and intragenic regions accounting for 2%, 95% and 3% respectively, of the differentially methylated microRNA peaks.

**Conclusion:** Genome-wide DNA methylation varied according to glucose tolerance. Methylation within non-promoter regions accounted for the majority of differentially methylated peaks identified, thus highlighting the importance of DNA methylation within these non-promoter regions in the pathogenesis of type 2 diabetes. This study suggests that DNA methylation within intergenic regions is a mechanism regulating microRNAs, another increasingly important epigenetic factor, during type 2 diabetes.



## 1. Introduction

Type 2 diabetes mellitus (T2D) continues to be a major source of morbidity and mortality worldwide [41]. In 2014, an estimated 387 million people worldwide had diabetes mellitus, with the figure projected to increase to 592 million by 2035 [17]. Epigenetics reflect the interplay between genetics and environmental factors, and has attracted considerable interest to explain the missing heritability of T2D [27] or to gain insight into the pathogenesis of the disease [7,38]. These include DNA methylation, histone modifications, chromatin remodeling and small non-coding RNAs, with the former being the most studied [23].

DNA methylation, which refers to the addition of methyl groups to the 5′ position on cytosine nucleotides, primarily in CpG islands in the promoter regions of genes [4], is catalyzed by the enzyme DNA methyltransferase (DNMT). This epigenetic modification alters chromatin structure and regulates gene expression by transcriptional activation or repression of genes, thus affecting phenotype. Aberrant DNA methylation leads to genomic and chromosomal instability and has been demonstrated in many diseases [13,35]. Characterization of altered DNA methylation during disease processes could thus give insight into the pathophysiology of the disease, and reveal novel diagnostic, prognostic and therapeutic targets [23,35].

Recent advances in techniques to study genome-wide methylation patterns have facilitated the identification of significant DNA

* Correspondence to: C. Pheiffer, Biomedical Research and Innovation Platform, South African Medical Research Council,PO Box 19070, Tygerberg, 7505 Cape Town, South Africa.
** Correspondence to: T. E. Matsha, Biomedical Sciences, Faculty of Health and Wellness Sciences, Cape Peninsula University of Technology, PO Box 1906, Bellville, 7530 Cape Town, South Africa.
*E-mail addresses:* carmen.pheiffer@mrc.ac.za (C. Pheiffer), matshat@cput.ac.za, tandimatsha@gmail.com (T.E. Matsha).
[1] CP and TEM contributed equally to the manuscript.

http://dx.doi.org/10.1016/j.clinbiochem.2015.11.021
0009-9120/

methylation in intergenic and intragenic (genebody) regions [6,19,36]. It is speculated that methylation within these non-promoter regions regulate alternative promoters, RNA processing, transposable elements such as long interspersed elements (LINEs) and short interspersed elements (SINEs), and non-coding RNAs [20].

Non-coding RNAs (ncRNAs) consist of long non-coding (lncRNAs) and microRNAs (miRNAs), and represent another important epigenetic mechanism, with the latter being the most studied. MicroRNAs are approximately 18 to 25 nucleotide RNA molecules that regulate gene expression at the post-transcriptional level by binding to the 3′ untranslated region (UTR) of their target messenger RNA (mRNA) resulting in mRNA degradation or translational repression [12]. Studies show that miRNA expression is dysregulated during disease, thus identifying these epigenetic modifications as key factors in disease pathogenesis [14]. More recently, a link between DNA methylation and miRNA expression was reported [15,22,29,33], aiding our understanding of how these two epigenetic mechanisms interact to regulate gene expression and contribute to disease.

In this study we investigated genome-wide DNA methylation in peripheral blood cell-derived DNA of type 2 diabetic, pre-diabetic and non-diabetic individuals. DNA methylation was characterized according to genomic location (promoter, intergenic, intragenic) and further quantitatively compared between the three groups. MiRNA-differentially methylated regions were further analyzed.

## 2. Methods

### 2.1. Study setting and subjects

The study setting has been described elsewhere [25]. Briefly, participants were members of a cohort study conducted in Bellville-South, Cape Town, a mixed ancestry township formed in the late 1950s. The cohort from which three screen detected diabetic, three pre-diabetic and three normo-glycemic age–gender–body mass index (BMI)–duration of residency matched participants were selected for the current study was initiated in April 2014. Participants for this study had no history of doctor diagnosed diabetes mellitus, thus underwent a 75 g oral glucose tolerance test (OGTT) as prescribed by the WHO [1]. Pre-diabetes and diabetes were diagnosed according to World Health Organization (WHO) guidelines [1]. Ethical approval for the study was obtained from the Research Ethics Committees of the Cape Peninsula University of Technology and Stellenbosch University (respectively, NHREC: REC-230 408–014 and N14/01/003).

### 2.2. DNA extraction

Genomic DNA was extracted from peripheral blood using the Wizard® Genomic DNA Purification Kit (Promega, Madison, WI, USA) according to the manufacturer's instructions. Briefly, white blood cells were lysed, thereafter cellular proteins were removed by salt precipitation, and high molecular weight genomic DNA left in solution was then concentrated and desalted by isopropanol precipitation. At least 2 μg of DNA (concentrations ranging between 70 ng/μL and 130 ng/μL) with A260/A280 and A260/A230 ratios ≥1.8 was shipped frozen on dry ice, as instructed by Arraystar Inc. (Rockville, MD, USA).

### 2.3. Sequencing library preparation and sequencing

Methylated DNA immunoprecipitation (MeDIP) was performed by Arraystar Inc. according to Down et al. [10] with minor modifications. Briefly, genomic DNA was sonicated to ~200–900 bp with a Bioruptor sonicator (Diagenode, Denville, NJ, USA). Thereafter, 800 ng of sonicated DNA was end-repaired, A-tailed, and ligated to single-end adapters following the standard Illumina protocol. After agarose size selection to remove unligated adapters, the adaptor-ligated DNA was used for immunoprecipitation using a human monoclonal anti-5-methylcytosine antibody (Diagenode). DNA was heat-denatured at 94 °C for 10 min, rapidly cooled on ice, and immunoprecipitated with 1 μL of primary antibody overnight at 4 °C with rocking agitation in 400 μL of immunoprecipitation buffer (0.5% BSA in PBS). Immunoprecipitated DNA fragments were recovered by adding 100 μL of protein G magnetic beads (Life Technologies, Carlsbad, CA, USA) and incubated for an additional 2 h at 4 °C with agitation. Thereafter, a total of five immunoprecipitation washes with ice-cold immunoprecipitation buffer were performed. As a negative control, nonspecific human IgG immunoprecipitation was performed in parallel to methyl DNA immunoprecipitation. Washed beads were resuspended in TE buffer with 0.25% SDS and 0.25 mg/mL proteinase K for 2 h at 65 °C and allowed to cool down to room temperature. MeDIP and supernatant DNA were purified using Qiagen MinElute columns and eluted in 16 μL EB (Qiagen, Germantown, MD, USA). Fourteen cycles of PCR were performed on 5 μL of the immunoprecipitated DNA using the single-end Illumina PCR primers. Amplicons were purified with Qiagen MinElute columns, after which a final size selection (300–1000 bp) was performed by electrophoresis in 2% agarose. Libraries were quality controlled with the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). An aliquot of each library was diluted in EB (Qiagen) to 5 ng/μL and 1 μL was used in real-time PCR reactions to confirm enrichment of methylated region. The enrichment of DNA immunoprecipitation was analyzed by qPCR using specific methylated sites at H19 locus and non-methylated sites at GAPDH. The library was denatured with 0.1 M NaOH to generate single-stranded DNA molecules, and loaded onto channels of the flow cell at 8 pM concentration, amplified in situ using TruSeq Rapid SR Cluster Kit (Illumina, San Diego, CA, USA). Sequencing was performed by running 100 cycles on the Illumina HiSeq 2000 according to the manufacturer's instructions. The Agilent 2100 Bioanalyzer was used for assessment of the quality and concentration of the sequencing library, while the size and concentration of each sample was determined after sequencing library preparation.

### 2.4. Data analysis

After the sequencing platform generated the sequencing images, the stages of image analysis and base calling were performed using Off-Line Basecaller software (OLB V1.8). After passing Solexa CHASTITY quality filter, the clean reads were aligned to the human genome (UCSC HG19) using BOWTIE software (V2.1.0). MeDIP peaks were identified by MACS2 and differentially methylated regions (DMRs) were identified by MANorm. Statistically significant MeDIP-enriched regions (peaks) detected by MACS2 were identified by comparison to a Poisson background model, using a q-value threshold of $10^{-2}$. The peaks in samples were annotated by the nearest gene (the nearest TSS to the center of the peak region) using the newest UCSC RefSeq database (UCSC Genome Browser, University of California, Santa Cruz). Peaks were divided into three classes on the basis of their distances to UCSC RefSeq genes:

1. Promoter peaks: Promoters were defined as 2000 bp upstream and downstream from the transcription start site (TSS). Peaks whose centers were located in these promoter regions were defined as promoter peaks.

2. Intragenic (genebody) peaks: The genebody region was defined as +2000 bp downstream of the transcription start site (TSS) to the transcription termination site (TTS).
3. Intergenic peaks: Intergenic peaks were defined as the other genomic regions not included in the above two regions. Peaks whose centers were located in these intergenic regions were defined as intergenic peaks.

Please cite this article as: C. Pheiffer, et al., Differential DNA methylation of microRNAs within promoters, intergenic and intragenic regions of type 2 diabetic, pre-diabetic and..., Clin Biochem (2015), http://dx.doi.org/10.1016/j.clinbiochem.2015.11.021