Author:
Alrashidi, Fatma

Title:
Beta cell loss in autoimmune diabetes

a useful clinical biomarker of disease progression?

‘Beta cell loss in autoimmune diabetes: a useful clinical biomarker of disease progression?’

Fatimah T. AlRashidi

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of PhD in the Faculty of Health Science

School of Translational Health Sciences Submitted July 2018

Words: 78,771

Abstract

Type 1 diabetes (T1D) is a chronic autoimmune condition hallmarked by a decline in the mass of insulin-secreting beta cells and resultant hyperglycaemia when approximately 70-80% of beta cells have been destroyed. The condition can be accurately predicted by the presence of multiple islet autoantibodies, but these are not directly related to the rate of beta cell destruction and have not proved useful as a measure of therapeutic benefit in clinical trials. Methods measuring the level of cell-free (cf) DNA containing fragments of the beta cell-specific insulin gene, released from dying beta cells into the circulation have been reported as a potential method for detecting real-time beta cell death and may represent an effective strategy for monitoring beta cell destruction in individuals at risk of future T1D, those with T1D and recipients of islet transplants.

Assays targeting the insulin gene have been reported but have proven difficult to replicate. The aim of this study, therefore, was to identify additional differentially methylated loci specific for the beta cell genome which could be multiplexed to an assay targeting exon 2 of the Insulin gene previously developed in the Bristol laboratory for high throughput analysis by droplet digital PCR. Initial approaches involved targeting *PDX-1,* but ultimately a genome-wide approach was used to obtain DNA methylation signatures in islets compared with peripheral blood mononuclear cells (PBMC) using the Illumina Infinium Human Methylation EPIC BeadChip. Through bioinformatic analyses, 3866 CpG loci were differentially hypomethylated, and 4990 CpG loci were differentially hypermethylated in islets when compared to the PBMC methylome. Combining the EPIC analysis with methylation signatures from other tissues conducted by the Roadmap and ENCODE studies resulted in 425 hypomethylated and 228 hypermethylated CpG targets that are differentially methylated in islets, but not in liver, lung, thymus, pancreas, spleen and peripheral blood mononuclear cells. Subsequently, the multiplex methylation sensitive assay was developed based on a combination of beta cell-specific differentially methylated regions (DMRs) of the combined promoter for *MIR-200c/MIR-141* and exon 2 of the insulin gene. This multiplex biomarker discriminated circulating beta cell DNA fragments in the circulation of recently diagnosed paediatric T1D patients, individuals at increased risk of developing T1D and in the circulation of T1D patients post islet transplantation therapy. Multiplexing several beta cell-specific targets increased the ability of the assay to discriminate between cases and controls.

The data presented in this thesis demonstrates that DMR based biomarkers could provide highly specific real-time detection of dying beta cells in the early stages of T1D pathology and potentially measure responses to clinical therapy. The newly developed assay will be further validated in the future by testing larger cohorts of patients and “at-risk individuals”, with longitudinal sampling to determine the usefulness of the assay for measuring rates of progression to T1D.

Acknowledgement

"In the Name of Allah, the Most Gracious, the Most Merciful"

I am deeply grateful to Allah for everything in my life, under his light, all my dreams come true. I am also grateful for giving me the inspiration, patience, time, and the ability to finish this work. With his will and mercy, I have been guided to all those great people who helped me to finish this work.

I wish to express sincere gratitude for my supervisor Dr Kathleen Gillespie, for her patient, understanding and continuing support, help, and guidance throughout my research work. Also, I would further like to thank my co-supervisor Mr Alistair Williams for all the support I got. I would also like to express my thanks to all the supporting staff the Diabetes and Metabolism, Technical Support team and the TrialNet nurse Staff for their assistance in many aspects of the PhD.

This thesis would never have been accomplished if it relied only on my efforts, but it was the result of the collective efforts of several valued and extraordinary people who directly or indirectly supported and assisted me in many ways during this long and challenging journey. To all these great people, I owe my sincere gratitude and thanks.

I also wish to express my gratitude to the University of Bristol for awarding me with this degree. My endless gratitude is also due to Kuwait University for awarding this scholarship.

I would like to offer my deepest and warmest thanks to my parents, Turki and Neema, for their endless prayers, encouragement, and unconditional love, for all of that, I am dedicating this PhD for you. Special thanks and appreciations are due to my brothers Ahmed, Yousif, Homod, Abdelaziz and Tareq. Ahmed, it all started with you, and I cannot thank you enough for being there when I needed you. Homod, there are no words to express my gratitude for your endless compassion. Finally, my biggest love for my niece "baby Fatma" who enlighten me and my family life.

"Praise is to Allah by Whose grace good deeds are completed"

Author's Declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's *Regulations and Code of Practice for Research Degree Programmes* and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: ..... DATE: .....**July 2018**.....

Table of Contents

Abstract **1**

Acknowledgement **2**

Author's Declaration **3**

Table of Contents **4**

List of Figures **9**

List of Tables **13**

Publications and Presentations that Have Arisen from the PhD Studentship to Date **14**

Abbreviation **15**

Chapter 1: Introduction **19**

1. General Introduction 20

1.1. Type 1 Diabetes 20

1.2. Pancreatic Beta Cells 24

1.3. T1D Etiology 26

1.3.1. Genetic Susceptibility of T1D 27

1.3.2. Environmental Risk Factors 38

1.3.3. The Effect of Environmental Interaction with Genetic Traits in T1D Development (Twin Studies) 44

1.4. Autoimmunity in T1D 45

1.4.1. Beta Cell Death in T1D (Insulitis) 45

1.4.2. Autoimmune Destruction of Beta Cells in T1D 51

1.4.3. Mechanism of Beta Cell Death in T1D, Necrotic or Apoptotic Pathway? 51

1.4.4. Autoantibodies to Islet Antigens 52

1.5. The Progression of T1D 54

1.5.1. The Linear Beta Cell Declines Hypothesis 54

1.5.2. The Non-Linear Relapsing-Remitting Model 56

1.5.3. Fertile Field Model: 57

1.6. Epigenetics 58

1.6.1. General Introduction 58

1.6.2. DNA Methylation Modification 61

1.6.3. The DNA Methylome of the Pancreatic Beta Cell 69

1.7. Biomarkers of Predicting T1D 71

1.7.1. General Introduction 71

1.7.2. Discovering a Novel Biomarker 71

1.7.3. Prediction and Diagnostic Biomarkers for Overt Diabetes 73

1.8. DNA Methylation Profiling Technologies and Studies 84
1.8.1. DNA Methylome Profiling 84
1.8.2. DNA Methylation-Sensitive Assay Design 85
1.8.3. The Effect of Diabetes on the Methylome of Beta Cell-Specific Genes 85
Conclusion 90
Project Rationale 90
Hypothesis 90
Aims 90
Chapter 2: General Methods 92
2. General Methods 93
2.1. Samples 93
2.1.1. Human Tissue 93
2.1.2. Cell Lines 93
2.2. Cell Culture 94
2.2.1. EndoC-βH1 Cell Culture 94
2.2.2. Prostate and Breast Cell Lines Culture 95
2.3. Serum and Plasma Samples 96
2.3.1. Locally Collected Samples (Assay Optimising Experiments) 96
2.3.2. Recently Diagnosed Paediatric Cohort 96
2.3.3. Recently Diagnosed Adult Cohort 97
2.3.4. At Risk for Developing T1D Cohort 97
2.3.5. Islet Transplantation T1D Recipients 98
2.3.6. Urine Samples 98
2.4. DNA Methylation Public Data 106
2.5. Design of Oligonucleotides 106
2.6. Polymerase Chain Reaction (PCR) 108
2.6.1. Droplet Digital PCR (ddPCR) 108
2.6.2. Probe-Based ddPCR 108
2.6.3. Droplet Generating, PCR Cycle Conditions 109
Chapter 3: The Pancreas and Duodenum Homeobox Protein-1 110
3. Introduction 111
Methods 113
3.1.1. DNA Extraction from Human Cells and Tissues 113
3.1.2. DNA Quantification 113
3.1.3. Bisulfite Conversion 114
3.1.4. Gel Electrophoresis 115

3.1.5. Vector Amplification 116
3.1.6. Whole Genome Amplification 116
3.1.7. Bisulfite Sanger Sequencing 116
Results 118
3.2.1. The Pancreatic Duodenal Homeobox-1 118
3.2.2. *PDX-1* Promoter Region Sequencing in Islet and PBMC Genome 119
3.2.3. Developing a *PDX-1* Beta cell-Specific Biomarker 125
3.2.4. Defining the Threshold for Droplet Digital PCR Experiments 128
Discussion 131
Chapter 4: Developing Biomarkers for Monitoring Beta cell Death 133
4. Introduction 134
Methods 136
4.1.1. The Illumina Infinium Human Methylation EPIC 850K BeadChip Array 136
4.1.2. Targeted Next-Generation Bisulfite Sequencing (tNGBS): 142
4.1.3. Universal Human Methylated DNA Standard Controls 144
4.1.4. Droplet Generating and PCR Cycle Condition 144
4.1.5. Methylation Data Alignment and Statistical Analysis 147
Results 148
4.2.1. The Illumina Infinium® Methylation EPIC BeadChip Analysis 148
4.2.2. Developing DMR based Beta Cell-Specific Biomarkers 151
4.2.3. Developing Triplex Assay to Detect Beta Cell Death 166
4.2.4. The Effect of Adding EcoRI to Minimise Droplet Rain 170
4.2.5. Targeted Next-Generation Sequencing 172
4.2.6. The effect of Hyperglycemia and Proinflammatory Cytokines of the Methylation Stability in the Developed Multiplex Assay 176
Discussion 178
Chapter 5: Optimising Cell Free DNA Extraction Protocol 180
5. Introduction 181
Methods 183
5.1.1. Locally Collected Samples 183
5.1.2. Processing the Urine Samples (Experiment Design) 184
5.1.3. DNA Preservative 185
5.1.4. DNA Extraction 185
5.1.5. cfDNA Quantification in Biofluid Samples 185
5.1.6. Droplet Digital PCR 186
5.1.7. Statistical Analysis 186

Results 187
5.2.1. cfDNA extraction using different extraction methods 187
5.2.2. Comparing the Modified pH Column-Based Method to Other Published Protocol ... 189
5.2.3. The Optimum Plasma Volume Needed for cfDNA Extraction 191
5.2.4. The Effect of DNA Preservative on Protecting cfDNA from Degradation in Urine Samples 193
5.2.5. Effects of Adding EDTA and Storage Temperature on the Stability of cfDNA in Urine 195
5.2.6. Optimising of Urine Volume Required for Successful cfDNA Analysis 200
Discussion 202
Chapter 6: Monitoring Beta Cell Death in Clinical T1D Samples 204
6. Introduction 205
Results 206
6.1.1. *MIR-141*/*MIR-200c* Cluster 1/*INS* Exon 2 Assay Triplex as a Biomarker for Beta Cell Death 206
6.1.2. *MIR-200c* Cluster 2 Assay Triplex combination as a Biomarker for Beta Cell Death .. 210
6.1.3. Monitoring Beta Cell Death in the "At-Risk" Autoantibody Positive Cohort 213
6.1.4. Monitoring Beta Cell Death in Recently Diagnosed T1D Subjects. 216
6.1.5. Measuring the Beta Cell Death Rate in the Post Islet Allotransplant Patients 220
6.1.6. Developing a New Assay for Monitoring Beta Cell Death 222
Discussion 226
Chapter 7: Biomarkers of Prostate Cancer Progression 229
7. Introduction 230
Methods 231
7.1.1. MicroRNA Extraction 231
7.1.2. Quantification of microRNA 231
7.1.3. Reverse Transcription PCR (RT-PCR) and ddPCR Analysis of Prostate Cancer Cell Lines. 231
Results 233
7.2.1. Methylation Analysis of *MIR‑141* Promoter in Prostate and Breast Epithelial Cells 233
7.2.2. The Expression of miR-141-3p/5p and miR-200c-3p in Prostate Cell Lines 235
7.2.3. The Methylome Level of Prostate Cell Lines at the Promoter of *MIR-141* Gene 236
Discussion 238
Chapter 8: General Discussion and Future Work 239
8. General Discussion 240
8.1. Main Findings 240
8.2. General Discussion 242

8.3. Biomarkers of Beta Cell Death 245
8.4. Strengths and Limitations 247
8.5. Future Direction 249
8.6. Conclusion 250
Chapter 9: References 251
9. References 252
Appendix A 298
A. Appendix A 299
Detailed Protocols 299
Cell Free DNA Extraction from Serum/Plasma 299
Triton/Heat/Phenol (THP) Protocol for Extracting Cell Free DNA 300
Cloning Using pGEM-T Easy Vector System I and DH5α Competent Cells 301
Blood Collection and Processing using Roche Collection Tubes for cfDNA Study Experiments 303
DNA Quantification Using the QuantiFluor® dsDNA System (Promega). 304
Bisulfite Converted DNA Quantification Using the QuantiFluor® dsDNA System (Promega). 305
RNA Quantification using the QuantiFluor® RNA System (Promega). 306
Amplification of Purified Genomic DNA Using the REPLI-g Mini Kit 307
Tables 309
Table A.1: The Characterises of oligonucleotides used in this project. 309
Table A.2: The microRNA LNA PCR primer sets. 311
Table A.3: The differentially hypomethylated CpG sites in the islet genome. 312
Table A.4: The differentially hypermethylated CpG sites in the islet genome. 321
Table A.5: Summery of methylation percentage data derived from targeted next generation sequencing of *MIR-200c/MIR-141* mutual promoter as well as post MIR-141 (Chr12: 6,963,000-6,964,001). 336
Appendix B 338
Materials 339

List of Figures

Figure 1.1: The incidence T1D in children (0-14) years. 22
Figure 1.2: Clinical Stages of Type 1 Diabetes Progression. 23
Figure 1.3: The biogenesis of the insulin 25
Figure 1.4: The Etiology of T1D. 26
Figure 1.5: GWAS significant susceptible genes for T1D. 28
Figure 1.6: Gene map of the Human Leukocyte Antigen (HLA) Region, 6p21. 32
Figure 1.7: A model of CD8+ T cell-mediated destruction of human pancreatic beta cells. 49
Figure 1.8: The timeline model of type 1 diabetes progression. 55
Figure 1.9: Immune regulations imbalance in T1D. 57
Figure 1.10: The Structure of the Nucleosome. 59
Figure 1.11: The process of cytosine methylation. 63
Figure 1.12: The structure of the methyltransferase enzyme family. 63
Figure 1.13: The methylation process of DNA. 64
Figure 1.14: The methylation cycle in a somatic cell. 67
Figure 3.1: Bisulfite Conversion of DNA 115
Figure 3.2: Amplification of the PDX-1 promoter region 118
Figure 3.3: A representative example of a distal PDX-1 promoter sequencing chromatogram from one clone of bisulfite treated islet DNA. 120
Figure 3.4: A representative example of a distal PDX-1 promoter chromatogram sequencing from one clone of pancreatic PBMC bisulfite treated DNA. 120
Figure 3.5: The CpG methylation pattern of the distal PDX-1 promoter. 125
Figure 3.6: The methylation status of CpG -857 of PDX-1 gene (from TSS) in DNA from an “in house” tissue panel. 126
Figure 3.7: 2-D amplification plots from droplet digital data generated for tissue analysis of differentially methylated PDX-1 assay. 127
Figure 3.8: Defining the threshold in 1-plex an EVAGreen ddPCR experiment run on the same plate. 128
Figure 3.9: Defining the threshold in 1-plex an EVAGreen ddPCR experiment analysed on the same plate. 129
Figure 3.10: 2-D Plot of droplets of Fluorescence amplified multiplex assay. 130
Figure 4.1: Illumina Infinium Type I and Type II Probes. 137
Figure 4.2: The chromosomal distribution of the probes in the Methylation EPIC BeadChip. 138
Figure 4.3**:** Infinium HD Methylation Assay Protocol Workflow. 141

Figure 4.4: The analysed region by targeted next generation bisulfite sequencing technology. 142
Figure 4.5: The Standard PCR Protocol. 145
Figure 4.6: The Touchdown PCR Protocol. 146
Figure 4.7: Genomic and functional details of the analysed CpG sites in the Methylation EPIC 850K microarray. 149
Figure 4.8: Heatmap of the methylation status of islet and PBMC DNA related to methylomes of various tissues. 150
Figure 4.9: Optimal annealing temperatures for differentially methylated multiplexed NCOR2 assay. 153
Figure 4.10: The methylation sensitivity of the NCOR2 assay. 155
Figure 4.11: Representative amplification of 2-D plots from ddPCR data generated for tissue analysis with differentially methylated NCOR2 assay. 156
Figure 4.12: The methylation level of CpG +49678 of NCOR2 exon 2. 157
Figure 4.13: The methylation status of CpG +20459 from TSS of SLC45A4 gene in various tissues. 158
Figure 4.14: 1-D fluorescence amplification plotted against the annealing temperature gradient for the differentially methylated multiplexed MIR-141 assay (CpG -18/CpG-20). 161
Figure 4.15: 2-D amplification plots from droplet digital PCR data generated for tissue analysis with the differentially methylated MIR-141 assay (CpG -18/CpG-20). 162
Figure 4.16: The methylation status of CpG -20 (Chr12:7,073,239) and CpG -18 (Chr12:7,073,241) of MIR-141 gene. 163
Figure 4.17: The methylation status of CpG -20 (Chr12:7,073,239) and CpG -18 (Chr12:7,073,241) of MIR-141 gene. 164
Figure 4.18: 1-D fluorescence amplification plots of two hypermethylated MIR-141 assays. 165
Figure 4.19: The sensitivity of the triplex assay for methylation detection by ddPCR 167
Figure 4.20: Representative 2-D amplification plots from droplet digital PCR data generated by the tissue analysis with the triplex assay. 168
Figure 4.21: The level of hypomethylated DNA copies of the triplex assay in different tissues available “in house”. 169
Figure 4.22: The effect of adding EcoRI on a ddPCR reaction. 171
Figure 4.23: The methylation status of a pancreatic islet sample and a beta cell line model. 173
Figure 4.24: The effect of high glucose exposure on the methylation profile of the EndoC-βH1 beta cell line model. 175
Figure 4.25: The effect of high glucose exposure on the methylation profile of the designed beta specific targets. 176

Figure 4.26: The effect of pro-inflammatory cytokines exposure on the methylation stability of the designed beta specific targets. 177
Figure 5.1: The experiment design of storage conditions and duration of Group A and Group B urine samples processing. 184
Figure 5.2: Comparison of beta-globin cell-free DNA (cfDNA) recovery from serum using three DNA isolation methods. 188
Figure 5.3: Comparison between two cell-free DNA (cfDNA) isolation protocols. 190
Figure 5.4: Comparison of the beta-globin yield extracted from different plasma volumes. 192
Figure 5.5: The PCR products of bisulfite converted cfDNA extracted from 4 mL control urine samples. 193
Figure 5.6: The absolute concentration of differentially methylated INS (CpG +367+374) in different volumes of a urine sample. 194
Figure 5.7: The effect of adding 40 mM EDTA on DNA stability in urine samples when stored at different temperatures and durations. 196
Figure 5.8: The effect of long-term storage on untreated and EDTA treated urine samples. 198
Figure 5.9: Comparison of the effect of two preservatives on the stability of cfDNA in urine samples with and without DNA preservative. 199
Figure 5.10: Comparison of beta-globin yield extracted from different volumes of urine from healthy volunteers. 201
Figure 6.1: Circulating hypomethylated MIR-200c cluster 1/ INS / MIR-141 cfDNA levels in recently diagnosed paediatric human cohort. 209
Figure 6.2: Circulating hypomethylated MIR-200c cluster 1/INS/MIR-200c cluster 2 cfDNA levels in recently diagnosed T1D pediatric human cohort. 212
Figure 6.3: Singleplex and multiplex analysis of the beta cell death biomarkers MIR-200c cluster 1/INS/MIR-200c cluster 2 in individuals at-risk of developing T1D. 215
Figure 6.4: Circulating hypomethylated MIR-200c cluster 1/INS/MIR-200c cluster 2 cfDNA levels in a recently diagnosed T1D adult cohort. 219
Figure 6.5: Change in hypomethylated MIR-200c cluster 1/INS/MIR-200c cluster 2 cfDNA pre and post-islet-infusion. 221
Figure 6.6: Singleplex and duplex analysis of the beta cell death hypomethylated MIR- 141 biomarker in individuals at-risk for T1D. 225
Figure 7.1: The EvaGreen ddPCR Protocol. 232
Figure 7.2: The methylation level of CpG -20 of MIR‐141 in Hormone sensitive and Hormone insensitive prostate and breast cell lines. 234

Figure 7.3: The miR-141-3p/5p and miR-200c-3 expression level in prostate cancer cell lines. 235
Figure 7.4: Targeted methylation analysis of the MIR‐200c/MIR‐141 promoter in prostate cancer cell lines 237
Figure 8.1: The DNA methylation of human glucagon and insulin genes in islet cells. 244
Figure 8.2: Expression levels of circulating plasma miR-375 post islet transplantation. 246
Figure A.1: The modified pH-column based method protocol 299
Figure A.2 The modified pH-column based method protocol 300

List of Tables

Table 1.1: Top genes variants reported to contribute to T1D susceptibility. 29
Table 1.2: The classical and non-classical loci of the Class I and Class II HLA region. 33
Table 1.3: Biomarker types based on their clinical applications 72
Table 1.4: Features and characterization of the ideal biomarker 73
Table 1.5: The sensitivity and the specificity of the primary antibodies and targets used as a predicting biomarker of T1D. 76
Table 1.6: Different DNA methylation-profiling technologies. 87
Table 2.1: Islet Cells Characteristics. 99
Table 2.2: Human Prostate Cell Lines Characterizes 100
Table 2.3: Human Breast Cell Lines Characterizes 101
Table 2.4: Recently Diagnosed T1D Paediatric Samples (UCL Cohort) 102
Table 2.5: The Characteristics of Paediatric Non-Diabetic Control Samples 102
Table 2.6: The Characteristics of Recently Diagnosed Adult Plasma Samples (Bristol Cohort) 103
Table 2.7: The Characteristics of Recently Diagnosed Adult Plasma Samples (StartRight Cohort) 103
Table 2.8: The non-Diabetic Healthy Samples Characteristics. 104
Table 2.9: The At-Risk for Developing T1D Samples Characteristics. 105
Table 2.10: The Non-Diabetic Urine Samples Characteristics 106
Table 3.1: The methylation status of the distal PDX-1 promoter in PBMC genome. 122
Table 3.2: The methylation status of the distal PDX-1 promoter in the islet genome. 123
Table 3.3: The methylation status of the distal PDX-1 promoter in the beta cell genome. 124
Table 4.1: The target region of the targeted bisulfite-sequencing assay. 143
Table 6.1: The effect of multiplexing in enhancing the discrimination of a beta cell death signal. 207
Table 6.2: The effect of multiplexing on enhancing the discriminating power of beta cell death signal. 210
Table 6.3: The genomic DNA yield extracted from the at-risk samples and non-diabetic control subjects. 223
Table A.1: The Characterises of oligonucleotides used in this project. 309
Table A.2: The microRNA LNA PCR primer sets. 311
Table A.3: The differentially hypomethylated CpG sites in the islet genome. 312
Table A.4: The differentially hypermethylated CpG sites in the islet genome. 321
Table A.5: Summery of methylation percentage data derived from targeted next generation sequencing of *MIR-200c/MIR-141* mutual promoter as well as post MIR-141 TSS (Chr12: 6,963,000-6,964,001). 336

Publications and Presentations that Have Arisen from the PhD Studentship to Date

Invited article

AlRashidi, Fatimah T., and Gillespie, Kathleen M. *J Current Diabetes Reports*. 2018. "Biomarkers in Islet Cell Transplantation for Type 1 Diabetes." 18 (10):94. doi: 10.1007/s11892-018-1059-4.

Abstracts

Jody Ye, **Fatma Alrashidi**, Matthew Suderman, Michelle Curran, Parth Narendran, Chloe Bulwer, Rakesh Amin, Caroline Relton, Colin Dayan, Alistair Williams, Kathleen Gillespie. Development of a multiplex methylation sensitive digital PCR assay to monitor islet autoimmunity in type 1 diabetes. (Talk Presentation given by Dr Jody Ye). *The 15th International Congress of the Immunology of Diabetes Society*. January 19-23 (2017). San Francisco, USA

AlRashidi F., Ye J., Suderman M., Relton C. and Gillespie KM., Development of multiplex sensitive strategy to detect beta cell death in type 1 diabetes. (Poster presentation) *The Diabetes UK professional Congress*, March 14-16, 2018, London, UK

AlRashidi F., Ye J., Tatovic D., Dayan CM., Williams A. and Gillespie KM. *A* methylation-sensitive multiplex droplet digital PCR assay to quantify beta-cell destruction in type 1 diabetes. (Poster presentation) *The 16th International Congress of the Immunology of Diabetes Society*. October 25-29 (2018). London, UK

Manuscript in preparation

Fatma Al Rashidi*, Jody Ye*, Mathew Suderman, Claire Caygill, Chloe Bulwer, Rakesh Amin, Michelle Curran, Parth Narendran, Colin M Dayan, Aniko Viradi, Rob Andrews, Angus Jones, Caroline Relton, Alistair J K Williams, and Kathleen M Gillespie. (2018). Methylation sensitive multiplex droplet digital PCR to quantify beta cell destruction in type 1 diabetes. **Target Journal** *The Journal of Clinical Investigation (JCI)*

Abbreviation

5caC	5-carboxylCytosine
5fC	5-formylCytosine
5hmC	5-hydroxymethylCytosine
5mC	5-methylCytosines
AICD	Activation-Induced Cell Death
APC	Antigen Presenting Cell
BCR	B Cell Receptors
CD	Cluster of Differentiation
cfDNA	Cell Free DNA
CpG	5'—C—phosphate—G—3'
CTLA4	Cytotoxic T-Lymphocyte-Associated Protein 4
CVB4	Enteroviruses, Coxsackie B4
DAISY	The Diabetes and Autoimmunity Study in the Young
ddPCR	Droplet Digital PCR
DISS	Diabetes Incidence Study in Sweden
DiViD	The Diabetes Virus Detection Study
DMR	Differentially Methylated Regions
Dnmt	DNA Methyltransferase
DZ	Dizygotic
EDTA	Ethylene-Diamine-Tetraacetic Acid
ENDIA	The Environmental Determinants of Islet Autoimmunity
ERK	Extracellular signal-Regulated Kinase
ESC	Embryonic Stem Cell
GAD65A	Glutamic Acid Decarboxylase 65 Autoantibodies
GWAS	Genome-Wide Association Studies
HLA	Human Leukocyte Antigen
IA-2A	Insulinoma-Associated Antigen 2 Autoantibodies
IAA	Insulin Autoantibodies
ICA	Islet cell Cytoplasmic Antibody

ICRs	Imprinted Control Regions
IDF	International Diabetes Federation
IFIH1	The interferon induced with helicase C domain 1 gene
IGRP	Islet-specific Glucose-6-phosphatase catalytic subunit-Related Protein
IHEC	International Human Epigenome Consortium
IL-1R AcP	IL-1R Accessory Protein
IL2RA	Interleukin 2 Receptor Alpha
INFγ	Interferon γ
INS	Human Insulin Gene
IRAKs	IL-1 Receptor-Associated Kinases
JDFR	The Juvenile Diabetes Research Foundation
JNK	c-Jun N-terminal Kinase
LADA	Latent autoimmune diabetes in adults
LINEs	Long Interspersed Nuclear Elements
LTRs	Long Terminal Repeats
LYP	Lymphoid protein tyrosine phosphatase
MAPK	Mitogen-activated protein kinase
MDA	Multiple Displacement Amplification
MDA5	Melanoma differentiation associated protein 5
MIR-141	microRNA-141
MIR-200c	microRNA-200c
MyD88	Myeloid differentiation primary response gene 88
MZ	Monozygotic
NCOR2	Nuclear Receptor CoRepressor 2
NF-κB	Nuclear Factor κB
NGMS	Next-Generation Methylation Sequencing
NK	Natural Killer
NOD	Non-obese Diabetic
nPOD	Network for Pancreatic Organ Donors with Diabetes
OGTT	Oral Glucose Tolerance Testing
PBMC	Peripheral Blood Mononuclear Cell

PCR	Polymerase Chain Reaction
PDX-1	The Pancreatic Duodenal Homeobox-1
PKC	Protein Kinase C
PTPN22	Protein Tyrosine Phosphatase, Non-receptor type 22
rpm	Revolutions per minute
SCID	Severe Combined Immunodeficiency
SINEs	Short Interspersed Nuclear Elements
SLC45A4	Solute carrier family 45 member 4
SLMV	Synaptic-Like MicroVesicles
SRA	SET and RING-Associated
SWENDIC	The Southwest Newly Diagnosed Diabetes Collection
T1D	Type 1 Diabetes
TCR	T Cell Receptors
TE	Transposable Element
TEDDY	The Environmental Determinants of Diabetes in the Young
TET	Ten-Eleven Translocation Protein
Tfh	Follicular T helper
Th	T helper cell
THP	Triton/Heat/Phenol
TNF	Tumor Necrosis Factor
tNGBS	Targeted Next-Generation Bisulfite Sequencing
Tollip	Toll-interacting protein
TPIAT	Total Pancreatectomy with Islet Auto-transplantation
Treg	T regulatory cells
TSPAN7	TetraSpanin Protein Family member 7
TSS	Transcriptional Start Site
UBASH3A	Ubiquitin-Associated and SH3 domain-containing protein A
UHRF1	Ubiquitin-like containing PHD and RING finger domains 1
VDR	Vitamin D receptor
VNTR	Variable Number of Tandem Repeat
vp1	Enteroviral Capsid Protein vp1

WGA	Whole Genome Amplification
WHO	World Health Organization
ZnT8A	Zinc transporter 8 autoantibodies

Chapter 1: Introduction

1. General Introduction

1.1. Type 1 Diabetes

Type 1 Diabetes (T1D) is a chronic autoimmune disease that results from specific immune-mediated destruction of the insulin-producing beta cells in the Langerhans islets of the pancreas. Uncontrolled hyperglycemia characterises this autoimmune disorder, due to insufficient insulin production. Although not the most common form of diabetes (accounting for 7-12% of all diabetes), T1D is the most common type of diabetes among children and adolescents under the age of 20 (IDF 2017). According to the last report of the International Diabetes Federation (IDF) in 2017, the incidence rate of T1D is increasing by 3% annually; approximately 132,600 new cases are registered each year. There are about 1,106,200 child and adolescents under the age of 20 years living with this condition globally, half of whom live in Europe (28.4%) and North American/Caribbean regions (both account for 21.5%). The report has also informed that Finland, Kuwait and Sweden were the countries with the highest incidence of T1D in children under the age of 15 (Figure 1.1) (IDF 2017). Moreover, data collected from two major international T1D registries (DIAMOND[1] & EURODIAB[2]) showed that there is a wide variation in T1D incidence between different geographic populations. They reported that the incidence rate was highest in Finland and Sardinia and was lowest in other populations such as China and Venezuela (Soltesz et al. 2007). Predictions based on EURODIAB registered children suggested that the incidence of T1D would double in European children under the age of 5 years between the years 2005-2020, and T1D prevalence will increase up to 70% in children under age 15 years (Patterson et al. 2007). Additionally, an analysis of time trends by the Nationwide Diabetes

1 Multinational Project for Childhood Diabetes (Diabetes Mondiale) DIAMOND: The DIAMOND project was initiated by the World Health Organization in 1990 to address the public health implications of T1D with a main objective to describe the incidence of T1D in children.

2 EURODIAB collaborative group established in 1988 prospective geographically-defined registers of new cases diagnosed under 15 years of age, the aim of this study group is to study the epidemiology of childhood-onset type 1 insulin-dependent diabetes in Europe.

Incidence Study in Sweden (DISS)[3] showed that the incidence of T1D onset in individuals aged 0-34 years from 1983 to 2007 shifted to younger ages (<14 years) (Dahlquist et al. 2011).

T1D is often diagnosed when vague symptoms appear such as fatigue and lack of energy, night sweating, sudden weight loss, frequent thirst and urination, blurred vision and constant hunger (IDF 2017). These symptoms start to appear when almost 70-80% of the individual's beta cells are already destroyed (Cnop et al. 2005, Notkins and Lernmark 2001). At this point, the only option is lifelong insulin replacement and blood glucose monitoring. Insulin replacement treatment does not work equally for all recipients. Exogenous insulin is not as exquisite as the endogenous insulin in controlling blood sugar levels, leaving the patient at risk of life-threatening hypoglycemic episodes, kidney failure, stroke and death. The time window between the onset of seroconversion, as indicated by the appearance of islet autoantibodies, and the appearance of the first clinical symptom of T1D can be rapid; for instance, in the case of children diagnosed at six months of age but may take decades (Bingley et al. 2006, Ziegler et al. 2012). Meanwhile, persistent destruction of beta cells occurs silently and undetected until hyperglycaemia is diagnosed. The progress to overt T1D is described as occurred through three stages (Figure 1.2): the first stage (the seroconversion stage) is characterised by the appearance of the islet autoantibodies in the circulation with the absence of any other clinical symptoms. The second stage is also asymptomatic, but dysglycaemia can be detected, and the third and final stage is when the clinical symptoms described above appear (Insel et al. 2015). The diagnosis of T1D is confirmed by measurement of beta cell function with tests identifying the glycated haemoglobin A1C (HbA1C), blood C- peptide levels, fasting plasma glucose, oral glucose tolerance testing (OGTT) and proinsulin level (Figure 1.2) (Lebastchi and Herold 2012). According to the World Health Organization (WHO), diabetes is diagnosed when the level of fasting plasma glucose exceeds 7 mmol/L (126 mg/dL), the level of plasma glucose is ≥ 11mmol/L (200 mg/dL) two hours following a

3 DISS is a nation-wide study based on the six major health regions of Sweden that follows the incidence of diabetes and its complications among young adults (15-34 years of age) in Sweden.

75g oral glucose intake, the level of random glucose exceeds 11 mmol/L (200 mg/dL) or the HbA1c is ≥ 48 mmol/mol (equivalent to 6.5%) (WHO 2006).

Figure 1.1: The incidence T1D in children (0-14) years.

The estimated number of new cases of T1D in children (0-14) years per 100,000 per year. The map is conducted from the International Diabetes Federation 8th Edition. http://www.diabetesatlas.org/across-the-globe.html

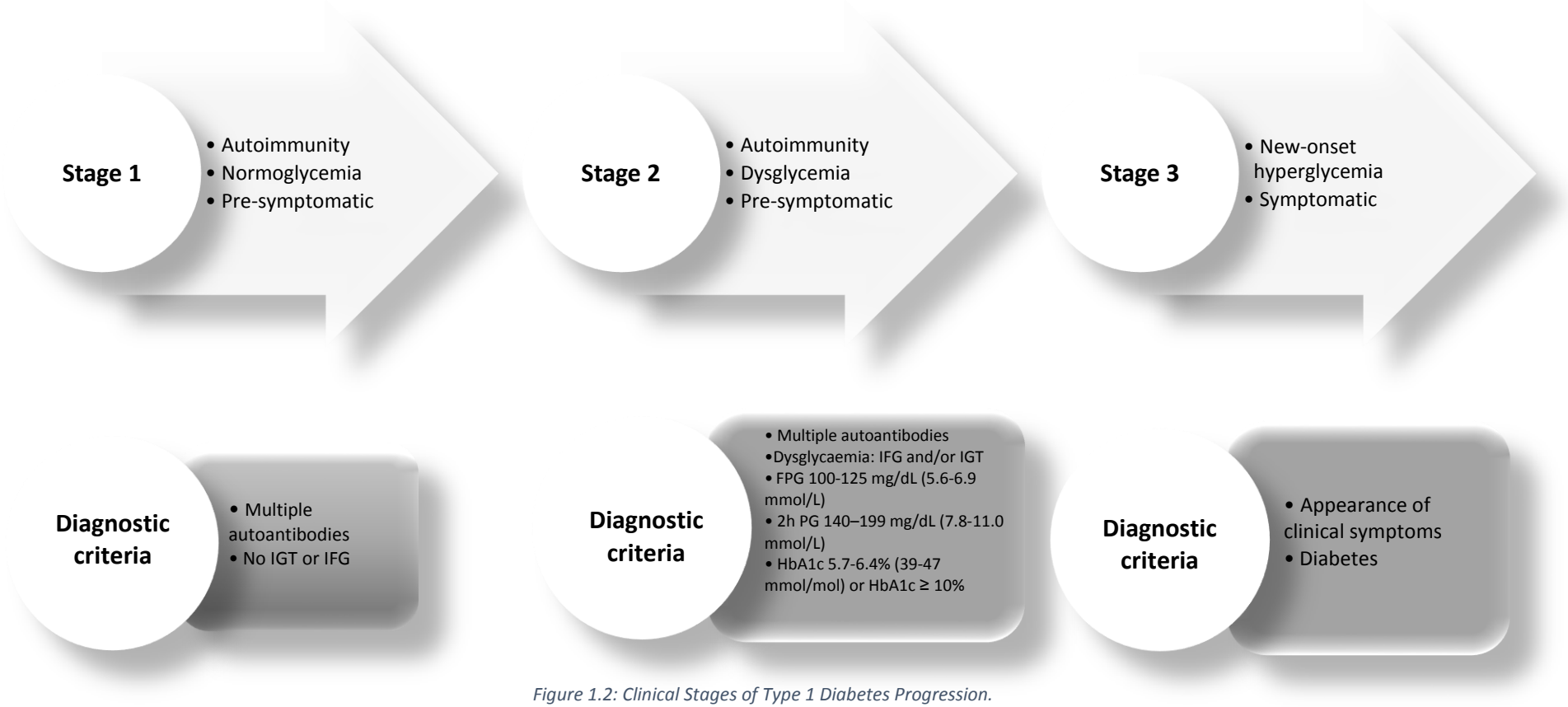


Figure 1.2: Clinical Stages of Type 1 Diabetes Progression.

Clinically, type 1 diabetes can be categorised into three stages: stage 1 is asymptomatic and is characterized by the appearance of islet autoantibodies with normal blood glucose levels; stage 2 is also asymptomatic and is characterized by the presence of autoantibodies combined with hyperglycaemia, Stage 3 is when the clinical diabetes symptoms appear.

PG: Post-load Glucose, IFG: Impaired Fast Glucose, IGT: Impaired Glucose Tolerance, FPG: Fasting Plasma Glucose

1.2. Pancreatic Beta Cells

The beta cell is a hormone-secreting cell that forms a significant component (about 60-80 %) of the islets of Langerhans in the pancreas. The pancreas is an organ of mixed cell types including exocrine, endocrine and endothelial cells as well as neurons. In addition to the beta cells, the islets consist of at least another four types of endocrine hormone-producing cells: 30% alpha-cells (secreting glucagon), <10% delta-cells (secreting somatostatin), <5% gamma-cells (secreting pancreatic polypeptide) and epsilon cells (ghrelin secreting) (Cabrera et al. 2006, Orci et al. 1976). The islets of Langerhans represent the endocrine section of the pancreas and account for only 1-2 % of the total mass of the human pancreas (Ionescu-Tirgoviste et al. 2015). An anatomical and three-dimensional (3D) virtualisation study showed that the endocrine pancreas is uniquely heterogeneous with respect to islet number and size. Most of the islets are ellipsoid or spherical. However, some, under the effect of the pressure of the adjacent structure, are irregular in shape (Ionescu-Tirgoviste et al. 2015). The total number of islets were approximately between 3.6-14.8 million, and their size varied from 0.5 to 1.3 cm^3 (Hellman 1959a, b, Ionescu-Tirgoviste et al. 2015, Saito et al. 1978). Although islets are more densely localised in the head of the pancreas, the body and the tail have a higher number of islets (Ionescu-Tirgoviste et al. 2015).

The beta cell is highly specialised with the ability to sense alterations in the level of circulating blood glucose and in response, it synthesises, stores and secretes insulin. The mature insulin is a 51-amino acid molecule that is processed from its precursors, the pre-proinsulin (in the endoplasmic reticulum) and the proinsulin (in the Golgi apparatus) in a multistep process. The biogenesis of insulin starts when an inactive precursor single chain pre-proinsulin is translated from insulin mRNA. The pre-proinsulin is converted to proinsulin when pre-proinsulin enters the lumen of the rough endoplasmic reticulum, and its N-terminal signal peptide is removed via signal peptidases. Next, proinsulin is transported to the Golgi apparatus where it is arranged into an immature secretory granule. Finally, the mature insulin protein is generated by cleaving proinsulin into insulin and C-peptide by the effect of the carboxypeptidase E and the prohormone convertases PC1/2 enzymes (Figure 1.3). Insulin, C-

peptide along with islet amyloid polypeptide and other less abundant beta cell secretory products are stored in the secretory granules until they are needed (Halban 1991, Steiner et al. 1986, Steiner and Oyer 1967).

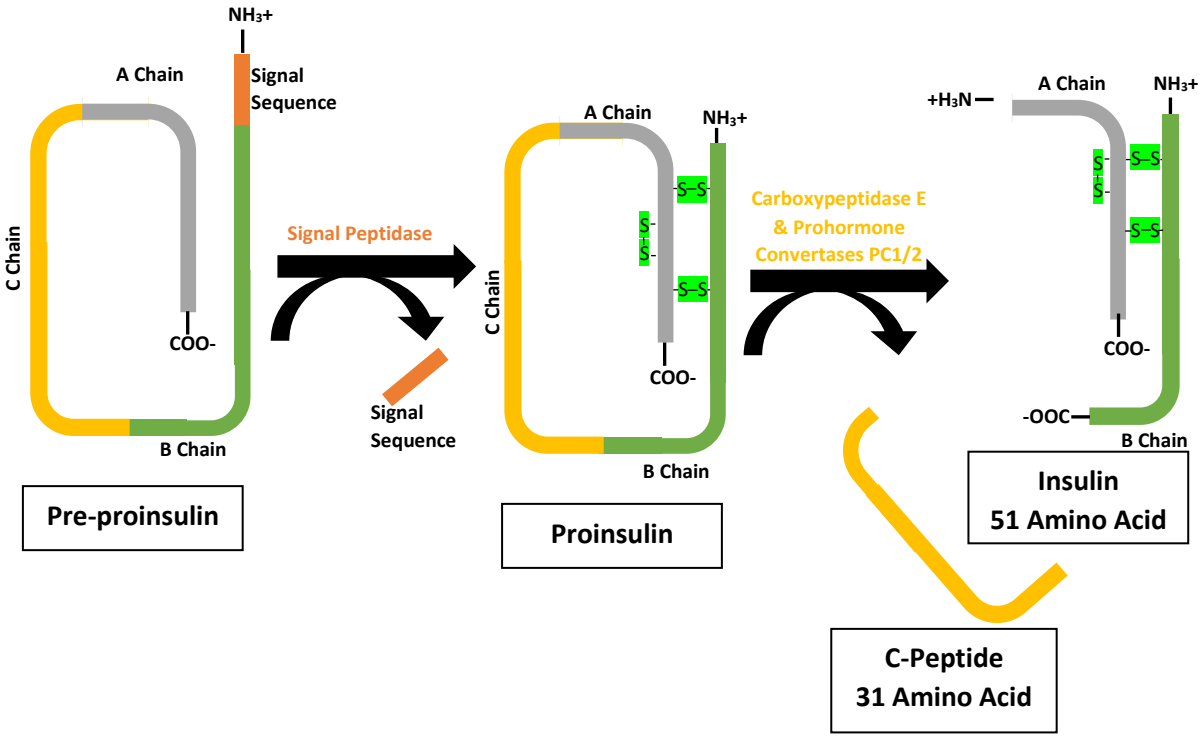


Figure 1.3: The biogenesis of the insulin

1.3. T1D Etiology

The etiology of T1D is complex; caused by a combination of the environmental trigger(s) and genomic susceptibility that leads to activation of autoimmune responses selectively against the pancreatic beta cells (van Belle et al. 2011) (Figure 1.4). Multiple genes have been shown to contribute to T1D susceptibility (Pociot et al. 2010, Pociot and McDermott 2002). Environmental factors have long been suggested to influence T1D progression as well as potentiate beta cell destruction (Akerblom et al. 2002, Knip et al. 2005, Lernmark 2016, Peng and Hagopian 2006). Therefore, the next section will briefly discuss the genetic susceptibility factors and some of the environmental triggers that are widely believed to accelerate the appearance of T1D pathology.

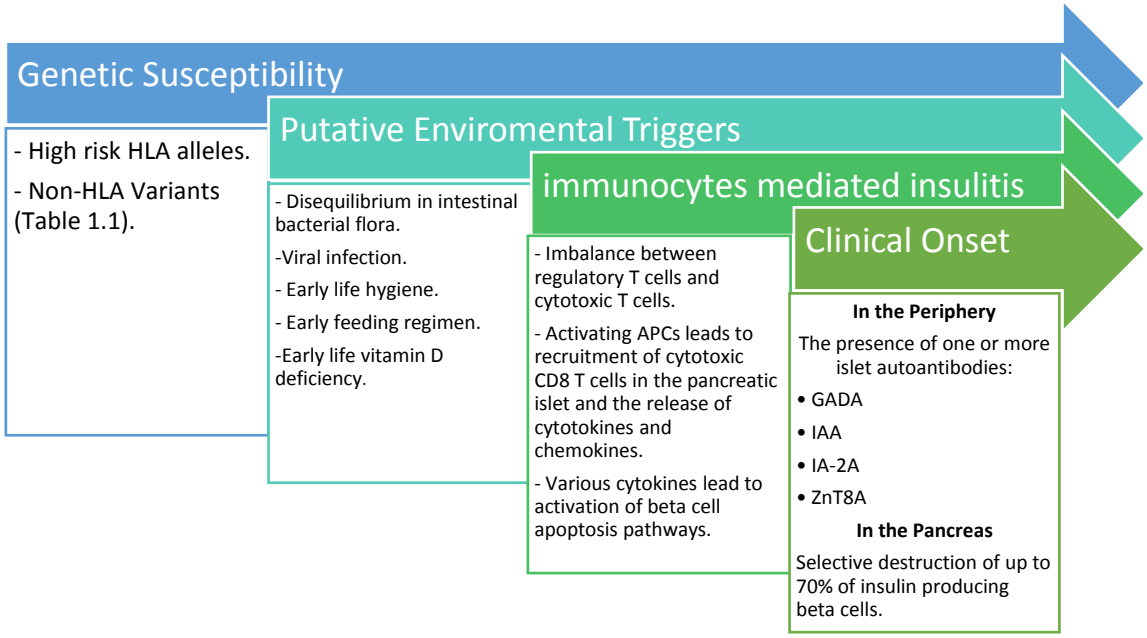


Figure 1.4: The Etiology of T1D.

When an individual with one or more genetic susceptibility variants is exposed to one or more environmental risk factors, the activation of chronic islet autoimmunity occurs. There is a preclinical gap between the initiation of autoimmunity toward beta cells and the appearance of the first symptoms of T1D which can vary from months to decades. The autoreactive response results in depletion of beta cell mass and eventually the appearance of clinical symptoms due to lack of insulin production.

All information was based on (Van Belle et al. 2011).

1.3.1. Genetic Susceptibility of T1D

Much evidence underlies the robust association between genetic elements and T1D susceptibility. This was shown in genetically identical monozygotic (MZ) siblings who have a higher risk rate for developing T1D (Redondo et al. 1999, Redondo et al. 2001) and in families of individuals with T1D where there is a high sibling concordance rate (Clayton 2009). Genome-wide association studies (GWAS) have linked more than 50 genetic variants to T1D (Barrett et al. 2009, Barrett et al. 2011). Among studied loci, the human leukocyte antigen (*HLA*) chromosomal region is confirmed as the most important genetic susceptibility factor for T1D (Table 1.1). It has been well acknowledged that about 50% of the T1D genetic susceptible factors are accounted for the highly polymorphic *HLA* class II regions (mainly *HLA-DRB1*, *-DQA1* and *-DQB1* genes) (Cudworth and Woodrow 1974, Nerup et al. 1974, Noble et al. 1996, Pociot et al. 2010, Singal and Blajchman 1973). Other significant non-*HLA* susceptibility gene loci include those for insulin (*INS*) (Barratt et al. 2004, Bell et al. 1984, Bennett et al. 1995, Vafiadis et al. 1997), cytotoxic T-lymphocyte-associated protein 4 (*CTLA4*) (Anjos et al. 2004, Kristiansen et al. 2000, Nistico et al. 1996, Ueda et al. 2003), protein tyrosine phosphatase, non-receptor type 22 (*PTPN22*) (Bottini et al. 2004, Smyth et al. 2004), interleukin 2 receptor alpha (*IL2RA*) (Lowe et al. 2007, Qu et al. 2007, Vella et al. 2005) and the ubiquitin-associated and SH3 domain-containing protein A (*UBASH3A*) (Figure 1.5) (Concannon et al. 2008, Ge et al. 2017).

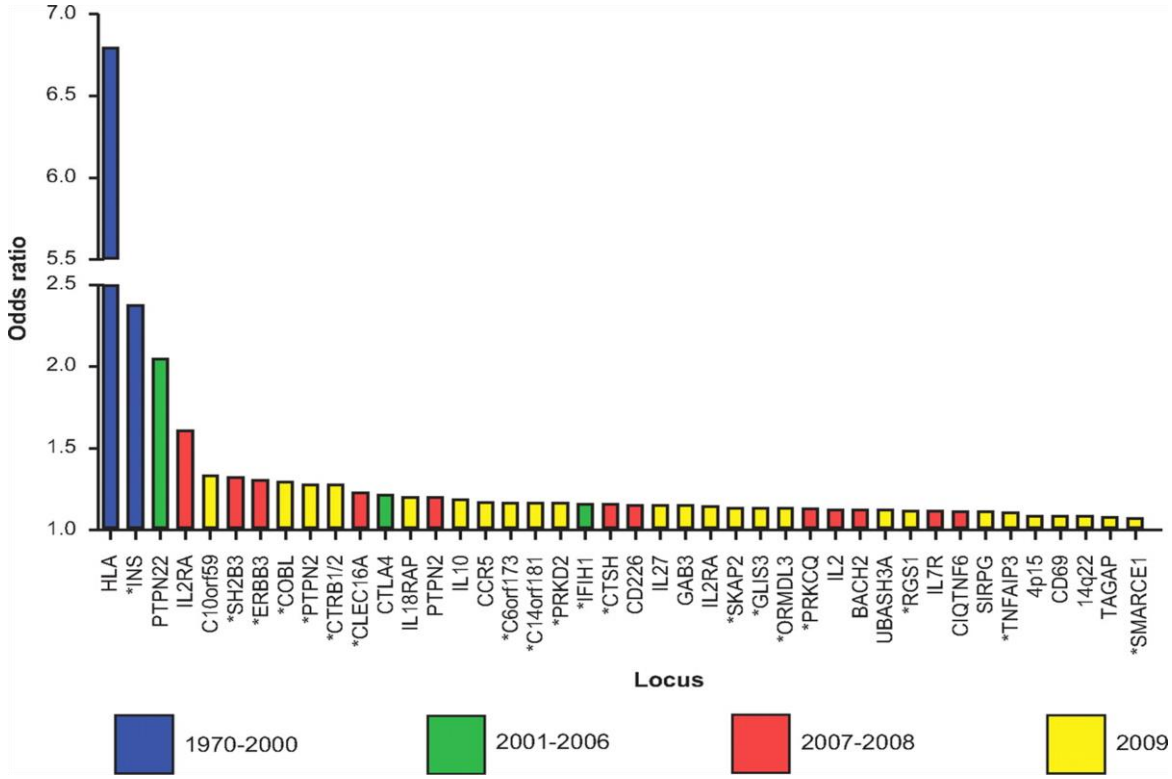


Figure 1.5: GWAS significant susceptible genes for T1D.
The figure was adapted from (Pociot et al. 2010)

Table 1.1: Top genes variants reported to contribute to T1D susceptibility.

Name	candidate Gene	Marker	Position	Odds Ratio	Full Name	References
6p21.3	*HLA-DQB1*	HLA	Chr6: 32659464..32666689	6.8	Human Leukocyte Antigen	GWAS database, (Cudworth and Woodrow 1975, 1976, Nerup et al. 1979, Nerup et al. 1974)
11p15.5	*INS*	rs7928968 A>T	chr11:2047640..2072504	1.25	Insulin Gene	(Bell et al. 1984, Bradfield et al. 2011)
1p13.2	*PTPN22*	rs2476601 C>T	chr1:113839149..114551845	2.05	Protein Tyrosine Phosphatase, Non-Receptor Type 22 (Lymphoid)	(Smyth et al. 2008)
		rs2476601 G>A	chr1:113839149..114551845			(Barrett et al. 2009)
		rs2476601 G>A	chr1:113839149..114551845	1.96		(Bradfield et al. 2011)
4p15.2		rs10517086 G>A	chr4:26027797..26132152	1.09		(Barrett et al. 2009)
4q27	*IL2*	rs4505848 A>G	chr4:122938634..123565302		Interleukin 2	(Barrett et al. 2009)
6q15	*BACH2*	rs11755527 C>G	chr6:90806835-91030155	1.13	BTB and CNC Homology 1, Basic Leucine Zipper Transcription Factor 2	(Barrett et al. 2009, Cooper et al. 2008, Smyth et al. 2008)
6q22.32		rs9388489 A>G	chr6:126457495..127419697	1.17		(Barrett et al. 2009)
9p24.2	*GLIS3*	rs7020673 G>C	chr9:4238265..4315101	1.14	GLIS Family Zinc Finger 3	(Barrett et al. 2009)
		rs10758593 G>A	chr9:4238265..4315442	1.15		(Bradfield et al. 2011)
12p13.31	*CD69*	rs4763879 G>A	chr12:9583216..9972763	1.09	CD69 Molecule	(Barrett et al. 2009)
12q13.2	*ERBB3*	rs705704 G>A	chr12:56351937..56791230	1.35	v-erb-b2 Avian Erythroblastic Leukemia Viral Oncogene Homolog 3	(Bradfield et al. 2011)
12q24.12	*SH2B3*	rs3184504 G>A	chr12:111716376..113030487	1.28	SH2B Adaptor Protein 3	(Smyth et al. 2008)
13q32.3	*GPR183*	rs9585056 T>C	chr13:99892888..100186578	1.15	G Protein-Coupled Receptor 183	(Heinig et al. 2010)
14q32.2		rs4900384 A>G	chr14:98363512..98604701	1.09		(Barrett et al. 2009)
16p11.2	*IL-27*	rs9924471 G>A	chr16:28295306..29025978	1.24	Interleukin 27	(Barrett et al. 2009)
16q23.1		rs7202877 T>G	chr16:75216240..75521030	1.28		(Barrett et al. 2009)
		rs8056814 G>A		1.25		(Bradfield et al. 2011)
18p11.21	*PTPN2*	rs1893217 A>G	chr18:12738413..12925253	1.20	Protein Tyrosine Phosphatase, Non-Receptor Type 2	(Bradfield et al. 2011)
18q22.2	*CD226*	rs763361 G>A	chr18:67498673..67570160	1.16	CD226 Molecule	(Smyth et al. 2008)
21q22.3	*UBASH3A*	rs11203203 G>A	chr21:43809418..43878660	1.14	Ubiquitin Associated and SH3 Domain Containing A	(Bradfield et al. 2011)
22q12.2		rs5753037 C>T	chr22:30073905..30669187	1.10		(Barrett et al. 2009)
22q12.3		rs229541 C>T	chr22:37568670..37660255	1.11		(Cooper et al. 2008, Smyth et al. 2008)
Xq28		rs2664170 A>G	chrX: 153654122..154285851	1.16		(Barrett et al. 2009)

1.3.1.1. The Human Leukocyte Antigen (*HLA*)

The *HLA* region represents a cluster of genes that are located on chromosome 6p21 (previously known as the IDDM1 locus). The composition of the HLA system is complex, consisting of large numbers of highly polymorphic genes, inherited as co-dominant haplotypes. The *HLA* gene family is arranged in 3 subgroups, class I, class II and class III. The class II loci are located at the centromeric end of the q arm, while class I is at the telomeric end, and both are flanking the Class III loci (Figure 1.6). Among the three classes of *HLA*, class II genes showed the most significant association with T1D as their alleles were found to be ranging from protective to strongly associated (Howson et al. 2009, Lambert et al. 2004). The *HLA* is stated to account for about 40-50% of the familial aggregation of T1D (Lambert et al. 2004). The *HLA* complex was first linked to T1D when serological typing showed an association between several *HLA* class I antigens such as *HLA-B8*, *HLA-B18* and *HLA-B15* in affected sibling pairs (Cudworth and Woodrow 1975, 1976, Nerup et al. 1979). Later, with the development of advanced analysis methods, a stronger association was discovered between the *HLA* class II genes (DQ, DR and DP) and T1D (Noble and Valdes 2011). Nonetheless, more loci within or near the *HLA* complex were found to be associated with the risk of T1D which add more complexity for analysing the contribution of *HLA* genes to the T1D genetic susceptibility for T1D.

1.3.1.1.1. *HLA* Class I

The *HLA* class I region is 1.8 Mb with 128 loci: 42 protein genes, 12 gene candidates, 10 non-coding genes and 64 pseudogenes, which all clustered in three separate duplication blocks: α, β and κ blocks (Shiina et al. 2009). The *HLA* molecule is a heterodimeric glycoprotein consisting of *HLA*-encoded 44 KD heavy α chains of and *HLA*-independent 12 KD light β chain (β2-microglobulin) (encoded in chromosome 15) (Mehra and Kaur 2016, Strominger 1987). Within the coding genes, six genes are sub-grouped into classical *HLA* genes (A, B and C) and three non-classical genes (E, F and G). The classical *HLA* class I antigens are expressed on most nucleated cells and platelets while non-classical *HL*A class I antigen expression is limited and appeared to have defined functions in transplantation (Shiina et al. 2009). The most significantly T1D associated alleles are *HLA-B*39:06* as

a predisposing allele and *HLA-B*57:01* as a protective allele (Noble et al. 2010). Other reported T1D susceptible haplotypes include *HLA-A*24:02*, *HLA-A*02:01*, *HLA-B*18:01* and protective alleles *HLA-C*05:01*, *HLA-A*11:01*, *HLA-A*32:01*, *HLA-A*66:01*, *HLA-B*07:02*, *HLA-B*44:03*, *HLA-B*35:02*, *HLA-C*16:01*, and *HLA-C*04:01* (Noble et al. 2010). *HLA-A*02* is also reported to be directly associated with T cell autoreactivity to insulin (Nejentsev et al. 2007, Pinkse et al. 2005, Valdes et al. 2005).

1.3.1.1.2. *HLA* class II

The class II *HLA* region is 700 Kb with 27 defined genes region, 17 protein-coding genes, 7 pseudogenes and 3 non-coding genes. The *HLA* class II region D, consists of three classical loci R, Q and P, and three non-classical loci M, N and O. Being classical or non-classical is based on several criteria such as the degree of gene polymorphism (high polymorphisms in classical genes) and the expression on tissues (Table 1.2). Class II antigens are expressed on some immunocytes such as B-cells, macrophages, dendritic cells and activated T cells (Cresswell 1994, Holling et al. 2002). B cells and dendritic cells have a role in activating $CD4^+$ T cells by presenting HLA class II antigens thereby initiating the immune response (Kambayashi and Laufer 2014). T1D is strongly linked to several *HLA class II* genes, the haplotypes DR-DQ: *DRB1*03:01-DQA1*05:01-DQB1*02:01* (*DR3*) and *DRB1*04:01/02/04/05/08-DQA1*03:01-DQB1*03:02/04* (*DR4*) were shown to confer the highest risk for T1D (Noble and Valdes 2011) while *DRB1*15:01-DQA1*01:02-DQB1*06:02* (*DR2*) protects from the disease (Graham et al. 1999, Noble and Valdes 2011, Pugliese et al. 1995). *HLA* class II genetic susceptibility differs from one ethnic group to another; for example, in Caucasians, the *DRB1*03-DQB1*0201 and/or *DRB1*04-DQB1*0302* haplotypes are predominant in T1D individuals, but they are rare in Japanese populations (Ikegami et al. 2006). The *DRB1*0405-DQB1*0401* and *DRB1*0901-DQB1*0303 HLA* class II haplotypes are more frequently associated with T1D in Japanese populations (Kawa et al. 2002, Kawabata et al. 2002).

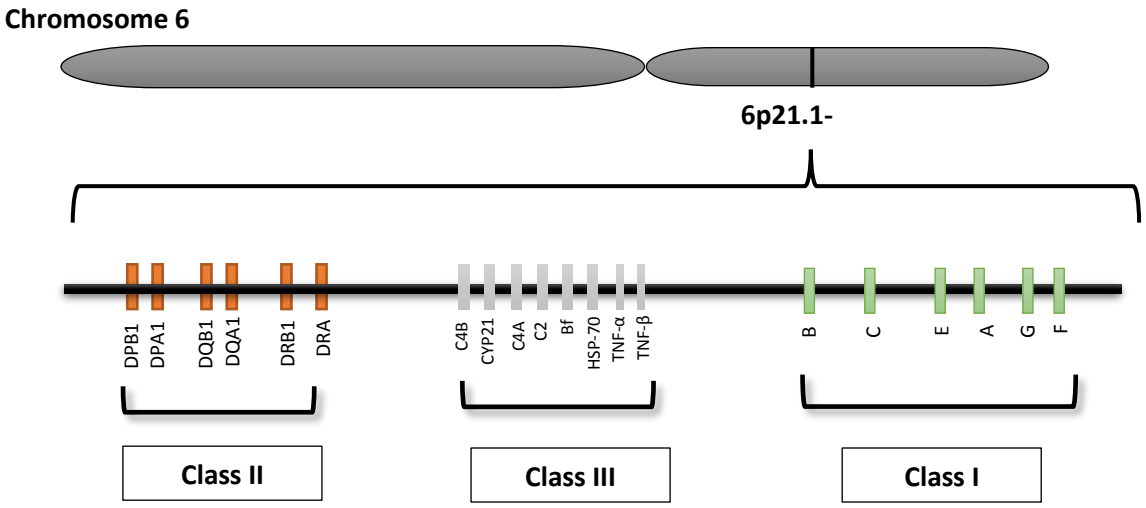


Figure 1.6: Gene map of the Human Leukocyte Antigen (HLA) Region, 6p21.

The HLA is the most polymorphic genomic region, which spans 3,807,696 nucleotides. The gene is divided into 3 main groups, Class I, Class II and Class III. Each of the classes is subdivided into specific genetic loci. Adapted from (Undlien et al. 2001).

Table 1.2: The classical and non-classical loci of the Class I and Class II HLA region.

Based on the EMBL-EBI database release 3.32, April 2018. http://www.ebi.ac.uk/ipd/imgt/hla/stats.html

Locus		Gene Type	Alleles	Notes/ Function
Classical *HLA* Class I				
HLA-A		coding gene	4,200	• Presenting peptides to T cells. • Inhibit the activity of NK cells.
HLA-B		coding gene	5,091	
HLA-C		coding gene	3,854	
Non-Classical *HLA* Class I				
HLA-E		coding gene	27	Indirect "detector" of *HLA* class Ia molecules allowing NK cells to monitor many *HLA* allelic variants with a single receptor.
HLA-F		coding gene	30	Mainly expressed in tonsil, spleen and thymus.
HLA-G		coding gene	60	• Predominantly expressed on placental trophoblast. • Might have a role in protecting the foetus from maternal NK cell attack.
Classical *HLA* Class II				
HLA-DR	*HLA-DRA*	coding gene	7	Comprise 1 α gene (*DRA*) and up to 9 β genes (*DRB*).
	HLA-DRB	*DRB* 1,3,4,5 (Coding genes) *DRB* 2,6,7,8,9 (Pseudogenes)	2,464	
HLA-DQ	*HLA DQA1*	coding gene	94	Constituted by 1 expressed α gene (*DQA*) and 1 β gene (*DQB*).
	HLA-DQB1	coding gene	1,196	
HLA-DP	*HLA-DPA1*	coding gene	65	
	HLA-DPB1	coding gene	975	
Non-Classical *HLA* Class II				
HLA-DM	*HLA-DMA*	coding gene	7	Help in constructing the antigenic peptides for class II *HLA* molecules.
	HLA-DMB	coding gene	13	
HLA-DO	*HLA-DOA*	coding gene	12	Regulation of class II *HLA* antigen processing.
	HLA-DOB	coding gene	13	

1.3.1.2. The Insulin Gene *(INS)*

INS is located on the short (p) arm of chromosome 11q15.5 and is associated with about 10% of genetic susceptibility to T1D (Steck and Rewers 2011). The promoter of *INS* is characterised by the presence of "the variable number of tandem repeat" (VNTR), which is a linear arrangement of multiple copies of short repeated DNA sequence "ACAGGGGTGTGGGG" and is known to be highly polymorphic amongst the same species. The sequence is repeated 28-44 times in the class I VNTR allele, and 138-159 times in the class III VNTR allele. The Class II VNTR allele has intermediately sized repeats and is rare (Stead et al. 2000). The *INS*-VNTR is located 0.5 kb upstream the transcription initiation site and has been shown to affect transcription rates. The association of *INS* with T1D was reported for the first time in the mid of 1980s when the length of the VNTR was linked to T1D progression in Caucasoids, showing that T1D is susceptible to the class I VNTR (Bell et al. 1984). Conversely, the inheritance of at least one copy of the class III VNTR allele reduced the probability of developing T1D 3-5-fold compared to the common I/I homozygous VNTR genotype (Bennett et al. 1995). Furthermore, expression of the class III allele was found to be associated with a higher (15-20 %) insulin mRNA secretion in adults and foetal human pancreatic cells when compared to those who express class I VNTR (Vafiadis et al. 1996). Higher expression of proinsulin mRNA in the thymus in early life is thought to induce central immune tolerance and eliminate T cell autoreactivity (Pugliese et al. 1997, Vafiadis et al. 1997).

1.3.1.3. The cytotoxic T-lymphocyte antigen-4 (*CTLA-4*)

CTLA-4, also known as *CD152*, is another confirmed T1D risk gene located on chromosome 2q33. *CTLA-4* consists of four exons and three introns that encode an immunoglobulin receptor that is expressed on the surface of active T helper cells (Dariavach et al. 1988, Harper et al. 1991). The CTLR-4 receptor has a critical role in T cell homeostasis, acting as an "off" switch for the signals mediated by CD80 or CD86 on the surface of active antigen-presenting cells. CD80 is a protein found on the surface of the dendritic cells, activated B cells and monocytes and provides a co-stimulatory signal necessary for T cell activation and survival. The CTLR-4 receptor maintains immune tolerance

through the inhibition of T cell activation, IL-2 production, and cell cycle progression (Eagar et al. 2002, Linsley et al. 1991). The link between *CTLA-4* and T1D was demonstrated in a Non-Obese Diabetic (NOD) model as *CTLA-4* knockout mice develop severe lymphoproliferative disorders (Waterhouse et al. 1995). There are at least three known polymorphisms in the *CTLA-4* which are associated with immune diseases: the promoter C-328T SNP (Deichmann et al. 1996), the A49G SNP in exon 1 (Nistico et al. 1996), and the (AT)n dinucleotide repeat polymorphism in the 3' untranslated region of exon 4 (Harper et al. 1991, Polymeropoulos et al. 1991). The A49G polymorphism has shown a strong association with T1D (Nistico et al. 1996), which causes changes in the primary amino acid sequence of CTLA-4 signal peptide resulting in a reduction in T helper cell immunoglobulin surface expression, and therefore, is likely to play a role in the central immune tolerance (Anjos et al. 2002). Further, the long (AT)n dinucleotide repeat has been shown to be associated with increased genetic risk for islet autoimmunity (de Jong et al. 2016).

1.3.1.4. *Protein tyrosine phosphatase, non-receptor type 22* (*PTPN22*)

The PTPN22 gene is located on chromosome 1q13 and encodes the lymphoid protein tyrosine phosphatase (LYP) which functions as a negative regulator of spontaneous T cell activation. LYP controls T cell activation through dephosphorylation of T cell trapping substrates such as Lck, Zap70, TCRζ/CD3ε complex, Vav and valosin-containing proteins (Wu et al. 2006). A substitution of the nucleotide C1858T within *PTPN22*, which changes codon 620 from arginine (CGG) into tryptophan (TGG), has been reported to be associated with T1D. This SNP disrupts the formation of the LYP-Csk complex, which usually shuts down the hyper-reactivity of T cells, making individuals who are heterozygous or homozygous for the tryptophan variant more predisposed to developing T1D (Bottini et al. 2004, Steck et al. 2006, Zoledziewska et al. 2008). Trp620 is also linked to a gain-of-function mutation that downregulates T cell receptor signalling (Vang et al. 2005).

1.3.1.5. *The Ubiquitin-associated and SH3 domain-containing protein A* (*UBASH3A*)

UBASH3A is a protein encoded by the gene *UBASH3A* that is located on chromosome 21q22.3 (Wattenhofer et al. 2001) and is expressed mainly on T cells (Carpino et al. 2004). UBASH3A is one of two family protein members belonging to the T cell ubiquitin ligand family that are known to negatively regulate T cell signalling (Collingwood et al. 2007). UBASH3A induces apoptosis in T cells by a growth factor withdrawal mechanism through its SH3 domain interaction with c-CBL through binding to ubiquitinated and ubiquitylated proteins via its UBA domain (Feshchenko et al. 2004, Tsygankov 2008). *UBASH3A* is considered to be T1D risk candidate loci (Barrett et al. 2009, Concannon et al. 2008, Grant et al. 2009, Johnson et al. 2012, Onengut-Gumuscu et al. 2015, Plagnol et al. 2011). The SNP rs876498, located in the 6th intron of *UBASH3A*, was shown to be associated with developing T1D. This has been shown through studying SNP genotyping data from a linkage study of T1D sibling pairs in about 2,500 multiplex families (Concannon et al. 2008). Follow-up of a 1715 SNPs study from the Wellcome Trust Cases Control Consortium genome-wide association study in T1D families confirmed *UBASH3A* as a genetic risk factor for T1D (Cooper et al. 2009). *UBASH3A* AA genotype could be useful for T1D risk prediction as it was shown to be an independent predictor of persistent islet autoimmunity and T1D in children who do not have a family history for T1D but carry the high-risk *HLA* genotype *HLA-DR3/4*, *DQB1*0302* (Johnson et al. 2012).

1.3.1.6. *IL-2 receptor complex* (*IL-2Rα* "*CD25*")

The *IL-2* gene is located on chromosome 10p15.1, and its protein IL-2 has been shown to have broad effects on T cells including survival, proliferation, activation-induced cell death (AICD), T cell differentiation, cytokine production and immune tolerance (Malek and Bayer 2004, Malek and Castro 2010). Low concentrations of IL-2 induce effector or memory T cell differentiation, while high concentrations of this cytokine induce the apoptosis of conventional T cells through AICD sensitisation (Bachmann and Oxenius 2007). The interleukin-2 receptor alpha subunit (IL-2R*α* "CD25") is expressed on activated T cells and is found to be associated with T1D (Lowe et al. 2007). T1D has been linked to

two SNPs located at the 5' end of intron 1 (rs706778 and rs3118470) (Qu et al. 2007). Further, downregulation of soluble IL-2Rα levels in the circulation is associated with susceptibility genotypes in T1D patients (Maier et al. 2009). Incongruously, none of the polymorphic variants of the *IL2Rα* gene has been shown to have a functional effect. Further, Belot and colleagues showed that there is a significant association between some differentially methylated CpG islands in the proximal promoter of *IL-2Rα* (CpGs -373 and -456) and the *IL2Rα* SNPs associated with T1D. They also showed an association between DNA methylation level at CpG -373 with SNPs within the promoter and intron 1 of *IL-2Rα* gene with T1D (Belot et al. 2013).

1.3.1.7. *The Interferon Induced with Helicase C Domain 1 Gene* (*IFIH1*)

IFIH1 is located on chromosome 2q24.2 and encodes the protein melanoma differentiation associated protein 5 (MDA5). As MDA5 plays an important role in recognising viral dsRNA during viral infection, *IFIH1* could potentially link environmental triggers and genetic susceptibility for T1D (Lincez et al. 2015). It was hypothesised that viral infections in patients with elevated MDA5 levels are mainly activated the MDA5 pathway, which leads to escalating antiviral responses and the secretion of interleukins and interferons (Lincez et al. 2015, van Belle et al. 2011). Remarkably, MDA5 signalling downregulation prompts regulatory T cell levels at the site of autoimmunity suggesting a potential protective mechanism (Downes et al. 2010, Lincez et al. 2015, Nejentsev et al. 2009). Furthermore, the minor allele of rs1990760 polymorphism in the *IFIH1* gene was shown to be strongly associated with T1D (Smyth et al. 2006, Todd et al. 2007). In the Chinese Han population, the allele rs3747517 but not rs1990760 is associated with developing T1D (Yang et al. 2012b).

1.3.2. Environmental Risk Factors

As discussed previously, the genes involved in T1D susceptibility explain why some individuals are more likely to develop T1D. However, it does not answer a question as to why one monozygotic twin develops T1D while the other sibling develops the condition later in life or, in some cases, does not? Defining the environmental trigger is not a straightforward task. The long gap between initiation and diagnosis of T1D (decades in slow progressors) have made it challenging to define environmental factors that can trigger autoimmunity. During this time window, a genetically susceptible individual may be exposed to multiple environmental factors. Different combinations of triggers may be required to develop autoimmunity toward the pancreatic beta cell. Therefore, in the next section, the important environmental factors that have been implicated in the pathogenesis of T1D will be summarised briefly.

1.3.2.1. Migration Studies

T1D epidemiology migration studies focused on migrants who moved from countries that have a low incidence of T1D such as Asian countries to high incidence areas such as Scandinavian countries. The main finding of these studies revealed that the migrants from low-incidence countries to high-incidence countries adopt the local risk rate of the country they migrated to (Oilinki et al. 2012). A study in Bradford, UK focused on Asian population children aged 0-16 years showed that the incidence of T1D had increased from 3.1: 100,000/ year in the period 1978-1981 to 11.7: 100000/ year in the period 1988-1990 among this Asian cohort. This incidence rate was higher than T1D rates of the native population children which were 10.5: 100,000/ year (Bodansky et al. 1992). Another migration study reached the same conclusion showing that Sweden born children who originated from low incidence countries had a higher risk of developing T1D (Hjern et al. 2012, Soderstrom et al. 2012). Additional evidence of the effect of early exposure to environmental factors on T1D incidence rates was from a study comparing the age of T1D onset in the Swedish population to those with non-Swedish born parents (Hussen et al. 2013). The study included individuals aged 0-30 years and showed that the onset of T1D shifted towards a younger age in both groups, emphasising the role of early life

environmental exposure to influence T1D initiation. On the other hand, some other studies failed to reach the same conclusion. A study done on Japanese children living in Hawaii who were diagnosed with T1D showed them to have the same incidence rates to their native Japanese (Patrick et al. 1997). Nonetheless, comparing T1D incidence in different ethnic groups in the Hawaiian population showed a variance. Japanese children had the lowest incidence rate (2.85-3.08: 100,000) while the half Hawaiian/half Japanese children had the highest incidence rates (15.34-16.58: 100,000). Caucasian children were the next highest with incidence rates (6.21-6.71: 100,000), followed by the Filipino children (3.66-3.96: 100,000) and finally the Japanese. The authors indicated that the variation in incidence rates might suggest that genetics play a more significant role in T1D development than environmental effectors (Patrick et al. 1997).

1.3.2.2. The Virus Hypothesis

The role of viruses in susceptibility to T1D is the subject of ongoing debate. The most extensive study to date, "The Environmental Determinants of Diabetes in the Young (TEDDY)"[4] failed to find a link between viral infections and T1D in young children (Lee et al. 2013). Meanwhile, other major studies reported the opposite conclusion. The Diabetes and Autoimmunity Study in the Young (DAISY)[5] found that enterovirus infections "accelerate" progression from the preclinical to the clinical stage in children (Stene et al. 2010). A nationwide study of children up to age 18 years in the period 2000-2008 in Taiwan compared the incidence rate of T1D among children who were diagnosed with enterovirus infections and reported a positive association between enterovirus infection and developing T1D (Lin et al. 2015). In parallel, a meta-analysis study had also confirmed a strong association between enterovirus infection and T1D. The study demonstrated that the risk of detecting enterovirus infection in children at the onset of T1D is almost ten-fold higher. They also found that T1D incidence is almost four-fold higher in pre-clinical T1D children when compared to control subjects (Yeung et al. 2011). The Diabetes Virus Detection study (DiViD) study also reported the presence of enterovirus in all

4 TEDDY is a longitudinal study was established in 2002 and are examining the environmental causes of T1D. The study follows children at high genetic risk for T1D from birth to 15 years of age.

5 DAISY is a study established in 2000 that study the causes and the risks of T1D

examined pancreatic islets samples (*n*=6) of T1D patients but in very few (2 of 9) in control islet samples (Krogvold et al. 2015). More recent data had shown that T1D children had been infected with enteroviruses three times more than control children samples (Honkanen et al. 2017)

The influence of viral infection on the progression of T1D is widely studied. The viral infection is believed to occur several years before the onset of T1D (Lonnrot et al. 2000). Rubella and enteroviruses are viruses that have been most closely linked to T1D development (Hyoty and Taylor 2002, Yeung et al. 2011). Enterovirus has been detected in the serum (Nairn et al. 1999), pancreas (Dotta et al. 2007) and small intestine (Oikarinen et al. 2008, Oikarinen et al. 2012) of the prediabetic paediatric T1D patients. Further, enterovirus was detected in sera of 64% of children who were diagnosed with T1D before the age of six (Clements et al. 1995). Among enteroviruses, Coxsackievirus B4 (CVB4) has been actively associated with T1D (Andreoletti et al. 1998, Hyoty and Taylor 2002). Coxsackievirus antibodies have been shown to induce apoptosis of pancreatic beta cells (Bason et al. 2013). The CVB4 infection has also been detected in islets of pancreatic samples from 3 of 6 T1D patients but not in the 26 tested control pancreatic tissue samples (Dotta et al. 2007). Furthermore, the enteroviral capsid protein vp1 (vp1) has been reported to be present at approximately 61% (44 of 72) in multiple islets of recently diagnosed T1D paediatric patients but in very few (3 out of 50 islets) age-matched controls (Richardson et al. 2009). Moreover, the presence of immunoreactive vp1 in beta cells of T1D patients has been shown to change the cellular phenotype in a way consistent with the activation of pathways of antiviral response and apoptosis (Richardson et al. 2013). Recently, data had shown that detecting enterovirus in young children stool samples is associated with the appearance of seroconversion within approximately 1 year post the infection (Honkanen et al. 2017). Interestingly, children carrying the high-risk haplotype *HLA-DR3/DR4* were shown to have a higher CVB4 antibody titre when compared with children who are carrying the protective *HLA-DR2* haplotype (Sadeharju et al. 2003). Overall, these finding support the hypothesis suggesting that enterovirus infections may play a role in initiating beta cell damage during T1D pathology.

1.3.2.3. Gut Microbiome and the Hygiene Hypothesis

The role of the intestinal flora in causing and preventing diseases is acknowledged. The development and maturation of the mucosal immune system is essential for the maintenance of immune homeostasis (Knip and Honkanen 2017). With regard to T1D, it has been shown that the normal mucosal immune system is capable of producing regulatory T cells which are known to have a critical role in preventing the development of T1D (Locke et al. 2006). Additionally, the composition of intestinal microflora of children at risk of T1D was found to be different from healthy controls. Intestinal microflora of children at high risk for T1D was found to be composed of non-butyrate-producing lactate-utilising bacteria that are known to prevent optimal mucin synthesis and as a consequence, form a "leaky gut" phenotype which is a commonly seen in T1D patients. In contrast, the gut microflora of healthy controls was composed of lactate and butyrate-producing bacteria that induce a sufficient amount of mucin that maintains gut integrity (Brown et al. 2011). Another interesting observation is that intestinal microbial antibodies are reported to be present in the circulation of Russian Karelian children at a higher concentration than in a matched population in Finland (Seiskari et al. 2007). Finland and Russian Karelia are two neighbouring countries whose populations share genetic background, yet the incidence of T1D in Finland is six times higher than in Russian Karelia (41.1/100000 VS 7.4/100000 child under 15 years) (Kondrashova et al. 2005). An investigation of school children in both countries showed that Karelian children who are living in a poorer hygiene environment developed a higher titre of microbial antibodies (antibodies against *Coxsackievirus B4*, *Helicobacter pylori, Toxoplasma gondii* and *Hepatitis A virus*) in their system (Seiskari et al. 2007). This observation suggests that children who were exposed to microbes in their early childhood developed better gut immunity and may be protected from autoimmune disorders. Again, this is supported by two observations 1) infants born by Caesarean section are at increased risk of T1D and 2) it has been shown that NOD mice housed under pathogen-free condition developed T1D spontaneously at a faster rate when compared to NOD mice which were housed in a conventional

environment (King and Sarvetnick 2011). Collectively, these findings suggest an important role of the microbiome as well as living in conventional environment in T1D prevention and/or development.

1.3.2.4. Vitamin D Hypothesis

Vitamin D refers to a group of fat-soluble secosteroids that function mainly as mineral absorption enhancers in the intestine. In humans, vitamin D3 (cholecalciferol) is the most abundant subtype (Bikle 2014). The activation of cholecalciferol occurs in two steps. The first step occurs in the liver in which the enzyme 25-hydroxylase, (CYP2R1) converts the cholecalciferol into 25-hydroxyvitamin D3-25-OHD3. Another hydroxylation takes place in the kidneys, where the enzyme 1-α hydroxylase (CYB27B1) converts 25-hydroxyvitamin D3-25-OHD3 into the active secosteroid hormone, 1α,25-dihydroxy vitamin D3 ($1\alpha,25(OH)_2D3$; vitamin D). Vitamin D exerts its action via binding to the nuclear vitamin D receptor (VDR). The *VDR* gene is located on chromosome 12q13.11 and is located very near to the vitamin D hormone hydrolysing gene *CYB27B1* (located 10 Mb centromeric of the *VDR* gene at 12q13.1-13.3) (Bikle 2014).

Vitamin D is considered a possible environmental susceptibility factor that can influence the immune attack in T1D patients. An association between T1D and vitamin D was first proposed when a significant relationship between vitamin D deficiency and pancreatic beta cell malfunction was reported (Norman et al. 1980). This observation has been confirmed in rodents by inducing mutations in *VDR*, which leads to insulin secretion reduction (Zeitz et al. 2003). Moreover, it has been shown that rendering NOD mice vitamin D deficient in early life induced impaired glucose tolerance within the first 100 days of life and a doubling of the prevalence of T1D in the following 100 days (Giulietti et al. 2004). This effect of vitamin D appears to be strain specific, as knocking *VDR* out in a different mouse strain (Zeitz) did not affect the function of beta cells (Mathieu et al. 2001). Conflicting conclusions of the effect of vitamin D on the progression of T1D has been reported by leading diabetes research groups around the world. The DAISY study group failed to find an association between vitamin D intake and 25(OH) D levels during childhood with the risk of islet autoimmunity or progression to T1D in their

population (Simpson et al. 2011). Another study in Denmark also failed to find evidence as exposing mothers during gestation and the child at first year of life to low dose of vitamin D can affect the risk of developing T1D before the age of 15 years (Jacobsen et al. 2015). On the other hand, the DISS study showed that patients at the onset of T1D have lower levels of plasma 25(OH)D when compared with a matched control group. Vitamin D was also shown to function as a protective factor from T1D in infants from seven European countries (EURODIAB Substudy 2 Study Group) (Group 1999). This protective role of vitamin D was confirmed by a meta-analysis of five retrospective studies (Zipitis and Akobeng 2008). A Finnish birth-cohort study concluded that the intake of dietary vitamin D in the first year of life decreases T1D risk in their birth-cohort study (Hypponen et al. 2001). Another interesting observation is the combined effect of two environmental factors in boosting the incidence of T1D; enterovirus infection has been reported to be more frequent in recent-onset T1D patients with co-existing vitamin D deficiency (unpublished data) (Penno et al. 2013).

1.3.3. The Effect of Environmental Interaction with Genetic Traits in T1D Development (Twin Studies)

Twins studies are valuable in medical research as they helped in understanding the effect of the environment on different human disease. Monozygotic (MZ) and dizygotic (DZ) twin studies provide an assessment of the effect and/or the contribution of genetic or environmental factors on the etiology of the disease. Despite sharing identical genetic polymorphisms in MZ twins (as they were developed from a single fertilised oocyte), they show significant phenotypic differences implicating the significant involvement of epigenetic and environmental factors (Fraga et al. 2005). A DZ twin is developed from two fertilised oocytes, and are shown to share 50% of genes, as non-twin siblings do (Boomsma et al. 2002).

Classic MZ twin and population-based registry studies showed that the concordance rate of T1D developing within identical twins is 50% and this percentage decreases in DZ twins down to 10 % (Hyttinen et al. 2003, Kaprio et al. 1992, Kyvik et al. 1995, Redondo et al. 2001). The study also revealed that there is a variation in the timing of seroconversion and development of overt T1D among MZ siblings who were followed-up for an extended period. In some cases, the seroconversion occurred in the second twin, three decades after T1D development in the first sibling, suggesting a considerable impact of the environmental factors in developing T1D. In another study, siblings of 53 MZ and 30 DZ twin pairs who had T1D were tested for the presence of autoantibodies (Redondo et al. 1999). The data showed the rate of having multiple autoantibodies in MZ twins is higher in comparison to DZ twin siblings. These data also revealed that having multiple autoantibodies in MZ twins is linked to the expression of the high-risk *HLA* class II haplotypes (*HLA DQ8/DQ2*).

1.4. Autoimmunity in T1D

1.4.1. Beta Cell Death in T1D (Insulitis)

The infiltration of the immune and inflammatory cells into and around the beta cells in the pancreatic islets "insulitis" is a hallmark phenomenon in autoimmune beta cell-specific destruction (Gepts 1965, Itoh et al. 1993). According to the consensus guidelines for the diagnosis of insulitis in humans (2013), insulitis is defined as follows: 1) "*the presence of a predominantly lymphocytic infiltration specifically targeting the islets of Langerhans*". 2) "*The lesion should be established in a minimum of three islets, with a threshold level of ≥ 15 CD45$^+$ cells/islet*". 3) "*The infiltrating cells may be found in the islet periphery (peri-insulitis) or diffuse and present throughout the islet parenchyma (intra-insulitis)*" 4) "*The presence of pseudoatrophic islets (insulin-negative)*" (Campbell-Thompson et al. 2013, Campbell-Thompson et al. 2017, In't Veld 2014). Up to date, our most extensive knowledge about beta cell inflammation is derived from studying the histopathology of pancreatic samples obtained from patients near the time of diagnosis (Foulis et al. 1986, Imagawa et al. 2001, Krogvold et al. 2014, Willcox et al. 2009).

Insulitis was described for the first time over a century ago when a characteristic inflammatory infiltration was documented in the islets of Langerhans of a child that died from diabetic ketoacidosis (Schmidt 1902). This immunological cell infiltration was referred to as "insulitis" (Von Meyenburg 1940). In a landmark paper published in 1965, Willy Gepts had described the insulitis as infiltration of lymphatic cells to the pancreatic islet. He documented the presence of insulitis in 68% (15 out of 22) of at onset T1D individuals while it was absent in long-duration T1D patients (Gepts 1965). Two decades later, Bottazzo and his group had observed that the CD8$^+$ T cell infiltration in and around the islets is associated with a dramatic upregulation in the expression of MHC class I on the pancreatic beta cell surface (Bottazzo et al. 1985). Later, in 1986, a landmark study done by Foulis and his group who used samples obtained from 119 young patients who died before the age of 20 in the UK, had shown that insulitis was present in 78% (47 out of 60) among recent onset of T1D young patients (<1 year). The study also found that the insulitis was present in 23% of insulin-containing islets and present

in only 1% of insulin-deficient islets. This latter observation had suggested that insulitis represents an immune-mediated process that results in pancreatic beta cell destruction (Foulis et al. 1986).

Foulis and his colleagues, in a separate study, confirmed the hyperexpression of MHC class I on the surface of beta cells and showed that interferon γ (INFγ) levels were elevated as well (Foulis et al. 1991). The hyper-expression of HLA class I antigen on islet cells is associated with STAT1 expression upregulation, which uniquely occurs in T1D patients and thereby contributing to their selective susceptibility to autoimmune-mediated destruction (Richardson et al. 2016). The characterisation of the immune cells infiltrating indicated that cytotoxic $CD8^+$ T cells are the most abundant cell type, followed by $CD20^+$ B cells, macrophages ($CD68^+$) and finally $CD4^+$ T cells. The $CD138^+$ plasma cells, Treg cells and natural killer (NK) cells were rarely detected (Willcox et al. 2009). Collectively, the data confirm the involvement of both innate and adaptive immune systems in the insulitis process.

The presence of insulitis seems to be dependent on both the onset age and the duration of the disease. For instance, the average onset age among those who were insulitis positive in Gepts' study was 7.9 years while the average age of the cases without insulitis was 17.1 years (Gepts 1965), suggesting that insulitis is present in patients in whom T1D onset occurred at a young age. A meta-analysis of 213 T1D cases had shown that insulitis was present in 73% of young (≤ 14 years) T1D patients who had short disease duration (≤ 1 month); in 60% of young patients with a disease duration of 1 month to 1 year, and only in 4% of young patients with long T1D duration (>1 year) (In't Veld 2014). Our limited understanding of human insulitis and beta cell death has come from the histopathological analysis of pancreatic autopsy specimens collected from T1D patients near the time of diagnosis, e.g. Foulis Study 1986. Fresh biopsy samples from new onset T1D patients taken recently in the Diabetes Virus Detection Study from Norway and in earlier studies from Japan provided additional information, but they stopped recruiting patients following unexpectedly high complication rates (Imagawa et al. 2001, Krogvold et al. 2014). During the last ten years, the Juvenile Diabetes Research Foundation (JDRF) has supported the initiation of an organised effort to study T1D patients in the USA (Network for Pancreatic Organ Donors with Diabetes; nPOD). The nPOD aimed to launch a

research resource for T1D by obtaining samples (such as the pancreas, pancreatic lymph nodes, spleen, thymus and blood) from organ donors with diagnosed and pre-clinical T1D patients. These samples are distributed to worldwide nPOD-approved investigators, to allow the broad and diverse study of human insulitis as well as the causes and the mechanisms of beta cell death and T1D. The Initiative also endorsed collaboration between different T1D research groups via sharing tissue and data, and through initiating and managing cooperative projects. Collectively, nPOD efforts should facilitate a comprehensive understanding of human T1D (Campbell-Thompson et al. 2012, Pugliese et al. 2014).

1.4.1.1. T Cells in T1D

In humans, the development of self-recognition of T cells occurs in the thymus. During early life, T cells are negatively and positively selected in the thymic medulla and cortex, respectively (Kappler et al. 1987, Takaba and Takayanagi 2017). The mechanism of self-recognition is essential to minimise autoreactivity of T cells in the periphery later in life. Mature T cells express unique T cell receptors (TCR) and are divided into two distinct populations upon their surface expression: the co-receptor molecules $CD4^+$ and $CD8^+$. Each population binds to antigens presented by different classes of MHC molecules and once activated, each subtype differentiates into functionally distinct effector cells. $CD8^+$ T cells have a cytotoxic role, while $CD4^+$ T cells "help" other immune cells during adaptive immunity. The cluster of differentiation (CD) receptors have a more important role than just being a phenotypic marker; they stabilise the interactions between the peptide-MHC complexes on the surface of antigen presenting cells (APC) and the TCR during immune responses (Chaplin 2010, Parkin and Cohen 2001). Evidence shows that T cells play an important role in T1D pathogenesis, for instance, treatment with immunosuppressive drugs that target T cells, e.g. cyclosporine A, although it did not affect the level of circulating autoantibodies but it delayed the progression to clinical diabetes (Bougneres et al. 1988, Mandrup-Poulsen et al. 1990). Further, treating recently diagnosed T1D patients with anti-T cell monoclonal antibodies improved the insulin production in these patients (Herold et al. 2002, Killestein 2002). Furthermore, selective beta cell destruction in a segment of

healthy pancreas that was transplanted in a T1D patient from his non-diabetic identical twin strongly suggests the involvement of immunological memory in islet-specific T cells (Sibley et al. 1985).

1.4.1.1.1. $CD4^+$ T Cell

The $CD4^+$ helper (Th) cells have been shown to be present at low levels during all phases of insulitis (Richardson et al. 2011). Previously, Th cells have been subdivided into two lineages, Th1 and Th2, based on the expression of molecules on their surface, the cytokines they release and their transcription factors (Mosmann et al. 1986, 2005). Now, there are at least five identified Th subpopulations including Th1, Th2, Th17, inducible T regulatory cells (iTregs) and follicular T helper cells (Tfh) (Walker and von Herrath 2016). Th1 cells have been shown to be involved in the pathogenesis of T1D in mice, as T cells from neonatal NOD mice express a diabetogenic TCR when the T cell is differentiated into a Th1, but not a Th2 phenotype (Katz et al. 1995). Th2 cells were also reported to be involved in T1D progression (Almawi et al. 1999, Poulin and Haskins 2000). Although still not fully understood, it has been recently suggested that Th17, as well as Tfh subpopulations contribute to T1D development (Heuts et al. 2017, Kenefeck et al. 2015, Walker and von Herrath 2016).

1.4.1.1.2. $CD8^+$ T Cell

Cytotoxic ($CD8^+$) T cells are believed to have an essential role in insulitis initiation and beta cell destruction (Coppieters et al. 2012, Roep 2003, Roep and Peakman 2011, van der Torren et al. 2016) as they are the most abundant cell type found in the insulitis lesion (Willcox et al. 2009), and they are found in the circulation of T1D patients at onset (Roep et al. 1990). The exact mechanisms by which $CD8^+$ T cell-mediated beta cell destruction are still not fully clear. Nevertheless, the process is suggested to involve: 1) direct induction of cell death though releasing perforin and granzymes and the binding of Fas Ligand to the Fas receptor. 2) Indirect cytotoxicity by secretion of inflammatory cytokines such as TNF and INFγ, with the generation of free radicals that induce endoplasmic reticulum stress (Figure 1.7) (Kaech et al. 2002, Thomas et al. 2010, Trapani and Smyth 2002).

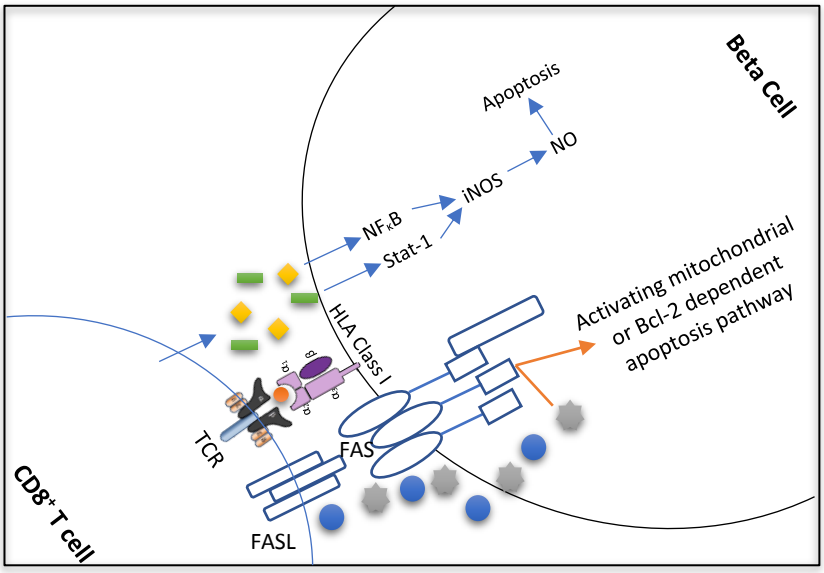


Figure 1.7: A model of CD8+ T cell-mediated destruction of human pancreatic beta cells.

The TCR of CD8+ T cell identifies beta cells via its expression of MHC class I. Mechanisms of beta cell destruction include: production of perforin and granzymes, Fas/FasL interaction that activates the Bcl-2-regulated or mitochondrial apoptosis pathway. Other mechanisms include releasing of cytokines that induce free radical production leading to endoplasmic reticulum stress and cell apoptosis.

Granzymes *Perforin* *INFγ* *TNF*

The figure was adapted from (Brissova et al. 2005, Thomas et al. 2010)

1.4.1.2. B Cells in T1D

Although the specific pathogenic role of B cells in beta cell mediated destruction is not apparent, the presence of activated B cells in sections of islets near to T1D diagnosis suggest their role in the pathogenesis of T1D (Boldison and Wong 2016; Richardson et al. 2011; Willcox et al. 2009; Wong and Wen 2012). It has been documented that $CD20^+$ B cells were present in only small numbers in early insulitis (Willcox et al. 2009). However, in cross-sectional studies, their infiltration was found to peak during the late-stage of insulitis in parallel with the time when beta cell mass declines (Richardson et al. 2011; Willcox et al. 2009). Additionally, variable infiltrations of $CD20^+$ B cells phenotypes (CD20Hi and CD20Lo) were reported to be related to the age of T1D diagnosis. Leete and colleagues reported that CD20Hi B cells predominate within B cell infiltration of the islets of children who were diagnosed before the age of 7 years, while the CD20Lo B cell phenotype was predominant

in those diagnosed after the age 13 years. They also showed that The CD20Hi phenotype was associated with more aggressive and rapid T1D onset (Leete et al. 2016).

The role of B cells in the process of beta cell destruction involves producing pro-inflammatory cytokines and presenting the beta cell-specific antigens to $CD4^+$ T cells (Boldison and Wong 2016; Silveira and Grey 2006). The elimination of B cells via anti-CD20 monoclonal treatment (Rituximab) partially reverses diabetes in recent-onset T1D human patients as well as attenuating beta cell loss (Pescovitz et al. 2009; Pescovitz et al. 2014). Treating transgenic NOD mice (expressing human $CD20^+$) with a single cycle of anti-human $CD20^+$ monoclonal antibodies (temporarily eliminating B cells) significantly delayed the onset of diabetes in mice and was shown to reverse diabetes in approximately one-third of diabetic NOD mice (Hu et al. 2007). Furthermore, the study showed that treatment with anti-$CD20^+$ monoclonal antibodies partially enhanced the diabetes outcome by influencing Treg clones during the period of B cell reconstitution in NOD mice (Hu et al. 2007; Xiu et al. 2008). Similar results were observed when B cells were eliminated via anti-CD22 monoclonal treatment in the NOD model (Fiorina et al. 2008). Depleting B cells by inactivating the μ immunoglobulin gene in the NOD mice also showed to protect from insulitis and T1D (Serreze et al. 1996). Additionally, the increase in CD19 expression on the surface of B cells enhanced the presentation of islet-specific glucose-6-phosphatase catalytic subunit-related protein (IGRP), leading to proliferation of reactive T cells specific for beta cells antigens (Ziegler et al. 2013a). The report also demonstrated that downregulation of CD19 had significantly inhibited the proliferation of $CD8^+$ T cells, suggesting that altering the CD19 signalling pathway may have a role in preventing islet-specific autoreactive T cell proliferation (Ziegler et al. 2013b). Together, these strands of evidence suggest the critical role of B cells in the pathogenesis of T1D. More evidence is required to confirm these data. It would be interesting to study the effect of B cell elimination in individuals “at risk” for T1D who carry high genetic risk alleles and/or are positive for islet autoantibodies.

1.4.2. Autoimmune Destruction of Beta Cells in T1D

During the early stages of insulitis, activated local APCs migrate to the pancreatic lymph node, recruit, and activate $CD4^+$ helper T cells by presenting beta cells antigens, leading to the release of chemokines and cytokines. $CD4^+$ helper T cells then stimulate APC to secrete more cytokines and nitric oxide. The released cytokines induce the secretion of chemokines by surrounding endothelial cells, which enhance the recruitment of more immune cells into the islets and, together with cytokines, activate cytotoxic $CD8^+$ T cells. Binding of the cytokine IL-1β to its receptor on the surface of beta cells activates the Nuclear Factor κB (NF-κB), Protein Kinase C (PKC), p38, and c-Jun N-terminal kinase (JNK). IL-1β binding also prompts the construction of a multiple-protein complex: IL-1R accessory protein (IL-1R AcP), Toll-interacting protein (Tollip), Myeloid differentiation primary response gene 88 (MyD88) and the IL-1 receptor-associated kinases (IRAKs): IRAK-1 and IRAK-4. The cytokine TNF α activates caspase-8, NF-κB and the mitogen-activated protein kinase (MAPK) p38 and JNK, whereas the cytokine IFNγ activates Stat-1 and the extracellular signal-regulated kinase (ERK). Collectively these activation pathways induce beta cell apoptosis (Pirot et al. 2008, Pugliese 2016).

1.4.3. Mechanism of Beta Cell Death in T1D, Necrotic or Apoptotic Pathway?

Necrotic cell death is considered to be a possible mechanism of beta cell damage in T1D by which cytolytic $CD8^+$ T cells mediate the cell death (Figure 1.7). As mentioned earlier, predominance of $CD8^+$ T and macrophage cell infiltration might mediate necrosis (Eizirik and Darville 2001, Kurrer et al. 1997). Necrotic cell death can result as a consequence of releasing granules that contain granzymes and perforin by the T cells. Perforin acts on the membrane of the beta cell forming a pore in Ca^{2+} dependent mechanisms. The pore allows the access of the serine protease granzyme into the cell which results in cleavage and activation of several targets, such as effector caspase-3 and the BH3 protein Bid (Eizirik and Mandrup-Poulsen 2001). Other necrotic cell death mediators are involved including Ca^{2+} and reactive oxygen species which lead to mitochondrial injury and affect both cell membrane integrity and ion balance (Jorns et al. 2014, Zhao et al. 2015). Another factor that may initiate necrotic cell death in beta cells is the activation of serine/threonine kinase receptor-interacting

protein 1 (RIP1) which is prompted by ligand-receptor interactions and ATP depletion (Thomas and Kay 2011). Necrotic beta cells release their cellular contents including high mobility group box 1 proteins, heat shock proteins, uric acid and pro-inflammatory cytokines, which trigger phagocytosis and immunological activity. Genetic pathway studies have suggested the involvement of apoptotic beta cell death pathways (Marroqui et al. 2015). Beta cell apoptotic pathways are possibly activated by TNF Fas/FasL pathway (Figure 1.7). Fas (CD95), which is a member of the TNF family, is activated through its binding to Fas L (CD 178) at the surface of beta cells. Upon activation, Fas recruits the Fas-associated death domain which initiates recruitment and activation of pro-caspase-8, cleaving the effector caspase-3 or activating the BH3 interacting domain death agonist which leads eventually to beta cell apoptosis (Pirot et al. 2008). Caspase induced apoptosis results in the emergence of the apoptotic cell features such as chromatin condensation and nuclear fragmentation (Elmore 2007). In T1D, beta cell death is most probably induced through the activation of both necrotic and apoptotic pathways, as both have been documented in cytokine-mediated islet cell death and post-death receptors FasL and TNFR activation (Stephens et al. 1999, Walter et al. 2000).

1.4.4. Autoantibodies to Islet Antigens

Islet autoantibodies are detected in the circulation of preclinical T1D individuals (Bingley et al. 2001, Pihoker et al. 2005). Although islet antigen-specific autoantibodies are not believed to be involved directly in beta cell destruction, they reflect the presence of beta cell damage, and therefore, are considered to be diagnostic rather than causative factors in T1D (De Filippo et al. 1997, Wong et al. 2004). By analogy to the TCR of T cells, B cells express receptors known as B cell receptor (BCR). The BCR components that are responsible for the recognition of antigens are known as immunoglobulins. When an antigen binds to the BCR, the B cells differentiate into an antibody-producing $CD138^{+}$ plasma cell. Consequently, plasma cells produce antibodies with a specificity equivalent to the immunoglobulin that recognised the antigen. The discovery of the first islet antigen autoantibody, the islet cell cytoplasmic antibody (ICA) (Bottazzo et al. 1974), contributed greatly to the understanding of the pathogenesis of T1D and in the prediction of overt diabetes. The main

disadvantage of ICA screening, at the time, was the difficulty in standardising the detection assay (Lernmark et al. 1991). Currently, there are four main autoantibodies assays that predict T1D with a high specificity; insulinoma-associated antigen 2 autoantibodies (IA-2A, ICA512) (Bonifacio et al. 1995, Rabin et al. 1992), insulin autoantibodies (IAA) (Palmer et al. 1983, Williams et al. 1997), glutamic acid decarboxylase 65 autoantibodies (GAD65A) (Baekkeskov et al. 1990) and the zinc transporter 8 autoantibodies (ZnT8A) (Long et al. 2012, Wenzlau et al. 2007). More recently, a fifth autoantibody which recognises a 38 kDa pancreatic islet glycosylated membrane-associated protein has been detected in >30 % of T1D patients: the tetraspanin protein family member (TSPAN7) known as glima38 (McLaughlin et al. 2016). The IA-2, ZnT8, and glima38 antigens exist on the membrane of secretory granules of the beta cell, insulin is contained within the granules, while the GAD65 is found in the cytoplasm of synaptic-like microvesicles (SLMV) (Lampasona and Liberati 2016). Although autoantibodies are considered to be diagnostic, several reports had suggested that autoantibodies specific to islet antigens can induce T1D by inducing reactive $CD4^+$ T cell (Silva et al. 2011) and $CD8^+$ T cell expansion (Reijonen et al. 2000).

1.5. The Progression of T1D

The chronic progressive autoimmune disease model of T1D was suggested by Johnny Ludvigsson at a Nordic Symposium and later, by George Eisenbarth in 1986 (Eisenbarth 1986). Although this model has been subjected to several modifications over the years as our understanding of the etiology of T1D enhanced, it remained the reference model of describing the natural history of T1D (Figure 1.8).

1.5.1. The Linear Beta Cell Declines Hypothesis

The linear model was suggested by Eisenbarth in 1986 (Eisenbarth 1986). The hypothesis implies that at some point, the genetically susceptible individuals are exposed to the environmental agent(s) that trigger islet autoimmune reaction initiation, which will result in the development of islet-specific autoantibodies, a linear decay in beta cell mass, a complete loss of C-peptide and the appearance of hyperglycaemia as the endpoint. Although the suggested linear model explains the sequence of events that occur during T1D pathogenesis, it failed to explain the variable prediabetic time window seen in T1D patients. This led to the conclusion that the process of progression to clinical T1D is non-linear, but instead, occurs as variable steps (Chatenoud and Bluestone 2007).

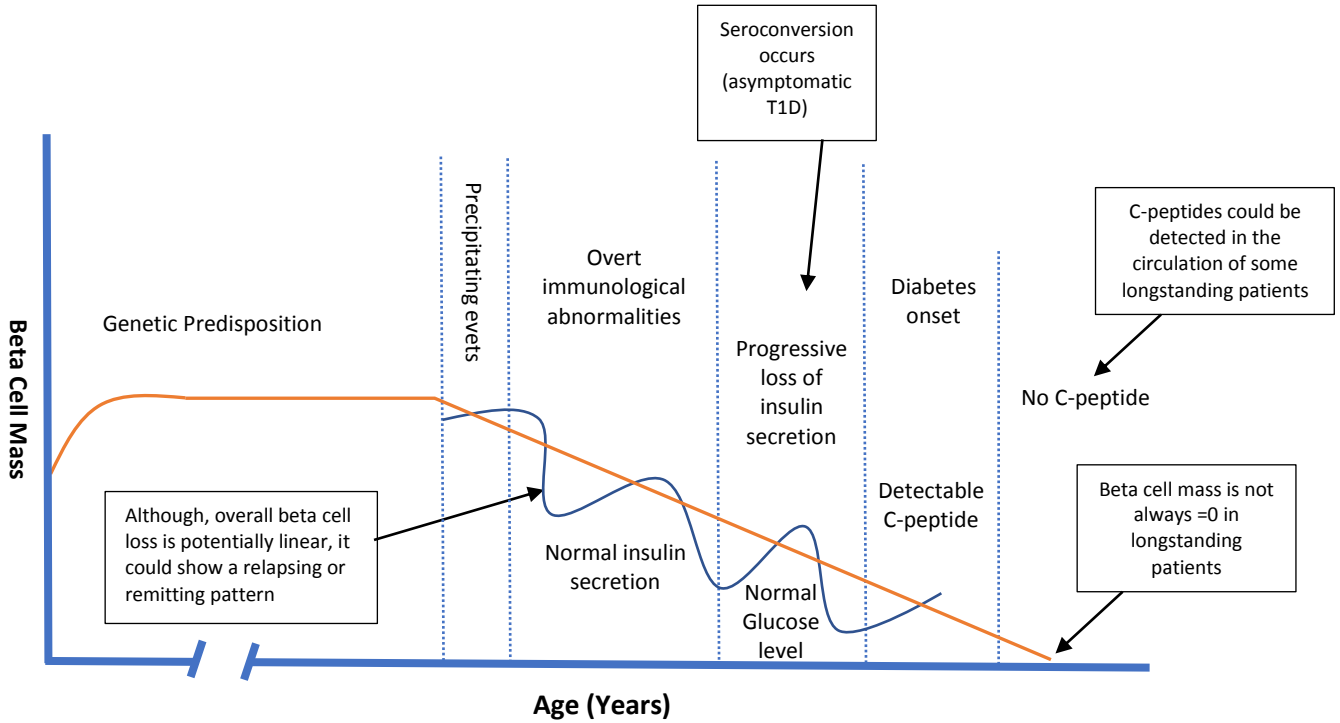


Figure 1.8: The timeline model of type 1 diabetes progression.

1.5.2. The Non-Linear Relapsing-Remitting Model

The relapsing-remitting model, proposed by Van Belle in 2011, is a very informative version of the nonlinear model (Bonifacio et al. 1999, van Belle et al. 2011, von Herrath et al. 2007). This model suggests that the selective autoimmune destruction of pancreatic beta cells in T1D results from failure of central and/or peripheral immune tolerance (Figure 1.9). Generally speaking, central tolerance originates from sites of lymphocyte development such as the thymus (for T cells) and bone marrow (for B cell) while peripheral tolerance is related to sites of antigen recognition and/or processing. The elimination of T cell auto-reactivity takes place in the thymus early in life, failure to correctly complete this process results in the release of autoreactive T cells to the circulation that recognise self-antigens as foreign molecules. In the case of T1D, failure of adverse selection of insulin-responsive T cells in the thymus results in a breakdown in central immune tolerance and consequently, T cells attack self-islet antigens (Klein et al. 2014, Kronenberg and Rudensky 2005). Analysis of islet infiltration in pre-diabetic NOD mice revealed the domination of $CD4^+$ and $CD8^+$ T cell subtypes (Wegmann et al. 1994). Later, studies showed that among all T cell subtypes, $CD8^+$ T cells have the most exclusive specificity towards islet autoantigens in both humans and the NOD mouse (Coppieters et al. 2012). Remarkably, these autoreactive effector T cells are not exclusively found in the periphery of T1D patients, but they are detectable in the circulation of healthy individuals as well (Arif et al. 2004, Peterson et al. 1999, Roep et al. 1999a). So, why are they pathogenic in T1D and not in healthy individuals? Healthy individuals are protected from pathogenicity of effector T cells by the action of other populations of specialised T lymphocytes called regulatory T cells (Treg) (Figure 1.9). Treg cells are a minor subpopulation of T cells that develop in the thymus and migrate to the periphery, where they maintain tolerance to self-antigens by suppressing self-reactive T cells that escape adverse selection in the thymus (Groux et al. 1997, Roncarolo and Levings 2000). IL-2/IL-2R signalling is vital for Treg development in the thymus, as well as their function and maintenance in the periphery (Chinen et al. 2016, Jenkinson et al. 1987, Raulet 1985, Tentori et al. 1988). This observation links three elements seen in T1D patients: low IL-2 concentration → IL-2Rα expression downregulation → Treg dysfunction leading to the breaking of

peripheral immune tolerance and the dominance of effector T cells. Studies on NOD mice have supported this conclusion. In NOD mice, it has been observed that Treg cell apoptosis peaks at the onset of T1D due to defective IL-2 production (Tang et al. 2008).

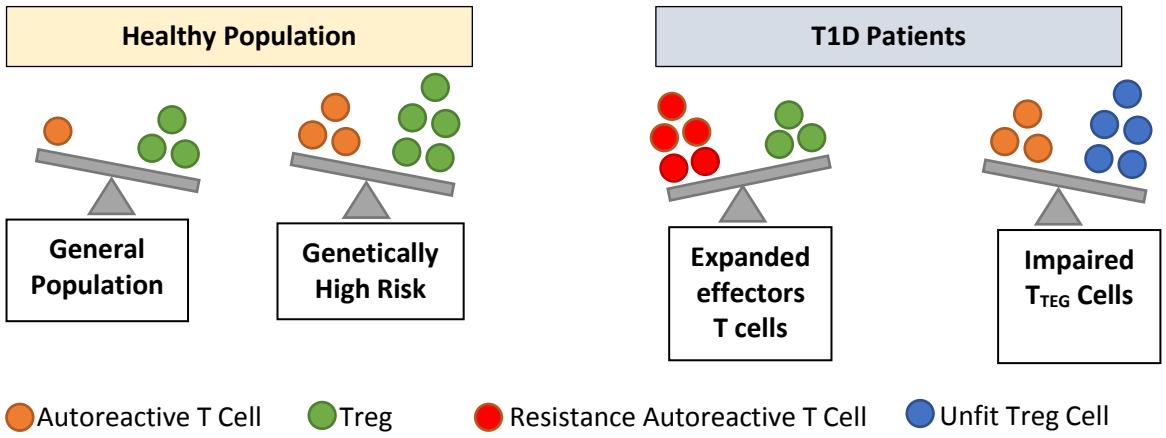


Figure 1.9: Immune regulations imbalance in T1D.

Treg cells have impaired function in patients with T1D, thereby failing to suppress autoreactive effector T cells. Alternatively, immune regulation might not suffice if effector T cells are resistant to immune regulation, as reported in patients with T1D. The information in the figure is based on (Roep and Tree 2014).

1.5.3. Fertile Field Model:

Von Herrath suggested the fertile field model of T1D progression in 2003. This model describes factors that lead to T1D autoimmunity during the time window post multiple viral insults. The hypothesis proposed that post each viral infection; T cell autoreactivity is induced by molecular mimicry or bystander activation, resulting in initiation of the autoimmune reaction and progression to clinical T1D (von Herrath et al. 2003).

1.6. Epigenetics

1.6.1. General Introduction

The human body contains approximately $3.7X10^{13}$ cells (Bianconi et al. 2013), each of which contains about two meters of DNA (Alberts et al. 2002). In a typical eukaryotic cell, the DNA and its associated protein (histone) are packaged inside the nucleus. The DNA is wrapped around the histone composing of repetitive units called nucleosomes. Each nucleosome consists of duplicates of the core histone octamers H2A, H2B, H3, and H4, which are wrapped 1.67 times in left-handed turns by 147 bp of DNA. The wrapper DNA enters and leaves the nucleosomes at points close to each other, a point where a fifth core histone H1 (the linker core histone) attaches to the DNA between two neighbouring histones (Kouzarides 2007) (Figure 1.10). The nucleosomes are the building unit of chromatin, the compact form of DNA inside the nucleus. The DNA sequences of all human body cells are similar, yet, each specialised cell (e.g. hepatocyte, keratinocytes, neurons, etc.) have different structures and functions. The diversity in the structure and the function of different cells results from the diversity of genes expression in different cell types. In the human genome, The DNA and its associated histones can be tagged by chemical molecules that have been shown to control the destiny of gene expression in the cell (Bird and Wolffe 1999, Dong and Weng 2013), these modifications are called Epigenetics. The Greek words (Epi) means “above” the genome; epigenetics are defined as modifications that change the function of the molecule (phenotype) without altering its primary structure (genotype) (Dupont et al. 2009). Several epigenetic modifications are known to control gene expressions such as the covalent modifications of DNA and histones, chromatin remodelling and the non-coding RNAs (Collins et al. 2017, Portela and Esteller 2010). The epigenetic modifications are also fundamental in controlling many other cell regulatory functions such as DNA-protein interactions, suppression of transposable element mobility, cellular differentiation, X-chromosome inactivation and genomic imprinting (Handy et al. 2011). Dysregulation of the epigenetic factors was shown to be associated with the development of vital disorders such as cancer (Jones and Baylin 2002), heart failure (DiSalvo

2015), learning difficulties (Franklin and Mansuy 2011) and autoimmune disorders (Farh et al. 2015, Quintero-Ronderos and Montoya-Ortiz 2012).

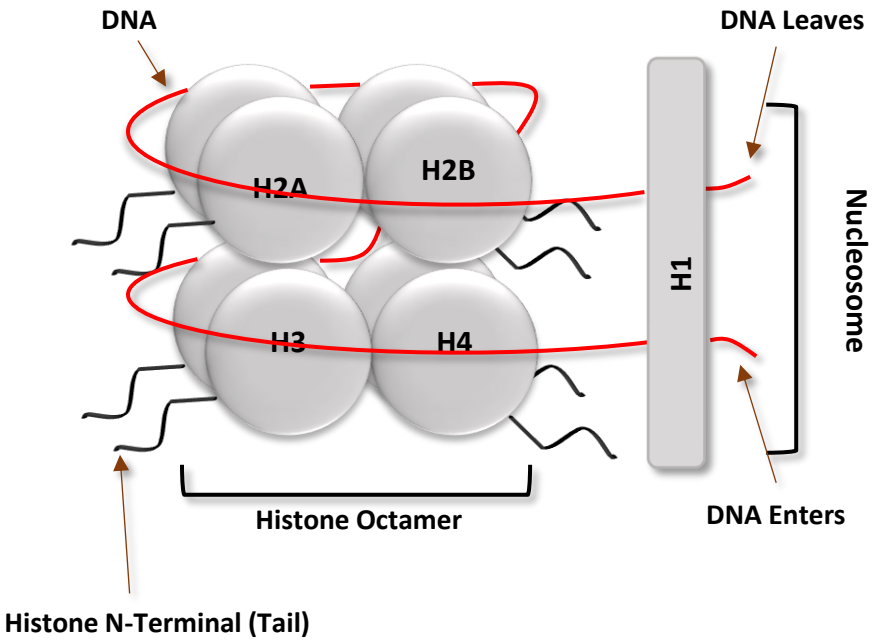


Figure 1.10: The Structure of the Nucleosome.

1.6.1.1. microRNA (miRNA)

MicroRNAs (miRNAs) are short non-coding RNA molecules (20-22 nucleotides) which are known to have a post-transcriptional gene regulating function. To date, there are 38589 identified miRNAs in 271 species (Entry number 22, March 2018, www.mirbase.org). The main regulating mechanism of miRNA is to "silence genes", a function that is reported to be initiated partially through binding to the 3'-untranslated end of a gene transcript (mRNA) to prevent its translation (Bartel 2004). Other regulating functions of miRNAs are activating the translation process of some genes (Vasudevan 2012) and heterochromatin formation (Kim et al. 2009). These regulating functions of miRNAs are described to be 'multiple' and 'cooperative', which means that one miRNA can target multiple mRNAs, meanwhile, one mRNA can be targeted by several miRNAs (Peter 2010). The majority of miRNAs are transcribed by the enzyme RNA polymerase II (Pol II), others, especially those which are lined within the repetitive elements, are transcribed by the RNA polymerase III (Pol III) (Borchert et al. 2006, Lee

et al. 2004). The translation of miRNAs undergoes several steps. First, the enzyme RNA Pol II transcribes the stem looped pri-miRNA. Then Drosha (ribonuclease III enzyme) cleaves pri-miRNA into hairpin structured pre-miRNA. Next, the pre-miRNA is transported into the cytoplasm by exportin-5 where a second cleavage occurs. The endoribonuclease Dicer cleaves the pre-miRNA into imperfect duplex strands of about 22 nucleotides, the mature miRNA and the passenger miRNA (Zamore et al. 2000). Finally, the mature strand either exerts its function inside the cells by fusing with Argonaute proteins (mainly Aro1 and Aro2) and integrating with the RNA-induced silencing complex (RISC) (Chendrimada et al. 2005) or is secreted into the circulation as 'circulating miRNAs'. Circulating miRNAs are either secreted as a vesicle-free form coupled with RNA binding proteins such as lipoproteins and Aro2 or encapsulated in vesicles as exosomes, microvesicles or apoptotic bodies (Kosaka et al. 2010).

miRNAs play important roles in many diseases including cancer (Jansson and Lund 2012), cardiovascular (Small and Olson 2011) and autoimmune diseases (Qu et al. 2014, Singh et al. 2013). miRNAs are also central regulators of beta cell biology as they are functionally associated with multiple mechanisms including the immune system (O'Connell et al. 2010), beta cell differentiation, function and survival (in particular MIR-375) (Lahmy et al. 2014). Furthermore, miRNAs function in response to viral infections (Ho et al. 2016), particularly in response to enterovirus infection. Enterovirus infection significantly dysregulates the miRNAs in human pancreatic islets (Nair et al. 2013) implicating that miRNA dysregulation could be a key mechanism of enterovirus-induced beta cell damage. Indeed, dysfunction to specific miRNAs has been linked to T1D development (Assmann et al. 2017).

1.6.1.2. The Epigenetic Covalent Modifications

The epigenetic covalent modifications occur when a chemical group is added directly to the DNA (commonly the methylation of cytosine nucleotides) or the histone tail (lysine acetylation, lysine and arginine methylation, serine and threonine phosphorylation, lysine ubiquitination and

sumoylation) (Bannister and Kouzarides 2011). Histones are essential key elements in epigenetics as they have important roles in transcriptional regulation (Karlic et al. 2010), DNA repair (Huertas et al. 2009) alternative splicing (Luco et al. 2010, Zhou et al. 2014) and chromosome condensation (Kouzarides 2007). The covalent modifications of the histone tail regulate gene expression, for instance, the acylation of H3K9, H3K27 and H3K18 in the promoter region is correlated with transcription activation (Kurdistani et al. 2004) while trimethylation of the histone H3K9, H3K27 and H3K79 is associated with transcription repression (Vakoc et al. 2006). The second epigenetic covalent modification (DNA methylation) will be discussed in more detail in the next section.

1.6.2. DNA Methylation Modification

In addition to its well-acknowledged role in gene expression, DNA methylation has an important role in inducing cell apoptosis. Through each DNA replication, the cell accumulates errors in DNA methylation resulting in inactivation of several important genes which eventually induces apoptosis (Mazin 1994, 2009). Further, DNA methylation is thought to have an evolutionary function; it has been reported that the mammalian genome has lost about 3% of 5-methylCytosines (5mC) and 6% of GC pairs in comparison to our ancestors (Mazin 1994). Consequently, cytosine mutation induces spontaneous changes in genome structure and function which influence the genetic and biological diversity among species (Cooper and Gerber-Huber 1985). Moreover, DNA methylation has been shown to have multiple roles in mammals' health and disease such as:

- X-chromosome inactivation:
    - The human male germ cell has only one X chromosome and one Y chromosome, meanwhile, the female germ cell has two X chromosomes. DNA methylation randomly inactivates one of the female X chromosomes, so that, individuals of both genders will have the same amount of X-related gene products. The X-chromosome inactivation process is considered to be a dosage compensation mechanism that starts in early mammalian life (Heard and Disteche 2006, Kaslow and Migeon 1987).

- Genomic imprinting:
    - Genomic imprinting is an epigenetic phenomenon that results in expressing the genes in a paternal-/maternal-origin-specific manner. Normally, only one allele of a gene, either paternal or maternal origin, is expressed. For instance, the gene encoding insulin-like growth factor 2 is only expressed from the allele inherited from the father (Court et al. 2014). Genomic imprinting is controlled by imprinted control regions (ICRs). The ICRs have clustered organisation allowing *cis*-regulation of expressing parental-specific DNA methylation and/or histone modifications resulting in monoallelic expression. A CpG rich ICR which is differentially methylated on one allele (hypermethylated in one allele and hypomethylated in the other) is associated with silencing of the gene of that allele (Li et al. 1993).
- Suppression of transposable element (TE):
    - TEs are DNA sequences that can mobile from one position to another in the genome. About 45% of the human genome consists of transposable elements (Lander et al. 2001). Due to their dynamic transport nature, TE integration could lead to gene inactivation, control gene expression or induce illegitimate recombination (Rebollo et al. 2012). Furthermore, foreign TEs, such as parasitic and retroviral TEs, can integrate into the host genome of other species. Suppression of the expression of the TEs by DNA hypermethylation is shown to be a defensive mechanism (Barlow 1993, Yoder et al. 1997).
- Development and cell/tissue differentiation:
    - DNA methylation (along with histone modifications) define the fate of the cell differentiation at early stages of embryogenesis. DNA Re-methylation (at the blastocyst stage) occurs in a tissue-specific manner allowing the expression of genes vital for the cell type, creating differentially methylated regions (DMR) among different cell types (Cantone and Fisher 2013, Reik et al. 2001).

1.6.2.1. DNA Methyltransferase

The process of DNA methylation is established in early life by a family of DNA methyltransferase enzymes, including Dnmt1, Dnmt2, and Dnmt3 (Figure 1.12). Dnmt2 is the most conserved and the less understood enzyme among the methyltransferase enzymes. Dnmt2 is now known as tRNA aspartic acid methyltransferase 1, which catalyses the methylation of position 38 in tRNA to yield 5-methylcytosine (Goll et al. 2006). Knocking-out *Dnmt2* results in no phenotype in mice (Okano et al. 1998). The Dnmt1 enzyme is mainly involved in methylation of hemimethylated DNA strands during cell division, while Dnmt3 are more involved in the *de novo* DNA methylation during embryogenesis. The process of DNA methylation involves the adding of a methyl group to the fifth carbon atom in the cytosine nucleotide which is donated from a methyl donor molecule (S-adenosylmethionine; SAM) under the effect of DNA methyltransferase enzyme (Figure 1.13).

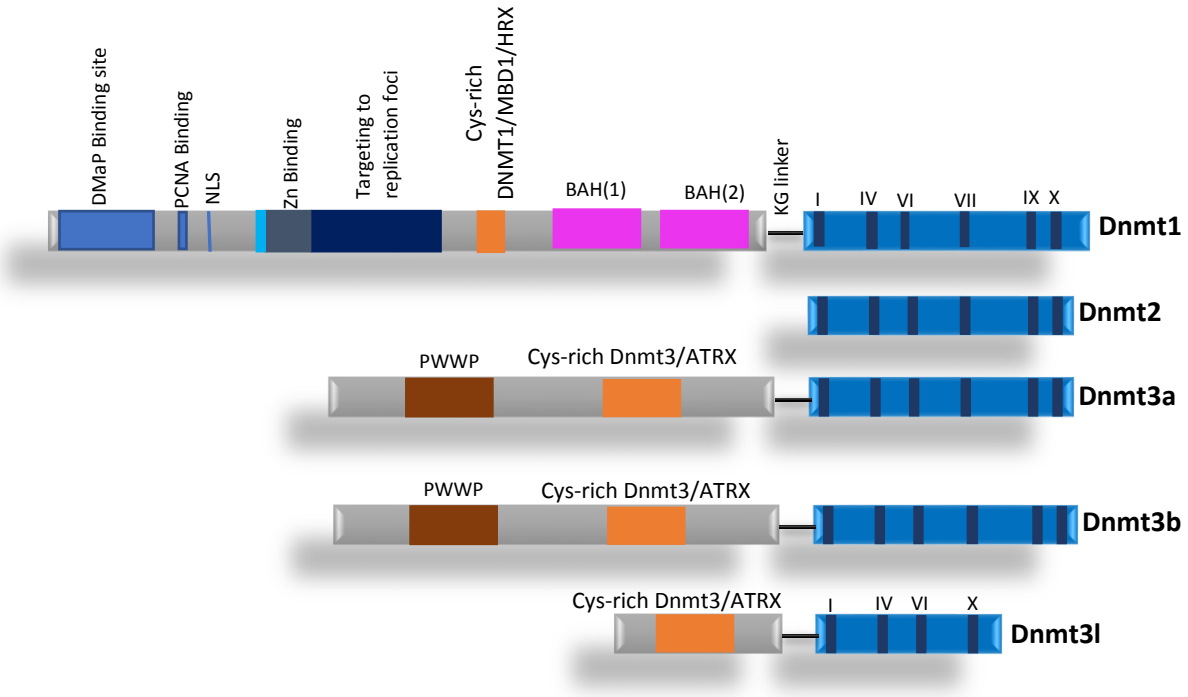


Figure 1.12: The structure of the methyltransferase enzyme family.

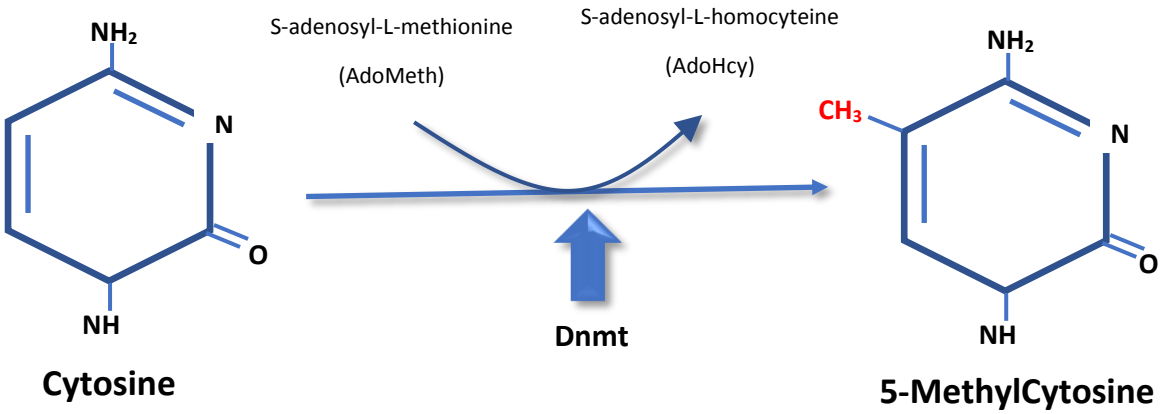


Figure 1.13: The methylation process of DNA.

The reaction is catalysed by DNA methyltransferase (Dnmt) that transfers a methyl group from the donor S-adenosyl-methionine (SAM) to the 5' carbon of an unmethylated Cytosine nucleotide.

1.6.2.2. Methylation Reprogramming

1.6.2.2.1. DNA Demethylation

DNA demethylation occurs by either a passive or an active mechanism. Passive demethylation occurs in cases when the Dnmt1 enzyme fails to methylate the daughter DNA strand during cell division (Hashimoto et al. 2012). Active demethylation has been described during embryogenesis. The Ten-eleven translocation protein (TET) family, including TET1, TET2 and TET3, function as an eraser of cytosine methylation by catalysing the 5mC into intermediate oxidative products such as 5hmC, 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC) (He et al. 2011, Ito et al. 2011, Tahiliani et al. 2009), which eventually are converted to unmodified cytosines as part of the active DNA demethylation pathway (Rasmussen and Helin 2016).

The maternal oocyte and the paternal sperm each have a different genome methylation pattern. Once fertilisation occurs, both genomes are contained in one cell (zygote; Figure 1.14). Most recently, it has been shown that the DNA contained in the zygote undergoes three waves of demethylation in a stepwise manner (Zhu et al. 2018). Within twelve hours post fertilisation, the zygote genome undergoes the first wave of demethylation targeting CpGs located in enhancers and gene bodies. This wave affects the oocyte-derived and sperm-derived pronucleus differently; the oocyte-derived DNA loses about 4% of its genomic methylation (decreased from 54.5% to 50.7%) while the sperm-derived genome loses approximately 29% (decreased from 82.0% to 52.9%). Both the second and the third global demethylation waves target intron regions and short interspersed nuclear elements (SINEs), such as the evolutionarily younger subfamily of Alu elements. The second wave of global demethylation occurs in the stage between the late zygote and the 4-cell stage embryo, decreasing the global genomic methylation levels by approximately 10% (from 49.9% to 40.4%), while the third wave of demethylation occurs between the stage of 8-cell embryo and morula stage, lowering the global methylation levels by about 11% (from 47.0% to 35.1%). At the end of the morula stage, the global genomic methylation level reached the minimum (Zhu et al. 2018). The mechanism of active global genome demethylation is still not fully understood. TET proteins are believed to be

primarily involved. TET3 protein actively demethylates the DNA by oxidising the 5mC into 5hmC, 5fC or 5caC (Gu et al. 2011, Inoue et al. 2011, Inoue and Zhang 2011, Ito et al. 2011). Finally, the intermediate oxidative products are replaced by unmodified cytosines via the base excision repair pathway (He et al. 2011, Zhang et al. 2012). Previously, the maternal genome was thought to undergo global demethylation differently to the paternal genome demethylation. The Maternal genome was thought to be globally demethylated through gradual passive demethylation; however, recent studies have shown both maternal and paternal demethylation undergo similar mechanism of embryogenesis demethylation (Figure 1.14) (Guo et al. 2014, Shen et al. 2014, Zhu et al. 2018).

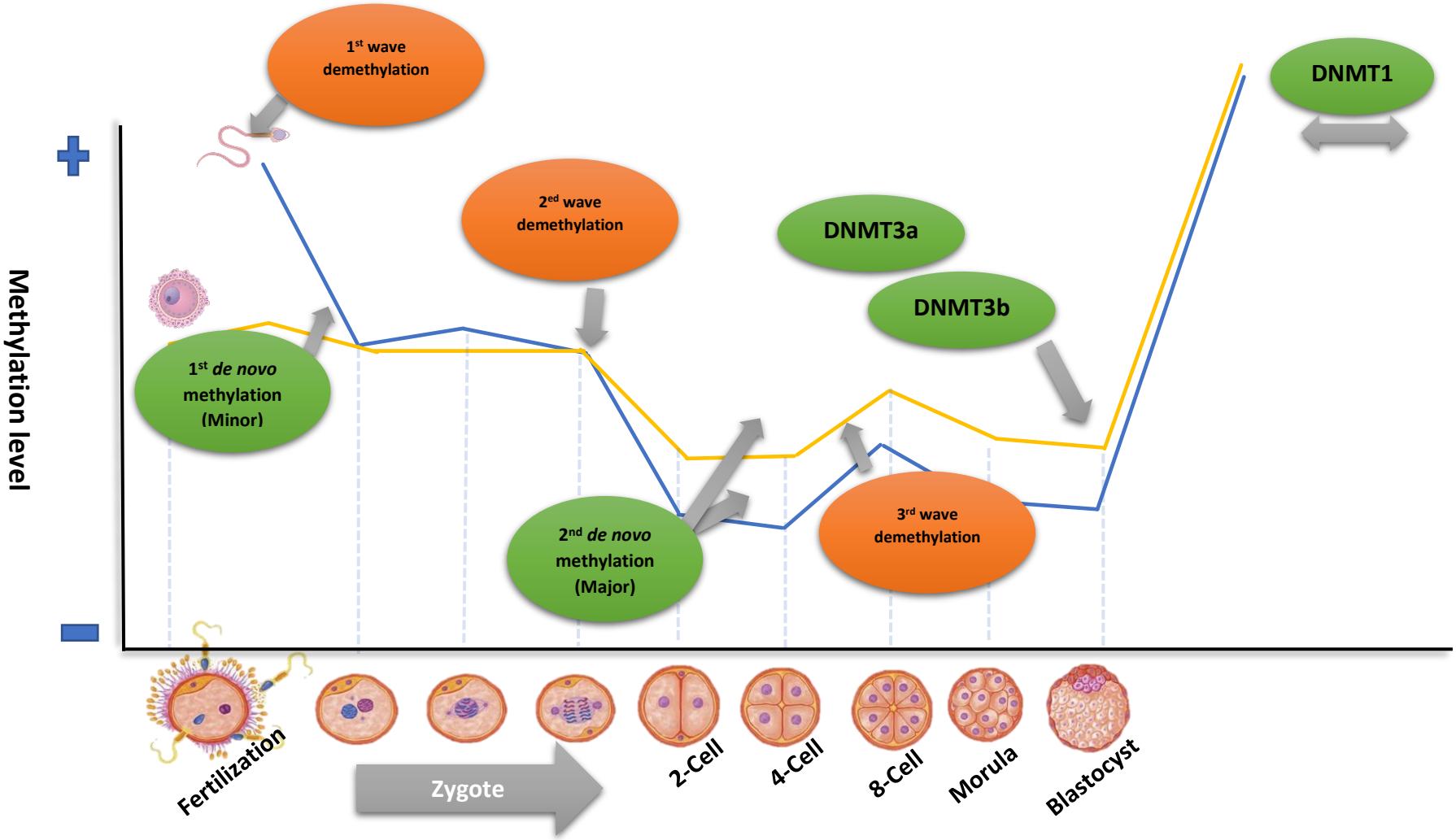


Figure 1.14: The methylation cycle in a somatic cell.

1.6.2.2.2. *De novo* DNA Methylation and the Maintenance of DNA Methylation

During human embryogenesis, two waves of *de novo* methylation occur, a major and minor wave. Both methylation waves target the major families of repeat elements, such as SINEs, long interspersed nuclear elements (LINEs), and long terminal repeats (LTRs). Methylation waves also target evolutionarily younger subfamilies, such as Alu and LINE-1 (L1) retroelements. The first *de novo* methylation wave is minor but significant, occurs in the spermatozoa genome between the early to the mid-pronuclei stage of the zygote. The second *de novo* methylation wave is major and occurs in the 4-cell to 8-cell stage of the embryo (Zhu et al. 2018) (Figure 1.14). Interestingly, a similar *de novo* DNA methylation wave has been reported to occur in the 2-cell to 8-cell embryo progression in monkeys (Gao et al. 2017).

The cycles of demethylation and *de novo* methylation that occur in the embryogenesis demonstrate the nature of DNA methylation reprogramming. It shows that the methylation process is dynamic and transient, as the strong genome-wide demethylation is balanced by targeted DNA methylation (Figure 1.14). The DNA demethylation and methylation cycles also demonstrate that the process may occur as the gene regulation requires (Gao et al. 2017, Zhu et al. 2018). The process of *de novo* DNA methylation is accomplished by the effect of Dnmt3a and Dnmt3b enzymes; both enzymes have been shown to be the most highly expressed enzymes in undifferentiated human embryonic stem cells (Liao et al. 2015). It has also been shown that Dnmt3a methylate the cytosine nucleotides at a faster rate in comparison to Dnmt3b, but slightly slower than Dnmt1. Dnmt3a establishes differential DNA methylation at the ICRs in gametes (Kaneda et al. 2004, Kato et al. 2007, Nishikawa et al. 2015). Combined or individual deletion of human *Dnmt3b* or *Dnmt3a* were shown to cause embryonic lethality (Liao et al. 2015, Okano et al. 1999).

Dnmt1 enzyme carries out the process of DNA methylation maintenance during the cell division and DNA replication. The process occurs passively or actively in a symmetrical manner where unmodified cytosine nucleotides on both DNA strands are converted into 5-methylcytosines. During

the synthesis phase of the DNA replication, the SET and RING-associated (SRA) domain of the protein ubiquitin-like containing PHD and RING finger domains 1 (UHRF1), identifies and binds to newly synthesised hemimethylated DNA strand (Jones and Liang 2009). SRA-DNA interaction induces Dnmt1 recruitment, which copies the DNA methylation pattern of the parental DNA strand to the daughter strand (Frauer et al. 2011, Jones and Liang 2009). Furthermore, the Dnmt3a and Dnmt3b (responsible for *de novo* DNA methyltransferases) have also shown to be associated with the DNA methylation maintenance (Chen et al. 2003, Hashimoto et al. 2012, Ji et al. 2014, Jones and Liang 2009).

The Dnmt3-like enzyme, known as Dnmt3l, is mainly expressed during the development (Aapola et al. 2000). Although Dnmt3l has been shown to have no catalytic activity itself, its expression is crucial for gene imprinting. The genetic deletion of *Dnmt3l* results in the death of the foetus and the co-expression of paternal genes in mice (Bourc'his and Bestor 2004, Bourc'his et al. 2001). Dnmt3l also functions as a cofactor for Dnmt3a and Dnmt3B to bind to their catalytic domain and thereby accelerating their ability to attach to SAM and therefore stimulating their ability to methylate DNA in the methylation *de novo* pathway (Chedin et al. 2002, Gowher et al. 2005, Jia et al. 2007). Another Dnmt3 isoform (Dnmt3c) was shown to be responsible for the methylation of evolutionarily young retrotransposon promoters in the male germ line which is required for mouse fertility (Barau et al. 2016).

1.6.3. The DNA Methylome of the Pancreatic Beta Cell

The methylation pattern of pancreatic beta cell specific genes e.g. insulin offers a classic example of the role of a gene promoter methylome in determining cell identity. This was reported for the first time in a study by Gilbert and colleagues in 1983 on the rat genome (Cate et al. 1983). They demonstrated that the promoters of the rat *insulin-I* and *insulin-II* genes in insulin expressing cell lines are differentially hypomethylated when compared with non-insulin expressing cells. Later, insertion of a methylated and mock-methylated reporter in the promoter of *INS* gene constructs in the insulin-producing cell line (NIT-1) results in insulin gene suppression (Kuroda et al. 2009). A similar conclusion was reported by another group showing that DNA hypermethylation of the *INS* promoter (CpG -234, -

180, -102 and +63) correlates negatively with an expression in the *INS* gene in human pancreatic islets (Yang et al. 2011). Collectively, these observations suggest that differential hypomethylation of the *INS* gene promoter is a unique feature of beta cell identity.

1.7. Biomarkers of Predicting T1D

1.7.1. General Introduction

A biomarker is an indicator that objectively measures a normal or pathological body state as well as evaluating biological responses to a therapeutic intervention (Biomarkers Definitions Working Group 2001). Ideally, a biomarker should have some specific criteria to be of clinical value; it should be highly selective and specific to a target, have a detectable threshold in the circulation and be reproducible. Above all, a biomarker should be non-invasive, acceptable to the patient and easily interpreted by the clinician (LaBaer 2005). Biomarkers could be powerful tools for monitoring health/disease states and predicting future changes in health due to treatments or environmental alteration. Different biomarkers have different functions, and accordingly, they are classified into different categories according to their proposed application. Thus, the adoption of unified biomarker classification will provide a more natural way to describe a biomarker and define its function. Many attempts had been made to categorise the biomarkers into different subgroups mainly according to their clinical application or according to their physiological type (genomic, proteomic, or their physiological products). Categorising the biomarker based on its clinical function is the most common way for subgrouping it (Table 1.3). Nevertheless, this approach to subgrouping has a primary disadvantage because a biomarker could belong to more than one subgroup, e.g. diagnostic and staging biomarkers. Hence, adopting a more generalised classification such as the one suggested by Mayeux would be more convenient (Mayeux 2004). He divided biomarkers into two major types: biomarkers of exposure (predicting the risk of development of a disease and assessment of a drug effect) and biomarkers of disease (utilised to screen, diagnose and monitor disease progression).

1.7.2. Discovering a Novel Biomarker

The first step in the process of developing a novel biomarker is defining its clinical purpose (Table 1.3). Then, a biomarker should ensure to meet the clinical qualities (Table 1.4). An acceptable biomarker should be sensitive and specific for the targeted disease/tissue, applicable to different populations, have a series of cut-off values in the general population and is detectable in the early

stage of the disease. Extra factors should be considered for diagnostic and therapeutic biomarkers: the expense, the difficulty of sample collection and patient disturbance during sample collection. For example, a biomarker destined to be used in routine screening tests should have the advantage of both easy access and minimum cost per sample. On the other hand, diagnostic and research biomarkers should be highly specific for the targeted specimen such as a tissue biopsy. In that case, the number of the processed samples will be limited, but the specificity and the sensitivity will be at their maximum (LaBaer 2005).

Table 1.3: Biomarker types based on their clinical applications.

Type of Biomarker	The Purpose
Antecedent	Identifying the risk of developing an illness
Screening	Screening for the presence of suspected subclinical disease
Monitoring	Monitoring functions of organs/tissue states
Diagnostic	Confirm the presence of an illness
Staging	Tracking disease severity
Endpoint Analysis	Monitoring the efficacy of a therapeutic agent

In a disease state, where there is a high rate of specific cell type death, the contents of the dying cells are released into the body's circulation. The contents of these cells, which include proteins, lipids, nucleic acids or gene products, can provide important biomarker readouts specific to those cells. Therefore, various assay techniques are required to detect the changes in these biological parameters. Genomic and epigenomic biomarkers are an advanced approach that has been widely used in cancer research recently (Boutros 2015, Leygo et al. 2017, Voyias et al. 2016). The advantage of the epigenomic biomarkers over conventional techniques is their ability to predict diseases in preclinical stages as well as accurate identification of drug target and treatment response (Heyn and Esteller 2012, Simon 2011).

Table 1.4: Features and characterization of the ideal biomarker.

Characterization	Comments
Specific	Have the ability to differentiate pathological from normal state
	Specific to the affected cells or tissue
Sensitive	Rapid and significant release upon disease development
Anticipative	Long half-life in biological sample
	Its release is proportional to the degree of disease severity
Robust	Rapid, simple, accurate and inexpensive detection
	Not changed by environmental factors and other diseases
Non-invasive	Present in easily accessible biological fluid samples
	Minimum risk to the patient at the time of sample collection

1.7.3. Prediction and Diagnostic Biomarkers for Overt Diabetes

In T1D, the ideal biomarker should have the capability to detect early insulitis, predict diabetes development in "at risk" individuals, predict therapy outcomes by measuring beta cell death and mass (Tooley and Herold 2014). Our current capability for direct quantification of insulin secretion, beta cell mass or death is limited due to the difficult anatomical position of the islet in the pancreas as well as their low abundance (only 1-2% of the whole pancreas mass) (Ionescu-Tirgoviste et al. 2015). It is not considered ethical to acquire a biopsy sample and applying imaging techniques to study them *in vivo* directly.

1.7.3.1. Autoantibodies to Islet Antigens as Predictive Biomarkers for T1D Development:

Islet autoantibodies are highly stable biomarkers, and they can be measured in the serum, plasma or saliva. In the last twenty years, a considerable effort was made by the Diabetes Autoantibody Standardization Program (DASP) to identify the most sensitive and specific assays, determine the ideal antigens and reference standards, and underline the weakness of different methods (Bingley et al. 2003; Lampasona et al. 2011). These efforts have contributed to standardising islet autoantibody measurement, prediction and diagnostic assays (Table 1.5). By the time of disease

onset, 60-80% of T1D patients are positive for GADA, 40-50% are positive for IA-2A, 75%–85% are positive for ICA. It is also reported that 70-80% of the patients who were diagnosed before the age of 5 years with T1D are positive for IAA, but the percentage drops up to <40% when the diabetes is diagnosed after the age of 12 years (Sacks et al. 2011, Ziegler et al. 1999). It was shown that 60-80% of the newly diagnosed with T1D patients were positive for ZnT8A (Wenzlau et al. 2007). Although, IAA are more commonly detected (>70%) in the sera of young children in comparison to adult samples, screening for IAA becomes invalid once exogenous insulin treatment is started. This is because exogenous insulin might induce antibody secretion against it, making differentiating newly developed antibodies from pre-existing IAA difficult (Winter and Schatz 2011). IAA is usually the first autoantibody to be detected, followed by GADA, while both IA-2A and ZnT8A, are considered as "late" autoantibodies that are rarely detected at the time of initial seroconversion. The peak of islet antigen-specific autoantibody seroconversion has been reported to occur between the age of 9 months and two years (Giannopoulou et al. 2015, Ziegler et al. 1999). Islet antibody seroconversion can be detected in sera as early as six months of age, and it is rare to occur before that age (Krischer et al. 2015, Parikka et al. 2012, Ziegler et al. 2012). The detection of autoantibodies is useful as a predictive biomarker for T1D progression as seroconversion to multi antibodies increases the lifetime risk of developing stage 3 T1D to 100% in children (Ziegler et al. 2013a). The presence of one or more islet autoantibodies in the circulation has been shown to be a very accurate predictor for T1D. Although different studies have given different risk results, screening for GAD, IA-2 and/or insulin antibodies using sensitive radio-binding assays can identify 85% of cases of newly diagnosed or future T1D with 98% specificity (Bingley et al. 2001). The presence of multiple islet autoantibodies in the sera of high-risk children increases the risk for developing T1D by 43.5%, 69.7% and 84.2% in 5, 10 and 15 years of follow-up, respectively (Steck et al. 2015, Ziegler et al. 2013a). Remarkably, the number of detected autoantibodies is associated with T1D risk. In high-risk infants, the detection of one, two or three autoantibodies increases the risk of development to stage 3 T1D to 12.7%, 61.6% and 79.1%, respectively, by the age of 15 years (Ziegler et al. 2013a). The TEDDY study concluded that the risk of

developing symptomatic T1D is elevated by 11%, 36% 47% in individuals who have one, two and three autoantibodies, respectively, by the age of 5 years (Steck et al. 2015). Seroconversion at a young age is another risk factor for stage 3 T1D progression. Development of multiple autoantibodies at the age 1-2 years increases the rate of progression to symptomatic T1D by 75% within six years of life (Steck et al. 2015; Ziegler et al. 2013). Further, children who had seroconversion before the age of 3 years had a risk of progression to symptomatic T1D of 74.9%, while those who had seroconversion after the age of 3 years had a risk of 60.9% for progression to symptomatic T1D by the age of 10 years (Ziegler et al. 2013). It has been well acknowledged that there is an association between specific *HLA* alleles and the development of islet autoantibodies. For example, GADA is strongly associated with the *HLA DR3-DQ2* allele, while IAA and IA-2A are more related to the *HLA DR4-DQ8* allele and *INS* VNTR (Giannopoulou et al. 2015, Graham et al. 2002). Islet autoantibody development is also controlled by other polymorphisms in the *PTPN22*, *ERBB3*, *SH2B3*, and *INS* genes according to the TEDDY study (Table 1.5) (Torn et al. 2015).

Table 1.5: The sensitivity and the specificity of the primary antibodies and targets used as a predicting biomarker of T1D.

Assay	Age of Onset	% of Newly Diagnosed	Genetic Association	Sensitivity	Specificity	κ (95% CI) of the Assay	Comment
GAD65A	Various sex and age	60%–80%[6]	HLA DR3–DQA1*0501–DQB1*0201[9]	88%[6]	98%[6]		
				74%[7]	85%[7]		
				83%[8]	92%[8]		
IA-2A	Young	40%–50%[9]	HLA DR4–DQA1*0301–DQB1*0302[9]	70%[6]	99%[6]		
				52%[7]	97%[7]		
				68%[8]	98%[8]		
IAA	Children < 5years	>70%–80%[9]	HLA DR4–DQA1*0301–DQB1*0302 & *INS* VNTR4			0.7 (0.58-0.82) [8]	occurs in <40% of individuals who develop diabetes after the age of 12 years
ICA		75%–85%[9]		85%[8]	99%[8]		Difficult to Standardize
ZnT8A		60-80%					Found in 26% of patients who are negative for GAD65A, IA-2A and IAA
Combined GAD/IA2	N/A			96%[6]	98%[6]		The presence of three or more autoantibodies increases the risk.
				48%[7]	97%[7]		
				83%[8]	92%[8]		
T cell proliferation				58%[7]	91%[7]	0.52 (0.30-0.73) [8]	Require fresh cells, antigens can be grouped.
				60%[8]	69%[8]		
Cellular Immunoblot				94%[7]	83%[7]	0.63 (0.41-0.86) [8]	Require fresh cells, large volumes, and multiple antigens.
				74%[8]	88%[8]		

[6] (Torn et al. 2015)
[7] (Seyfert-Margolis et al. 2006)
[8] (Herold et al. 2009)
[9] (Sacks et al. 2011)

1.7.3.2. Islet Autoantibodies and C-Peptide in the Assessment of T1D Clinical Therapies

Pancreatic islet transplantation is a conventional T1D therapy used to reinstate insulin secretion and production in individuals with ongoing issues with controlling blood glucose level. Islet transplantation is usually rejected due to autoimmunity relapses even after the administration of immunosuppressive drugs. In combination with the indices of glucose homeostasis and C-peptide levels, islet autoantibodies are used to predict pancreatic allograph rejection. They were found to reappear, or their titres increase when the grafted islets are rejected (Bosi et al. 2001, Braghi et al. 2000, Roep et al. 1999b). Although GADA were the first autoantibodies to reappear, screening for IA-2A and ZnT8A are found to be the more robust islet transplant failure predictors (Piemonti et al. 2013).

Several reports had shown the shortcomings of using islet autoantibodies as a biomarker for predicting and assessing the T1D Therapies. First of all, anti-islet autoantibodies are not the primary effectors of insulitis and beta cell destruction; instead, they are an indirect reflection of the process. Therefore, autoantibody titres post immunotherapy are probably not an indication of immune efficacy (Rekers et al. 2015). Secondly, as explained earlier, T1D is a T cell-mediated disorder so naturally, most therapeutic strategies target T cells. Meanwhile, autoantibodies can be secreted independently by long-lived plasma cells that are not affected by immunosuppressive therapies, and from new plasma cells that can be further differentiated from memory B cells under the stimulation of persistent or recurrent T cells. Furthermore, T cells have a faster dynamic response to external modifications than changes in autoantibody titres. This has been shown in new-onset T1D patients followed up for both autoantibodies and T cell responses from diagnosis for a median of 11 months. It has been demonstrated that administration of pre-proinsulin and GAD epitopes decreased the number of reactive T cell responses (from 60 to 20% for pre-proinsulin and 67 to 20% for GAD) meanwhile, the titre of GAD and IA-2 autoantibodies remained unchanged in 75% of cases (Martinuzzi et al. 2008). Furthermore, anti-CD3 immunotherapy (which targets reactive T cells and partially depletes B cells) in new onset T1D patients demonstrated better preservation of residual insulin secretion but did not show a significant change in autoantibody titres (Demeester et al. 2015, Keymeulen et al. 2005,

Madaschi et al. 2010). Unchanged autoantibody levels can be solely because anti-CD3 immunotherapy targets T cells but not B cells. Another example is the immunotherapy trial on IAA positive T1D patients with anti-CD20 rituximab, which targets B cells resulting in the elimination of IAA in only 40% of treated group patients in comparison to the placebo-treated group (Pescovitz et al. 2014, Pescovitz et al. 2009, Yu et al. 2011). Finally, islet autoantibody measurements are not valid in a non-autoimmune condition that may need beta cell grafting such as type 2 diabetes.

C-peptide is another robust biomarker that reflects beta cell function. C-peptide is the part of proinsulin which is cleaved and co-secreted with insulin from the secretory granules of the beta cell (Figure 1.3). Following proinsulin cleavage, insulin (51-amino-acid peptide) and the C-peptide (31-amino-acid peptide) are produced in a 1:1 ratio (Marques et al. 2004, Steiner et al. 1967). The measurement of C-peptide levels is a useful indicator to distinguish between the different types of diabetes, e.g. type 1 diabetes from type 2 diabetes or Maturity onset diabetes of the young (MODY) as a very low C-peptide confirms type 1 diabetes (Jones and Hattersley 2013). Measuring the secretion levels of C-peptide in type 1 diabetes can assist in estimating the secretion of endogenous insulin especially when the patient receives insulin therapy and when measuring endogenous insulin levels becomes difficult (Clark 1999, Shapiro et al. 1988). Like islet autoantibodies, C-peptide measurement showed several limitations as an indicator for the outcomes of clinical trails. Although it is not affected by exogenous insulin, C-peptide is usually influenced by factors that have an impact on insulin secretion, such as insulin resistance (Saisho 2016). Further, the level of C-peptide essentially reflects the remaining beta cell function rather than measuring the dying beta cells. Further, C-peptide often fails to reflect the presence of beta cell destruction during the pre-diabetic stage, particularly at the very early period of T1D development and in latent autoimmune diabetes in adults (LADA) when the beta cell injury is slow. Moreover, C-peptide metabolic clearance rate may fluctuate from one individual to another, and due to the long half-life of C-peptide, false positive results could be caused by short time intervals of the tests (Polonsky et al. 1986a, Polonsky et al. 1986b, Polonsky and Rubenstein 1986). Therefore, the lack of current real-time biomarkers to measure the efficacy of

preventative and interventional therapeutic approaches still poses an immense challenge. Recently, evidence has accumulated showing that epigenomic biomarkers such as DNA methylation of key beta cell genes have the potential to be powerful tools to detect beta cell injury in the early preclinical stage of T1D.

1.7.3.3. Differentially Methylated Circulating DNA as an Indicator of Beta Cell Death

During the pathogenesis of a disease, injured cells release their DNA into the circulation. Isolating and quantifying such cell-free DNA (cfDNA) has been used recently to study the rate of cell turnover and/or cell apoptosis/necrosis (Schwarzenbach et al. 2011). cfDNA is fragmented gDNA which originates from multiple possible sources such as normal cell turnover, the breakdown of blood cells, leukocyte surface DNA, newly synthesised DNA or from released DNA/RNA-lipoprotein complexes from healthy cells. When the size of the cell-free nucleic acids was measured with gel electrophoresis, the result was a ladder-like pattern where the nucleic acid fragments were mainly around 180-1000 bp in size. Massively parallel sequencing is a more advanced technology with higher resolution that was used recently to confirm the size of those cell-free nucleic acids. It showed that the majority of cfDNA fragments are around 166 bp which suggests that apoptosis is the major source of cfDNA in the circulation (Jiang and Lo 2016). In healthy individuals, the mean quantity of plasma cfDNA is from 10 ng/µL - 1.5 µg/ µL (Elshimali et al. 2013, Fleischhacker and Schmidt 2007). This low concentration of cfDNA is due to the efficient removal of dead cells by circulating phagocytes. It has been shown that in the case of diseases such as cancer and developing autoimmunity, the level of cfDNA is dramatically increased due to the rapid destruction of cells (Jiang and Lo 2016). One of the approaches to identify such circulating cfDNA is through detecting differentially methylated sites specific to the injured cell. This approach is particularly promising in biomarker development research, and it has been applied successfully in cancer research (Han et al. 2017, Wang et al. 2017).

Beta cells have uniquely hypomethylated CpG sites within the promoter of the *INS* gene which is found to be predominately hypermethylated elsewhere (Akirav et al. 2011, Kuroda et al. 2009,

Lehmann-Werman et al. 2016). Targeting differentially hypomethylated CpG sites within *INS* seems to be a promising strategy for monitoring beta cell death. Multiple groups have made an effort to develop a unique DMR-based biomarker for monitoring beta cell injury when required. Ferreri and colleagues reported that specific CpG sites within the insulin gene promoter are uniquely hypomethylated in both human (*INS*) and mouse (*Ins2*) beta cells (Husseiny et al. 2012, Kuroda et al. 2009). They reported that CpG sites (−357, −345, −234, −206, −180, −135, −102, −69 and −19 relative to the transcription start site, TSS) were uniquely hypomethylated in the human beta cell *INS* promoter but not in other tested control tissues, such as liver and kidney. Further analysis had shown that only CpG -19, -69, -135, -206 and -357 from TSS were significantly hypomethylated in beta cells but not elsewhere (Husseiny et al. 2014). They further showed that hypermethylation induction in specific CpG sites resulted in significantly decreasing insulin gene expression in the NOD mouse model. They were also able to reprogram mouse embryonic stem cells to differentiate into insulin-producing cells by demethylating specific CpG sites in the *Ins2* gene promoter *in vitro* (Kuroda et al. 2009). A differentially methylated assay whose design was based on the CpG sites -206 and-135 was able to successfully detect hypomethylated *INS* beta cell-derived DNA in the circulation of human islet transplant recipients (Husseiny et al. 2014). Another differentially methylated assay designed by another group based on the CpG -69 was successful in differentiating beta cell-derived DNA in the circulation of recently diagnosed T1D patients as well as in the circulation of a NOD mouse model which underwent xeno-islet transplantation rejection (Fisher et al. 2015).

Another group, Herold and colleagues (Akirav et al. 2011) were also able to confirm the presence of two differentially hypomethylated CpG sites located outside the promoter region of human *INS* gene (CpG +273 and +399 relatives to TSS). Herold and his group suggested that CpG +273 and +399 are potentially better targets for discriminating beta cell-derived DNA from non-beta cell DNA. Based on which, a beta cell discriminating assay was designed and used to monitor beta cell death in recent onset and at risk T1D individuals (Akirav et al. 2011, Herold et al. 2015a, Usmani-Brown et al. 2014), post-immune therapy (Lebastchi et al. 2013) and post islet replacement therapy (Bellin et al.

2017). Despite the success of the assay in detecting beta cell-derived DNA, a considerable background level of hypomethylated *INS* DNA was detected in control samples (Bellin et al. 2017, Herold et al. 2015a). The authors referred to the source of hypomethylated *INS* background levels as of non-beta cell origin, since the same CpG site was previously reported to be hypomethylated in non-beta cells (Husseiny et al. 2014). The sensitivity and specificity of the detection assay were shown to be increased when multiple adjacent CpGs were multiplexed together (Lehmann-Werman et al. 2016). High-throughput methylation sequencing of six individual CpG sites within the promoter of the *INS* gene in multiple tissues revealed that these CpG sites are 90-95% hypomethylated in beta cell genome and 5-15% in other studied tissues. The specificity of these CpG sites to discriminate beta cell from non-beta cell DNA was increased by combining the six sites (multiplexing), decreasing the detection of hypomethylated CpG sites in the non-beta cell genome to 0.01% (Lehmann-Werman et al. 2016). The power of multiplexing was also manifested in an experiment which showed that the analysis of the concentration of individual hypomethylated CpG sites of the *INS* promoter in DNA fragments failed to differentiate between T1D and control serum samples. However, when the result was multiplexed, a clear signal discriminating T1D from controls was observed (beta cell-derived cfDNA level was 1.9-5.5% in T1D patients and 0-0.1% in control)(Lehmann-Werman et al. 2016).

The pancreatic beta cell is also known to expresses multiple genes almost exclusively, which hence could be regulated by a unique methylome. Therefore, other beta cell-specific genes might be used to design methylation sensitive assays. For instance, specific differentially hypomethylated CpG sites within the islet-specific gene amyloid polypeptide (amylin) had shown the ability to discriminate beta cell-derived cfDNA in the circulation of both humans and a mouse model (Olsen et al. 2016). The study demonstrated that a methylation sensitive assay design based on CpG +5414 and +5419 successfully discriminated beta cell-derived DNA in the circulation of recently onset T1D individuals (Olsen et al. 2016). Another example is the use of differentially methylated CpG sites in exon 1 of the gene glucokinase (*GCK*), which was shown to be predominantly hypomethylated in the islet genome but not elsewhere (Sklenarova et al. 2017). The study aimed to compare the efficiency of methylation

specific biomarkers based on the CpGs +13, +17 and +19 of exon 1 of the gene *GCK* and the biomarker based on the CpG -69 of *INS* promoter designed by (Fisher et al. 2015). Although both assays were not statistically significant in differentiating between recently diagnosed T1D patients or those who are at risk of developing T1D from non-diabetic matched controls (Sklenarova et al. 2017), *GCK* and amylin based biomarkers were a step forward toward employing beta cell-specific genes, apart from *INS*, for designing differentially methylated biomarkers that detect beta cell death in the circulation.

1.7.3.4. DNA Methylation in Assessing the Outcomes of Clinical Practice

A differentially methylated based biomarker has been used for assessing beta cell survival following innovative therapies. It has proved to be of benefit in the evaluation of beta cell responses to the anti-CD3 immune therapy (teplizumab) in recent-onset T1D patients (Lebastchi et al. 2013). Differentially hypomethylated *INS* based assays were also shown to be of benefit in assessing the outcomes of allogeneic islet transplantation in humans. Following up islet transplant patients (*n*=6) showed that hypomethylated *INS* levels significantly increased a day after transplantation, the elevated level persisted for at least fourteen days (Husseiny et al. 2014). The hypomethylated *INS* promoter CpG was also used to interpret the outcomes of cadaveric allogeneic islet transplantation in longstanding T1D patients (*n*=10) (Lehmann-Werman et al. 2016). Hypomethylated *INS* promoter cfDNA was detected in the plasma of islet recipients 1-2 hours post the transplantation; the level declined sharply in the following hours and days in a manner that was in parallel with the result of imaging studies done on islet transplantation recipients (Eich et al. 2007). Increased levels of hypomethylated *INS* promoter fragments were re-detected in the circulation of the recipients 7-days to 1-month post the transplantation, indicating ongoing cell injury (Lehmann-Werman et al. 2016). More recently, it was possible to correlate between beta cell-specific cfDNA post clinical islet allograft and recipient outcomes (Gala-Lopez et al. 2018). Another example of the visibility of using differentially hypomethylated circulating *INS* DNA fragments to assess the outcomes of a clinical trial was a study of the *INS* cfDNA level in patients undergoing pancreatectomy with islet autotransplantation (TPIAT). A study done by the Herold group (Bellin et al. 2017) showed that the

demethylation *INS* ratio (demethylated *INS* counts / demethylated INS counts + methylated *INS* counts (U/[M+U])) was significantly raised in the first three hours post islet infusion in TPIAT patients. Variable late elevation (between the day 3 and 30 post islet infusion) in the methylated *INS* cfDNA ratio was noticed, suggesting inconsistency in graft loss among recipients (Bellin et al. 2017). Collectively, these findings suggest that differentially methylated circulating DNA is a potentially powerful and highly specific biomarker for monitoring beta cell death during and after islet graft. Despite the number of described assays that can discriminate T1D from controls, none are used clinically and workshops have not so far reported cfDNA biomarkers that are proven to be effective.

1.8. DNA Methylation Profiling Technologies and Studies

1.8.1. DNA Methylome Profiling

Next-generation methylation sequencing (NGMS) technologies are non-Sanger-based, massively parallel high-throughput DNA sequencing technologies. Genome-wide methylation testing permits the analysis of global epigenetic changes and therefore the opportunity to assess the whole genome to identify more gene-specific DMRs (Beck and Rakyan 2008, Bock et al. 2010). There are several different experimental technologies for profiling genome-wide DNA methylation, summarised in Table 1.6. Microarrays are one of the widely used techniques that allow studying genome-wide epigenetic modifications at a molecular-level among healthy and disease samples. Infinium Human Methylation 450 BeadChip and the more recently Illumina MethylationEPIC BeadChip are a method of choice for epigenome-wide association studies (Rakyan et al. 2011b). The EPIC microarray provides an assessment for more than 850,000 CpG loci, covering key features of the human genome including enhancers, CpG islands, shores, promoters, gene bodies, intergenic, and imprinted regions (Moran et al. 2016).

Large volumes of data have been generated from microarrays, the first genome-wide, single-base resolution sequencing of the mammalian genome was published in 2009 (Lister et al. 2009). Large international initiatives are contributing to the mapping of the human epigenome as the International Human Epigenome Consortium (IHEC) (Bujold et al. 2016; Stunnenberg et al. 2016). IHEC coordinates the contribution from seven international consortia; ENCODE (Consortium 2012), NIH Roadmap (Bernstein et al. 2010, Kundaje et al. 2015), CEEHRC, BLUEPRINT (Adams et al. 2012), DEEP, AMED-CREST, and KNIH; that aims to sequence and decipher 1000 human epigenomes of various cell types. The Roadmap Epigenomics consortium recently published a landmark paper of post-translational modification profiles for 111 reference human epigenomes such as histone modification, DNA accessibility, DNA methylation and RNA expression by using high-throughput sequencing techniques (Kundaje et al. 2015). Other key studies include, the DNA methylation profiling of 17 human somatic tissues (Lokk et al. 2014); in-depth analysis of 42 whole-genome bisulphite sequencing datasets across

30 diverse human cell and tissue types (Ziller et al. 2013), analysis of the methylation signature of 1628 human samples (Fernandez et al. 2012) and profiling methylation of 82 human cell lines (Varley et al. 2013).

1.8.2. DNA Methylation-Sensitive Assay Design

The first designed generation of methylation-sensitive assays used SYBR Green (Akirav et al. 2011, Husseiny et al. 2012). The SYBR Green-based technology had some limitations, firstly, SYBR green technology has the potential for high background signals, off-target annealing of the methylation-sensitive primers which could decrease the assay sensitivity and the inability to multiplex multiple targets in one reaction. In 2013, Mirmira and colleagues introduced a dual-labelled probe-based methodology (TaqMan) in a semi-quantitative manner, which allowed multiplexing the methylated and unmethylated CpG -182 (relative to TSS) of the *Ins2* DNA probe in one reaction (Fisher et al. 2013). The proposed method obviated the need for the nested PCR step and significantly decreased the background signal. The dual-labelled probe-based methodology was the basis of the development of the second generation of absolute quantitative assays using the droplet digital PCR (ddPCR). ddPCR is a technique that involves a water-oil emulsion droplet (microfluidics) technology that partitions the sample into 20,000 droplets. The fractionated samples are then amplified by thermal cycling, and then Poisson statistics are used to calculate the absolute copy numbers of DNA fragments (Bio-Rad droplet digital application guide). Utilising ddPCR has led to the development of two insulin based methylation-sensitive assays to monitor beta cell death (Fisher et al. 2015, Usmani-Brown et al. 2014).

1.8.3. The Effect of Diabetes on the Methylome of Beta Cell-Specific Genes

T1D-specific methylation has been reported to be present at variable positions in the DNA of monocytes isolated from MZ and discordant healthy controls. T1D MZ and pre-T1D individuals who were positive for the diabetes-associated autoantibodies (GAD65, IA-2 and Islet Cell Antibodies), had healthy blood glucose levels and were not receiving insulin treatment showed a higher level of DNA methylation within the *INS* promoter. This indicates that changes in DNA methylation are due to T1D

progression and is not due to post-disease associated factors such as hyperglycaemia (Rakyan et al. 2011a). Furthermore, islets isolated from subjects with type 2 diabetes displayed higher methylation levels in the *INS* gene in comparison to islets isolated from control subjects (Yang et al. 2011). Additionally, the CpG hypermethylation at defined positions in the *INS* proximal promoter (-69, -102, -108, -206) were strongly correlated with the *INS* cis-genotype at the rs689 polymorphism, which is known to be strongly associated with T1D (Fradin et al. 2012). More recently, a report confirmed the effect of hyperglycaemia on the methylome of the beta cell in the NOD model (Rui et al. 2016). The presented data showed that *Ins1/Ins2* gene was becoming hypermethylated as T1D progressed in the NOD mouse model. Therefore, studying the effect of hyperglycaemia on the methylome of a CpG is highly recommended when a methylation sensitive assay is designed.

Table 1.6: Different DNA methylation-profiling technologies.

CHARM: Comprehensive high-throughput array for relative methylation. MBDCap-Seq: Methyl-CpG binding domain-based capture and sequencing. MeDIP: Methylated DNA immunoprecipitation sequencing. WGBS: Whole-genome bisulfite sequencing. RRBS: Reduced-representation bisulfite sequencing. TAB-Seq: Tet-assisted bisulfite sequencing. The information was based on (Plongthongkum et al. 2014).

(1) (Irizarry et al. 2009); (2) (Rauch et al. 2008); (3) (Weber et al. 2005); (4) (Cardenas et al. 2016); (5) (Lister et al. 2009); (6) (E. J. Lee et al. 2015); (7) (Zhang et al. 2016); (8) (Guo et al. 2015)

	Description	Coverage	DNA input	Advantages	Disadvantages	Resolution	Cost	Examples
CHARM	Array-based technique for methylation profiling using *Mcr*BC	19% of genomic CG	2-10 µg	-Cost-effective. -Highly quantitative.	A significant amount of DNA input is required.		Low	Most tissue-specific DMR occurs at CpG island shores but not at CpG islands. (1)
MBDCap-Seq	A sequencing technology that relays on proteins to capture hypermethylated DNA in the genome.	60% of genomic CG	1 µg	-Cost-effective. -Allow detecting methylated cytosine in dense CpG-dense regions and regions with lower CpG density. -MBD proteins can discriminate 5mC from 5hmC.	- Low resolution in comparison to other single base resolution technologies. -Detect the hypermethylated cytosine only. -DNA region with low-density methylation can be missed.	~150 bp	Moderate	Identified and confirmed 11 CpG islands that were methylated in 80-100% of the Squamous Cell Carcinoma. (2)
MeDIP	5mC-specific antibodies are used to study cytosine modifications in isolated, methylated DNA from genomic DNA by immunoprecipitation.	50-90% of genomic CG	0.1-5 µg	- Covers CpG and non-CpG methylated cytosines throughout the genome. - Detect methylation in dense, less dense, and repeat regions. - Antibody-based selection is independent of DNA sequence.	-Specific-antibodies are biased towards hypermethylated cytosines only. -Antibody specificity and selectivity should carefully test to avoid nonspecific interaction.	~150 bp	Moderate	Analysis of 6,000 CpG islands in the inactive X chromosome showed that aberrant methylation of CpG island promoters in malignancy might be less frequent than previously hypothesised. (3)

Illumina's Infinium Methylation assay	A BeadChip technology that uses Infinium I and Infinium II probes to generate comprehensive genome-wide profiling of human DNA methylation.	~2% of genomic CG	250-500 ng	-Cost-effective. -Do not require a significant amount of input DNA. -No PCR is needed, which means that there will be no selection bias towards shorter fragments. - Process up to 8 samples per chip allowing high throughput processing. - The ability to integrate data between other platforms (E.g. gene expression and microRNA profiling). - Provide genome-wide methylation patterns coverage as the method looks at ~2 CpG sites per CpG island.	-Available for the human genome only. -The CPG coverage is dependent on the array design (27K, 450K, and 850K). -Risk of DNA degradation post bisulfite conversion.	Single base	Low	Estimate of nucleated cells types in cord blood is a DNA methylation reference panel. (4)
WGBS	A protocol to detect methylated cytosines in bisulfite converted DNA by sequencing	>90%	0.5-1 µg	- Covers CpG and non-CpG methylated cytosine throughout the genome. - Detect methylation in dense, less dense, and repeat regions.	-High cost. -Risk of DNA degradation due to bisulfite treatment. -Does not distinguish between 5mC and 5hmC.	Single base	High	Identify differentially methylated regions proximal to genes involved in pluripotency and differentiation in human embryonic stem cells. (5)
RRBS	A protocol that uses one or multiple restriction enzymes on the genomic DNA to produce sequence-specific fragmentation	20-60% of genomic CpGs	100 µg	-The technique of choice to study specific regions of interest. -Useful in high dense methylation region as in promoters and repeat regions. - Genome-wide covers CpGs in islands at single-base resolution.	- Limited methylation analysis. -Risk of DNA degradation due to bisulfite treatment. -Do not distinguish between 5mC and 5hmC. - Does not cover non-CpG areas, genome-wide CpGs or CpGs in regions with no enzyme restriction site.	Single base	Moderate	Glioblastoma-derived cancer stem cells possess unique epigenetic signatures that may play essential roles in the pathogenesis of Glioblastoma. (6)

Targeted bisulfite-seq	A technology for DNA methylation analysis of target regions using a hybridisation-based step on platforms containing pre-designed oligos to the target region.	<<<1%	250 ng	- Analyse the methylation of DNA at single-base resolution DNA of particular regions. - Cost effective. - Differentiate between SNPs and methylation status.	-Risk of DNA degradation due to bisulfite treatment. -Do not distinguish between 5mC and 5hmC.	Single base	Moderate	Epigenetic dysregulation of Wnt signalling and the pathogenesis are linked to the progression of Parkinson disease. (7)
TAB-seq	A novel method that uses bisulfite conversion and Tet proteins to study 5hmC.	>90%	1–3 µg	- Distinguish 5hmC from 5mC in single base resolution. - Detect 5hmC in dense, less dense, and repeat sites.	-Substantial DNA degradation after bisulfite treatment. -Tet protein with low efficiency might miss methylated bases. -Deep sequencing is required to cover the low abundant 5hmC in the whole genome. -Harsh oxidation can lead to substantial loss of DNA.	Single base	High	Profile the methylation of primordial germ cells from the migrating stage to the gonadal stage at single-cell and single-base resolutions. (8)

Conclusion

Islet autoantibody titration and the measurement of beta cell function are the best current methods for T1D diagnosis and progression prediction. These biomarkers still have some limitations particularly in clinical practice trials as they have been shown to be indirect and not specific for measuring beta cell mass. Therefore, the lack of current biomarkers to measure the efficacy of preventative and interventional therapeutic approaches still poses an immense challenge. Recently, accumulating evidence suggesting that epigenomic biomarkers such as differentially methylated *INS* promoter assays have shown themselves to potentially be powerful tools to detect real-time beta cell injury. Multiplexing several DMR targets had also shown to increase the sensitivity and the specificity of the assay during the preclinical stage of T1D that may cure T1D. In this research project, we identify multiple novel targets to facilitate designing a multiplex assay that monitors real-time beta cell injury.

Project Rationale

Hypothesis

Differential methylation patterns of beta cell-specific CpG sites can be used as biomarkers to detect dying beta cells in the circulation. The sensitivity and specificity of the designed assay can be increased by multiplexing several different differentially methylated targets.

Aims

1. Identify novel differentially methylated CpG sites within beta cell-specific genes **(Chapter 4)**.
    a. Verify the specificity of the identified targets in other somatic tissues methylome conducted from the Roadmap and ENCODE Studies.
    b. Design and optimise methylation sensitive probe-based assays for the identified CpG targets.
    c. Optimise a multiplex assay of the designed probe-based assays and confirm the sensitivity and the specificity of the developed multiplex assay.

2. Test the effect of hyperglycaemia and proinflammatory cytokines on the methylation of the identified CpG targets in the genome of the beta cell line (EndoC-βH1) **(Chapter 4)**.
3. Optimise the method of cfDNA extraction from human liquid biopsy samples, i.e. plasma/ serum and urine **(Chapter 5)**.
4. Test the performance of the developed methylation sensitive multiplex assay in detecting beta cell DNA in the circulation of children and adults with recently diagnosed T1D, at increased risk for developing T1D **(Chapter 6)**.

Chapter 2: General Methods

2. General Methods

2.1. Samples

2.1.1. Human Tissue

Islet samples, which was used for the whole genome methylation sequencing (Illumina EPIC BeadChip), were either provided by Dr Stephen Hughes from the Oxford Centre for Diabetes (JDRF, Oxford, UK) or by Professor James Shaw from the University of Newcastle. Details of the islets are provided in Table 2.1 (Purity: 40-75%, mean Purity: 61%). Peripheral blood mononuclear cell (PBMC) samples (*n*=3) were obtained locally from healthy adult female donors from the Department of Diabetes and Metabolism, University of Bristol (Bristol, UK). Healthy human cortex kidney tissue samples (*n*=2) were provided by Dr Gavin Welsh, (Bristol Renal Group, University of Bristol, UK). The normal human spleen, heart, liver, and lung tissue samples obtained from an anonymised male cadaver with scleroderma were provided by Professor Lee Nelson (Fred Hutchinson Cancer Research Centre, Seattle, US). All samples were collected with informed consent.

2.1.2. Cell Lines

2.1.2.1. EndoC-βH1

For experiments involving human cell beta line, we used EndoC-βH1 cells, which were kindly provided by Professor Aniko Varadi, (Department of Applied Sciences, University of West England, Bristol, UK). EndoC-βH1 is a genetically engineered human pancreatic beta cell line derived from human foetal pancreatic buds. Briefly, human foetal pancreatic ducts were transduced with a lentiviral vector expressing the oncogene Simian Virus 40 large T antigen (SV40LT) under the control of the rat insulin promoter. The foetal ducts were then grafted into severe combined immunodeficiency (SCID) mice to allow cell expansion. Prior to re-grafting into other SCID mice to allow proliferation, beta cells that express SV40LT were isolated and transduced with human telomerase reverse transcriptase to induce cell immortalisation. The resulting beta cells were further expanded in culture to form cell lines (Ravassard et al. 2011).

2.1.2.2. Prostate and Breast Cancer cell lines

Both prostate and breast cancer cell lines were kindly provided by Professor Jeffrey Holly (IGFs and Metabolic Endocrinology Group (IMEG), University of Bristol, UK). All cell lines were provided as a cell pellet from which the DNA/microRNA were extracted. All Prostate cell lines (LNcaP, DU145, VcaP and PC3) were obtained from the American Type Culture Collection (ATCC) except for the differentiated prostate epithelial cell line (PNT2) which was obtained from the European Collection of Authenticated Cell Cultures (ECACC). Breast cell lines HS 578T, MCF7, MDA-MB-231, T47D and ZR-751 were obtained from the American Type Culture Collection (ATCC). Details of all the human prostate cell lines are summarised in Table 2.2; the breast cell lines details are summarised in Table 2.3.

2.2. Cell Culture

2.2.1. EndoC-βH1 Cell Culture

EndoC-βH1 cells were cultured in Dulbecco's modified Eagles Medium (DMEM) low glucose (1 g/L) medium containing 2% foetal bovine serum (FBS) fraction V, 50 µM 2-mercaptoethanol, 10 mM nicotinamide, 5.5 µg/mL transferrin, 6.7 ng/mL sodium selenite and Penicillin (100 units/mL)/Streptomycin (100 µg/mL). Under normal circumstances, the cells were incubated in a humidified atmosphere of 95% air and 5% CO_2 at 37°C in which they were routinely passaged when they reached 70-80% confluence. The cells were fed every two days.

For the experiment of examining the effect of hyperglycaemia on the methylome of the beta cell, EndoC-βH1 cells were grown in T25 flasks for 5 days until confluent. Then each flask was grown in normal media with 25 mM glucose instead of the usual 5.5 mM up to 48 hours. One control flask was kept in 5.5 mM glucose throughout the experiment. After the treatments, the flasks were washed with 1X Phosphate-buffered saline (PBS) and trypsinised to remove cells from the flask. Neutralisation medium was then added (80% of 1X PBS and 20% Bovine serum albumin (BSA)), and the cells were spun down at 700x g for 5 minutes. EndoC-βH1 cells were then resuspended in 200 µL PBS and frozen at -80^{o}C for DNA extraction.

For the experiment studying the effect of the proinflammatory cytokines on methylation stability, TNF-α, IFN-γ and IL-1β were used as these three cytokines are widely utilised to mimic T1D and induce beta cell death *in vitro* (Liu et al. 2000, Rabinovitch and Suarez-Pinzon 1998). The effect of cytokines was examined after 1, 3 and 7 days of culture in the presence of recombinant murine IFN- γ (14 U/mL "10.3 ng/mL", eBioscience), recombinant human IL-1β (60 U/mL, "4.4 ng/mL", Abcam), and recombinant murine TNF-α (185 U/mL, "8.7 ng/mL", Abcam). These concentrations of cytokines are commonly used in beta cell research (Lortz et al. 2000, Mandrup-Poulsen et al. 1987, Nerup et al. 1994). Control EndoC-βH1 cells were grown in the absence of test proinflammatory cytokines.

2.2.2. Prostate and Breast Cell Lines Culture

The DU145, PC3, and PNT2 cell lines were maintained in 25 mM, 4.5g glucose/L DMEM, (BioWhittaker, Verviers, Belgium) supplemented with 10% heat-inactivated FBS (Gibco, Paisley, UK) and 5% L- Glutamine (LG, Lonza Ltd). The LNCaP cell line was maintained in 25 mM, 4.5 g glucose/L RPMI-BW12-702F media, with 5% L-glutamine and 10% FBS. Stock cultures of DU145 and PC3 cells were grown in T175 cm^2 flasks (T175, Greiner Bio-one) while PNT2 and LNCaP cell lines were grown in T75 flasks (T75, Corning) to approximately 80-90% confluency before passaging. Breast cell lines were maintained in either normal (5 mM, 1 g/L) or high (25 mM, 4.5 g/L) glucose-containing DMEM (Sigma, Dorset, UK). The media were supplemented with 10% of FBS and 1% L-glutamine. The normal human mammary epithelial cell line (MCF10A) was maintained in Ham's F12 medium and DMEM. The media, which was supplemented with 1% L-glutamine (DMEM: F12, Gibco), 5% horse serum (v/v HS, Gibco), 20 ng/mL EGF (Calbiochem, Nottingham, UK), 100 ng/mL cholera toxin (Sigma), 10 µg/mL insulin (Novo Nordisk, West Sussex, UK) and 0.5 µg/mL hydrocortisone (Sigma), were maintained in 5.5 mM, 4.5g glucose/L DMEM, (BioWhittaker, Verviers, Belgium) supplemented with 10% heated inactivated FBS (Gibco, Paisley, UK) and 5% LG (Lonza Ltd).

2.3. Serum and Plasma Samples

2.3.1. Locally Collected Samples (Assay Optimising Experiments)

Peripheral whole blood samples were collected in a 5 mL SST™ tube (Becton Dickinson) containing a gel barrier to separate the serum after centrifugation. All samples were processed at room temperature within 2 hours from the time of venepuncture. Serum was separated from the cellular fraction by centrifugation at 2,000xg for 15 minutes. After centrifugation, each serum sample was transferred into a sterile Lobind Eppendorf tube and re-centrifuged at 16,000 g for 10 minutes. Afterwards, the supernatant was divided into aliquots of 300 µL. Aliquots were stored immediately at -80°C until cfDNA was extracted. Plasma was separated from the cellular fraction of peripheral whole blood samples collected in 8.5 mL cell-free DNA collection tubes (Roche) by centrifugation at 1,600xg for 15 minutes. The separated plasma samples were divided into aliquots of 1 mL in an appropriate Lobind Eppendorf tube and were stored immediately at -80°C until cfDNA was extracted.

2.3.2. Recently Diagnosed Paediatric Cohort

Serum samples from a small cohort of children with recent onset T1D (*n*=9, age range: 4-14 years) was provided by Dr Chloe Bulwer (University College London Hospital and the Institute of Child Health, University College London, UK). Each sample was collected into serum-separation tubes (SST^{Tm}, BD, Oxford, UK). Serum was separated within 1-4 hours post venepuncture by centrifuging at 1000xg for 10 minutes and were stored at -80^{o}C until they were shipped to the University of Bristol for further analysis. The data for each participant is summarised in Table 2.4. Serum samples from a control group of non-diabetic children (*n*=12, age range: 6-15 years) were provided by Dr Danijela Tatovic (Brecon Cohort, Diabetes Research Group, University of Cardiff). Each sample was collected into SST^{TM} tubes and was separated within 2 hours post venepuncture by centrifuging at 2000xg for 15 minutes. Serum samples were stored at -80^{o}C until they were shipped to the University of Bristol for further analysis. The data of each participant is summarised in Table 2.5.

2.3.3. Recently Diagnosed Adult Cohort

2.3.3.1. Bristol Cohort

Plasma samples were collected from males who were recently diagnosed with T1D (*n*=7, median duration of 4 months, range 1-7.9 months) as well as from healthy controls (*n*=4) (Table 2.6). The median age of patients at the time of sampling was 26.9 year (range 22.3 - 37.0 years) and of 35.0 years in controls (range 25.2 - 44.5 years). Patients were recruited from the Southwest Newly Diagnosed Diabetes Collection (SWENDIC) and had a clinical diagnosis of T1D according to World Health Organization criteria. Control Subjects lived in the same region and were recruited between October 2007 and October 2010. Up to 70 mL of venous blood was drawn from each individual and processed within one-hour post venepuncture.

2.3.3.2. StartRight (Exeter) Cohort

We studied the level of hypomethylated singleplex, and multiplex DNA in plasma samples (*n*=9, female = 4, male = 5) collected from individuals participating in StartRight Young Adults with diabetes Study (University of Exeter). The median age of patients at venepuncture was 26 years (range 18 - 26 years) and of controls 25.5 years (range 22 - 49 years). Samples were collected within 3 months of diabetes diagnosis (Table 2.7). Plasma from age-matched nondiabetic control subjects (*n*= 12) were collected from the local bank of control samples (Table 2.8). Peripheral blood was collected in 8.5 mL Roche cfDNA tubes; plasma was separated within 1-hour post venepuncture by spinning the sample at 1600xg for 10 minutes. Plasma samples were then stored at -80^{o}C until further analysis.

2.3.4. At Risk for Developing T1D Cohort

Peripheral blood samples were obtained from blindly selected relatives of individuals with T1D (*n*= 15) participating in the TrialNet Natural History study as part of an Ancillary study (https://www.trialnet.org/). These subjects were classified as “at risk” of future developing T1D, as they were positive for at least one autoantibody against GAD65, ICA512/IA-2, ZnT8, insulin or ICA. All the subjects had a normal level of HbA1c (<6.5%; <48 mmol/mol), did not show clinical symptoms of

diabetes and samples were collected between 17/08/2017 and 16/05/2018 (Table 2.8). The samples were collected in Cell-Free DNA Collection Tube (Roche, Roche Diagnostics GmbH, Germany). Plasma was separated by centrifugation at 1,600xg for 15 minutes, and then frozen at -80^{o}C until further analysis. Sera from age-matched nondiabetic control subjects were collected from the local bank of control samples.

2.3.5. Islet Transplantation T1D Recipients

Plasma samples from longstanding T1D patients (*n*= 2) who underwent allo-islet transplantation were provided by Dr Vito Lampasona (IRCCS San Raffaele Scientific Institute, Italy). The patients were participating in either the Bone Marrow vs Liver as a Site for Islet Transplantation (IsletBOM2) or in a study assessing the efficacy and safety of Reparixin in pancreatic islet transplantation (REP0211). From each participant, whole blood samples were drawn in EDTA collection tube before the islet infusion, and then at 1, 3, 6, 12, and 24 hours, and 2, 3, 5 and 7 days after the end of islet infusion. Plasma was separated by 2500xg for 10 minutes at 4°C; Samples were then frozen at -80^{o}C until further analysis.

2.3.6. Urine Samples

Urine samples used for studying cfDNA stability were collected by taking midstream of the first-morning urine sample from healthy volunteers and were processed within 2 hours post collection (Table 2.10). Upon collection, urine samples were spun at maximum speed for 10 minutes to precipitate the cellular component. The Supernatant was transferred into a new sterile 50 mL centrifuge tube. The sample was then aliquoted (1 mL, in appropriate Eppendorf tube) and incubated as indicated in the protocol (Chapter 5).

Table 2.1: Islet Cells Characteristics.

Islet ID	Form	Source	Date of Receiving	Cell count& Storage method	Purity	Viability	Donor	Age	BMI
LD117	islet pellet	Newcastle			NA	NA	NA	NA	NA
HP14-50	islet pellet	Oxford	26/01/2015	~3000 islets one aliquot extracted -20^{o}C	60%	80%	M	49	26
HP15-20	islet pellet	Oxford	10/04/2015	17,000 islets divided into 5 aliquots in 100 µL PBS -20^{o}C	70%		F	56	29.76
HP15-35	islet pellet	Oxford	18/06/2015	10,000 islets divided into 2 aliquots in 200 µL PBS -20^{o}C	60%	75%	F	41	34
HP1557	islet pellet	Oxford	22/10/2015	10,000 islets divided into 3 aliquots in 200 µL PBS -20^{o}C storage	75%	70%	M	51	31

Table 2.2: Human Prostate Cell Lines Characterizes

Cell Line	Source	Catalogue Number	Provided Cell Passage
DU 145	Organ: Prostate Disease: Carcinoma Tissue derived from: Brain Age: 69 Gender: Male Ethnicity: Caucasian	ATCC: HTB-81	26
PC-3	Organ: Prostate Disease: Adenocarcinoma Tumour Grade: IV Tissue derived from: Bone Age: 62 Gender: Male Ethnicity: Caucasian	ATCC: CRL-1435	24
LNCaP	Organ: Prostate Disease: Carcinoma Tissue derived from Left supraclavicular lymph node Age: 50 Gender: Male Ethnicity: Caucasian	ATCC: CRL-1740	24
PNT2	Organ: Prostate Cell Type: Epithelial Growth mode: Adherent Tissue derived from Prostate Age: 33 Gender: Male	ECACC: 95012613	18
VCaP	Organ: Prostate Disease: Cancer Tissue derived from: Vertebral Metastasis Age: 59 Gender: Male Ethnicity: Caucasian	ATCC: CRL-2876	28

Table 2.3: Human Breast Cell Lines Characterizes

Cell Line	Source	Catalogue Number	Provided Passage
MCF-7	Organ: Breast Disease: breast adenocarcinoma Tissue derived from: Pleural Effusion Age: 69 Gender: Female Ethnicity: Caucasian	ECACC: 86012803	18
MCF-10A	Organ: Breast Tissue derived from non-Tumorigenic adherent cells Age: 36 Gender: Female Ethnicity: Caucasian	ATCC: CRL-10317	16
MDA-MB-231	Organ: Breast Disease: metastatic breast cancer Tissue derived from: Pleural Effusion Age: 51 Gender: Female Ethnicity: Not specified	ECACC: 92020424	26
T47D	Organ: Breast Disease: ductal carcinoma of the breast Tissue derived from Pleural Effusion Age: 54 Gender: Female Ethnicity: Not specified	ECACC: 85102201	23
HS-578-T	Organ: Breast Disease: breast carcinoma Tumorigenic: No Age: 74 Gender: Female Ethnicity: Caucasian	ATCC: HTB-126	24
ZR-75-1	Organ: Breast Disease: ductal carcinoma of the breast Tissue derived from ascites Tumorigenic: Yes Age: 63 Gender: Female Ethnicity: Not specified	CRL-1500	26

Table 2.4: Recently Diagnosed T1D Paediatric Samples (UCL Cohort)

Sample ID	Sample	Gender	Age (year)	Duration of Diabetes (Days)	Blood processed
BCD1	Serum	F	10	315	within 1 hour – 1000x g for 10 minutes and frozen to -80^{o}C
BCD2	Serum	F	13	280	within 1 hour -1000x g for 10 minutes and frozen to -80^{o}C
BCD3	Serum	F	7	372	within 1 hour -1000x g for 10 minutes and frozen to -80^{o}C
BCD4	Serum	M	9.5	25	within 4 hour -1000x g for 10 minutes and frozen to -80^{o}C
BCD5	Serum	M	8	58	within 4 hour -1000x g for 10 minutes and frozen to -80^{o}C
BCD6	Serum	F	6	296	within 4 hour -1000x g for 10 minutes and frozen to -80^{o}C
BCD8	Serum	F	14	388	within 1 hour -1000x g for 10 minutes and frozen to -80^{o}C
BCD9	Serum	M	11.5	191	within 1 hour -1000x g for 10 minutes and frozen to -80^{o}C
BCD 10	Serum	F	4	8	within 2.5 hour -1000x g for 10 minutes and frozen to -80^{o}C

Table 2.5: The Characteristics of Paediatric Non-Diabetic Control Samples

Study ID	Sample	Gender	Age (Years)
UHWBC002	Serum	Female	6
UHWBC005	Serum	Male	6
UHWBC006	Serum	Male	15
UHWBC007	Serum	Female	15
UHWBC010	Serum	Female	8
UHWBC011	Serum	Female	6
UHWBC012	Serum	Female	9
UHWBC013	Serum	Male	9
UHWBC016	Serum	Female	7
UHWBC023	Serum	Male	14
UHWBC024	Serum	Female	7
UHWBC026	Serum	Female	15

Table 2.6: The Characteristics of Recently Diagnosed Adult Plasma Samples (Bristol Cohort)

Sample	Age	Gender	HbA1c	Diabetes	Fasting C-peptide	Duration (years)	Duration (months)
PATSI 1	25.2	M	5.2	No	0.23		
PATSI 2	27.7	M	5.5	No	0.62		
PATSI 3	44.5	M	5.4	No	0.52		
PATSI 4	42.2	M	7.4	T2D	0.52	0.17	2.0
PATSI 5	24.1	M	5.6	T1D	0.09	0.20	2.4
PATSI 6	25.9	M	7.7	T1D	0.2	0.08	1.0
PATSI 7	26.9	M	7.1	T1D	0.28	0.15	1.8
PATSI 8	37.0	M	7.3	T1D	0.28	0.66	7.9
PATSI 9	22.3	M	13.5	T1D	0.23	0.46	5.5
PATSI 10	34.1	M	11	T1D	0.04	0.52	6.3
PATSI 11	36.7	M	6	T1D	0.1	0.47	5.6

Table 2.7: The Characteristics of Recently Diagnosed Adult Plasma Samples (StartRight Cohort)

Sample ID	Age	Gender	Ethnic Origin
SR110293	21	Male	White British
SR110453	18	Male	White British
SR110122	48	Female	White British
SR110123	23	Male	Other Mixed
SR110126	26	Female	White Irish
SR110130	21	Female	White British
SR110285	35	Male	Any Other Group
SR110468	39	Female	Other White
SR110469	28	Male	White British

Table 2.8: The non-Diabetic Healthy Adult Samples Characteristics.

CFDNA_ID	Gender	Age
SCF2101	M	37
SCF2102	F	30
SCF2103	F	25
SCF2104	M	36
SCF2105	F	49
SCF2106	M	29
SCF2107	F	26
SCF2108	M	33
SCF2109	M	28
SCF2110	F	26
SCF2111	F	22
SCF2112	M	24
SCF2113	M	15
SCF2114	M	24
SCF2115	F	31

Table 2.9: The At-Risk for Developing T1D Samples Characteristics.

Patient IDs	TrialNet ID	Age	Gender	GAD65/GAD65H	ICA512/IA-2H	MIAA	ICA	ZnT8A	OGTT	HbA1c	HLA Typing	
											DQA1*0102 DQB1*0602	DQB1*0302
TNID0001	**189896 PH1**	50	Male	Positive	Negative	Negative	Negative	Negative	Indeterminate	Normal	Absent	Absent
TNID0002	**251453 HAF**	15	Male	Positive	Negative	Negative	Negative	Negative	Normal	Normal	Absent	Present
TNID0003	**251743 EMA**	16	Female	Positive	Negative	Negative	Negative	Negative	Normal	Normal	Present	Absent
TNID0004	**205410 SAO**	17	Male	Positive	Positive	Negative	Positive	Positive	Normal	Normal	Absent	Absent
TNID0005	**2323200 EF**	12	Female	Positive	Negative	Positive	Positive	Negative	Normal	Normal	Absent	Present
TNID0006	**427709 RUS**	9	Female	Positive	Negative	Positive	Negative	Negative	clinical Alter	Normal	Absent	Present
TNID0007	**428879 SAA**	47	Female	Positive	Negative	Positive	Negative	Negative	Normal	Normal	Absent	Present
TNID0008	**189896 PHI**	50	Male	Positive	Negative	Negative	Positive	Negative	Indeterminate	Normal	Absent	Absent
TNID0009	**432363 GIJ**	9	Male	Negative	Negative	Positive	Negative	Negative	Normal	Normal	Absent	Present
TNID0010	**447149 AIN**	6	Female	Positive	Negative	Negative	Negative	Negative	Impaired	Normal	Absent	Present
TNID0011	**433709 PAF**	19	Male	Positive	Negative	Negative	Positive	Negative	Normal	Normal	Absent	Present
TNID0012	**251453 HAF**	16	Male	Positive	Negative	Negative	Negative	Negative	Normal	Normal	Absent	Present
TNID0013	**347666 HES**	9	Male	Negative	Negative	Negative	Negative	Positive	Normal	Normal	Absent	Present
TNID0014	**433478 OWF**	6	Male	Positive	Negative	Positive	Negative	Positive	Impaired	Normal	Absent	Absent
TNID0015	**127439 GAA**	43	Male	Positive	Negative	Negative	Positive	Negative	Normal	Normal	Absent	Present

Table 2.10: The Non-Diabetic Urine Samples Characteristics.

Control ID	Gender	Age	Time of sample collection
001	F	30	08:40 AM
002	M	25	10:45 AM
003	M	58	10:00 AM
004	F	25	08:00AM
005	F	24	06:20 AM
006	M	34	09: 00 AM
007	M	32	08:40 AM
008	M	32	08:12 AM
009	F	32	08:30 AM
010	F	24	06:32 AM
011	F	24	07:30 AM

2.4. DNA Methylation Public Data

DNA methylome data for adult human liver, lung, thymus, left and right ventricle, adipose tissue, bladder cells, placenta primary tissue and spleen primary tissue were downloaded from Epigenomics Roadmap database (Kundaje et al. 2015). The methylome data was used as a guidance to exclude the CpG sites that are not exclusively hypomethylated in the islet genome. The methylation level and the genomic sequence data were aligned using the Integrated Genome Browser (IGB, http://bioviz.org/).

2.5. Design of Oligonucleotides

The bisulfite specific primers that were used in this project were designed using EpiDesign online software (Agena Bioscience, https://www.epidesigner.com/), with human sequences available from the Genome Reference Consortium Human Build 37 database (GRCh37/hg19). The primer pairs were selected using the following criteria: primer Tm (54-58°C), primer length (18-30 bp), and amplicon size (80-120bp). Probes were designed manually around the desired CpG site. Primers/probes compatibly were checked using Primer Express™ Software (version 3.0.1). All primers were obtained from Sigma Aldrich (Synthesis scale 0.025 µM, Desalted purification). Probes, which were HPLC purified, were

obtained from Life Technologies. The full details for all oligonucleotides used in this project are listed in Appendix A, Table A.1.

2.6. Polymerase Chain Reaction (PCR)

PCR is a technique that amplifies a single copy of DNA into thousands of folds of copies *in vitro* (Mullis and Faloona 1987). During a typical PCR reaction, a new (daughter) DNA strand is synthesised from the template DNA strand under the action of the polymerase. The reaction is completed in the presence of a buffer that contains primers, flanking the desired DNA segment, polymerase enzyme, and deoxynucleotide triphosphates (dNTPs). Once the template DNA strands are denatured, the primers anneal to complementary sequences on the DNA template; then the DNA polymerase adds nucleotides in the 5'→ 3' DNA direction to create a new strand in each PCR cycle.

2.6.1. Droplet Digital PCR (ddPCR)

The idea of the droplet digital PCR was first defined in the late 1990's (Vogelstein and Kinzler 1999). Similar to the quantitative PCR (qPCR), ddPCR involves adding primers and probes to detect the targeted sequences. However, ddPCR is more sensitive and precise in comparison to qPCR. The major difference between the two methods, in the ddPCR, the samples are randomly distributed using a water-oil emulsion droplet (microfluidics) technology that partitions the sample into nanoliter 20,000 droplets before PCR cycling to the endpoint. After endpoint PCR, the presence or absence of template DNA in each droplet is determined individually. Since the partitioning is random, the absolute copy numbers of DNA fragments can be calculated using Poisson statistics (Hindson et al. 2011) (Bio-Rad droplet digital application guide).

2.6.2. Probe-Based ddPCR

The ddPCR reaction mixture contains 1X ddPCR Supermix for Probes (12.5 µL, Bio-Rad), appropriate concentration of primers and probes, 0.4X Q-solution (Qiagen), the bisulfite converted DNA sample and nuclease-free water (Qiagen) to a total volume of 25 µL. To minimise the pipetting errors, all components, except the bisulfite converted DNA sample, were premixed in LoBind Eppendorf tubes (Sigma Aldrich), and the final PCR mix was prepared individually by combining the bisulfite converted DNA sample with the pre-sample mixture. Twenty microliters of the PCR mix and

70 µL of the Droplet Generation Oil for Probes (Bio-Rad) were pipetted into the appropriate wells in the Droplet Generator DG8™ Cartridge (Bio-Rad). For each sample, the assay was run in duplicate.

2.6.3. Droplet Generating, PCR Cycle Conditions

Once samples and appropriate oil was loaded in the cartridge, the DG8 cartridges were covered by the Gaskets (Bio-Rad) and then placed in a QX200™ Droplet Generator (Bio-Rad). Next, the droplets were gently transferred into a semi-skirted and PCR clean 96-well PCR plate (Eppendorf) using an E4 XLS Multichannel pipette with the range 5-50 µL (Rainin). The PCR plate was sealed with pierceable foil (Bio-Rad) using a PX1™ PCR Plate Sealer for 10 seconds (Bio-Rad). Afterwards, the PCR amplification was performed by placing the plate in a C1000 Touch™ Thermal Cycler (Bio-Rad). For the probe-based droplets the PCR, the standard PCR protocol was used in the single plex and duplex experiments (Figure 4.5); the touchdown PCR protocol was used in the multiplex probe experiments (Figure 4.6). For the EvaGreen based ddPCR, the PCR cycle conditions were as follows: 95°C for 5 minutes; 40 cycles of 95°C for 30 s, 58°C for 1 minute (Ramp-Rate 1.6 °C /s), 4°C for 5 minute and a final step at 90°C for 5 minutes (Figure 7.1). The reaction plates were then held at 4°C until reading. The droplet readings and analyses were performed with the QX™ 200 Droplet Reader (Bio-Rad) with a ddPCR™ Droplet Reader Oil (Bio-Rad). The data analyses were performed using the software package QuantaSoft™ (Bio-Rad, V 1.6.6.0320). The threshold of fluorescence amplitude, which distinguishes the positive from the negative droplets, was set manually. The same threshold was applied to all the wells of one PCR plate. The average droplet number that was accepted for valid measurements was around 17,000.

Chapter 3: The Pancreas and Duodenum Homeobox Protein-1

3. Introduction

Although developed methylation sensitive insulin promoter assays have been shown to be a potential biomarker measure that measure beta cell destruction during the pathogenesis of T1D, they were difficult to replicate in other laboratories. Therefore, the main aim of this project was to identify other differentially methylated CpG sites in beta cell-specific genes (other than *INS*). The identified CpG targets will be used to develop singleplex and multiplex assays that would be able to discriminate circulating cfDNA of beta cell-origin with high sensitivity and specificity. In this chapter, the initial approach that was used to identify differentially methylated CpG sites will be described.

The methylome of the gene the Pancreatic Duodenal Homeobox-1 *(PDX-1*) (Chr13:28,494,168-28,500,451), was studied. *PDX-1*, which is also known as insulin promoter factor 1, is a member of homeobox factors that plays a key role in pancreas development (Jonsson et al. 1994). During the early embryonic period of life, *PDX-1* is expressed in the gut region when the foregut endoderm becomes committed to common pancreatic precursor cells. This expression persists in pancreatic tissue with high levels at around 4 weeks of gestation, but then declines until it becomes restricted to adult human beta cells (Lyttle et al. 2008). It is now well acknowledged that *PDX-1* plays a crucial role in pancreas development, differentiation and in maintaining mature beta cell function, and more importantly, its central role in glucose-dependent insulin secretion (Hui and Perfetti 2002). *PDX-1* is also essential for beta cell phenotype maintenance as *PDX-1* knockout specifically in beta cells results in phenotype loss (Ahlgren et al. 1998) and pancreas agenesis (Stoffers et al. 1997). Furthermore, *PDX-1* is required for beta cell survival as cells with reduced *PDX-1* were shown to have a higher rate of apoptosis (Johnson et al. 2003, Johnson et al. 2006). Treating diabetic mice with recombinant Pdx-1 protein induces beta cell regeneration, transient hepatocyte reprogramming and prevented the onset of T1D in pre-diabetic NOD mice (Koya et al. 2008). Moreover, a successful acinar cell reprogramming in mice into functional insulin-producing beta cells required acinar cells viral transduction infection with *PDX-1*, *Neurog3*, and *MafA* (Zhou et al. 2008). These data were successfully replicated *in vitro* as

another group reprogramed a rat pancreatic exocrine cell line (AR42J) into a beta cell using *Pdx-1*, *Ngn3* and *MafA* (Akinci et al. 2012).

In order to fulfil our goal, we studied the methylation pattern of the distal promoter region (Chr13: 28,493,260 - 28,493,696) of *PDX-1*. As this assay will be used in the future to discriminate beta cell fragments in the circulation, it was important to consider the sources of circulating DNA contamination that is derived from other cells and tissues. The primary source of contamination would be from PBMC; other circulating DNA contamination would potentially originate from apoptotic cells. Consequently, the pattern of *PDX-1* promoter methylation in islets was compared to the PBMC genome. Next, the specificity of the identified differentially methylated CpG sites toward the islet genome was examined by comparing its methylation level to a panel of DNAs from locally available tissues. It was also important to test the methylation level of the biomarker developed in the duodenal genome as the gene has been reported to be expressed in the duodenum.

Methods

3.1.1. DNA Extraction from Human Cells and Tissues

Genomic DNA was extracted from frozen tissue samples (described in General Methods, Chapter 2: Section 2.1.1) using a QIAamp DNA mini kit (Qiagen) following the manufacturer's instructions. Approximately 50 mg of each sample was thawed at room temperature, and then 20 µL of Qiagen Proteinase K and 200 µL of lysis buffer was added to break down the proteins. The samples were mixed by vortexing for 15 seconds and then incubated at 56°C for 10 minutes. 200 µL of 100% ethanol was added to the samples, mixed well and transferred to the QIAamp Mini spin column. The columns were spun for 60 seconds then the flow-through was discarded. Next, 500 µL of Buffer AW1 was added to wash the columns, and the tubes were spun for 60 seconds. The columns were placed into a new 2 mL collection tube and were rewashed with 500 µL Buffer AW2. The tubes were spun at maximum speed for three minutes. The columns were placed in new 1.5 mL tubes and centrifuged at maximum speed for 60 seconds to remove any remaining buffer left in the columns. Next, the columns were transferred to a new 1.5 mL tube and 50 µL buffer AE (pre-warmed 45°C) was added directly to the column membranes to elute the DNA. The tubes were incubated at 45^{o}C for 10 minutes and then centrifuged for 2 minutes at maximum speed. The DNA was recovered and measured using the Quantus™ Fluorometer instrument according to manufacturer's instructions.

3.1.2. DNA Quantification

Genomic double-stranded DNA was quantified using the QuantiFluor® dsDNA system (Promega) following the manufacturer's protocol. The Standard curve was constructed using Lambda DNA (provided within the kit) in the range of 0-200 ng/µL from which the concentration of unknown samples was extrapolated. The DNA concentration readings were performed using the Quantus™ Fluorometer (Promega).

3.1.3. Bisulfite Conversion

After DNA extraction of the samples, 500 ng of genomic DNA from each sample was bisulfite converted using the EZ DNA Methylation-Lightning™ Kit (Zymo Research) in accordance with the manufacturer's protocol. In summary, the 500 ng of genomic DNA sample was made up to a volume of 20 µL with nuclease-free water. 130 µL of CT conversion reagent was then added to the genomic DNA sample and mixed, and the mixture was run in a thermal cycler using the following programme: 98°C for 10 minutes, 54°C 1 hours, then 4°C up to 20 hours. The sample was subsequently loaded into a Zymo spin column with 600 µL of M-binding buffer, mixed and centrifuged at maximum speed for 30 seconds and the flow-through discarded. Next, 100 µL of M-wash buffer was added to the column, followed by a further centrifuge at maximum speed for 30 seconds. The column was then incubated for 20 minutes with 200 µL of M-desulphonation buffer protected from direct light. This step was followed by two additional wash steps with 200 µL of M-wash buffer. Finally, the columns were transferred into a new 1.5 mL tube, 10 µL of M-elution buffer was added to the column and incubated at room temperature for 2 minutes. Then columns were centrifuged at maximum speed for 1 minute and stored at -20^{o}C until further analysis.

Typically, post sodium bisulfite treatment, all cytosines are converted to uracil bases except for those located within methylated CpG dinucleotides which remain intact. In subsequent amplification reactions, unmethylated cytosines produce thymine whereas the protected methylated cytosines retain the cytosine (Figure 3.1).

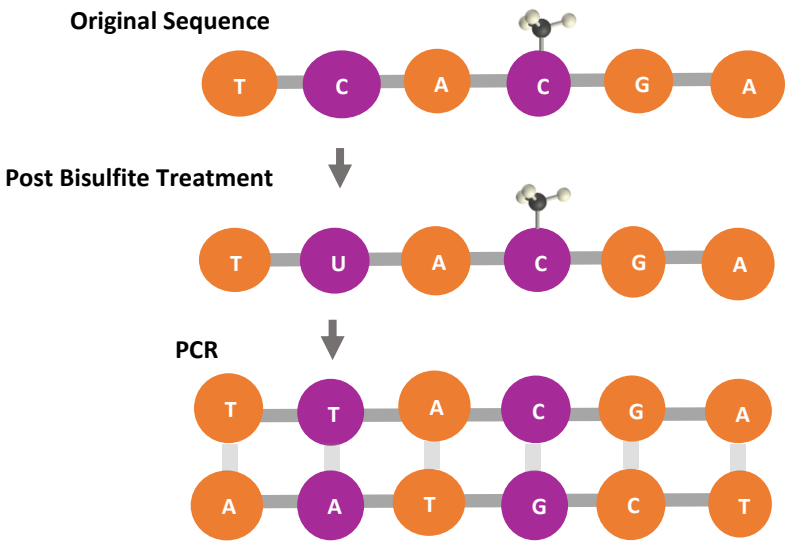


Figure 3.1: Bisulfite Conversion of DNA

DNA with methylated cytosine nucleotide was not affected by the sodium bisulfite treatment whereas the unmethylated cytosine nucleotides have been changed chemically into uracil. In the subsequent PCR, the converted uracil bases were detected as thymine and the methylated cytosine remained unchanged.

Following sodium bisulfite treatment of the DNA samples, PCR amplification was performed with primers that were designed to amplify bisulfite treated DNA strands (Bisulfite specific primers). The PCR master mix was prepared using reagents from the Hot Start Taq Polymerase kit (Qiagen). The 10 µL PCR mixture contained 1 µL of 10X PCR buffer, 500 µM $MgCl_2$, 200 nM dNTPs, 500 nM desired primers, 0.05 µL Qiagen HotStar Taq and 5 ng bisulfite treated DNA. The PCR cycle was: 95°C for 10 minutes, 50 cycles of 95°C for 30 seconds and 58°C for 50 seconds, 72°C for 50 seconds, followed by incubation step of 98°C for 10 minutes and then a final storage incubation at 4°C. PCR products were visualised under UV light after staining with Midori Green in a 1.8% agarose gel.

3.1.4. Gel Electrophoresis

In experiments that needed DNA products to visualise, PCR products were run on a 1.8 % agarose gel. The gel was prepared using molecular biology grade agarose (MelFord), 1X Tris/Borate/EDTA (TBE) buffer (OmniPure) and stained with MedoriGreen (Nippon Genetics). Ten µL of PCR product and 4 µL of 60% Sucrose Creosol Red (SCR) DNA loading dye were loaded into each gel well. A 50 bp DNA ladder

(size range 50 - 1350 bp; BioLabs) was run alongside the samples at 120 mV for 75 minutes. DNA fragments were visualised on a UV transilluminator (183 nm) and photographed with the ChemiDoc MP System and ImageLab software (Bio-Rad, Version 5.2). When needed, the desired sized DNA fragment was excised from the gel using a clean scalpel. Purification was carried out using the MinElute Gel Extraction Kit (Qiagen) following the manufacturer's instructions. The purified DNA fragments were then stored at -20°C until the time of processing.

3.1.5. Vector Amplification

PCR products of the *PDX-1* promoter were cloned into DH5α competent bacterial cells using the pGEM-T Easy vector system I (Promega, detailed protocol in appendix A). A single colony was inoculated in 5 mL of Terrific broth containing 100 µg/mL Ampicillin (Sigma) shaking at 170 rpm overnight at 37°C. On the following day, the cells were pelleted at >8000 rpm for 3 minutes. The plasmids were then purified using the QIAprep Spin miniprep kit (Qiagen) in accordance with the manufacturer's protocol. 50-100 ng/µL of purified plasmid was sent to the Eurofins Genetic Service (https://www.eurofins.co.uk/genomic-services/) for sequencing.

3.1.6. Whole Genome Amplification

Whole genome amplification of the PCR products was initially carried out using QIAGEN's REPLI-g Mini Kit according to the manufacturer's instructions. REPLI-g kit works on a principle is known as Multiple Displacement Amplification (MDA) method, where random primers (hexamers) bind to the template and generate fragments at a constant temperature with the help of a high-fidelity DNA polymerase enzyme, usually 'Φ29 DNA polymerase'. The resulting fragments of DNA are larger than conventional PCR products and with lower error frequency (Dean et al. 2002).

3.1.7. Bisulfite Sanger Sequencing

Sanger sequencing was performed using Eurofins Genetic services (Eurofins Genomic, UK). The PCR, DNA fragment visualisation and DNA cloning were performed following the standard protocols as described above. Then, 50-100 ng/µL of purified vector cloning products (in 15 µL nuclease-free

H_2O) along with 2 µL of 10 µM primer targeted at the desired region of sequencing were sent to Eurofins Genetic service for sequencing. All sequences were generated using BigDye terminator (version 3.1) of Applied Biosystems (Foster City CA, US) following standard protocols. For sequencing reactions, Primus 96 HPL Thermocyclers (MWG AG, Ebersberg, Germany) or DNA engine Tetrad 2 cyclers (Bio-Rad, Munich, Germany) were used. Finally, all reactions were run on ABI3730xl capillary sequencers. Quality values were assigned to each base using KB basecaller software (version 1.4).

Results

3.2.1. The Pancreatic Duodenal Homeobox-1

We aimed to investigate whether islet-specific, differentially methylated CpG sites exist within the *PDX-1* promoter that might be targeted in a cfDNA assay. To achieve this aim, it was first necessary to bisulfite sequence the promoter region of the gene for which the sequence had been previously defined (Yang et al. 2012). Consequently, bisulfite specific primers flanking the proximal (Chr13: 28,494,048 - 28,494,141) (P1), distal (Chr13: 28,493,260 - 28,493,696) (P2) promoter and the enhancer (Chr13: 28,490,368 - 28,490,973) (E) regions were designed. Bisulfite-treated genomic DNA from PBMCs and islets was used as the DNA template for validating the primers. The PCR products were then visualised on a 1.8 % agarose gel with a 50 bp DNA marker for size verification. The result showed that the PCR products yielded a single specific product post PCR amplification (Figure 3.2).

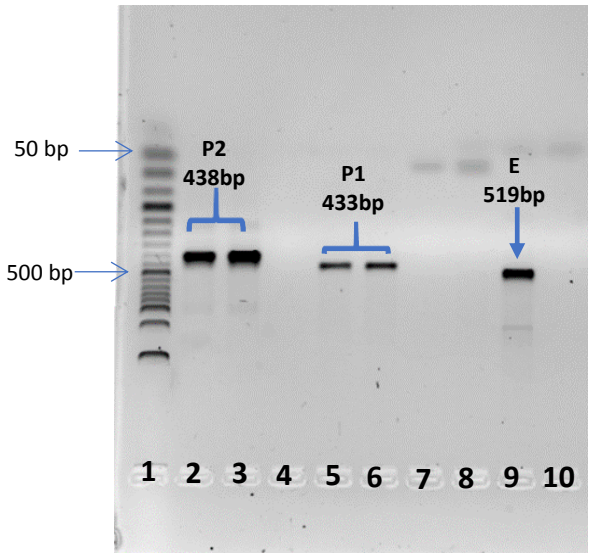


Figure 3.2: Amplification of the PDX-1 promoter region

The Figure is a representative of the PCR products from BS-DNA extracted from healthy PBMCs donors. Lane 1 represents the 50 bp ladder, lanes 2+3 show the PCR product of PDX-1 distal promoter, lane 5+6 represents the proximal promoter of the PDX-1 gene and lane 9 is represents the enhancer region of the PDX-1 gene.

3.2.2. *PDX-1* Promoter Region Sequencing in Islet and PBMC Genome

The next goal was to identify specific CpG sites that are differentially methylated in islet compared with PBMC DNA. This was achieved following the approach used by two leading groups (Akirav et al. 2011; Kuroda et al. 2009). Briefly, extracted DNA was subjected to bisulfite conversion, which was used as a template for PCR amplification with bisulfite specific primers, designed to target the region of interest. The PCR amplicons are then used as a template for Sanger sequencing.

This approach proved to be ineffective as the PCR products failed to produce readable signals. We noticed that extracting DNA from the desired band of the agarose gel results in severely decreasing the DNA concentration. Therefore, we assumed that could be causing the improper sequencing signal. Consequently we tried to troubleshoot the low DNA concentration of the PCR product post gel extraction by amplifying the gel extracted DNA with a whole genome amplification method. This approach also failed to produce good quality sequencing. An alternative strategy was amplifying the gel extracted DNA by cloning in pGEM-T vectors. Although this approach resulted in informative sequencing data, the signals were relatively weak (Figure 3.3 and Figure 3.4).

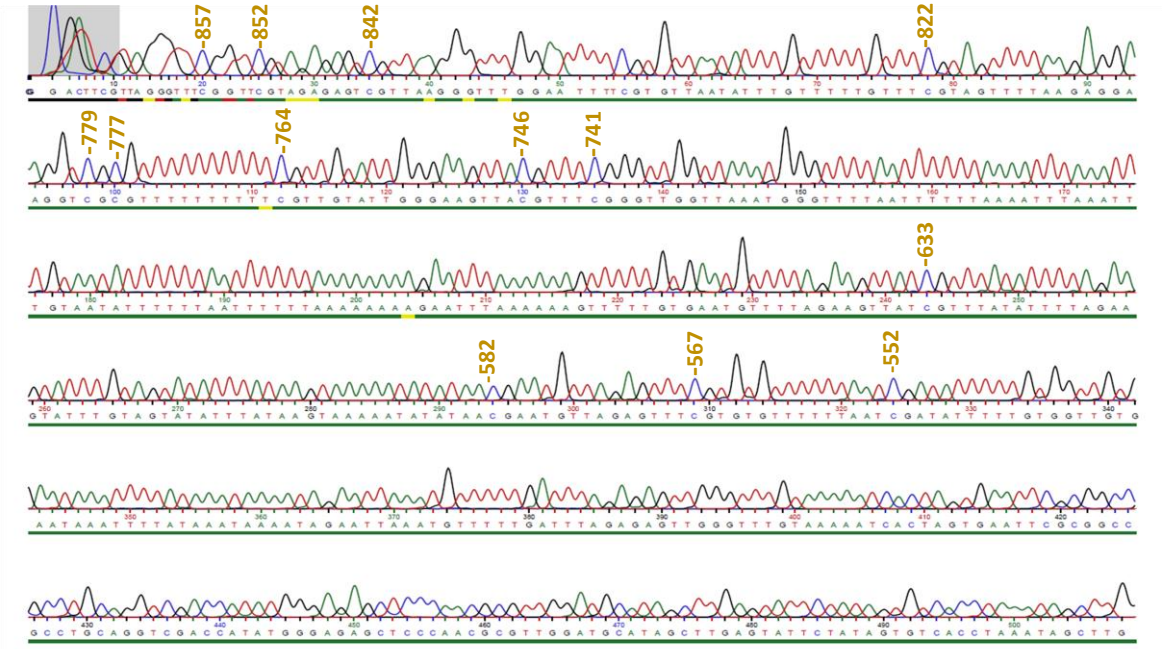


Figure 3.3: A representative example of a distal PDX-1 promoter sequencing chromatogram from one clone of bisulfite treated islet DNA.

The Chromatogram illustrates that most CpG sites within the distal promoter of PDX-1 (-908: -472 bp upstream TSS) are hypermethylated.

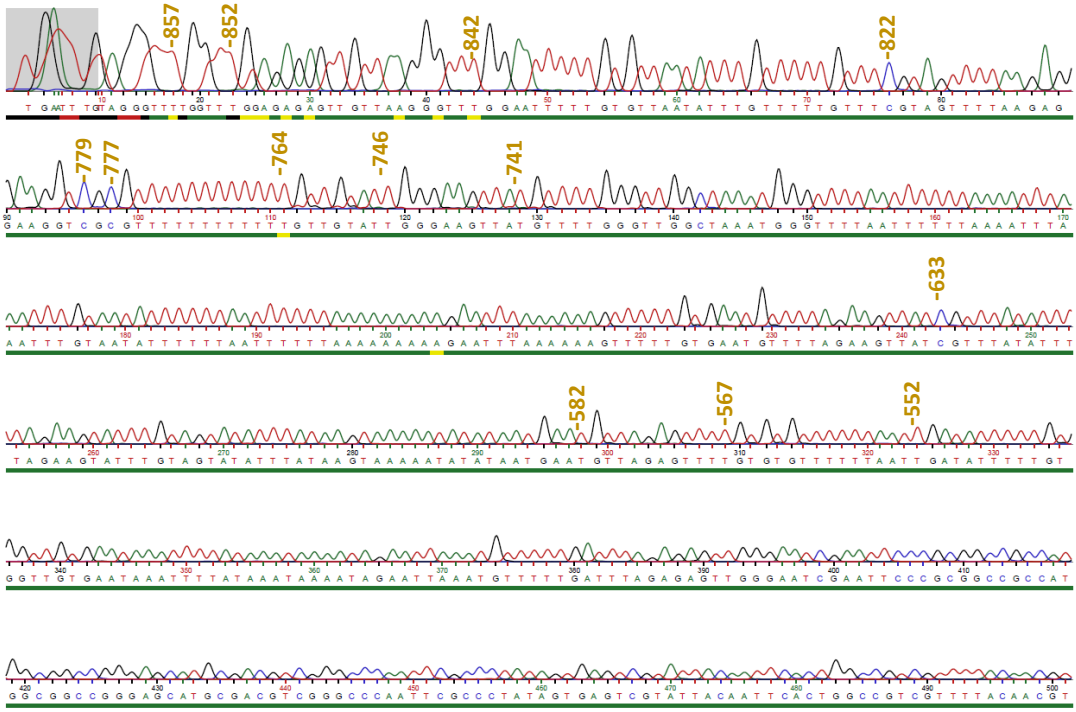


Figure 3.4: A representative example of a distal PDX-1 promoter chromatogram sequencing from one clone of pancreatic PBMC bisulfite treated DNA.

The sequence chromatogram here shows that the majority of the CpG sites within the distal promoter of PDX-1 (-908: -472 bp upstream TSS) are hypomethylated.

The bisulphite sequencing results suggested that CpG -857 (Chr13:28,493,310) was predominantly hypomethylated in the PBMC genome while it was 66% hypermethylated in the islet genome. Acknowledging that islets consist 60-80% beta cells, the 34% hypomethylation could have originated from the non-beta cell genome (e.g. alpha and delta cells). Furthermore, a comprehensive look at the results suggested that there is a similar trend in the methylation status of the neighbouring CpG sites. When CpG -857 is hypomethylated, the neighbouring CpG sites -852, -822, -799, -777, -764, -746 and -741 upstream TSS, were also hypomethylated and vice versa, suggesting that these CpG sites are acting in clusters and are mostly hypomethylated or hypermethylated in tandem (Table 3.1 and Table 3.2). To test this hypothesis, the beta cell DNA sample was used as the template for amplifying the *PDX-1* distal promoter. The PCR products were cloned and bisulfite sequenced as described previously. The beta cell DNA was obtained in small quantities from a historical experiment where freshly obtained islets were FACS sorted following collagenase treatment and Newport Green staining. Twenty-one independent beta cell clones showed that CpG sites -857, -852, -822, -799, -777, -764, -746 and -741 are predominately hypermethylated, together, in the beta cell genome (Table 3.3), confirming our assumption that the 34% hypomethylation signal of CpG-857 observed in the islet sequencing results came from other cells of non-beta cell origin. Collectively the data suggested that CpG -857 is differentially hypermethylated in islet and beta cell genomes when compared to the PBMC genome (Figure 3.5).

Table 3.1: The methylation status of the distal PDX-1 promoter in PBMC genome.

The rows represent the methylation status in each clone. The columns represent the methylation status of each CpG dinucleotide in each clone. Genomic DNA samples obtained from human PBMC were analysed for methylation of the PDX-1 distal promoter (-908 to -472 bp upstream TSS), which contains 14 CpG sites. Note that 0 of 26 (0%) of PBMC clones exhibited CpG hypomethylation. ● Indicates hypermethylated C base, ○ indicates hypomethylated C base.

CpG	-857	-852	-842	-822	-799	-779	-777	-764	-746	-741	-633	-582	-567	-552
Clone 1	○	○	○	○	○	○	○	○	○	○	UKN	UKN	UKN	UKN
Clone 2	○	○	○	○	○	○	○	○	○	○	UKN	UKN	UKN	UKN
Clone 3	○	○	○	○	○	○	○	●	○	○	○	○	○	UKN
Clone 4	○	○	○	○	●	●	●	○	○	○	●	○	○	○
Clone 5	○	○	○	○	○	○	○	○	○	○	UKN	UKN	UKN	UKN
Clone 6	○	○	●	○	●	●	○	○	●	●	○	○	○	○
Clone 7	○	○	○	●	●	○	○	●	●	○	○	○	○	○
Clone 8	○	○	○	○	○	○	●	○	○	○	UKN	UKN	UKN	UKN
Clone 9	○	●	○	○	○	○	●	○	○	○	○	○	○	○
Clone 10	○	○	○	○	○	●	●	○	○	○	●	○	○	○
Clone 11	○	○	○	○	●	UKN	●	●	○	○	○	○	○	○
Clone 12	○	○	○	○	○	○	○	○	○	○	UKN	UKN	UKN	UKN
Clone 13	○	○	○	○	●	○	○	○	○	○	UKN	UKN	UKN	UKN
Clone 14	○	○	○	○	○	○	○	○	○	○	UKN	UKN	UKN	UKN
Clone 15	○	○	○	○	○	○	○	○	○	○	○	UKN	UKN	UKN
Clone 16	○	○	○	○	●	●	●	○	○	○	○	○	○	○
Clone 17	○	○	○	○	○	○	○	○	○	○	UKN	UKN	UKN	UKN
Clone 18	○	○	○	○	●	○	○	○	○	○	○	○	○	○
Clone 19	○	○	○	○	○	○	○	○	○	○	UKN	UKN	UKN	UKN
Clone 20	○	○	○	○	○	○	○	○	○	○	UKN	UKN	UKN	UKN
Clone 21	○	○	○	○	○	UKN	○	UKN	○	○	UKN	UKN	UKN	UKN
Clone 22	○	○	○	○	○	○	○	UKN	○	○	UKN	UKN	UKN	UKN
Clone 23	○	○	○	○	○	●	○	○	○	○	○	○	○	○
Clone 24	○	○	○	○	○	○	○	○	○	○	●	●	○	●
Clone 25	○	○	○	○	UKN	○	○	UKN	UKN	UKN	○	○	○	○
Clone 26	○	○	○	○	○	○	UKN	UKN	○	○	UKN	UKN	UKN	UKN
methylation	**0%**	**4%**	**4%**	**4%**	**28%**	**21%**	**24%**	**14%**	**8%**	**4%**	**23%**	**8%**	**0%**	**9%**

Table 3.2: The methylation status of the distal PDX-1 promoter in the islet genome.

The rows represent the methylation status in each clone. The columns represent the methylation status of each CpG dinucleotide in each clone. Genomic DNA samples obtained from human islet were analysed for methylation of the PDX-1 distal promoter (-908 to -472 bp upstream TSS bp), which contains 14 CpG sites. Note that 10 of 29 clones (34%) of the clones in the islet exhibited CpG hypomethylation, which is consistent with the percentage of non-beta cell content (60-80% beta cell, 40-20% non-beta cell). The orange boxes highlighted the trend in the methylation status of the neighbouring CpG sites. ● *indicates hypermethylated C base,* ○ *indicates hypomethylated C base.*

CpG	-857	-852	-842	-822	-799	-779	-777	-764	-746	-741	-633	-582	-567	-552
Clone 1	●	○	●	●	●	●	●	●	○	●	●	●	●	●
Clone 2	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Clone 3	●	●	●	●	●	●	●	●	○	●	●	○	●	●
Clone 4	○	○	○	○	○	○	○	○	○	○	○	○	○	●
Clone 5	●	●	●	●	●	●	●	○	○	○	●	○	○	●
Clone 6	●	●	○	○	○	○	●	●	●	●	●	○	○	●
Clone 7	UKN	●	○	○	○	○	●	○	○	○	UKN	UKN	UKN	UKN
Clone 8	○	○	○	○	○	●	●	○	UKN	UKN	UKN	UKN	UKN	UKN
Clone 9	○	○	○	○	○	○	○	○	○	○	UKN	UKN	UKN	UKN
Clone 10	○	○	○	○	○	○	○	○	○	○	UKN	UKN	UKN	UKN
Clone 11	●	●	●	●	●	●	●	●	●	●	○	●	●	●
Clone 12	●	●	●	○	●	●	●	●	○	○	○	○	○	○
Clone 13	○	○	●	○	●	●	●	●	●	●	●	●	●	●
Clone 14	○	○	○	○	○	○	○	○	○	○	UKN	UKN	UKN	UKN
Clone 15	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Clone 16	●	●	●	○	●	○	○	●	●	●	○	○	●	●
Clone 17	●	●	●	●	●	●	●	●	●	●	●	●	○	●
Clone 18	○	○	○	○	●	○	○	○	○	○	○	●	●	○
Clone 19	●	●	●	○	●	●	●	○	●	●	●	○	●	●
Clone 20	●	●	●	●	●	○	○	●	●	●	○	○	●	●
Clone 21	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Clone 22	●	○	●	●	●	●	●	●	○	●	●	●	●	●
Clone 23	○	○	○	○	○	○	○	○	○	○	UKN	UKN	UKN	UKN
Clone 24	○	○	○	○	○	○	○	○	○	○	UKN	UKN	UKN	UKN
Clone 25	●	●	●	○	●	●	●	●	●	○	●	●	●	●
Clone 26	○	○	○	○	○	○	○	○	○	○	UKN	UKN	UKN	UKN
Clone 27	●	●	●	●	●	○	○	●	●	●	○	○	●	●
Clone 28	●	●	●	●	●	○	○	●	●	●	○	○	●	●
Clone 29	●	●	●	●	●	●	●	●	○	●	●	●	●	●
Clone 30	●	●	●	●	●	●	●	●	●	●	●	●	●	●
methylation	66%	60%	63%	50%	67%	53%	60%	60%	48%	59%	64%	56%	77%	91%

Table 3.3: The methylation status of the distal PDX-1 promoter in the beta cell genome.

The rows represent the methylation status in each clone. The columns represent the methylation status of each CpG dinucleotide of each clone. Genomic DNA samples obtained from human, FACS sorted, beta cells were analysed for methylation of the PDX-1 distal promoter (-908 to -472 bp upstream TSS), which contains 14 CpG sites. Note that 19 of 20 clones (95%) of the clones in the islet exhibited CpG hypermethylation.

CpG	-857	-852	-842	-822	-799	-779	-777	-764	-746	-741	-633	-582	-567	-552
Clone 1	●	○	●	●	●	●	●	●	○	●	●	●	●	●
Clone 2	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Clone 3	●	●	●	●	●	●	●	●	○	●	●	○	●	●
Clone 4	●	●	●	●	●	●	●	○	○	○	●	○	○	●
Clone 5	●	●	○	○	○	○	●	●	●	●	●	○	○	●
Clone 6	UKN	●	○	○	○	○	●	○	○	○	UKN	UKN	UKN	UKN
Clone 7	●	●	●	●	●	●	●	●	●	●	○	●	●	●
Clone 8	●	●	●	○	●	●	●	●	○	○	○	○	○	○
Clone 9	○	○	●	○	●	●	●	●	●	●	●	●	●	●
Clone 10	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Clone 11	●	●	●	○	●	○	○	●	●	●	○	○	●	●
Clone 12	●	●	●	●	●	●	●	●	●	●	●	●	○	●
Clone 13	●	●	●	○	●	●	●	○	●	●	●	○	●	●
Clone 14	●	●	●	●	●	○	○	●	●	●	○	○	●	●
Clone 15	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Clone 16	●	○	●	●	●	●	●	●	○	●	●	●	●	●
Clone 17	●	●	●	○	●	●	●	●	●	○	●	●	●	●
Clone 18	●	●	●	●	●	○	○	●	●	●	○	○	●	●
Clone 19	●	●	●	●	●	○	○	●	●	●	○	○	●	●
Clone 20	●	●	●	●	●	●	●	●	○	●	●	●	●	●
Clone 21	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Methylation	95%	86%	91%	67%	91%	72%	81%	86%	67%	81%	70%	55%	80%	95%

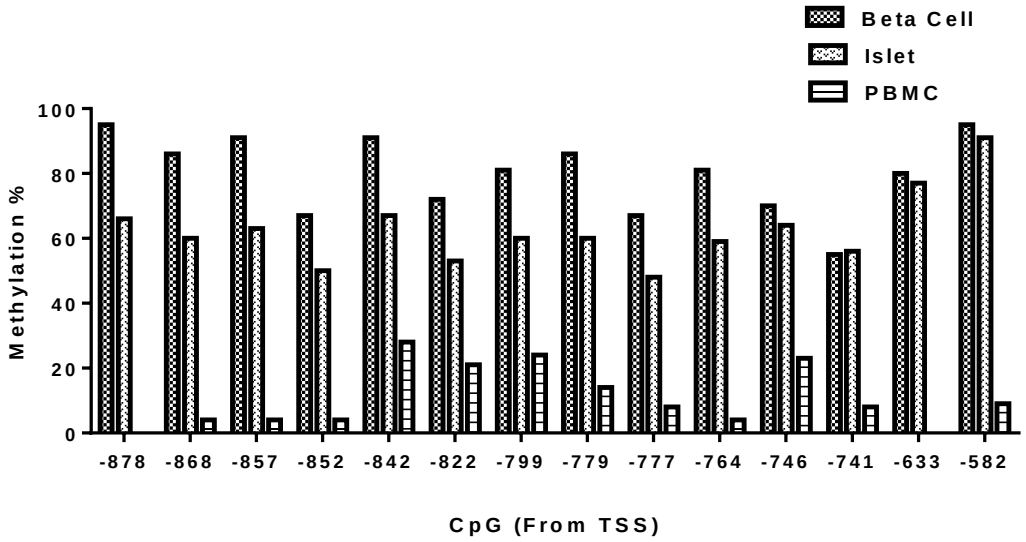


Figure 3.5: The CpG methylation pattern of the distal PDX-1 promoter.

The column represents the hypermethylation percentage of each CpG site within the distal promoter. The result shows that most CpG sites are predominantly hypermethylated (>70%) in the distal promoter region of the PDX-1 .

3.2.3. Developing a *PDX-1* Beta cell-Specific Biomarker

The direct bisulfite sequencing results from PBMC, islet and beta cell genomes indicated that CpG-857, relative to the TSS, could be a potential differentially methylated biomarker candidate. Although CpG -878 and CpG -868 were shown to be potential targets, we needed to have two neighbouring CpG sites that have the similar hypermethylation levels in order to design a stable probe (Technical requirement for probe design). As CpG -857 and CpG -852 were within 5 bp of one another, we decided to choose these sites for targeted probe design. Our next goal was to develop a probe-based assay that amplifies either hypomethylated CpG -857/CpG-852 or hypermethylated CpG -857/ CpG-852 amplicons (as described in General Methods, Chapter2, section 2.5). To fulfill this goal, two differentially methylated dual labelled probes were designed, one of which detected hypermethylated CpG -857 and CpG -852 (labelled with VIC dye) and the other detects the hypomethylated DNA fragments at the same CpG sites (labelled with FAM dye) (The sequence of the primers and probe are listed in Appendix A, Table A.1). The methylation status of these CpG sites was further examined in a panel of locally available tissues: Liver, Kidney, Lung, Heart, Spleen, PBMC, pancreatic islet and human beta cell line (EndoC-βH1) (Figure 3.6 and Figure 3.7). The results indicated that contrary to the data

generated earlier in this study, hypermethylated CpG -857 and CpG -852 are not specific to the beta cell genome. The results show that the CpG -857 is 50% hypermethylated in both liver and islet genome. Counterintuitively, the data also showed that CpG -857 and CpG -852 are hypomethylated in the beta cell line (EndoC βH1) DNA.

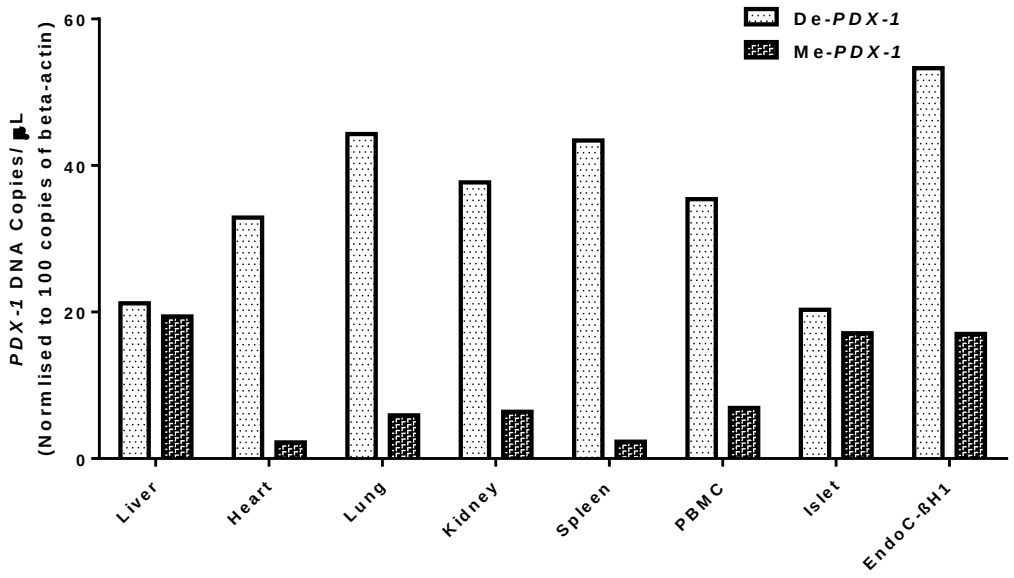


Figure 3.6: The methylation status of CpG -857 of PDX-1 gene (from TSS) in DNA from an "in house" tissue panel.

DNA samples were extracted by QIAamp DNA tissue kit and subjected to bisulfite conversion. Next 5 ng of each DNA sample was analysed by droplet digital PCR utilizing beta-actin and differentially methylated PDX-1 probe assays. The DNA copies in each sample was normalised to 100 copies of beta-actin. The columns represent the DNA copy number per µL in the PCR reaction mix after amplifying 5 ng DNA sample. The result shows that islet and beta cell line genome are not completely hypermethylated in the tested CpG sites. It shows that CpG -857 and CpG -852 are almost 50% hypermethylated in islet and liver genome and they are predominately hypomethylated in beta cell line genome.

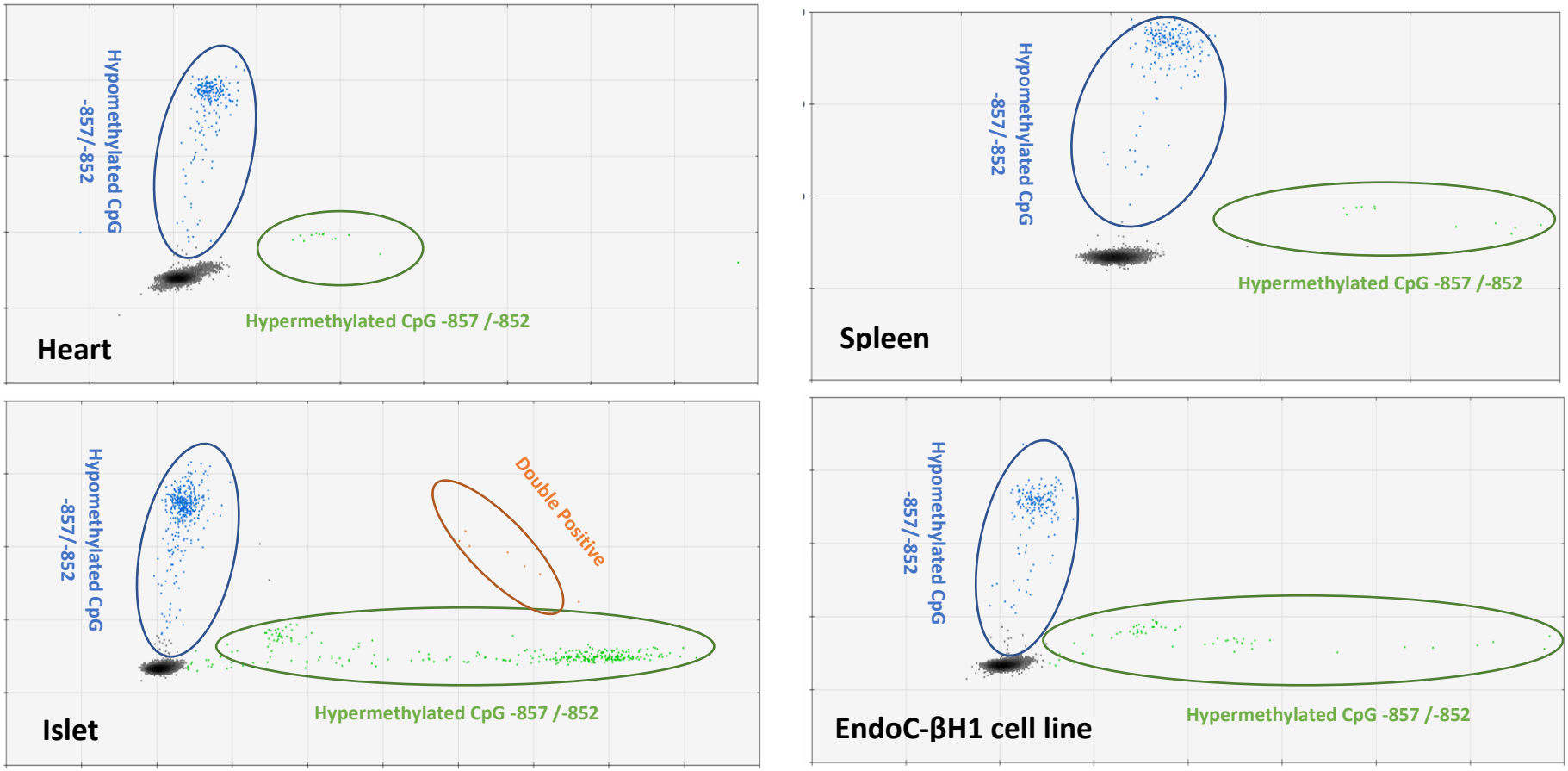


Figure 3.7: 2-D amplification plots from droplet digital data generated for tissue analysis of differentially methylated PDX-1 assay.

The FAM channel represents the hypomethylated DNA copies of CpG -852/-857 (blue cluster) while the VIC channel (green cluster) represents the hypermethylated DNA copies of CpG -852/-857. The data here reflects figure 3.6 showing that CpG sites -852/-857 are non-specifically hypermethylated in islets and beta cell model.

3.2.4. Defining the Threshold for Droplet Digital PCR Experiments

During this project, ddPCR data analysis was viewed as the 1-D plot in single plex probe/EvaGreen experiments or 2-D plot in duplex and multiplex probe experiments. In 1-D plots (Figure 3.8 A &B), droplets were plotted in a graph of fluorescent intensity vs droplet number. All positive droplets, those above the pink threshold line are considered as "positive", and each is assigned a value of 1. All negative droplets, those below the threshold line are scored as negative with score value of 0. ddPCR data from duplex (two targets amplified in the same reaction) and triplex (three targets amplified in the same reaction) assays were viewed in a 2-D plot where channel 1 fluorescence (FAM) is plotted against channel 2 fluorescence (VIC/HEX) for each sample. The threshold in these experiments was adjusted manually across an entire plate in comparison to their parallel cluster position in positive (synthesised DNA standard) and negative (no template) samples (Figure 3.9 and Figure 3.10).

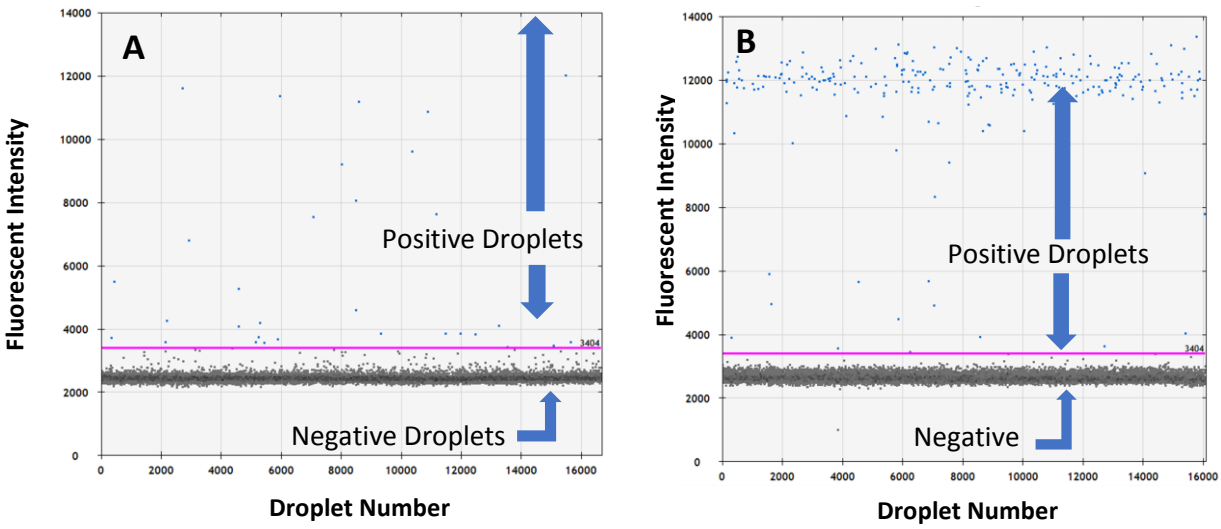


Figure 3.8: Defining the threshold in 1-plex an EVAGreen ddPCR experiment run on the same plate.

The blue dots represent the positive amplified droplets; the grey dots represent the negative droplets. A) 1-D plot of a test sample amplified by Evagreen dye. B) 1-D plot of a positive control sample amplified with Evagreen dye.

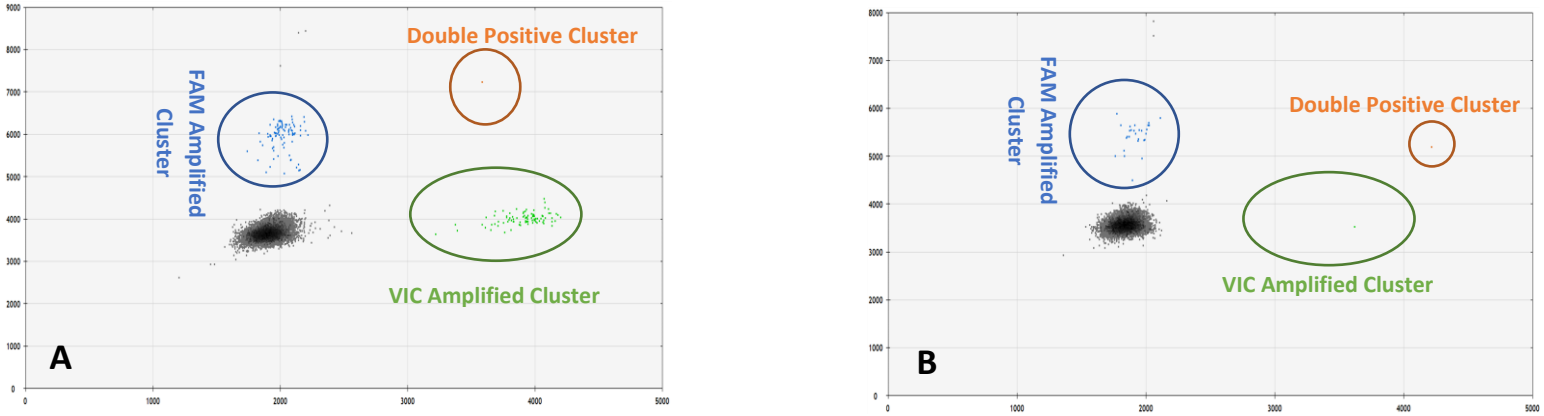


Figure 3.9: Defining the threshold in 1-plex an EVAGreen ddPCR experiment analysed on the same plate.

The plots represent an experiment of amplifying a multiplex assay of the duplex probes. The Blue cluster represents FAM amplified probe, the green cluster represents the VIC amplified probe and the orange cluster represents the double positive (FAM + VIC) amplified probes. The grey cluster represents the negative empty droplets. A) 2-D plot of a duplex-probe assay amplifying a positive control DNA template. B) 2-D plot of a duplex-probe assay amplifying a test sample. Both samples were run on the same plate.

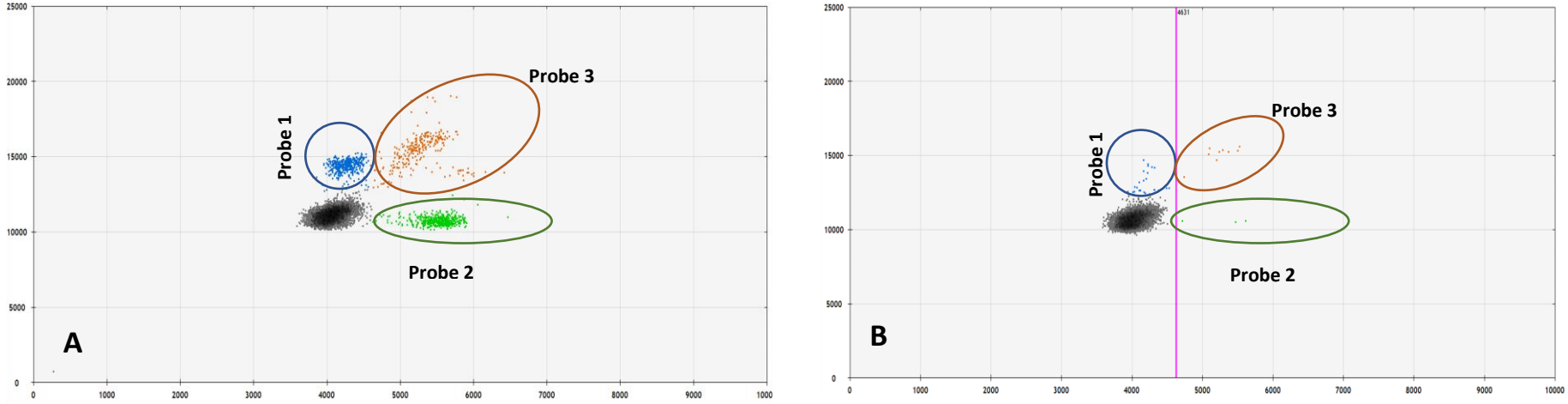


Figure 3.10: 2-D Plot of droplets of Fluorescence amplified multiplex assay.

The plots represent an experiment of amplifying a multiplex assay of the triple probes. The Blue cluster represents FAM amplified probe 1, the green cluster represents the VIC amplified probe 2, the orange cluster represents the amplified (70 %FAM + 30 %VIC) of probe 3 and the grey cluster represent the negative empty droplets. A) 2-D plot of a triplex-probe assay amplifying a positive control DNA template. B) 2-D plot of a triplex-probe assay amplifying an unknown sample. Both samples were run on the same plate.

Discussion

The *INS* gene was the first target for developing DMR based biomarkers for monitoring beta cell death. It was a logical choice because *INS* is expressed exclusively in beta cells and it has an essential role in beta cell function. The first year of this PhD was, therefore, dedicated to identifying DMR with other genes that are known to be vital for the specific function of beta cells such as the *genes for PDX-1, NeuroD1* and *MafA*. To achieve this aim, we followed the approach used by two leading groups in the field (Akirav et al. 2011, Husseiny et al. 2014, Husseiny et al. 2012). In summary, they isolated genomic DNA from islet cells and other control tissues, e.g. kidney, liver, PBMC, and the DNA was bisulfite treated and sequenced by the Sanger method for verification and identification of methylation differences at specific CpG sites. Based on the results obtained, the bisulfite specific primer set and methylation specific primers were designed for a nested PCR approach. We chose to replace the nested PCR method with a probe-based ddPCR technique because of the enhanced sensitivity of this approach. Generally, healthy individuals have a low level of circulating cfDNA due to the efficient cfDNA clearance from the circulation by phagocytes (Fleischhacker and Schmidt 2007). Usually, cfDNA is eliminated from the periphery by phagocytosis, kidney and liver and has a half-life of 10 - 15 minutes (Khier and Lohan 2018, Rumore et al. 1992, Schwarzenbach et al. 2011). Conversely, due to rapid destruction, the level of cfDNA is dramatically increased in the case of diseases such as cancer and developing autoimmunity (Jiang and Lo 2016). Nevertheless, the methylation level in other tissues was studied to eliminate possible false positive results that could be generated due to cell turnover.

Our initial investigations focused on *PDX-1,* Although Yang *et al*. have previously described the DNA methylation of *PDX-1* promoters and enhancer, he did not compare it to other non-islet tissues. Their research aimed to compare the *PDX-1* methylation pattern of type 2 diabetic population to normal controls (Yang et al. 2012a). They showed that in islets, the methylation of CpG islets within the distal promoter in a healthy population is 40-60% hypermethylated in the CpG islet: -857, -852, -

822, -799, -777, -764, -746 and -741, relative to TSS. We were able to confirm their published results in our healthy population. Here we were aiming to compare the methylation profile of *PDX-1* in pancreatic islets to other tissues such as PBMC, spleen, liver and kidney tissue in healthy individuals. Among the predominantly hypermethylated CpG sites, we chose to design a differentially methylated assay based on CpG-857 and CpG -852 as a stable probe design for ddPCR requires including more than one SNP (based on initial optimising experiments). The data showed that hypermethylated CpG -852/-857 was not specific for the islet and beta cell genome as it has a similar methylation level in liver.

PDX-1 expression is also essential for the development of the proximal duodenum and maintenance of the duodenal enterocytes and entero-endocrine cells, gastro-duodenal junction and Brunner's glands (Chen et al. 2009, Holland et al. 2013). Therefore, it was necessary to study the methylation of targeted CpG sites in the duodenum genome and compare it to the methylation level in islets. Unfortunately, the available in-house human tissues do not include the duodenum, but as the CpG -857 was shown not to be unique for the islet genome, further study of CpG -857 methylation levels in the duodenum genome was not further investigated.

Overall, we concluded that while using the Sanger bisulfite sequencing method to identify novel differentially methylated CpG sites can work, it has proven to be very laborious and time-consuming for a time-limited project, and a weakness of the approach is that the best potential biomarkers to measure beta cell death may be missed. It was therefore decided to adopt a hypothesis-free, genome-wide methylation sequencing approach: the Illumina Infinium® Methylation EPIC BeadChip to identify novel DMR to compare methylation profiles in islets with PBMC and to develop assays based on the targets identified.

Chapter 4: Developing Biomarkers for Monitoring Beta cell Death

4. Introduction

Bisulfite genomic sequencing of multiple clones of genes is considered to be the gold standard method to study the CpG methylation status of the underlying DNA sequence. Nevertheless, this strategy proved to be impractical for obtaining results for whole-genome bisulfite sequences as it requires huge efforts in time, labour, and budget. DNA methylation microarrays are an alternative strategy for studying genome-wide methylation of DNA and now are widely adopted in many genome-wide association studies. The Illumina Infinium Human Methylation EPIC BeadChip array (EPIC) involves hybridisation of bisulfite converted DNA into BeadChip array. The array generates a methylation profile of 853,307 CpG sites distributed among the 22 autosomes as well as the two sex chromosomes (X & Y chromosomes) per sample at single-nucleotide resolution (Moran et al. 2016, Pidsley et al. 2016). The beadchip can analyse up to 8 samples per array and only requires 250 ng of bisulfite-converted DNA per sample.

In this chapter, the Illumina Infinium Human Methylation EPIC BeadChip was used to identify loci that are uniquely hypo-/hyper-methylated in the pancreatic islet. As the differentially methylated biomarker would ultimately be used in peripheral samples, and as the main source of genomic contamination in the peripheral samples is the DNA from PBMC, PBMC DNA samples were also included in the EPIC array. Genes that have differentially methylated CpG regions and reached genome-wide significance (p < 10E-10) were considered to be potential targets for differentially methylated biomarker design. Among these target CpG sites, we decided to focus on those which lay within beta cell-specific genes only. Then the methylation level of the chosen CpG targets will be compared in a variety of other tissues (data conducted from public data; ENCODE annotation data) in order to rule out contamination of serum with genomic DNA resulting from normal cell turnover. Finally, the results will be confirmed by testing the methylation level of the target CpG site in a locally available tissue panel. Based on these results, methylation sensitive dual-probe assays were designed for the approved target genes.

Furthermore, an interesting differentially methylated locus identified by EPIC array; the mutual promoter of the adjacent genes *MIR-200c*/*MIR-141,* was further studied using a high throughput methylation sequencing technique: The targeted Next-Generation Bisulfite Sequencing (tNGBS). The tNGBS sequencing allows the study of specific regions or genes of interest with higher quality data because of the significantly higher coverage of each single CpG site within the region of interest. This high throughput technique allows measurement of the methylation level percentage of several target regions and multiple samples in a single run. Here, the tNGBS was performed to assess the methylation pattern of all CpG dinucleotides in the promoter region of *MIR-200c/MIR-141* in the beta cell line genome (EndoC-βH1) as the whole genome methylation sequencing, EPIC array was performed on an islet genome. Further, we intended to examine the effect of high glucose exposure on the methylation level of CpG loci within the *MIR-200c/MIR-141* promoter.

Methods

4.1.1. The Illumina Infinium Human Methylation EPIC 850K BeadChip Array

The Illumina Infinium EPIC array is a methylation sequencing platform that uses bead technology to perform high-resolution multiplexed methylation sequencing for > 850,000 CpGs in enhancer regions, gene bodies, promoters, and CpG islands. The protocol involves bisulfite conversion and whole genome amplification of genomic DNA, followed by hybridisation of the converted DNA to arrays containing predesigned probes that discriminate hypermethylated from hypomethylated cytosine nucleotides. The bead-chip comprised two bead types, Infinium type I and type II (Figure 4.1), each has a 50 bp specific probe that is complementary to specific human bisulphite converted DNA with a CpG site at the 3' end of the probe. Once hybridised to the specific region, single base extension occurs assembling a sequence of fluorescent-labelled ddNTPs at the 3' end of the CpG site. The majority of the probes are designed to target CpG sites within the gene body (35.9%), 32.3% of the probes are directed to CpG sites located in the promoter area (9.4% of the probes are within the proximal promoter[10], 13.9% are within the distal promoter[11] of these genes; Figure 4.2A). Genomically, the targeted sites include 9.5% on chromosome 1, the rest of the probes are shown to be distributed almost equally amongst the autosomal chromosomes (Figure 4.2B) (Moran et al. 2016; Pidsley et al. 2016). The fluorescent signal (beta value "β") is calculated as the ratio of hypermethylated (C) to hypo-methylated (T) signal using the following equation:

$$\beta = \frac{\text{The Intensity of the Hypermethylated Signal}}{\text{The Intensity of Hypomethylated Signal} + \text{The Intensity of Hypermethylated Signal} + 100}$$

The completely hypermethylated site is defined as β = 1, whereas the fully hypomethylated sites are defined as β = 0.

[10] Proximal Promoter: DNA region located 200 bp upstream of TSS
[11] Distal promoter: DNA region located 1500 bp upstream of transcription start sites

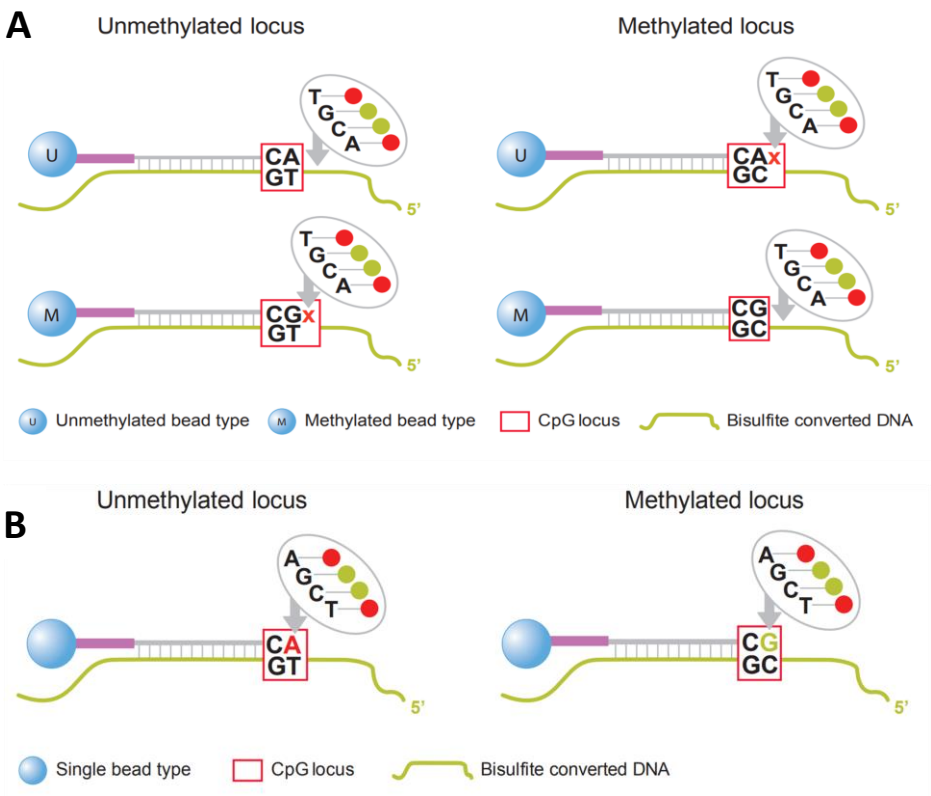


Figure 4.1: Illumina Infinium Type I and Type II Probes.

A) The Illumina Infinium Type I probe has two separate probe versions per CpG site, a probe for a hypermethylated sequence and another for hypomethylated CpG sites. B) The Illumina Infinium Type II probe has only one probe sequence specific for a CpG site. Each probe type has an advantage over the other. Compared to type II probes, Type I probes use the double physical space on the Bead Chip but its design (with two sequence versions) is more efficient in sequencing the high density CpG regions.

Embedded from: Illumina.com (https://www.illumina.com/content/dam/illumina-marketing/documents/products/datasheets/humanmethylationepic-data-sheet-1070-2015-008.pdf)

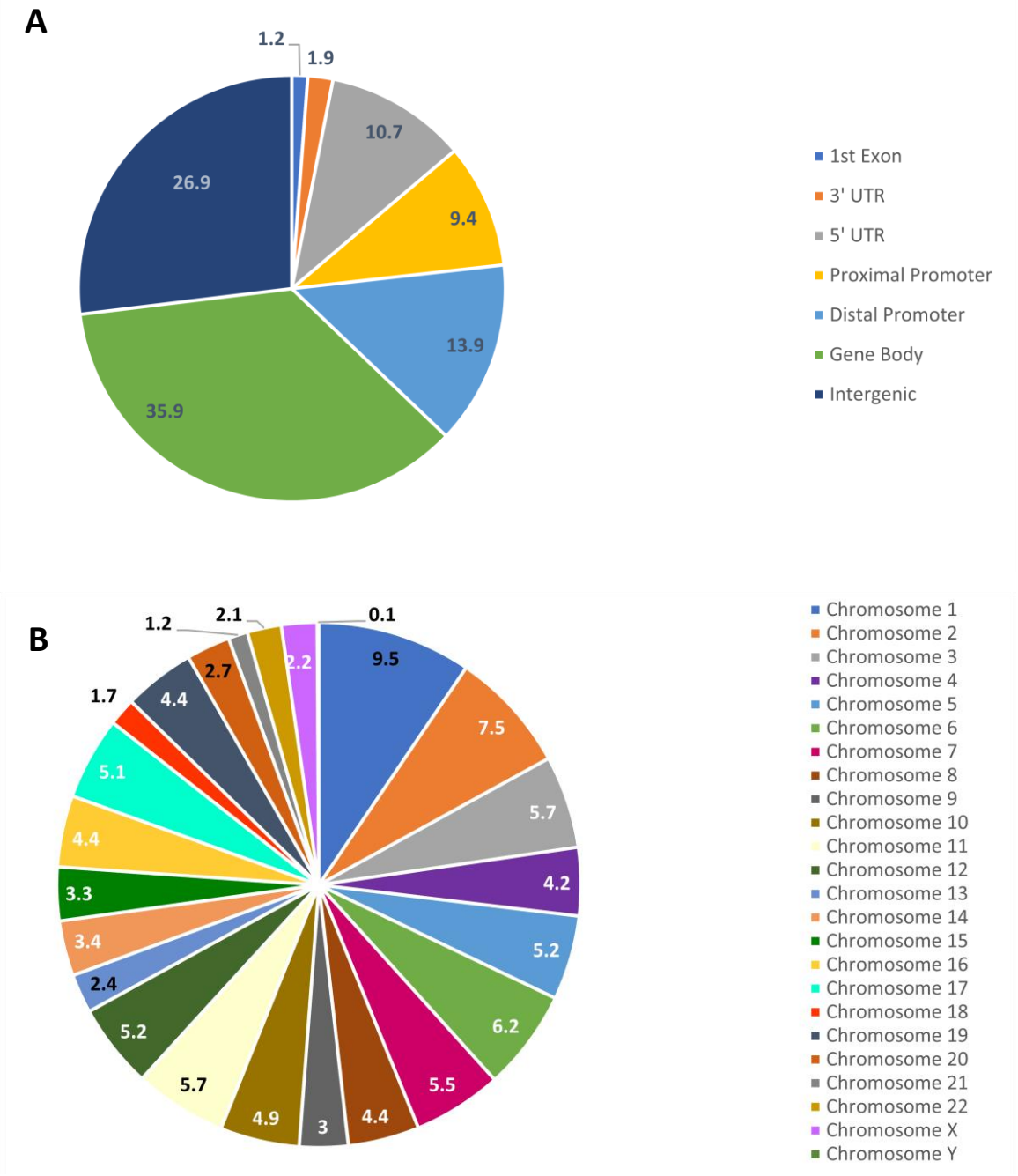


Figure 4.2: The chromosomal distribution of the probes in the Methylation EPIC BeadChip.
A) The distribution of EPIC array probes in gene regions. B) The genomic distribution of the probes location across the chromosomes. The data is based on the information introduced in (Pidsley et al. 2016)

4.1.1.1. Sample Processing for the EPIC Array

DNA extraction of the samples used in the array was carried out in-house (University of Bristol). The purity of the DNA was confirmed by measuring the sample's absorbance spectrum using the NanoPhotometer® (IMPLEN, P330) and calculating the A_{260}/A_{280} ratio which was in the range of 1.8 and 2.0. The sample concentration was quantified using the QuantiFluor® dsDNA System. Then 500 ng of genomic DNA (20 ng/ 45 µL H_2O) was submitted to AROS Applied Biotechnology A/S (Aarhus N, Denmark) where the sample was bisulfite converted and whole genome methylation sequenced.

Bisulfite conversion of genomic DNA carried out using the EZ DNA Methylation Kit (Zymo Research). The PCR conditions for bisulfite conversion were as follow (95^{o}C for 30 sec., 50^{o}C for 60 minutes) for 16 cycles, then "hold" at 4^{o}C. Immediately following elution, all samples were processed through the EPIC array protocol (Figure 4.3). Briefly, 7 µL of converted DNA was denatured with 1 µL of 0.1N sodium hydroxide before whole genome amplification was performed on the MSA4 plate. All other steps were carried out as recommended by the manufacturer's protocol. Array bead chips were scanned on Illumina iScan or HiScan system (Qiagen).

4.1.1.2. EPIC Array Data Normalisation and Statistics

The resulting raw data (IDATs) were represented as a beta value, as mentioned earlier, of fluorescence intensity ratio for each CpG site between 0 (unmethylated) and 1 (fully methylated). IDATs were further analysed by Dr Matthew Suderman (Lecturer in Genetic Epidemiology, School of Social and Community Medicine, University of Bristol). Data were normalised, and background corrected using the functional normalisation (Fortin et al. 2014). The DMRs were identified using the meffil R package (Ritchie et al. 2015).

4.1.1.3. Bioinformatics Analysis of Microarray Data and Data Visualising (Heat map)

The data visualisation was done in collaboration with Dr Jody Ye (the University of Bristol, Bristol/Albert Einstein College of Medicine, New York). The methylome signature of liver, aorta and spleen (Roadmap study), as well as pancreas (ENCODE study), were downloaded. The R 3.2.2 (x64 bit)

statistical programme (http://www.r-project.org) associated with packages including ggplot2 (Wickham 2009), and reshape2 (Wickham 2007) were used to compare the methylation level of islet and PBMC methylome (EPIC data) to liver, aorta, spleen and pancreas. Using R, scripts of tissues methylome were saved in comma-separated values format (.csv). After installing and loading the appropriate R package, .csv files were loaded and read into R. The scripts were then applied to the data; heatmaps were exported as a Portable Network Graphic format (.png) file.

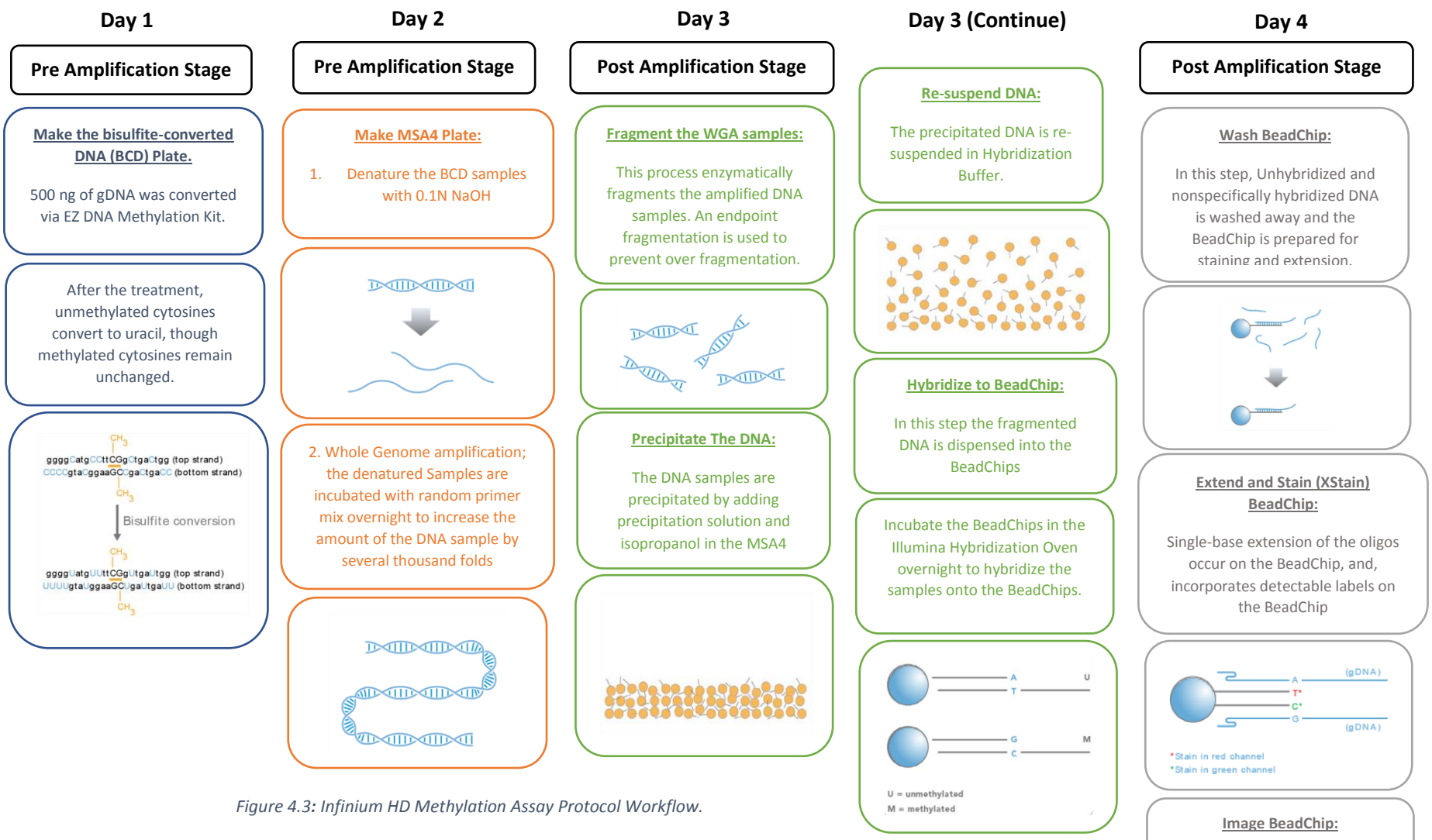


Figure 4.3: Infinium HD Methylation Assay Protocol Workflow.

The following figure graphically represents the Infinium HD Methylation Assay protocol as recommended by the manufacturer. The figures were adopted from the Infinium Methylation Assay manual (https://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/infinium_assays/infinium_hd_methylation/infinium-hd-methylation-guide-15019519-01.pdf)

4.1.2. Targeted Next-Generation Bisulfite Sequencing (tNGBS):

Human *MIR-200c* and *MIR-141* genes share a promoter with an Ensembl regulatory built ID: ENSR00000048377 (Chr12: 6,963,000-6,964,001) (Figure 4.4). Seven bisulfite sequencing assays were designed to cover this 1002-bp region, in addition to a region around the transcriptional start.

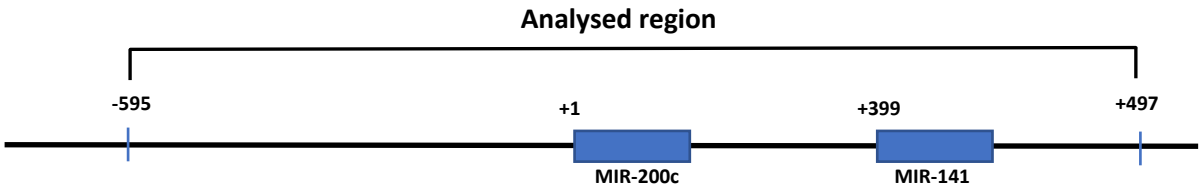


Figure 4.4: The analysed region by targeted next generation bisulfite sequencing technology.

4.1.2.1. DNA Isolation and Bisulfite Conversion for tNGBS:

Genomic DNA was isolated in-house from tissue and cell lines (described in General methods, Chapter 2, Section 2.1.2) using the QiaAmp Blood and Tissue Kit (Qiagen). Next, 500 ng DNA samples (40 ng/µL) were shipped to EpigenDX (Hopkinton, Massachusetts, USA) where samples were sodium bisulfite treated using the EZ DNA Methylation Kit (Zymo Research). Nucleotide sequences for the human *MIR-141* gene (Gene ID: 406933) were obtained from GenBank. tNGBS was used to assess the methylome at the targeted DNA region extracted from multiple cells/ tissues.

4.1.2.2. Multiplex PCR Optimisation for tNGBS

All assays were initially divided into multiple groups based on the amplicon size, primer Tm, GC content, and ΔG values, while also avoiding overlapping primer pairs. A Gradient PCR was performed at two different magnesium concentrations (1.5 mM and 3.0 mM), followed by capillary electrophoresis (CE) of the PCR products using the QIAxcel Advanced System (Qiagen). Final multiplex PCR conditions were determined based on the optimal annealing temperature (Ta), magnesium concentration, and overall amplification efficiency. Based on the CE results, 8 PCR amplicons were grouped into three simplex PCRs, one duplex PCR, and one triplex PCR.

4.1.2.3. tNGBS PCR Protocol

For PCR amplification, the bisulfite-converted DNA was PCR amplified in a 20 µL reaction as following, 0.5 U of HotStar Taq Polymerase (Qiagen), 2 µL 10× PCR buffer (Contains 15 mM $MgCl_2$), 200 µM dNTPs, 4 pmoL of each Primer pair and 1-5 µL of bisulfite treated DNA. The cycling condition for the reaction was 95°C for 15 minutes; 45 x (95°C 30s; proper annealing temperature 30 s; 68°C 30s); 68°C for 5 minute and 4°C forever.

4.1.2.4. tNGBS Library Preparation and Sequencing

Libraries were prepared using the KAPA Library Preparation Kit for Ion Torrent platforms (Cat# KK8310) and Ion Xpress™ Barcode Adapters (Thermo Fisher, Table 4.1). Next, library molecules were purified using Agencourt AMPure XP beads (Beckman Coulter) and quantified using the Qiagen QIAxcel Advanced System. Barcoded samples were then pooled in an equimolar fashion before template preparation and enrichment were performed on the Ion Chef™ system (Thermo Fisher) using Ion 520™ & Ion 530™ Chef reagents. Following this, enriched, template-positive library molecules were then sequenced on the Ion S5™ sequencer using Ion 530™ sequencing chips (Thermo Fisher).

Table 4.1: The DNA target region sequenced by the targeted bisulfite-sequencing assay.

Seven assays were designed to cover the MIR-200c/MIR-141 mutual promoter region as well as MIR-141 exon (Chr12: 6,963,000-6,964,001).

Assay ID	Assay Location	Location From TSS	GRCh38/hg38	# of CpGs
ADS5466	5-Upstream of *MIR-200c*& *MIR-141*	-595 to -586	chr12:6963104-6963113	2
ADS5467		-524 to -359	chr12:6963175-6963340	12
ADS5465		-280 to -135	chr12:6963419-6963564	14
ADS5468		-97	chr12:6963602	5
ADS995- FS1	5-Upstream of *MIR-200c* & *MIR-141* & *MIR-2000C* Exon	-56 to +30	chr12:6963643-6963728	5
ADS995BFS2	*MIR-200c* Exon & 5'-Upstream of *MIR-141*	+53 to +166	chr12:6963751-6963864	5
ADS996	5'-Upstream of *MIR-141* & *MIR-141* Exon	+261 to +403	chr12:6963959-6964101	8
ADS5469	*MIR-141* Exon & Downstream of *MIR-200c/MIR-141*	+483 to +580	chr12:6964181-6964278	3

4.1.2.5. tNGBS Data Analysis

FASTQ files from the Ion Torrent S5 server were aligned to the local reference database using open-source Bismark Bisulfite Read Mapper with the Bowtie2 alignment algorithm. Methylation levels were calculated in Bismark by dividing the number of methylated reads by the total number of reads.

4.1.3. Universal Human Methylated DNA Standard Controls

A predominately hypermethylated and hypomethylated human DNA set (Zymo Research) was used to generate a standard methylation curve that was used in assay optimisation experiments.

4.1.4. Droplet Generating and PCR Cycle Condition

Once samples and appropriate oil was loaded in the cartridge, the DG8 cartridges were covered by the Gaskets (Bio-Rad) and then placed in a QX200™ Droplet Generator (Bio-Rad). Next, the droplets were gently transferred into a semi-skirted and PCR clean 96-well PCR plate (Eppendorf) using an E4 XLS Multichannel pipette with the range 5-50 µL (Rainin). The PCR plate was sealed with pierceable foil (Bio-Rad) using a PX1™ PCR Plate Sealer (Bio-Rad). Afterwards, the PCR amplification was performed by placing the plate in a C1000 Touch™ Thermal Cycler (Bio-Rad). For the probe-based droplet PCR, the **standard PCR protocol** was used in the single plex and duplex experiment (Figure 4.5) and **touch down PCR** was used in the multiplex experiments (Figure 4.6). The Touchdown PCR is a modified PCR in which the initial annealing temperature is higher than the optimal temperature and is gradually decreased in the subsequent cycles until reaching the optimal temperature. Amplification is then continued using this annealing temperature (Don et al. 1991). The reaction plates were then held at 4°C until reading. Droplet reading and analysis were performed with the QX™ 200 Droplet reader (Bio-Rad) via ddPCR™ Droplet Reader Oil (Bio-Rad). The data analyses were performed using the software package QuantaSoft™ (Bio-Rad, V 1.6.6.0320). The threshold of fluorescence, which distinguishes the positive from the negative droplets, was set manually. The same threshold was applied to all the wells of one PCR plate. The average droplet number that was accepted for the valid measurement was around 17,000.

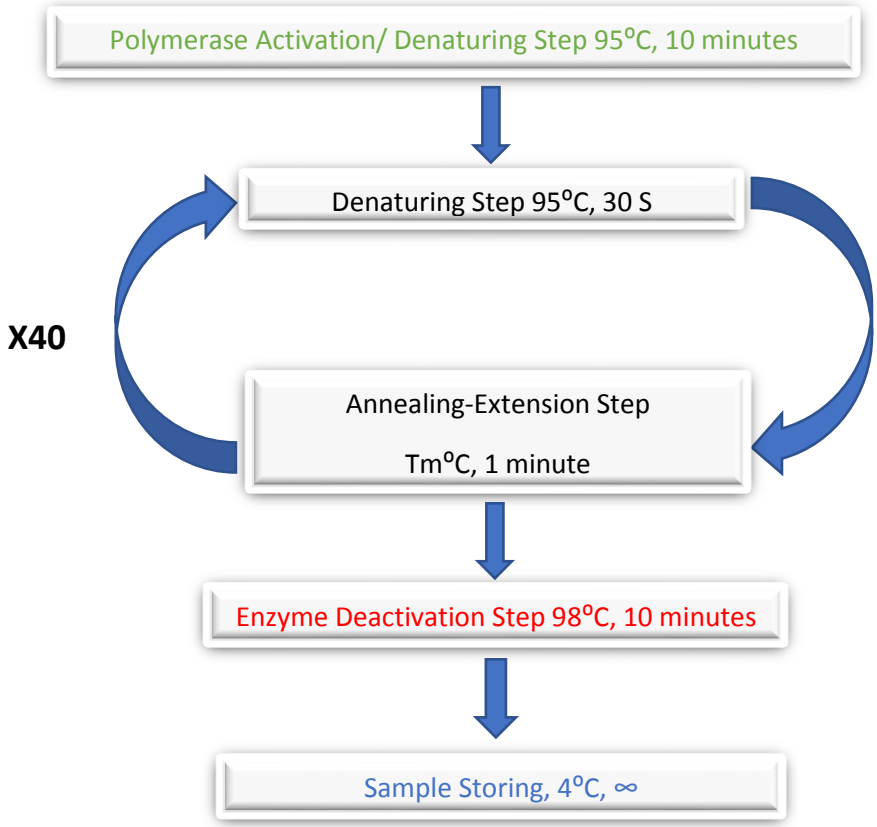


Figure 4.5: The Standard PCR Protocol.

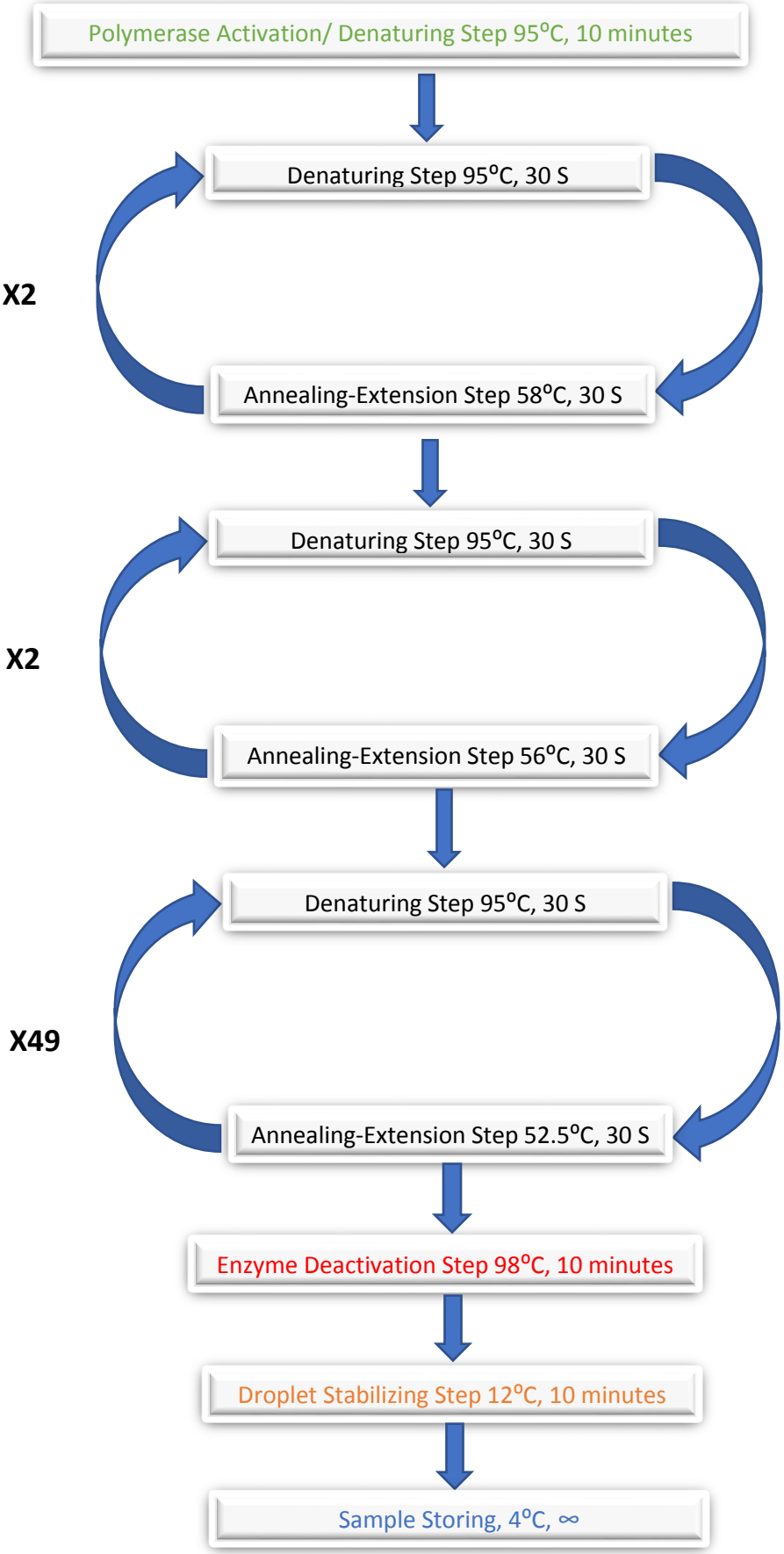


Figure 4.6: The Touchdown PCR Protocol.

4.1.5. Methylation Data Alignment and Statistical Analysis

In all candidate CpG sites, the methylation state of the candidate CpG sites in multiple tissues were aligned for comparison using the Integrated Genome Browser software (Nicol et al. 2009) (http://bioviz.org/). The number of copies per 1 µL of the target was analysed by the Quantasoft software (Bio-Rad, version 1.6.6.0320). The direct comparison between the levels of circulating hypomethylated beta cell-specific targets in test VS control group was achieved by Mann-Whitney testing for nonparametric data. For all statistical comparisons, a *p* value of < 0.05 was considered to be statistically significant. Statistical analyses were performed with GraphPad/Prism 6 (version 6.07).

The relative methylation/unmethylation level in DNA extracted from tissues was determined by normalising input DNA against the beta-actin gene as follows.

$$\text{DNA}\,\frac{\text{copies}}{\mu\text{L}}\,(\text{in 5 ng DNA sample}) = \frac{\text{gene DNA copies per 1 µL}}{Beta\ actin\ DNA\ copies\ per\ 1\ \mu L} * 100 * 25$$

The absolute level of beta cell-specific targets cfDNA in this thesis has been presented as the cfDNA copies in 1 mL of serum/plasma sample. The level of circulating hypomethylated DNA was calculated as follows:

$$\underline{\text{Number of copies in 1 mL sample}} =$$

$$\frac{\text{Copies}}{\mu\text{L}}\,\text{in duplicate 1} + \frac{\text{Copies}}{\mu\text{L}}\,\text{in duplicate 2} * 25 * \left(\frac{1000}{\text{The volume of original sample (µL)}}\right)$$

Results

4.2.1. The Illumina Infinium® Methylation EPIC BeadChip Analysis

In order to gain a more comprehensive insight into the methylation pattern of the islet genome, the methylome profile of islet tissue samples from five brain-stem death adult donors was compared to the methylome of PBMC samples obtained from three healthy adult donors. The methylation signature of both tissues was profiled in the Illumina methylation EPIC array. The analysis was performed using meffil package in R software. Quality control and normalisation of the EPIC array were performed and resulted in the final probe set of 867,926 of a total 867,926 (100%). Genome-wide comparison of methylation signature between islets (*n*=5) and PBMC (*n*=3) at the Bonferroni threshold of 5.9e-8 identified 8856 CpG sites that reached epigenome-wide significance (P<10E-10). The analysis was not based on gender or batch variability of the samples. Of DMR identified in the islet genome, 3866 CpGs were hypomethylated, and 4990 were hypermethylated (Figure 4.7A). Another comparison was made within islet-DMS, which showed that the majority of the identified DMS (37%) were located within the gene body, 15% were located at the proximal or distal promoters (Figure 4.7B).

The methylation profile of islet and PBMC genomes were further compared to the methylation signature data of liver, lung, thymus, spleen, pancreas as well as aorta that were identified from the Roadmap and ENCODE methylation signature studies (Figure 4.8). Aligning data in Heatmap showed that the overall methylome of islets and pancreas are opposite to those in PBMC, thymus and spleen. The data also showed that the methylome of liver, lung and aorta had a considerable level of homogeneity to the methylation level of the islet genome. Selecting for islet differentially methylated CpG sites versus other non-beta cell tissue methylomes had reduced the number of islet-specific hypomethylated CpG sites into 425 CpG targets (Appendix A, Table A.3) and islet-specific hypermethylated CpG targets into 228 (Appendix A, Table A.4).

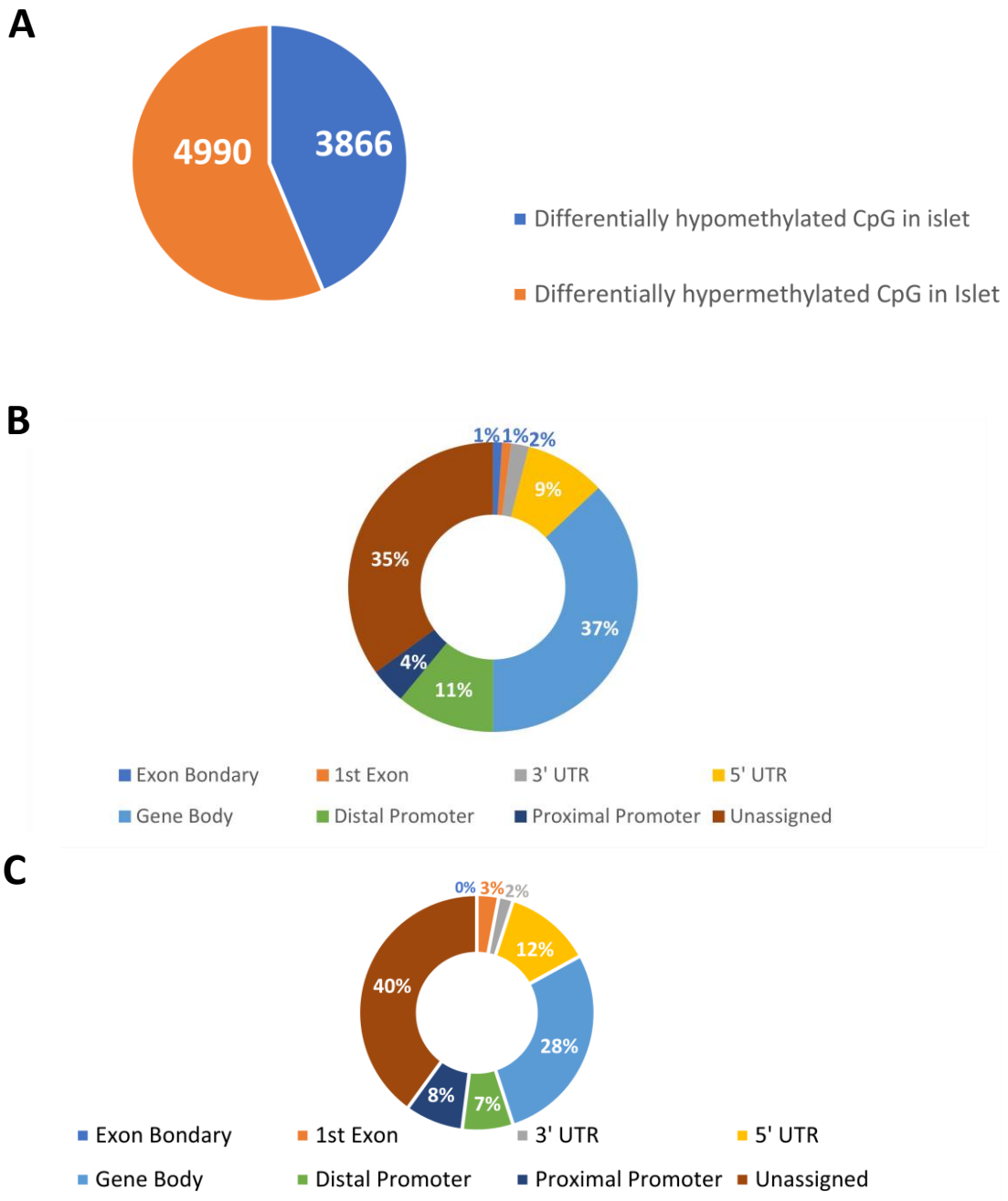


Figure 4.7: Genomic and functional details of the analysed CpG sites in the Methylation EPIC 850K microarray. A) The quantification of hypomethylated vs hypermethylated CpG sites in the islet genome; B) and C) UCSC gene region feature categories (proximal promoter, distal promoter, 5´UTR, 1st Exon, Gene Body, 3´UTR) of the hypomethylated CpG sites (B) and the hypermethylated sites (C).

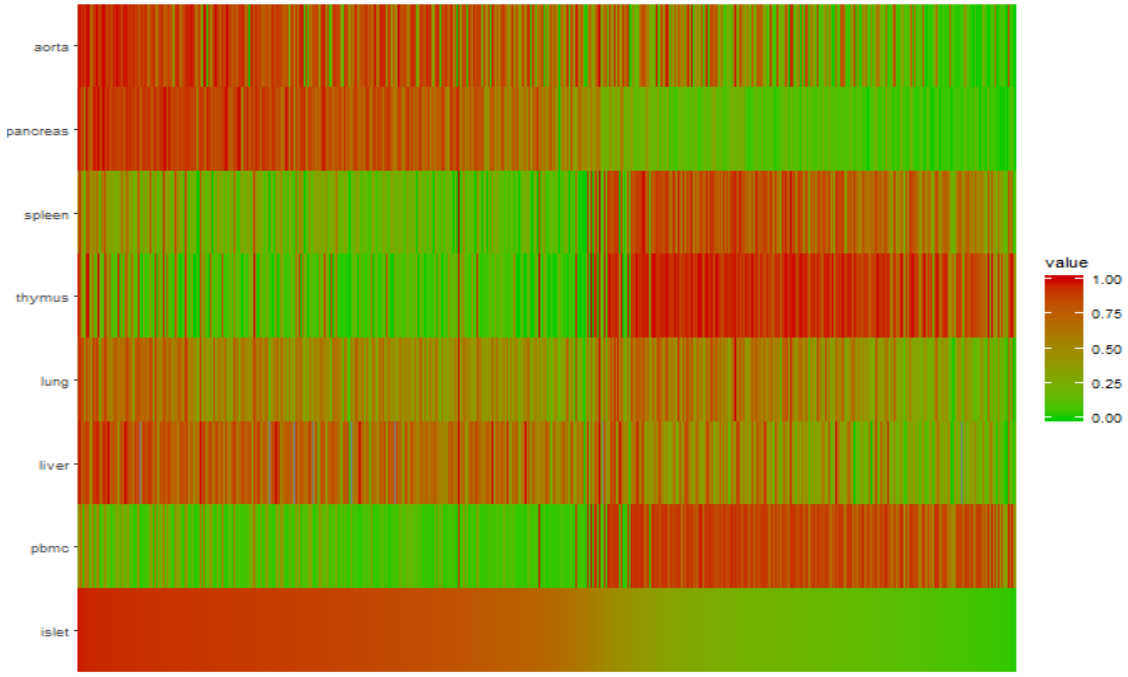


Figure 4.8: Heatmap of the methylation status of islet and PBMC DNA related to methylomes of various tissues. Data on the methylation status of DNA from liver, lung, thymus, spleen, pancreas and aorta was provided by the roadmap study. Each cell represents the CpG level of methylation for one site in one sample. The methylation scale indicates the level of methylation: green=sample predominantly hypomethylated (0-49% methylated); orange=50% of the sample is hypermethylated; red=sample predominantly hypermethylated (51-100%). As expected, the pancreas methylome was similar to the islets, whereas the methylome of PBMC, spleen and thymus was mostly the opposite of the islet's methylome. Those data further illustrate that there is considerable methylome homology between the islet, aorta, liver and lung.

***The Heatmap was constructed by Dr Jody Ye*

4.2.2. Developing DMR based Beta Cell-Specific Biomarkers

As mentioned previously, the analysis of the EPIC data resulted in identifying 3866 CpG sites that were differentially hypomethylated in islets when compared to the PBMC methylation signature. Among those, the CpG +49678, relative to TSS, in the gene *nuclear receptor corepressor 2* (*NCOR2*), and the CpG +20459 in the gene *solute carrier family 45 member 4* (*SLC45A4*) were in the top ten differentially hypomethylated sites. Therefore, these two sites were considered to candidate targets for designing DMR assays that could be specific for islet genome.

Initially, we tried to compare the methylation level of these specific sites (CpG +49678 of *NCOR2* and CpG +20459 of *SLC45A4*) to their methylation level in other tissues from published databases downloaded from Human Epigenome Atlas library (http://www.genboree.org/epigenomeatlas/multiGridViewerPublic.rhtml) and the genome browser of University of California Santa Cruz library (UCSC; http://smithlab.usc.edu/trackdata/methylation/hub.txt). Unfortunately, we found some conflicting data across different studies, most probably due to the use of different sequencing platforms and the adoption of different normalisation and analytic methods. Consequently, we were unable to compare the islet methylation level at these two sites, to other tissues. Instead, differentially methylated dual-labelled probes were designed to confirm the methylation specificity on a locally available tissue panel that includes a beta cell line (EndoC-βH1) sample.

4.2.2.1. Nuclear Receptor Corepressor 2 *(NCOR2)*

NCOR2 (Chr12: 124,808,957-125,052,079) is a pancreatic beta cell-specific gene (Bramswig et al. 2013). According to EPIC sequencing data, *NCOR2* was shown to be specifically differentially methylated between PBMC and islets at several CpG sites. Among the significantly hypomethylated sites, CpG +49678 in exon 2 (Chr12: 125,002,332), relative to TSS, was shown to be predominantly hypomethylated in islet DNA (89% hypomethylated in islet and only 11% in PBMC genome). Bisulfite specific primers and a pair of methylation differentiating probes were designed. The assay was used to investigate the methylation status of CpG +49678 and CpG +49671 in various locally available tissue samples (primer and probe design is described General Methods, Chapter 2, section 2.5). The assay amplifies a 270 bp DNA fragment that flanks a hypermethylated (VIC) and a hypomethylated (FAM) CpG +49678 and CpG +49671 probes. The optimum annealing temperature was identified by incorporating a temperature gradient into thermal cycling conditions for the annealing-extension step of amplification. The PCR was performed on eight samples containing the same final concentrations (400 pg/µL) of a template of 50% methylated bisulfite converted human DNA standard, following the standard ddPCR assay conditions (Figure 4.9). The sensitivity of the assay was confirmed when a linear correlation between the DNA methylation concentration and the expected positive droplets was observed (Figure 4.10 E & F).

Subsequently, the assay was run on a panel of 12 locally available tissue samples containing the same concentration of bisulfite converted DNA (5 ng/µL, Figure 4.11 and Figure 4.12). In parallel, the beta-actin assay was analysed as a normalising internal control. The results conducted from this tissue panel confirms both 1) the methylation differences between the islet and PBMC genome suggested by the EPIC array and 2) the non-specificity of the hypomethylated CpG + 49678 towards the islet genome as it was hypomethylated elsewhere.

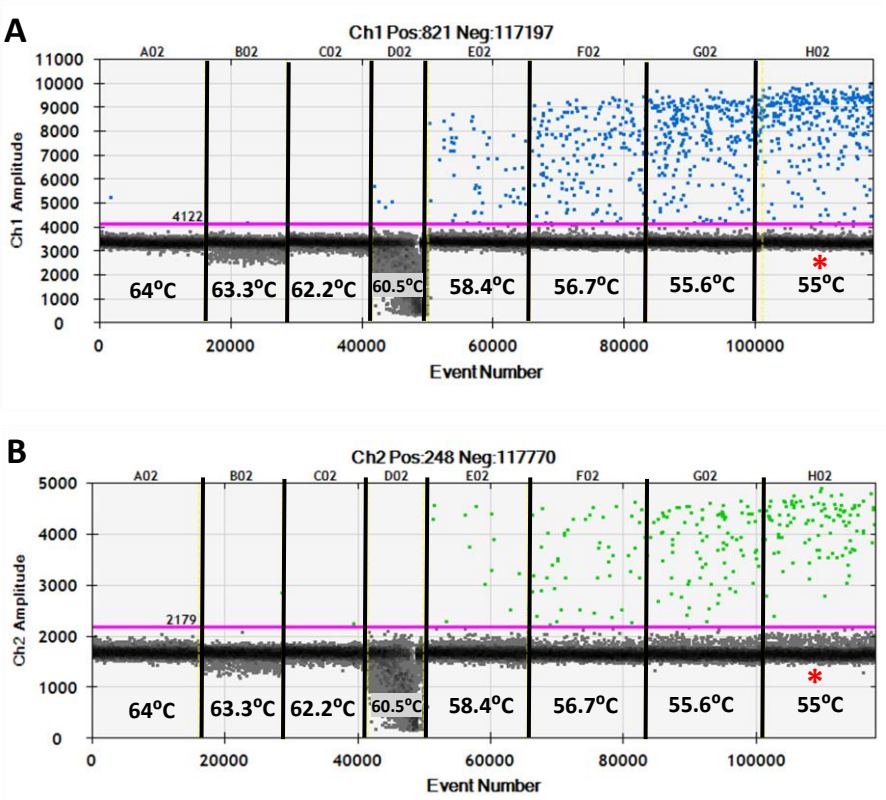
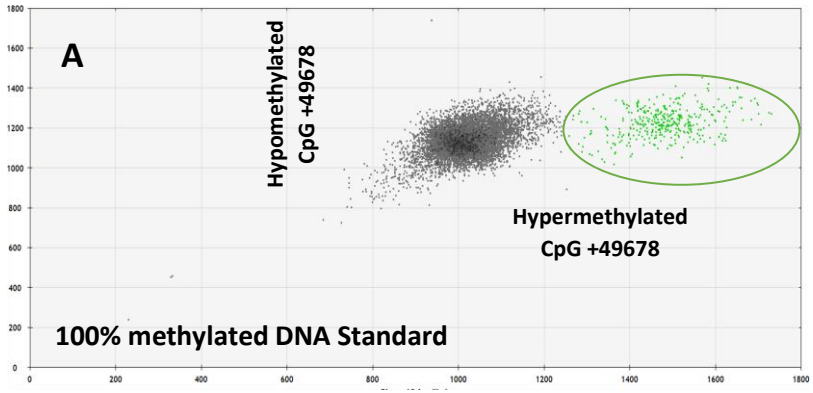
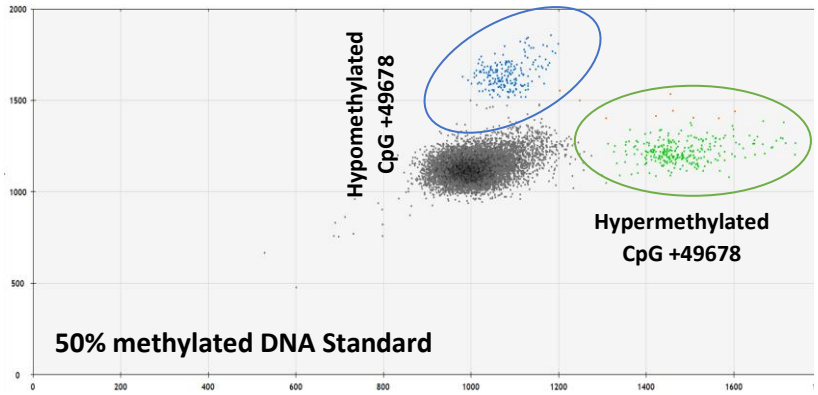
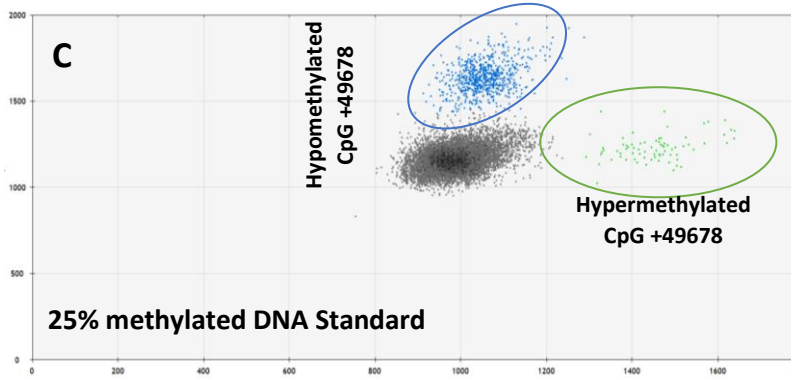


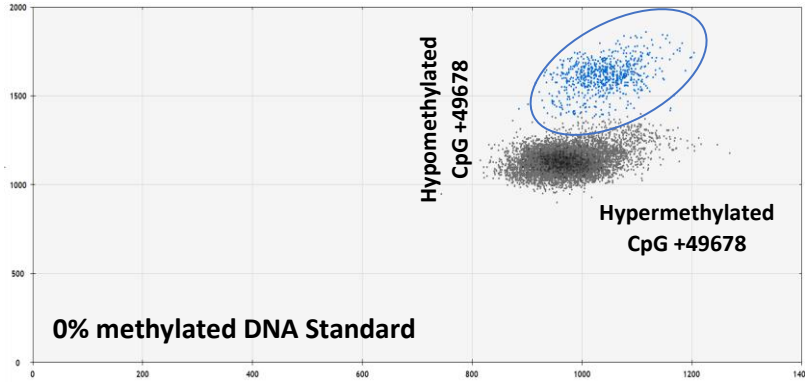
Figure 4.9: Optimal annealing temperatures for differentially methylated multiplexed NCOR2 assay.

A temperature gradient PCR was used to determine the optimal annealing temperature for the multiplexed assay. The tested annealing temperature per sample ranged from 55-64 degrees and was indicated in each ddPCR well in degrees Celsius (°C). An asterisk indicates the optimal ddPCR temperature. A) 1-D plot of 8 ddPCR amplification results of optimal annealing temperature for hypomethylated NCOR2 CpG +49678 assay. B) 1-D plot of 8 ddPCR amplification results of optimal annealing temperature for hypermethylated NCOR2 CpG +49678 assay.


A
Hypomethylated
CpG +49678
Hypermethylated
CpG +49678
100% methylated DNA Standard


Hypomethylated
CpG +49678
Hypermethylated
CpG +49678
50% methylated DNA Standard


C
Hypomethylated
CpG +49678
Hypermethylated
CpG +49678
25% methylated DNA Standard


Hypomethylated
CpG +49678
Hypermethylated
CpG +49678
0% methylated DNA Standard

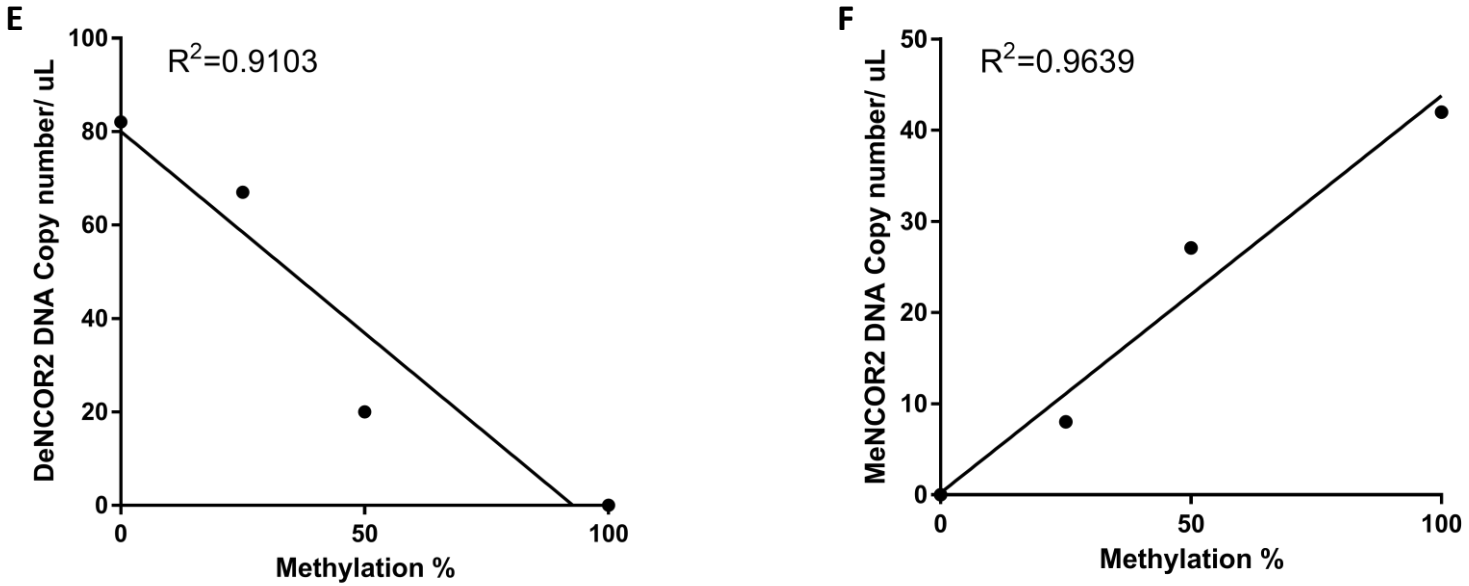


Figure 4.10: The methylation sensitivity of the NCOR2 assay.

*Synthesised hyper- and hypo-methylated human genomic DNA control was used as DNA template in this experiment. Both DNA standards were subjected to bisulfite conversion. The bisulfite converted DNA controls were mixed according to the percentage indicated in B) and C) and then amplified and analysed using ddPCR. A-D) Amplitude 2-D analysis plot of the results. The FAM channel represents the hypomethylated DNA copies of CpG +49678 (blue cluster) while the VIC channel (green cluster) represents the hypermethylated DNA copies of CpG +49678. E) &F) Linear regression analysis - graph plotted as DNA copies values in 1 µL ddPCR reaction (x-axis) and input methylated DNA (y-axis), as a way of evaluating the assay sensitivity. The hypomethylated assay line has the equation Y = -0.8629*X + 80.00 with the R2 value of 0.9103 (p = 0.0459). The hypermethylated assay line has the equation Y = 0.4355*X + 0.2200 with the R^2 value of 0.9639 (p = 0.0182). The plots show that there is a good relationship between the expected and the measured percentage of hypo- and hyper-methylated DNA.*

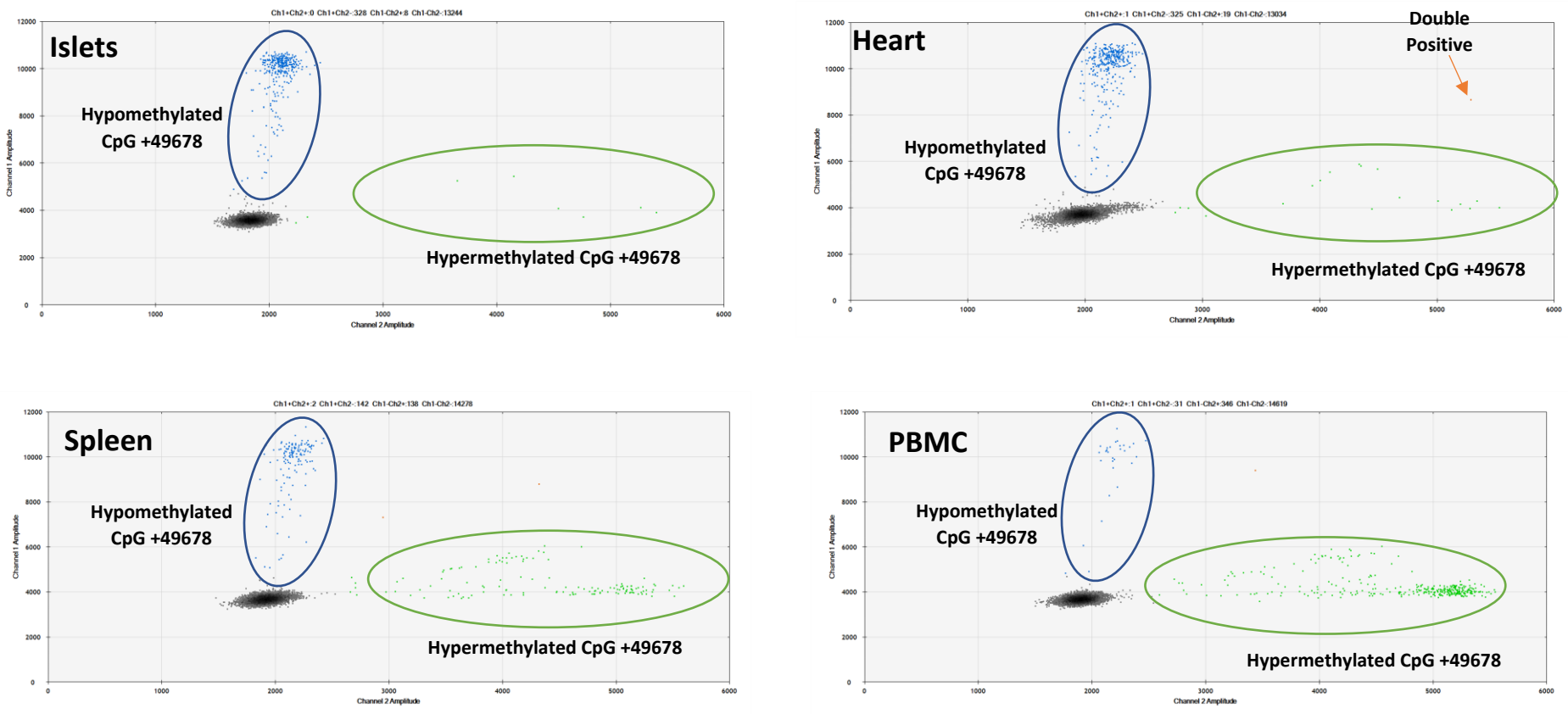


Figure 4.11: Representative amplification of 2-D plots from ddPCR data generated for tissue analysis with differentially methylated NCOR2 assay.

The FAM channel represents the hypomethylated DNA copies of CpG +49678 (blue cluster) while the VIC channel (green cluster) represents the hypermethylated DNA copies of CpG +49678. The data here reflect graph 4.12 showing that CpG +49678 is predominately hypomethylated in many tissues, which confirms the difference in the islet genome.

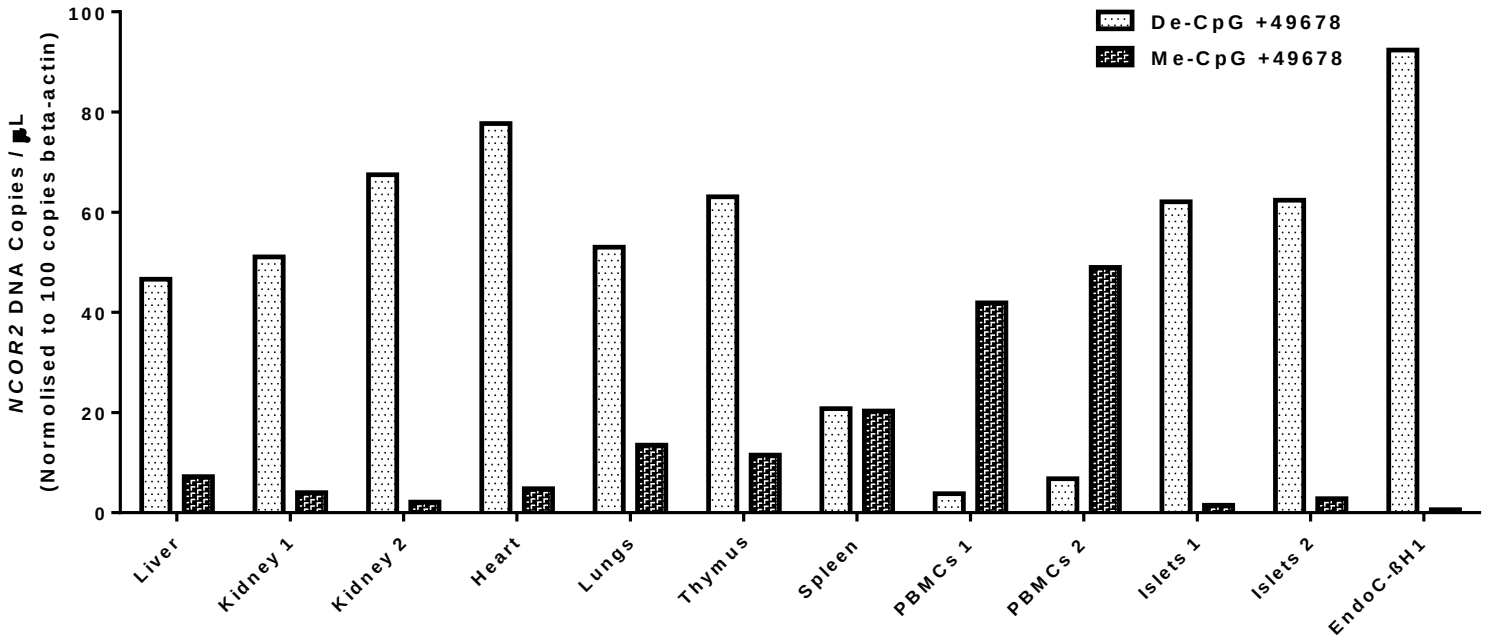


Figure 4.12: The methylation level of CpG +49678 of NCOR2 exon 2.

This CpG site was chosen based on EPIC array data that showed that CpG +49678 is 89% hypomethylated in islets. The graph shows the number of hypomethylated and hypermethylated DNA copies of CpG +49678 of NCOR2 exon 2 in 5ng of DNA samples extracted from different locally available tissue. The total copy number of NCOR2 was normalized to 100 copies of beta-actin. The data confirm the conclusion of EPIC array data which suggested that there is a differential methylation between the PBMC and islet methylomes. The data of both graphs illustrate that hypomethylated CpG +49678 of NCOR2 exon 2 is not specific for the pancreatic islets. The data were aligned using Integrated Genome Browser (IGB).

4.2.2.2. *Solute Carrier Family 45 Member 4* (*SLC45A4*)

SLC45A4 (Chr8: 142,217,273-142,310,241) was another discriminating biomarker candidate. This gene is beta cell-specific based on published data (Bramswig et al. 2013). The CpG +20459 (Chr8: 142,289,782) was shown to be predominantly hypomethylated in the islet genome (83% hypomethylated in islet and only 17% hypomethylated in the PBMC genome). Following the same steps to investigate specificity as for *NCOR2*, the methylation rate of this CpG site was studied in nine locally available tissue samples containing the same concentration of bisulfite converted DNA (5 ng/µL, Figure 4.13). In parallel, the beta-actin assay was run as a normalising internal control. The results from this tissue panel confirm the methylation differences between the islet and PBMC genome suggested by the EPIC array, and it further indicates the non-specificity of the hypomethylated CpG + 49678 toward the islet genome as it was hypomethylated in other tissues including lung.

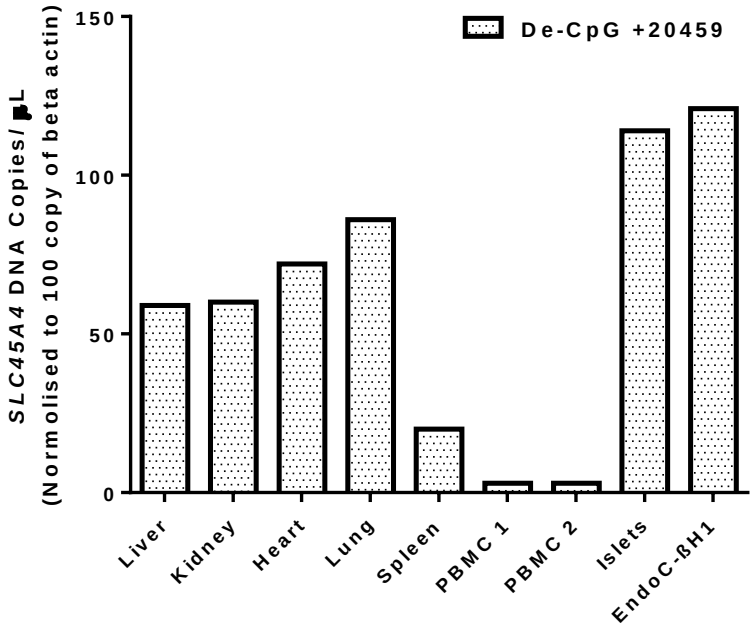


Figure 4.13: The methylation status of CpG +20459 from TSS of SLC45A4 gene in various tissues.

According to the EPIC array data, CpG +20459 is predominately hypomethylated in the islet genome. The columns in this graph represent the number of hypomethylated DNA copies of CpG +20459 of SLC45A4 in 5 ng of DNA samples in different locally available tissue DNA samples. The total copy number of SLC45A4 was normalized to 100 copies of beta-actin. The data confirm the conclusion of EPIC array data, which suggested that this CpG site is differentially hypomethylated in islets. Yet, the degree of methylation difference of this CpG in the other tested tissues was not high enough to be considered as specific for islet and EndoC-ßH1 cell line methylomes.

4.2.2.3. *microRNA-141* (*MIR-141*)

The series of experiments validating the methylation specificity of *NCOR2* (CpG +49678) and *SLC45A4* (CpG +20459) had led us to conclude that the data obtained from both the RoadMap study (Kundaje et al. 2015) and ENDCODE methylation library (Consortium 2012) were in agreement with our methylation data results. Therefore, we downloaded the liver, thymus, spleen and aorta methylation sequencing data from the Roadmap study database and the pancreas methylome data from the ENCODE methylation database. We then compared the methylome of these tissues to the islets in order to further exclude CpG sites that are not uniquely hypomethylated in the islet. The comparison resulted in a reduction in the number of islet-specific hypomethylated CpG sites from 3866 to only 425 (Appendix A, Table A.3). Among these sites, the CpG -20 of the gene *MIR-141* (relative to the TSS) had been shown to be predominantly hypomethylated in the islet genome and therefore was considered to be a candidate site for developing a DMR islet-specific assay.

MIR-141 (Chr12: 7,073,260-7,073,354) is a member of a gene family that was shown to induce apoptosis in a beta cell (Belgardt et al. 2015, Filios et al. 2014). The EPIC data suggested that CpG -20 (Chr12: 7,073,239) of the *MIR-141* gene promoter is predominantly hypomethylated in islets (72% hypomethylated in islet and 23% hypomethylated in the PBMC genome). Therefore, a methylation-sensitive probe-based assay was designed to detect hypermethylated and hypomethylated *MIR-141* at CpG -20 and CpG -18, relative to TSS. The annealing temperature for the assay was optimised, as for the previous two biomarkers, by incorporating a temperature gradient into thermal cycling conditions for the annealing-extension step of amplification. The PCR reaction was performed for eight samples containing the same concentrations (400 pg) of the template of bisulfite converted 50% methylated human DNA standard following the ddPCR conditions (Figure 4.14). Next, the assay specificity was tested on ten locally available tissue samples containing the same concentration of bisulfite converted DNA (5 ng/µL, Figure 4.15 & 4.16). In parallel, the beta-actin assay was run as a normalising internal control. The data shows that CpG -20 is specifically hypomethylated in the islet and beta cell line genomes. The assay specificity was also confirmed by other tissue methylome data

that were taken from the public Roadmap and ENCODE methylation databases (as described in General Methods, Chapter 2, section 2.4 of). The data were in agreement with the conclusion from the tissue panel experiment, confirming the specificity of CpG-20 hypomethylation to the pancreatic islets and suggesting that CpG -20 of *MIR-141* is a potential valid biomarker candidate (Figure 4.17).

During the experiments optimising the hypermethylated CpG -20 *MIR-141* probe (FAM), we noticed that the amplification of the FAM probe resulted in the production of two distinct clusters in the ddPCR 2-D results. The reason why two droplet clusters were generated was further investigated, and consequently, a new set of probes that amplify either the hypermethylated CpG -18 or the hypermethylated CpG -20 were designed. The PCR reaction was performed for both probes containing the same concentrations of DNA template (50% hypermethylated bisulfite converted human DNA standard, 400 pg). The data showed that the predominant amplification was to CpG -20 (Figure 4.18). Nevertheless, due to technical issues with the amplification stability of the new probes, the first designed FAM probe (Hypermethylated at CpG -18 and CpG -20) was used in the subsequent experiments during this project.

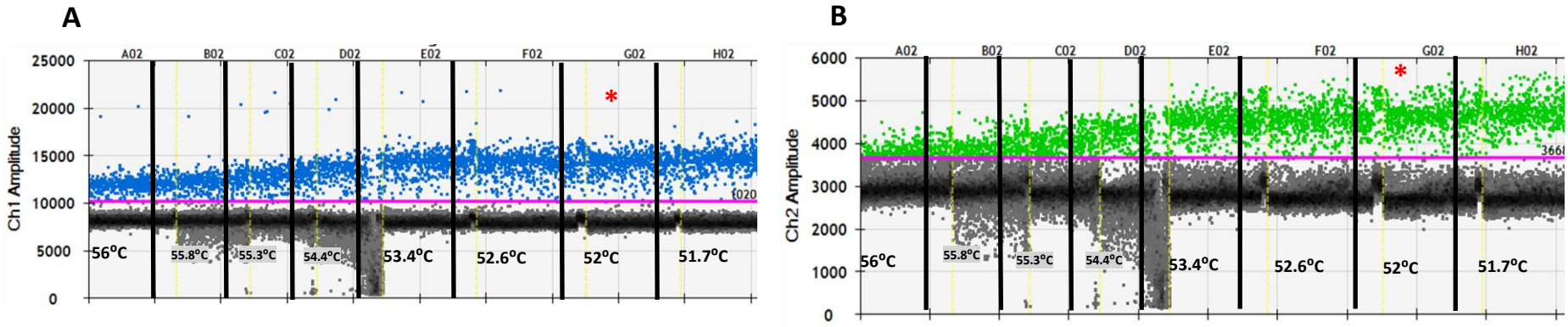


Figure 4.14: 1-D fluorescence amplification plotted against the annealing temperature gradient for the differentially methylated multiplexed MIR-141 assay (CpG -18/CpG-20).

The figure shows the droplet populations for duplex of fluorescence amplitude for each droplet. Blue (FAM) and green (VIC) dots represent the positive droplets, generated from amplifying 400 pg of 50% methylated human DNA standard, for hypermethylated and hypomethylated CpG-18/-20 probes, respectively. The grey dots below the pink horizontal threshold represent the negative droplets. A temperature gradient PCR was used to determine the optimal annealing temperature for the multiplexed assay. The annealing temperature tested per sample ranged from 51.7-56 degrees that was indicated in each ddPCR well in degrees Celsius (°C). An asterisk indicates optimal ddPCR temperature. A) 1-D plot of 8 ddPCR amplification results of optimal annealing temperature for the hypermethylated MIR-141 CpG -18 and CpG-20 assay. B) 1-D plot of 8 ddPCR amplification results of optimal annealing temperature for hypomethylated MIR-141 CpG -18 and CpG-20 assay. The data indicate that the temperature of 52°C is the optimum temperature that sufficiently separates the positive signal from the negative in both assays.

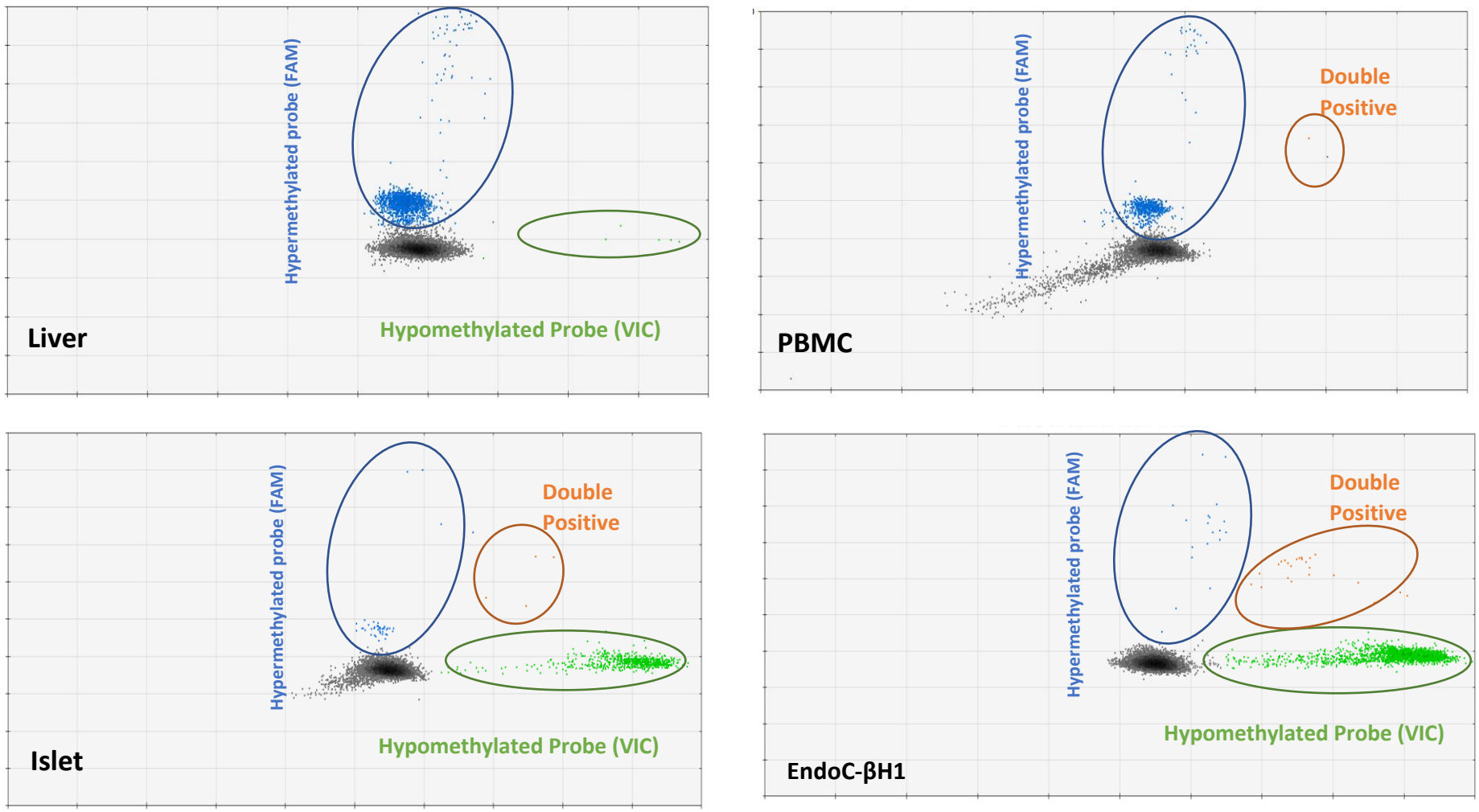


Figure 4.15: 2-D amplification plots from droplet digital PCR data generated for tissue analysis with the differentially methylated MIR-141 assay (CpG -18/CpG-20)..

The FAM channel represents the hypermethylated DNA copies of CpG -18/-20 (blue cluster) while the VIC channel (green cluster) represents the hypomethylated DNA copies of CpG -18/-20. The data here reflect the graph 4.16 showing that CpG -20 is specifically and predominately hypomethylated in islets and beta cell line model.

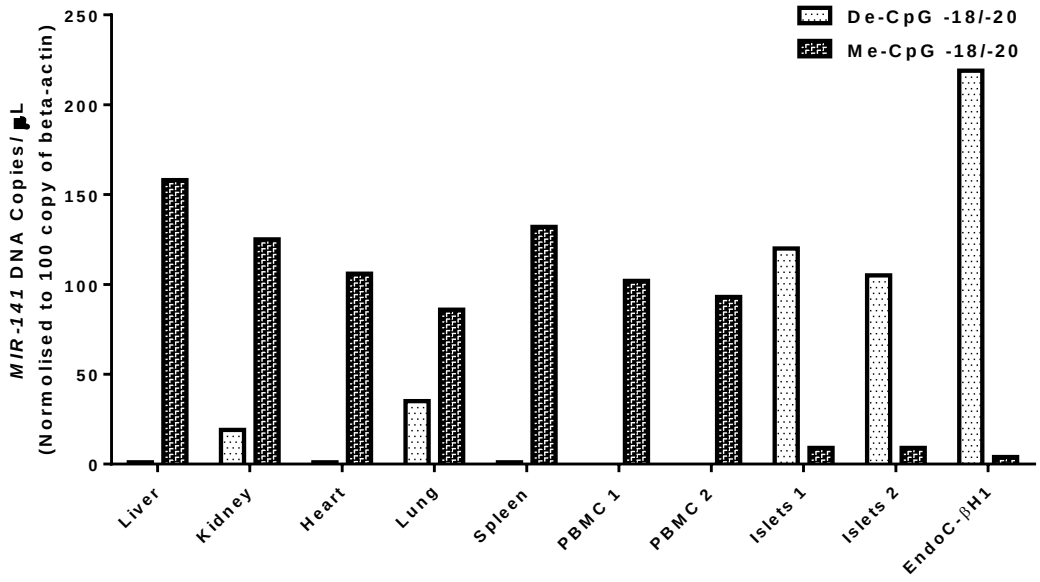


Figure 4.16: The methylation status of CpG -20 (Chr12:7,073,239) and CpG -18 (Chr12:7,073,241) of MIR-141 gene.

These CpG sites were chosen based on EPIC array data, which showed that CpG -20 is 72% hypomethylated in islets. The level of methylation of the MIR-141 assay in a panel of different locally available tissues. The data presented here are the copy number of MIR-141 in 200 pg of each sample that was normalised to 100 copies of beta-actin. The data supports the conclusion from the EPIC array data, which suggested that this CpG site is differentially hypomethylated in islets and suggests that hypomethylated CpG -18 and CpG -20 of MIR-141 promoter is specific for islets as it is predominately hypermethylated elsewhere.

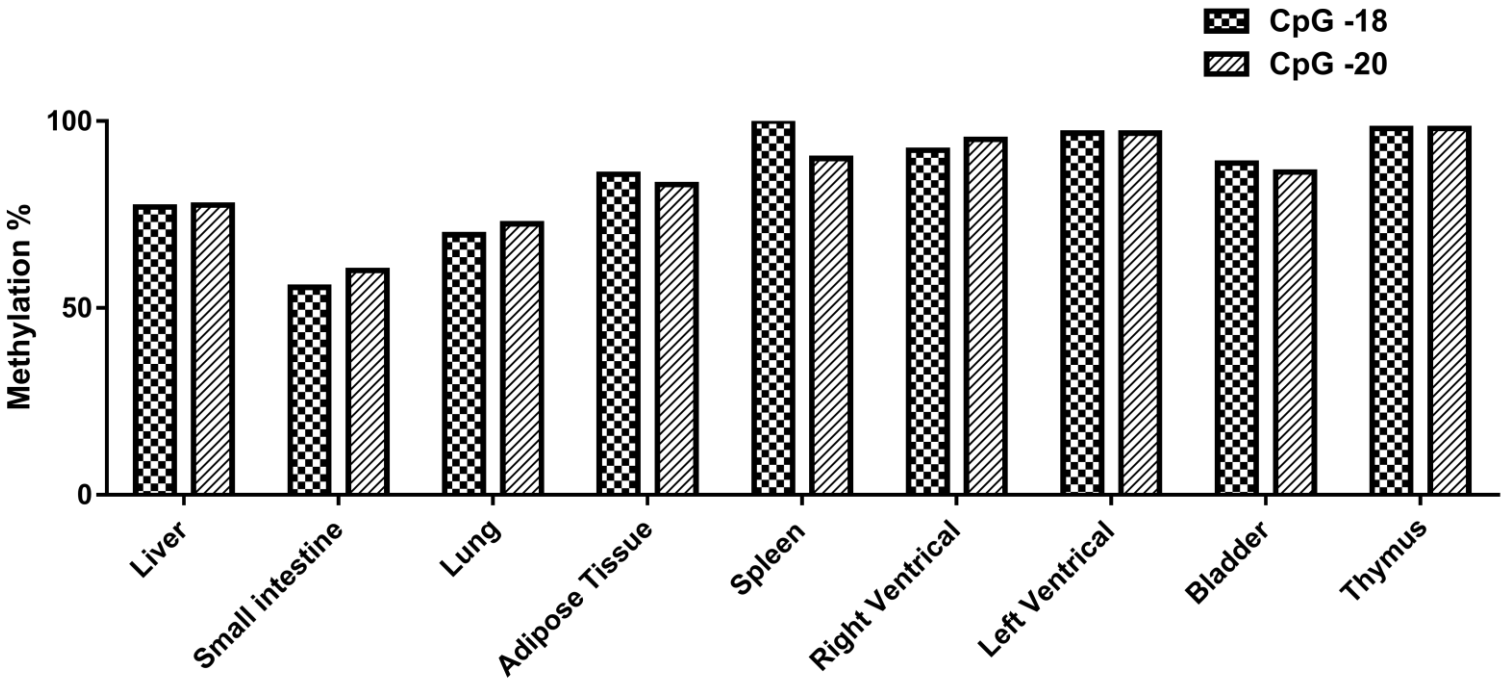


Figure 4.17: The methylation status of CpG -20 (Chr12:7,073,239) and CpG -18 (Chr12:7,073,241) of MIR-141 gene.

These CpG sites were chosen based on EPIC array data, which showed that CpG -20 is 72% hypomethylated in islets. The data were aligned using the Integrated Genome Browser (IGB). The methylation rate of CpG -18 and CpG -20 of MIR-141 in different tissue methylomes obtained from the Roadmap methylation signature study. The data suggest that both CpG sites are highly hypermethylated in the tested tissues.

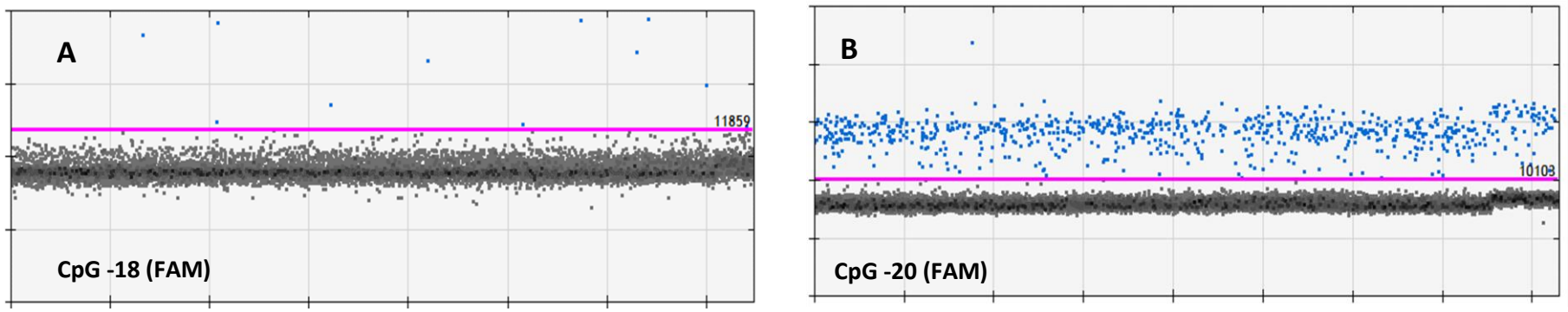


Figure 4.18: 1-D fluorescence amplification plots of two hypermethylated MIR-141 assays.

The Blue (FAM) dots represent the positive droplets, generated by amplifying 400 pg of 50% methylated human DNA standard, while the grey dots below the pink horizontal threshold represent the negative droplets. 400 pg of bisulfite converted 50% methylated human DNA standard was used as the DNA template A) The hypermethylated FAM probe amplifies MIR-141 at CpG -18 (from TSS). B) The hypermethylated FAM probe amplifies MIR-141 at CpG -20 (from TSS). The plots confirm that the dense cluster that was seen in the previous experiment represents hypermethylated CpG-20 amplification.

4.2.3. Developing Triplex Assay to Detect Beta Cell Death

In addition to the *MIR-141* assay, another two single plex assays had been developed locally. They were the CpG +367/+374 *INS* exon 2 and CpG -1052 of *MIR-200c* (from TSS) assays which were developed and validated by Dr Jody Ye. The gene for *MIR-200c* is adjacent to *MIR-141,* and both share the same promoter (Neves et al. 2010). The next experiments aimed to multiplex these assays with the hypomethylated *MIR-141* probe to increase the sensitivity of the beta cell death assay.

4.2.3.1. Optimising the Triplex Assay

A touchdown PCR reaction (as described in section 4.1.4) was used in the amplification of multiplex assay in order to increase the specificity of the PCR. The temperature gradient ddPCR experiment had shown that 52.5^{o}C was the optimum annealing temperature for the triplex assay. Next, the sensitivity of the multiplexed assay toward the hypomethylated DNA fragments was evaluated. A serial concentration of hypomethylated synthesised DNA (0, 12.5, 25, 50, 75, 100 pg) was added to the reaction in the presence of a fixed concentration of hypermethylated synthesised DNA standard (100 pg). The ddPCR analysis showed that the positive droplets number declined in parallel with the hypomethylated DNA concentration decreases and was not affected by the presence of the hypermethylated DNA (Figure 4.19). Additionally, the R^2 value (R^2 = 0.9972) shows a linear correlation between expected and measured percentage of the hypomethylation (Figure 4.19B). Subsequently, the specificity of the assay was tested by running the assay on the locally available tissue samples (*n*=11) containing the same concentration of bisulfite converted DNA (5 ng/µL, Figure 4.20 and Figure 4.21). In Parallel, the beta-actin assay was run as a normalising internal control, and each sample was normalised against 100 DNA copies of beta-actin. The results conducted from the tissue panel experiment confirms the observation of previous experiments, which indicates that hypomethylated CpG -1052 of the *MIR-200c* promoter, CpG -20 of the *MIR-141* promoter and CpG +367+374 of *INS* exon 2 are beta cell and islet-specific.

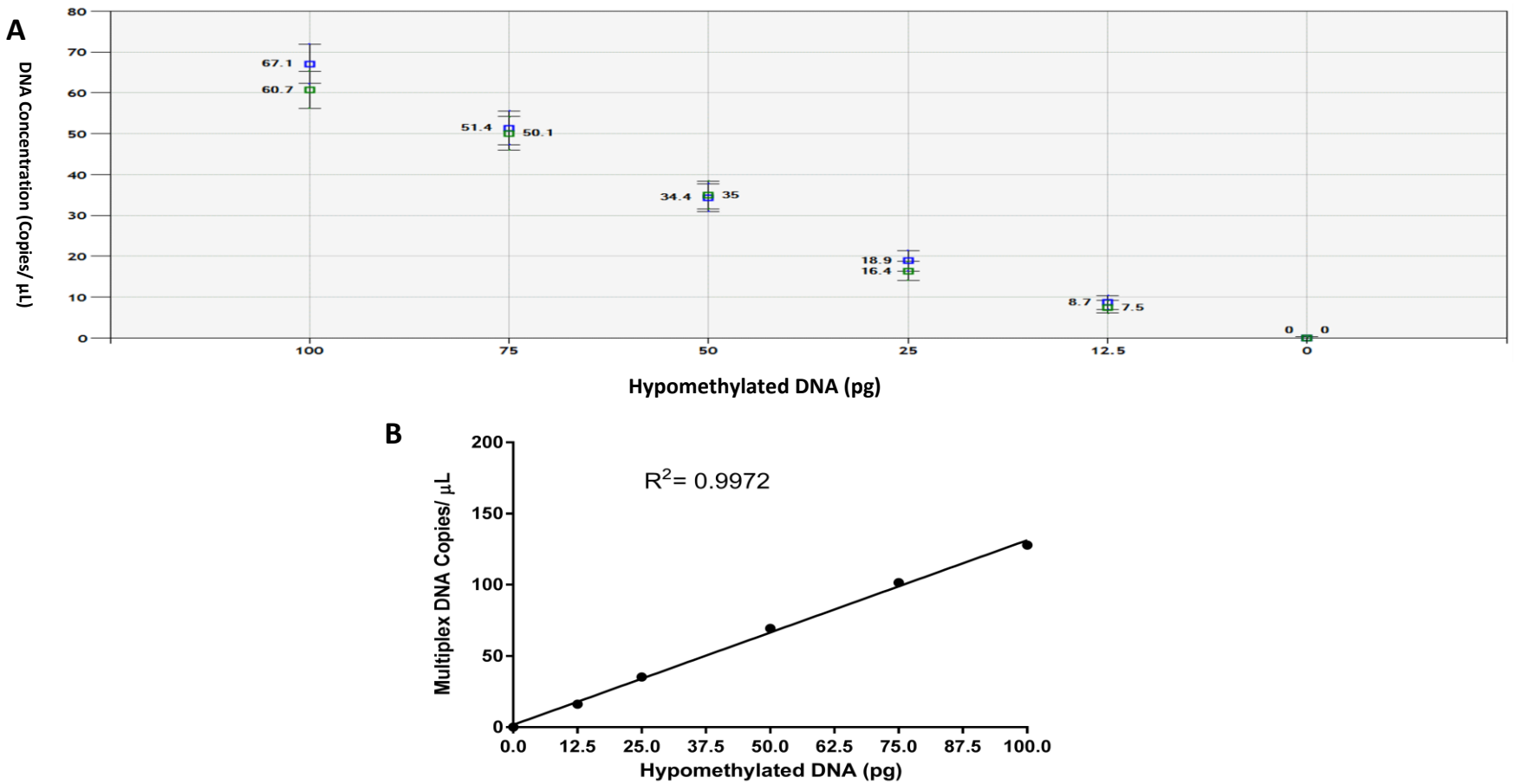


Figure 4.19: The sensitivity of the triplex assay for methylation detection by ddPCR

*A gradient of hypomethylated bisulfite converted DNA controls were mixed with a fixed concentration of hypermethylated bisulfite converted DNA control (100 pg) as indicated in figure A. A) The raw plot of gradient hypomethylated DNA yield (0, 12.5, 25, 50, 75, 100 pg) in the presence of fixed hypermethylated DNA yield (100 pg). The data are presented as the DNA concentration per 1 µL reaction. B) Linear regression analysis – the graph plotted as DNA copies values per 1 uL ddPCR reaction (x-axis) and the percentage of input methylated DNA (y-axis), as a way of evaluating the assay sensitivity. The triplex hypomethylated assay line has the equation Y = 1.297*X + 1.622with the R2 value of 0.9972 (p <0.0001). The plots show that there is a linear relationship between the expected and the measured percentage of hypomethylated DNA.*

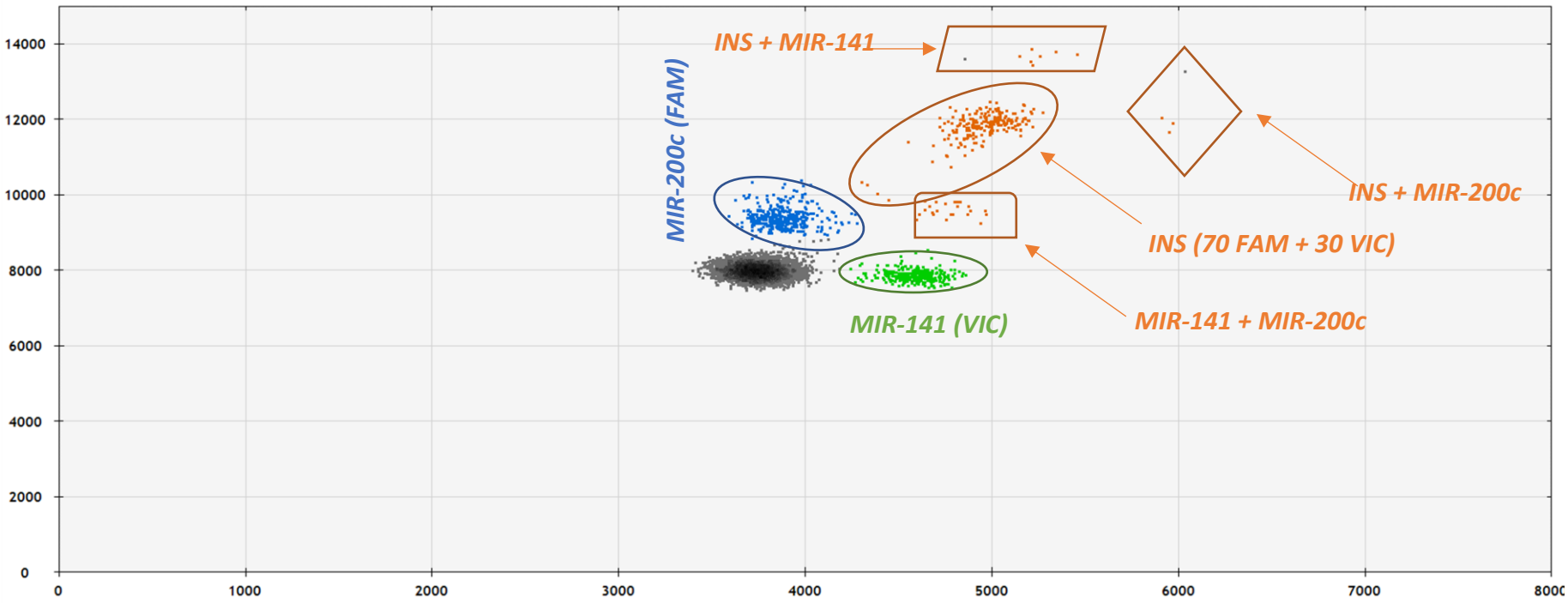


Figure 4.20: Representative 2-D amplification plots from droplet digital PCR data generated by the tissue analysis with the triplex assay.

Amplification of 5 ng bisulfite converted islet's DNA. The FAM channel represents the hypomethylated DNA copies of CpG -1052 (blue cluster) of MIR-200c, the VIC channel (green cluster) represents the hypomethylated DNA copies of CpG -20 of MIR-141 while the central orange cluster represents a mixed probe (70% FAM + 30% VIC) of the hypomethylated CpG +367+374 of INS exon 2.

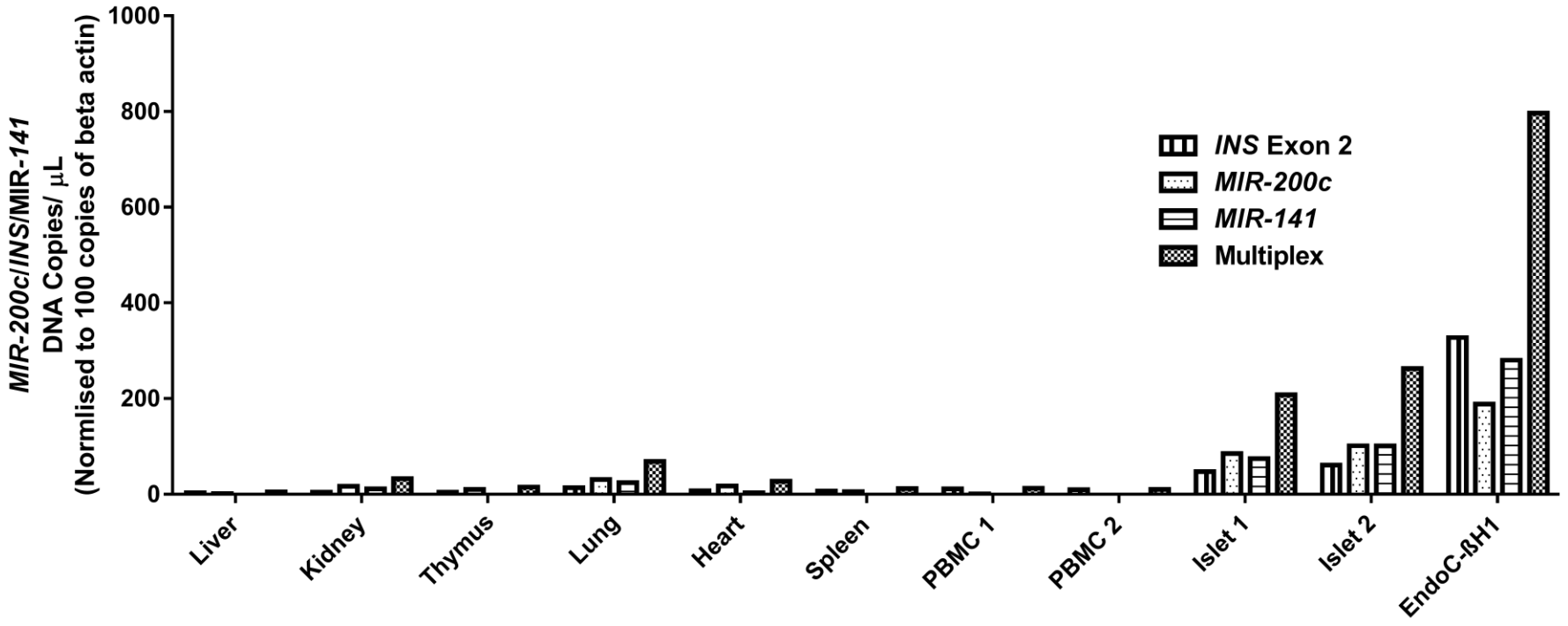


Figure 4.21: The level of hypomethylated DNA copies of the triplex assay in different tissues available "in house".

The columns represent the normalised DNA copy number of hypomethylated CpG -20 MIR-141, CpG -1052 of MIR-200c and CpG +367+374 *of INS gene in 1 µL of DNA sample. The data presented here are the copy number of MIR-141 in 200 pg of each sample normalised to 100 copies of beta-actin. The data support previous conclusions confirming the specificity of the tested CpG of MIR-141 and MIR-200c assays toward human islet and beta cell genomes.*

4.2.4. The Effect of Adding EcoRI to Minimise Droplet Rain

Although positive and negative droplets are clustered separately from the negative (empty) droplets, some droplets fall between the negative and positive clusters and are usually referred to as droplet rain. Droplet rain is problematic sometimes because it makes drawing the threshold line challenging. The possible sources of such intermediate droplets vary. It could be due to damaged and irregular droplet size, coagulation of multiple droplets, partial PCR inhibition, non-specific or suboptimal amplification or suboptimal PCR amplification due to sequence variances (Witte et al. 2016). Adding restriction enzymes was recommended by Bio-Rad to minimise intermediate droplet rain as, according to their ddPCR guidance book: "*restriction digestion separates tandem gene copies, ensuring proper random partitioning into droplets, restriction digestion also reduces sample viscosity and improves assay performance by improving template accessibility*". Therefore, one unit of EcoRI (0.5 µL of 50X) was added to the multiplex ddPCR reaction (25 µL) and compared to a reaction without EcoRI. Adding a restriction enzyme was shown to have a "cluster condensing" effect but did not minimise the droplet rain (Figure 4.22). Therefore, the enzyme was not used in our clinical sample analysis.

Figure 4.22: The effect of adding EcoRI on a ddPCR reaction.

A fixed concentration of bisulfite converted 0% methylation DNA template was added to the reaction (350 ng). The reactions were amplified following a touchdown PCR protocol. A) 2-D plot of ddPCR reaction without EcoRI. B) 2-D plot of ddPCR reaction with 1x EcoRI. Adding the restriction enzyme resulted in narrowing the margin between MIR-200c and INS cluster, but no other effect e.g. minimizing the droplet rain was observed.

4.2.5. Targeted Next-Generation Sequencing

4.2.5.1. The level of Methylation Differences Between the Pancreatic Islet and EndoC-βH1 Cell Line

As our EPIC analysis was based on the methylation level of the entire islet genome, we were concerned that there might be important differences between the beta cell and islet methylomes (which is composed of 60-80% beta cells). Another reason for our query was raised in a the talk given by one of the leading researchers in the field, Dr Kevan Herold, Immunology of Diabetes Society Congress 2017 in San Francisco, indicating that based on their data, the methylome differences between islets and the beta cell is 5-10%. Therefore, we aimed to investigate the methylation differences in DNA between a pancreatic islet and a pure beta cell methylome at the mutual promoter of *MIR-200c/MIR-141*. To attain this goal, DNA was extracted from normal pancreatic islets as well as from the EndoC-βH1 cell line (described in sections in General Materials, Chapter 2, section 2.1 and 2.2.1). The DNA samples were then sent to EpigenDX for further analysis as described earlier in the Methods Section 4.1.2. The data showed that there is a 5.3 % methylation difference between the two samples (Table 4.2; Figure 4.23).

Table 4.2: Summary of methylation differences between EndoC-βH1 cell line and a pancreatic islet sample. The average difference was calculated as (average Islet methylation – average islet methylation).

	EndoC-βH1 cell line	**Pancreatic islet**
	Methylation Percentage	
Average	0.26 %	5.52 %
Maximum	2.89 %	9.94 %
Minimum	0 %	0 %
Median	0 %	5.55 %
AVG Difference	5.3 %	

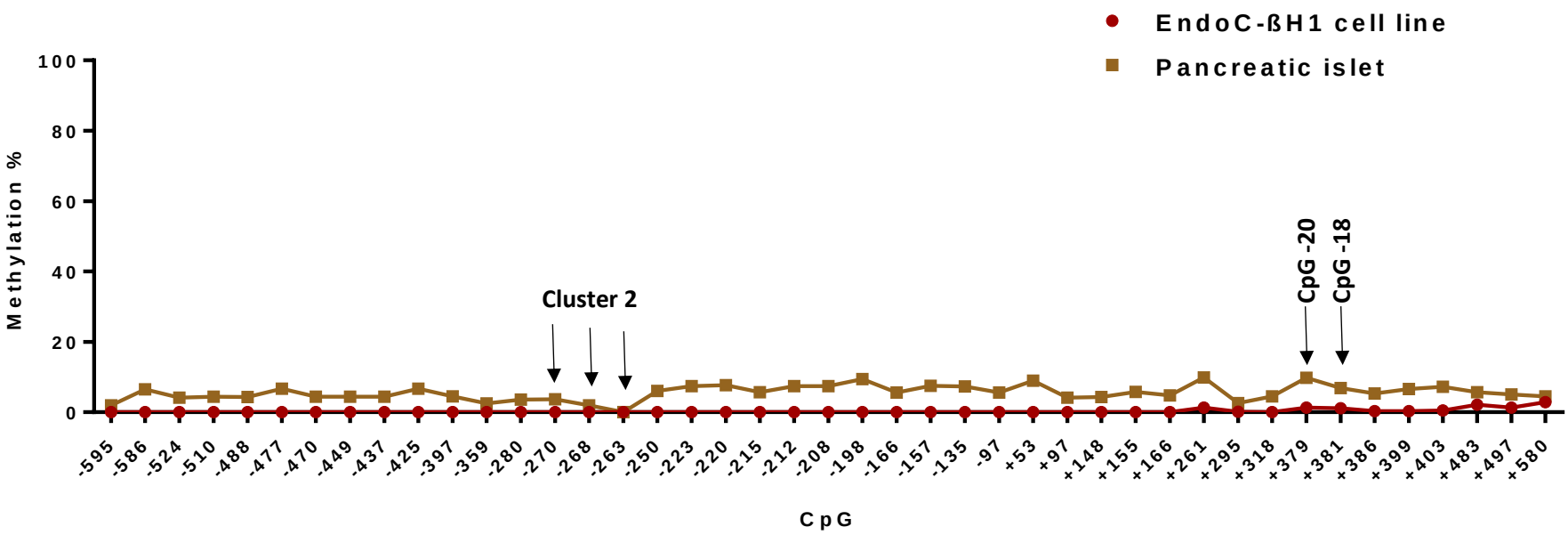


Figure 4.23: The methylation status of a pancreatic islet sample and a beta cell line model.

The line chart shows that there is a 5.3% margin in methylation difference between islet and EndoC-βH1 methylomes in the tested DNA region (Chr12: 6,963,000-6,964,001).

4.2.5.2. The Effects of High-Glucose Exposure on DNA Methylation in the *MIR-200c/MIR-141* Promoter Region

The influence of hyperglycaemia on the methylation profile of the beta cell line in the promoter region of *MIR-141* was investigated. To achieve this, a targeted next-generation bisulfite sequencing technology was used to assess methylome differences in the pancreatic beta cell line (EndoC-βH1) model, which was exposed to normal (5.5 mM) or high glucose (25 mM) concentration for 48 hours. The data showed that hyperglycaemia exposure has a minimum effect on the methylome profile of the targeted region (Figure 4.24). The average difference in methylation change of the treated cells was 0.5% (Appendix A, Table A.5).

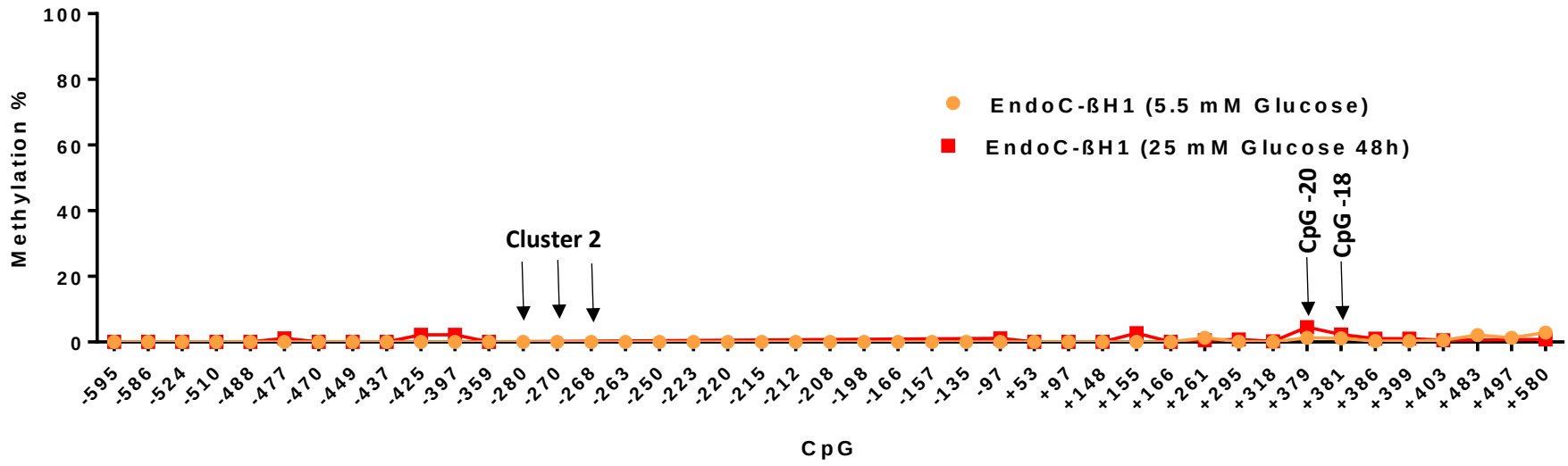


Figure 4.24: The effect of high glucose exposure on the methylation profile of the EndoC-ßH1 beta cell line model.

The figure illustrates that there is a minimum effect of hyperglycaemia on the methylation profile of CpG sites within the MIR-200c/MIR-141 promoter region (Chr12: 6,963,000-6,964,001).

4.2.6. The effect of Hyperglycemia and Proinflammatory Cytokines of the Methylation Stability in the Developed Multiplex Assay

To investigate whether high glucose and proinflammatory cytokines exposure could affect the methylation stability of the developed multiplex assay, we treated EndoC-βH1 cells with high glucose concentration up to 48 hours, as described in General Methods, Chapter 2, section 2.2.1. The methylation level of the treated cells was measured relative to the control EndoC-βH1 cells that were grown in the presence of physiological glucose concentration (5.5 mM) and the absence of test cytokines. As shown in Figure 4.25 and Figure 4.26, compared to the control cells, there were no significant changes ($p > 0.05$) in the relative hypomethylation levels for all three cfDNA targets when either treated with hyperglycemic nor proinflammatory conditions.

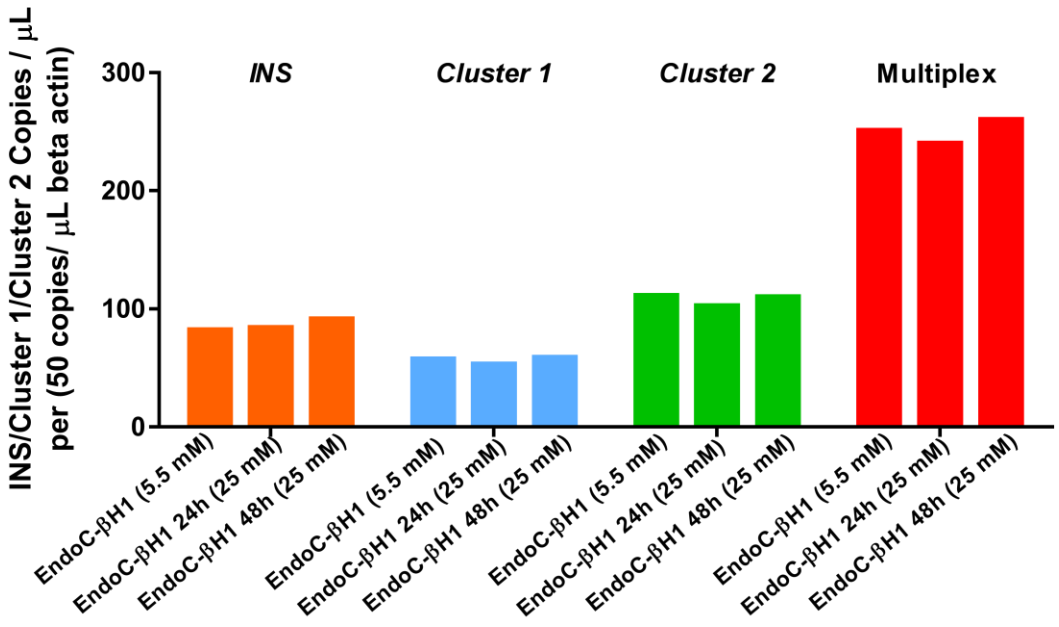


Figure 4.25: The effect of high glucose exposure on the methylation profile of the designed beta specific targets. The figure illustrates that there is no significant ($p > 0.05$) change in the relative methylation level of the INS, Cluster 1 or Cluster 2 single plex assays nor in the multiplex assay when exposed to hyperglycaemia condition in vitro.

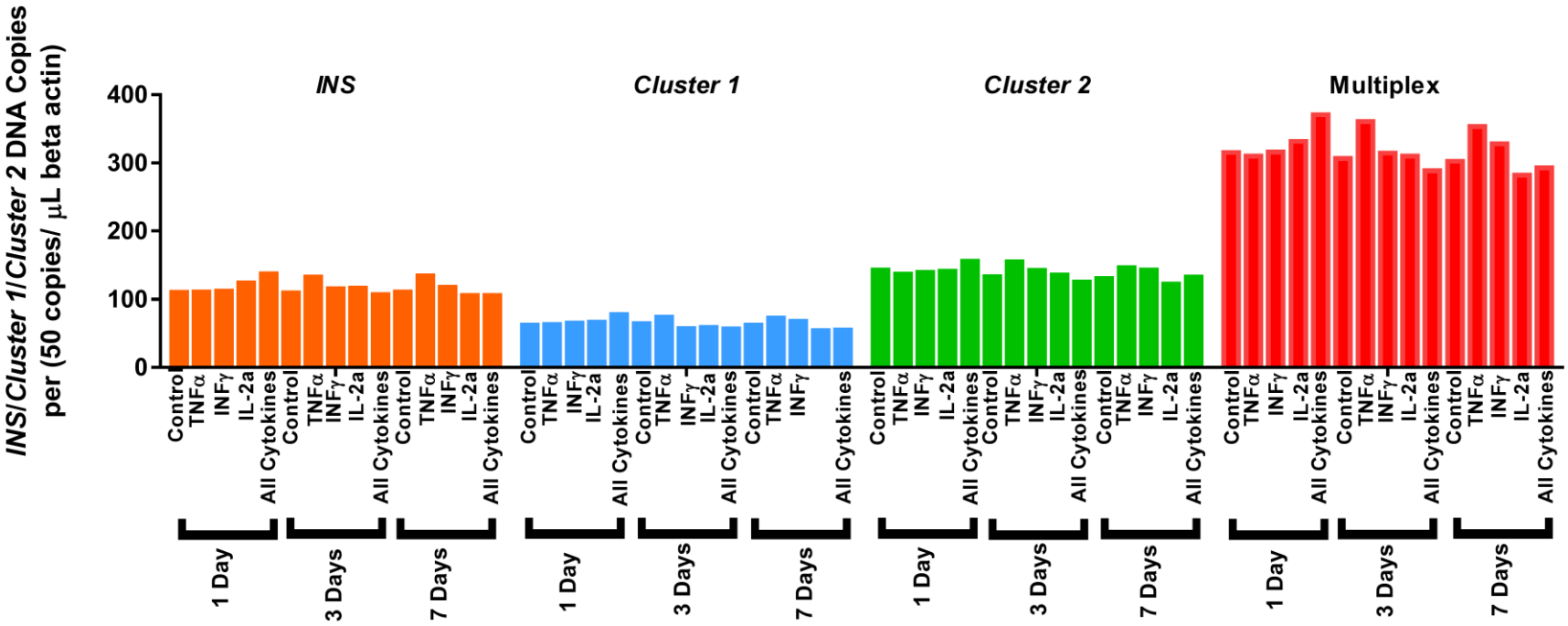


Figure 4.26: The effect of pro-inflammatory cytokines exposure on the methylation stability of the designed beta specific targets.

The figure illustrates that there is no significant change ($p > 0.05$) in the relative methylation level of the INS, Cluster 1 or Cluster 2 single plex assays nor in the multiplex assay when exposed to proinflammatory cytokines in vitro.

Discussion

The genome-wide methylation sequencing approach resulted in a library of differentially methylated CpG loci that are specific for the islet genome. Comparing the islet and PBMC methylation signature data to other tissue methylation signatures available from the Roadmap and ENCODE studies revealed that there are 425 hypomethylated and 228 hypermethylated CpG targets specific to the islet genome. It is important to highlight here that there were some difficulties in comparing tissue methylome data available from the UCSC Genome Database (https://genome.ucsc.edu/index.html) with the data generated by the EPIC array analysis. There was variation in the methylation levels of the tissue methylomes reported by different studies, for instance, the liver methylome. This variation most probably arose due to the use of different sequencing techniques as well as different analysis and annotation methods in representing the methylation of the tissue. Comparing our in-house sequencing of different CpG sites, e.g. CpG +49678 in exon 2 of the *NCOR2* gene, CpG +20459 of *SLC45A4*, led us to conclude that the Roadmap methylation signature data was the most reliable database that could be used as a reference database. It is also worth highlighting here that none of the DMRs identified as beta cell-specific were located within genes that are well-known to be crucial for beta cell function (*PDX-1, MafA, and NeuroD1*). This could suggest an issue with the original candidate approach used in this project. However just because they were not identified, it does not mean that there are no differentially methylated CpG sites within these genes. The EPIC BeadChip covers only 3% of the total number of human genome CpG dinucleotides (total number of human CpG dinucleotides is 27,999,538, and the EPIC BeadChip array sequenced only 850,000 CpG sites). There is, therefore, a significant chance that there are DMRs within these genes that were not sequenced via the EPIC array.

Among the uniquely hypomethylated 425 CpG sites, several CpG loci within the mutual promoter of *MIR-200c*/*MIR-141* shown to be possible candidates for designing beta cell-specific biomarkers (they are uniquely hypomethylated in islet and beta cell line genome). Further characterisation

suggested that the identified candidates may be able to detect and differentiate beta cell fragments in the circulation. The *MIR-200c/MIR-141* genes are reported to play an important role in beta cell survival and the pathophysiology of diabetes. Multiple studies have shown that the *MIR-200c*/*MIR-141* genes are actively involved in apoptosis of islet cells through pathways that induce endoplasmic reticulum stress (Belgardt et al. 2015, Filios and Shalev 2015, Filios et al. 2014, Klein et al. 2013). Two combinations of the human hypomethylated sensitive multiplexed assay were developed (*MIR-200c cluster 1* + *INS* exon 2 + *MIR-141*) and (*MIR-200c cluster 1*+ *INS exon 2* and *MIR-200c cluster 2;* developed by Dr Jody Ye). Both combinations showed a high degree of specificity towards the islet and human EndoC-βH1 genomes. The sensitivity of both assays was confirmed when the amplification preserved a linear pattern throughout a different range of artificially hypomethylated DNA concentrations in the presence of a fixed concentration of hypermethylated DNA control. Further, the specificity of the multiplex assays toward beta cells was not affected when the cells were exposed to hyperglycaemia or to a mixture of pro-inflammatory cytokines that mimic the microenvironment of T1D. As these biomarker assays will be used on biofluid samples collected from patients having a condition involving beta cell death, e.g. preclinical T1D, islet transplantation and at risk for developing T1D, it was essential to define the optimum the volume of biofluid sample required to perform downstream analysis, e.g. ddPCR. Optimising the ideal volume of starting material was important due to the limited number and volume of samples available. It was also important to optimising the volume of the starting material in-house as conflicting starting material volumes was reported by other groups (Akirav et al. 2011, Lehmann-Werman et al. 2016, Tersey et al. 2016).

Chapter 5: Optimising Cell Free DNA Extraction Protocol

5. Introduction

At the time of T1D onset, 70-80 % of beta cells will be destroyed by autoimmune mechanisms (Cnop et al. 2005; Notkins and Lernmark 2001). Consequently, the nucleic acid content of dying beta cells is released into the peripheral circulation. One of the approaches for discriminating beta cell-derived DNA fragments is by designing an assay that amplifies DNA fragments that have a distinct differentially methylated nucleotide. These types of assays are known as methylation sensitive assays. The employment of methylation-sensitive biomarkers has already been used to study cancer (Delpu et al. 2013; Esteller 2007; Fukushige and Horii 2013; Leygo et al. 2017; Szyf 2012; Verma et al. 2006; J. Wang et al. 2017b; Wang et al. 2017a; Yi et al. 2012) and recently in T1D (Akirav et al. 2011, Husseiny et al. 2014, Lehmann-Werman et al. 2016). Estimates of the concentration of cfDNA in the circulation of healthy individuals vary wildly as it is highly dependent on the used extraction method (Fleischhacker et al. 2011; Sharma et al. 2011; Xue et al. 2009). Exploring the published literature discussing cfDNA extraction revealed that the most commonly used extraction methods relied on commercial kits (e.g. QIAamp Circulating Nucleic Acid and QIAmp DNA blood kits from QIAGEN) that employed column-based extraction chemistry (Sharma et al. 2011). Additionally, we found that the volume of the starting extraction material (plasma/serum) varied in different studies that investigated similar hypotheses, from 50 µL of serum (Tersey et al. 2016), to 200 µL serum (Akirav et al. 2011) and ultimately to 0.2-1 mL of serum/plasma (Lehmann-Werman et al. 2016). It was, therefore, essential to standardise an internal protocol for isolating cfDNA from biofluid samples.

In this Chapter, we aimed to compare the absolute cfDNA yield extracted from serum/plasma samples using different DNA isolation protocols; the QiaAmp mini blood DNA extraction kit according to the manufacturer's protocol, a modified pH-column based method and a modified Triton/Heat/Phenol (THP) protocol. Then, we aimed to determine the efficiency and reproducibility of the chosen extraction method using different volumes of starting extraction material.

Urine is a non-invasive and easy to collect clinical sample (Cannas et al. 2009; Streleckiene et al. 2018), so we further aimed to, in parallel with optimising cfDNA extraction from human sera, study the viability of extracting cfDNA from urine samples. Urinary cfDNA, is another important source of cfDNA that has attaracted significant attention in the past decade. Urinary DNA is derived either from cells shedding from the urogenital tract or from the peripheral circulation after ultrafiltration by the renal glomeruli. Urinary short sequence cfDNAs, which are typically between 150-250 base pairs (bp) in length, are known as "transrenal DNA" (Su et al. 2004). Transrenal DNA contains important genomic information from various positions all over the body. The presence of transrenal DNA has been confirmed by various studies (Bryzgunova and Laktionov 2015, Bryzgunova et al. 2006, Melkonyan et al. 2008, Su et al. 2004, Tsui et al. 2012). The clinical utility of urinary cfDNA as a potential diagnostic biomarker for a variety of cancers and infectious diseases has been evaluated by studying the genetic variation of transrenal DNA such as mutations and methylation status, as well as DNA integrity and concentration (Bryzgunova and Laktionov 2015). Most recently, transrenal DNA was shown to yield sufficient concentrations of high-quality DNA for DNA methylation profiling in a study of diabetic chronic kidney disease (Lecamwasam et al. 2018). Therefore, in this chapter, we sought to optimise the required conditions for transporting urine samples (from the source of collection) and the storage conditions for the samples without losing the cfDNA content to be used in future projects.

Methods

5.1.1. Locally Collected Samples

Serum: Peripheral whole blood samples from randomly selected healthy volunteers were collected in a 5 mL SST™ tube (Becton Dickinson) containing a gel barrier to separate the serum after centrifugation. All samples were processed at room temperature within 2 hours from the time of venipuncture. Serum was separated from the cellular fraction by centrifugation at 2,000 g for 15 minutes. After centrifuging, each serum sample was transferred into a sterile Lobind Eppendorf tube and re-centrifuged at 16,000x g for 10 minutes. Afterwards, the supernatant was divided into aliquots of 300 µL. Aliquots were stored immediately at -80°C until cfDNA was extracted. The details of volunteers are summarised in General Methods, Chapter 2, Table 2.4.

Plasma: Peripheral whole blood samples from randomly selected healthy volunteers and were collected in 6 mL tri-potassium ethylenediamine tetra-acetic acid (K_3EDTA, VWR) tubes. Plasma was separated from the cellular fraction by centrifugation at 2,000 g for 15 minutes. After centrifuging, each plasma sample was transferred into a sterile Lobind Eppendorf tube and re-centrifuged at 16,000x g for 10 minutes. Afterwards, the supernatant was divided into aliquots of 300 µL. Aliquots were stored immediately at -80°C until cfDNA was extracted. The details of volunteers are summarised in General Methods, Chapter 2, Tables 2.6-2.9.

Urine: All urine samples used in this project were collected locally from healthy adult donors, University of Bristol, (Bristol, UK). 4 samples were used (males *n* = 2, females *n* = 2) in the experiment investigating the effect of DNA preservative on the stability of cfDNA. In the experiment comparing the effect of a commercial preservative with 40 mM EDTA on the stability of the cfDNA in urine, seven urine samples were used (males *n*=3, females *n*=4). The details of volunteers are summarised in General Methods, Chapter 2, Table 2.10.

5.1.2. Processing the Urine Samples (Experiment Design)

In order to study the effect of adding EDTA on cfDNA preservation in urine samples, the samples were divided into two groups; a control group (group A), where the samples were stored without any additions, and a test group (group B) where a final concentration of 40 mM EDTA was added to the samples. The final concentration of the preservative was achieved by adding 3.2 mL of sterile 0.5 M EDTA stock (pH 8.5) to exactly 40 mL urine sample. Next, 1 mL of sample from both groups were then aliquoted into sterile 1.5 mL Eppendorf tubes and stored as required and this is detailed in Figure 5.1.

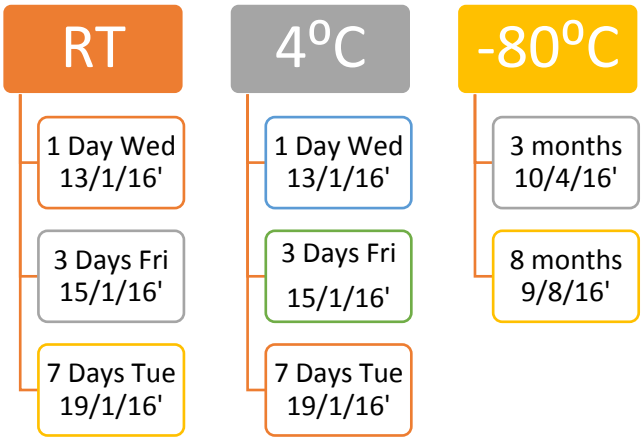


Figure 5.1: The experiment design of storage conditions and duration of Group A and Group B urine samples processing.

1 mL from both groups were stored at either room temperature (RT), 4°C for 1, 3 and 7 days. cfDNA sample from each group was extracted immediately and was used as a control for comparing the level of beta-globin cfDNA in the studied time points. Samples from both groups were stored in -80°C for long-term studies and was analysed after 3 and 8 months of storage.

In the experiment comparing efficiency of a commercial cfDNA urine preservative (Streck, Omaha, USA) to the 40 mM EDTA preservative in stabilising the cfDNA in urine samples, 5 mL of the Streck cfDNA urine preservative was added to 25-100 mL of urine sample as recommended by the manufacturer's instructions. The urine samples + streck preservative (Group C), Urine samples + 40 mM EDTA (Group B) were compared to a control group in which urine samples had no additives (Group A). One mL samples from groups A, B and C were aliquoted and stored at ambient temperature for cfDNA extraction at different time points (Immediately and then after 1, 3, 5 and 8 days).

5.1.3. DNA Preservative

5.1.3.1. Ethylene-Diamine-Tetra-Acetic Acid (EDTA)

A stock solution of 0.5M EDTA was prepared as described in the materials section in Appendix B. A final concentration of 40 mM EDTA was added to the urine samples.

5.1.3.2. Commercial cfDNA Urine Preservative

A liquid reagent that stabilises cfDNA in urine specimens was purchased from Streck (Omaha, USA). Five mL was added to 25-100 mL urine specimens as recommended by the manufacturer's instructions.

5.1.4. DNA Extraction

Plasma/ Serum

cfDNA was extracted from 0.2-2 mL of serum/plasma samples (as available for optimisation experiments) using three cfDNA extraction methods; the QiaAmp mini blood DNA extraction kit according to the manufacturer's protocol, the modified pH-column based method and the modified Triton/Heat/Phenol (THP) protocol (Keshavarz et al. 2015). The exact steps of the latter two methods are explained in Appendix A, Figure A.1 and Figure A.2.

Urine

In the first experiment, both QiaAmp Blood mini Kit (Qiagen) and QIAamp Circulating Nucleic Acid Kit (Qiagen) were used to extract cfDNA from urine samples following the manufacturer's protocol. In later experiments of optimising the cfDNA preservative, the cfDNA was extracted from 1mL urine sample using the QiaAmp Blood mini Kit only. Extracted cfDNA samples were stored at -20°C for later use.

5.1.5. cfDNA Quantification in Biofluid Samples

During this project, the assessment of cfDNA quantity in biofluid samples was carried out using either a **beta-actin assay** (for bisulfite converted samples) or a **beta-globin assay** (for non-bisulfite

treated), samples. Both the bisulfite specific beta-actin assay and the beta-globin assay were available locally from previous projects. The full details of the assays are summarised in Appendix A, Table A.1.

5.1.6. Droplet Digital PCR

The DNA concentration was measured using a standard droplet ddPCR protocol as described previously (Chapter 4, Section 4.5). The amplification conditions were as follows:

Polymerase Activation/ Denaturing	95°C	10 minutes	X1
Denaturing	94°C	30 seconds	X40
Annealing-Extension Step	Optimum Tm	1 minute	
Enzyme Deactivation	98°C	10 minutes	X1
Droplet Stabilizing	12°C	10 minutes	X1
Incubation	4°C	∞	∞

5.1.7. Statistical Analysis

The number of DNA copies per 1 µL of the target was analysed by the Quantasoft software (Bio-Rad, version 1.6.6.0320). The direct comparison between the levels of circulating hypomethylated beta cell-specific targets in test VS control group was achieved by Mann-Whitney testing for nonparametric data. For all statistical comparisons, a *p*-value of < 0.05 was considered to be statistically significant. Statistical analyses were performed with GraphPad/Prism 6 (version 6.07). The absolute level of beta cell-specific target cfDNA was presented in this thesis as the cfDNA copies in 1 mL of serum/plasma sample. The absolute level of circulating hypomethylated DNA was calculated as follows:

$$\textbf{Number of copies in 1 mL sample} = \frac{\text{Copies}}{\mu\text{L}}\text{in duplicate 1} + \frac{\text{Copies}}{\mu\text{L}}\text{in duplicate 2} * 25 * \left(\frac{1000}{\text{The volume of the original sample in } \mu\text{L}}\right)$$

Results

5.2.1. cfDNA extraction using different extraction methods

Serum samples were collected from healthy local adult control volunteers (*n*=7) and were used for cfDNA extraction. In this experiment, we were interested in quantifying the concentration of cfDNA in 300 µL serum samples. Therefore, the extracted cfDNA was analysed by ddPCR utilising a dual fluorescence beta-globin assay. Three extraction methods, THP and QIAamp DNA Blood Mini Kit, modified pH-column based methods were employed to compare cfDNA concentration in adult non-diabetic healthy individuals' samples. cfDNA was subjected to ddPCR to detect beta-globin concentration in the serum of non-diabetic healthy volunteers (*n*=7). Quantification of ddPCR data was fulfilled using cfDNA copies exploited to 1 mL sample. The mean ± SD of beta-globin cfDNA copies gained by the THP protocol was 2342 ± 1330 (range: 119-3967), the QIAamp DNA Blood Mini Kit was 5143 ± 3006 (range: 2158-9975), while for the pH-column based method was 6625 ± 4089 (range: 2358-13,418). Evaluation of the results of three extraction methods showed the higher efficiency of cfDNA extraction was gained when the modified pH-column based protocol for cfDNA extraction was used ($p < 0.05$, Kruskal-Wallis test) (Figure 5.2).

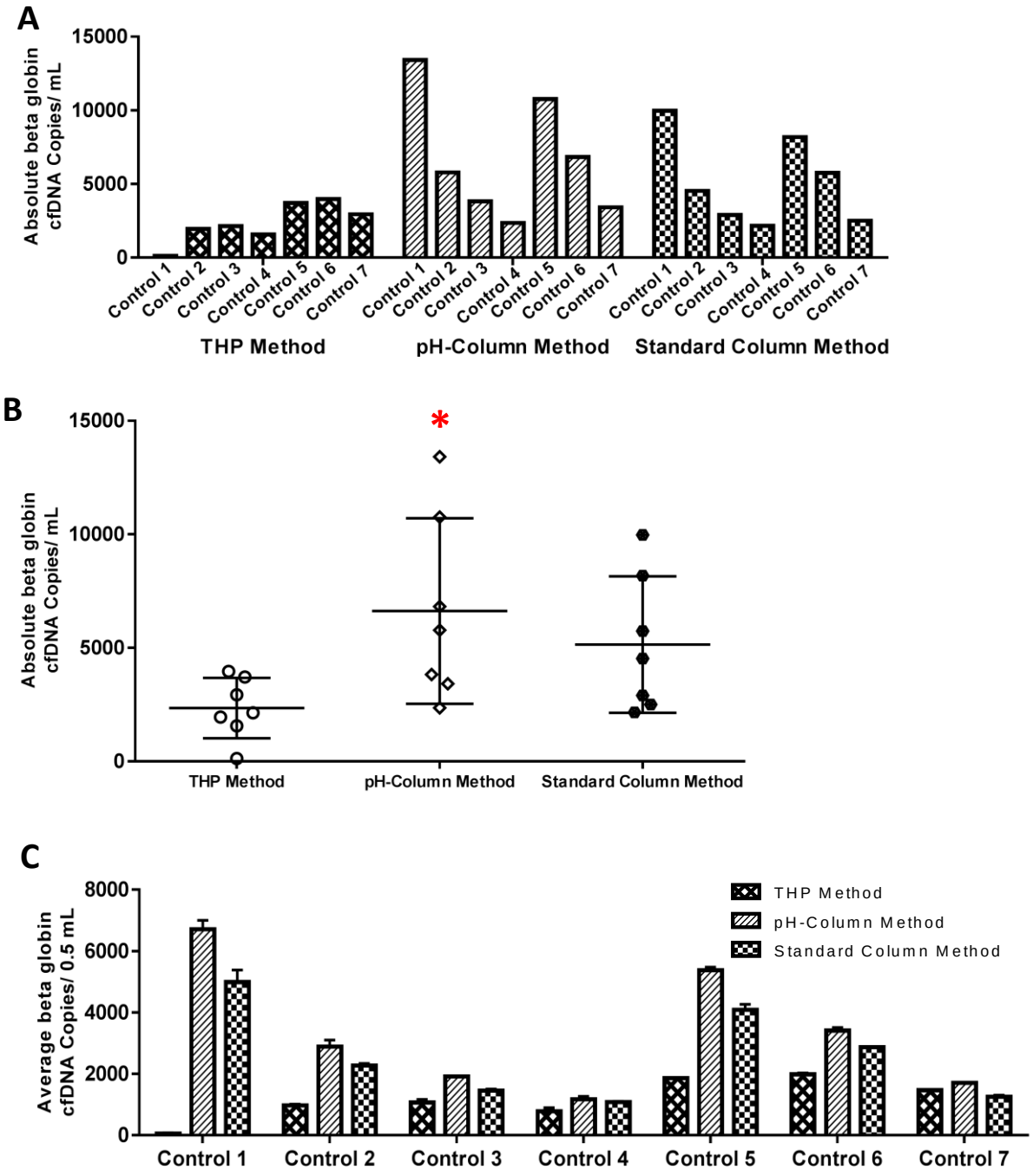


*Figure 5.2: Comparison of beta-globin cell-free DNA (cfDNA) recovery from serum using three DNA isolation methods. cfDNA was extracted from 300 µL serum collected from adult healthy volunteers (n=7). cfDNA was extracted via the THP protocol, the pH modified column protocol and the standard protocol of QIAamp blood mini kit. Each sample was quantified by droplet digital PCR utilizing beta-globin assay in duplicates. A) The columns represent the absolute quantification of beta-globin cfDNA copies per 1 mL serum sample. B) The scatter dot plot represents the absolute number of cfDNA copies in all tested samples, the middle line indicates the mean value ± SD (error bar). C) The Data are presented as mean of the duplicate value (column) ± SD (error bar). The absolute yields of cfDNA generated by the pH-column method were significantly higher (*p < 0.05) than the yields extracted by the THP protocol (Kruskal-Wallis test).*

5.2.2. Comparing the Modified pH Column-Based Method to Other Published Protocol

Recently, an enhanced Qiagen cfDNA isolation protocol has been described by (Tersey et al. 2016). The authors reported that this enhanced protocol was able to extract cfDNA in detectable yield from as low as 50 µL starting material. We, therefore, sought to compare the cfDNA yield from our pH modified protocol to the published enhanced DNA extraction protocol using two plasma volumes, 0.5 and 2.5 mL, collected from local healthy adult volunteers (*n*=3), samples were described in General Methods, Chapter 2, section 2.3.6. Due to limited availability, plasma samples were used in this experiment. Based on previous optimising experiments (carried out by Dr Jody Ye), we did not observe a significant difference in cfDNA quantity between plasma and serum samples that were separated within 2 hours post venepuncture. The extracted cfDNA by both methods were subjected to ddPCR in a beta-globin probe-based assay. The mean ± SD of cfDNA concentration corrected to 1 mL plasma obtained by the modified pH-column based protocol was 197 ± 59 (range: 125-215), while for the published enhanced DNA extraction protocol the yield was 44 ± 17 (range: 19-42) from 0.5 mL plasma. The beta-globin cfDNA obtained by modified the pH-column based protocol was 220 ± 98 (range: 159-265), while for the published enhanced DNA extraction protocol was 32 ± 17 (range: 25-59) from 2.5 mL plasma. Collectively, the data show that the pH modified protocol was significantly more efficient in extracting cfDNA from both tested plasma volumes (Wilcoxon matched-pairs signed rank test, Figure 5.3).

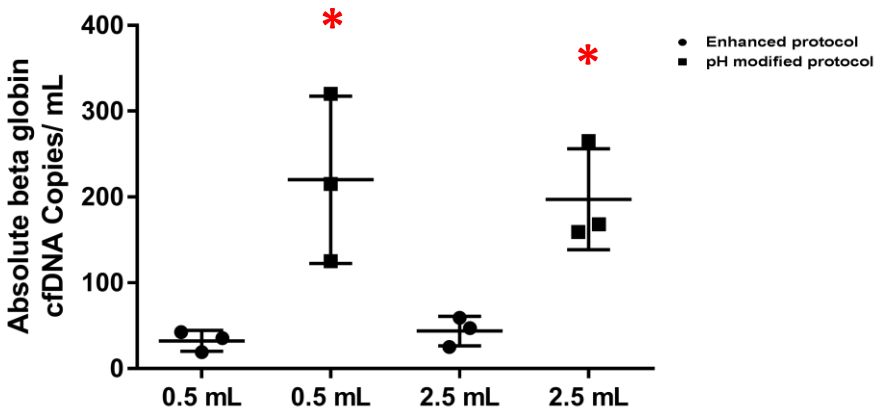


Figure 5.3: Comparison between two cell-free DNA (cfDNA) isolation protocols.

The data comparing the cfDNA extracted from 0.5 mL and 2.5 mL of plasma samples using two modified protocols for the QIAamp blood mini DNA extraction kit; enhanced column protocol and our developed pH column-based protocol. Plasma was isolated from healthy blood samples donated by 3 volunteers; cfDNA was quantified by droplet digital PCR utilizing the beta-globin assay. The dot plot represents the absolute beta-globin concentration in 1 mL of plasma sample. The results suggest that cfDNA extraction using the pH-modified column protocol significantly enhanced ($p < 0.05$) cfDNA recovery when compared to the other cfDNA extraction protocols (Wilcoxon matched-pairs signed rank test).*

5.2.3. The Optimum Plasma Volume Needed for cfDNA Extraction

Next, we aimed to determine the effect of starting material volume on extracted cfDNA yield and whether there is a linear correlation between the cfDNA concentration and the initial plasma volume extracted. To achieve this goal, three different plasma volumes of 2.5 mL, 1 mL and 0.5 mL were obtained from the same healthy donors used in the previous experiment. The cfDNA yield extracted by each method was analysed by ddPCR utilising a dual fluorescence-based human beta-globin assay. The results showed that the beta-globin cfDNA average was 110± 49, 201± 30 and 493± 98 copies /µL in 0.5, 1.0 and 2.5 mL gained from 3 plasma samples respectively (Figure 5.4A). Next, we tested the consistency of DNA extraction from each plasma volume. Therefore, the absolute beta-globin cfDNA copy number/1 mL was determined for each sample in each group (Figure 5.4B). Interestingly, the average of the absolute beta-globin copy number was around 200 copies/mL in all tested samples (Figure 5.4B). The ddPCR analysis also revealed a linear correlation (R^2= 0.9998) between the cfDNA concentration obtained and the plasma volume extracted. The beta-globin cfDNA copy number increased in parallel with the increased plasma volume (Figure 5.4 C).

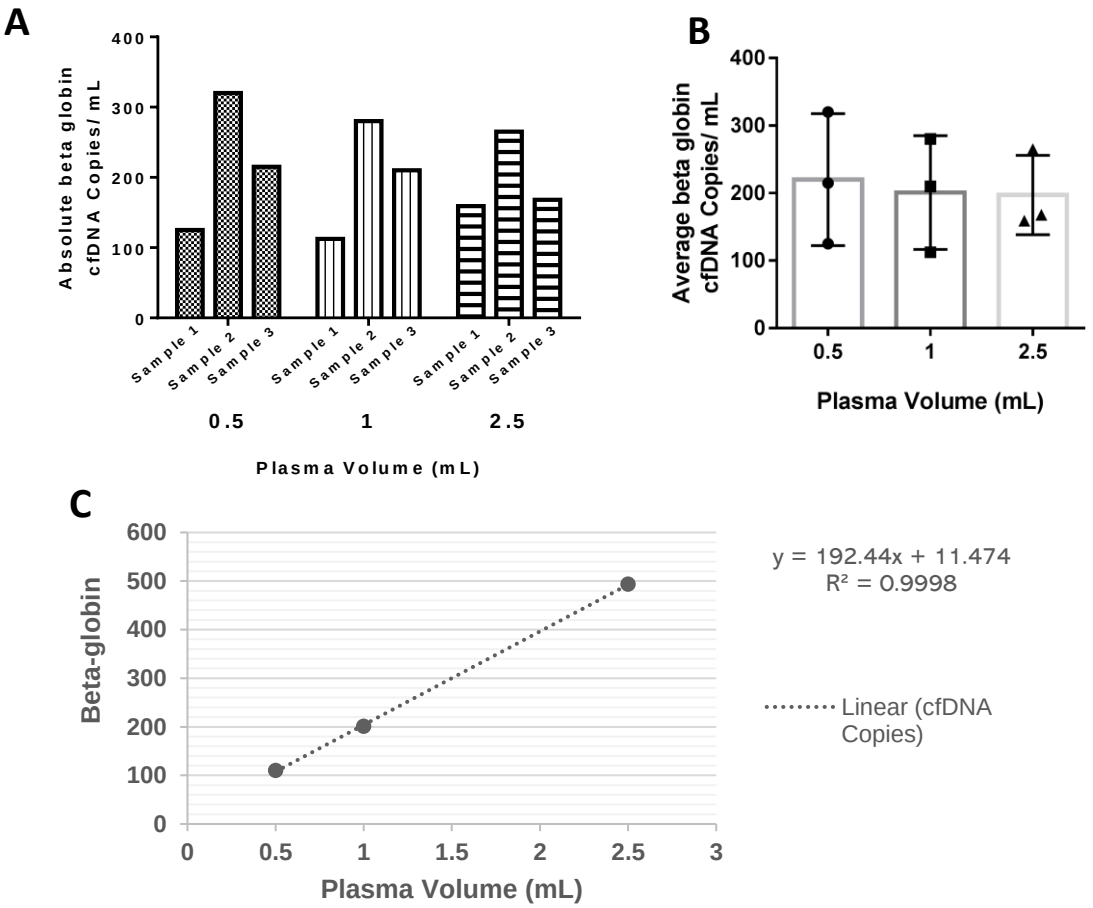


Figure 5.4: Comparison of the beta-globin yield extracted from different plasma volumes.

cfDNA was extracted from either 0.5, 1 or 2 mL plasma samples donated by 3 healthy volunteers. cfDNA was extracted using the pH modified column-based protocol and each sample was quantified by droplet digital PCR utilizing the beta-globin assay. A) The column represents the absolute beta-globin cfDNA copy number in 1 mL plasma sample B) The columns represent mean ± SD (error bar) *of absolute beta-globin cfDNA copy number per 1 mL in each tested plasma volume. C) The dot plot represents the mean value of the absolute beta-globin copy number of 3 samples in each tested group. The data here indicates the productivity of cfDNA extraction using a pH modified column-based protocol.*

5.2.4. The Effect of DNA Preservative on Protecting cfDNA from Degradation in Urine Samples

5.2.4.1. The Integrity of Extracted cfDNA in Urine Sample

Morning and mid-day urine samples (4 mL) collected from a single adult healthy donor were used for cfDNA extraction. The extracted cfDNA was then subjected to bisulfite treatment and amplified with bisulfite specific beta-actin primers. The PCR products were then visualised on a 1.8 % agarose gel (Figure 5.5). The result shows that cfDNA extracted from urine was intact and confirms that urine samples contain sufficient cfDNA concentration that could be used in downstream analysis.

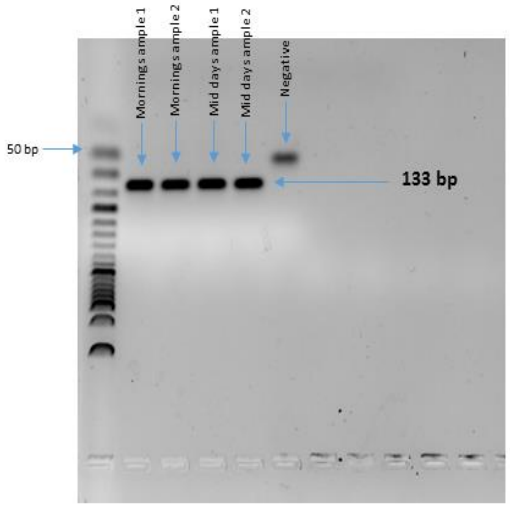


Figure 5.5: The PCR products of bisulfite converted cfDNA extracted from 4 mL control urine samples.

The urine sample was collected from a healthy donor (n=1) and cfDNA was extracted from 4 mL samples using the QIAamp Circulating Nucleic Acid Kit. cfDNA was then subjected to bisulfite treatment and was amplified with beta-actin primers (133 bp). The samples were visualized in a 1.8% agarose gel by staining them with Midori Green and the PCR product sized with an adjacent 50 bp ladder. These data indicate that cfDNA extracted from urine samples could be a successful alternative source of cfDNA to serum or plasma for analysis of beta cell death.

5.2.4.2. The Need for a Preservative in the Collected Urine Samples

An aliquot of urine specimens that had been stored in -80^{o}C for ten days was used in this experiment. cfDNA was extracted from different urine volumes: 0.2, 0.4, 0.6, 0.8, 1 mL and 4 mL in duplicate using a QiaAmp Mini Blood Kit. The cfDNA samples were bisulfite treated and then analysed by ddPCR utilising the beta-actin assay (FAM), and a previously developed differentially methylated *INS* assay at CpG +367+374 (FAM/VIC) by Dr Jody Ye (described in Appendix A, Table A.1). The result shows that both assays, beta-actin and the *INS* had a very low yield in all tested volumes (Figure 5.6). This experiment suggests a possibility of cfDNA degradation occurrence in the urine sample and highlights the possible need to add DNA preservative to the urine samples at the time of collection.

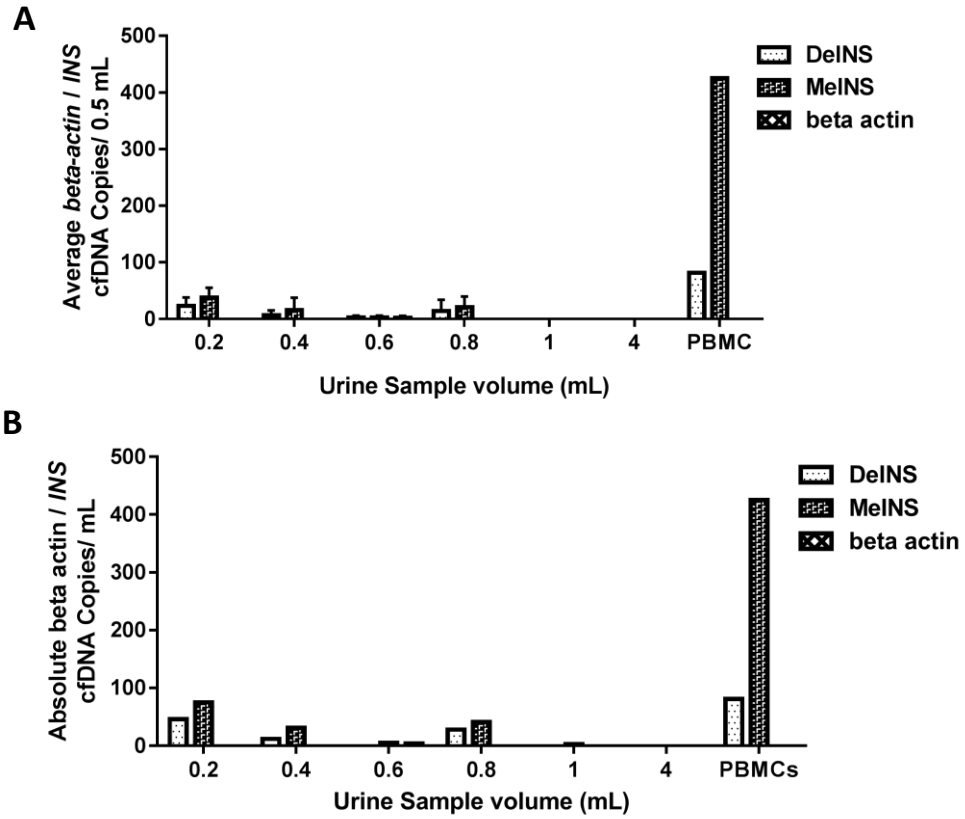


Figure 5.6: The absolute concentration of differentially methylated INS (CpG +367+374) in different volumes of a urine sample.

The cfDNA was extracted from either 0.2, 0.4, 0.6, 0.8, 1 or 4 mL urine samples via QIAamp DNA blood mini kit. The urine sample was collected from a healthy donor (n=1) and was quantified by a droplet digital PCR differentially methylated INS (CpG +367+374) probe assay in duplicate. A) The data in this graph are presented as mean of the duplicate value (column) ± SD (error bar) of each tested sample. B) The graph represents the absolute INS and beta-actin cfDNA copy number per 1 mL of the extracted cfDNA from different urine volumes. Although some amplification was seen in INS, no amplification of beta-actin suggests cfDNA degradation in the stored urine sample may have occurred.

5.2.5. Effects of Adding EDTA and Storage Temperature on the Stability of cfDNA in Urine

When the urine from the previous experiment was stored at -80^{o}C for ten days without the addition of DNA preservative, a decrease in the beta-actin signal was observed (the copy number of beta-actin decreased from 200 copies/mL to 4 copies/mL). We sought to examine whether that reduction was an effect of nucleic acid degradation, especially when previous reports had reported degradation occurrence in urine samples when stored without the addition of nucleic acid preservative (Cannas et al. 2009). To investigate this, we studied the effect of adding a DNA preservative (40 mM EDTA) to a morning first-urine sample collected from a single healthy adult donor which was stored at different storage temperatures and duration conditions that would mimic sample transportation collection for the Bart's Oxford study of type 1 diabetes coordinated from this research group.

5.2.5.1. The Effect of 40 mM EDTA on Short-Term Storage

To investigate the effect of adding DNA preservative to urine samples which will be stored for a short time (≤ 7 days). In this experiment, we were interested in quantifying the cfDNA concentration in 1 mL urine sample. Therefore, the extracted cfDNA was analysed by ddPCR utilising a dual fluorescence beta-globin assay. Aliquots of a urine sample (1 mL), conducted from a single healthy adult volunteer, were stored either at room temperature or 4^{o}C with and without the addition of 40 mM EDTA. cfDNA was extracted at baseline, 3 and 7 days using a QiaAmp mini blood kit and the beta-globin yield quantified by ddPCR. The results show that adding 40 mM EDTA was essential to prevent cfDNA degradation in samples stored in conditions mimicking sample transport (Figure 5.7).

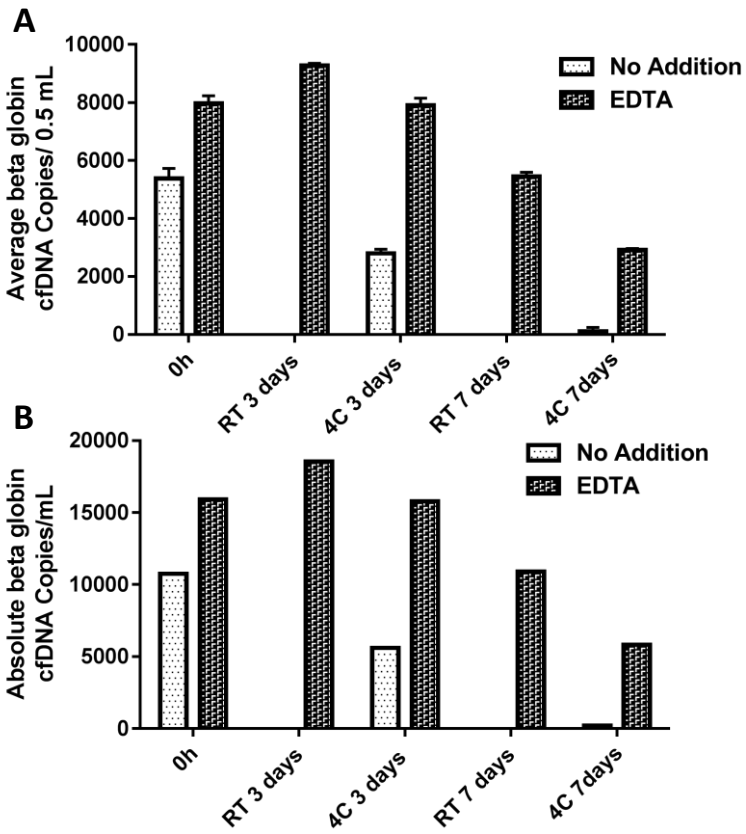


Figure 5.7: The effect of adding 40 mM EDTA on DNA stability in urine samples when stored at different temperatures and durations.

The samples were collected from a single healthy donor and were stored either with or without the addition of DNA preservative (EDTA, 40 mM) at ambient temperature (RT) or 4^{o}C for 1, 3 and 7 days. The cfDNA was extracted from 1 mL urine samples via the QIAamp DNA blood mini kit and were quantified in duplicate by droplet digital PCR utilizing the beta-globin assay. A) The data in this graph represents the mean of the duplicate value (column) ± SD (error bar) of each tested sample per 1 mL urine sample. B) The graph represents the absolute beta-globin cfDNA copy number per 1 mL urine samples. The data show that adding 40 mM EDTA to the urine sample was essential for protecting cfDNA from degradation for up to 7 days when urine samples were stored at ambient temperature or at 4^{o}C.

5.2.5.2. The Effect of 40 mM EDTA on Long-Term Storage

Next, we sought to test the effect of adding EDTA to urine samples stored at -80^{o}C and for a long-term period (3 and 8 months). Therefore, we stored urine samples collected locally from healthy volunteers (*n*=4; male= 2, female= 2) with and without adding 40 mM EDTA for up to 8 months. As we were interested in comparing the cfDNA concentration in the presence or absence of DNA preservative, no baseline was included. Instead the cfDNA concentration in Group A samples (no preservative) was compared to those in Group B (40 mM EDTA added) at each time point. At the appropriate time, cfDNA was extracted from a 1 mL urine sample using the QiaAmp mini blood kit. Then, cfDNA samples were amplified with ddPCR utilising the beta-globin assay. The absolute beta-globin cfDNA copies/ mL of treated samples were compared to the control (No EDTA) samples' (Figure 5.8). Although we observed a decrease in the concentration of beta-globin in samples that were stored at -80^{o}C for 8 months, the data show that there was no statically significant difference between the two test groups ($p > 0.05$, Mann–Whitney U test).

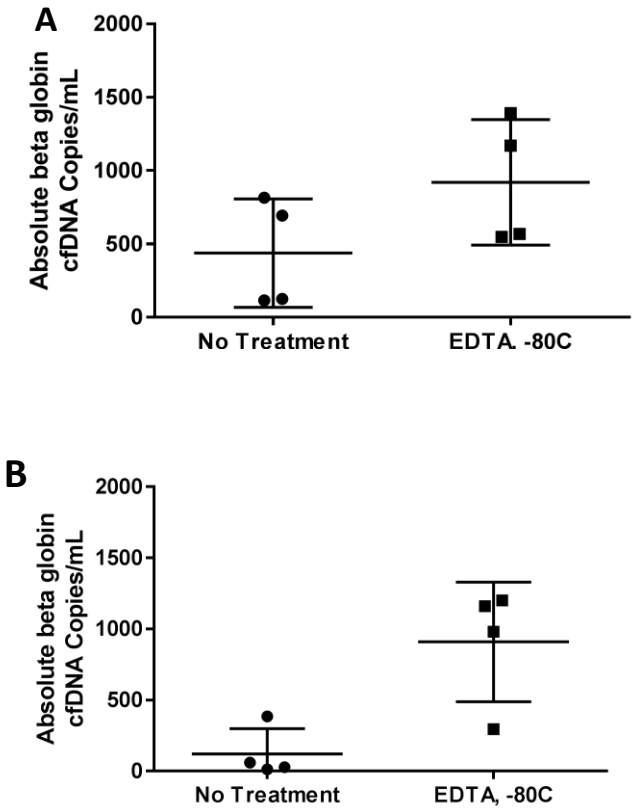


Figure 5.8: The effect of long-term storage on untreated and EDTA treated urine samples.

Urine samples were collected from healthy adult volunteers (n=4), 1 mL of each samples (with or without EDTA addition) was stored for 3 and 8 months at -80^{o}C. At the appropriate time, cfDNA was extracted from 1 mL sample using the QiaAmp mini blood kit, the total cfDNA was used as a template for beta-globin quantification via ddPCR. A) The scatter plot represents the absolute beta-globin cfDNA copy number in 1 mL urine sample after storage for 3 months at *-80^{o}C*. B) scatter plot represents the absolute beta-globin cfDNA copy number in 1 mL urine sample after storage for 8 months at *-80^{o}C*. The line *indicates the mean value and ± SD (error bar) of each tested sample.*

5.2.5.3. The Effect of Adding a Commercial Preservative product on cfDNA Stability in Urine Samples

Our next goal was to test a substitute preservative that can be safely delivered to the patients at home for urine collection. Therefore, 5 mL commercial preservative from Streck (Omaha, USA) was added to 25-100 mL freshly collected urine samples, as recommended by the manufacturer's instructions. Urine samples collected from local adult healthy volunteers (*n*= 8; male= 4, female= 4) with and without adding 40 mM EDTA or the commercial preservative. The samples were divided into 1 mL aliquots and stored at ambient temperature for 1, 3, 5, 8 days. At appropriate time points, total cfDNA was extracted using a QiaAmp mini blood kit and amplified by ddPCR utilising the beta-globin assay. The results showed both cfDNA preservatives maintained the stability of the cfDNA in the urine for up to 5 days at ambient temperature. The results also showed that the commercial preservative failed to stabilise the cfDNA in the sample when stored for a period exceeding 5 days (Figure 5.9). However, more samples will be tested to confirm this conclusion.

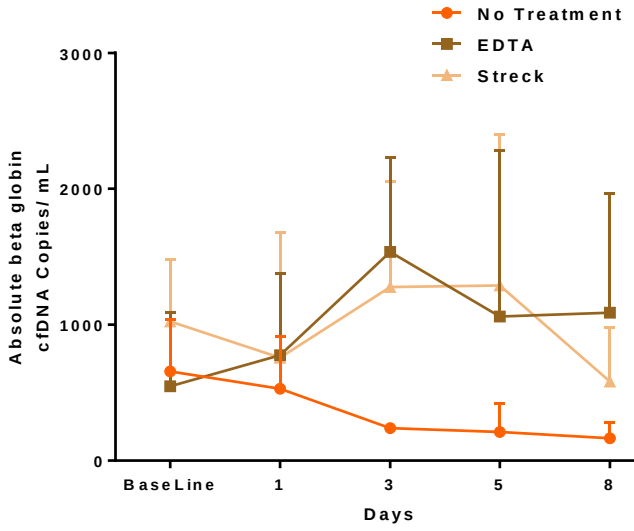


Figure 5.9: Comparison of the effect of two preservatives on the stability of cfDNA in urine samples with and without DNA preservative.

Urine samples were collected from healthy volunteers (n=8), and 1 mL of each sample, with EDTA (40 mM), the commercial preservative or without preservative was stored for 0, 1, 3, 5 or 8 days at ambient temperature. At the appropriate time, cfDNA was extracted from 1 mL sample using the QiaAmp mini blood kit, the cfDNA was used then as a template for beta-globin quantification via ddPCR. The graph represents the absolute quantification of beta-globin per 1 mL urine sample, each dot represents the mean of beta-globin absolute value in 8 samples. The data showed that adding 40 mM EDTA or a commercial preservative to urine samples stabilized the cfDNA for up to 5 days. Only 40 mM EDTA prevented cfDNA degradation for up to 8 days when the samples were stored at room temperature.

5.2.6. Optimising of Urine Volume Required for Successful cfDNA Analysis

Next, we tried to identify the optimum volume of a urine sample that could be used for cfDNA extraction. Another fresh mid-day DNA sample was collected from a single healthy volunteer (80 mL), 14 mL of the sample was used in this experiment, the rest of the sample was aliquoted into 15 mL volumes and stored at -80°C. cfDNA was extracted from 200 µL, 400 µL, 600 µL, 800 µL, 1 mL and 4 mL [12] in duplicate using two DNA extraction kits; the QiaAmp Circulating Nucleic Acid Kit (NA) and the QiaAmp Mini Blood Kit (QBL). The extracted cfDNA was bisulfite treated and amplified by ddPCR utilising the beta-globin assay. The result suggests that extraction of cfDNA from urine was more efficient when the QiaAmp mini blood kit was used (Figure 5.10). The data also indicated that using volumes of less than 0.5 mL could have a positive impact on the resulting cfDNA yield (Figure 5.10D). Therefore, a volume of 400 µL will be used in subsequent experiments.

[12] cfDNA was isolated from 4 mL urine using the QiaAmp Circulating Nucleic Acid Kit only.

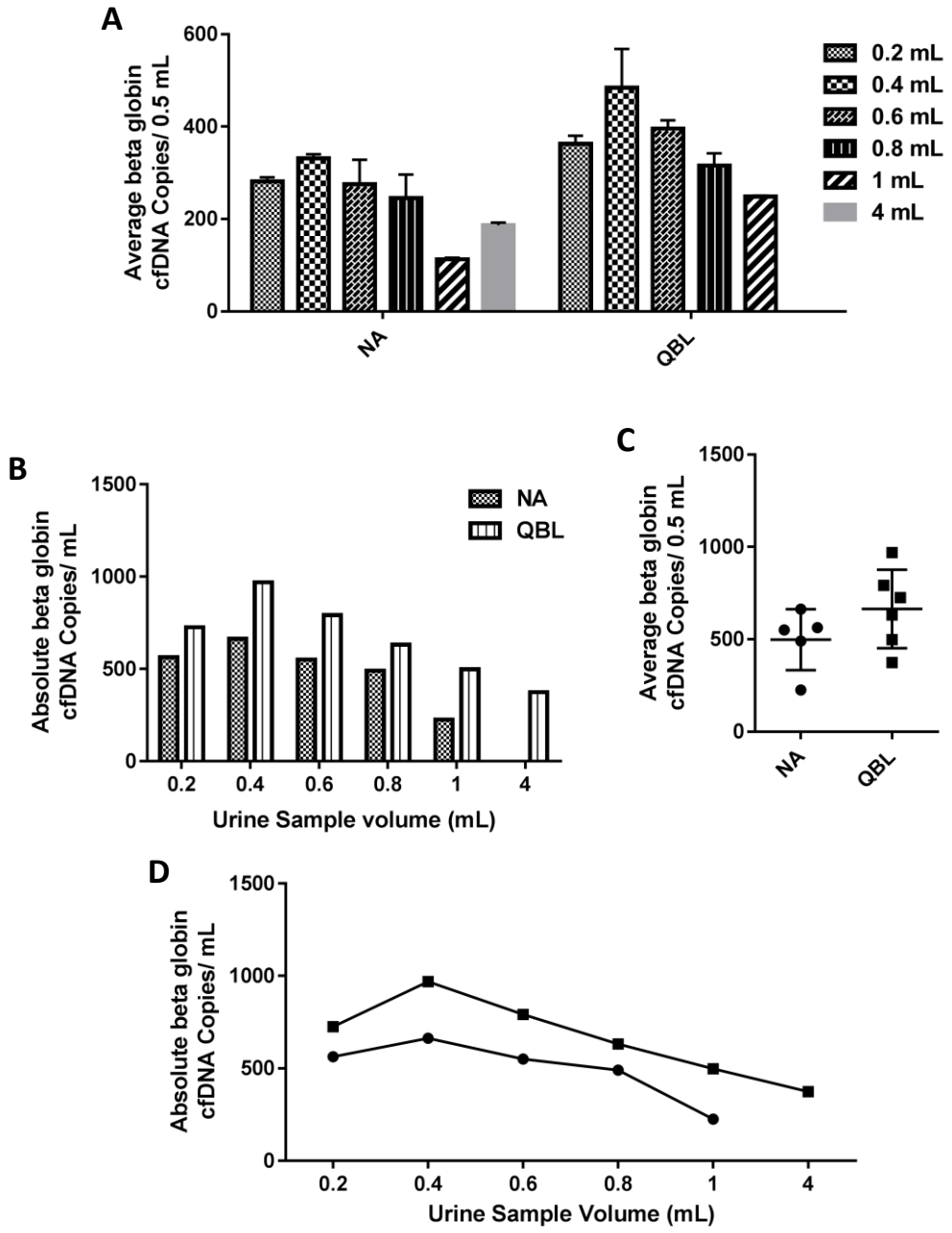


Figure 5.10: Comparison of beta-globin yield extracted from different volumes of urine from healthy volunteers.

cfDNA was extracted from either 0.2, 0.4, 0.6, 0.8, 1 or 4 mL urine samples via the QIAamp Circulating Nucleic Acid Kit (NA) or QIAamp mini blood nucleic acid kit (QBL). The urine sample was collected from a healthy donor (n=1) and quantified by droplet digital PCR utilizing the beta-globin probe assay in duplicate. A) The data in this graph is presented as mean of the duplicate values (column) ± SD (error bar) of each tested sample. B), C) & D) The graph represents the absolute beta-globin cfDNA copy number per 1 mL of the extracted cfDNA from different urine volumes. The data suggests that the QiaAmp blood mini kit extracted cfDNA from urine samples more efficiently in comparison to the QIAamp Circulating Nucleic Acid Kit. The data also suggest that extracting cfDNA in volumes <0.5 ml could result in higher cfDNA yield.

Discussion

This chapter summarises the work performed to optimise the conditions for cfDNA extraction from biofluid samples, e.g. serum, plasma, which are known to be challenging biological samples for DNA extraction. The results suggest that a large volume of starting biological material may be necessary in order to detect low concentration signals such as fragmented DNA released from dying beta cells. Therefore, where possible 1 mL of plasma/ serum will be used during this project for extracting DNA using the modified pH-column based method.

A series of experiments was also set up to optimise the isolation and analysis of urinary DNA. The urinary DNA is another important source of cfDNA, and it has been attracting significant attention in the past decade due to the large volume of sample that can be collected from the patient and being an easy to access source of sample. In the longer term, it was hoped that the strategies developed could be used to collect samples from participants in the Bart's-Oxford (BOX) family study, the longest-running family study of type 1 diabetes, which is co-ordinated from Bristol. The data in this chapter showed clearly that adding cfDNA preservative is essential to protect the cfDNA from degradation during sample transport from the source of the collection to the laboratory. One study which examined the optimum conditions for storing transrenal DNA at different temperatures, e.g. room temperature, 4^{o}C and - 80^{o}C was carried out by Cannans and colleagues (Cannas et al. 2009). They reported that site-specific differences in urine composition significantly affected the stability of urinary cfDNA during storage at all temperatures tested. Additionally, they highlighted the importance of adding DNA preservative agents, e.g. EDTA at a high concentration (40 mM) to protect the urinary DNA from degradation. Therefore, we tested the viability of adding 40 mM of EDTA to 1 mL urine samples collected from healthy volunteers and showed that adding 40 mM EDTA was essential to protect the urinary DNA from degradation when the samples were stored at ambient and 4^{o}C for ≤ 7 days. Although not statistically different, we noticed that a there was a difference in the number of beta-globin copies between the EDTA treated and control groups when the samples were stored for

8 months at -80°C compared with controls (*n*=4, Figure 5.8). This experiment will, therefore, be followed up in the future with more samples. As adding 40 mM EDTA was shown to be essential for stabilising transrenal DNA in the urine samples, it raises an issue for the Bart's Oxford team in terms of delivering tubes with 40 mM EDTA to the general population who participate in this study. Another preservative reagent was therefore researched that could be safely delivered to patients. According to the Safety Data Sheet of Cell-Free DNA urine preservative (Streck, Omaha, USA), the product does not contain any hazardous chemicals at concentrations of 1% or higher, does not contain any known carcinogens, category 1 germ cell mutagens, reproductive toxicants, respiratory or skin sensitizers at concentrations of 0.1% or greater (https://www.streck.com/wp-content/uploads/sync/Collection/Preservatives/Cell-Free_DNA_Urine_Preserve/04_SDS/01_Cell-Free_DNA_Urine_Preserve_SDS_Exemption_Letter.pdf). Therefore, we tested the efficiency of this commercial reagent in stabilizing transrenal DNA when stored at ambient temperature. Our data indicated that adding the Streck preservative stabilised the urinary DNA for up to 5 days when stored at ambient temperature. The data in this chapter also indicate that using a volume of 0.4 mL is recommended as high urine volume e.g. 1 and 4 mL resulted in decreased cfDNA yield, which could due to the nature of the urine samples (high salt concentration) that could interact with the column during the cfDNA extraction using the column based method.

Chapter 6: Monitoring Beta Cell Death in Clinical T1D Samples

6. Introduction

As shown previously in chapter 4, hypomethylated CpG sites within *MIR-200c, MIR-141* and *INS* were found specifically in beta cells. Therefore, any beta cell destruction would release hypomethylated DNA (at these specific CpG sites) into the circulation. Circulating cell-free DNA (cfDNA) has been used as a biomarker for disease development (Swarup and Rajeswari 2007). Serum and plasma are non-invasive, easy to obtain sources of DNA, and when using a highly sensitive analysis method such as ddPCR, beta cell-derived cfDNA can be detected at low concentrations. As the primary goal of this project was to develop a methylation-based biomarker for monitoring beta cell death, we aimed in this chapter to assess whether the developed multiplex assay could detect beta cell genome in the circulation of recently diagnosed T1D adult and children patients, those who were at risk of developing the condition and T1D patients who received clinical islet transplantation.

Results

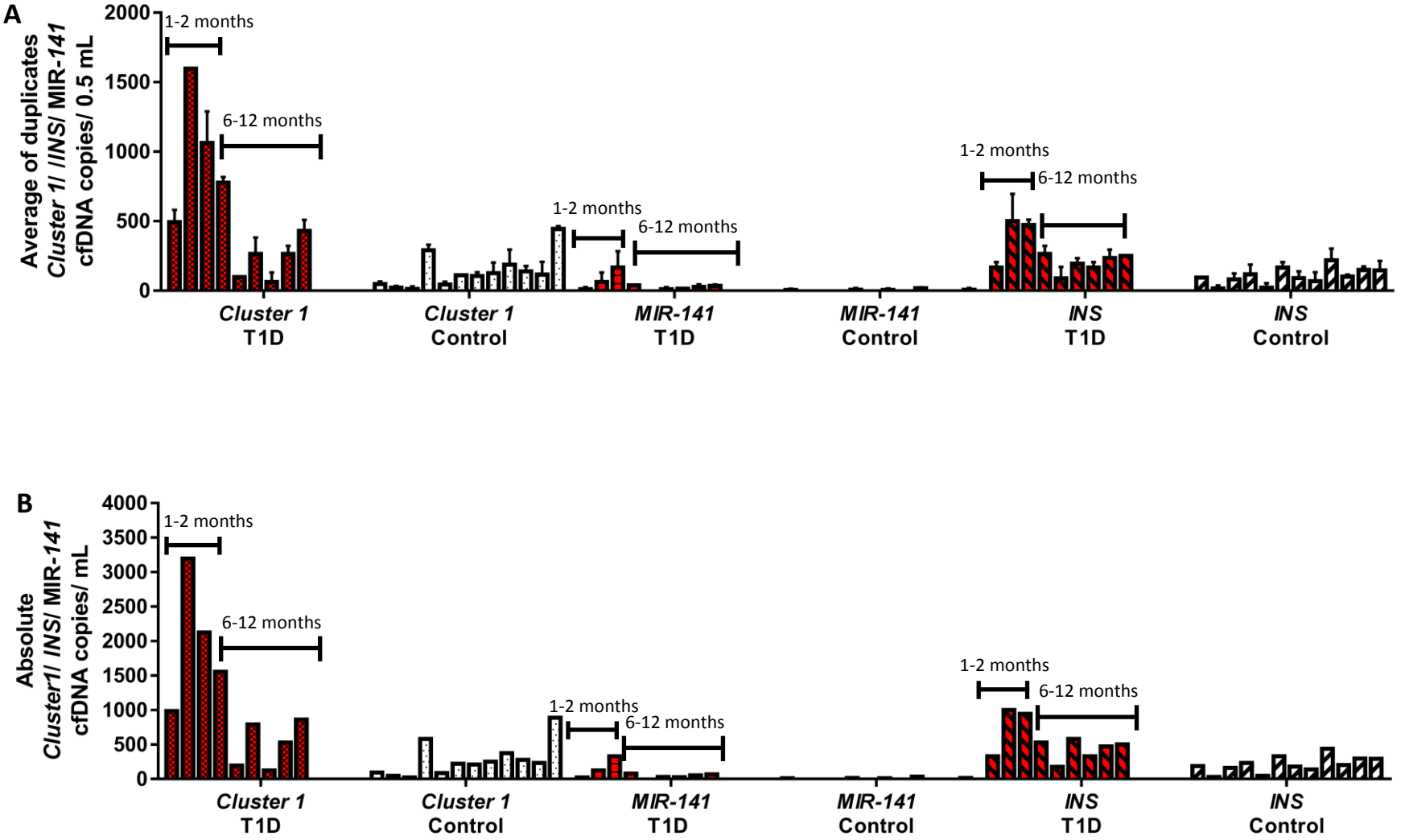
6.1.1. *MIR-141*/*MIR-200c* Cluster 1/*INS* Exon 2 Assay Triplex as a Biomarker for Beta Cell Death

The viability of the triplex assay to detect beta cell death in the circulation was tested on recently diagnosed paediatric subjects (*n*=9) who have T1D duration of 1 year or less against non-diabetic subjects (*n*=12) who were age and gender-matched, detailed characterization of the participant is described in General Methods, Chapter 2 Tables 2.4 & 2.5. Crucially these samples were collected, and serum isolated as rapidly as possible after venepuncture to prevent contamination with genomic DNA fragments from PBMC which causes false positive results with the insulin cfDNA assay (data not shown). The level of hypomethylated *MIR-200c* Cluster 1, *MIR-141* and *INS* were analysed in bisulfite-converted DNA that was isolated from 900 µL serum via ddPCR. Further, the total number of positive droplets in each sample was calculated and referred to as the "multiplex". The data showed that there was a significant difference between the level of circulating beta cell cfDNA in participants who were close to diagnosis and their corresponding control subjects ($p < 0.01$, Mann-Whitney U-test). A comprehensive look at the data shows that there was a higher level of beta cell cfDNA in children who had a disease duration of < 6 months in comparison with those who were sampled post six months of diagnosis (Figure 6.1). This trend was consistent in the single plex (Figure 6.1B) as well as in the multiplex (Figure 6.1C) results. Next, we tested the ability of the assay to discriminate between participants who were sampled after 6 months post-diagnosis (*n*=6), the Mann-Whitney U-test shows that there were no significant differences in beta cell fragments level between the tested group and their age and gender-matched control subjects ($p > 0.05$). Further, we sought to investigate the effect of multiplexing of several targets in discriminating beta cell cfDNA in circulation. The power of discrimination (PD) was calculated by dividing the number of cfDNA copies resulting from the single plex assay in all tested samples by those resulting from the multiplex assay. The highly efficient assay is expected to have power of discrimination near to the value =1. The data show that multiplexing

several beta cell-specific targets increased the power of discrimination three times in comparison with targeting *INS* in single plex alone (Table 6.1).

Table 6.1: The effect of multiplexing several targets in enhancing the discrimination power of a beta cell death signal. Multiplexing several beta cell specific targets had enhanced cluster 1 assay by 1.5 times (PD= 0.65) and almost 3.3 times in the INS (DP= 0.30). The effect of multiplexing was highest with MIR-141 as it enhanced the assay by 21.6 times (DP= 0.046).

Sample	**T1D Duration (Months)**	***MIR-200c Cluster 1* Copies/mL**	***MIR-141* Copies/mL**	***INS* Copies/mL**	**Multiplex**
BCD10	0.263555251	986.1111	20.8333333	333.3333	1340.278
BCD4	0.823610158	3194.444	126.388889	1000	4320.833
BCD5	1.943719973	2125	333.333333	944.4444	3402.778
BCD9	6.358270419	1555.556	77.7777778	527.7778	2161.111
BCD2	9.323266987	194.4444	0	177.7778	372.2222
BCD6	9.883321894	791.6667	31.25	583.3333	1406.25
BCD1	10.50926561	125	27.7777778	330.5556	483.3333
BCD3	12.38709677	527.7778	54.1666667	472.2222	1054.167
BCD8	12.94715168	861.1111	68.0555556	500	1429.167
	Sum	10361.11	739.583333	4869.444	15970.14
	Power of Discrimination	0.65	0.046	0.30	


A
Average of duplicates
Cluster 1/ /INS/ MIR-141
cfDNA copies/ 0.5 mL
2000
1500
1000
500
0
1-2 months
6-12 months
1-2 months
6-12 months
1-2 months
6-12 months
Cluster 1
T1D
Cluster 1
Control
MIR-141
T1D
MIR-141
Control
INS
T1D
INS
Control
B
Absolute
Cluster1/ INS/ MIR-141
cfDNA copies/ mL
4000
3500
3000
2500
2000
1500
1000
500
0
1-2 months
6-12 months
1-2 months
6-12 months
1-2 months
6-12 months
Cluster 1
T1D
Cluster 1
Control
MIR-141
T1D
MIR-141
Control
INS
T1D
INS
Control

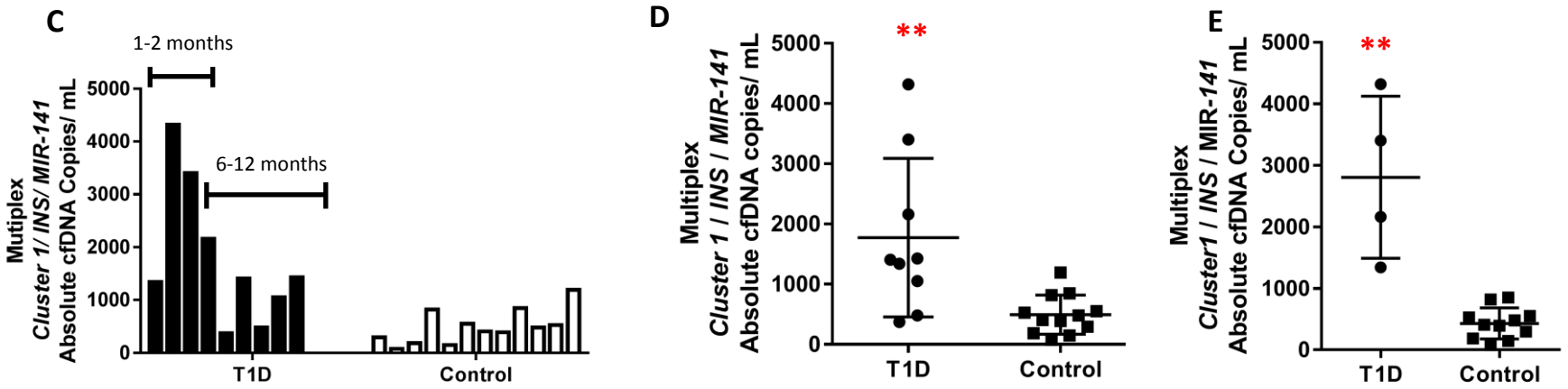


Figure 6.1: Circulating hypomethylated MIR-200c cluster 1/ INS / MIR-141 cfDNA levels in recently diagnosed paediatric human cohort.

The level of hypomethylated multiplex assay of MIR-200c cluster 1/INS/MIR-141 DNA was measured via ddPCR in serum samples from recently diagnosed T1D (≤ 1-year diagnosis, n=9) and age-matched nondiabetic control subjects (n=12) in duplicates. Serum samples were separated rapidly after venepuncture to avoid PBMC genomic DNA contamination. The data are represented as the level of hypomethylated cfDNA copies in 1 mL of serum. A) The columns are presented as mean of the duplicate values (column) ± SD (error bar). B) The absolute level of each hypomethylated target in recently diagnosed T1D subjects and non-diabetic controls. C) The level of absolute multiplex targets in 1 mL of T1D and non-diabetic serum samples. D) Comparison between the absolute level of circulating hypomethylated cfDNA in recently diagnosed T1D, within 1 year, and healthy control subjects ($p = 0.0013$, Mann Whitney U-test; $^{**}p < 0.01$). E) Comparison between the absolute level of circulating hypomethylated cfDNA in the recently diagnosed T1D, within 6 months, and healthy control subjects ($p = 0.0015$, Mann Whitney U-test; $^{**}p < 0.01$). The data shows that the level of circulating hypomethylated MIR-200c cluster 1/INS/MIR-141 cfDNA was significantly elevated in recently diagnosed T1D especially in those who were diagnosed within 6 months. The data also demonstrated that the level of hypomethylated MIR-200c cluster 1/INS/MIR-141 declined post 6 months of diagnosis.

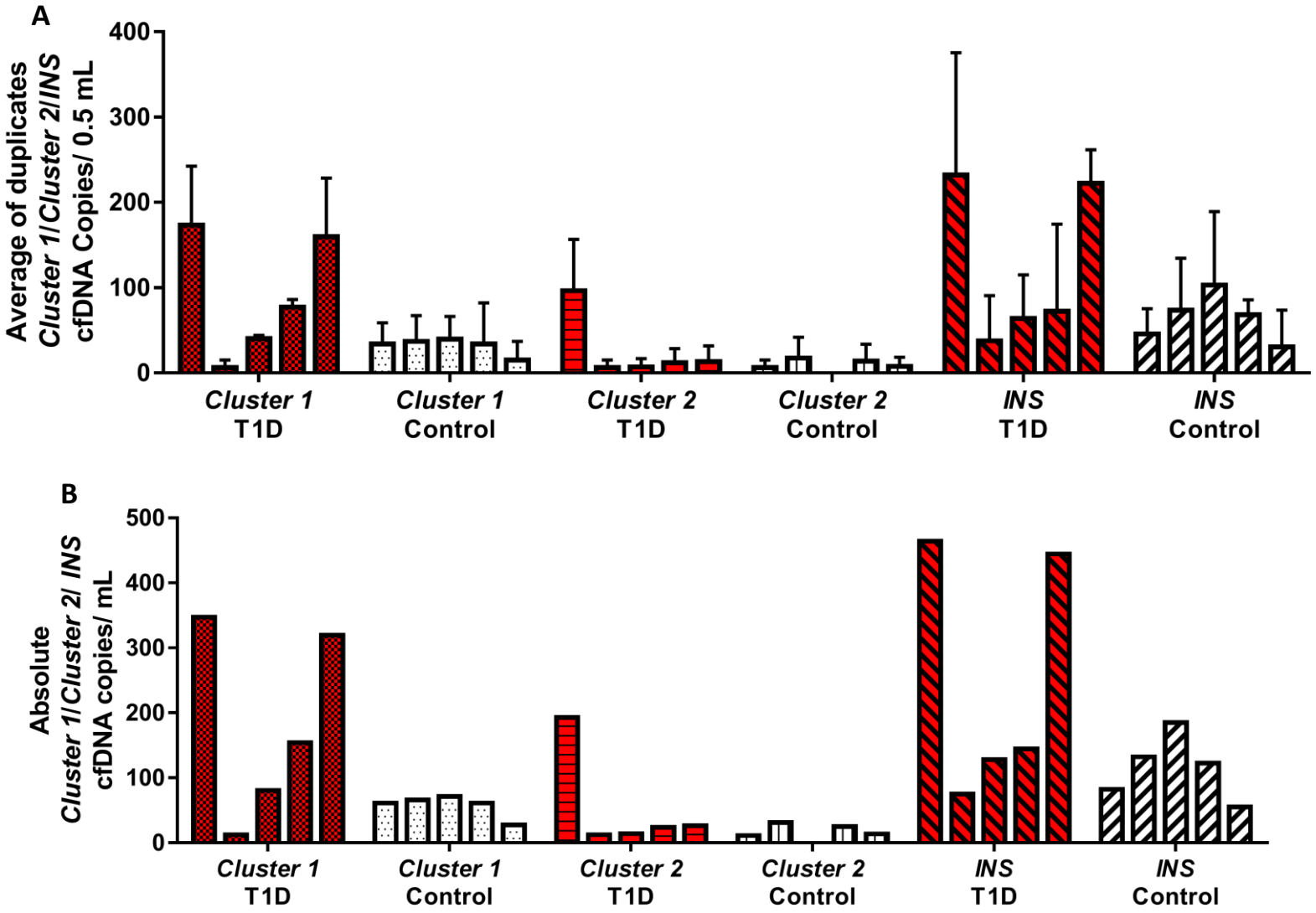
6.1.2. *MIR-200c* Cluster 2 Assay Triplex combination as a Biomarker for Beta Cell Death

Although the hypomethylated *MIR-141* (CpG -20, from TSS) single plex discriminated recently diagnosed T1D children from their corresponding control subjects ($p < 0.01$), the level of discrimination was not as efficient as that observed with *MIR-200c cluster 1* or *INS* singleplex. Therefore, we sought to replace the *MIR-141* assay with the *MIR-200c cluster 2* single plex (VIC, developed by Dr Jody Ye). Then, we examined the viability of this combination triplex assay (*Cluster1/INS/Cluster2*) to detect beta cell death in the selected recently diagnosed paediatric group (n=4, subject to availability as most of the samples had been used up in previous experiments) in comparison with an age and gender-matched non-diabetic control group (n=6). The sample of new onset T1D children were collected 6-12 months post-diagnosis. As expected, the assay showed no significant discrimination between the test and the control groups as the tested subjects were diagnosed post six months of T1D onset (Figure 6.2). Nevertheless, *Cluster 2* single plex showed better discrimination performance (0.072) (Table 6.2) in comparison with the *MIR-141* (0.046) (Table 6.1). Consequently, we decided to use *cluster 2* instead of *MIR-141* in the upcoming experiments.

Table 6.2: The effect of multiplexing on enhancing the discriminating power of beta cell death signal.

The multiplexing of several beta cell specific targets enhanced the discrimination power by 2 times over cluster 1 and INS single plex. The effect of multiplexing was the highest in cluster 2 as it elevated the discriminating power by almost 14 times.

Sample	Duration (months)	*Cluster 1*	*Cluster 2*	*INS*	Multiplex
BCD 9	6.36	154.17	23.61	144.44	322.22
BCD 2	9.32	12.50	12.50	75.00	100.00
BCD 1	10.51	12.50	12.50	75.00	100.00
BCD 3	12.39	181.94	26.39	194.44	402.78
BCD 8	12.95	430.56	54.17	380.56	865.28
	Sum	791.67	129.17	869.44	1790.28
	Power of Discrimination	0.44	0.072	0.485	


A
Average of duplicates
Cluster 1/Cluster 2/INS
cfDNA Copies/ 0.5 mL
400
300
200
100
0
Cluster 1
T1D
Cluster 1
Control
Cluster 2
T1D
Cluster 2
Control
INS
T1D
INS
Control
B
Absolute
Cluster 1/Cluster 2/ INS
cfDNA copies/ mL
500
400
300
200
100
0
Cluster 1
T1D
Cluster 1
Control
Cluster 2
T1D
Cluster 2
Control
INS
T1D
INS
Control

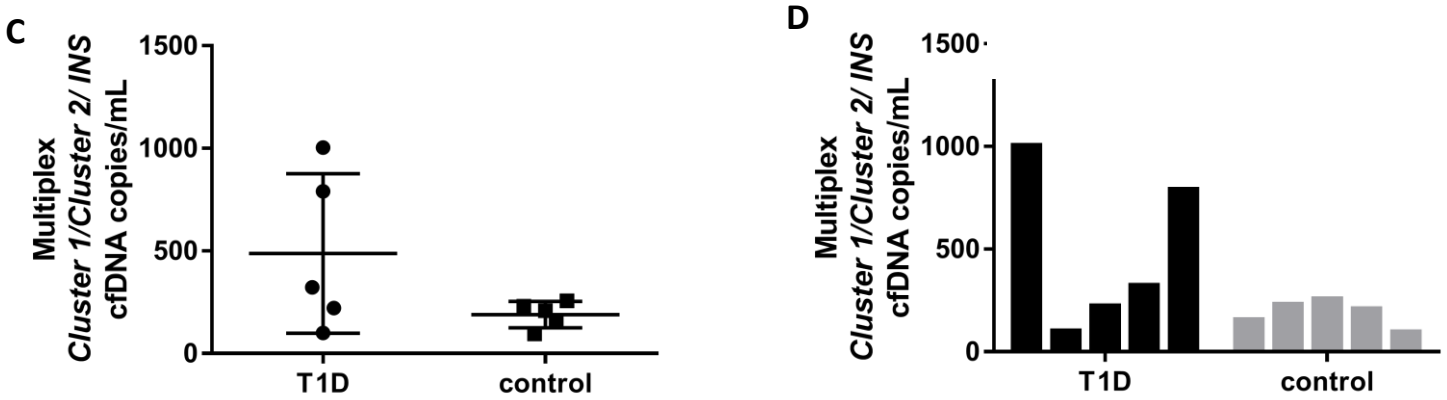


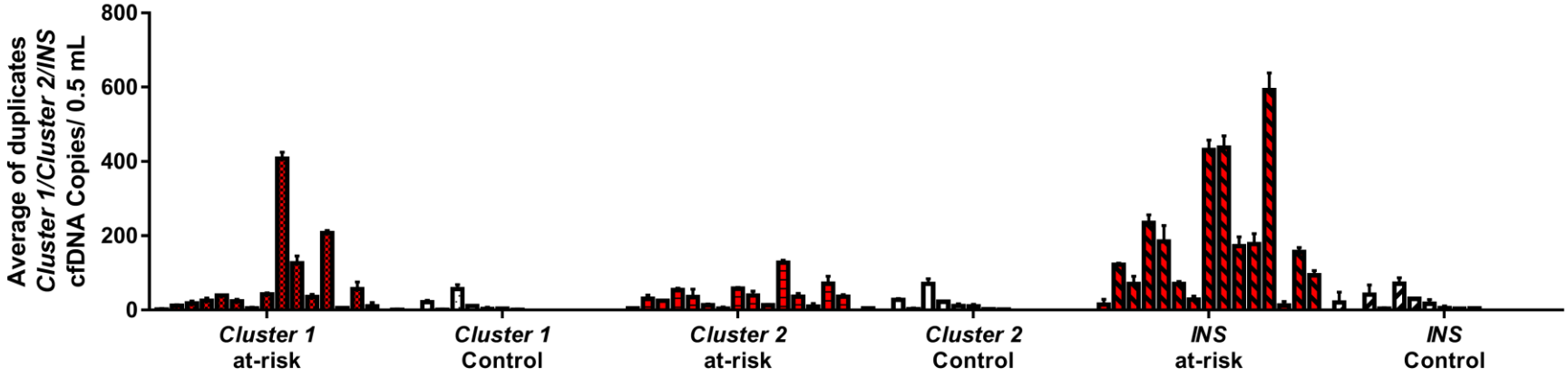
Figure 6.2: Circulating hypomethylated MIR-200c cluster 1/INS/MIR-200c cluster 2 cfDNA levels in recently diagnosed T1D pediatric human cohort.

The serum samples were collected 6-12 months post T1D diagnosis (n=6) and age/gender matched of nondiabetic control subjects (n=6). Extracted cfDNA were subjected to bisulfite treatment and analysed by ddPCR in duplicates. The columns represent the absolute copy number of hypomethylated cfDNA targets in 1 mL sample. A) The columns are presented as mean of the duplicate values (column) ± SD (error bar). B) Comparison of the absolute level of each hypomethylated target in recently diagnosed T1D subjects and non-diabetic controls. C) A comparison of the multiplex level of circulating hypomethylated beta cell targets in T1D and non-diabetic 1 mL serum samples. D) Comparison between the absolute level of circulating hypomethylated cfDNA of the multiplex assay in the recently diagnosed T1D and healthy control subjects. The data show that there is no significant difference in the level of circulating hypomethylated MIR-200c cluster 1/INS/MIR-200c cluster 2 cfDNA between recently diagnosed T1D (>6 months of diagnosis) and their corresponding control subjects ($p > 0.05$ by Mann-Whitney U-test).

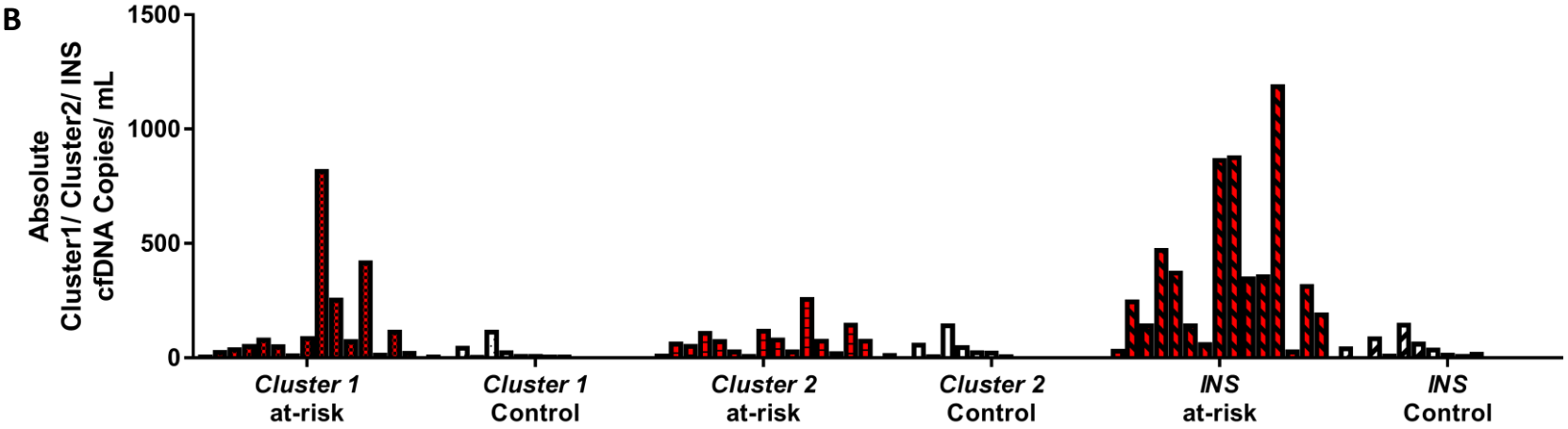
6.1.3. Monitoring Beta Cell Death in the "At-Risk" Autoantibody Positive Cohort

In addition to measuring beta cell death in recently diagnosed T1D patients, we aimed to test the aptitude of the assay in detecting beta cell death in pre-clinical individuals who were at risk for developing the disease. To achieve this goal, we studied the level of beta cell-specific hypomethylated single plex and multiplex cfDNA in plasma samples blindly recruited from subjects who are at-risk for progression to T1D (n=15). The participants were relatives of patients with T1D, participating in the Type 1 Diabetes TrialNet Pathway to Prevention trial, and were positive for at least one islet autoantibody. The results show that the at-risk subjects had a higher level of hypomethylated *cluster1/INS/cluster 2* in comparison to control subjects ($p < 0.05$ by Mann-Whitney U-test, Figure 6.3).

A


Average of duplicates
Cluster 1/Cluster 2/INS
cfDNA Copies/ 0.5 mL
800
600
400
200
0
Cluster 1
at-risk
Cluster 1
Control
Cluster 2
at-risk
Cluster 2
Control
INS
at-risk
INS
Control

B


Absolute
Cluster1/ Cluster2/ INS
cfDNA Copies/ mL
1500
1000
500
0
Cluster 1
at-risk
Cluster 1
Control
Cluster 2
at-risk
Cluster 2
Control
INS
at-risk
INS
Control

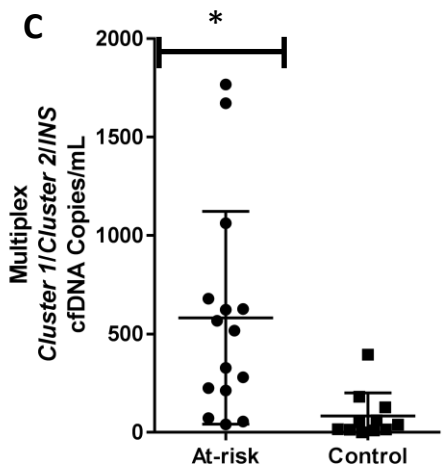
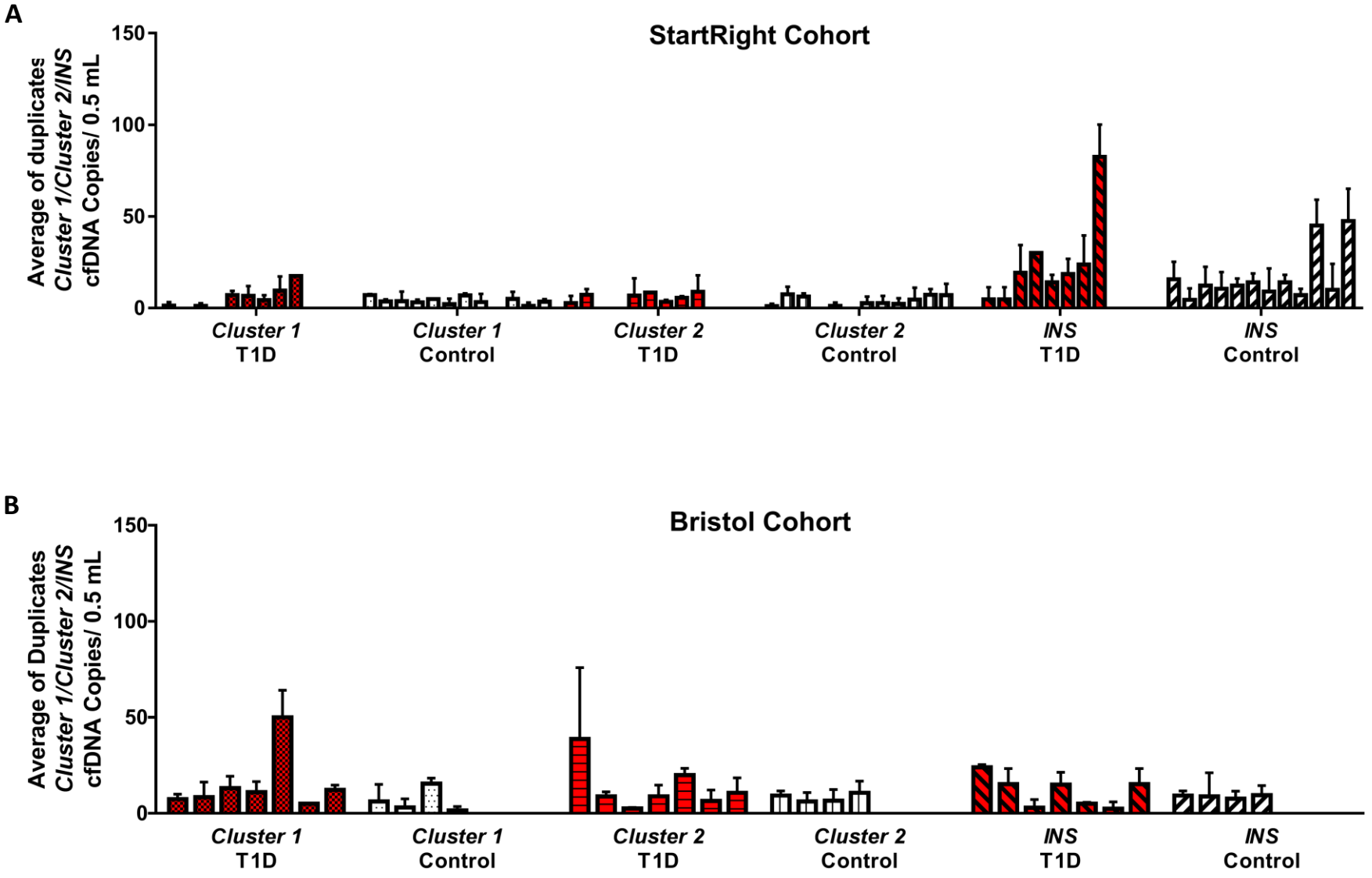


Figure 6.3: Singleplex and multiplex analysis of the beta cell death biomarkers MIR-200c cluster 1/INS/MIR-200c cluster 2 in individuals at-risk of developing T1D.

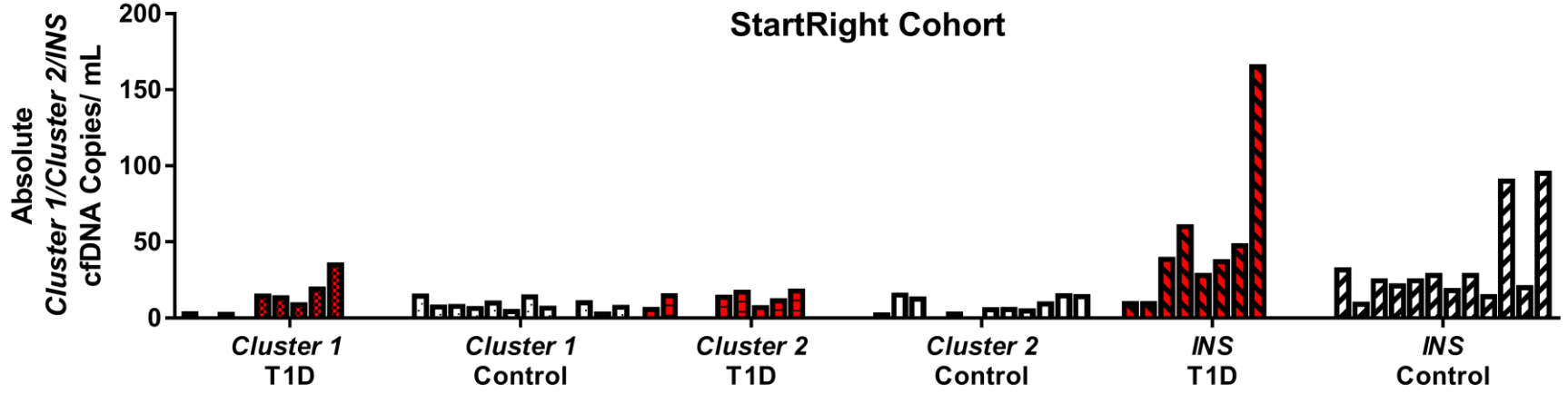
The columns represent the level of hypomethylated MIR-200c cluster 1/INS/MIR-200c cluster 2 DNA in 1 mL plasma samples collected from at-risk group and age and gender matched non-diabetic subjects. A) The columns are presented as mean of the duplicate value (column) ± SD (error bar). B) Comparing the absolute singleplex level of beta cell death biomarkers in each at-risk samples and non-diabetic control subjects. C) The level of multiplex beta cell death biomarkers in at-risk and non-diabetic control subjects. The mean and SD interquartile range is shown. The multiplex assay significantly discriminated the at-risk individuals from the non-diabetic control subjects ($p < 0.05$ by Mann-Whitney U-test).

6.1.4. Monitoring Beta Cell Death in Recently Diagnosed T1D Subjects.

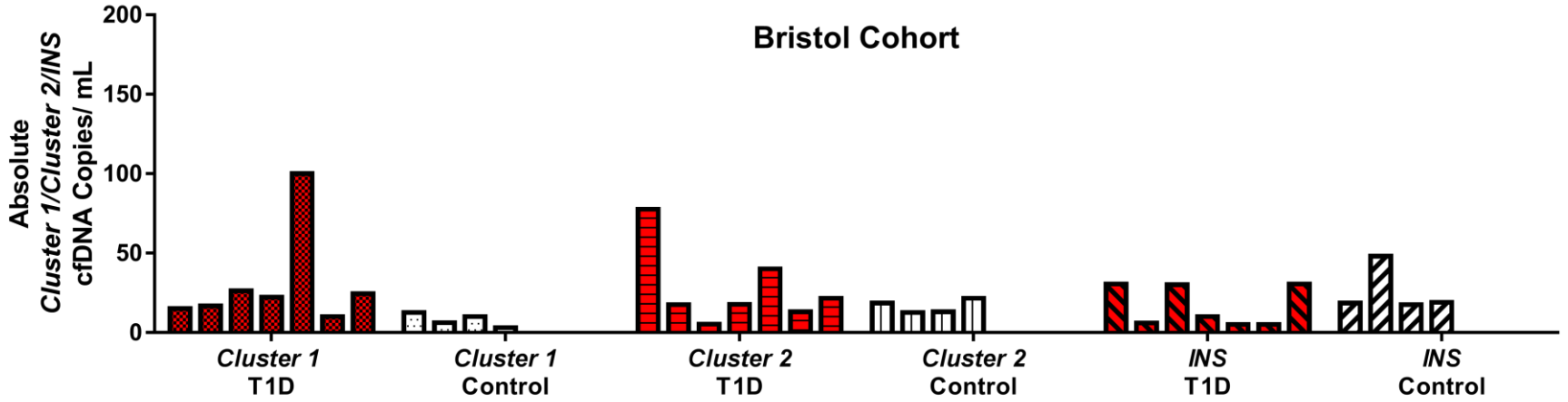
In addition to detect ongoing beta cell death in recently diagnosed and at-risk cohorts, we aimed to examine the level of beta cell cfDNA in recently diagnosed adult T1D patients. To achieve this aim, we studied the level of hypomethylated single plex and multiplex DNA in plasma samples available from the StartRight study (*n*=9) compared to locally available age and gender-matched non-diabetic controls (*n*=12), and in addition to plasma samples (T1D *n*=7, Control *n*=4) collected from the Bristol cohort. The clinical characteristics available for these samples are shown in General Methods, Chapter 2, section 2.3.3. When the rate of beta cell death was compared to age and gender-matched non-diabetic control subjects, both singleplex and multiplex analysis of *MIR-200c cluster 1/INS/MIR-200c cluster 2* showed no significant differences between recently diagnosed adult T1D patients and the control group ($p > 0.05$) (Figure 6.4).


A
StartRight Cohort
Average of duplicates
Cluster 1/Cluster 2/INS
cfDNA Copies/ 0.5 mL
150
100
50
0
Cluster 1
T1D
Cluster 1
Control
Cluster 2
T1D
Cluster 2
Control
INS
T1D
INS
Control
B
Bristol Cohort
Average of Duplicates
Cluster 1/Cluster 2/INS
cfDNA Copies/ 0.5 mL
150
100
50
0
Cluster 1
T1D
Cluster 1
Control
Cluster 2
T1D
Cluster 2
Control
INS
T1D
INS
Control

C


StartRight Cohort
Absolute
Cluster 1/Cluster 2/INS
cfDNA Copies/ mL
200
150
100
50
0
Cluster 1
T1D
Cluster 1
Control
Cluster 2
T1D
Cluster 2
Control
INS
T1D
INS
Control

D


Bristol Cohort
Absolute
Cluster 1/Cluster 2/INS
cfDNA Copies/ mL
200
150
100
50
0
Cluster 1
T1D
Cluster 1
Control
Cluster 2
T1D
Cluster 2
Control
INS
T1D
INS
Control

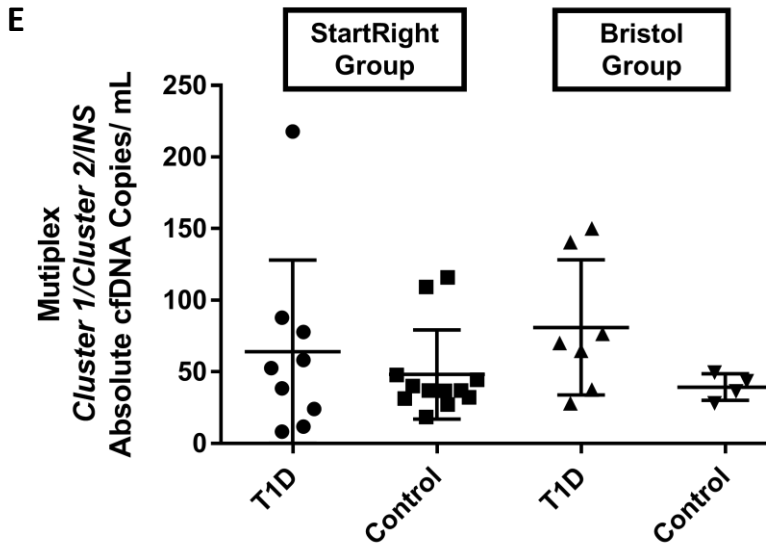


Figure 6.4: Circulating hypomethylated MIR-200c cluster 1/INS/MIR-200c cluster 2 cfDNA levels in a recently diagnosed T1D adult cohort.

The plasma samples were collected from T1D patients who had the condition for 3 months (n=9) and gender/age-matched nondiabetic control subjects (n=12). The columns represent the absolute copy number of hypomethylated cfDNA targets in a 1 mL sample. A) and B) The average level of duplicates of each single circulating hypomethylated beta cell target (MIR-200c cluster 1/INS/MIR-200c cluster 2 cfDNA in a 0.5 mL plasma samples. C) and D) A comparison of the absolute level of each circulating hypomethylated beta cell target in T1D and non-diabetic serum samples. E) Comparison between the absolute levels of multiplex circulating hypomethylated cfDNA in the multiplex assay of the T1D adult patients and healthy control subjects. The mean and SD interquartile range is shown. The data show that the was no significant difference in the level of circulating hypomethylated MIR-200c cluster 1/INS/ MIR-200c cluster 2 cfDNA ($p > 0.05$ by Mann-Whitney U-test) between recently diagnosed T1D adult patients and their corresponding control subjects in both single plex and multiplex assay.

6.1.5. Measuring the Beta Cell Death Rate in the Post Islet Allotransplant Patients

We extended our tested groups to include plasma samples (*n*=2) collected from clinical islet transplant patients. One of the samples was recruited from a study to assess the efficacy and safety of Reparixin in pancreatic islet transplantation (REP0211), and the other was from the bone marrow vs liver as the site for islet transplantation study (IsletBOM2). As predicted, the level of circulating beta cell cfDNA was elevated immediately after islet infusion in both recipients (1-hour post-infusion), the peak level declined sharply afterwards, reaching the baseline level after 12 hours post-infusion in the IsletBOM2 samples and after 48 hours in the REP0211 sample (Figure 6.5 A and B). The baseline level of the hypomethylated circulating cfDNA persists up to one-week post-infusion. The results demonstrate that the developed multiplex assay can be used to monitor beta cell death in human clinical transplant samples.

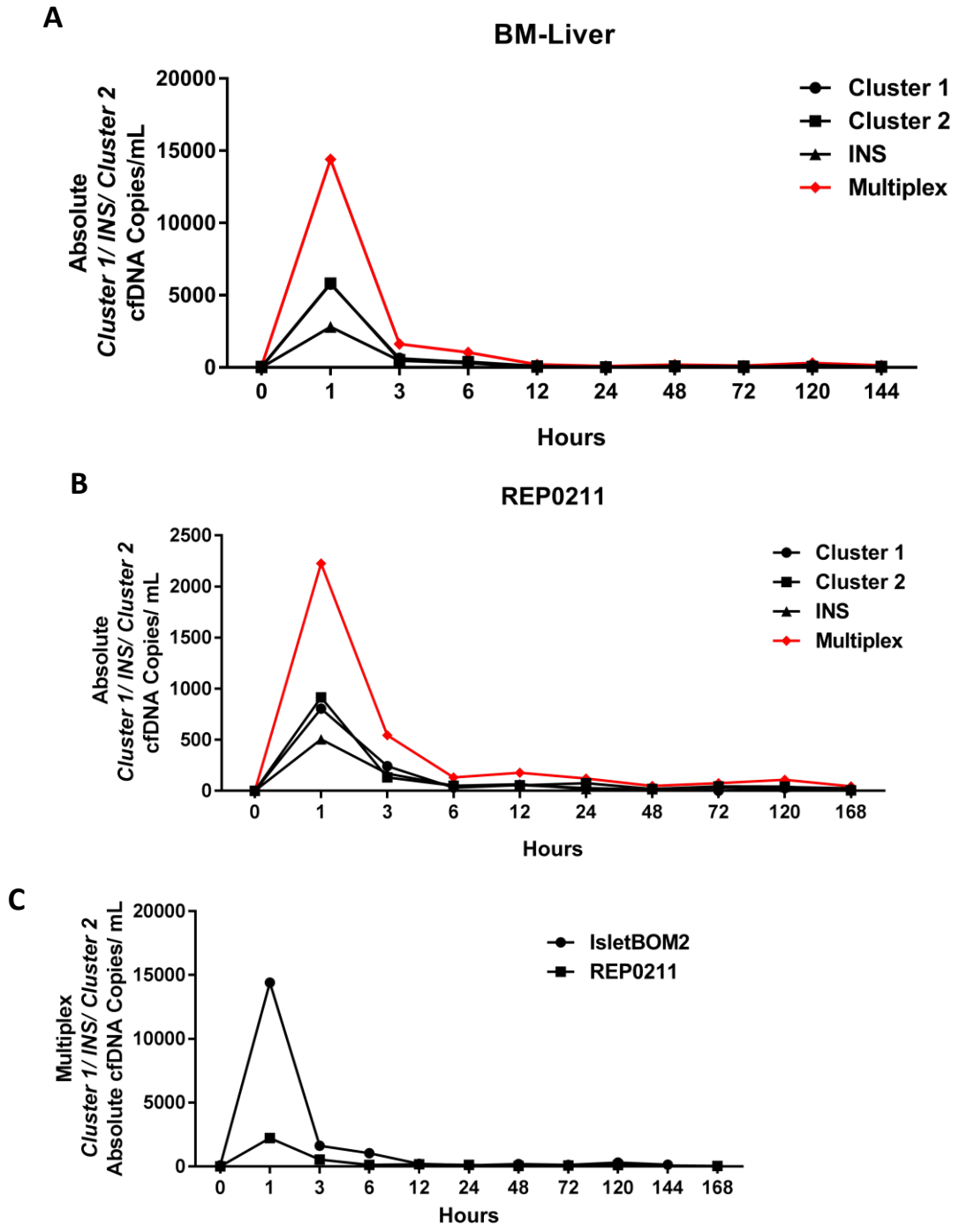


Figure 6.5: Change in hypomethylated MIR-200c cluster 1/INS/MIR-200c cluster 2 cfDNA pre and post-islet-infusion.

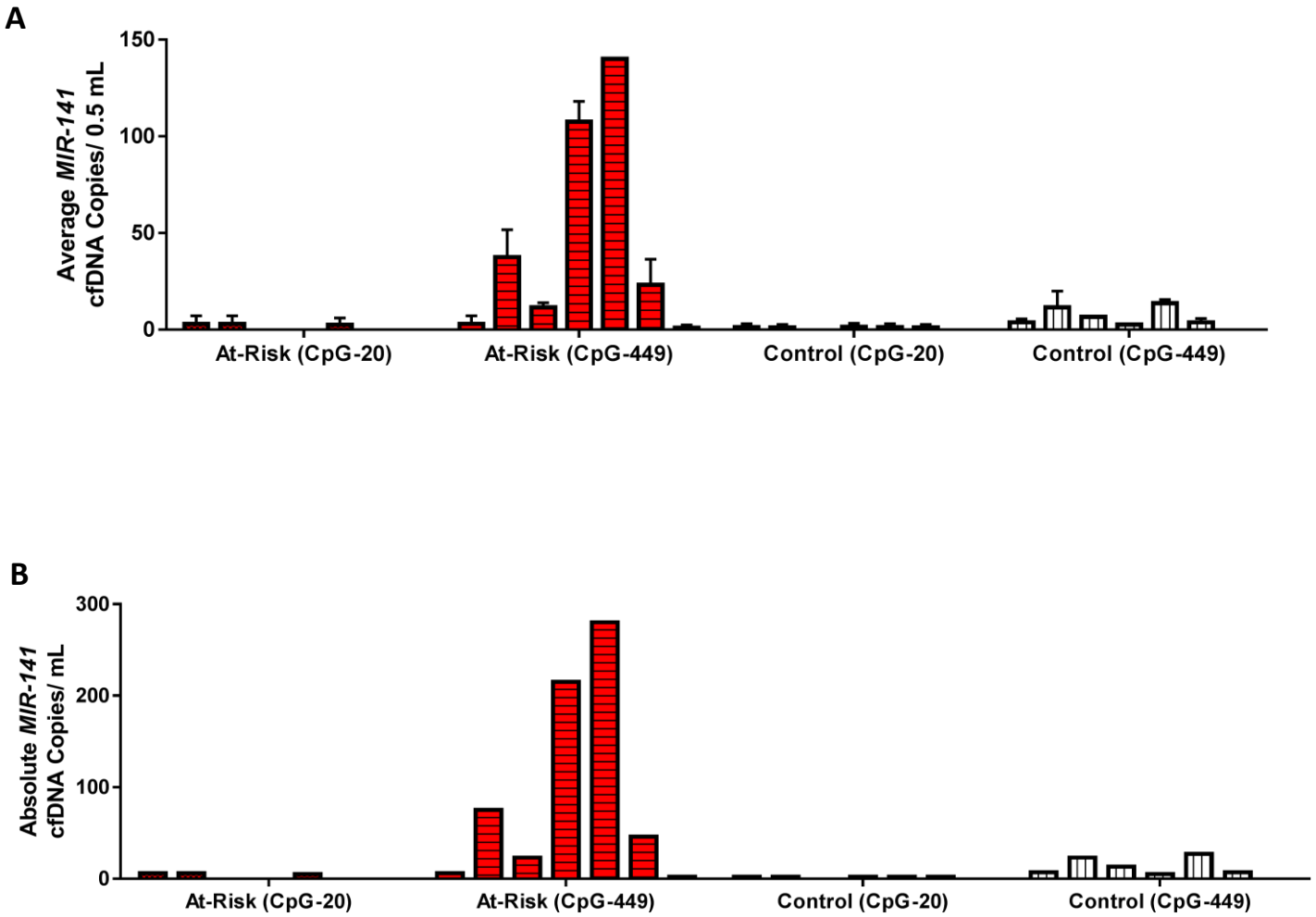
Genomic DNA from plasma samples was bisulfite-converted and used for multiplex ddPCR assays in duplicate. The single plex and multiplex beta cell death biomarker level changes post islet-infusion in patient 1 (A) and patient 2 (B). C) The absolute cfDNA copy number of the multiplex assays in 1 mL plasma samples. The data show that the level of beta cell cfDNA rises and peaks 1-hour post-infusion, the signal declines through the following hours and reaches the baseline level after 12 hours in patient 1 and after 48 hours in patient 2. Although low, there was a fluctuation in the level of beta cell cfDNA in patient 2.

6.1.6. Developing a New Assay for Monitoring Beta Cell Death

Although the *MIR-200c cluster 2* had a better discrimination power in comparison to the CpG -20 *MIR-141* assays, it still lower in comparison to the *cluster 1* and *INS* assay. Therefore, we sought to replace the *cluster 2* assay with another assay, which would enhance the total discriminating power of the multiplex assay. The data for in-depth analysis of the *MIR-141* promoter methylome had suggested CpG -449, from TSS, as a potential beta cell-specific DMR biomarker. In order to determine whether the level of hypomethylated *MIR-141* (CpG-449) singleplex would increase in at-risk individuals we examined the level of a hypomethylated duplex of both *MIR-141* CpG -20 and CpG -449 in second elution DNA samples (*n*=7) isolated from the previous experiment studying the performance of a beta cell-specific multiplex assay in age/gender match non-diabetic controls. The genomic DNA concentration was relatively low in comparison to the first elusion sample (Table 6.3). The results were consistent with the observations seen in the multiplex assay. The single plex and duplex assays were not significantly different between the control and test group ($p > 0.05$ by Mann-Whitney U-test) (Figure 6.6); samples that had sufficient DNA yield recovery (Samples number 3 and 4) resulted in DNA copy numbers of 200-300. Nevertheless, more experiments will be required to determine the usefulness of this assay.

Table 6.3: The genomic DNA yield extracted from the at-risk samples and non-diabetic control subjects.

		gDNA 1st elution (ng/µL) in 40 µL elusion	gDNA 2nd elution (ng/µL) in 25 µL elusion
1 Ab +ve	**TNID0001**	4.9	1.19
1 Ab +ve	**TNID0002**	6.4	2.57
1 Ab +ve	**TNID0003**	2.92	1.75
4 Ab +ve	**TNID0004**	6.1	7.1
3 Ab +ve	**TNID0005**	4.95	7
2 Ab +ve	**TNID0006**	6.1	2.8
2 Ab +ve	**TNID0007**	3.3	1.28
	SCF2105	1.77	1.32
	UHWBC006	1.23	0.927
	UHWBC010	1.57	1.08
	UHWBC026	2.32	1.46


A
Average MIR-141 cfDNA Copies/ 0.5 mL
150
100
50
0
At-Risk (CpG-20)
At-Risk (CpG-449)
Control (CpG-20)
Control (CpG-449)
B
Absolute MIR-141 cfDNA Copies/ mL
300
200
100
0
At-Risk (CpG-20)
At-Risk (CpG-449)
Control (CpG-20)
Control (CpG-449)

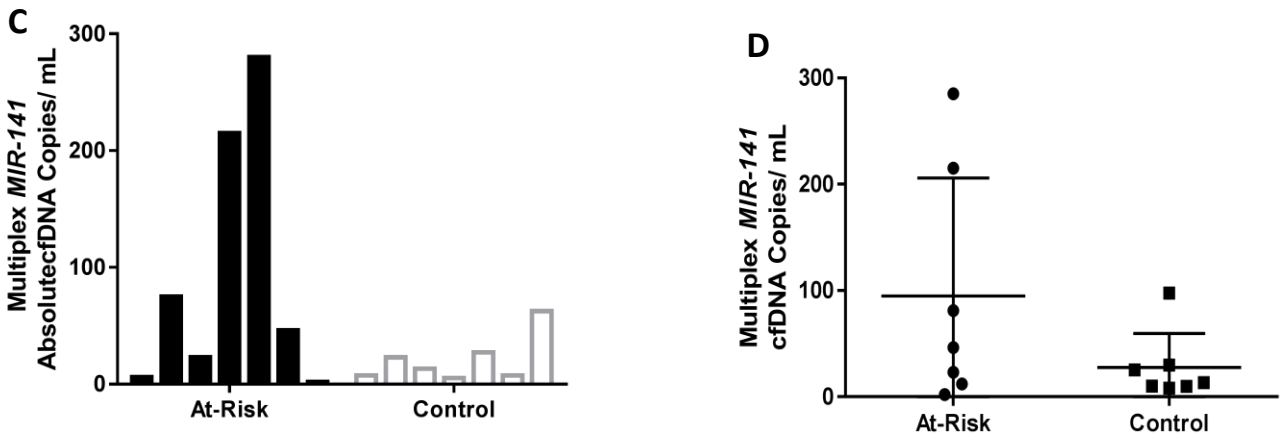


Figure 6.6: Single plex and duplex analysis of the beta cell death hypomethylated MIR- 141 biomarker in individuals at-risk for T1D.

A) The columns are presented as the mean of the duplicate values (column) ± SD (error bar). B) The columns represent the absolute level of hypomethylated MIR-141 (CpG-20 and CpG-449, from TSS) DNA in 1 mL plasma samples collected from the at-risk group and age/gender matched nondiabetic subjects. C) The rate of beta cell death duplex biomarker in at-risk individuals and non-diabetic control subjects. D) A comparison of the MIR-141 (CpG-20 and CpG -449, from TSS) biomarker level in at-risk against non-diabetic control subjects. The mean and SD interquartile range is shown. The data suggest that the duplex biomarker assay does not significantly discriminate the at-risk individuals from the non-diabetic control subjects ($p > 0.05$ by Mann-Whitney U-test).

Discussion

Despite the limited number of samples available, the work presented in this chapter suggests that young children at diagnosis onset, those who are at-risk of developing T1D and islet transplantation recipients are the most promising cohorts to be targeted for further analysis. The results from islet transplant recipients, although limited, has shown the validity of our initial hypothesis, that multiplex DMR based biomarkers show promise for detecting real-time beta cell death. Previously, DMR biomarkers based on differentially methylated CpG sites within the *INS* gene were successfully developed (Akirav et al. 2011, Bellin et al. 2017, Herold et al. 2015a). Although those authors showed this biomarker was significantly discriminating beta cell fragments in the periphery, the authors reported the presence of detectable hypomethylated *INS* signals in the control samples that could have originated (even partially) from non-beta cell origin. Detecting hypomethylated signals in control samples highlights the issue of limited sensitivity. To address this concern, Lehmann-Werman and colleagues studied the effect of multiplexing more than one *INS* CpG target on the sensitivity and specificity of DMR based biomarkers. They showed that multiplexing six specific CpG sites in the promoter region of the *INS* gene achieved maximum specificity and decreased the detected hypomethylated *INS* signal from 15-20% to 0.01% in non-beta cell tissue genomes. Consequently, multiplexing allowed discrimination of individuals with recent-onset T1D from healthy controls with 100% sensitivity and specificity (Lehmann-Werman et al. 2016). Collectively, the data highlighted the power of multiplexing several adjacent CpGs. In parallel, we showed that multiplexing several CpG sites within the promoter region of *MIR-200c*/*MIR-141* enhanced the sensitivity of our assay. We further observed that the discriminating power of the *MIR-141* assay was low in comparison with that of the other singlpex assays developed, so we decided to carry on the project using only the *MIR-200c cluster 1* plus the *INS exon 2* and *MIR-200c cluster 2* multiplex combination. Although *MIR-141* detected only low levels of beta cell death in cases, it was significantly different from the control samples ($p < 0.05$ Mann Whitney U Test); we may, therefore, revisit this assay in the future when more samples are available for analysis.

The cfDNA assays developed appear to successfully discriminate beta cell-derived cfDNA in serum from a small cohort of recently diagnosed children. The level of hypomethylated cfDNA was observed to be lower in samples taken six months after diagnosis, and this low level seems to persist. There are however some limitations to this analysis. The number of serum samples available taken within one hour of venepuncture was small and longitudinal follow up of individuals was not possible as these samples were obtained through colleagues at University College London. The experiments require relatively large volumes of serum and therefore repeated analysis of individual samples was also not possible. Moving forward, samples collected from the Bart's Oxford study of type 1 diabetes (http://www.bristol.ac.uk/translational-health-sciences/research/diabetes/research/box/) will allow the number of samples tested to be increased and provide longitudinal follow up. It is worth noting that the observation that cfDNA levels are increased close to diagnosis was also described by three independent reports that observed a significant upregulation of the level of circulating hypomethylated *INS* level in recently diagnosed T1D patients (Akirav et al. 2011, Fisher et al. 2015, Lehmann-Werman et al. 2016). Further, our finding was consistent with the observation reported by Fisher and colleagues who noticed that beta cell-specific cfDNA returns to basal levels by one-year post-diagnosis (Fisher et al. 2015), even though their samples were also cross-sectional.

When the multiplex assay (*MIR-200c cluster 1*+ *INS exon 2* and *MIR-200c cluster 2*) was tested in the blinded at-risk cohort, we observed a significant upregulation in hypomethylated cfDNA levels in those who were at risk of developing T1D (autoantibody positive) compared to the non-diabetic control subjects. This observation confirmed a previous report showing that the beta cell death biomarker (hypomethylated *INS* fragments) was high in 70 % of those at-risk of T1D (7 of 10 individuals who progress to clinical T1D) (Herold et al. 2015a). A workshop set up to examine cfDNA levels in blinded samples from "at risk" children and age matched-controls based on existing assays targeting the *INS* methylation variants alone, however, failed to replicate the increased levels in "at risk" children (Professor Raghu Mirmira, Personal communication). It is worth highlighting that we noticed a correlation between the age of seroconversion and the number of circulating autoantibodies in the

first experimental sample batch. The data conducted from that experiment suggest that seroconversion at a younger age as well as having multiple autoantibodies in the circulation are two factors that may contribute to more aggressive destruction of beta cells. Unfortunately, we did not observe such correlation in the second sample batch tested. Nevertheless, this observation will be further explored when a larger number of samples are analysed. Next, the level of beta cell destruction in recently diagnosed adult T1D patients was analysed. The analysis of plasma samples showed that there was no significant difference between the levels of *MIR-200c cluster 1/INS/MIR-200c cluster 2* in adult T1D patients in comparison to non-diabetic control subjects. A larger sample size is required to confirm this conclusion.

In addition to the early detection of ongoing beta cell destruction, circulating beta cell-specific DNA could be of clinical benefit for other purposes such as evaluating the outcome of islet transplantation and responses to clinical therapy. Beta cell-specific cfDNA signals were detected in the circulation of islet transplantation recipients, provided kindly by Dr Vito Lampasona (IRCCS San Raffaele Scientific Institute, Italy), as early as 1-hour post islet infusion when a pulse of beta cell death was observed. The striking elevation in hypomethylated cfDNA that was observed in the first-hour has also been reported by other groups (Bellin et al. 2017, Lehmann-Werman et al. 2016) and attributed to the rapid beta cell death immediately post-transplantation caused by the instant blood-mediated inflammatory reaction (IBMIR) and/or beta cell hypoxia (Kanak et al. 2014, Nilsson et al. 2011, Olsson et al. 2011). The level of hypomethylated cfDNA dropped sharply within 3 hours and reached baseline level 12 hours after infusion. It is postulated that close monitoring of the beta cell death signal in the days after islet transplantation will inform future outcomes. What the experiments carried out in this study highlighted is the added sensitivity of the multiplex assay developed as part of this project compared with using insulin methylome targets alone, as other studies have done.

Chapter 7: Biomarkers of Prostate Cancer Progression

7. Introduction

During the in-depth study of the mutual promoter regions of the genes *MIR-200c*/*MIR-141* as a potential differentially methylated region in islets and a beta cell line (EndoC-βH1), we found out that this region had also been reported previously to be predominantly hypomethylated in prostate and breast epithelial cells (Lynch et al. 2016; Neves et al. 2010; Vrba et al. 2010). This encouraged us to investigate whether our developed assay could differentiate the normal genome from the genome of benign and malignant cells and whether our developed assay could be used in future as a staging biomarker for prostate and breast cancer. We initially studied the methylation level of CpG -20, from TSS, of the *MIR-141* gene in well-established normal and malignant breast and prostate cell lines (described in more details in General Methods, Chapter 2, section 2.1.2.2). We further compared the level of miR-141-3p/5p expression to the methylation level of CpG-20 (*MIR-141*) in prostate epithelial cell lines.

Methods

7.1.1. MicroRNA Extraction

Total RNA extraction (including microRNA) was performed using the miRNeasy Serum/Plasma Kit (Qiagen) according to the manufacturer's protocol. RNA was eluted in 20 µL of nuclease-free H_2O (Qiagen) and stored at -80^{o}C.

7.1.2. Quantification of microRNA

The concentration of all RNA samples was measured using a Quantus™ Fluorometer (Promega) in combination with the QuantiFluor® RNA System (Promega) according to the manufacturer's instruction. The High (7.8–500 ng/µL) or low (0.16–10 ng/µL) standard curve was constructed via the RNA standard (included in the kit) from which the concentration (ng/ µL) of unknown RNA was calculated. Oligonucleotides of miRNA Amplification

Locked nucleic acid (LNA™) primers that amplify miR-141-3p, miR-141-5p and miR-200c-3p were purchased from EXIQON (Qiagen, UK). LNA bases are a class of high-affinity RNA analogues in which the ribose ring is "locked" the 3' end conformation by a bridge of methylene between the 2' oxygen and the 4' carbon in the idyllic structure for Watson-Crick binding. The "Locked" structure improves the organisation of the base backbone and stacking, elevating the melting temperature, the stability and the binding affinity (Braasch and Corey 2001). The primer sequences are described in Table A.2, Appendix 1).

7.1.3. Reverse Transcription PCR (RT-PCR) and ddPCR Analysis of Prostate Cancer Cell Lines.

Ten ng of total RNA was reverse transcribed using a Universal cDNA Synthesis Kit II (Exiqon) according to the manufactural protocol. The ddPCR mixture reaction (20 µL) containing 10 µL of ddPCR 1× EvaGreen ddPCR Supermix (Bio-Rad), a final concentration of 1X LNA^{TM} primers (Qiagen) and 9 µL of 1:1000 diluted cDNA template. No template (NTC) were used as a negative control. The 20 µL Reactions were premixed in LoBind Eppendorf tubes and dispensed along with 70 µL of droplet generation oil for EvaGreen (Bio-Rad) into the appropriate well of the droplet generator DG8 cartridge

(Bio-Rad). The PCR cycles were as following; 95°C for 5 minutes, 40 cycles of 95°C for 30 seconds and 58°C for 1 minute (Ramp-Rate 1.6°C/s), 4°C for 5 minutes, 90°C for another 5 minutes and then a final incubation at 4°C (Figure 7.1). Data were analysed using the QuantaSoft software (Bio-Rad, version 1.6.6.0320).

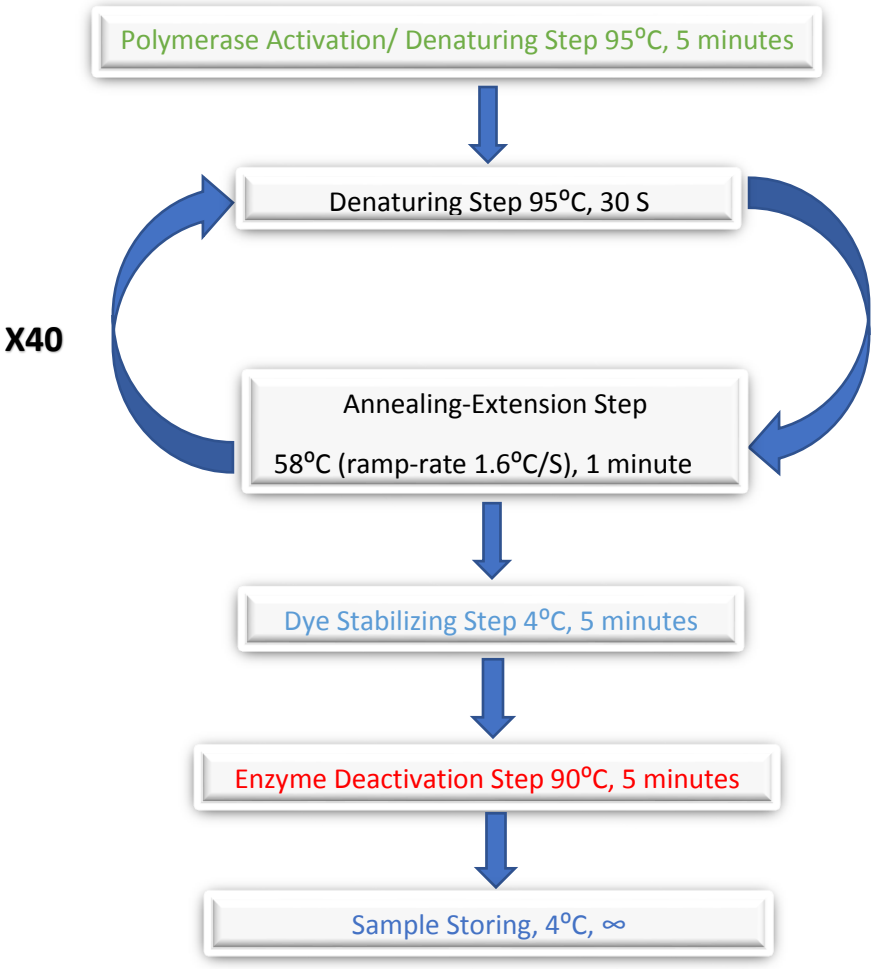


Figure 7.1: The EvaGreen ddPCR Protocol.

Results

7.2.1. Methylation Analysis of *MIR-141* Promoter in Prostate and Breast Epithelial Cells

We developed a methylation-specific assay to simultaneously quantitate hypermethylation or hypomethylation of *MIR-141* DNA at position CpG -20 bp from TSS. We showed previously in chapter 4 that CpG -20 bp is preferentially hypomethylated in islet and beta cell line (EndoC-βH1). Other studies have reported that the *MIR-141* promoter is also hypomethylated in breast and prostate epithelial cells (Lynch et al. 2016; Vrba et al. 2010). Therefore, we tested our differentially methylated developed *MIR-141* in a panel of prostate epithelial cell lines that either expressed hormonal androgen receptors (AR) such as PNT2 and LNcaP or had lost expression of the hormonal AR such as PC3 and DU145. The normal prostate cell line (PNT2) showed a predominant hypomethylation at CpG -20. The level of hypomethylation was shown to be increased in the benign (grade I) cell line (LNcaP). As the cells became malignant and lost expression of the AR, as in the DU145 cell line (grade II), we noticed that CpG-20 was shifting from hypomethylation into hypermethylation. Furthermore, analysing grade IV metastasis adenocarcinoma (PC3), showed that CpG -20 lost the hypomethylation completely. Collectively, the data suggest that CpG -20 in *MIR-141* is hypomethylated in hormone-sensitive prostate cell lines and the level of hypomethylation is shifted toward hypermethylation as the hormonal sensitivity is lost (Figure 7.2A).

Next, we aimed to test more cell lines that were or were not hormone responsive. Again, we tested our methylation sensitive assay on a prostate cancer grade III cell line (Vcap) and well-established breast cancer cell lines that either express the progesterone receptor (PR) such as MCF10A, MCF7, T47D, ZR751 or do not express the receptors such as HS578T and MDA-MB231. ddPCR analysis of CpG -20 showed a similar conclusion (Figure 7.2B). The CpG -20 was hypomethylated in all cell lines that express ERs; the methylation status is shifted into hypermethylation when the receptor expression is lost.

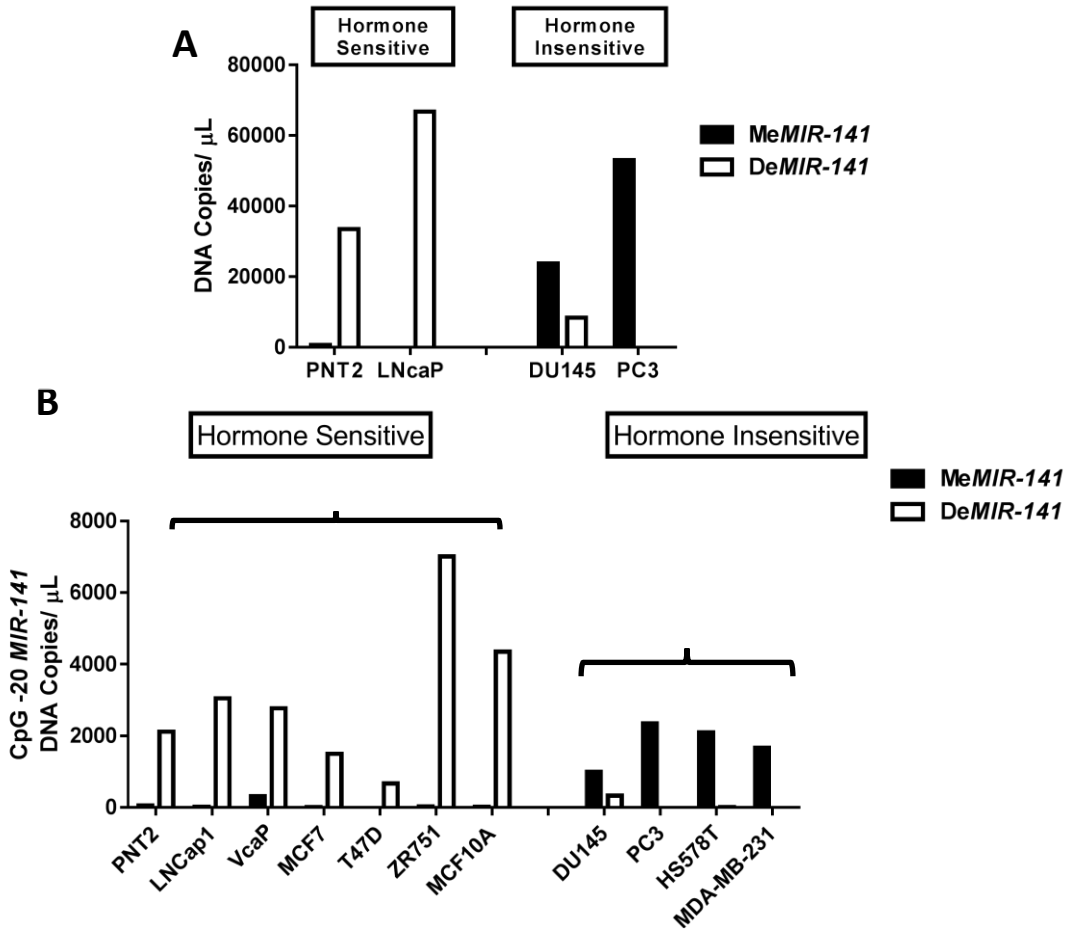


Figure 7.2: The methylation level of CpG -20 of MIR-141 in Hormone sensitive and Hormone insensitive prostate and breast cell lines.

The columns represent the level of methylated CpG -20 DNA copies in 1 µL (5 ng DNA sample). Each graph represents a single analysis. The DNA copies were normalised to 100 copies of beta-actin. A) ddPCR analysis of CpG-20 of MIR-141 methylation in prostate cancer cell lines LNCaP, DU145, and PC3 and normal prostate epithelial cell line PNT2. The data indicates that there are shifts from hypomethylation into hypermethylation at the CpG-20 in parallel with cancer progression and as hormone sensitivity is lost. B) ddPCR analysis of CpG-20 of MIR-141 methylation in prostate and breast hormone sensitive cell lines (LNCaP, VcaP, MCF7, T46T and ZR-751) and hormone insensitive cell lines (DU145, PC3, HS-578-T and MDA-MB-231). The data confirm the findings in the first experiment which showed that the methylation shift from hypomethylation to hypermethylation at CpG -20 of MIR-141 is linked to the hormone sensitivity of the cell.

7.2.2. The Expression of miR-141-3p/5p and miR-200c-3p in Prostate Cell Lines

In addition to measuring the level of methylation in prostate and breast cell lines, we sought to discover whether measuring the expression of the miR-141 and its neighbouring gene miR-200c would add greater confirmation value to the assay. To fulfil this goal, we extracted total miRNA from cell pellets provided kindly by Professor Jeffrey Holly. The miRNA was converted into complementary DNA, which was used as a template for ddPCR analysis. Interestingly, we observed an upregulation in the level of expression of both miR-141-3p and miR-200c-3p in the benign cell line LNcaP, which had been shown previously (Figure 7.2) to have a higher level of hypomethylation level in comparison to the normal prostate epithelial cell line PNT2. The data also show complete suppression of the miR-141-3p and miR-200c-3p expression in the PC3 cell line, which had shown a complete shift from hypomethylation to hypermethylation at CpG -20. We did not detect any change in the level of miR-141-5p (Figure 7.3).

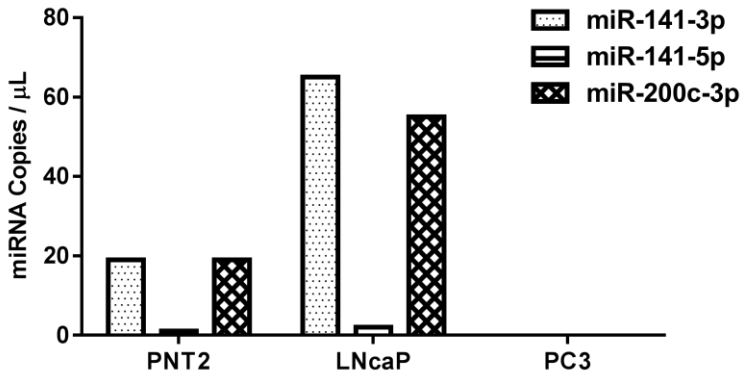


Figure 7.3: The miR-141-3p/5p and miR-200c-3p expression level in prostate cancer cell lines.

The columns represent DNA copy number in 1 µL DNA sample. The data illustrate that miR-141-3p but not miR-141-5p is correlated to the methylation status of CpG -20 of MIR-141. The data also show that the expression of miR-200c is correlated with the methylation of the CpG -20 as its expression is upregulated in PNT2 and LNcap cells (CpG -20 is hypomethylated) and it is supressed in PC3 where CpG -20 is completely hypermethylated.

7.2.3. The Methylome Level of Prostate Cell Lines at the Promoter of *MIR-141* Gene

There was a correlation between the methylation status of CpG -20 prostate cell lines and the expression of miR-141-3p and miR-200c. Therefore, we thought to investigate whether this correlation is specific to the CpG -20 site or it was related to the neighbouring CpG sites of the *MIR-200c/MIR-141* promoter. Therefore, a targeted next-generation bisulfite sequencing (tNGBS) was performed on PNT2, LNcaP and PC3 cell lines to sequence the *MIR-200c*/*MIR-141* mutual promoter area as well as the encoding region of *MIR-141* Chr12: 6,963,000-6,964,001 (described in Chapter 4; Section 4.1.2). The results showed that the *MIR-200c*/*MIR-141* promoter was significantly hypermethylated in PC3 while LNcaP and the normal epithelial (PNT2) cells were predominantly hypomethylated (Figure 7.4). Therefore, concluding that the methylation pattern of *MIR-200c*/*MIR-141* could be linked to the hormonal sensitivity of the malignant cell, which may allow us to use the *MIR-141* assay as a staging biomarker for prostate and breast cancer.

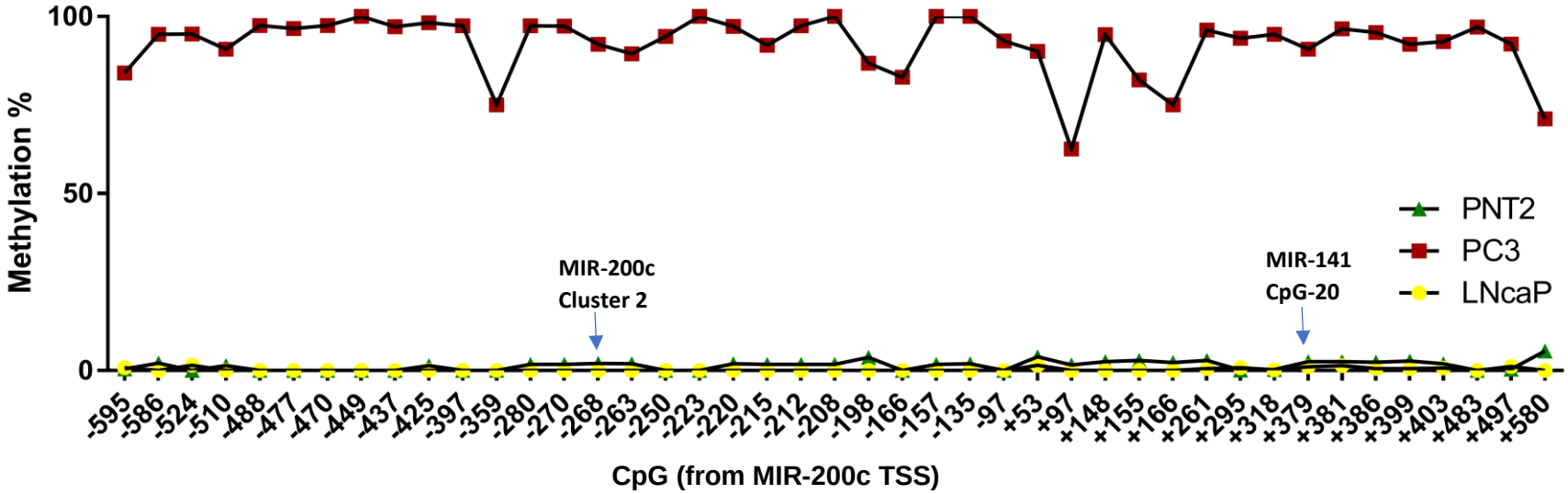


Figure 7.4: Targeted methylation analysis of the MIR-200c/MIR-141 promoter in prostate cancer cell lines.

The line chart presents the methylation status of the region Chr12: 6,963,000-6,964,001 where the CpG numbering is related to the transcription start site of MIR-200c. The data here show that the DNA methylation is completely shifted from hypomethylation in the normal PNT2 cells and the hormone sensitive LNcaP cells into predominantly hypermethylation in hormone insensitive PC3 cells in the promoter of MIR-200c/MIR-141.

Discussion

The methylation pattern of the *MIR-141* promoter has been reported previously to be predominantly hypomethylated in prostate and breast epithelial cells (Lynch et al. 2016; Neves et al. 2010; Vrba et al. 2010). Further, miR-200c and miR-141 had also been reported previously to be upregulated in breast and prostate cancer (Brase et al. 2011, Li et al. 2016, Lynch et al. 2016, Vrba et al. 2010). Vrba and colleagues had profiled the methylation status of *MIR-200c* in multiple breast and prostate cell lines. Their data suggested that each of the adenocarcinoma cell lines that express miR-200c and miR-141 (UACC-893, MB-468, UACC-1179, MCF-7, MB-453 and UACC-2087) has a hypomethylated *MIR-200c* promoter. Other malignant breast cancer cell lines where miR-200c and miR-141 were suppressed (MB-157, MB-231, HS-578T and BT-549) were predominantly hypermethylated at the promoter region of the *MIR-200c* gene. The study reached a similar conclusion in malignant prostate cancer cell lines (PC3 and PC3 B1) in which each cell line that had suppressed miR-200c and miR-141 levels are hypermethylated at the *MIR-200c* promoter. In contrast, benign and less malignant prostate cancer cell lines (LNCaP and DU145, respectively) which is express miR-141 and miR-200c had a hypomethylated *MIR-200c* promoter (Vrba et al. 2010). Our observation was in parallel with the findings by Vrba *et al*; the high throughput pyrosequencing data showed that the prostate cancer cell line (LNCaP) is predominently hypomethylated at *MIR-200c*/*MIR-141* promoter and they express miR-141-3p (but not miR-141-5p) and miR-200c-3p according to our ddPCR analysis. On the other hand, PC3 cells that did not express either miR-141-3p or miR-200c-3p had a predominantly hypermethylated *MIR-200c*/*MIR-141* promoter. Aberrant DNA methylation was hypothesised to co-occur with the cell losing expression of the androgen receptor (prostate cancer cells) and oestrogen/progesterone receptor (breast cancer cells). This hypothesis arose from the observation that, as mentioned, the expression of miR-141 and miR-200c is lost only in cell lines that have no steroid hormone receptors and are hypermethylated at CpG -20 of *MIR-141*. It would be very interesting to investigate this hypothesis in the future by expanding this study into populations of patients with different grades of breast and prostate cancer and with patients in remission.

Chapter 8: General Discussion and Future Work

8. General Discussion

8.1. Main Findings

1. The novel finding from studies conducted during this PhD was the identification and confirmation of differentially methylated beta cell-specific CpG sites that can be targeted in multiplex assays which can be used to monitor beta cell death rate. In total, 867,926 CpG sites were analysed, of which, 3866 were differentially hypomethylated, and 4990 were differentially hypermethylated in islets when compared to PBMC genome. Comparing the genome-wide methylation EPIC analysis carried out in this study with methylation data from other tissues obtained from the Roadmap and ENCODE studies narrowed the CpG sites of interest to 425 hypomethylated and 228 hypermethylated CpG targets that are uniquely differentially methylated in islet but not in liver, lung, thymus, pancreas, spleen and PBMC.

2. The promoter region of the *MIR-200c*/*MIR-141* gene (Chr12: 6,963,000-6,964,001) was one of the top beta cell methylation discriminating regions. In-depth analysis using a high throughput targeted bisulfite sequencing technique confirmed that CpG dinucleotides, within the *MIR-200c*/*MIR-141* promoter region, are predominately hypomethylated in the beta cell line genome. The targeted bisulfite sequencing data have also confirmed that the methylation level of these CpG sites is minimally affected by hyperglycaemia (~5%) and they are therefore potentially useful for monitoring beta cell death in individuals with diabetes.

3. Multiplex DMR-based cfDNA assays of *MIR-200c cluster 1/INS/ MIR-200c cluster 2* were designed and verified to be not affected by hyperglycaemia or proinflammatory environment. The designed assay shown to successfully discriminate cfDNA from the beta cell genome in the circulation of a small cohort of a) recently diagnosed paediatric cases, b) individuals “at-risk” of T1D and c) islet transplant recipients.

4. Urine specimens have the potential to be used as an alternative source for cfDNA extraction and are attractive because of the non-invasive nature of the samples. Our initial experiments demonstrate that adding nucleic acid stabilising reagents such as 40 mM EDTA or commercial preservatives, are essential for protecting cfDNA from degradation in conditions where transportation of sample to the laboratory is required. It also could be needed for samples that will be stored for a long time (more than a year). A volume of < 0.5 mL is recommended for cfDNA extraction from urine for upcoming experiments.

8.2. General Discussion

Biomarkers based on differential methylation CpG sites take advantage of two well-acknowledged principles of biology. First, cells injured *in vivo* release their DNA into the circulation, and second, each cell type has a unique methylation signature, usually within the genes that are vital for cellular identity and functions. Combining these two principles allowed development of a biomarker that can potentially discriminate the cfDNA derived from a specific dying cell. A distinctive feature of this approach is the ability to discriminate cfDNA derived from normal tissues without the need to identify a mutation or rely on a genomic variation analysis as is often the case for cfDNA analysis in cancers. In this thesis, we aimed to identify a panel of the beta cell-specific DMRs by generating genome-wide methylation profiles of the islet genome and comparing it to the methylome of PBMCs, the primary source of genomic contamination in peripheral blood. From these studies, we further aimed to establish a reliable DMR based multiplex biomarker assay to detect beta cell injury in the periphery for future use in individuals with preclinical T1D and clinical trial participants or those at risk of diabetes. The promoter region of the human *MIR-200c/MIR-141* genes was found to contain several DMRs which were specific for the islet and beta cell line genome. Based on these results, several single plex assays that discriminate the beta cell genome were developed and verified. Multiplexing several assays novel beta cell-specific CpG targets, along with the previously designed exon 2 *INS* assay, greatly improved discrimination of beta cell cfDNA in the circulation of relatively small cohorts of first-degree relatives "at-risk" of T1D, recently diagnosed paediatric patients, and islet transplant recipients.

Almost all other reports in the literature describing differentially methylated beta cell-specific CpG sites were targeting the *INS* gene, with the exception of one study based on the Amylin gene (Olsen et al. 2016). One of the first reports discussing beta cell differentially methylated CpG sites, (Kuroda et al. 2009), had shown the CpG[13] −19, CpG −69, CpG −135, CpG −206, and CpG −357, from TSS of the

[13] The position of the CpG are related to the Transcription Start Site of the human insulin gene (upstream) and based on NUCL:52339278 (http://genome.ucsc.edu/cgi-bin/hgGateway, Feb 2009 GRCh37/hg19).

human *INS* promoter, were significantly hypomethylated in the beta cell but not elsewhere. Another report published similar evidence; revealing that CpGs +396 and +399 in the *INS* gene exon 1 were also specific for the pancreatic islet (Akirav et al. 2011).

It is worth highlighting here that defining the location where DMRs likely lies within a gene was a subject of some controversy. For instance, the CpG dinucleotides within the *INS* promoter were chosen as potential targets for developing DMR based biomarkers for beta cell death (Husseiny et al. 2014, Husseiny et al. 2012). Sequencing analysis carried out by the same group showed that only CpG sites within the promoter region could be used as targets for developing DMR biomarkers, while other CpG sites fall within the intron 1, intron 2 and exon 2 of the *INS* gene were not explicitly methylated in the beta cell genome (Husseiny et al. 2014). A DMR biomarker based on these findings was developed and successfully discriminated cfDNA released from dying beta cell in serum/plasma samples isolated from T1D patients who received an islet graft (Husseiny et al. 2014) and recently diagnosed T1D patients (Fisher et al. 2015). Nevertheless, other groups were able to develop a DMR biomarker based on 2 CpG sites within the exon 1 of the gene (Akirav et al. 2011). The assay was successfully used to discriminate *INS* fragments released from dying beta cell in at-risk and recently diagnosed T1D patients (Akirav et al. 2011, Herold et al. 2015a, Usmani-Brown et al. 2014), patients who received an islet graft (Bellin et al. 2017), and in studying the outcomes of an immune therapy trial (teplizumab) (Lebastchi et al. 2013). Despite the fact that methylation sensitive *INS* assays have potential as powerful tools for monitoring beta cell death, a considerable amount of hypomethylated cfDNA was reported in control samples, which was not of beta cell origin (Bellin et al. 2017, Herold et al. 2015b, Husseiny et al. 2014, Lehmann-Werman et al. 2016).

Neiman and colleagues have shown recently that the promoter of the *INS* gene was not solely hypomethylated in beta cells even within islets. They demonstrated that it was also hypomethylated in other islet cells such as alpha and delta cells. Using high throughput pyrosequencing technology, they showed that alpha and beta cells share the same promoter methylation pattern in several genes

including insulin and glucagon. Additionally, their data suggest that CpG dinucleotides located downstream of the TSS, but not those that are located in the promoter (upstream of the TSS), are more associated with the expression of the hormone, e.g. glucagon in alpha cell and insulin in the beta cell (Figure 8.1). Furthermore, global methylation analysis of over 450,000 CpG sites in alpha and beta cells using the Illumina Methylation 450K array revealed that 75% of DMRs lie within the distal regulatory regions (enhancer region) such as gene bodies or intergenic regions rather than in the promoter regions. This suggests that cell-type-specific gene expression might be associated with the enhancer methylome rather than the promoter methylome (Neiman et al. 2017). In agreement with this report, a comprehensive analysis of our EPIC data showed that the majority of DMRs lie within the gene body (hypomethylated: 37% and hypermethylated: 28%) while only 15% lie within the promoter region (Figure 4.1 in chapter 4). Taken together, these data could be the basis of changing the common belief that the promoter region methylome is the controller of tissue-specific gene expression.

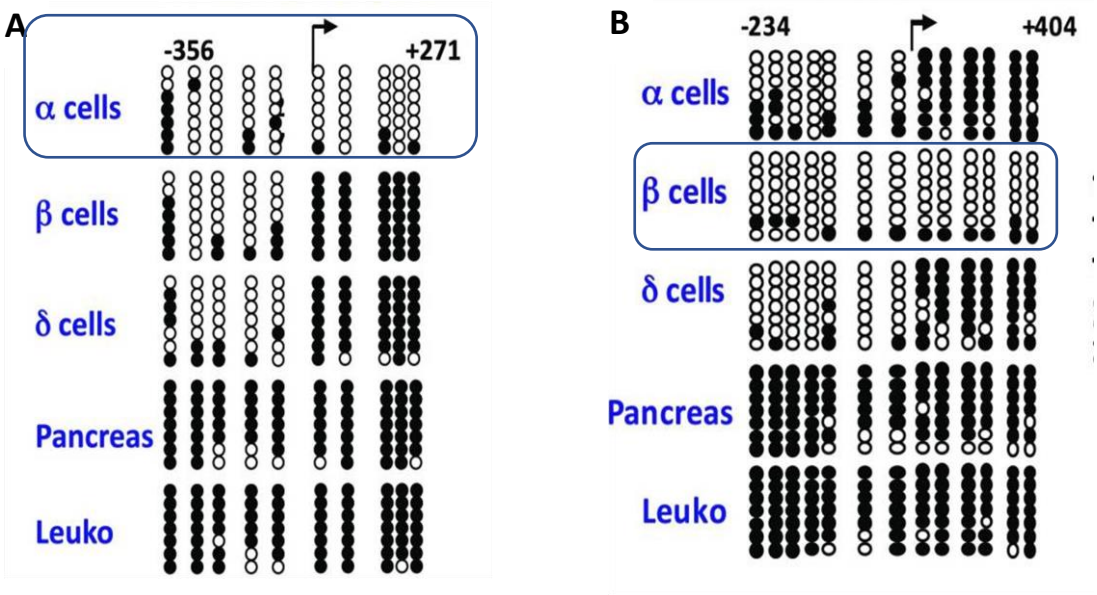


Figure 8.1: The DNA methylation of human glucagon and insulin genes in islet cells.

A) Bisulfite sequencing of CpG sites in the human glucagon promoter, from positions -356 to +271 relative to the transcription start site (arrow). B) Bisulfite sequencing of the human insulin gene promoter, containing 13 CpG dinucleotides from positions -234 to +404 relative to the transcription start site (arrow). Adapted from (Neiman et al. 2017). The figure showed that insulin gene is not only hypomethylated in beta cell but also found to be hypomethylated in alpha cell.

○ *indicates unmethylated CpG site* ● *Indicates methylated CpG site*

8.3. Biomarkers of Beta Cell Death

Although discriminating beta cell-derived cfDNA in multiple cohorts, DNA methylation-based biomarkers have some limitation that are worth to be highlighted. This approach of identifying beta cell death is mainly based on releasing beta cell DNA in the circulation during beta cell damage. As described in section 1.4.3 in the Introduction Chapter, beta cell destruction is most likely mediated by both necrotic and apoptotic mechanisms. Beta cell death by necrosis results in releasing the cellular content in the circulation and thereby, DMR assay can be applied to measure the rate of beta cell death. This may not be applied if the cell underwent apoptosis. Apoptosis involves the fragmentation of the genome, and therefore, beta cell cfDNA may not be detectable. Another limitation of using the differentially methylated biomarker is that circulating cfDNAs have a relatively short half-life in the periphery. Herold and colleges have estimated that the half-life of *INS* cfDNA is 117 ± 37 minutes (Herold et al. 2015a) Similar conclusion was also reported for circulating fetal cfDNA (Lo et al. 1999). Short half-life means that there is a chance for a pulse of beta cell death missing due to the need for rapid sample collection from the patient. Although considered a limitation, having a short half-life could be considered as a real-time reflection to beta cell death. Having a biomarker that survives for 2 hours before clearance means that the levels of this cfDNA at any time could afford an almost "real-time" measure of loss of beta cell mass in T1D in a non-invasive manner.

Another approach of detecting beta cell death is the measuring the level of circulating beta cell-specific miRNAs. miR-375 is one of the first identified and most abundant islet-specific miRNAs and also one of the best-characterised microRNAs regarding its function (Poy et al. 2009). miR-375 plays a fundamental role in normal glucose homeostasis, alpha- and beta cell turnover, and adaptive beta cell expansion in response to increasing insulin demand in insulin resistance (Dotta et al. 2018, Joglekar et al. 2009, Poy et al. 2004, Poy et al. 2009). Dr Vito Lampasona's laboratory had analysed the same plasma samples used detecting beta cell death in the circulation of islet graft recipient (samples were described in General Methods, Chapter 2, Section 2.1.2). They found a similar pattern in the levels of the circulating beta cell-specific miR-375 as we found for cfDNA with the same two samples tested in

our laboratory for beta cell death (Figure 8.2). We can see from the figure that the level of the miR-375 peaked 1-hour post the infusion, was similar to the elevated level of circulating cfDNA (Figure 6.5), and the level dropped after that, reaching the baseline level three days post infusion. We hypothesised that the longer time that miR-375 took to reach the baseline level is due to the stable nature of microRNAs in biological samples when compared with cfDNA. Although limited, the analysis showed similar results, and therefore, further studies will be needed to compare the advantages and disadvantages of cfDNA and microRNAs as circulating biomarkers in T1D.

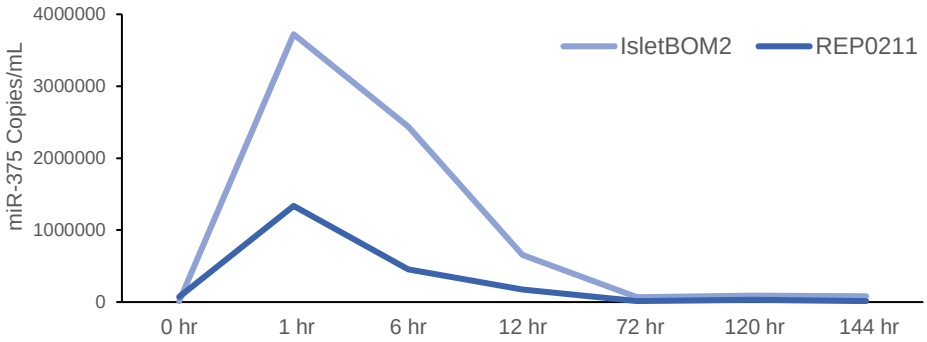


Figure 8.2: Expression levels of circulating plasma miR-375 post islet transplantation.

8.4. Strengths and Limitations

One of the strengths of this study was the genome-wide bisulfite sequencing approach, and the combination of different advanced methods, such as DNA bisulfite conversion and targeted bisulfite DNA sequencing and droplet digital PCR analysis. Another strength was the ability of the multiplex assay method to identify beta cell fragments in different T1D cohorts, e.g. recently diagnosed T1D children, at-risk individuals and acute beta cell death post islet infusion in islet graft recipients. Further, the developed assay allowed to measure the rate of beta cell death directly. Multiplexing the *INS* assay with additional different gene was an innovative approach. To the best of our knowledge, all other researchers had multiplexed several CpG sites lying within the same gene (e.g. *INS*). We were able to increase the sensitivity of the developed assay through multiplexing CpG sites within the exon 2 of the *INS* gene with other beta cell methylation specific CpG sites in the mutual promoter region of *MIR-200c*/*MIR-141*.

The following limitations are worth noting for this project:

1. Although it is the best available beta cell model (Andersson et al. 2015, Gurgul-Convey et al. 2015, Ravassard et al. 2011, Tsonkova et al. 2018), EndoC-βH1 cells were derived from human foetal pancreas tissue which may have a different methylation profile in comparison to adult beta cells. Our in-depth comparison of the methylation pattern in the *MIR-200c*/*MIR-141* promoter region in adult islet and EndoC-βH1 is ≤ 5.5 %, however; the methylation difference in other regions is unknown.
2. While it is the most widely used approach for most whole-genome methylation analysis studies, the utmost advanced available Illumina Methylation EPIC array still covers only 3% of total human CpG sites.
3. The sample size in each of clinical diabetes and “at risk” groups is limited although sample collection continues. More samples need to be measured to set the threshold level of

hypomethylated cfDNA resulting from normal beta cell turnover. This is now ongoing within the BOX and TrialNet studies.

4. Due to the limited number and volume of available clinical samples, we were unable to study the level of hypermethylated *MIR-200c*, *MIR-141* and *INS* cfDNA.

8.5. Future Direction

The data in this work provides support for a novel biomarker that discriminates beta cell fragments in the peripheral of the T1D patients. In order to take this study forward, ongoing experiments have been established. The developed assays are being repeated with larger samples size, including more control samples which will allow the definition of a threshold for positivity in the biofluid samples. In parallel with increasing the sample size of each tested cohort, available islet metabolic biomarker data, such as islet autoantibodies, C-peptide levels and the genetic risk test-1 (GRS) score, will be integrated with hypomethylation level data to give a better understanding of the processes underlying T1D.

Regarding technical approaches, the single plex assay CpG-449 of *MIR-141* will be further optimised, and more experiments performed to test the impact of replacing the *MIR-200c* cluster 2 assay by CpG-449 assay on the total discriminating power of the multiplex assay. It would also be useful to investigate the level of circulating miR-200c-3p and miR-141-3p in the sample of controls and T1D patients. Therefore the assay for extracting and quantifying miRNAs is undergoing optimisation. Another future aim is to optimise and validate a multiplex assay based on *MIR-200c/MIR-141* promoter CpG assays on prostate and/or breast cell lines with knockout hormonal receptors to test the hypothesis of changing the methylation pattern of *MIR-200c*/*MIR-141* promoter as the hormonal receptor is lost. Furthermore, the multiplex assay will be tested on serum/plasma samples obtained from patients diagnosed with benign and malignant prostate and/or breast samples to confirm the viability of using *MIR-200c/MIR-141* promoter methylation sensitive multiplex assay as a diagnostic and staging biomarker for breast and prostate cancer.

Once the multiplex assay optimisation is finalised, it will be put forward as a candidate biomarker for evaluation in an international workshop, e.g. Biomarkers of Beta Cell Stress in Type 1 Diabetes (BetaMarker), Human Islet Research Network, for discussion. If proven to be effective in the workshop setting, the multiplex cfDNA assay would provide a high throughput test for studies of progression and therapeutic intervention trials in T1D.

8.6. Conclusion

In recent years, circulating beta cell-specific, DMR based biomarkers are recognised as potentially powerful tools to detect real-time beta cell death in multiple type 1 diabetes cohorts. Although highly promising, these differentially methylated cfDNA biomarkers are still in their infancy. The multiplex DMR-based, the beta cell-specific assay described here has the potential to track the rate of beta cell death in multiple cohorts with high specificity and sensitivity. Data generated by the assay may help predict the rate of progression of T1D and the survival of beta cells post-transplantation. Further studies are required to validate and evaluate our findings in another biofluid sample, such as urine, and in larger cohorts as well as confirming the utility of the assay in clinical settings such as immune interventions.

Chapter 9: References

9. References

Aapola, U., K. Kawasaki, H. S. Scott, J. Ollila, M. Vihinen, M. Heino, A. Shintani, K. Kawasaki, S. Minoshima, K. Krohn, S. E. Antonarakis, N. Shimizu, J. Kudoh, and P. Peterson. 2000. "Isolation and initial characterization of a novel zinc finger gene, DNMT3L, on 21q22.3, related to the cytosine-5-methyltransferase 3 gene family." *Genomics* 65 (3):293-8. doi: 10.1006/geno.2000.6168.

Adams, D., L. Altucci, S. E. Antonarakis, J. Ballesteros, S. Beck, A. Bird, C. Bock, B. Boehm, E. Campo, A. Caricasole, F. Dahl, E. T. Dermitzakis, T. Enver, M. Esteller, X. Estivill, A. Ferguson-Smith, J. Fitzgibbon, P. Flicek, C. Giehl, T. Graf, F. Grosveld, R. Guigo, I. Gut, K. Helin, J. Jarvius, R. Kuppers, H. Lehrach, T. Lengauer, A. Lernmark, D. Leslie, M. Loeffler, E. Macintyre, A. Mai, J. H. Martens, S. Minucci, W. H. Ouwehand, P. G. Pelicci, H. Pendeville, B. Porse, V. Rakyan, W. Reik, M. Schrappe, D. Schubeler, M. Seifert, R. Siebert, D. Simmons, N. Soranzo, S. Spicuglia, M. Stratton, H. G. Stunnenberg, A. Tanay, D. Torrents, A. Valencia, E. Vellenga, M. Vingron, J. Walter, and S. Willcocks. 2012. "BLUEPRINT to decode the epigenetic signature written in blood." *Nat Biotechnol* 30 (3):224-6. doi: 10.1038/nbt.2153.

Ahlgren, U., J. Jonsson, L. Jonsson, K. Simu, and H. Edlund. 1998. "beta-cell-specific inactivation of the mouse Ipf1/Pdx1 gene results in loss of the beta-cell phenotype and maturity onset diabetes." *Genes Dev* 12 (12):1763-8.

Akerblom, H. K., O. Vaarala, H. Hyoty, J. Ilonen, and M. Knip. 2002. "Environmental factors in the etiology of type 1 diabetes." *Am J Med Genet* 115 (1):18-29. doi: 10.1002/ajmg.10340.

Akinci, E., A. Banga, L. V. Greder, J. R. Dutton, and J. M. Slack. 2012. "Reprogramming of pancreatic exocrine cells towards a beta (beta) cell character using Pdx1, Ngn3 and MafA." *Biochem J* 442 (3):539-50. doi: 10.1042/bj20111678.

Akirav, E. M., J. Lebastchi, E. M. Galvan, O. Henegariu, M. Akirav, V. Ablamunits, P. M. Lizardi, and K. C. Herold. 2011. "Detection of beta cell death in diabetes using differentially methylated circulating DNA." *Proc Natl Acad Sci U S A* 108 (47):19018-23. doi: 10.1073/pnas.1111008108.

Alberts, B, A Johnson, and J Lewis. 2002. "Chromosomal DNA and Its Packaging in the Chromatin Fiber." In *Molecular Biology of the Cell*. New York: Garland Science.

Almawi, W. Y., H. Tamim, and S. T. Azar. 1999. "Clinical review 103: T helper type 1 and 2 cytokines mediate the onset and progression of type I (insulin-dependent) diabetes." *J Clin Endocrinol Metab* 84 (5):1497-502. doi: 10.1210/jcem.84.5.5699.

Andersson, L. E., B. Valtat, A. Bagge, V. V. Sharoyko, D. G. Nicholls, P. Ravassard, R. Scharfmann, P. Spegel, and H. Mulder. 2015. "Characterization of stimulus-secretion coupling in the human

pancreatic EndoC-betaH1 beta cell line." *PLoS One* 10 (3):e0120879. doi: 10.1371/journal.pone.0120879.

Andreoletti, L., D. Hober, C. Hober-Vandenberghe, I. Fajardy, S. Belaich, V. Lambert, M. C. Vantyghem, J. Lefebvre, and P. Wattre. 1998. "Coxsackie B virus infection and beta cell autoantibodies in newly diagnosed IDDM adult patients." *Clin Diagn Virol* 9 (2-3):125-33.

Anjos, S., A. Nguyen, H. Ounissi-Benkalha, M. C. Tessier, and C. Polychronakos. 2002. "A common autoimmunity predisposing signal peptide variant of the cytotoxic T-lymphocyte antigen 4 results in inefficient glycosylation of the susceptibility allele." *J Biol Chem* 277 (48):46478-86. doi: 10.1074/jbc.M206894200.

Anjos, S., M. C. Tessier, and C. Polychronakos. 2004. "Association of the cytotoxic T lymphocyte-associated antigen 4 gene with type 1 diabetes: evidence for independent effects of two polymorphisms on the same haplotype block." *J Clin Endocrinol Metab* 89 (12):6257-65. doi: 10.1210/jc.2004-0881.

Arif, S., T. I. Tree, T. P. Astill, J. M. Tremble, A. J. Bishop, C. M. Dayan, B. O. Roep, and M. Peakman. 2004. "Autoreactive T cell responses show proinflammatory polarization in diabetes but a regulatory phenotype in health." *J Clin Invest* 113 (3):451-63. doi: 10.1172/JCI19585.

Assmann, T. S., M. Recamonde-Mendoza, B. M. De Souza, and D. Crispim. 2017. "MicroRNA expression profiles and type 1 diabetes mellitus: systematic review and bioinformatic analysis." *Endocr Connect* 6 (8):773-790. doi: 10.1530/EC-17-0248.

Bachmann, M. F., and A. Oxenius. 2007. "Interleukin 2: from immunostimulation to immunoregulation and back again." *EMBO Rep* 8 (12):1142-8. doi: 10.1038/sj.embor.7401099.

Baekkeskov, S., H. J. Aanstoot, S. Christgau, A. Reetz, M. Solimena, M. Cascalho, F. Folli, H. Richter-Olesen, and P. De Camilli. 1990. "Identification of the 64K autoantigen in insulin-dependent diabetes as the GABA-synthesizing enzyme glutamic acid decarboxylase." *Nature* 347 (6289):151-6. doi: 10.1038/347151a0.

Bannister, A. J., and T. Kouzarides. 2011. "Regulation of chromatin by histone modifications." *Cell Res* 21 (3):381-95. doi: 10.1038/cr.2011.22.

Barau, J., A. Teissandier, N. Zamudio, S. Roy, V. Nalesso, Y. Herault, F. Guillou, and D. Bourc'his. 2016. "The DNA methyltransferase DNMT3C protects male germ cells from transposon activity." *Science* 354 (6314):909-912. doi: 10.1126/science.aah5143.

Barlow, D. P. 1993. "Methylation and imprinting: from host defense to gene regulation?" *Science* 260 (5106):309-10.

Barratt, B. J., F. Payne, C. E. Lowe, R. Hermann, B. C. Healy, D. Harold, P. Concannon, N. Gharani, M. I. McCarthy, M. G. Olavesen, R. McCormack, C. Guja, C. Ionescu-Tirgoviste, D. E. Undlien, K. S.

Ronningen, K. M. Gillespie, E. Tuomilehto-Wolf, J. Tuomilehto, S. T. Bennett, D. G. Clayton, H. J. Cordell, and J. A. Todd. 2004. "Remapping the insulin gene/IDDM2 locus in type 1 diabetes." *Diabetes* 53 (7):1884-9.

Barrett, J. C., D. G. Clayton, P. Concannon, B. Akolkar, J. D. Cooper, H. A. Erlich, C. Julier, G. Morahan, J. Nerup, C. Nierras, V. Plagnol, F. Pociot, H. Schuilenburg, D. J. Smyth, H. Stevens, J. A. Todd, N. M. Walker, S. S. Rich, and Consortium Type 1 Diabetes Genetics. 2009. "Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes." *Nat Genet* 41 (6):703-7. doi: 10.1038/ng.381.

Barrett, J. H., M. M. Iles, M. Harland, J. C. Taylor, J. F. Aitken, P. A. Andresen, L. A. Akslen, B. K. Armstrong, M. F. Avril, E. Azizi, B. Bakker, W. Bergman, G. Bianchi-Scarra, B. Bressac-de Paillerets, D. Calista, L. A. Cannon-Albright, E. Corda, A. E. Cust, T. Debniak, D. Duffy, A. M. Dunning, D. F. Easton, E. Friedman, P. Galan, P. Ghiorzo, G. G. Giles, J. Hansson, M. Hocevar, V. Hoiom, J. L. Hopper, C. Ingvar, B. Janssen, M. A. Jenkins, G. Jonsson, R. F. Kefford, G. Landi, M. T. Landi, J. Lang, J. Lubinski, R. Mackie, J. Malvehy, N. G. Martin, A. Molven, G. W. Montgomery, F. A. van Nieuwpoort, S. Novakovic, H. Olsson, L. Pastorino, S. Puig, J. A. Puig-Butille, J. Randerson-Moor, H. Snowden, R. Tuominen, P. Van Belle, N. van der Stoep, D. C. Whiteman, D. Zelenika, J. Han, S. Fang, J. E. Lee, Q. Wei, G. M. Lathrop, E. M. Gillanders, K. M. Brown, A. M. Goldstein, P. A. Kanetsky, G. J. Mann, S. Macgregor, D. E. Elder, C. I. Amos, N. K. Hayward, N. A. Gruis, F. Demenais, J. A. Bishop, D. T. Bishop, and M. E. L. Consortium Geno. 2011. "Genome-wide association study identifies three new melanoma susceptibility loci." *Nat Genet* 43 (11):1108-13. doi: 10.1038/ng.959.

Bartel, D. P. 2004. "MicroRNAs: genomics, biogenesis, mechanism, and function." *Cell* 116 (2):281-97.

Bason, C., R. Lorini, C. Lunardi, M. Dolcino, A. Giannattasio, G. d'Annunzio, A. Rigo, N. Pedemonte, R. Corrocher, and A. Puccetti. 2013. "In type 1 diabetes a subset of anti-coxsackievirus B4 antibodies recognize autoantigens and induce apoptosis of pancreatic beta cells." *PLoS One* 8 (2):e57729. doi: 10.1371/journal.pone.0057729.

Beck, S., and V. K. Rakyan. 2008. "The methylome: approaches for global DNA methylation profiling." *Trends Genet* 24 (5):231-7. doi: 10.1016/j.tig.2008.01.006.

Belgardt, B. F., K. Ahmed, M. Spranger, M. Latreille, R. Denzler, N. Kondratiuk, F. von Meyenn, F. N. Villena, K. Herrmanns, D. Bosco, J. Kerr-Conte, F. Pattou, T. Rulicke, and M. Stoffel. 2015. "The microRNA-200 family regulates pancreatic beta cell survival in type 2 diabetes." *Nat Med* 21 (6):619-27. doi: 10.1038/nm.3862.

Bell, G. I., S. Horita, and J. H. Karam. 1984. "A polymorphic locus near the human insulin gene is associated with insulin-dependent diabetes mellitus." *Diabetes* 33 (2):176-83.

Bellin, M. D., P. Clark, S. Usmani-Brown, T. B. Dunn, G. J. Beilman, S. Chinnakotla, T. L. Pruett, P. Ptacek, B. J. Hering, Z. Wang, T. Gilmore, J. J. Wilhelm, J. S. Hodges, A. Moran, and K. C. Herold. 2017. "Unmethylated Insulin DNA Is Elevated After Total Pancreatectomy With Islet Autotransplantation: Assessment of a Novel Beta Cell Marker." *Am J Transplant* 17 (4):1112-1118. doi: 10.1111/ajt.14054.

Belot, M. P., D. Fradin, N. Mai, S. Le Fur, D. Zelenika, J. Kerr-Conte, F. Pattou, B. Lucas, and P. Bougneres. 2013. "CpG methylation changes within the IL2RA promoter in type 1 diabetes of childhood onset." *PLoS One* 8 (7):e68093. doi: 10.1371/journal.pone.0068093.

Bennett, S. T., A. M. Lucassen, S. C. Gough, E. E. Powell, D. E. Undlien, L. E. Pritchard, M. E. Merriman, Y. Kawaguchi, M. J. Dronsfield, F. Pociot, and et al. 1995. "Susceptibility to human type 1 diabetes at IDDM2 is determined by tandem repeat variation at the insulin gene minisatellite locus." *Nat Genet* 9 (3):284-92. doi: 10.1038/ng0395-284.

Bernstein, B. E., J. A. Stamatoyannopoulos, J. F. Costello, B. Ren, A. Milosavljevic, A. Meissner, M. Kellis, M. A. Marra, A. L. Beaudet, J. R. Ecker, P. J. Farnham, M. Hirst, E. S. Lander, T. S. Mikkelsen, and J. A. Thomson. 2010. "The NIH Roadmap Epigenomics Mapping Consortium." *Nat Biotechnol* 28 (10):1045-8. doi: 10.1038/nbt1010-1045.

Bianconi, E., A. Piovesan, F. Facchin, A. Beraudi, R. Casadei, F. Frabetti, L. Vitale, M. C. Pelleri, S. Tassani, F. Piva, S. Perez-Amodio, P. Strippoli, and S. Canaider. 2013. "An estimation of the number of cells in the human body." *Ann Hum Biol* 40 (6):463-71. doi: 10.3109/03014460.2013.807878.

Bikle, D. D. 2014. "Vitamin D metabolism, mechanism of action, and clinical applications." *Chem Biol* 21 (3):319-29. doi: 10.1016/j.chembiol.2013.12.016.

Bingley, P. J., E. Bonifacio, A. G. Ziegler, D. A. Schatz, M. A. Atkinson, G. S. Eisenbarth, and Society Immunology of Diabetes. 2001. "Proposed guidelines on screening for risk of type 1 diabetes." *Diabetes Care* 24 (2):398.

Bingley, P. J., E. A. Gale, and Group European Nicotinamide Diabetes Intervention Trial. 2006. "Progression to type 1 diabetes in islet cell antibody-positive relatives in the European Nicotinamide Diabetes Intervention Trial: the role of additional immune, genetic and metabolic markers of risk." *Diabetologia* 49 (5):881-90. doi: 10.1007/s00125-006-0160-4.

Biomarkers Definitions Working Group. 2001. "Biomarkers and surrogate endpoints: preferred definitions and conceptual framework." *Clin Pharmacol Ther* 69 (3):89-95. doi: 10.1067/mcp.2001.113989.

Bird, A. P., and A. P. Wolffe. 1999. "Methylation-induced repression--belts, braces, and chromatin." *Cell* 99 (5):451-4.

Bock, C., E. M. Tomazou, A. B. Brinkman, F. Muller, F. Simmer, H. Gu, N. Jager, A. Gnirke, H. G. Stunnenberg, and A. Meissner. 2010. "Quantitative comparison of genome-wide DNA methylation mapping technologies." *Nat Biotechnol* 28 (10):1106-14. doi: 10.1038/nbt.1681.

Bodansky, H. J., A. Staines, C. Stephenson, D. Haigh, and R. Cartwright. 1992. "Evidence for an environmental effect in the aetiology of insulin dependent diabetes in a transmigratory population." *BMJ* 304 (6833):1020-2.

Bonifacio, E., V. Lampasona, S. Genovese, M. Ferrari, and E. Bosi. 1995. "Identification of protein tyrosine phosphatase-like IA2 (islet cell antigen 512) as the insulin-dependent diabetes-related 37/40K autoantigen and a target of islet-cell antibodies." *J Immunol* 155 (11):5419-26.

Bonifacio, E., M. Scirpoli, K. Kredel, M. Fuchtenbusch, and A. G. Ziegler. 1999. "Early autoantibody responses in prediabetes are IgG1 dominated and suggest antigen-specific regulation." *J Immunol* 163 (1):525-32.

Boomsma, D., A. Busjahn, and L. Peltonen. 2002. "Classical twin studies and beyond." *Nat Rev Genet* 3 (11):872-82. doi: 10.1038/nrg932.

Borchert, G. M., W. Lanier, and B. L. Davidson. 2006. "RNA polymerase III transcribes human microRNAs." *Nat Struct Mol Biol* 13 (12):1097-101. doi: 10.1038/nsmb1167.

Bosi, E., S. Braghi, P. Maffi, M. Scirpoli, F. Bertuzzi, G. Pozza, A. Secchi, and E. Bonifacio. 2001. "Autoantibody response to islet transplantation in type 1 diabetes." *Diabetes* 50 (11):2464-71.

Bottazzo, G. F., B. M. Dean, J. M. McNally, E. H. MacKay, P. G. Swift, and D. R. Gamble. 1985. "In situ characterization of autoimmune phenomena and expression of HLA molecules in the pancreas in diabetic insulitis." *N Engl J Med* 313 (6):353-60. doi: 10.1056/NEJM198508083130604.

Bottazzo, G. F., A. Florin-Christensen, and D. Doniach. 1974. "Islet-cell antibodies in diabetes mellitus with autoimmune polyendocrine deficiencies." *Lancet* 2 (7892):1279-83.

Bottini, N., L. Musumeci, A. Alonso, S. Rahmouni, K. Nika, M. Rostamkhani, J. MacMurray, G. F. Meloni, P. Lucarelli, M. Pellecchia, G. S. Eisenbarth, D. Comings, and T. Mustelin. 2004. "A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes." *Nat Genet* 36 (4):337-8. doi: 10.1038/ng1323.

Bougneres, P. F., J. C. Carel, L. Castano, C. Boitard, J. P. Gardin, P. Landais, J. Hors, M. J. Mihatsch, M. Paillard, J. L. Chaussain, and et al. 1988. "Factors associated with early remission of type I diabetes in children treated with cyclosporine." *N Engl J Med* 318 (11):663-70. doi: 10.1056/NEJM198803173181103.

Bourc'his, D., and T. H. Bestor. 2004. "Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L." *Nature* 431 (7004):96-9. doi: 10.1038/nature02886.

Bourc'his, D., G. L. Xu, C. S. Lin, B. Bollman, and T. H. Bestor. 2001. "Dnmt3L and the establishment of maternal genomic imprints." *Science* 294 (5551):2536-9. doi: 10.1126/science.1065848.

Boutros, P. C. 2015. "The path to routine use of genomic biomarkers in the cancer clinic." *Genome Res* 25 (10):1508-13. doi: 10.1101/gr.191114.115.

Bradfield, J. P., H. Q. Qu, K. Wang, H. Zhang, P. M. Sleiman, C. E. Kim, F. D. Mentch, H. Qiu, J. T. Glessner, K. A. Thomas, E. C. Frackelton, R. M. Chiavacci, M. Imielinski, D. S. Monos, R. Pandey, M. Bakay, S. F. Grant, C. Polychronakos, and H. Hakonarson. 2011. "A genome-wide meta-analysis of six type 1 diabetes cohorts identifies multiple associated loci." *PLoS Genet* 7 (9):e1002293. doi: 10.1371/journal.pgen.1002293.

Braghi, S., E. Bonifacio, A. Secchi, V. Di Carlo, G. Pozza, and E. Bosi. 2000. "Modulation of humoral islet autoimmunity by pancreas allotransplantation influences allograft outcome in patients with type 1 diabetes." *Diabetes* 49 (2):218-24.

Bramswig, N. C., L. J. Everett, J. Schug, C. Dorrell, C. Liu, Y. Luo, P. R. Streeter, A. Naji, M. Grompe, and K. H. Kaestner. 2013. "Epigenomic plasticity enables human pancreatic alpha to beta cell reprogramming." *J Clin Invest* 123 (3):1275-84. doi: 10.1172/JCI66514.

Brase, J. C., M. Johannes, T. Schlomm, M. Falth, A. Haese, T. Steuber, T. Beissbarth, R. Kuner, and H. Sultmann. 2011. "Circulating miRNAs are correlated with tumor progression in prostate cancer." *Int J Cancer* 128 (3):608-16. doi: 10.1002/ijc.25376.

Brissova, M., M. J. Fowler, W. E. Nicholson, A. Chu, B. Hirshberg, D. M. Harlan, and A. C. Powers. 2005. "Assessment of human pancreatic islet architecture and composition by laser scanning confocal microscopy." *J Histochem Cytochem* 53 (9):1087-97. doi: 10.1369/jhc.5C6684.2005.

Brown, C. T., A. G. Davis-Richardson, A. Giongo, K. A. Gano, D. B. Crabb, N. Mukherjee, G. Casella, J. C. Drew, J. Ilonen, M. Knip, H. Hyoty, R. Veijola, T. Simell, O. Simell, J. Neu, C. H. Wasserfall, D. Schatz, M. A. Atkinson, and E. W. Triplett. 2011. "Gut microbiome metagenomics analysis suggests a functional model for the development of autoimmunity for type 1 diabetes." *PLoS One* 6 (10):e25792. doi: 10.1371/journal.pone.0025792.

Bryzgunova, O. E., and P. P. Laktionov. 2015. "Extracellular Nucleic Acids in Urine: Sources, Structure, Diagnostic Potential." *Acta Naturae* 7 (3):48-54.

Bryzgunova, O. E., T. E. Skvortsova, E. V. Kolesnikova, A. V. Starikov, E. Y. Rykova, V. V. Vlassov, and P. P. Laktionov. 2006. "Isolation and comparative study of cell-free nucleic acids from human urine." *Ann N Y Acad Sci* 1075:334-40. doi: 10.1196/annals.1368.045.

Cabrera, O., D. M. Berman, N. S. Kenyon, C. Ricordi, P. O. Berggren, and A. Caicedo. 2006. "The unique cytoarchitecture of human pancreatic islets has implications for islet cell function." *Proc Natl Acad Sci U S A* 103 (7):2334-9. doi: 10.1073/pnas.0510790103.

Campbell-Thompson, M. L., M. A. Atkinson, A. E. Butler, N. M. Chapman, G. Frisk, R. Gianani, B. N. Giepmans, M. G. von Herrath, H. Hyoty, T. W. Kay, O. Korsgren, N. G. Morgan, A. C. Powers, A. Pugliese, S. J. Richardson, P. A. Rowe, S. Tracy, and P. A. In't Veld. 2013. "The diagnosis of insulitis in human type 1 diabetes." *Diabetologia* 56 (11):2541-3. doi: 10.1007/s00125-013-3043-5.

Campbell-Thompson, M. L., M. A. Atkinson, A. E. Butler, B. N. Giepmans, M. G. von Herrath, H. Hyoty, T. W. Kay, N. G. Morgan, A. C. Powers, A. Pugliese, S. J. Richardson, and P. A. In't Veld. 2017. "Re-addressing the 2013 consensus guidelines for the diagnosis of insulitis in human type 1 diabetes: is change necessary?" *Diabetologia* 60 (4):753-755. doi: 10.1007/s00125-016-4195-x.

Campbell-Thompson, M. L., C. Wasserfall, J. Kaddis, A. Albanese-O'Neill, T. Staeva, C. Nierras, J. Moraski, P. Rowe, R. Gianani, G. Eisenbarth, J. Crawford, D. Schatz, A. Pugliese, and M. Atkinson. 2012. "Network for Pancreatic Organ Donors with Diabetes (nPOD): developing a tissue biobank for type 1 diabetes." *Diabetes Metab Res Rev* 28 (7):608-17. doi: 10.1002/dmrr.2316.

Cannas, A., G. Kalunga, C. Green, L. Calvo, P. Katemangwe, K. Reither, M. D. Perkins, L. Maboko, M. Hoelscher, E. A. Talbot, P. Mwaba, A. I. Zumla, E. Girardi, J. F. Huggett, and T. B. trDNA consortium. 2009. "Implications of storing urinary DNA from different populations for molecular analyses." *PLoS One* 4 (9):e6985. doi: 10.1371/journal.pone.0006985.

Cantone, I., and A. G. Fisher. 2013. "Epigenetic programming and reprogramming during development." *Nat Struct Mol Biol* 20 (3):282-9. doi: 10.1038/nsmb.2489.

Carpino, N., S. Turner, D. Mekala, Y. Takahashi, H. Zang, T. L. Geiger, P. Doherty, and J. N. Ihle. 2004. "Regulation of ZAP-70 activation and TCR signaling by two related proteins, Sts-1 and Sts-2." *Immunity* 20 (1):37-46.

Cate, R. L., W. Chick, and W. Gilbert. 1983. "Comparison of the methylation patterns of the two rat insulin genes." *J Biol Chem* 258 (10):6645-52.

Chaplin, D. D. 2010. "Overview of the immune response." *J Allergy Clin Immunol* 125 (2 Suppl 2):S3-23. doi: 10.1016/j.jaci.2009.12.980.

Chatenoud, L., and J. A. Bluestone. 2007. "CD3-specific antibodies: a portal to the treatment of autoimmunity." *Nat Rev Immunol* 7 (8):622-32. doi: 10.1038/nri2134.

Chedin, F., M. R. Lieber, and C. L. Hsieh. 2002. "The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a." *Proc Natl Acad Sci U S A* 99 (26):16916-21. doi: 10.1073/pnas.262443999.

Chen, C., R. Fang, C. Davis, C. Maravelias, and E. Sibley. 2009. "Pdx1 inactivation restricted to the intestinal epithelium in mice alters duodenal gene expression in enterocytes and enteroendocrine cells." *Am J Physiol Gastrointest Liver Physiol* 297 (6):G1126-37. doi: 10.1152/ajpgi.90586.2008.

Chen, T., Y. Ueda, J. E. Dodge, Z. Wang, and E. Li. 2003. "Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b." *Mol Cell Biol* 23 (16):5594-605.

Chendrimada, T. P., R. I. Gregory, E. Kumaraswamy, J. Norman, N. Cooch, K. Nishikura, and R. Shiekhattar. 2005. "TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing." *Nature* 436 (7051):740-4. doi: 10.1038/nature03868.

Chinen, T., A. K. Kannan, A. G. Levine, X. Fan, U. Klein, Y. Zheng, G. Gasteiger, Y. Feng, J. D. Fontenot, and A. Y. Rudensky. 2016. "An essential role for the IL-2 receptor in Treg cell function." *Nat Immunol* 17 (11):1322-1333. doi: 10.1038/ni.3540.

Clark, P. M. 1999. "Assays for insulin, proinsulin(s) and C-peptide." *Ann Clin Biochem* 36 (Pt 5):541-64. doi: 10.1177/000456329903600501.

Clayton, D. G. 2009. "Prediction and interaction in complex disease genetics: experience in type 1 diabetes." *PLoS Genet* 5 (7):e1000540. doi: 10.1371/journal.pgen.1000540.

Clements, G. B., F. McGarry, C. Nairn, and D. N. Galbraith. 1995. "Detection of enterovirus-specific RNA in serum: the relationship to chronic fatigue." *J Med Virol* 45 (2):156-61.

Cnop, M., N. Welsh, J. C. Jonas, A. Jorns, S. Lenzen, and D. L. Eizirik. 2005. "Mechanisms of pancreatic beta-cell death in type 1 and type 2 diabetes: many differences, few similarities." *Diabetes* 54 Suppl 2:S97-107.

Collingwood, T. S., E. V. Smirnova, M. Bogush, N. Carpino, R. S. Annan, and A. Y. Tsygankov. 2007. "T-cell ubiquitin ligand affects cell death through a functional interaction with apoptosis-inducing factor, a key factor of caspase-independent apoptosis." *J Biol Chem* 282 (42):30920-8. doi: 10.1074/jbc.M706870200.

Collins, Lesley J., Barbara Schönfeld, and Xiaowei Sylvia Chen4. 2017. "The Epigenetics of Non-coding RNA." In *Handbook of Epigenetics: The New Molecular and Medical Genetics*, edited by Tollefsbol T., 49-61. USA: Academic Press 2017.

Concannon, P., S. Onengut-Gumuscu, J. A. Todd, D. J. Smyth, F. Pociot, R. Bergholdt, B. Akolkar, H. A. Erlich, J. E. Hilner, C. Julier, G. Morahan, J. Nerup, C. R. Nierras, W. M. Chen, S. S. Rich, and

Consortium Type 1 Diabetes Genetics. 2008. "A human type 1 diabetes susceptibility locus maps to chromosome 21q22.3." *Diabetes* 57 (10):2858-61. doi: 10.2337/db08-0753.

Consortium, Encode Project. 2012. "An integrated encyclopedia of DNA elements in the human genome." *Nature* 489 (7414):57-74. doi: 10.1038/nature11247.

Cooper, D. N., and S. Gerber-Huber. 1985. "DNA methylation and CpG suppression." *Cell Differ* 17 (3):199-205.

Cooper, J. D., D. J. Smyth, A. M. Smiles, V. Plagnol, N. M. Walker, J. E. Allen, K. Downes, J. C. Barrett, B. C. Healy, J. C. Mychaleckyj, J. H. Warram, and J. A. Todd. 2008. "Meta-analysis of genome-wide association study data identifies additional type 1 diabetes risk loci." *Nat Genet* 40 (12):1399-401. doi: 10.1038/ng.249.

Cooper, J. D., N. M. Walker, D. J. Smyth, K. Downes, B. C. Healy, J. A. Todd, and I. Diabetes Genetics Consortium Type. 2009. "Follow-up of 1715 SNPs from the Wellcome Trust Case Control Consortium genome-wide association study in type I diabetes families." *Genes Immun* 10 Suppl 1:S85-94. doi: 10.1038/gene.2009.97.

Coppieters, K. T., F. Dotta, N. Amirian, P. D. Campbell, T. W. Kay, M. A. Atkinson, B. O. Roep, and M. G. von Herrath. 2012. "Demonstration of islet-autoreactive CD8 T cells in insulitic lesions from recent onset and long-term type 1 diabetes patients." *J Exp Med* 209 (1):51-60. doi: 10.1084/jem.20111187.

Court, F., C. Tayama, V. Romanelli, A. Martin-Trujillo, I. Iglesias-Platas, K. Okamura, N. Sugahara, C. Simon, H. Moore, J. V. Harness, H. Keirstead, J. V. Sanchez-Mut, E. Kaneki, P. Lapunzina, H. Soejima, N. Wake, M. Esteller, T. Ogata, K. Hata, K. Nakabayashi, and D. Monk. 2014. "Genome-wide parent-of-origin DNA methylation analysis reveals the intricacies of human imprinting and suggests a germline methylation-independent mechanism of establishment." *Genome Res* 24 (4):554-69. doi: 10.1101/gr.164913.113.

Cresswell, P. 1994. "Assembly, transport, and function of MHC class II molecules." *Annu Rev Immunol* 12:259-93. doi: 10.1146/annurev.iy.12.040194.001355.

Cudworth, A. G., and J. C. Woodrow. 1974. "Letter: HL-A antigens and diabetes mellitus." *Lancet* 2 (7889):1153.

Cudworth, A. G., and J. C. Woodrow. 1975. "Evidence for HL-A-linked genes in "juvenile" diabetes mellitus." *Br Med J* 3 (5976):133-5.

Cudworth, A. G., and J. C. Woodrow. 1976. "Genetic susceptibility in diabetes mellitus: analysis of the HLA association." *Br Med J* 2 (6040):846-8.

Dahlquist, G. G., L. Nystrom, and C. C. Patterson. 2011. "Incidence of type 1 diabetes in Sweden among individuals aged 0-34 years, 1983-2007: an analysis of time trends." *Diabetes Care* 34 (8):1754-9. doi: 10.2337/dc11-0056.

Dariavach, P., M. G. Mattei, P. Golstein, and M. P. Lefranc. 1988. "Human Ig superfamily CTLA-4 gene: chromosomal localization and identity of protein sequence between murine and human CTLA-4 cytoplasmic domains." *Eur J Immunol* 18 (12):1901-5. doi: 10.1002/eji.1830181206.

De Filippo, G., N. Pozzi, E. Cosentini, M. Cavalcanti, J. C. Carel, S. Tamasi, A. Franzese, and C. Pignata. 1997. "Increased CD5+CD19+ B lymphocytes at the onset of type 1 diabetes in children." *Acta Diabetol* 34 (4):271-4.

de Jong, V. M., A. Zaldumbide, A. R. van der Slik, S. Laban, B. P. Koeleman, and B. O. Roep. 2016. "Variation in the CTLA4 3'UTR has phenotypic consequences for autoreactive T cells and associates with genetic risk for type 1 diabetes." *Genes Immun* 17 (1):75-8. doi: 10.1038/gene.2015.51.

Dean, F. B., S. Hosono, L. Fang, X. Wu, A. F. Faruqi, P. Bray-Ward, Z. Sun, Q. Zong, Y. Du, J. Du, M. Driscoll, W. Song, S. F. Kingsmore, M. Egholm, and R. S. Lasken. 2002. "Comprehensive human genome amplification using multiple displacement amplification." *Proc Natl Acad Sci U S A* 99 (8):5261-6. doi: 10.1073/pnas.082089499.

Deichmann, K., A. Heinzmann, E. Bruggenolte, J. Forster, and J. Kuehr. 1996. "An Mse I RFLP in the human CTLA4 promotor." *Biochem Biophys Res Commun* 225 (3):817-8. doi: 10.1006/bbrc.1996.1256.

Demeester, S., B. Keymeulen, L. Kaufman, A. Van Dalem, E. V. Balti, U. Van de Velde, P. Goubert, K. Verhaeghen, H. W. Davidson, J. M. Wenzlau, I. Weets, D. G. Pipeleers, and F. K. Gorus. 2015. "Preexisting insulin autoantibodies predict efficacy of otelixizumab in preserving residual beta-cell function in recent-onset type 1 diabetes." *Diabetes Care* 38 (4):644-51. doi: 10.2337/dc14-1575.

DiSalvo, T. G. 2015. "Epigenetic regulation in heart failure: part II DNA and chromatin." *Cardiol Rev* 23 (6):269-81. doi: 10.1097/CRD.0000000000000074.

Don, R. H., P. T. Cox, B. J. Wainwright, K. Baker, and J. S. Mattick. 1991. "'Touchdown' PCR to circumvent spurious priming during gene amplification." *Nucleic Acids Res* 19 (14):4008.

Dong, X., and Z. Weng. 2013. "The correlation between histone modifications and gene expression." *Epigenomics* 5 (2):113-6. doi: 10.2217/epi.13.13.

Dotta, F., S. Censini, A. G. van Halteren, L. Marselli, M. Masini, S. Dionisi, F. Mosca, U. Boggi, A. O. Muda, S. Del Prato, J. F. Elliott, A. Covacci, R. Rappuoli, B. O. Roep, and P. Marchetti. 2007.

"Coxsackie B4 virus infection of beta cells and natural killer cell insulitis in recent-onset type 1 diabetic patients." *Proc Natl Acad Sci U S A* 104 (12):5115-20. doi: 10.1073/pnas.0700442104.

Dotta, F., G. Ventriglia, I. V. Snowhite, and A. Pugliese. 2018. "MicroRNAs: markers of beta-cell stress and autoimmunity." *Curr Opin Endocrinol Diabetes Obes*. doi: 10.1097/MED.0000000000000420.

Downes, K., M. Pekalski, K. L. Angus, M. Hardy, S. Nutland, D. J. Smyth, N. M. Walker, C. Wallace, and J. A. Todd. 2010. "Reduced expression of IFIH1 is protective for type 1 diabetes." *PLoS One* 5 (9). doi: 10.1371/journal.pone.0012646.

Dupont, C., D. R. Armant, and C. A. Brenner. 2009. "Epigenetics: definition, mechanisms and clinical perspective." *Semin Reprod Med* 27 (5):351-7. doi: 10.1055/s-0029-1237423.

Eagar, T. N., N. J. Karandikar, J. A. Bluestone, and S. D. Miller. 2002. "The role of CTLA-4 in induction and maintenance of peripheral T cell tolerance." *Eur J Immunol* 32 (4):972-81. doi: 10.1002/1521-4141(200204)32:4<972::AID-IMMU972>3.0.CO;2-M.

Eich, T., O. Eriksson, T. Lundgren, and Transplantation Nordic Network for Clinical Islet. 2007. "Visualization of early engraftment in clinical islet transplantation by positron-emission tomography." *N Engl J Med* 356 (26):2754-5. doi: 10.1056/NEJMc070201.

Eisenbarth, G. S. 1986. "Type I diabetes mellitus. A chronic autoimmune disease." *N Engl J Med* 314 (21):1360-8. doi: 10.1056/NEJM198605223142106.

Eizirik, D. L., and M. I. Darville. 2001. "beta-cell apoptosis and defense mechanisms: lessons from type 1 diabetes." *Diabetes* 50 Suppl 1:S64-9.

Eizirik, D. L., and T. Mandrup-Poulsen. 2001. "A choice of death--the signal-transduction of immune-mediated beta-cell apoptosis." *Diabetologia* 44 (12):2115-33. doi: 10.1007/s001250100021.

Elmore, S. 2007. "Apoptosis: a review of programmed cell death." *Toxicol Pathol* 35 (4):495-516. doi: 10.1080/01926230701320337.

Elshimali, Y. I., H. Khaddour, M. Sarkissyan, Y. Wu, and J. V. Vadgama. 2013. "The clinical utilization of circulating cell free DNA (CCFDNA) in blood of cancer patients." *Int J Mol Sci* 14 (9):18925-58. doi: 10.3390/ijms140918925.

Farh, K. K., A. Marson, J. Zhu, M. Kleinewietfeld, W. J. Housley, S. Beik, N. Shoresh, H. Whitton, R. J. Ryan, A. A. Shishkin, M. Hatan, M. J. Carrasco-Alfonso, D. Mayer, C. J. Luckey, N. A. Patsopoulos, P. L. De Jager, V. K. Kuchroo, C. B. Epstein, M. J. Daly, D. A. Hafler, and B. E. Bernstein. 2015. "Genetic and epigenetic fine mapping of causal autoimmune disease variants." *Nature* 518 (7539):337-43. doi: 10.1038/nature13835.

Fernandez, A. F., Y. Assenov, J. I. Martin-Subero, B. Balint, R. Siebert, H. Taniguchi, H. Yamamoto, M. Hidalgo, A. C. Tan, O. Galm, I. Ferrer, M. Sanchez-Cespedes, A. Villanueva, J. Carmona, J. V.

Sanchez-Mut, M. Berdasco, V. Moreno, G. Capella, D. Monk, E. Ballestar, S. Ropero, R. Martinez, M. Sanchez-Carbayo, F. Prosper, X. Agirre, M. F. Fraga, O. Grana, L. Perez-Jurado, J. Mora, S. Puig, J. Prat, L. Badimon, A. A. Puca, S. J. Meltzer, T. Lengauer, J. Bridgewater, C. Bock, and M. Esteller. 2012. "A DNA methylation fingerprint of 1628 human samples." *Genome Res* 22 (2):407-19. doi: 10.1101/gr.119867.110.

Feshchenko, E. A., E. V. Smirnova, G. Swaminathan, A. M. Teckchandani, R. Agrawal, H. Band, X. Zhang, R. S. Annan, S. A. Carr, and A. Y. Tsygankov. 2004. "TULA: an SH3- and UBA-containing protein that binds to c-Cbl and ubiquitin." *Oncogene* 23 (27):4690-706. doi: 10.1038/sj.onc.1207627.

Filios, S. R., and A. Shalev. 2015. "beta-Cell MicroRNAs: Small but Powerful." *Diabetes* 64 (11):3631-44. doi: 10.2337/db15-0831.

Filios, S. R., G. Xu, J. Chen, K. Hong, G. Jing, and A. Shalev. 2014. "MicroRNA-200 is induced by thioredoxin-interacting protein and regulates Zeb1 protein signaling and beta cell apoptosis." *J Biol Chem* 289 (52):36275-83. doi: 10.1074/jbc.M114.592360.

Fisher, M. M., C. N. Perez Chumbiauca, K. J. Mather, R. G. Mirmira, and S. A. Tersey. 2013. "Detection of islet beta-cell death in vivo by multiplex PCR analysis of differentially methylated DNA." *Endocrinology* 154 (9):3476-81. doi: 10.1210/en.2013-1223.

Fisher, M. M., R. A. Watkins, J. Blum, C. Evans-Molina, N. Chalasani, L. A. DiMeglio, K. J. Mather, S. A. Tersey, and R. G. Mirmira. 2015. "Elevations in Circulating Methylated and Unmethylated Preproinsulin DNA in New-Onset Type 1 Diabetes." *Diabetes* 64 (11):3867-72. doi: 10.2337/db15-0430.

Fleischhacker, M., and B. Schmidt. 2007. "Circulating nucleic acids (CNAs) and cancer--a survey." *Biochim Biophys Acta* 1775 (1):181-232. doi: 10.1016/j.bbcan.2006.10.001.

Fortin, J. P., E. Fertig, and K. Hansen. 2014. "shinyMethyl: interactive quality control of Illumina 450k DNA methylation arrays in R." *F1000Res* 3:175. doi: 10.12688/f1000research.4680.2.

Foulis, A. K., C. N. Liddle, M. A. Farquharson, J. A. Richmond, and R. S. Weir. 1986. "The histopathology of the pancreas in type 1 (insulin-dependent) diabetes mellitus: a 25-year review of deaths in patients under 20 years of age in the United Kingdom." *Diabetologia* 29 (5):267-74.

Foulis, A. K., M. McGill, and M. A. Farquharson. 1991. "Insulitis in type 1 (insulin-dependent) diabetes mellitus in man--macrophages, lymphocytes, and interferon-gamma containing cells." *J Pathol* 165 (2):97-103. doi: 10.1002/path.1711650203.

Fradin, D., S. Le Fur, C. Mille, N. Naoui, C. Groves, D. Zelenika, M. I. McCarthy, M. Lathrop, and P. Bougneres. 2012. "Association of the CpG methylation pattern of the proximal insulin gene promoter with type 1 diabetes." *PLoS One* 7 (5):e36278. doi: 10.1371/journal.pone.0036278.

Fraga, M. F., E. Ballestar, M. F. Paz, S. Ropero, F. Setien, M. L. Ballestar, D. Heine-Suner, J. C. Cigudosa, M. Urioste, J. Benitez, M. Boix-Chornet, A. Sanchez-Aguilera, C. Ling, E. Carlsson, P. Poulsen, A. Vaag, Z. Stephan, T. D. Spector, Y. Z. Wu, C. Plass, and M. Esteller. 2005. "Epigenetic differences arise during the lifetime of monozygotic twins." *Proc Natl Acad Sci U S A* 102 (30):10604-9. doi: 10.1073/pnas.0500398102.

Franklin, T. B., and I. M. Mansuy. 2011. "The involvement of epigenetic defects in mental retardation." *Neurobiol Learn Mem* 96 (1):61-7. doi: 10.1016/j.nlm.2011.04.001.

Frauer, C., T. Hoffmann, S. Bultmann, V. Casa, M. C. Cardoso, I. Antes, and H. Leonhardt. 2011. "Recognition of 5-hydroxymethylcytosine by the Uhrf1 SRA domain." *PLoS One* 6 (6):e21306. doi: 10.1371/journal.pone.0021306.

Gala-Lopez, B. L., D. Neiman, T. Kin, D. O'Gorman, A. R. Pepper, A. J. Malcolm, S. Pianzin, P. A. Senior, P. Campbell, B. Glaser, Y. Dor, R. Shemer, and A. M. J. Shapiro. 2018. "Beta Cell Death by Cell-Free DNA and Outcome after Clinical Islet Transplantation." *Transplantation*. doi: 10.1097/TP.0000000000002083.

Gao, F., Y. Niu, Y. E. Sun, H. Lu, Y. Chen, S. Li, Y. Kang, Y. Luo, C. Si, J. Yu, C. Li, N. Sun, W. Si, H. Wang, W. Ji, and T. Tan. 2017. "De novo DNA methylation during monkey pre-implantation embryogenesis." *Cell Res* 27 (4):526-539. doi: 10.1038/cr.2017.25.

Ge, Y., T. K. Paisie, J. R. B. Newman, L. M. McIntyre, and P. Concannon. 2017. "UBASH3A Mediates Risk for Type 1 Diabetes Through Inhibition of T-Cell Receptor-Induced NF-kappaB Signaling." *Diabetes* 66 (7):2033-2043. doi: 10.2337/db16-1023.

Gepts, W. 1965. "Pathologic anatomy of the pancreas in juvenile diabetes mellitus." *Diabetes* 14 (10):619-33.

Giannopoulou, E. Z., C. Winkler, R. Chmiel, C. Matzke, M. Scholz, A. Beyerlein, P. Achenbach, E. Bonifacio, and A. G. Ziegler. 2015. "Islet autoantibody phenotypes and incidence in children at increased risk for type 1 diabetes." *Diabetologia* 58 (10):2317-23. doi: 10.1007/s00125-015-3672-y.

Giulietti, A., C. Gysemans, K. Stoffels, E. van Etten, B. Decallonne, L. Overbergh, R. Bouillon, and C. Mathieu. 2004. "Vitamin D deficiency in early life accelerates Type 1 diabetes in non-obese diabetic mice." *Diabetologia* 47 (3):451-462. doi: 10.1007/s00125-004-1329-3.

Goll, M. G., F. Kirpekar, K. A. Maggert, J. A. Yoder, C. L. Hsieh, X. Zhang, K. G. Golic, S. E. Jacobsen, and T. H. Bestor. 2006. "Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2." *Science* 311 (5759):395-8. doi: 10.1126/science.1120976.

Gowher, H., K. Liebert, A. Hermann, G. Xu, and A. Jeltsch. 2005. "Mechanism of stimulation of catalytic activity of Dnmt3A and Dnmt3B DNA-(cytosine-C5)-methyltransferases by Dnmt3L." *J Biol Chem* 280 (14):13341-8. doi: 10.1074/jbc.M413412200.

Graham, J., W. A. Hagopian, I. Kockum, L. S. Li, C. B. Sanjeevi, R. M. Lowe, J. B. Schaefer, M. Zarghami, H. L. Day, M. Landin-Olsson, J. P. Palmer, M. Janer-Villanueva, L. Hood, G. Sundkvist, A. Lernmark, N. Breslow, G. Dahlquist, G. Blohme, Group Diabetes Incidence in Sweden Study, and Group Swedish Childhood Diabetes Study. 2002. "Genetic effects on age-dependent onset and islet cell autoantibody markers in type 1 diabetes." *Diabetes* 51 (5):1346-55.

Graham, J., I. Kockum, C. B. Sanjeevi, M. Landin-Olsson, L. Nystrom, G. Sundkvist, H. Arnqvist, G. Blohme, F. Lithner, B. Littorin, B. Schersten, L. Wibell, J. Ostman, A. Lernmark, N. Breslow, and G. Dahlquist. 1999. "Negative association between type 1 diabetes and HLA DQB1*0602-DQA1*0102 is attenuated with age at onset. Swedish Childhood Diabetes Study Group." *Eur J Immunogenet* 26 (2-3):117-27.

Grant, S. F., H. Q. Qu, J. P. Bradfield, L. Marchand, C. E. Kim, J. T. Glessner, R. Grabs, S. P. Taback, E. C. Frackelton, A. W. Eckert, K. Annaiah, M. L. Lawson, F. G. Otieno, E. Santa, J. L. Shaner, R. M. Smith, R. Skraban, M. Imielinski, R. M. Chiavacci, R. W. Grundmeier, C. A. Stanley, S. E. Kirsch, D. Waggott, A. D. Paterson, D. S. Monos, Dcct Edic Research Group, C. Polychronakos, and H. Hakonarson. 2009. "Follow-up analysis of genome-wide association data identifies novel loci for type 1 diabetes." *Diabetes* 58 (1):290-5. doi: 10.2337/db08-1022.

Group, The EURODIAB Substudy 2 Study. 1999. "Vitamin D supplement in early childhood and risk for Type I (insulin-dependent) diabetes mellitus. The EURODIAB Substudy 2 Study Group." *Diabetologia* 42 (1):51-4.

Groux, H., A. O'Garra, M. Bigler, M. Rouleau, S. Antonenko, J. E. de Vries, and M. G. Roncarolo. 1997. "A CD4+ T-cell subset inhibits antigen-specific T-cell responses and prevents colitis." *Nature* 389 (6652):737-42. doi: 10.1038/39614.

Gu, T. P., F. Guo, H. Yang, H. P. Wu, G. F. Xu, W. Liu, Z. G. Xie, L. Shi, X. He, S. G. Jin, K. Iqbal, Y. G. Shi, Z. Deng, P. E. Szabo, G. P. Pfeifer, J. Li, and G. L. Xu. 2011. "The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes." *Nature* 477 (7366):606-10. doi: 10.1038/nature10443.

Guo, F., X. Li, D. Liang, T. Li, P. Zhu, H. Guo, X. Wu, L. Wen, T. P. Gu, B. Hu, C. P. Walsh, J. Li, F. Tang, and G. L. Xu. 2014. "Active and passive demethylation of male and female pronuclear DNA in the mammalian zygote." *Cell Stem Cell* 15 (4):447-459. doi: 10.1016/j.stem.2014.08.003.

Gurgul-Convey, E., M. T. Kaminski, and S. Lenzen. 2015. "Physiological characterization of the human EndoC-betaH1 beta-cell line." *Biochem Biophys Res Commun* 464 (1):13-9. doi: 10.1016/j.bbrc.2015.05.072.

Halban, P. A. 1991. "Structural domains and molecular lifestyles of insulin and its precursors in the pancreatic beta cell." *Diabetologia* 34 (11):767-78.

Han, X., J. Wang, and Y. Sun. 2017. "Circulating Tumor DNA as Biomarkers for Cancer Detection." *Genomics Proteomics Bioinformatics* 15 (2):59-72. doi: 10.1016/j.gpb.2016.12.004.

Handy, D. E., R. Castro, and J. Loscalzo. 2011. "Epigenetic modifications: basic mechanisms and role in cardiovascular disease." *Circulation* 123 (19):2145-56. doi: 10.1161/circulationaha.110.956839.

Harper, K., C. Balzano, E. Rouvier, M. G. Mattei, M. F. Luciani, and P. Golstein. 1991. "CTLA-4 and CD28 activated lymphocyte molecules are closely related in both mouse and human as to sequence, message expression, gene structure, and chromosomal location." *J Immunol* 147 (3):1037-44.

Hashimoto, H., Y. Liu, A. K. Upadhyay, Y. Chang, S. B. Howerton, P. M. Vertino, X. Zhang, and X. Cheng. 2012. "Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation." *Nucleic Acids Res* 40 (11):4841-9. doi: 10.1093/nar/gks155.

He, Y. F., B. Z. Li, Z. Li, P. Liu, Y. Wang, Q. Tang, J. Ding, Y. Jia, Z. Chen, L. Li, Y. Sun, X. Li, Q. Dai, C. X. Song, K. Zhang, C. He, and G. L. Xu. 2011. "Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA." *Science* 333 (6047):1303-7. doi: 10.1126/science.1210944.

Heard, E., and C. M. Disteche. 2006. "Dosage compensation in mammals: fine-tuning the expression of the X chromosome." *Genes Dev* 20 (14):1848-67. doi: 10.1101/gad.1422906.

Heinig, M., E. Petretto, C. Wallace, L. Bottolo, M. Rotival, H. Lu, Y. Li, R. Sarwar, S. R. Langley, A. Bauerfeind, O. Hummel, Y. A. Lee, S. Paskas, C. Rintisch, K. Saar, J. Cooper, R. Buchan, E. E. Gray, J. G. Cyster, Consortium Cardiogenics, J. Erdmann, C. Hengstenberg, S. Maouche, W. H. Ouwehand, C. M. Rice, N. J. Samani, H. Schunkert, A. H. Goodall, H. Schulz, H. G. Roider, M. Vingron, S. Blankenberg, T. Munzel, T. Zeller, S. Szymczak, A. Ziegler, L. Tiret, D. J. Smyth, M. Pravenec, T. J. Aitman, F. Cambien, D. Clayton, J. A. Todd, N. Hubner, and S. A. Cook. 2010. "A trans-acting locus regulates an anti-viral expression network and type 1 diabetes risk." *Nature* 467 (7314):460-4. doi: 10.1038/nature09386.

Hellman, B. 1959a. "Actual distribution of the number and volume of the islets of Langerhans in different size classes in non-diabetic humans of varying ages." *Nature* 184(Suppl 19):1498-9.

Hellman, B. 1959b. "The frequency distribution of the number and volume of the islets Langerhans in man. I. Studies on non-diabetic adults." *Acta Soc Med Ups* 64:432-60.

Herold, K. C., B. Brooks-Worrell, J. Palmer, H. M. Dosch, M. Peakman, P. Gottlieb, H. Reijonen, S. Arif, L. M. Spain, C. Thompson, J. M. Lachin, and Group Type 1 Diabetes TrialNet Research. 2009. "Validity and reproducibility of measurement of islet autoreactivity by T-cell assays in subjects with early type 1 diabetes." *Diabetes* 58 (11):2588-95. doi: 10.2337/db09-0249.

Herold, K. C., W. Hagopian, J. A. Auger, E. Poumian-Ruiz, L. Taylor, D. Donaldson, S. E. Gitelman, D. M. Harlan, D. Xu, R. A. Zivin, and J. A. Bluestone. 2002. "Anti-CD3 monoclonal antibody in new-onset type 1 diabetes mellitus." *N Engl J Med* 346 (22):1692-8. doi: 10.1056/NEJMoa012864.

Herold, K. C., S. Usmani-Brown, T. Ghazi, J. Lebastchi, C. A. Beam, M. D. Bellin, M. Ledizet, J. M. Sosenko, J. P. Krischer, and J. P. Palmer. 2015a. "beta cell death and dysfunction during type 1 diabetes development in at-risk individuals." *J Clin Invest* 125 (3):1163-73. doi: 10.1172/jci78142.

Herold, K. C., S. Usmani-Brown, T. Ghazi, J. Lebastchi, C. A. Beam, M. D. Bellin, M. Ledizet, J. M. Sosenko, J. P. Krischer, J. P. Palmer, and Group Type 1 Diabetes TrialNet Study. 2015b. "beta cell death and dysfunction during type 1 diabetes development in at-risk individuals." *J Clin Invest* 125 (3):1163-73. doi: 10.1172/JCI78142.

Heuts, F., N. M. Edner, and L. S. Walker. 2017. "Follicular T Helper Cells: A New Marker of Type 1 Diabetes Risk?" *Diabetes* 66 (2):258-260. doi: 10.2337/dbi16-0062.

Heyn, H., and M. Esteller. 2012. "DNA methylation profiling in the clinic: applications and challenges." *Nat Rev Genet* 13 (10):679-92. doi: 10.1038/nrg3270.

Hindson, B. J., K. D. Ness, D. A. Masquelier, P. Belgrader, N. J. Heredia, A. J. Makarewicz, I. J. Bright, M. Y. Lucero, A. L. Hiddessen, T. C. Legler, T. K. Kitano, M. R. Hodel, J. F. Petersen, P. W. Wyatt, E. R. Steenblock, P. H. Shah, L. J. Bousse, C. B. Troup, J. C. Mellen, D. K. Wittmann, N. G. Erndt, T. H. Cauley, R. T. Koehler, A. P. So, S. Dube, K. A. Rose, L. Montesclaros, S. Wang, D. P. Stumbo, S. P. Hodges, S. Romine, F. P. Milanovich, H. E. White, J. F. Regan, G. A. Karlin-Neumann, C. M. Hindson, S. Saxonov, and B. W. Colston. 2011. "High-throughput droplet digital PCR system for absolute quantitation of DNA copy number." *Anal Chem* 83 (22):8604-10. doi: 10.1021/ac202028g.

Hjern, A., U. Soderstrom, and J. Aman. 2012. "East Africans in Sweden have a high risk for type 1 diabetes." *Diabetes Care* 35 (3):597-8. doi: 10.2337/dc11-1536.

Ho, B. C., P. C. Yang, and S. L. Yu. 2016. "MicroRNA and Pathogenesis of Enterovirus Infection." *Viruses* 8 (1). doi: 10.3390/v8010011.

Holland, A. M., S. Garcia, G. Naselli, R. J. Macdonald, and L. C. Harrison. 2013. "The Parahox gene Pdx1 is required to maintain positional identity in the adult foregut." *Int J Dev Biol* 57 (5):391-8. doi: 10.1387/ijdb.120048ah.

Holling, T. M., N. van der Stoep, E. Quinten, and P. J. van den Elsen. 2002. "Activated human T cells accomplish MHC class II expression through T cell-specific occupation of class II transactivator promoter III." *J Immunol* 168 (2):763-70.

Honkanen, H., S. Oikarinen, N. Nurminen, O. H. Laitinen, H. Huhtala, J. Lehtonen, T. Ruokoranta, M. M. Hankaniemi, V. Lecouturier, J. W. Almond, S. Tauriainen, O. Simell, J. Ilonen, R. Veijola, H. Viskari, M. Knip, and H. Hyoty. 2017. "Detection of enteroviruses in stools precedes islet autoimmunity by several months: possible evidence for slowly operating mechanisms in virus-induced autoimmunity." *Diabetologia* 60 (3):424-431. doi: 10.1007/s00125-016-4177-z.

Howson, J. M., N. M. Walker, D. Clayton, J. A. Todd, and Consortium Type 1 Diabetes Genetics. 2009. "Confirmation of HLA class II independent type 1 diabetes associations in the major histocompatibility complex including HLA-B and HLA-A." *Diabetes Obes Metab* 11 Suppl 1:31-45. doi: 10.1111/j.1463-1326.2008.01001.x.

Huertas, D., R. Sendra, and P. Munoz. 2009. "Chromatin dynamics coupled to DNA repair." *Epigenetics* 4 (1):31-42.

Hui, H., and R. Perfetti. 2002. "Pancreas duodenum homeobox-1 regulates pancreas development during embryogenesis and islet cell function in adulthood." *Eur J Endocrinol* 146 (2):129-41.

Husseiny, M. I., A. Kaye, E. Zebadua, F. Kandeel, and K. Ferreri. 2014. "Tissue-specific methylation of human insulin gene and PCR assay for monitoring beta cell death." *PLoS One* 9 (4):e94591. doi: 10.1371/journal.pone.0094591.

Husseiny, M. I., A. Kuroda, A. N. Kaye, I. Nair, F. Kandeel, and K. Ferreri. 2012. "Development of a quantitative methylation-specific polymerase chain reaction method for monitoring beta cell death in type 1 diabetes." *PLoS One* 7 (10):e47942. doi: 10.1371/journal.pone.0047942.

Hussen, H. I., M. Persson, and T. Moradi. 2013. "The trends and the risk of type 1 diabetes over the past 40 years: an analysis by birth cohorts and by parental migration background in Sweden." *BMJ Open* 3 (10):e003418. doi: 10.1136/bmjopen-2013-003418.

Hyoty, H., and K. W. Taylor. 2002. "The role of viruses in human diabetes." *Diabetologia* 45 (10):1353-61. doi: 10.1007/s00125-002-0852-3.

Hypponen, E., E. Laara, A. Reunanen, M. R. Jarvelin, and S. M. Virtanen. 2001. "Intake of vitamin D and risk of type 1 diabetes: a birth-cohort study." *Lancet* 358 (9292):1500-3. doi: 10.1016/S0140-6736(01)06580-1.

Hyttinen, V., J. Kaprio, L. Kinnunen, M. Koskenvuo, and J. Tuomilehto. 2003. "Genetic liability of type 1 diabetes and the onset age among 22,650 young Finnish twin pairs: a nationwide follow-up study." *Diabetes* 52 (4):1052-5.

IDF. 2017. "IDF Diabetes Atlas." *International Diabetes Federation* 8.

Ikegami, H., T. Fujisawa, Y. Kawabata, S. Noso, and T. Ogihara. 2006. "Genetics of type 1 diabetes: similarities and differences between Asian and Caucasian populations." *Ann N Y Acad Sci* 1079:51-9. doi: 10.1196/annals.1375.008.

Imagawa, A., T. Hanafusa, S. Tamura, M. Moriwaki, N. Itoh, K. Yamamoto, H. Iwahashi, K. Yamagata, M. Waguri, T. Nanmo, S. Uno, H. Nakajima, M. Namba, S. Kawata, J. I. Miyagawa, and Y. Matsuzawa. 2001. "Pancreatic biopsy as a procedure for detecting in situ autoimmune phenomena in type 1 diabetes: close correlation between serological markers and histological evidence of cellular autoimmunity." *Diabetes* 50 (6):1269-73.

In't Veld, P. 2014. "Insulitis in human type 1 diabetes: a comparison between patients and animal models." *Semin Immunopathol* 36 (5):569-79. doi: 10.1007/s00281-014-0438-4.

Inoue, A., L. Shen, Q. Dai, C. He, and Y. Zhang. 2011. "Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development." *Cell Res* 21 (12):1670-6. doi: 10.1038/cr.2011.189.

Inoue, A., and Y. Zhang. 2011. "Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos." *Science* 334 (6053):194. doi: 10.1126/science.1212483.

Insel, R. A., J. L. Dunne, M. A. Atkinson, J. L. Chiang, D. Dabelea, P. A. Gottlieb, C. J. Greenbaum, K. C. Herold, J. P. Krischer, A. Lernmark, R. E. Ratner, M. J. Rewers, D. A. Schatz, J. S. Skyler, J. M. Sosenko, and A. G. Ziegler. 2015. "Staging presymptomatic type 1 diabetes: a scientific statement of JDRF, the Endocrine Society, and the American Diabetes Association." *Diabetes Care* 38 (10):1964-74. doi: 10.2337/dc15-1419.

Ionescu-Tirgoviste, C., P. A. Gagniuc, E. Gubceac, L. Mardare, I. Popescu, S. Dima, and M. Militaru. 2015. "A 3D map of the islet routes throughout the healthy human pancreas." *Sci Rep* 5:14634. doi: 10.1038/srep14634.

Ito, S., L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He, and Y. Zhang. 2011. "Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine." *Science* 333 (6047):1300-3. doi: 10.1126/science.1210597.

Itoh, N., T. Hanafusa, A. Miyazaki, J. Miyagawa, K. Yamagata, K. Yamamoto, M. Waguri, A. Imagawa, S. Tamura, M. Inada, and et al. 1993. "Mononuclear cell infiltration and its relation to the expression of major histocompatibility complex antigens and adhesion molecules in pancreas biopsy specimens from newly diagnosed insulin-dependent diabetes mellitus patients." *J Clin Invest* 92 (5):2313-22. doi: 10.1172/JCI116835.

Jacobsen, R., E. Hypponen, T. I. Sorensen, A. A. Vaag, and B. L. Heitmann. 2015. "Gestational and Early Infancy Exposure to Margarine Fortified with Vitamin D through a National Danish Programme

and the Risk of Type 1 Diabetes: The D-Tect Study." *PLoS One* 10 (6):e0128631. doi: 10.1371/journal.pone.0128631.

Jansson, M. D., and A. H. Lund. 2012. "MicroRNA and cancer." *Mol Oncol* 6 (6):590-610. doi: 10.1016/j.molonc.2012.09.006.

Jenkinson, E. J., R. Kingston, and J. J. Owen. 1987. "Importance of IL-2 receptors in intra-thymic generation of cells expressing T-cell receptors." *Nature* 329 (6135):160-2. doi: 10.1038/329160a0.

Ji, D., K. Lin, J. Song, and Y. Wang. 2014. "Effects of Tet-induced oxidation products of 5-methylcytosine on Dnmt1- and DNMT3a-mediated cytosine methylation." *Mol Biosyst* 10 (7):1749-52. doi: 10.1039/c4mb00150h.

Jia, D., R. Z. Jurkowska, X. Zhang, A. Jeltsch, and X. Cheng. 2007. "Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation." *Nature* 449 (7159):248-51. doi: 10.1038/nature06146.

Jiang, P., and Y. M. D. Lo. 2016. "The Long and Short of Circulating Cell-Free DNA and the Ins and Outs of Molecular Diagnostics." *Trends Genet* 32 (6):360-371. doi: 10.1016/j.tig.2016.03.009.

Joglekar, M. V., V. M. Joglekar, and A. A. Hardikar. 2009. "Expression of islet-specific microRNAs during human pancreatic development." *Gene Expr Patterns* 9 (2):109-13. doi: 10.1016/j.gep.2008.10.001.

Johnson, J. D., N. T. Ahmed, D. S. Luciani, Z. Han, H. Tran, J. Fujita, S. Misler, H. Edlund, and K. S. Polonsky. 2003. "Increased islet apoptosis in Pdx1+/- mice." *J Clin Invest* 111 (8):1147-60. doi: 10.1172/jci16537.

Johnson, J. D., E. Bernal-Mizrachi, E. U. Alejandro, Z. Han, T. B. Kalynyak, H. Li, J. L. Beith, J. Gross, G. L. Warnock, R. R. Townsend, M. A. Permutt, and K. S. Polonsky. 2006. "Insulin protects islets from apoptosis via Pdx1 and specific changes in the human islet proteome." *Proc Natl Acad Sci U S A* 103 (51):19575-80. doi: 10.1073/pnas.0604208103.

Johnson, K., R. Wong, K. J. Barriga, G. Klingensmith, A. G. Ziegler, M. J. Rewers, and A. K. Steck. 2012. "rs11203203 is associated with type 1 diabetes risk in population pre-screened for high-risk HLA-DR,DQ genotypes." *Pediatr Diabetes* 13 (8):611-5. doi: 10.1111/j.1399-5448.2012.00888.x.

Jones, A. G., and A. T. Hattersley. 2013. "The clinical utility of C-peptide measurement in the care of patients with diabetes." *Diabet Med* 30 (7):803-17. doi: 10.1111/dme.12159.

Jones, P. A., and S. B. Baylin. 2002. "The fundamental role of epigenetic events in cancer." *Nat Rev Genet* 3 (6):415-28. doi: 10.1038/nrg816.

Jones, P. A., and G. Liang. 2009. "Rethinking how DNA methylation patterns are maintained." *Nat Rev Genet* 10 (11):805-11. doi: 10.1038/nrg2651.

Jonsson, J., L. Carlsson, T. Edlund, and H. Edlund. 1994. "Insulin-promoter-factor 1 is required for pancreas development in mice." *Nature* 371 (6498):606-9. doi: 10.1038/371606a0.

Jorns, A., T. Arndt, A. Meyer zu Vilsendorf, J. Klempnauer, D. Wedekind, H. J. Hedrich, L. Marselli, P. Marchetti, N. Harada, Y. Nakaya, G. S. Wang, F. W. Scott, C. Gysemans, C. Mathieu, and S. Lenzen. 2014. "Islet infiltration, cytokine expression and beta cell death in the NOD mouse, BB rat, Komeda rat, LEW.1AR1-iddm rat and humans with type 1 diabetes." *Diabetologia* 57 (3):512-21. doi: 10.1007/s00125-013-3125-4.

Kaech, S. M., S. Hemby, E. Kersh, and R. Ahmed. 2002. "Molecular and functional profiling of memory CD8 T cell differentiation." *Cell* 111 (6):837-51.

Kambayashi, T., and T. M. Laufer. 2014. "Atypical MHC class II-expressing antigen-presenting cells: can anything replace a dendritic cell?" *Nat Rev Immunol* 14 (11):719-30. doi: 10.1038/nri3754.

Kanak, M. A., M. Takita, F. Kunnathodi, M. C. Lawrence, M. F. Levy, and B. Naziruddin. 2014. "Inflammatory response in islet transplantation." *Int J Endocrinol* 2014:451035. doi: 10.1155/2014/451035.

Kaneda, M., M. Okano, K. Hata, T. Sado, N. Tsujimoto, E. Li, and H. Sasaki. 2004. "Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting." *Nature* 429 (6994):900-3. doi: 10.1038/nature02633.

Kappler, J. W., N. Roehm, and P. Marrack. 1987. "T cell tolerance by clonal elimination in the thymus." *Cell* 49 (2):273-80.

Kaprio, J., J. Tuomilehto, M. Koskenvuo, K. Romanov, A. Reunanen, J. Eriksson, J. Stengard, and Y. A. Kesaniemi. 1992. "Concordance for type 1 (insulin-dependent) and type 2 (non-insulin-dependent) diabetes mellitus in a population-based cohort of twins in Finland." *Diabetologia* 35 (11):1060-7.

Karlic, R., H. R. Chung, J. Lasserre, K. Vlahovicek, and M. Vingron. 2010. "Histone modification levels are predictive for gene expression." *Proc Natl Acad Sci U S A* 107 (7):2926-31. doi: 10.1073/pnas.0909344107.

Kaslow, D. C., and B. R. Migeon. 1987. "DNA methylation stabilizes X chromosome inactivation in eutherians but not in marsupials: evidence for multistep maintenance of mammalian X dosage compensation." *Proc Natl Acad Sci U S A* 84 (17):6210-4.

Kato, Y., M. Kaneda, K. Hata, K. Kumaki, M. Hisano, Y. Kohara, M. Okano, E. Li, M. Nozaki, and H. Sasaki. 2007. "Role of the Dnmt3 family in de novo methylation of imprinted and repetitive sequences

during male germ cell development in the mouse." *Hum Mol Genet* 16 (19):2272-80. doi: 10.1093/hmg/ddm179.

Katz, J. D., C. Benoist, and D. Mathis. 1995. "T helper cell subsets in insulin-dependent diabetes." *Science* 268 (5214):1185-8.

Kawa, S., M. Ota, K. Yoshizawa, A. Horiuchi, H. Hamano, Y. Ochi, K. Nakayama, Y. Tokutake, Y. Katsuyama, S. Saito, O. Hasebe, and K. Kiyosawa. 2002. "HLA DRB10405-DQB10401 haplotype is associated with autoimmune pancreatitis in the Japanese population." *Gastroenterology* 122 (5):1264-9.

Kawabata, Y., H. Ikegami, Y. Kawaguchi, T. Fujisawa, M. Shintani, M. Ono, M. Nishino, Y. Uchigata, I. Lee, and T. Ogihara. 2002. "Asian-specific HLA haplotypes reveal heterogeneity of the contribution of HLA-DR and -DQ haplotypes to susceptibility to type 1 diabetes." *Diabetes* 51 (2):545-51.

Kenefeck, R., C. J. Wang, T. Kapadi, L. Wardzinski, K. Attridge, L. E. Clough, F. Heuts, A. Kogimtzis, S. Patel, M. Rosenthal, M. Ono, D. M. Sansom, P. Narendran, and L. S. Walker. 2015. "Follicular helper T cell signature in type 1 diabetes." *J Clin Invest* 125 (1):292-303. doi: 10.1172/JCI76238.

Keymeulen, B., E. Vandemeulebroucke, A. G. Ziegler, C. Mathieu, L. Kaufman, G. Hale, F. Gorus, M. Goldman, M. Walter, S. Candon, L. Schandene, L. Crenier, C. De Block, J. M. Seigneurin, P. De Pauw, D. Pierard, I. Weets, P. Rebello, P. Bird, E. Berrie, M. Frewin, H. Waldmann, J. F. Bach, D. Pipeleers, and L. Chatenoud. 2005. "Insulin needs after CD3-antibody therapy in new-onset type 1 diabetes." *N Engl J Med* 352 (25):2598-608. doi: 10.1056/NEJMoa043980.

Khier, S., and L. Lohan. 2018. "Kinetics of circulating cell-free DNA for biomedical applications: critical appraisal of the literature." *Future Sci OA* 4 (4):FSO295. doi: 10.4155/fsoa-2017-0140.

Killestein, J. 2002. "Anti-CD3 monoclonal antibody in new-onset type 1 diabetes mellitus." *N Engl J Med* 347 (14):1116-7; author reply 1116-7. doi: 10.1056/NEJM200210033471416.

Kim, V. N., J. Han, and M. C. Siomi. 2009. "Biogenesis of small RNAs in animals." *Nat Rev Mol Cell Biol* 10 (2):126-39. doi: 10.1038/nrm2632.

King, C., and N. Sarvetnick. 2011. "The incidence of type-1 diabetes in NOD mice is modulated by restricted flora not germ-free conditions." *PLoS One* 6 (2):e17049. doi: 10.1371/journal.pone.0017049.

Klein, D., R. Misawa, V. Bravo-Egana, N. Vargas, S. Rosero, J. Piroso, H. Ichii, O. Umland, J. Zhijie, N. Tsinoremas, C. Ricordi, L. Inverardi, J. Dominguez-Bendala, and R. L. Pastori. 2013. "MicroRNA expression in alpha and beta cells of human pancreatic islets." *PLoS One* 8 (1):e55064. doi: 10.1371/journal.pone.0055064.

Klein, L., B. Kyewski, P. M. Allen, and K. A. Hogquist. 2014. "Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see)." *Nat Rev Immunol* 14 (6):377-91. doi: 10.1038/nri3667.

Knip, M., and J. Honkanen. 2017. "Modulation of Type 1 Diabetes Risk by the Intestinal Microbiome." *Curr Diab Rep* 17 (11):105. doi: 10.1007/s11892-017-0933-9.

Knip, M., R. Veijola, S. M. Virtanen, H. Hyoty, O. Vaarala, and H. K. Akerblom. 2005. "Environmental triggers and determinants of type 1 diabetes." *Diabetes* 54 Suppl 2:S125-36.

Kondrashova, A., A. Reunanen, A. Romanov, A. Karvonen, H. Viskari, T. Vesikari, J. Ilonen, M. Knip, and H. Hyoty. 2005. "A six-fold gradient in the incidence of type 1 diabetes at the eastern border of Finland." *Ann Med* 37 (1):67-72.

Kosaka, N., H. Iguchi, Y. Yoshioka, F. Takeshita, Y. Matsuki, and T. Ochiya. 2010. "Secretory mechanisms and intercellular transfer of microRNAs in living cells." *J Biol Chem* 285 (23):17442-52. doi: 10.1074/jbc.M110.107821.

Kouzarides, T. 2007. "Chromatin modifications and their function." *Cell* 128 (4):693-705. doi: 10.1016/j.cell.2007.02.005.

Koya, V., S. Lu, Y. P. Sun, D. L. Purich, M. A. Atkinson, S. W. Li, and L. J. Yang. 2008. "Reversal of streptozotocin-induced diabetes in mice by cellular transduction with recombinant pancreatic transcription factor pancreatic duodenal homeobox-1: a novel protein transduction domain-based therapy." *Diabetes* 57 (3):757-69. doi: 10.2337/db07-1441.

Krischer, J. P., K. F. Lynch, D. A. Schatz, J. Ilonen, A. Lernmark, W. A. Hagopian, M. J. Rewers, J. X. She, O. G. Simell, J. Toppari, A. G. Ziegler, B. Akolkar, E. Bonifacio, and Teddy Study Group. 2015. "The 6 year incidence of diabetes-associated autoantibodies in genetically at-risk children: the TEDDY study." *Diabetologia* 58 (5):980-7. doi: 10.1007/s00125-015-3514-y.

Kristiansen, O. P., Z. M. Larsen, and F. Pociot. 2000. "CTLA-4 in autoimmune diseases--a general susceptibility gene to autoimmunity?" *Genes Immun* 1 (3):170-84. doi: 10.1038/sj.gene.6363655.

Krogvold, L., B. Edwin, T. Buanes, G. Frisk, O. Skog, M. Anagandula, O. Korsgren, D. Undlien, M. C. Eike, S. J. Richardson, P. Leete, N. G. Morgan, S. Oikarinen, M. Oikarinen, J. E. Laiho, H. Hyoty, J. Ludvigsson, K. F. Hanssen, and K. Dahl-Jorgensen. 2015. "Detection of a low-grade enteroviral infection in the islets of langerhans of living patients newly diagnosed with type 1 diabetes." *Diabetes* 64 (5):1682-7. doi: 10.2337/db14-1370.

Krogvold, L., B. Edwin, T. Buanes, J. Ludvigsson, O. Korsgren, H. Hyoty, G. Frisk, K. F. Hanssen, and K. Dahl-Jorgensen. 2014. "Pancreatic biopsy by minimal tail resection in live adult patients at the

onset of type 1 diabetes: experiences from the DiViD study." *Diabetologia* 57 (4):841-3. doi: 10.1007/s00125-013-3155-y.

Kronenberg, M., and A. Rudensky. 2005. "Regulation of immunity by self-reactive T cells." *Nature* 435 (7042):598-604. doi: 10.1038/nature03725.

Kundaje, A., W. Meuleman, J. Ernst, M. Bilenky, A. Yen, A. Heravi-Moussavi, P. Kheradpour, Z. Zhang, J. Wang, M. J. Ziller, V. Amin, J. W. Whitaker, M. D. Schultz, L. D. Ward, A. Sarkar, G. Quon, R. S. Sandstrom, M. L. Eaton, Y. C. Wu, A. R. Pfenning, X. Wang, M. Claussnitzer, Y. Liu, C. Coarfa, R. A. Harris, N. Shoresh, C. B. Epstein, E. Gjoneska, D. Leung, W. Xie, R. D. Hawkins, R. Lister, C. Hong, P. Gascard, A. J. Mungall, R. Moore, E. Chuah, A. Tam, T. K. Canfield, R. S. Hansen, R. Kaul, P. J. Sabo, M. S. Bansal, A. Carles, J. R. Dixon, K. H. Farh, S. Feizi, R. Karlic, A. R. Kim, A. Kulkarni, D. Li, R. Lowdon, G. Elliott, T. R. Mercer, S. J. Neph, V. Onuchic, P. Polak, N. Rajagopal, P. Ray, R. C. Sallari, K. T. Siebenthall, N. A. Sinnott-Armstrong, M. Stevens, R. E. Thurman, J. Wu, B. Zhang, X. Zhou, A. E. Beaudet, L. A. Boyer, P. L. De Jager, P. J. Farnham, S. J. Fisher, D. Haussler, S. J. Jones, W. Li, M. A. Marra, M. T. McManus, S. Sunyaev, J. A. Thomson, T. D. Tlsty, L. H. Tsai, W. Wang, R. A. Waterland, M. Q. Zhang, L. H. Chadwick, B. E. Bernstein, J. F. Costello, J. R. Ecker, M. Hirst, A. Meissner, A. Milosavljevic, B. Ren, J. A. Stamatoyannopoulos, T. Wang, and M. Kellis. 2015. "Integrative analysis of 111 reference human epigenomes." *Nature* 518 (7539):317-30. doi: 10.1038/nature14248.

Kurdistani, S. K., S. Tavazoie, and M. Grunstein. 2004. "Mapping global histone acetylation patterns to gene expression." *Cell* 117 (6):721-33. doi: 10.1016/j.cell.2004.05.023.

Kuroda, A., T. A. Rauch, I. Todorov, H. T. Ku, I. H. Al-Abdullah, F. Kandeel, Y. Mullen, G. P. Pfeifer, and K. Ferreri. 2009. "Insulin gene expression is regulated by DNA methylation." *PLoS One* 4 (9):e6953. doi: 10.1371/journal.pone.0006953.

Kurrer, M. O., S. V. Pakala, H. L. Hanson, and J. D. Katz. 1997. "Beta cell apoptosis in T cell-mediated autoimmune diabetes." *Proc Natl Acad Sci U S A* 94 (1):213-8.

Kyvik, K. O., A. Green, and H. Beck-Nielsen. 1995. "Concordance rates of insulin dependent diabetes mellitus: a population based study of young Danish twins." *BMJ* 311 (7010):913-7.

LaBaer, J. 2005. "So, you want to look for biomarkers (introduction to the special biomarkers issue)." *J Proteome Res* 4 (4):1053-9. doi: 10.1021/pr0501259.

Lahmy, R., M. Soleimani, M. H. Sanati, M. Behmanesh, F. Kouhkan, and N. Mobarra. 2014. "MiRNA-375 promotes beta pancreatic differentiation in human induced pluripotent stem (hiPS) cells." *Mol Biol Rep* 41 (4):2055-66. doi: 10.1007/s11033-014-3054-4.

Lambert, A. P., K. M. Gillespie, G. Thomson, H. J. Cordell, J. A. Todd, E. A. Gale, and P. J. Bingley. 2004. "Absolute risk of childhood-onset type 1 diabetes defined by human leukocyte antigen class II

genotype: a population-based study in the United Kingdom." *J Clin Endocrinol Metab* 89 (8):4037-43. doi: 10.1210/jc.2003-032084.

Lampasona, V., and D. Liberati. 2016. "Islet Autoantibodies." *Curr Diab Rep* 16 (6):53. doi: 10.1007/s11892-016-0738-2.

Lebastchi, J., S. Deng, A. H. Lebastchi, I. Beshar, S. Gitelman, S. Willi, P. Gottlieb, E. M. Akirav, J. A. Bluestone, and K. C. Herold. 2013. "Immune therapy and beta-cell death in type 1 diabetes." *Diabetes* 62 (5):1676-80. doi: 10.2337/db12-1207.

Lebastchi, J., and K. C. Herold. 2012. "Immunologic and metabolic biomarkers of beta-cell destruction in the diagnosis of type 1 diabetes." *Cold Spring Harb Perspect Med* 2 (6):a007708. doi: 10.1101/cshperspect.a007708.

Lecamwasam, A., A. Sexton-Oates, J. Carmody, E. I. Ekinci, K. M. Dwyer, and R. Saffery. 2018. "DNA methylation profiling of genomic DNA isolated from urine in diabetic chronic kidney disease: A pilot study." *PLoS One* 13 (2):e0190280. doi: 10.1371/journal.pone.0190280.

Lee, H. S., T. Briese, C. Winkler, M. Rewers, E. Bonifacio, H. Hyoty, M. Pflueger, O. Simell, J. X. She, W. Hagopian, A. Lernmark, B. Akolkar, J. Krischer, A. G. Ziegler, and Teddy study group. 2013. "Next-generation sequencing for viruses in children with rapid-onset type 1 diabetes." *Diabetologia* 56 (8):1705-1711. doi: 10.1007/s00125-013-2924-y.

Lee, Y., M. Kim, J. Han, K. H. Yeom, S. Lee, S. H. Baek, and V. N. Kim. 2004. "MicroRNA genes are transcribed by RNA polymerase II." *EMBO J* 23 (20):4051-60. doi: 10.1038/sj.emboj.7600385.

Leete, P., A. Willcox, L. Krogvold, K. Dahl-Jorgensen, A. K. Foulis, S. J. Richardson, and N. G. Morgan. 2016. "Differential Insulitic Profiles Determine the Extent of beta-Cell Destruction and the Age at Onset of Type 1 Diabetes." *Diabetes* 65 (5):1362-9. doi: 10.2337/db15-1615.

Lehmann-Werman, R., D. Neiman, H. Zemmour, J. Moss, J. Magenheim, A. Vaknin-Dembinsky, S. Rubertsson, B. Nellgard, K. Blennow, H. Zetterberg, K. Spalding, M. J. Haller, C. H. Wasserfall, D. A. Schatz, C. J. Greenbaum, C. Dorrell, M. Grompe, A. Zick, A. Hubert, M. Maoz, V. Fendrich, D. K. Bartsch, T. Golan, S. A. Ben Sasson, G. Zamir, A. Razin, H. Cedar, A. M. Shapiro, B. Glaser, R. Shemer, and Y. Dor. 2016. "Identification of tissue-specific cell death using methylation patterns of circulating DNA." *Proc Natl Acad Sci U S A* 113 (13):E1826-34. doi: 10.1073/pnas.1519286113.

Lernmark, A. 2016. "Environmental factors in the etiology of type 1 diabetes, celiac disease, and narcolepsy." *Pediatr Diabetes* 17 Suppl 22:65-72. doi: 10.1111/pedi.12390.

Lernmark, A., J. L. Molenaar, W. A. van Beers, Y. Yamaguchi, S. Nagataki, J. Ludvigsson, and N. K. Maclaren. 1991. "The Fourth International Serum Exchange Workshop to standardize

cytoplasmic islet cell antibodies. The Immunology and Diabetes Workshops and Participating Laboratories." *Diabetologia* 34 (7):534-5.

Leygo, C., M. Williams, H. C. Jin, M. W. Y. Chan, W. K. Chu, M. Grusch, and Y. Y. Cheng. 2017. "DNA Methylation as a Noninvasive Epigenetic Biomarker for the Detection of Cancer." *Dis Markers* 2017:3726595. doi: 10.1155/2017/3726595.

Li, E., C. Beard, and R. Jaenisch. 1993. "Role for DNA methylation in genomic imprinting." *Nature* 366 (6453):362-5. doi: 10.1038/366362a0.

Li, Z., Y. Y. Ma, J. Wang, X. F. Zeng, R. Li, W. Kang, and X. K. Hao. 2016. "Exosomal microRNA-141 is upregulated in the serum of prostate cancer patients." *Onco Targets Ther* 9:139-48. doi: 10.2147/OTT.S95565.

Liao, J., R. Karnik, H. Gu, M. J. Ziller, K. Clement, A. M. Tsankov, V. Akopian, C. A. Gifford, J. Donaghey, C. Galonska, R. Pop, D. Reyon, S. Q. Tsai, W. Mallard, J. K. Joung, J. L. Rinn, A. Gnirke, and A. Meissner. 2015. "Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells." *Nat Genet* 47 (5):469-78. doi: 10.1038/ng.3258.

Lin, H. C., C. H. Wang, F. J. Tsai, K. P. Hwang, W. Chen, C. C. Lin, and T. C. Li. 2015. "Enterovirus infection is associated with an increased risk of childhood type 1 diabetes in Taiwan: a nationwide population-based cohort study." *Diabetologia* 58 (1):79-86. doi: 10.1007/s00125-014-3400-z.

Lincez, P. J., I. Shanina, and M. S. Horwitz. 2015. "Reduced expression of the MDA5 Gene IFIH1 prevents autoimmune diabetes." *Diabetes* 64 (6):2184-93. doi: 10.2337/db14-1223.

Linsley, P. S., W. Brady, L. Grosmaire, A. Aruffo, N. K. Damle, and J. A. Ledbetter. 1991. "Binding of the B cell activation antigen B7 to CD28 costimulates T cell proliferation and interleukin 2 mRNA accumulation." *J Exp Med* 173 (3):721-30.

Lister, R., M. Pelizzola, R. H. Dowen, R. D. Hawkins, G. Hon, J. Tonti-Filippini, J. R. Nery, L. Lee, Z. Ye, Q. M. Ngo, L. Edsall, J. Antosiewicz-Bourget, R. Stewart, V. Ruotti, A. H. Millar, J. A. Thomson, B. Ren, and J. R. Ecker. 2009. "Human DNA methylomes at base resolution show widespread epigenomic differences." *Nature* 462 (7271):315-22. doi: 10.1038/nature08514.

Liu, D., D. Pavlovic, M. C. Chen, M. Flodstrom, S. Sandler, and D. L. Eizirik. 2000. "Cytokines induce apoptosis in beta-cells isolated from mice lacking the inducible isoform of nitric oxide synthase (iNOS-/-)." *Diabetes* 49 (7):1116-22.

Lo, Y. M., J. Zhang, T. N. Leung, T. K. Lau, A. M. Chang, and N. M. Hjelm. 1999. "Rapid clearance of fetal DNA from maternal plasma." *Am J Hum Genet* 64 (1):218-24. doi: 10.1086/302205.

Locke, N. R., S. Stankovic, D. P. Funda, and L. C. Harrison. 2006. "TCR gamma delta intraepithelial lymphocytes are required for self-tolerance." *J Immunol* 176 (11):6553-9.

Lokk, K., V. Modhukur, B. Rajashekar, K. Martens, R. Magi, R. Kolde, M. Koltsina, T. K. Nilsson, J. Vilo, A. Salumets, and N. Tonisson. 2014. "DNA methylome profiling of human tissues identifies global and tissue-specific methylation patterns." *Genome Biol* 15 (4):r54. doi: 10.1186/gb-2014-15-4-r54.

Long, A. E., A. T. Gooneratne, S. Rokni, A. J. Williams, and P. J. Bingley. 2012. "The role of autoantibodies to zinc transporter 8 in prediction of type 1 diabetes in relatives: lessons from the European Nicotinamide Diabetes Intervention Trial (ENDIT) cohort." *J Clin Endocrinol Metab* 97 (2):632-7. doi: 10.1210/jc.2011-1952.

Lonnrot, M., K. Salminen, M. Knip, K. Savola, P. Kulmala, P. Leinikki, T. Hyypia, H. K. Akerblom, and H. Hyoty. 2000. "Enterovirus RNA in serum is a risk factor for beta-cell autoimmunity and clinical type 1 diabetes: a prospective study. Childhood Diabetes in Finland (DiMe) Study Group." *J Med Virol* 61 (2):214-20.

Lortz, S., M. Tiedge, T. Nachtwey, A. E. Karlsen, J. Nerup, and S. Lenzen. 2000. "Protection of insulin-producing RINm5F cells against cytokine-mediated toxicity through overexpression of antioxidant enzymes." *Diabetes* 49 (7):1123-30.

Lowe, C. E., J. D. Cooper, T. Brusko, N. M. Walker, D. J. Smyth, R. Bailey, K. Bourget, V. Plagnol, S. Field, M. Atkinson, D. G. Clayton, L. S. Wicker, and J. A. Todd. 2007. "Large-scale genetic fine mapping and genotype-phenotype associations implicate polymorphism in the IL2RA region in type 1 diabetes." *Nat Genet* 39 (9):1074-82. doi: 10.1038/ng2102.

Luco, R. F., Q. Pan, K. Tominaga, B. J. Blencowe, O. M. Pereira-Smith, and T. Misteli. 2010. "Regulation of alternative splicing by histone modifications." *Science* 327 (5968):996-1000. doi: 10.1126/science.1184208.

Lynch, S. M., K. M. O'Neill, M. M. McKenna, C. P. Walsh, and D. J. McKenna. 2016. "Regulation of miR-200c and miR-141 by Methylation in Prostate Cancer." *Prostate* 76 (13):1146-59. doi: 10.1002/pros.23201.

Lyttle, B. M., J. Li, M. Krishnamurthy, F. Fellows, M. B. Wheeler, C. G. Goodyer, and R. Wang. 2008. "Transcription factor expression in the developing human fetal endocrine pancreas." *Diabetologia* 51 (7):1169-80. doi: 10.1007/s00125-008-1006-z.

Madaschi, S., A. Rossini, I. Formenti, V. Lampasona, S. B. Marzoli, G. Cammarata, L. S. Politi, V. Martinelli, E. Bazzigaluppi, M. Scavini, E. Bosi, and R. Lanzi. 2010. "Treatment of thyroid-associated orbitopathy with rituximab--a novel therapy for an old disease: case report and literature review." *Endocr Pract* 16 (4):677-85. doi: 10.4158/EP09385.RA.

Maier, L. M., D. E. Anderson, C. A. Severson, C. Baecher-Allan, B. Healy, D. V. Liu, K. D. Wittrup, P. L. De Jager, and D. A. Hafler. 2009. "Soluble IL-2RA levels in multiple sclerosis subjects and the effect of soluble IL-2RA on immune responses." *J Immunol* 182 (3):1541-7.

Malek, T. R., and A. L. Bayer. 2004. "Tolerance, not immunity, crucially depends on IL-2." *Nat Rev Immunol* 4 (9):665-74. doi: 10.1038/nri1435.

Malek, T. R., and I. Castro. 2010. "Interleukin-2 receptor signaling: at the interface between tolerance and immunity." *Immunity* 33 (2):153-65. doi: 10.1016/j.immuni.2010.08.004.

Mandrup-Poulsen, T., K. Bendtzen, C. A. Dinarello, and J. Nerup. 1987. "Human tumor necrosis factor potentiates human interleukin 1-mediated rat pancreatic beta-cell cytotoxicity." *J Immunol* 139 (12):4077-82.

Mandrup-Poulsen, T., J. Molvig, H. U. Andersen, S. Helqvist, G. A. Spinas, and M. Munck. 1990. "Lack of predictive value of islet cell antibodies, insulin antibodies, and HLA-DR phenotype for remission in cyclosporin-treated IDDM patients. The Canadian-European Randomized Control Trial Group." *Diabetes* 39 (2):204-10.

Marques, R. G., M. J. Fontaine, and J. Rogers. 2004. "C-peptide: much more than a byproduct of insulin biosynthesis." *Pancreas* 29 (3):231-8.

Marroqui, L., R. S. Dos Santos, T. Floyel, F. A. Grieco, I. Santin, A. Op de Beeck, L. Marselli, P. Marchetti, F. Pociot, and D. L. Eizirik. 2015. "TYK2, a Candidate Gene for Type 1 Diabetes, Modulates Apoptosis and the Innate Immune Response in Human Pancreatic beta-Cells." *Diabetes* 64 (11):3808-17. doi: 10.2337/db15-0362.

Martinuzzi, E., G. Novelli, M. Scotto, P. Blancou, J. M. Bach, L. Chaillous, G. Bruno, L. Chatenoud, P. van Endert, and R. Mallone. 2008. "The frequency and immunodominance of islet-specific CD8+ T-cell responses change after type 1 diabetes diagnosis and treatment." *Diabetes* 57 (5):1312-20. doi: 10.2337/db07-1594.

Mathieu, C., E. Van Etten, C. Gysemans, B. Decallonne, S. Kato, J. Laureys, J. Depovere, D. Valckx, A. Verstuyf, and R. Bouillon. 2001. "In vitro and in vivo analysis of the immune system of vitamin D receptor knockout mice." *J Bone Miner Res* 16 (11):2057-65. doi: 10.1359/jbmr.2001.16.11.2057.

Mayeux, R. 2004. "Biomarkers: potential uses and limitations." *NeuroRx* 1 (2):182-8. doi: 10.1602/neurorx.1.2.182.

Mazin, A. L. 1994. "[Enzymatic DNA methylation as an aging mechanism]." *Mol Biol (Mosk)* 28 (1):21-51.

Mazin, A. L. 2009. "Suicidal function of DNA methylation in age-related genome disintegration." *Ageing Res Rev* 8 (4):314-27. doi: 10.1016/j.arr.2009.04.005.

Mehra, N. K., and G. Kaur. 2016. "Histocompatibility Antigen Complex of Man." In *eLS*. John Wiley & Sons.

Melkonyan, H. S., W. J. Feaver, E. Meyer, V. Scheinker, E. M. Shekhtman, Z. Xin, and S. R. Umansky. 2008. "Transrenal nucleic acids: from proof of principle to clinical tests." *Ann N Y Acad Sci* 1137:73-81. doi: 10.1196/annals.1448.015.

Moran, S., C. Arribas, and M. Esteller. 2016. "Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences." *Epigenomics* 8 (3):389-99. doi: 10.2217/epi.15.114.

Mosmann, T. R., H. Cherwinski, M. W. Bond, M. A. Giedlin, and R. L. Coffman. 1986. "Two types of murine helper T cell clone. I. Definition according to profiles of lymphokine activities and secreted proteins." *J Immunol* 136 (7):2348-57.

Mosmann, T. R., H. Cherwinski, M. W. Bond, M. A. Giedlin, and R. L. Coffman. 2005. "Two types of murine helper T cell clone. I. Definition according to profiles of lymphokine activities and secreted proteins. 1986." *J Immunol* 175 (1):5-14.

Nair, S., A. Al-Shabeeb, and M. E. Craig. 2013. "Enterovirus Infection, β-cell apoptosis and Type 1 Diabetes." *Microbiology Australia* 34 (3):153-156.

Nairn, C., D. N. Galbraith, K. W. Taylor, and G. B. Clements. 1999. "Enterovirus variants in the serum of children at the onset of Type 1 diabetes mellitus." *Diabet Med* 16 (6):509-13.

Neiman, D., J. Moss, M. Hecht, J. Magenheim, S. Piyanzin, A. M. J. Shapiro, E. J. P. de Koning, A. Razin, H. Cedar, R. Shemer, and Y. Dor. 2017. "Islet cells share promoter hypomethylation independently of expression, but exhibit cell-type-specific methylation in enhancers." *Proc Natl Acad Sci U S A* 114 (51):13525-13530. doi: 10.1073/pnas.1713736114.

Nejentsev, S., J. M. Howson, N. M. Walker, J. Szeszko, S. F. Field, H. E. Stevens, P. Reynolds, M. Hardy, E. King, J. Masters, J. Hulme, L. M. Maier, D. Smyth, R. Bailey, J. D. Cooper, G. Ribas, R. D. Campbell, D. G. Clayton, J. A. Todd, and Consortium Wellcome Trust Case Control. 2007. "Localization of type 1 diabetes susceptibility to the MHC class I genes HLA-B and HLA-A." *Nature* 450 (7171):887-92. doi: 10.1038/nature06406.

Nejentsev, S., N. Walker, D. Riches, M. Egholm, and J. A. Todd. 2009. "Rare variants of IFIH1, a gene implicated in antiviral responses, protect against type 1 diabetes." *Science* 324 (5925):387-9. doi: 10.1126/science.1167728.

Nerup, J., M. Christy, H. Kromann, P. Platz, L. P. Ryder, M. Thomsen, and A. Svejgaard. 1979. "HLA and insulin-dependent diabetes mellitus." *Postgrad Med J* 55 Suppl 2:8-13.

Nerup, J., T. Mandrup-Poulsen, S. Helqvist, H. U. Andersen, F. Pociot, J. I. Reimers, B. G. Cuartero, A. E. Karlsen, U. Bjerre, and T. Lorenzen. 1994. "On the pathogenesis of IDDM." *Diabetologia* 37 Suppl 2:S82-9.

Nerup, J., P. Platz, O. O. Andersen, M. Christy, J. Lyngsoe, J. E. Poulsen, L. P. Ryder, L. S. Nielsen, M. Thomsen, and A. Svejgaard. 1974. "HL-A antigens and diabetes mellitus." *Lancet* 2 (7885):864-6.

Nicol, J. W., G. A. Helt, S. G. Blanchard, Jr., A. Raja, and A. E. Loraine. 2009. "The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets." *Bioinformatics* 25 (20):2730-1. doi: 10.1093/bioinformatics/btp472.

Nilsson, B., K. N. Ekdahl, and O. Korsgren. 2011. "Control of instant blood-mediated inflammatory reaction to improve islets of Langerhans engraftment." *Curr Opin Organ Transplant* 16 (6):620-6. doi: 10.1097/MOT.0b013e32834c2393.

Nishikawa, K., Y. Iwamoto, Y. Kobayashi, F. Katsuoka, S. Kawaguchi, T. Tsujita, T. Nakamura, S. Kato, M. Yamamoto, H. Takayanagi, and M. Ishii. 2015. "DNA methyltransferase 3a regulates osteoclast differentiation by coupling to an S-adenosylmethionine-producing metabolic pathway." *Nat Med* 21 (3):281-7. doi: 10.1038/nm.3774.

Nistico, L., R. Buzzetti, L. E. Pritchard, B. Van der Auwera, C. Giovannini, E. Bosi, M. T. Larrad, M. S. Rios, C. C. Chow, C. S. Cockram, K. Jacobs, C. Mijovic, S. C. Bain, A. H. Barnett, C. L. Vandewalle, F. Schuit, F. K. Gorus, R. Tosi, P. Pozzilli, and J. A. Todd. 1996. "The CTLA-4 gene region of chromosome 2q33 is linked to, and associated with, type 1 diabetes. Belgian Diabetes Registry." *Hum Mol Genet* 5 (7):1075-80.

Noble, J. A., and A. M. Valdes. 2011. "Genetics of the HLA region in the prediction of type 1 diabetes." *Curr Diab Rep* 11 (6):533-42. doi: 10.1007/s11892-011-0223-x.

Noble, J. A., A. M. Valdes, M. Cook, W. Klitz, G. Thomson, and H. A. Erlich. 1996. "The role of HLA class II genes in insulin-dependent diabetes mellitus: molecular analysis of 180 Caucasian, multiplex families." *Am J Hum Genet* 59 (5):1134-48.

Noble, J. A., A. M. Valdes, M. D. Varney, J. A. Carlson, P. Moonsamy, A. L. Fear, J. A. Lane, E. Lavant, R. Rappner, A. Louey, P. Concannon, J. C. Mychaleckyj, H. A. Erlich, and Consortium Type 1 Diabetes Genetics. 2010. "HLA class I and genetic susceptibility to type 1 diabetes: results from the Type 1 Diabetes Genetics Consortium." *Diabetes* 59 (11):2972-9. doi: 10.2337/db10-0699.

Norman, A. W., J. B. Frankel, A. M. Heldt, and G. M. Grodsky. 1980. "Vitamin D deficiency inhibits pancreatic secretion of insulin." *Science* 209 (4458):823-5.

Notkins, A. L., and A. Lernmark. 2001. "Autoimmune type 1 diabetes: resolved and unresolved issues." *J Clin Invest* 108 (9):1247-52. doi: 10.1172/JCI14257.

O'Connell, R. M., D. S. Rao, A. A. Chaudhuri, and D. Baltimore. 2010. "Physiological and pathological roles for microRNAs in the immune system." *Nat Rev Immunol* 10 (2):111-22. doi: 10.1038/nri2708.

Oikarinen, M., S. Tauriainen, T. Honkanen, S. Oikarinen, K. Vuori, K. Kaukinen, I. Rantala, M. Maki, and H. Hyoty. 2008. "Detection of enteroviruses in the intestine of type 1 diabetic patients." *Clin Exp Immunol* 151 (1):71-5. doi: 10.1111/j.1365-2249.2007.03529.x.

Oikarinen, M., S. Tauriainen, S. Oikarinen, T. Honkanen, P. Collin, I. Rantala, M. Maki, K. Kaukinen, and H. Hyoty. 2012. "Type 1 diabetes is associated with enterovirus infection in gut mucosa." *Diabetes* 61 (3):687-91. doi: 10.2337/db11-1157.

Oilinki, T., T. Otonkoski, J. Ilonen, M. Knip, and P. J. Miettinen. 2012. "Prevalence and characteristics of diabetes among Somali children and adolescents living in Helsinki, Finland." *Pediatr Diabetes* 13 (2):176-80. doi: 10.1111/j.1399-5448.2011.00783.x.

Okano, M., D. W. Bell, D. A. Haber, and E. Li. 1999. "DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development." *Cell* 99 (3):247-57.

Okano, M., S. Xie, and E. Li. 1998. "Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells." *Nucleic Acids Res* 26 (11):2536-40.

Olsen, J. A., L. A. Kenna, M. G. Spelios, M. J. Hessner, and E. M. Akirav. 2016. "Circulating Differentially Methylated Amylin DNA as a Biomarker of beta-Cell Loss in Type 1 Diabetes." *PLoS One* 11 (4):e0152662. doi: 10.1371/journal.pone.0152662.

Olsson, R., J. Olerud, U. Pettersson, and P. O. Carlsson. 2011. "Increased numbers of low-oxygenated pancreatic islets after intraportal islet transplantation." *Diabetes* 60 (9):2350-3. doi: 10.2337/db09-0490.

Onengut-Gumuscu, S., W. M. Chen, O. Burren, N. J. Cooper, A. R. Quinlan, J. C. Mychaleckyj, E. Farber, J. K. Bonnie, M. Szpak, E. Schofield, P. Achuthan, H. Guo, M. D. Fortune, H. Stevens, N. M. Walker, L. D. Ward, A. Kundaje, M. Kellis, M. J. Daly, J. C. Barrett, J. D. Cooper, P. Deloukas, Consortium Type 1 Diabetes Genetics, J. A. Todd, C. Wallace, P. Concannon, and S. S. Rich. 2015. "Fine mapping of type 1 diabetes susceptibility loci and evidence for colocalization of causal variants with lymphoid gene enhancers." *Nat Genet* 47 (4):381-6. doi: 10.1038/ng.3245.

Orci, L., D. Baetens, M. Ravazzola, Y. Stefan, and F. Malaisse-Lagae. 1976. "Pancreatic polypeptide and glucagon : non-random distribution in pancreatic islets." *Life Sci* 19 (12):1811-5.

Palmer, J. P., C. M. Asplin, P. Clemons, K. Lyen, O. Tatpati, P. K. Raghu, and T. L. Paquette. 1983. "Insulin antibodies in insulin-dependent diabetics before insulin treatment." *Science* 222 (4630):1337-9.

Parikka, V., K. Nanto-Salonen, M. Saarinen, T. Simell, J. Ilonen, H. Hyoty, R. Veijola, M. Knip, and O. Simell. 2012. "Early seroconversion and rapidly increasing autoantibody concentrations predict prepubertal manifestation of type 1 diabetes in children at genetic risk." *Diabetologia* 55 (7):1926-36. doi: 10.1007/s00125-012-2523-3.

Parkin, J., and B. Cohen. 2001. "An overview of the immune system." *Lancet* 357 (9270):1777-89. doi: 10.1016/S0140-6736(00)04904-7.

Patrick, S. L., J. K. Kadohiro, S. H. Waxman, J. D. Curb, T. J. Orchard, J. S. Dorman, L. H. Kuller, and R. E. LaPorte. 1997. "IDDM incidence in a multiracial population. The Hawaii IDDM Registry, 1980-1990." *Diabetes Care* 20 (6):983-7.

Patterson, C. C., G. Dahlquist, V. Harjutsalo, G. Joner, R. G. Feltbower, J. Svensson, E. Schober, E. Gyurus, C. Castell, B. Urbonaite, J. Rosenbauer, V. Iotova, A. V. Thorsson, and G. Soltesz. 2007. "Early mortality in EURODIAB population-based cohorts of type 1 diabetes diagnosed in childhood since 1989." *Diabetologia* 50 (12):2439-42. doi: 10.1007/s00125-007-0824-8.

Peng, H., and W. Hagopian. 2006. "Environmental factors in the development of Type 1 diabetes." *Rev Endocr Metab Disord* 7 (3):149-62. doi: 10.1007/s11154-006-9024-y.

Penno, M. A., J. J. Couper, M. E. Craig, P. G. Colman, W. D. Rawlinson, A. M. Cotterill, T. W. Jones, L. C. Harrison, and Endia Study Group. 2013. "Environmental determinants of islet autoimmunity (ENDIA): a pregnancy to early life cohort study in children at-risk of type 1 diabetes." *BMC Pediatr* 13:124. doi: 10.1186/1471-2431-13-124.

Pescovitz, M. D., C. J. Greenbaum, B. Bundy, D. J. Becker, S. E. Gitelman, R. Goland, P. A. Gottlieb, J. B. Marks, A. Moran, P. Raskin, H. Rodriguez, D. A. Schatz, D. K. Wherrett, D. M. Wilson, J. P. Krischer, J. S. Skyler, and C. D. Study Group Type 1 Diabetes TrialNet Anti. 2014. "B-lymphocyte depletion with rituximab and beta-cell function: two-year results." *Diabetes Care* 37 (2):453-9. doi: 10.2337/dc13-0626.

Pescovitz, M. D., C. J. Greenbaum, H. Krause-Steinrauf, D. J. Becker, S. E. Gitelman, R. Goland, P. A. Gottlieb, J. B. Marks, P. F. McGee, A. M. Moran, P. Raskin, H. Rodriguez, D. A. Schatz, D. Wherrett, D. M. Wilson, J. M. Lachin, J. S. Skyler, and C. D. Study Group Type 1 Diabetes TrialNet Anti. 2009. "Rituximab, B-lymphocyte depletion, and preservation of beta-cell function." *N Engl J Med* 361 (22):2143-52. doi: 10.1056/NEJMoa0904452.

Peter, M. E. 2010. "Targeting of mRNAs by multiple miRNAs: the next step." *Oncogene* 29 (15):2161-4. doi: 10.1038/onc.2010.59.

Peterson, D. A., R. J. DiPaolo, O. Kanagawa, and E. R. Unanue. 1999. "Quantitative analysis of the T cell repertoire that escapes negative selection." *Immunity* 11 (4):453-62.

Pidsley, R., E. Zotenko, T. J. Peters, M. G. Lawrence, G. P. Risbridger, P. Molloy, S. Van Djik, B. Muhlhausler, C. Stirzaker, and S. J. Clark. 2016. "Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling." *Genome Biol* 17 (1):208. doi: 10.1186/s13059-016-1066-1.

Piemonti, L., M. J. Everly, P. Maffi, M. Scavini, F. Poli, R. Nano, M. Cardillo, R. Melzi, A. Mercalli, V. Sordi, V. Lampasona, A. Espadas de Arias, M. Scalamogna, E. Bosi, E. Bonifacio, A. Secchi, and P. I. Terasaki. 2013. "Alloantibody and autoantibody monitoring predicts islet transplantation outcome in human type 1 diabetes." *Diabetes* 62 (5):1656-64. doi: 10.2337/db12-1258.

Pihoker, C., L. K. Gilliam, C. S. Hampe, and A. Lernmark. 2005. "Autoantibodies in diabetes." *Diabetes* 54 Suppl 2:S52-61.

Pinkse, G. G., O. H. Tysma, C. A. Bergen, M. G. Kester, F. Ossendorp, P. A. van Veelen, B. Keymeulen, D. Pipeleers, J. W. Drijfhout, and B. O. Roep. 2005. "Autoreactive CD8 T cells associated with beta cell destruction in type 1 diabetes." *Proc Natl Acad Sci U S A* 102 (51):18425-30. doi: 10.1073/pnas.0508621102.

Pirot, P., A. K. Cardozo, and D. L. Eizirik. 2008. "Mediators and mechanisms of pancreatic beta-cell death in type 1 diabetes." *Arq Bras Endocrinol Metabol* 52 (2):156-65.

Plagnol, V., J. M. Howson, D. J. Smyth, N. Walker, J. P. Hafler, C. Wallace, H. Stevens, L. Jackson, M. J. Simmonds, Consortium Type 1 Diabetes Genetics, P. J. Bingley, S. C. Gough, and J. A. Todd. 2011. "Genome-wide association analysis of autoantibody positivity in type 1 diabetes cases." *PLoS Genet* 7 (8):e1002216. doi: 10.1371/journal.pgen.1002216.

Pociot, F., B. Akolkar, P. Concannon, H. A. Erlich, C. Julier, G. Morahan, C. R. Nierras, J. A. Todd, S. S. Rich, and J. Nerup. 2010. "Genetics of type 1 diabetes: what's next?" *Diabetes* 59 (7):1561-71. doi: 10.2337/db10-0076.

Pociot, F., and M. F. McDermott. 2002. "Genetics of type 1 diabetes mellitus." *Genes Immun* 3 (5):235-49. doi: 10.1038/sj.gene.6363875.

Polonsky, K., B. Frank, W. Pugh, A. Addis, T. Karrison, P. Meier, H. Tager, and A. Rubenstein. 1986a. "The limitations to and valid use of C-peptide as a marker of the secretion of insulin." *Diabetes* 35 (4):379-86.

Polonsky, K. S., J. Licinio-Paixao, B. D. Given, W. Pugh, P. Rue, J. Galloway, T. Karrison, and B. Frank. 1986b. "Use of biosynthetic human C-peptide in the measurement of insulin secretion rates in normal volunteers and type I diabetic patients." *J Clin Invest* 77 (1):98-105. doi: 10.1172/JCI112308.

Polonsky, K. S., and A. H. Rubenstein. 1986. "Current approaches to measurement of insulin secretion." *Diabetes Metab Rev* 2 (3-4):315-29.

Polymeropoulos, M. H., H. Xiao, D. S. Rath, and C. R. Merril. 1991. "Dinucleotide repeat polymorphism at the human CTLA4 gene." *Nucleic Acids Res* 19 (14):4018.

Portela, A., and M. Esteller. 2010. "Epigenetic modifications and human disease." *Nat Biotechnol* 28 (10):1057-68. doi: 10.1038/nbt.1685.

Poulin, M., and K. Haskins. 2000. "Induction of diabetes in nonobese diabetic mice by Th2 T cell clones from a TCR transgenic mouse." *J Immunol* 164 (6):3072-8.

Poy, M. N., L. Eliasson, J. Krutzfeldt, S. Kuwajima, X. Ma, P. E. Macdonald, S. Pfeffer, T. Tuschl, N. Rajewsky, P. Rorsman, and M. Stoffel. 2004. "A pancreatic islet-specific microRNA regulates insulin secretion." *Nature* 432 (7014):226-30. doi: 10.1038/nature03076.

Poy, M. N., J. Hausser, M. Trajkovski, M. Braun, S. Collins, P. Rorsman, M. Zavolan, and M. Stoffel. 2009. "miR-375 maintains normal pancreatic alpha- and beta-cell mass." *Proc Natl Acad Sci U S A* 106 (14):5813-8. doi: 10.1073/pnas.0810550106.

Pugliese, A. 2016. "Insulitis in the pathogenesis of type 1 diabetes." *Pediatr Diabetes* 17 Suppl 22:31-6. doi: 10.1111/pedi.12388.

Pugliese, A., R. Gianani, R. Moromisato, Z. L. Awdeh, C. A. Alper, H. A. Erlich, R. A. Jackson, and G. S. Eisenbarth. 1995. "HLA-DQB1*0602 is associated with dominant protection from diabetes even among islet cell antibody-positive first-degree relatives of patients with IDDM." *Diabetes* 44 (6):608-13.

Pugliese, A., M. Yang, I. Kusmarteva, T. Heiple, F. Vendrame, C. Wasserfall, P. Rowe, J. M. Moraski, S. Ball, L. Jebson, D. A. Schatz, R. Gianani, G. W. Burke, C. Nierras, T. Staeva, J. S. Kaddis, M. Campbell-Thompson, and M. A. Atkinson. 2014. "The Juvenile Diabetes Research Foundation Network for Pancreatic Organ Donors with Diabetes (nPOD) Program: goals, operational model and emerging findings." *Pediatr Diabetes* 15 (1):1-9. doi: 10.1111/pedi.12097.

Pugliese, A., M. Zeller, A. Fernandez, Jr., L. J. Zalcberg, R. J. Bartlett, C. Ricordi, M. Pietropaolo, G. S. Eisenbarth, S. T. Bennett, and D. D. Patel. 1997. "The insulin gene is transcribed in the human thymus and transcription levels correlated with allelic variation at the INS VNTR-IDDM2 susceptibility locus for type 1 diabetes." *Nat Genet* 15 (3):293-7. doi: 10.1038/ng0397-293.

Qu, H. Q., A. Montpetit, B. Ge, T. J. Hudson, and C. Polychronakos. 2007. "Toward further mapping of the association between the IL2RA locus and type 1 diabetes." *Diabetes* 56 (4):1174-6. doi: 10.2337/db06-1555.

Qu, Z., W. Li, and B. Fu. 2014. "MicroRNAs in autoimmune diseases." *Biomed Res Int* 2014:527895. doi: 10.1155/2014/527895.

Quintero-Ronderos, P., and G. Montoya-Ortiz. 2012. "Epigenetics and autoimmune diseases." *Autoimmune Dis* 2012:593720. doi: 10.1155/2012/593720.

Rabin, D. U., S. M. Pleasic, R. Palmer-Crocker, and J. A. Shapiro. 1992. "Cloning and expression of IDDM-specific human autoantigens." *Diabetes* 41 (2):183-6.

Rabinovitch, A., and W. L. Suarez-Pinzon. 1998. "Cytokines and their roles in pancreatic islet beta-cell destruction and insulin-dependent diabetes mellitus." *Biochem Pharmacol* 55 (8):1139-49.

Rakyan, V. K., H. Beyan, T. A. Down, M. I. Hawa, S. Maslau, D. Aden, A. Daunay, F. Busato, C. A. Mein, B. Manfras, K. R. Dias, C. G. Bell, J. Tost, B. O. Boehm, S. Beck, and R. D. Leslie. 2011a. "Identification of type 1 diabetes-associated DNA methylation variable positions that precede disease diagnosis." *PLoS Genet* 7 (9):e1002300. doi: 10.1371/journal.pgen.1002300.

Rakyan, V. K., T. A. Down, D. J. Balding, and S. Beck. 2011b. "Epigenome-wide association studies for common human diseases." *Nat Rev Genet* 12 (8):529-41. doi: 10.1038/nrg3000.

Rasmussen, K. D., and K. Helin. 2016. "Role of TET enzymes in DNA methylation, development, and cancer." *Genes Dev* 30 (7):733-50. doi: 10.1101/gad.276568.115.

Raulet, D. H. 1985. "Expression and function of interleukin-2 receptors on immature thymocytes." *Nature* 314 (6006):101-3.

Ravassard, P., Y. Hazhouz, S. Pechberty, E. Bricout-Neveu, M. Armanet, P. Czernichow, and R. Scharfmann. 2011. "A genetically engineered human pancreatic beta cell line exhibiting glucose-inducible insulin secretion." *J Clin Invest* 121 (9):3589-97. doi: 10.1172/JCI58447.

Rebollo, R., K. Miceli-Royer, Y. Zhang, S. Farivar, L. Gagnier, and D. L. Mager. 2012. "Epigenetic interplay between mouse endogenous retroviruses and host genes." *Genome Biol* 13 (10):R89. doi: 10.1186/gb-2012-13-10-r89.

Redondo, M. J., M. Rewers, L. Yu, S. Garg, C. C. Pilcher, R. B. Elliott, and G. S. Eisenbarth. 1999. "Genetic determination of islet cell autoimmunity in monozygotic twin, dizygotic twin, and non-twin siblings of patients with type 1 diabetes: prospective twin study." *BMJ* 318 (7185):698-702.

Redondo, M. J., L. Yu, M. Hawa, T. Mackenzie, D. A. Pyke, G. S. Eisenbarth, and R. D. Leslie. 2001. "Heterogeneity of type I diabetes: analysis of monozygotic twins in Great Britain and the United States." *Diabetologia* 44 (3):354-62. doi: 10.1007/s001250051626.

Reijonen, H., T. L. Daniels, A. Lernmark, and G. T. Nepom. 2000. "GAD65-specific autoantibodies enhance the presentation of an immunodominant T-cell epitope from GAD65." *Diabetes* 49 (10):1621-6.

Reik, W., W. Dean, and J. Walter. 2001. "Epigenetic reprogramming in mammalian development." *Science* 293 (5532):1089-93. doi: 10.1126/science.1063443.

Rekers, N. V., M. G. von Herrath, and J. D. Wesley. 2015. "Immunotherapies and immune biomarkers in Type 1 diabetes: A partnership for success." *Clin Immunol* 161 (1):37-43. doi: 10.1016/j.clim.2015.05.021.

Richardson, S. J., P. Leete, A. J. Bone, A. K. Foulis, and N. G. Morgan. 2013. "Expression of the enteroviral capsid protein VP1 in the islet cells of patients with type 1 diabetes is associated with induction of protein kinase R and downregulation of Mcl-1." *Diabetologia* 56 (1):185-93. doi: 10.1007/s00125-012-2745-4.

Richardson, S. J., T. Rodriguez-Calvo, I. C. Gerling, C. E. Mathews, J. S. Kaddis, M. A. Russell, M. Zeissler, P. Leete, L. Krogvold, K. Dahl-Jorgensen, M. von Herrath, A. Pugliese, M. A. Atkinson, and N. G. Morgan. 2016. "Islet cell hyperexpression of HLA class I antigens: a defining feature in type 1 diabetes." *Diabetologia* 59 (11):2448-2458. doi: 10.1007/s00125-016-4067-4.

Richardson, S. J., A. Willcox, A. J. Bone, A. K. Foulis, and N. G. Morgan. 2009. "The prevalence of enteroviral capsid protein vp1 immunostaining in pancreatic islets in human type 1 diabetes." *Diabetologia* 52 (6):1143-51. doi: 10.1007/s00125-009-1276-0.

Richardson, S. J., A. Willcox, A. J. Bone, N. G. Morgan, and A. K. Foulis. 2011. "Immunopathology of the human pancreas in type-I diabetes." *Semin Immunopathol* 33 (1):9-21. doi: 10.1007/s00281-010-0205-0.

Ritchie, M. E., B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, and G. K. Smyth. 2015. "limma powers differential expression analyses for RNA-sequencing and microarray studies." *Nucleic Acids Res* 43 (7):e47. doi: 10.1093/nar/gkv007.

Roep, B. O. 2003. "The role of T-cells in the pathogenesis of Type 1 diabetes: from cause to cure." *Diabetologia* 46 (3):305-21. doi: 10.1007/s00125-003-1089-5.

Roep, B. O., S. D. Arden, R. R. de Vries, and J. C. Hutton. 1990. "T-cell clones from a type-1 diabetes patient respond to insulin secretory granule proteins." *Nature* 345 (6276):632-4. doi: 10.1038/345632a0.

Roep, B. O., M. A. Atkinson, P. M. van Endert, P. A. Gottlieb, S. B. Wilson, and J. A. Sachs. 1999a. "Autoreactive T cell responses in insulin-dependent (Type 1) diabetes mellitus. Report of the first international workshop for standardization of T cell assays." *J Autoimmun* 13 (2):267-82. doi: 10.1006/jaut.1999.0312.

Roep, B. O., and M. Peakman. 2011. "Diabetogenic T lymphocytes in human Type 1 diabetes." *Curr Opin Immunol* 23 (6):746-53. doi: 10.1016/j.coi.2011.10.001.

Roep, B. O., I. Stobbe, G. Duinkerken, J. J. van Rood, A. Lernmark, B. Keymeulen, D. Pipeleers, F. H. Claas, and R. R. de Vries. 1999b. "Auto- and alloimmune reactivity to human islet allografts transplanted into type 1 diabetic patients." *Diabetes* 48 (3):484-90.

Roep, B. O., and T. I. Tree. 2014. "Immune modulation in humans: implications for type 1 diabetes mellitus." *Nat Rev Endocrinol* 10 (4):229-42. doi: 10.1038/nrendo.2014.2.

Roncarolo, M. G., and M. K. Levings. 2000. "The role of different subsets of T regulatory cells in controlling autoimmunity." *Curr Opin Immunol* 12 (6):676-83.

Rui, J., S. Deng, J. Lebastchi, P. L. Clark, S. Usmani-Brown, and K. C. Herold. 2016. "Methylation of insulin DNA in response to proinflammatory cytokines during the progression of autoimmune diabetes in NOD mice." *Diabetologia* 59 (5):1021-9. doi: 10.1007/s00125-016-3897-4.

Rumore, P., B. Muralidhar, M. Lin, C. Lai, and C. R. Steinman. 1992. "Haemodialysis as a model for studying endogenous plasma DNA: oligonucleosome-like structure and clearance." *Clin Exp Immunol* 90 (1):56-62.

Sacks, D. B., M. Arnold, G. L. Bakris, D. E. Bruns, A. R. Horvath, M. S. Kirkman, A. Lernmark, B. E. Metzger, and D. M. Nathan. 2011. "Guidelines and recommendations for laboratory analysis in the diagnosis and management of diabetes mellitus." *Clin Chem* 57 (6):e1-e47. doi: 10.1373/clinchem.2010.161596.

Sadeharju, K., M. Knip, M. Hiltunen, H. K. Akerblom, and H. Hyoty. 2003. "The HLA-DR phenotype modulates the humoral immune response to enterovirus antigens." *Diabetologia* 46 (8):1100-5. doi: 10.1007/s00125-003-1157-x.

Saisho, Y. 2016. "Postprandial C-Peptide to Glucose Ratio as a Marker of beta Cell Function: Implication for the Management of Type 2 Diabetes." *Int J Mol Sci* 17 (5). doi: 10.3390/ijms17050744.

Saito, K., N. Iwama, and T. Takahashi. 1978. "Morphometrical analysis on topographical difference in size distribution, number and volume of islets in the human pancreas." *Tohoku J Exp Med* 124 (2):177-86.

Schmidt, MB. 1902. "Ueber die beziehung der langenhans'schen inseln des pankreas zum diabetes mellitus. ." *München Med Wochenschr* 49:51–54.

Schwarzenbach, H., D. S. Hoon, and K. Pantel. 2011. "Cell-free nucleic acids as biomarkers in cancer patients." *Nat Rev Cancer* 11 (6):426-37. doi: 10.1038/nrc3066.

Seiskari, T., A. Kondrashova, H. Viskari, M. Kaila, A. M. Haapala, J. Aittoniemi, M. Virta, M. Hurme, R. Uibo, M. Knip, H. Hyoty, and Epivir study group. 2007. "Allergic sensitization and microbial load--a comparison between Finland and Russian Karelia." *Clin Exp Immunol* 148 (1):47-52. doi: 10.1111/j.1365-2249.2007.03333.x.

Seyfert-Margolis, V., T. D. Gisler, A. L. Asare, R. S. Wang, H. M. Dosch, B. Brooks-Worrell, G. S. Eisenbarth, J. P. Palmer, C. J. Greenbaum, S. E. Gitelman, G. T. Nepom, J. A. Bluestone, and K. C. Herold. 2006. "Analysis of T-cell assays to measure autoimmune responses in subjects with type 1 diabetes: results of a blinded controlled study." *Diabetes* 55 (9):2588-94. doi: 10.2337/db05-1378.

Shapiro, E. T., H. Tillil, A. H. Rubenstein, and K. S. Polonsky. 1988. "Peripheral insulin parallels changes in insulin secretion more closely than C-peptide after bolus intravenous glucose administration." *J Clin Endocrinol Metab* 67 (5):1094-9. doi: 10.1210/jcem-67-5-1094.

Shen, L., A. Inoue, J. He, Y. Liu, F. Lu, and Y. Zhang. 2014. "Tet3 and DNA replication mediate demethylation of both the maternal and paternal genomes in mouse zygotes." *Cell Stem Cell* 15 (4):459-471. doi: 10.1016/j.stem.2014.09.002.

Shiina, T., K. Hosomichi, H. Inoko, and J. K. Kulski. 2009. "The HLA genomic loci map: expression, interaction, diversity and disease." *J Hum Genet* 54 (1):15-39. doi: 10.1038/jhg.2008.5.

Sibley, R. K., D. E. Sutherland, F. Goetz, and A. F. Michael. 1985. "Recurrent diabetes mellitus in the pancreas iso- and allograft. A light and electron microscopic and immunohistochemical analysis of four cases." *Lab Invest* 53 (2):132-44.

Silva, D. G., S. R. Daley, J. Hogan, S. K. Lee, C. E. Teh, D. Y. Hu, K. P. Lam, C. C. Goodnow, and C. G. Vinuesa. 2011. "Anti-islet autoantibodies trigger autoimmune diabetes in the presence of an increased frequency of islet-reactive CD4 T cells." *Diabetes* 60 (8):2102-11. doi: 10.2337/db10-1344.

Simon, R. 2011. "Genomic biomarkers in predictive medicine: an interim analysis." *EMBO Mol Med* 3 (8):429-35. doi: 10.1002/emmm.201100153.

Simpson, M., H. Brady, X. Yin, J. Seifert, K. Barriga, M. Hoffman, T. Bugawan, A. E. Baron, R. J. Sokol, G. Eisenbarth, H. Erlich, M. Rewers, and J. M. Norris. 2011. "No association of vitamin D intake or 25-hydroxyvitamin D levels in childhood with risk of islet autoimmunity and type 1 diabetes: the Diabetes Autoimmunity Study in the Young (DAISY)." *Diabetologia* 54 (11):2779-88. doi: 10.1007/s00125-011-2278-2.

Singal, D. P., and M. A. Blajchman. 1973. "Histocompatibility (HL-A) antigens, lymphocytotoxic antibodies and tissue antibodies in patients with diabetes mellitus." *Diabetes* 22 (6):429-32.

Singh, R. P., I. Massachi, S. Manickavel, S. Singh, N. P. Rao, S. Hasan, D. K. Mc Curdy, S. Sharma, D. Wong, B. H. Hahn, and H. Rehimi. 2013. "The role of miRNA in inflammation and autoimmunity." *Autoimmun Rev* 12 (12):1160-5. doi: 10.1016/j.autrev.2013.07.003.

Sklenarova, J., L. Petruzelkova, S. Kolouskova, J. Lebl, Z. Sumnik, and O. Cinek. 2017. "Glucokinase Gene May Be a More Suitable Target Than the Insulin Gene for Detection of beta Cell Death." *Endocrinology* 158 (7):2058-2065. doi: 10.1210/en.2016-1923.

Small, E. M., and E. N. Olson. 2011. "Pervasive roles of microRNAs in cardiovascular biology." *Nature* 469 (7330):336-42. doi: 10.1038/nature09783.

Smyth, D., J. D. Cooper, J. E. Collins, J. M. Heward, J. A. Franklyn, J. M. Howson, A. Vella, S. Nutland, H. E. Rance, L. Maier, B. J. Barratt, C. Guja, C. Ionescu-Tirgoviste, D. A. Savage, D. B. Dunger, B.

Widmer, D. P. Strachan, S. M. Ring, N. Walker, D. G. Clayton, R. C. Twells, S. C. Gough, and J. A. Todd. 2004. "Replication of an association between the lymphoid tyrosine phosphatase locus (LYP/PTPN22) with type 1 diabetes, and evidence for its role as a general autoimmunity locus." *Diabetes* 53 (11):3020-3.

Smyth, D. J., J. D. Cooper, R. Bailey, S. Field, O. Burren, L. J. Smink, C. Guja, C. Ionescu-Tirgoviste, B. Widmer, D. B. Dunger, D. A. Savage, N. M. Walker, D. G. Clayton, and J. A. Todd. 2006. "A genome-wide association study of nonsynonymous SNPs identifies a type 1 diabetes locus in the interferon-induced helicase (IFIH1) region." *Nat Genet* 38 (6):617-9. doi: 10.1038/ng1800.

Smyth, D. J., V. Plagnol, N. M. Walker, J. D. Cooper, K. Downes, J. H. Yang, J. M. Howson, H. Stevens, R. McManus, C. Wijmenga, G. A. Heap, P. C. Dubois, D. G. Clayton, K. A. Hunt, D. A. van Heel, and J. A. Todd. 2008. "Shared and distinct genetic variants in type 1 diabetes and celiac disease." *N Engl J Med* 359 (26):2767-77. doi: 10.1056/NEJMoa0807917.

Soderstrom, U., J. Aman, and A. Hjern. 2012. "Being born in Sweden increases the risk for type 1 diabetes - a study of migration of children to Sweden as a natural experiment." *Acta Paediatr* 101 (1):73-7. doi: 10.1111/j.1651-2227.2011.02410.x.

Soltesz, G., C. C. Patterson, G. Dahlquist, and Eurodiab Study Group. 2007. "Worldwide childhood type 1 diabetes incidence--what can we learn from epidemiology?" *Pediatr Diabetes* 8 Suppl 6:6-14. doi: 10.1111/j.1399-5448.2007.00280.x.

Stead, J. D., J. Buard, J. A. Todd, and A. J. Jeffreys. 2000. "Influence of allele lineage on the role of the insulin minisatellite in susceptibility to type 1 diabetes." *Hum Mol Genet* 9 (20):2929-35.

Steck, A. K., S. Y. Liu, K. McFann, K. J. Barriga, S. R. Babu, G. S. Eisenbarth, M. J. Rewers, and J. X. She. 2006. "Association of the PTPN22/LYP gene with type 1 diabetes." *Pediatr Diabetes* 7 (5):274-8. doi: 10.1111/j.1399-5448.2006.00202.x.

Steck, A. K., and M. J. Rewers. 2011. "Genetics of type 1 diabetes." *Clin Chem* 57 (2):176-85. doi: 10.1373/clinchem.2010.148221.

Steck, A. K., K. Vehik, E. Bonifacio, A. Lernmark, A. G. Ziegler, W. A. Hagopian, J. She, O. Simell, B. Akolkar, J. Krischer, D. Schatz, M. J. Rewers, and Teddy Study Group. 2015. "Predictors of Progression From the Appearance of Islet Autoantibodies to Early Childhood Diabetes: The Environmental Determinants of Diabetes in the Young (TEDDY)." *Diabetes Care* 38 (5):808-13. doi: 10.2337/dc14-2426.

Steiner, D. F., S. J. Chan, J. M. Welsh, D. Nielsen, J. Michael, H. S. Tager, and A. H. Rubenstein. 1986. "Models of peptide biosynthesis: the molecular and cellular basis of insulin production." *Clin Invest Med* 9 (4):328-36.

Steiner, D. F., D. Cunningham, L. Spigelman, and B. Aten. 1967. "Insulin biosynthesis: evidence for a precursor." *Science* 157 (3789):697-700.

Steiner, D. F., and P. E. Oyer. 1967. "The biosynthesis of insulin and a probable precursor of insulin by a human islet cell adenoma." *Proc Natl Acad Sci U S A* 57 (2):473-80.

Stene, L. C., S. Oikarinen, H. Hyoty, K. J. Barriga, J. M. Norris, G. Klingensmith, J. C. Hutton, H. A. Erlich, G. S. Eisenbarth, and M. Rewers. 2010. "Enterovirus infection and progression from islet autoimmunity to type 1 diabetes: the Diabetes and Autoimmunity Study in the Young (DAISY)." *Diabetes* 59 (12):3174-80. doi: 10.2337/db10-0866.

Stephens, L. A., H. E. Thomas, L. Ming, M. Grell, R. Darwiche, L. Volodin, and T. W. Kay. 1999. "Tumor necrosis factor-alpha-activated cell death pathways in NIT-1 insulinoma cells and primary pancreatic beta cells." *Endocrinology* 140 (7):3219-27. doi: 10.1210/endo.140.7.6873.

Stoffers, D. A., N. T. Zinkin, V. Stanojevic, W. L. Clarke, and J. F. Habener. 1997. "Pancreatic agenesis attributable to a single nucleotide deletion in the human IPF1 gene coding sequence." *Nat Genet* 15 (1):106-10. doi: 10.1038/ng0197-106.

Strominger, J. L. 1987. "Structure of class I and class II HLA antigens." *Br Med Bull* 43 (1):81-93.

Su, Y. H., M. Wang, D. E. Brenner, A. Ng, H. Melkonyan, S. Umansky, S. Syngal, and T. M. Block. 2004. "Human urine contains small, 150 to 250 nucleotide-sized, soluble DNA derived from the circulation and may be useful in the detection of colorectal cancer." *J Mol Diagn* 6 (2):101-7. doi: 10.1016/S1525-1578(10)60497-7.

Swarup, V., and M. R. Rajeswari. 2007. "Circulating (cell-free) nucleic acids--a promising, non-invasive tool for early detection of several human diseases." *FEBS Lett* 581 (5):795-9. doi: 10.1016/j.febslet.2007.01.051.

Tahiliani, M., K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind, and A. Rao. 2009. "Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1." *Science* 324 (5929):930-5. doi: 10.1126/science.1170116.

Takaba, H., and H. Takayanagi. 2017. "The Mechanisms of T Cell Selection in the Thymus." *Trends Immunol* 38 (11):805-816. doi: 10.1016/j.it.2017.07.010.

Tang, Q., J. Y. Adams, C. Penaranda, K. Melli, E. Piaggio, E. Sgouroudis, C. A. Piccirillo, B. L. Salomon, and J. A. Bluestone. 2008. "Central role of defective interleukin-2 production in the triggering of islet autoimmune destruction." *Immunity* 28 (5):687-97. doi: 10.1016/j.immuni.2008.03.016.

Tentori, L., D. L. Longo, J. C. Zuniga-Pflucker, C. Wing, and A. M. Kruisbeek. 1988. "Essential role of the interleukin 2-interleukin 2 receptor pathway in thymocyte maturation in vivo." *J Exp Med* 168 (5):1741-7.

Tersey, S. A., J. B. Nelson, M. M. Fisher, and R. G. Mirmira. 2016. "Measurement of Differentially Methylated INS DNA Species in Human Serum Samples as a Biomarker of Islet beta Cell Death." *J Vis Exp* (118). doi: 10.3791/54838.

Thomas, H. E., and T. W. Kay. 2011. "Intracellular pathways of pancreatic beta-cell apoptosis in type 1 diabetes." *Diabetes Metab Res Rev* 27 (8):790-6. doi: 10.1002/dmrr.1253.

Thomas, H. E., J. A. Trapani, and T. W. Kay. 2010. "The role of perforin and granzymes in diabetes." *Cell Death Differ* 17 (4):577-85. doi: 10.1038/cdd.2009.165.

Todd, J. A., N. M. Walker, J. D. Cooper, D. J. Smyth, K. Downes, V. Plagnol, R. Bailey, S. Nejentsev, S. F. Field, F. Payne, C. E. Lowe, J. S. Szeszko, J. P. Hafler, L. Zeitels, J. H. Yang, A. Vella, S. Nutland, H. E. Stevens, H. Schuilenburg, G. Coleman, M. Maisuria, W. Meadows, L. J. Smink, B. Healy, O. S. Burren, A. A. Lam, N. R. Ovington, J. Allen, E. Adlem, H. T. Leung, C. Wallace, J. M. Howson, C. Guja, C. Ionescu-Tirgoviste, Finland Genetics of Type 1 Diabetes in, M. J. Simmonds, J. M. Heward, S. C. Gough, Consortium Wellcome Trust Case Control, D. B. Dunger, L. S. Wicker, and D. G. Clayton. 2007. "Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes." *Nat Genet* 39 (7):857-64. doi: 10.1038/ng2068.

Tooley, J. E., and K. C. Herold. 2014. "Biomarkers in type 1 diabetes: application to the clinical trial setting." *Curr Opin Endocrinol Diabetes Obes* 21 (4):287-92. doi: 10.1097/MED.0000000000000076.

Torn, C., D. Hadley, H. S. Lee, W. Hagopian, A. Lernmark, O. Simell, M. Rewers, A. Ziegler, D. Schatz, B. Akolkar, S. Onengut-Gumuscu, W. M. Chen, J. Toppari, J. Mykkanen, J. Ilonen, S. S. Rich, J. X. She, A. K. Steck, J. Krischer, and Teddy Study Group. 2015. "Role of Type 1 Diabetes-Associated SNPs on Risk of Autoantibody Positivity in the TEDDY Study." *Diabetes* 64 (5):1818-29. doi: 10.2337/db14-1497.

Trapani, J. A., and M. J. Smyth. 2002. "Functional significance of the perforin/granzyme cell death pathway." *Nat Rev Immunol* 2 (10):735-47. doi: 10.1038/nri911.

Tsonkova, V. G., F. W. Sand, X. A. Wolf, L. G. Grunnet, A. Kirstine Ringgaard, C. Ingvorsen, L. Winkel, M. Kalisz, K. Dalgaard, C. Bruun, J. J. Fels, C. Helgstrand, S. Hastrup, F. K. Oberg, E. Vernet, M. P. B. Sandrini, A. C. Shaw, C. Jessen, M. Gronborg, J. Hald, H. Willenbrock, D. Madsen, R. Wernersson, L. Hansson, J. N. Jensen, A. Plesner, T. Alanentalo, M. B. K. Petersen, A. Grapin-Botton, C. Honore, J. Ahnfelt-Ronne, J. Hecksher-Sorensen, P. Ravassard, O. D. Madsen, C. Rescan, and T. Frogne. 2018. "The EndoC-betaH1 cell line is a valid model of human beta cells

and applicable for screenings to identify novel drug target candidates." *Mol Metab* 8:144-157. doi: 10.1016/j.molmet.2017.12.007.

Tsui, N. B., P. Jiang, K. C. Chow, X. Su, T. Y. Leung, H. Sun, K. C. Chan, R. W. Chiu, and Y. M. Lo. 2012. "High resolution size analysis of fetal DNA in the urine of pregnant women by paired-end massively parallel sequencing." *PLoS One* 7 (10):e48319. doi: 10.1371/journal.pone.0048319.

Tsygankov, A. Y. 2008. "Multidomain STS/TULA proteins are novel cellular regulators." *IUBMB Life* 60 (4):224-31. doi: 10.1002/iub.36.

Ueda, H., J. M. Howson, L. Esposito, J. Heward, H. Snook, G. Chamberlain, D. B. Rainbow, K. M. Hunter, A. N. Smith, G. Di Genova, M. H. Herr, I. Dahlman, F. Payne, D. Smyth, C. Lowe, R. C. Twells, S. Howlett, B. Healy, S. Nutland, H. E. Rance, V. Everett, L. J. Smink, A. C. Lam, H. J. Cordell, N. M. Walker, C. Bordin, J. Hulme, C. Motzo, F. Cucca, J. F. Hess, M. L. Metzker, J. Rogers, S. Gregory, A. Allahabadia, R. Nithiyananthan, E. Tuomilehto-Wolf, J. Tuomilehto, P. Bingley, K. M. Gillespie, D. E. Undlien, K. S. Ronningen, C. Guja, C. Ionescu-Tirgoviste, D. A. Savage, A. P. Maxwell, D. J. Carson, C. C. Patterson, J. A. Franklyn, D. G. Clayton, L. B. Peterson, L. S. Wicker, J. A. Todd, and S. C. Gough. 2003. "Association of the T-cell regulatory gene CTLA4 with susceptibility to autoimmune disease." *Nature* 423 (6939):506-11. doi: 10.1038/nature01621.

Undlien, D. E., B. A. Lie, and E. Thorsby. 2001. "HLA complex genes in type 1 diabetes and other autoimmune diseases. Which genes are involved?" *Trends Genet* 17 (2):93-100.

Usmani-Brown, S., J. Lebastchi, A. K. Steck, C. Beam, K. C. Herold, and M. Ledizet. 2014. "Analysis of beta-cell death in type 1 diabetes by droplet digital PCR." *Endocrinology* 155 (9):3694-8. doi: 10.1210/en.2014-1150.

Vafiadis, P., S. T. Bennett, E. Colle, R. Grabs, C. G. Goodyer, and C. Polychronakos. 1996. "Imprinted and genotype-specific expression of genes at the IDDM2 locus in pancreas and leucocytes." *J Autoimmun* 9 (3):397-403. doi: 10.1006/jaut.1996.0054.

Vafiadis, P., S. T. Bennett, J. A. Todd, J. Nadeau, R. Grabs, C. G. Goodyer, S. Wickramasinghe, E. Colle, and C. Polychronakos. 1997. "Insulin expression in human thymus is modulated by INS VNTR alleles at the IDDM2 locus." *Nat Genet* 15 (3):289-92. doi: 10.1038/ng0397-289.

Vakoc, C. R., M. M. Sachdeva, H. Wang, and G. A. Blobel. 2006. "Profile of histone lysine methylation across transcribed mammalian chromatin." *Mol Cell Biol* 26 (24):9185-95. doi: 10.1128/MCB.01529-06.

Valdes, A. M., H. A. Erlich, and J. A. Noble. 2005. "Human leukocyte antigen class I B and C loci contribute to Type 1 Diabetes (T1D) susceptibility and age at T1D onset." *Hum Immunol* 66 (3):301-13. doi: 10.1016/j.humimm.2004.12.001.

van Belle, T. L., K. T. Coppieters, and M. G. von Herrath. 2011. "Type 1 diabetes: etiology, immunology, and therapeutic strategies." *Physiol Rev* 91 (1):79-118. doi: 10.1152/physrev.00003.2010.

van der Torren, C., A. Zaldumbide, D. L. Roelen, G. Duinkerken, S. H. Brand-Schaaf, M. Peakman, P. Czernichow, P. Ravassard, R. Scharfmann, and B. O. Roep. 2016. "Innate and adaptive immunity to human beta cell lines: implications for beta cell therapy." *Diabetologia* 59 (1):170-5. doi: 10.1007/s00125-015-3779-1.

Vang, T., M. Congia, M. D. Macis, L. Musumeci, V. Orru, P. Zavattari, K. Nika, L. Tautz, K. Tasken, F. Cucca, T. Mustelin, and N. Bottini. 2005. "Autoimmune-associated lymphoid tyrosine phosphatase is a gain-of-function variant." *Nat Genet* 37 (12):1317-9. doi: 10.1038/ng1673.

Varley, K. E., J. Gertz, K. M. Bowling, S. L. Parker, T. E. Reddy, F. Pauli-Behn, M. K. Cross, B. A. Williams, J. A. Stamatoyannopoulos, G. E. Crawford, D. M. Absher, B. J. Wold, and R. M. Myers. 2013. "Dynamic DNA methylation across diverse human cell lines and tissues." *Genome Res* 23 (3):555-67. doi: 10.1101/gr.147942.112.

Vasudevan, S. 2012. "Posttranscriptional upregulation by microRNAs." *Wiley Interdiscip Rev RNA* 3 (3):311-30. doi: 10.1002/wrna.121.

Vella, A., J. D. Cooper, C. E. Lowe, N. Walker, S. Nutland, B. Widmer, R. Jones, S. M. Ring, W. McArdle, M. E. Pembrey, D. P. Strachan, D. B. Dunger, R. C. Twells, D. G. Clayton, and J. A. Todd. 2005. "Localization of a type 1 diabetes locus in the IL2RA/CD25 region by use of tag single-nucleotide polymorphisms." *Am J Hum Genet* 76 (5):773-9. doi: 10.1086/429843.

Vogelstein, B., and K. W. Kinzler. 1999. "Digital PCR." *Proc Natl Acad Sci U S A* 96 (16):9236-41.

von Herrath, M. G., R. S. Fujinami, and J. L. Whitton. 2003. "Microorganisms and autoimmunity: making the barren field fertile?" *Nat Rev Microbiol* 1 (2):151-7. doi: 10.1038/nrmicro754.

von Herrath, M., S. Sanda, and K. Herold. 2007. "Type 1 diabetes as a relapsing-remitting disease?" *Nat Rev Immunol* 7 (12):988-94. doi: 10.1038/nri2192.

Von Meyenburg, M. 1940. " Ueber “Insulitis” bei Diabetes." *Schweiz Med Wochenschr* 21:554–557.

Voyias, P.D., A. A. Patel, and R. P. Arasaradnam. 2016. "Chapter 10: Epigenetic Biomarkers of Disease." In *Medical Epigenetics*, edited by Tollefsbol T. O., 159-176. Boston: Academic Press.

Vrba, L., T. J. Jensen, J. C. Garbe, R. L. Heimark, A. E. Cress, S. Dickinson, M. R. Stampfer, and B. W. Futscher. 2010. "Role for DNA methylation in the regulation of miR-200c and miR-141 expression in normal and cancer cells." *PLoS One* 5 (1):e8697. doi: 10.1371/journal.pone.0008697.

Walker, L. S., and M. von Herrath. 2016. "CD4 T cell differentiation in type 1 diabetes." *Clin Exp Immunol* 183 (1):16-29. doi: 10.1111/cei.12672.

Walter, U., A. Franzke, A. Sarukhan, C. Zober, H. von Boehmer, J. Buer, and O. Lechner. 2000. "Monitoring gene expression of TNFR family members by beta-cells during development of autoimmune diabetes." *Eur J Immunol* 30 (4):1224-32. doi: 10.1002/1521-4141(200004)30:4<1224::AID-IMMU1224>3.0.CO;2-B.

Wang, J., X. Han, and Y. Sun. 2017. "DNA methylation signatures in circulating cell-free DNA as biomarkers for the early detection of cancer." *Sci China Life Sci* 60 (4):356-362. doi: 10.1007/s11427-016-0253-7.

Waterhouse, P., J. M. Penninger, E. Timms, A. Wakeham, A. Shahinian, K. P. Lee, C. B. Thompson, H. Griesser, and T. W. Mak. 1995. "Lymphoproliferative disorders with early lethality in mice deficient in Ctla-4." *Science* 270 (5238):985-8.

Wattenhofer, M., K. Shibuya, J. Kudoh, R. Lyle, J. Michaud, C. Rossier, K. Kawasaki, S. Asakawa, S. Minoshima, A. Berry, B. Bonne-Tamir, N. Shimizu, S. E. Antonarakis, and H. S. Scott. 2001. "Isolation and characterization of the UBASH3A gene on 21q22.3 encoding a potential nuclear protein with a novel combination of domains." *Hum Genet* 108 (2):140-7.

Wegmann, D. R., M. Norbury-Glaser, and D. Daniel. 1994. "Insulin-specific T cells are a predominant component of islet infiltrates in pre-diabetic NOD mice." *Eur J Immunol* 24 (8):1853-7. doi: 10.1002/eji.1830240820.

Wenzlau, J. M., K. Juhl, L. Yu, O. Moua, S. A. Sarkar, P. Gottlieb, M. Rewers, G. S. Eisenbarth, J. Jensen, H. W. Davidson, and J. C. Hutton. 2007. "The cation efflux transporter ZnT8 (Slc30A8) is a major autoantigen in human type 1 diabetes." *Proc Natl Acad Sci U S A* 104 (43):17040-5. doi: 10.1073/pnas.0705894104.

WHO. 2006. "Definition and diagnosis of diabetes mellitus and intermediate hyperglycemia : report of a WHO/ IDF consultation." *World Health Organization*.

Wickham, H. 2007. "Reshaping data with the reshape package." *Journal of Statistical Software* 21 (12).

Wickham, H. 2009. *ggplot2: Elegant Graphics for Data Analysis*: Springer Publishing Company.

Willcox, A., S. J. Richardson, A. J. Bone, A. K. Foulis, and N. G. Morgan. 2009. "Analysis of islet inflammation in human type 1 diabetes." *Clin Exp Immunol* 155 (2):173-81. doi: 10.1111/j.1365-2249.2008.03860.x.

Williams, A. J., P. J. Bingley, E. Bonifacio, J. P. Palmer, and E. A. Gale. 1997. "A novel micro-assay for insulin autoantibodies." *J Autoimmun* 10 (5):473-8. doi: 10.1006/jaut.1997.0154.

Winter, W. E., and D. A. Schatz. 2011. "Autoimmune markers in diabetes." *Clin Chem* 57 (2):168-75. doi: 10.1373/clinchem.2010.148205.

Wong, F. S., L. Wen, M. Tang, M. Ramanathan, I. Visintin, J. Daugherty, L. G. Hannum, C. A. Janeway, Jr., and M. J. Shlomchik. 2004. "Investigation of the role of B-cells in type 1 diabetes in the NOD mouse." *Diabetes* 53 (10):2581-7.

Wu, J., A. Katrekar, L. A. Honigberg, A. M. Smith, M. T. Conn, J. Tang, D. Jeffery, K. Mortara, J. Sampang, S. R. Williams, J. Buggy, and J. M. Clark. 2006. "Identification of substrates of human protein-tyrosine phosphatase PTPN22." *J Biol Chem* 281 (16):11002-10. doi: 10.1074/jbc.M600498200.

Yang, B. T., T. A. Dayeh, C. L. Kirkpatrick, J. Taneera, R. Kumar, L. Groop, C. B. Wollheim, M. D. Nitert, and C. Ling. 2011. "Insulin promoter DNA methylation correlates negatively with insulin gene expression and positively with HbA(1c) levels in human pancreatic islets." *Diabetologia* 54 (2):360-7. doi: 10.1007/s00125-010-1967-6.

Yang, B. T., T. A. Dayeh, P. A. Volkov, C. L. Kirkpatrick, S. Malmgren, X. Jing, E. Renstrom, C. B. Wollheim, M. D. Nitert, and C. Ling. 2012a. "Increased DNA methylation and decreased expression of PDX-1 in pancreatic islets from patients with type 2 diabetes." *Mol Endocrinol* 26 (7):1203-12. doi: 10.1210/me.2012-1004.

Yang, H., Z. Wang, K. Xu, R. Gu, H. Chen, D. Yu, C. Xing, Y. Liu, L. Yu, J. Hutton, G. Eisenbarth, and T. Yang. 2012b. "IFIH1 gene polymorphisms in type 1 diabetes: genetic association analysis and genotype-phenotype correlation in Chinese Han population." *Autoimmunity* 45 (3):226-32. doi: 10.3109/08916934.2011.633134.

Yeung, W. C., W. D. Rawlinson, and M. E. Craig. 2011. "Enterovirus infection and type 1 diabetes mellitus: systematic review and meta-analysis of observational molecular studies." *BMJ* 342:d35. doi: 10.1136/bmj.d35.

Yoder, J. A., C. P. Walsh, and T. H. Bestor. 1997. "Cytosine methylation and the ecology of intragenomic parasites." *Trends Genet* 13 (8):335-40.

Yu, L., K. Herold, H. Krause-Steinrauf, P. L. McGee, B. Bundy, A. Pugliese, J. Krischer, G. S. Eisenbarth, and C. D. Study Group Type 1 Diabetes TrialNet Anti. 2011. "Rituximab selectively suppresses specific islet antibodies." *Diabetes* 60 (10):2560-5. doi: 10.2337/db11-0674.

Zamore, P. D., T. Tuschl, P. A. Sharp, and D. P. Bartel. 2000. "RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals." *Cell* 101 (1):25-33. doi: 10.1016/S0092-8674(00)80620-0.

Zeitz, U., K. Weber, D. W. Soegiarto, E. Wolf, R. Balling, and R. G. Erben. 2003. "Impaired insulin secretory capacity in mice lacking a functional vitamin D receptor." *FASEB J* 17 (3):509-11. doi: 10.1096/fj.02-0424fje.

Zhang, L., X. Lu, J. Lu, H. Liang, Q. Dai, G. L. Xu, C. Luo, H. Jiang, and C. He. 2012. "Thymine DNA glycosylase specifically recognizes 5-carboxylcytosine-modified DNA." *Nat Chem Biol* 8 (4):328-30. doi: 10.1038/nchembio.914.

Zhao, Y., N. A. Scott, S. Fynch, L. Elkerbout, W. W. Wong, K. D. Mason, A. Strasser, D. C. Huang, T. W. Kay, and H. E. Thomas. 2015. "Autoreactive T cells induce necrosis and not BCL-2-regulated or death receptor-mediated apoptosis or RIPK3-dependent necroptosis of transplanted islets in a mouse model of type 1 diabetes." *Diabetologia* 58 (1):140-8. doi: 10.1007/s00125-014-3407-5.

Zhou, H. L., G. Luo, J. A. Wise, and H. Lou. 2014. "Regulation of alternative splicing by local histone modifications: potential roles for RNA-guided mechanisms." *Nucleic Acids Res* 42 (2):701-13. doi: 10.1093/nar/gkt875.

Zhou, Q., J. Brown, A. Kanarek, J. Rajagopal, and D. A. Melton. 2008. "In vivo reprogramming of adult pancreatic exocrine cells to beta-cells." *Nature* 455 (7213):627-32. doi: 10.1038/nature07314.

Zhu, P., H. Guo, Y. Ren, Y. Hou, J. Dong, R. Li, Y. Lian, X. Fan, B. Hu, Y. Gao, X. Wang, Y. Wei, P. Liu, J. Yan, X. Ren, P. Yuan, Y. Yuan, Z. Yan, L. Wen, L. Yan, J. Qiao, and F. Tang. 2018. "Single-cell DNA methylome sequencing of human preimplantation embryos." *Nat Genet* 50 (1):12-19. doi: 10.1038/s41588-017-0007-6.

Ziegler, A. G., E. Bonifacio, and Babydiab-Babydiet Study Group. 2012. "Age-related islet autoantibody incidence in offspring of patients with type 1 diabetes." *Diabetologia* 55 (7):1937-43. doi: 10.1007/s00125-012-2472-x.

Ziegler, A. G., M. Hummel, M. Schenker, and E. Bonifacio. 1999. "Autoantibody appearance and risk for development of childhood diabetes in offspring of parents with type 1 diabetes: the 2-year analysis of the German BABYDIAB Study." *Diabetes* 48 (3):460-8.

Ziegler, A. G., M. Rewers, O. Simell, T. Simell, J. Lempainen, A. Steck, C. Winkler, J. Ilonen, R. Veijola, M. Knip, E. Bonifacio, and G. S. Eisenbarth. 2013a. "Seroconversion to multiple islet autoantibodies and risk of progression to diabetes in children." *Jama* 309 (23):2473-9. doi: 10.1001/jama.2013.6285.

Ziegler, A. I., M. A. Le Page, M. J. Maxwell, J. Stolp, H. Guo, A. Jayasimhan, M. L. Hibbs, P. Santamaria, J. F. Miller, M. Plebanski, P. A. Silveira, and R. M. Slattery. 2013b. "The CD19 signalling molecule is elevated in NOD mice and controls type 1 diabetes development." *Diabetologia* 56 (12):2659-68. doi: 10.1007/s00125-013-3038-2.

Ziller, M. J., H. Gu, F. Muller, J. Donaghey, L. T. Tsai, O. Kohlbacher, P. L. De Jager, E. D. Rosen, D. A. Bennett, B. E. Bernstein, A. Gnirke, and A. Meissner. 2013. "Charting a dynamic DNA

methylation landscape of the human genome." *Nature* 500 (7463):477-81. doi: 10.1038/nature12433.

Zipitis, C. S., and A. K. Akobeng. 2008. "Vitamin D supplementation in early childhood and risk of type 1 diabetes: a systematic review and meta-analysis." *Arch Dis Child* 93 (6):512-7. doi: 10.1136/adc.2007.128579.

Zoledziewska, M., C. Perra, V. Orru, L. Moi, P. Frongia, M. Congia, N. Bottini, and F. Cucca. 2008. "Further evidence of a primary, causal association of the PTPN22 620W variant with type 1 diabetes." *Diabetes* 57 (1):229-34. doi: 10.2337/db07-0289.

Appendix A

A. Appendix A

Detailed Protocols

Cell Free DNA Extraction from Serum/Plasma

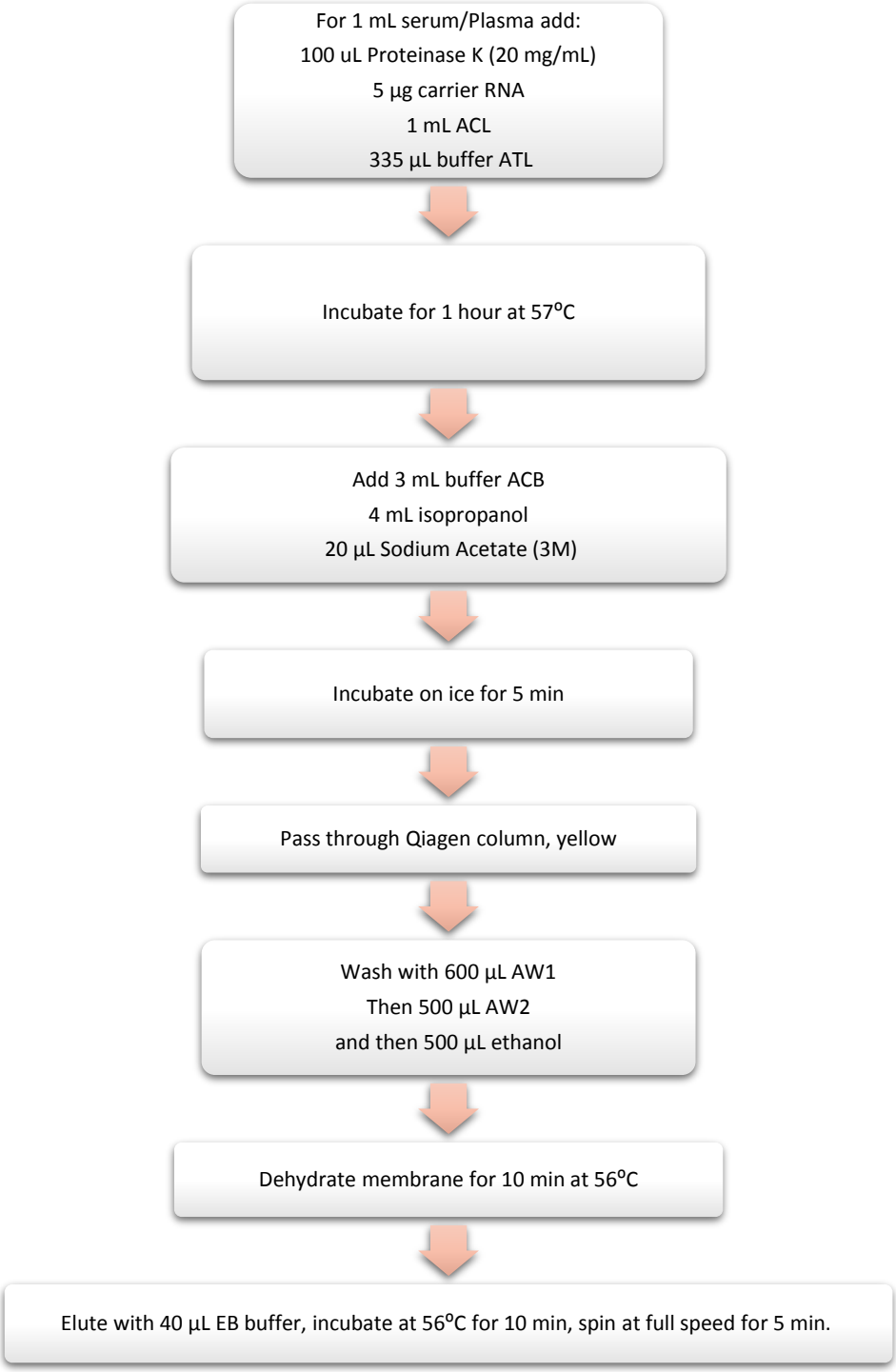


Figure A.1: The modified pH-column based method protocol

Triton/Heat/Phenol (THP) Protocol for Extracting Cell Free DNA

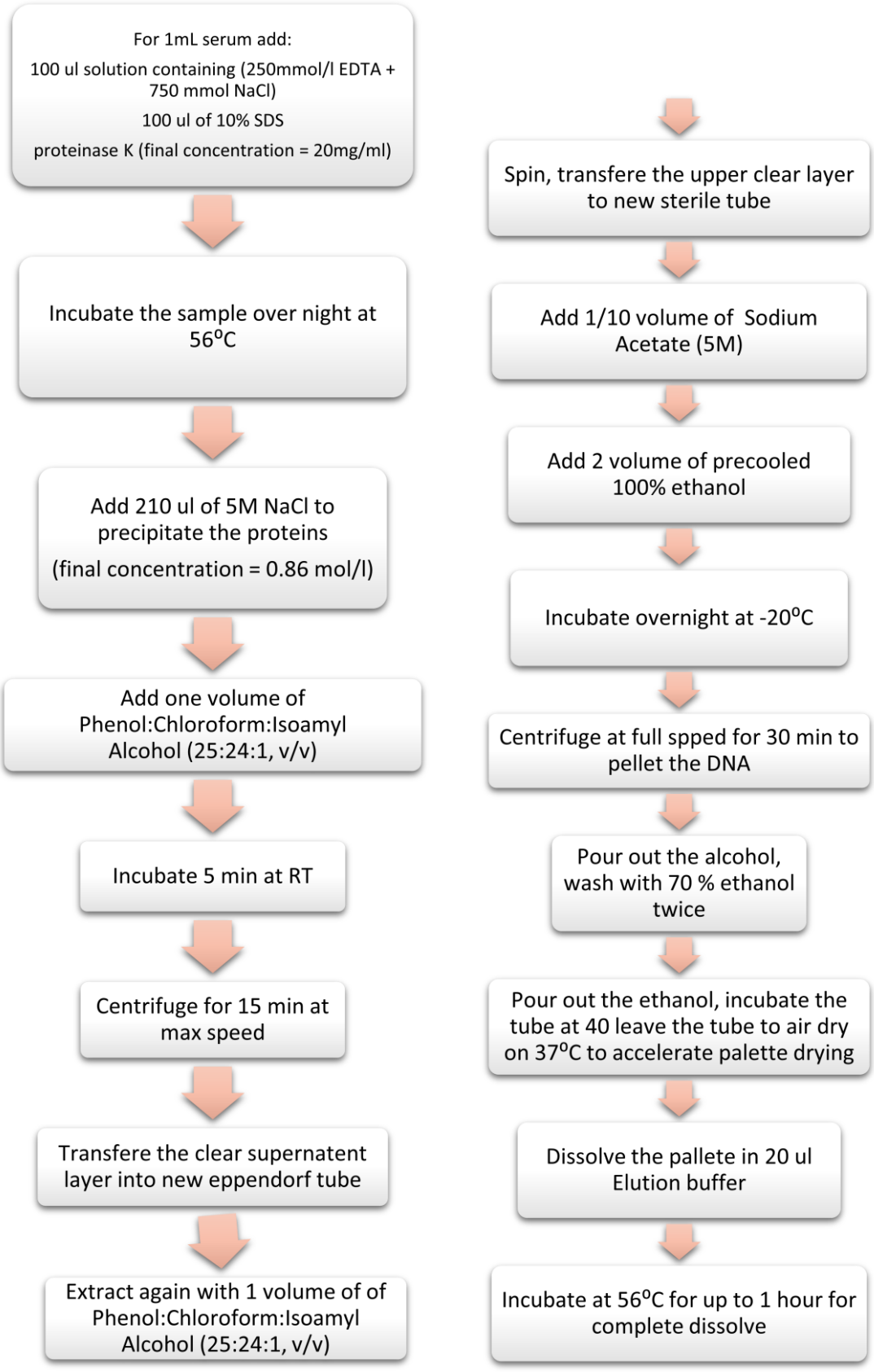


Figure A.2 The modified pH-column based method protocol

Cloning Using pGEM-T Easy Vector System I and DH5α Competent Cells

-80 Reagents (upstairs):

- Competent Cells (Invitrogen Cat# 18263-012)
- SOC Medium (Invitrogen, Cat#15544-034)

-20 Reagents (In Jody Box in post PCR Lab):

- pGEM-T Easy Vector system I (Promega, Cat# A1360)
- Ampicillin 100mg/mL (Sigma, Cat# A5354-10ML)

RT Reagents:

- imMedia Amp Blue agar media (Invitrogen, Cat# Q602-20)
- Terrific Broth (Sigma, Cat#T5574-500 mL)
- QiaPrep Spin miniprep Kit (Qiagen, Cat#27106)
- Petri dishes, polystyrene (Sigma, Cat# P5731-500EA)
- Spreader (Fisher Scientifics, Cat# 12322048)

Day 1:

A. Ligation Reaction:

1. Briefly Centrifuge the pGEM-T Vector and control insert DNA tubes.
2. Vortex the 2X rapid ligation buffer variously before each use
3. Set up the 10 µL reaction in 0.5 mL tube

Reagent	Single Reaction volume
2X Rapid Ligation Buffer	5 µL
pGEM-T Easy Vector	1 µL
T4 DNA Ligase	1 µL
PCR Product[14] / H_2O	3 µL

4. Set up the control reaction (Total 10 µL)

Reagent	Single Reaction volume
2X Rapid Ligation Buffer	5 µL

[14] DNA Product: is the PCR product band extracted from gel (agarose gel); when the band is cut from the gel keep the gel tray to minimize the gel exposure to the UV light to avoid Pyrimidine dimer.

pGEM-T Easy Vector	1 µL
T4 DNA Ligase	1 µL
Control Insert DNA	2 µL
Nuclease Free H_2O	1 µL

5. Mix the reaction by pipetting, incubate overnight @ 4°C

Day 2:

B. Prepare the agar plates:

1. Mix the imMedia™ pouch contents with 200 mL water.
2. Microwave on MEDIUM setting for 2-3 minutes. Then mix the solution and reheat for 30 seconds.
3. Pour 20 mL into 8-10 agar plates, allow them to solidify at RT.

C. Transformation Reaction:

1. Thaw the competent cells in ice, once thawed, flick the tube to mix (Do not vortex).
2. Add 3 µL of ligated reaction mix into ice cold 1.5 mL sterile Eppendorf tube (alternatively you can use the 15 mL sterile falcon tubes)
3. Gently, transform 50 µL of competent cells into the 1.5 tube.
4. Gently flick the tube to mix and incubate it on ice for 20 min
5. Heat-shock the reaction @ EXACTLY 42°C for 45-50 sec (DO NOT SHACK)
6. Immediately return the tubes into the ice and incubate it for 2 min.
7. Add 250 µL of RT SOC medium to each ligation reaction transformation.
8. Incubate @ 37°C for 1.5 hrs shaking (250 rpm)
9. Plate 100 µL of each transformation culture onto duplicate agar plate.
10. Incubate the plates overnight @ 37°C

Day 3:

1. In 15 mL sterile falcon tube add 5 mL of terrific broth + 100 µg/mL Amp (5 mL broth + 5 µL 100mg/mL Ampicillin)
2. Transfer 1 white colonies with 200 µL tip into the sterile 15 mL falcon tube.
3. Incubate the plates overnight @ 37°C shaking @ 170 rpm

Day 4:

1. Pellet the 5mL bacterial overnight culture at >8000 rpm (6800 x g) for 3 min @ RT
2. Follow the QiaPrep Spin miniprep Kit protocol
3. Measure the eluted DNA concentration using the nanodrop.
4. Send the proper concentration mixed with forward primer for sequencing
    a. Follow the instruction of company that will sequence the product.

Blood Collection and Processing using Roche Collection Tubes for cfDNA Study Experiments

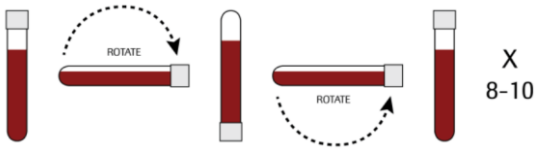
We collect 17mL of blood from patients who have consented to take part in the study. Blood is collected in two cell free DNA collection tubes (8.5 mL). Collected blood is stable at ambient temperature (18-25^{o}C) for 7 days.

Do not REFRIGERATE OR FREEZE.

Procedure

Safety Procedures Treat all blood samples as potentially infections.

1. Collect specimen (avoid backflow):
    a. Place the patient's arm in a downward position.
    b. Hold the tube with the stopper uppermost.
    c. Release tourniquet once blood starts to flow in the tube.
    d. Fill tube and confirm that the fill volume is between the minimum fill line and the nominal fill line.
    e. Ensure the tube contents do not touch the stopper or the end of the needle during collection.
    Do not fill the Cell-Free DNA Collection Tube above the nominal fill line.
2. Remove tube from adapter.
3. To avoid hemolysis, immediately mix by gentle inversion 8 to 10 times to ensure adequate mixing of the chemical additives with the blood specimen.



Separating the plasma

1. Separate plasma by centrifuging the sample 1600 x g for 15 minutes
    Do not exceed 1600 x g
    Do not Re-centrifuge.

2. Decap Cell-Free DNA Collection Tube: a. Hold the tube firmly in one hand, using a solid base to support the arm.
3. Twist the safety cap with the other hand to loosen.
4. Carefully open the tube with a gentle twist and pull motion.
5. Transfer 1 mL of plasma in 1.5 or 2 mL LoBind Eppendorf tubes.
6. Label it properly, store the samples in -80^{o}C until the time of DNA extraction.

DNA Quantification Using the QuantiFluor® dsDNA System (Promega).

Materials Required

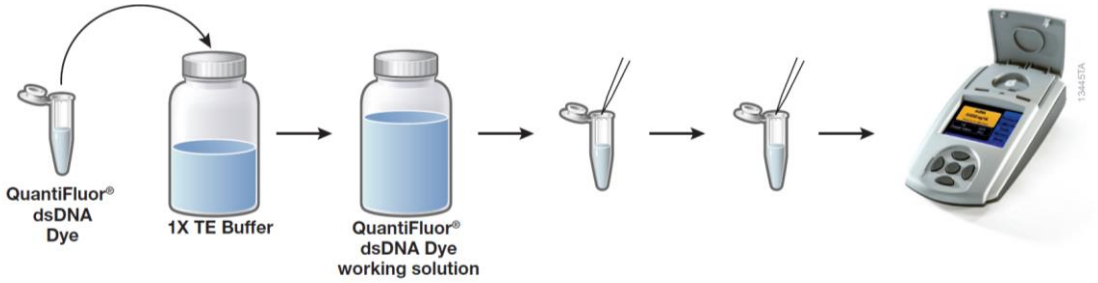
- QuantiFluor® dsDNA System (Cat. # E2670)
- Quantus™ Fluorometer (Cat. # E6150)
- Thin-walled 0.5 mL PCR tubes
- Nuclease-free water

Warm all assay components to room temperature before use.

Note: If the Quantus™ Fluorometer was previously calibrated, you may not need to calibrate it again. Therefore, do not prepare blank and standard samples.

Procedure

1. Prepare 1X TE Buffer: Dilute the 20X TE Buffer 20-fold with nuclease-free water.
2. Prepare Working Solution: Dilute the QuantiFluor® dsDNA Dye 1:400 in 1X TE buffer, and mix.
3. Prepare Blank: Mix 2 µL of 1X TE buffer with 200 µL of QuantiFluor® dsDNA Dye working solution in an empty 0.5 mL PCR tube. Vortex well and protect tube from light.
4. Prepare Standard: Add 2 µL of the provided DNA Standard (100 ng/µL) to 200 µL of QuantiFluor® dsDNA Dye working solution in an empty 0.5 mL PCR tube. Vortex well and protect tube from light.
5. Prepare Unknown(s): Add 1 µL of unknown samples to 200 µL of QuantiFluor® dsDNA Dye working solution in 0.5 mL PCR tubes. Vortex well and protect tube from light.
6. Incubate the prepared samples at room temperature for 5 minutes, protected from light.
7. Select the dsDNA protocol on the Quantus™ Fluorometer.
8. If needed, calibrate the Quantus™ Fluorometer by reading the blank (Step 3) and standard (Step 4) samples in the Calibration screen, then select “Save”.
9. Enter the volume of the unknown sample (1 µL used in Step 5) and desired concentration units.
10. Measure fluorescence of the unknown sample and record the final sample concentration results.



Bisulfite Converted DNA Quantification Using the QuantiFluor® dsDNA System (Promega).

Materials Required

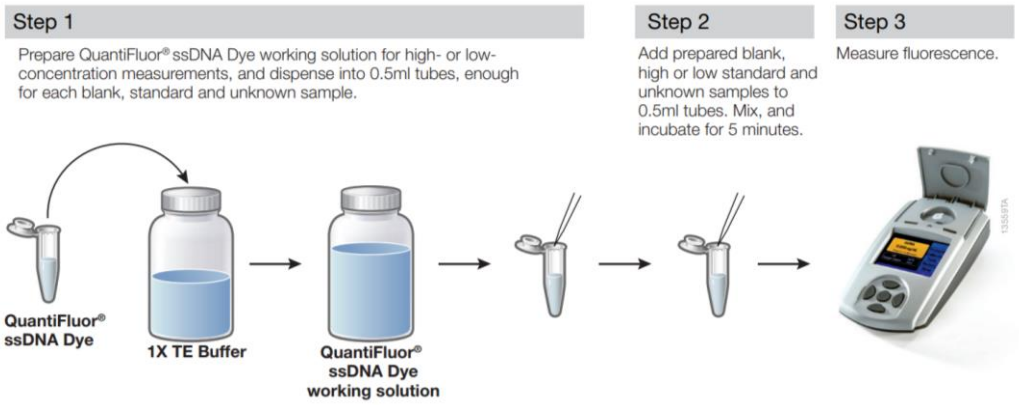
- QuantiFluor® ssDNA System (Cat. # E3190)
- Quantus™ Fluorometer (Cat. # E6150)
- thin-walled 0.5mL PCR tubes (Cat. # E4941 or Axygen Cat. # PCR-05-C)
- nuclease-free water

Warm all assay components to room temperature before use.

Note: If the Quantus™ Fluorometer was previously calibrated, you may not need to calibrate it again. Therefore, do not prepare blank and standard samples.

Procedure

1. Prepare 1X TE Buffer: Dilute the 20X TE Buffer 20-fold with nuclease-free water.
2. Prepare Working Solution: High Standard Calibration: Dilute the QuantiFluor® ssDNA Dye 1:400 in 1X TE buffer and mix thoroughly. Low Standard Calibration: Dilute the QuantiFluor® ssDNA Dye 1:2,000 in 1X TE buffer, and mix.
3. Prepare Blank: Add 200 µL of QuantiFluor® ssDNA Dye working solution in an empty 0.5 mL PCR tube. Protect tube from light.
4. Prepare Standard: High Standard Calibration: Prepare a 400-ng standard by adding 4 µL of the provided ssDNA Standard to 200 µL of QuantiFluor® ssDNA Dye working solution in an empty 0.5 mL PCR tube. Mix, and protect tube from light. Low Standard Calibration: Prepare a 10-ng standard by diluting the provided ssDNA Standard 1:100 in 1X TE buffer. Next, add 10 µL of diluted standard to 200 µL of QuantiFluor® ssDNA Dye working solution in a 0.5 mL PCR tube. Mix, and protect tube from light.
5. Prepare Unknown(s): Add 1 µL of unknown samples to 200 µL of QuantiFluor® ssDNA Dye working solution in 0.5 mL PCR tubes. Vortex well and protect tube from light.
6. Incubate the prepared samples at room temperature for 5 minutes, protected from light.
7. Select the ssDNA protocol on the Quantus™ Fluorometer. If needed, calibrate the Quantus™ Fluorometer by reading the blank (Step 3) and standard (Step 4) samples in the Calibration screen, then select “Save”.
8. Enter the volume of the unknown sample (1 µL used in Step 5) and desired concentration units.
9. Measure fluorescence of the unknown sample and record the final sample concentration results.



RNA Quantification using the QuantiFluor® RNA System (Promega).

Materials Required

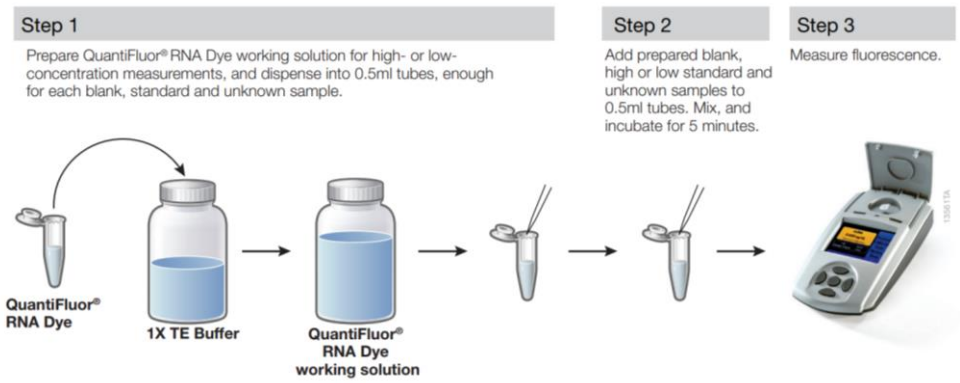
- QuantiFluor® RNA System (Cat. # E3310)
- Quantus™ Fluorometer (Cat. # E6150)
- Thin-walled 0.5mL PCR tubes (Cat. # E4941 or Axygen Cat. # PCR-05-C)
- Nuclease-free water

Warm all assay components to room temperature before use.

Note: If the Quantus™ Fluorometer was previously calibrated, you may not need to calibrate it again. Therefore, do not prepare blank and standard samples.

Procedure

1. Prepare 1X TE Buffer: Dilute the 20X TE Buffer 20-fold with nuclease-free water.
2. Prepare Working Solution: High Standard Calibration: Dilute the QuantiFluor® RNA Dye 1:400 in 1X TE buffer and mix thoroughly. Low Standard Calibration: Dilute the QuantiFluor® RNA Dye 1:2,000 in 1X TE buffer, and mix.
3. Prepare Blank: Add 200 µL of QuantiFluor® RNA Dye working solution in an empty 0.5 mL PCR tube. Protect tube from light.
4. Prepare Standard: High Standard Calibration: Prepare a 500-ng standard by adding 5 µL of the provided RNA Standard to 200 µL of QuantiFluor® RNA Dye working solution in an empty 0.5mL PCR tube. Mix, and protect tube from light. Low Standard Calibration: Prepare a 10-ng standard by diluting the provided RNA Standard 1:100 in 1X TE buffer. Next, add 10 µL of diluted standard to 200 µL of QuantiFluor® RNA Dye working solution in a 0.5 mL PCR tube. Mix, and protect tube from light.
5. Prepare Unknown(s): Add 1 µL of unknown samples to 200 µL of QuantiFluor® RNA Dye working solution in 0.5 mL PCR tubes. Vortex well and protect tube from light.
6. Incubate the prepared samples at room temperature for 5 minutes, protected from light.
7. Select the RNA protocol on the Quantus™ Fluorometer. If needed, calibrate the Quantus™ Fluorometer by reading the blank (Step 3) and standard (Step 4) samples in the Calibration screen, then select “Save”.
8. Enter the volume of the unknown sample (1 µL used in Step 5) and desired concentration units.
9. Measure fluorescence of the unknown sample and record the final sample concentration results.



Amplification of Purified Genomic DNA Using the REPLI-g Mini Kit

Procedure

1. Prepare sufficient Buffer D1 (denaturation buffer) and Buffer N1 (neutralization buffer) for the total number of whole genome amplification reactions.
    a. Note: The total volumes of Buffer D1 and Buffer N1 given in Tables 1 and 2 are suitable for up to ▲ 15 or ● 7 reactions. Buffer D1 and Buffer N1 should not be stored longer than 3 months

Preparation of Buffer D1	
Component	**Volume (µL)**
Reconstituted Buffer DLB	9
Nuclease-free water	32
Total Volume	41

Preparation of Buffer N1	
Component	**Volume (µL)**
Stop Solution	12
Nuclease-free water	68
Total Volume	80

2. Place ▲ 2.5 µL or ● 5 µL template DNA into a microcentrifuge. The amount of template DNA should be >10 ng. A DNA control reaction can be set up using 10 ng (1 µl) control genomic DNA (e.g., REPLI-g Human Control Kit, cat. no. 150090). Adjust the volume with TE to the starting volume of your sample.
3. Add ▲ 2.5 µL or ● 5 µL Buffer D1 to the DNA. Mix by vortexing and centrifuge briefly.
4. Incubate the samples at room temperature for 3 min.
5. Add ▲ 5 µL or ● 10 µL Buffer N1 to the samples. Mix by vortexing and centrifuge briefly.
6. Thaw REPLI-g Mini DNA Polymerase on ice. Thaw all other components at room temperature, vortex, then centrifuge briefly. The REPLI-g Mini Reaction Buffer may form a precipitate after thawing. The precipitate will dissolve by vortexing for 10 s.
7. Prepare a master mix on ice according to following Table. Mix and centrifuge briefly.
    a. Important: Add the master mix components in the order listed in the Table. After the addition of water and REPLI-g Mini Reaction Buffer, briefly vortex and centrifuge the mixture before the addition of REPLI-g Mini DNA Polymerase. The master mix should be kept on ice and used immediately upon addition of the REPLI-g Mini DNA Polymerase.

Component	Volume (µL)
Nuclease-free water	▲ 10 µl; ● 0 µl
REPLI-g Mini Reaction Buffer	29
REPLI-g Mini DNA Polymerase	1
Total Volume	▲ 40 µl; ● 30 µl

8. Add ▲ 40 µL or ● 30 µL of the master mix to ▲ 10 µL or ● 20 µL of denatured DNA (step 5).
9. Incubate at 30°C for 10–16 h. Maximum DNA yield is achieved using an incubation time of 16 h. After incubation at 30°C, heat the water bath or heating block up to 65°C if the same water bath or heating block will be used in step 10. Note: If a thermal cycler is used with a heated lid, temperature of the lid should be set to 70 °C.
10. Inactivate REPLI-g Mini DNA Polymerase by heating the sample for 3 min at 65°C.
11. If performing PCR analysis, dilute the amplified DNA 1:20 and use 3 µL of diluted DNA for each PCR.
    a. Note: For dilution, add 2 µL amplified to 38 µL water or TE. Use 3 µL of the diluted DNA for each PCR. Optical density (OD) measurements do not accurately quantify double-stranded DNA.
12. Store amplified DNA at 4°C for short-term storage or –20°C for long-term storage. DNA amplified using the REPLI-g kit should be treated as genomic DNA with minimal freeze-thaw cycles. Storage of nucleic acids at low concentration over a long period of time may result in acid hydrolysis. We therefore recommend storage of nucleic acids at a concentration of at least 100 ng/µl.

Tables

Table A.1: The Characterises of oligonucleotides used in this project.

Gene	GRCh37/hg19	Amplicon size (bp)	Forward Primer	Final Concentration (nM)	Reverse Primer	Final Concentration (nM)	Probe (FAM)	Final Concentration (nM)	Probe (VIC)	Final Concentration (nM)
PDX-1	Chr13: 28,493,258-28,493,533	275	5'TGTAAGTAGAAGAGAGTGAGTGTTTTTTG3'	900	5'ATAACTTCTAAAACATTCACAAAAACTTT3'	900	5'FAM-TCTACAAACCAAAACC-MGB3'	200	5'VIC-TCTACGAACCGAAACC-MGB3'	200
NCOR2	Chr12: 125,002,279-125,002,466	270	5'GGGAATAGAAGTTTTTTGAGTTTTTAAGG3'	150	5'ACCAATAAAAACAACTAAACCCTCCC3'	100	5'FAM-AAACACTAACCAAACAA-MGB3'	200	5'VIC-GTTTGTTCGGTTAGCGT-MGB3'	200
MIR-141	Chr12: 6,963,000-6,964,001	98	5'TGGTTTGTGGTTAGGGTTTTT3'	100	5'CCATCCAACACTATACTAAAAAATAAA3'	600	5'FAM-TAAACTACAATACGCGCTCA-MGB3'	300	5'VIC-AATAAACTACAATACACACTCA-MGB3'	300
SLC45A4	Chr8: 142,289,746-142,289,846	100	5'TGGTTGTTTTTGTAAAGATTTTTTG3'	600	5'CACCCCTCCCTAACCAAACTAC3'	600	5'FAM-CTCCCTAACACACACCC-MGB3'	300		
Beta-actin	Chr7: 5,566,779-5,570,232	130	5'TGGTGATGGAGGAGGTTTAGTAAGT3'	300	5'AACCAATAAAACCTACTCCTCCCTTAA3'	300	5'FAM-ACCACCACCCAACACACAATAACAAACACA-BHQ1'3	300	This assay was conducted from **MethyLight: a high-throughput assay to measure DNA methylation**	
Beta-globin	Chr.11: 5244554 - 5245546	75	5'-TGCACGTGGATCCTGAGAACT-3'	900	5'-AATTCTTTGCCAAAGTGATGGG-3'	900	5'FAM-CAGCACGTTGCCCAGGAGCCTG- MGB 3'	200	This assay was conducted from ThermoFisher Assay ID: Hs01629437_s1 Catalogue # 4331182	

MIR-200c Cluster 1	Chr12: 7,072,862-7,072,929	101	5'CCGAAACATCAACCTTACCCC3'	600	5'GGTTGTGTTGTGAGGTGGGTTT3'	50	5'FAM-AAAAAACAATCACACTCCAC-MGB3'	300	Designed and optimised by Dr Jody Ye
MIR-200c Cluster 2	Chr12: 7,072,862-7,072,929	101	5'TCTAAACCACCTTCCCCTACCC3'	300	5'TTGGGTTGAGTTTGGGATTGTA3'	50	5'VIC-AACCTTTCCCAACCCACAA-MGB3'	300	
INS	Chr11: 2150342-2182439	96	5'CTACAAATCCTCTACCTCCCAACA3'	600	5'TGTGGTTTATATTTGGTGGAAGTTT3'	50	5'FAM/VIC-AAAAACCTCATTCCCCA-MGB3' (70% FAM+30% VIC)	400	
PDX-1 proximal promoter (P1)	Chr13: 28,493,258-28,493,533	433	5'AAGTAGAAGAGAGTGAGTGTTTTTTGT3'	5000	5'TTTACAAACCCAACTCTCTAAATCAA3'	5000			
PDX-1 distal promoter (P2)	Chr13: 28,493,258-28,493,533	483	5'AATGTTTTGATTAGAGAGTTGGGT3'	5000	5'ACTAATCTCAAAAAAAACCCACAAC3'	5000			

Table A.2: The microRNA LNA PCR primer sets.

Primer Sets	Target Sequence	Product Number
hsa-miR-200c-3p	5'UAAUACUGCCGGGUAAUGAUGGA’3	YP00204482
hsa-miR-141-3p	5'UAACACUGUCUGGUAAAGAUGG’3	YP00204504
hsa-miR-141-5p	5'CAUCUUCCAGUACAGUGUUGGA’3	YP00206088

Table A.3: The differentially hypomethylated CpG sites in the islet genome.

	CpG ID	chromosome	position	Gene Name	islet	PBMC	aorta	liver	lung	pancreas	thymus	spleen
1	cg14516100	chr4	186560083	*SORBS2*	0.97	0.03	0.08	0.47	0.46	0.76	0.00	0.08
2	cg01241974	chr13	36045613	*NBEA*	0.93	0.05	0.31	0.27	0.10	0.57	0.05	0.06
3	cg03953626	chr6	30652396	*KIAA1949*	0.92	0.03	0.13	0.29	0.30	0.50	0.00	0.00
4	cg16863795	chr17	79259536	*SLC38A10*	0.92	0.03	0.43	0.40	0.48	0.52	0.04	0.00
5	cg04416414	chr19	14260587	*LPHN1*	0.92	0.09	0.05	0.44	0.23	0.82	0.08	0.03
6	cg02332696	chr10	21783091	*DNAJC1*	0.91	0.13	0.31	0.24	0.47	0.94	0.00	0.21
7	cg24540114	chr7	156812941	*LMBR1*	0.91	0.15	0.18	0.09	0.12	0.83	0.24	0.15
8	cg01978215	chr2	145078303	*GTDC1*	0.90	0.09	0.07	0.32	0.27	0.81	0.17	0.06
9	cg10781408	chr1	151810899	*C2CD4D; LOC100132111*	0.90	0.03	0.03	0.02	0.26	0.85	0.00	0.03
10	cg19112359	chr2	169346613	*LASS6*	0.90	0.05	0.37	0.32	0.35	0.38	0.00	0.07
11	cg02599716	chr2	10231486		0.89	0.05	0.26	0.38	0.23	0.81	0.05	0.13
12	cg08004073	chr3	169377725	*MECOM*	0.89	0.21	0.04	0.16	0.43	0.35	0.07	0.11
13	cg01156663	chr4	154360462	*DCHS2*	0.89	0.04	0.30	0.00	0.47	0.67	0.02	0.23
14	cg24717695	chr7	156809644	*LMBR1*	0.89	0.15	0.13	0.08	0.05	0.90	0.05	0.09
15	cg14565903	chr19	14260591	*LPHN1*	0.89	0.17	0.04	0.45	0.23	0.81	0.00	0.03
16	cg24922129	chr11	129242530		0.89	0.04	0.06	0.03	0.14	0.70	0.00	0.06
17	cg01244015	chr2	225266843	*FAM124B*	0.88	0.08	0.34	0.21	0.17	0.71	0.03	0.02
18	cg15487334	chr11	130536852		0.88	0.08	0.32	0.30	0.30	0.58	0.16	0.04
19	cg27603467	chr3	71276631	*FOXP1*	0.88	0.07	0.03	0.29	0.27	0.77	0.35	0.05
20	cg00580562	chr10	21789124	*DNAJC1*	0.88	0.08	0.45	0.11	0.41	0.85	0.05	0.08
21	cg02789362	chr6	151937260	*C6orf97*	0.87	0.23	0.45	0.34	0.49	0.81	0.24	0.47
22	cg09780180	chr16	73097903		0.87	0.06	0.18	0.16	0.46	0.67	0.04	0.18
23	cg17734243	chr6	16701191	*ATXN1*	0.87	0.04	0.16	0.26	0.45	0.81	0.06	0.25
24	cg03232056	chr7	157461310	*PTPRN2*	0.87	0.03	0.36	0.35	0.30	0.77	0.00	0.07

25	cg02244556	chr13	23721125		0.87	0.11	0.30	0.25	0.47	0.79	0.02	0.22
26	cg06844968	chr18	24131604	*KCTD1*	0.86	0.09	0.43	0.25	0.43	0.64	0.15	0.13
27	cg06881639	chr1	89752648		0.86	0.05	0.28	0.03	0.08	0.68	0.00	0.00
28	cg20318143	chr19	17516370	*BST2; MVB12A; BISPR*	0.86	0.10	0.26	0.36	0.37	0.74	0.03	0.00
29	cg11288670	chr2	54823611	*SPTBN1*	0.86	0.09	0.19	0.49	0.23	0.75	0.36	0.06
30	cg25500021	chr7	31990121	*PDE1C*	0.86	0.29	0.45	0.47	0.41	0.93	0.38	0.37
31	cg22393436	chr1	168489851		0.86	0.07	0.20	0.38	0.28	0.59	0.03	0.20
32	cg16672535	chr1	6510652	*ESPN*	0.86	0.14	0.15	0.09	0.33	0.78	0.15	0.00
33	cg24944092	chr12	12325234	*LRP6*	0.85	0.08	0.47	0.47	0.37	0.51	0.04	0.31
34	cg15703773	chr7	156813342	*LMBR1*	0.85	0.12	0.20	0.25	0.15	0.59	0.05	0.12
35	cg15007954	chr3	189348674	*TP63*	0.85	0.13	0.41	0.08	0.00	0.89	0.08	0.06
36	cg06864895	chr12	46767683	*SLC38A2*	0.85	0.05	0.48	0.21	0.15	0.73	0.00	0.06
37	cg15718448	chr8	111930817		0.85	0.07	0.44	0.48	0.25	0.76	0.03	0.08
38	cg04026169	chr4	85422265		0.84	0.34	0.41	0.41	0.44	0.71	0.30	0.38
39	cg12993990	chr1	224023540	*TP53BP2*	0.84	0.07	0.05	0.29	0.19	0.85	0.38	0.09
40	cg05021743	chr1	151810893	*C2CD4D; LOC100132111*	0.84	0.02	0.03	0.05	0.08	0.74	0.00	0.00
41	cg04511697	chr4	173707334	*GALNTL6*	0.84	0.09	0.02	0.44	0.25	0.50	0.05	0.16
42	cg20138749	chr1	1370013	*VWA1; LOC102724312*	0.83	0.07	0.36	0.17	0.00	0.50	0.00	0.08
43	cg12095712	chr6	53803789	*LOC101927189; LRRC1*	0.83	0.05	0.47	0.45	0.44	0.60	0.03	0.39
44	cg08424184	chr1	36043633	*TFAP2E*	0.83	0.08	0.10	0.33	0.20	0.77	0.00	0.08
45	cg02330078	chr5	163704924		0.83	0.16	0.35	0.49	0.43	0.60	0.23	0.15
46	cg02082588	chr3	130569372	*ATP2C1*	0.83	0.05	0.03	0.48	0.12	0.52	0.02	0.17
47	cg00056468	chr3	88489988		0.83	0.24	0.11	0.33	0.44	0.63	0.20	0.12
48	cg18517818	chr4	10956577		0.83	0.03	0.30	0.48	0.33	0.39	0.02	0.27
49	cg05861512	chr14	94442960	*SERPINA11*	0.83	0.11	0.04	0.39	0.20	0.68	0.06	0.02
50	cg26704078	chr18	24131115	*KCTD1*	0.83	0.21	0.00	0.20	0.23	0.72	0.10	0.09

51	cg16576285	chr12	719645	*NINJ2*	0.83	0.08	0.38	0.48	0.40	0.77	0.10	0.06
52	cg04296699	chr1	151810904	*C2CD4D; LOC100132111*	0.83	0.02	0.03	0.00	0.21	0.60	0.00	0.00
53	cg15413430	chr3	122017315	*ILDR1*	0.82	0.04	0.47	0.35	0.44	0.53	0.07	0.24
54	cg21429982	chr6	25138047	*CMAHP*	0.82	0.04	0.33	0.48	0.29	0.67	0.30	0.12
55	cg20316549	chr7	156810507	*LMBR1*	0.82	0.15	0.00	0.02	0.03	0.54	0.00	0.09
56	cg07585069	chr17	27039120	*PROCA1*	0.82	0.04	0.19	0.29	0.07	0.55	0.12	0.00
57	cg18910298	chr14	101471487		0.81	0.12	0.46	0.17	0.22	0.58	0.27	0.03
58	cg15015892	chr1	151810887	*C2CD4D; LOC100132111*	0.81	0.02	0.05	0.07	0.22	0.83	0.00	0.00
59	cg08703055	chr1	2232481	*SKI*	0.81	0.05	0.04	0.43	0.22	0.84	0.00	0.00
60	cg02805890	chr6	27778836		0.81	0.07	0.16	0.13	0.45	0.58	0.04	0.04
61	cg00449021	chr1	234658467		0.81	0.11	0.44	0.24	0.45	0.82	0.05	0.10
62	cg22028698	chr6	24229032	*DCDC2*	0.81	0.06	0.40	0.38	0.25	0.52	0.31	0.02
63	cg11920269	chr12	12107334		0.81	0.06	0.39	0.11	0.15	0.63	0.03	0.03
64	cg23015118	chr1	2232527	*SKI*	0.81	0.01	0.08	0.41	0.28	0.90	0.05	0.00
65	cg07188513	chr7	156807474	*LMBR1*	0.81	0.18	0.29	0.36	0.32	0.63	0.18	0.11
66	cg01413582	chr6	30652688	*KIAA1949*	0.81	0.03	0.00	0.30	0.21	0.71	0.05	0.00
67	cg27083040	chr5	148810061	*MIR145; LOC728264*	0.80	0.29	0.03	0.29	0.29	0.83	0.35	0.30
68	cg05267204	chr2	55360999	*CCDC88A*	0.80	0.07	0.15	0.01	0.27	0.66	0.41	0.10
69	cg09742918	chr22	39570729	*CACNA1I*	0.80	0.04	0.47	0.48	0.27	0.39	0.05	0.02
70	cg17564741	chr22	50060970	*MLC1*	0.80	0.05	0.24	0.33	0.28	0.45	0.00	0.00
71	cg12099669	chr7	134035671	*EXOC4*	0.80	0.11	0.12	0.42	0.34	0.72	0.04	0.06
72	cg11828180	chr1	41848985	*HIVEP3*	0.80	0.05	0.24	0.33	0.00	0.44	0.00	0.09
73	cg19255477	chr1	151810586	*LOC100132111; C2CD4D*	0.80	0.03	0.08	0.00	0.08	0.49	0.06	0.00
74	cg24044052	chr2	231191662	*SP140L*	0.79	0.05	0.32	0.00	0.10	0.65	0.20	0.00
75	cg16498647	chr8	21638104	*GFRA2*	0.79	0.15	0.33	0.35	0.45	0.32	0.13	0.09
76	cg07215736	chr12	121829653	*SETD1B*	0.79	0.10	0.48	0.02	0.21	0.59	0.18	0.06

77	cg11041686	chr2	6780727		0.78	0.11	0.00	0.09	0.41	0.68	0.06	0.10
78	cg04161194	chr18	42686603	*LIN00907*	0.78	0.06	0.21	0.38	0.26	0.64	0.09	0.16
79	cg21753092	chr2	238071771	*SCLY; UBE2F-SCLY*	0.77	0.08	0.47	0.43	0.15	0.37	0.13	0.30
80	cg08636385	chr6	71816668		0.77	0.12	0.05	0.46	0.40	0.53	0.23	0.07
81	cg01210113	chr16	11352835		0.77	0.14	0.25	0.10	0.18	0.79	0.12	0.06
82	cg23243378	chr11	65406311	*SIPA1*	0.77	0.08	0.24	0.16	0.38	0.74	0.33	0.03
83	cg06144905	chr17	27369780	*PIPOX*	0.77	0.31	0.46	0.13	0.24	0.49	0.15	0.11
84	cg10308253	chr6	149803111	*ZC3H12D*	0.77	0.08	0.02	0.45	0.25	0.58	0.23	0.02
85	cg01961252	chr12	54955143	*PDE1B*	0.77	0.24	0.04	0.44	0.24	0.65	0.04	0.04
86	cg12435551	chr1	118679290	*SPAG17*	0.76	0.06	0.00	0.43	0.20	0.66	0.10	0.00
87	cg15174393	chr12	125259081		0.76	0.07	0.02	0.14	0.29	0.46	0.21	0.03
88	cg05246303	chr2	182321354	*ITGA4*	0.76	0.14	0.23	0.24	0.36	0.29	0.03	0.08
89	cg15039365	chr12	54806893	*ITGA5; LOC102724050*	0.76	0.10	0.27	0.21	0.44	0.80	0.05	0.35
90	cg15402210	chr2	63274857	*LOC100132215; WDPCP*	0.76	0.03	0.43	0.39	0.33	0.76	0.00	0.09
91	cg26365776	chr10	23493947	*LINC01552*	0.75	0.12	0.05	0.20	0.00	0.44	0.08	0.00
92	cg14799482	chr1	90135931	*LRRC8C*	0.75	0.09	0.13	0.30	0.37	0.55	0.16	0.25
93	cg06492744	chr11	65406254	*SIPA1*	0.75	0.03	0.21	0.28	0.16	0.68	0.14	0.00
94	cg00308914	chr5	17436715		0.75	0.21	0.49	0.48	0.40	0.51	0.14	0.09
95	cg19246944	chr2	57926355	*VRK2*	0.75	0.06	0.43	0.40	0.18	0.42	0.00	0.14
96	cg02187406	chr5	159933519	*ADRA1B*	0.74	0.03	0.21	0.34	0.23	0.49	0.03	0.13
97	cg07991044	chr4	37782572		0.74	0.04	0.37	0.26	0.22	0.43	0.00	0.00
98	cg22463795	chr19	13215292	*CACNA1A*	0.74	0.04	0.35	0.47	0.39	0.56	0.05	0.11
99	cg15212266	chr1	21879311	*ALPL*	0.74	0.06	0.00	0.17	0.23	0.36	0.02	0.15
100	cg11122351	chr1	214731459		0.74	0.07	0.45	0.43	0.39	0.55	0.04	0.00
101	cg11543957	chr7	151536312	*PRKAG2*	0.73	0.03	0.25	0.33	0.40	0.50	0.00	0.06
102	cg14100181	chr3	120613445		0.73	0.05	0.06	0.36	0.11	0.26	0.00	0.02
103	cg13531667	chr5	112825132	*MCC*	0.73	0.04	0.03	0.48	0.14	0.51	0.02	0.33

104	cg03619352	chr5	179219098	*ADAMTS2*	0.73	0.15	0.19	0.43	0.28	0.47	0.15	0.03
105	cg12640176	chr13	59527967		0.72	0.05	0.05	0.29	0.18	0.37	0.00	0.03
106	cg02424365	chr16	73098356		0.72	0.15	0.04	0.26	0.28	0.68	0.00	0.04
107	cg03226454	chr1	198608163	*PTPRC*	0.72	0.06	0.36	0.25	0.48	0.35	0.02	0.13
108	cg20803910	chr17	6659375	*XAF1*	0.72	0.04	0.28	0.02	0.02	0.77	0.00	0.00
109	cg04211581	chr6	152011656	*ESR1*	0.72	0.08	0.10	0.37	0.23	0.27	0.09	0.08
110	cg03171003	chr2	111875934	*ANAPC1*	0.71	0.03	0.06	0.38	0.31	0.75	0.00	0.05
111	cg19282042	chr5	73872818	*ARHGEF28*	0.71	0.05	0.36	0.00	0.00	0.39	0.10	0.08
112	cg25951430	chr11	57192432	*SLC43A3*	0.71	0.04	0.15	0.35	0.16	0.63	0.03	0.08
113	cg06367135	chr1	3077835	*PRDM16*	0.71	0.13	0.06	0.48	0.11	0.66	0.19	0.17
114	cg12345460	chr9	131664788	*LRRC8A*	0.71	0.04	0.33	0.48	0.21	0.79	0.00	0.11
115	cg21583284	chr8	134517821	*ST3GAL1*	0.71	0.04	0.25	0.29	0.34	0.56	0.00	0.12
116	cg12441203	chr8	41121898	*SFRP1*	0.71	0.04	0.14	0.48	0.26	0.48	0.02	0.00
117	cg18614381	chr19	31844182		0.71	0.11	0.21	0.25	0.27	0.37	0.10	0.29
118	cg02852025	chr7	50348174	*IKZF1*	0.70	0.03	0.12	0.17	0.28	0.39	0.07	0.07
119	cg08144172	chr1	41849203	*HIVEP3*	0.70	0.12	0.19	0.14	0.00	0.85	0.03	0.05
120	cg07706375	chr19	13947646	*MIR23A; MIR24-2; MIR27A*	0.70	0.08	0.04	0.34	0.36	0.75	0.12	0.00
121	cg18848965	chr7	8402432		0.70	0.08	0.09	0.13	0.15	0.24	0.03	0.00
122	cg24570841	chr1	36043401	*TFAP2E*	0.69	0.03	0.05	0.04	0.22	0.50	0.00	0.03
123	cg22932808	chr7	41909842		0.69	0.14	0.12	0.34	0.44	0.22	0.17	0.15
124	cg02367708	chr2	136873789	*CXCR4*	0.69	0.07	0.11	0.25	0.12	0.10	0.00	0.10
125	cg21642108	chr8	25056560	*DOCK5*	0.69	0.05	0.40	0.15	0.29	0.64	0.06	0.08
126	cg13727629	chr1	2232469	*SKI*	0.69	0.05	0.05	0.47	0.16	0.71	0.00	0.00
127	cg18285788	chr7	157475692	*PTPRN2*	0.69	0.07	0.12	0.16	0.09	0.66	0.08	0.41
128	cg21603744	chr8	110020783		0.69	0.09	0.33	0.39	0.42	0.54	0.10	0.27
129	cg12265878	chr7	41995442		0.68	0.07	0.08	0.40	0.08	0.37	0.04	0.11
130	cg10611580	chr17	75319942	*SEPT9*	0.68	0.09	0.02	0.36	0.31	0.82	0.00	0.05
131	cg07838324	chr9	127048336	*NEK6*	0.68	0.05	0.32	0.06	0.11	0.70	0.04	0.07

132	cg01274170	chr19	19562378	*GATAD2A*	0.68	0.04	0.38	0.43	0.20	0.27	0.08	0.17
133	cg23974730	chr6	71022072		0.68	0.04	0.13	0.21	0.32	0.37	0.07	0.14
134	cg21393051	chr1	209585310		0.68	0.05	0.30	0.19	0.37	0.33	0.00	0.04
135	cg20047483	chr1	32732649	*LCK*	0.67	0.08	0.19	0.32	0.27	0.58	0.02	0.11
136	cg08273995	chr4	7526521	*SORCS2*	0.67	0.08	0.00	0.48	0.17	0.00	0.27	0.00
137	cg18107094	chr8	110020924		0.67	0.12	0.36	0.23	0.46	0.58	0.09	0.19
138	cg19978318	chr2	215430983	*VWC2L*	0.67	0.20	0.30	0.42	0.47	0.40	0.31	0.27
139	cg01592063	chr2	229338518		0.66	0.07	0.38	0.40	0.38	0.41	0.06	0.00
140	cg05226008	chr8	13372132	*DLC1*	0.66	0.13	0.01	0.49	0.34	0.35	0.18	0.11
141	cg23889684	chr22	36556813	*APOL3*	0.66	0.04	0.10	0.03	0.00	0.46	0.03	0.00
142	cg27391934	chr1	1369948	*VWA1*	0.66	0.09	0.11	0.41	0.11	1.00	0.10	0.18
143	cg15533448	chr15	66331737	*MIR4311; MEGF11*	0.66	0.25	0.00	0.18	0.23	0.75	0.16	0.15
144	cg11964154	chr1	225616987	*LBR*	0.66	0.09	0.36	0.21	0.40	0.30	0.03	0.13
145	cg17753212	chr17	6659562	*XAF1*	0.66	0.02	0.30	0.05	0.09	0.52	0.00	0.00
146	cg15272173	chr7	157585857	*PTPRN2*	0.65	0.35	0.29	0.20	0.27	0.68	0.23	0.43
147	cg20503768	chr18	45958576		0.65	0.09	0.24	0.30	0.34	0.52	0.09	0.33
148	cg04474049	chr7	116164533	*CAV1*	0.65	0.28	0.03	0.33	0.07	0.37	0.24	0.30
149	cg01141339	chr5	72530198		0.65	0.13	0.00	0.05	0.08	0.73	0.00	0.03
150	cg05135593	chr5	49726092	*EMB*	0.65	0.04	0.28	0.24	0.25	0.30	0.00	0.00
151	cg11132443	chr3	130569560	*ATP2C1*	0.65	0.04	0.02	0.24	0.05	0.38	0.03	0.03
152	cg25831439	chr2	161508799		0.64	0.03	0.44	0.39	0.30	0.46	0.00	0.14
153	cg18708013	chr6	149803206	*ZC3H12D*	0.64	0.07	0.06	0.47	0.29	0.45	0.09	0.03
154	cg07137934	chr5	180230911	*MGAT1*	0.64	0.08	0.38	0.20	0.24	0.61	0.00	0.00
155	cg12183993	chr19	36352005	*KIRREL2*	0.64	0.11	0.08	0.02	0.09	0.72	0.09	0.06
156	cg14972362	chr3	153346517		0.64	0.06	0.09	0.25	0.22	0.24	0.11	0.00
157	cg16220180	chr4	2789378		0.64	0.06	0.32	0.17	0.22	0.68	0.05	0.05
158	cg20713639	chr11	65185684		0.64	0.02	0.04	0.10	0.00	0.64	0.11	0.00
159	cg26747050	chr6	105153166		0.64	0.02	0.20	0.06	0.20	0.40	0.02	0.03

160	cg06697540	chr10	125532925	*CPXM2*	0.64	0.05	0.47	0.25	0.33	0.34	0.09	0.07
161	cg18241962	chr17	75385432	*SEPT9*	0.63	0.02	0.36	0.08	0.23	0.24	0.00	0.07
162	cg24725729	chr12	13741741	*GRIN2B*	0.63	0.06	0.11	0.43	0.18	0.40	0.06	0.15
163	cg06613765	chr17	8771055	*PIK3R6*	0.63	0.03	0.41	0.35	0.08	0.35	0.04	0.05
164	cg03819134	chr5	957584		0.62	0.02	0.22	0.18	0.33	0.55	0.00	0.03
165	cg19147390	chr19	4911058	*UHRF1*	0.62	0.12	0.33	0.48	0.10	0.50	0.00	0.08
166	cg10655396	chr19	1077244	*HMHA1*	0.62	0.07	0.47	0.13	0.31	0.74	0.00	0.00
167	cg09418290	chr1	247579319	*NLRP3*	0.62	0.02	0.37	0.48	0.38	0.55	0.00	0.13
168	cg08905567	chr14	76777176		0.62	0.05	0.06	0.23	0.10	0.12	0.00	0.00
169	cg20495843	chr6	106973166	*AIM1*	0.61	0.03	0.39	0.15	0.26	0.32	0.00	0.05
170	cg17175926	chr4	110280376		0.61	0.02	0.11	0.45	0.20	0.44	0.03	0.00
171	cg12645202	chr11	32442615	*WT1*	0.61	0.04	0.09	0.23	0.06	0.52	0.03	0.02
172	cg11953749	chr10	119590608		0.61	0.09	0.28	0.39	0.04	0.66	0.00	0.04
173	cg04951822	chr12	113345598	*OAS1*	0.61	0.20	0.09	0.00	0.18	0.50	0.17	0.10
174	cg24461557	chr6	43892169	*LINC01512*	0.61	0.03	0.03	0.38	0.40	0.65	0.00	0.22
175	cg24208921	chr3	107843436	*LINC01215*	0.60	0.03	0.32	0.38	0.05	0.32	0.03	0.00
176	cg14415629	chr11	33033607		0.60	0.05	0.27	0.23	0.47	0.45	0.00	0.24
177	cg12679308	chr7	101500014	*CUX1*	0.59	0.05	0.05	0.01	0.02	0.48	0.00	0.00
178	cg22074858	chr1	89488430	*GBP3*	0.59	0.04	0.26	0.20	0.15	0.35	0.03	0.03
179	cg05408831	chr5	179220797	*LTC4S*	0.59	0.02	0.00	0.43	0.29	0.60	0.07	0.00
180	cg16101156	chr2	6034874		0.59	0.11	0.33	0.37	0.44	0.38	0.07	0.23
181	cg06603374	chr12	54807242	*ITGA5; LOC102724050*	0.58	0.05	0.00	0.14	0.38	0.71	0.00	0.00
182	cg07385423	chr1	161697574	*FCRLB*	0.58	0.03	0.20	0.43	0.38	0.38	0.00	0.07
183	cg20836437	chr10	116561232		0.58	0.03	0.49	0.44	0.25	0.37	0.00	0.08
184	cg16549694	chr1	42204925	*HIVEP3*	0.58	0.05	0.14	0.20	0.07	0.45	0.00	0.00
185	cg09119967	chr20	9494838	*C20orf103*	0.58	0.13	0.20	0.00	0.23	0.24	0.04	0.10
186	cg23859630	chr6	32160455	*GPSM3*	0.58	0.05	0.33	0.27	0.00	0.42	0.17	0.00
187	cg25552098	chr1	31358506	*SDC3*	0.58	0.06	0.08	0.37	0.27	0.50	0.17	0.04

188	cg17095460	chr11	119468660		0.58	0.06	0.40	0.34	0.40	0.29	0.00	0.27
189	cg04841176	chr5	94363209	*MCTP1*	0.57	0.03	0.09	0.32	0.21	0.33	0.03	0.04
190	cg13492139	chr9	111063281		0.57	0.05	0.43	0.36	0.19	0.37	0.04	0.04
191	cg03317202	chr15	85873141		0.57	0.03	0.11	0.36	0.18	0.27	0.00	0.13
192	cg15898874	chr15	96515411		0.57	0.11	0.45	0.47	0.11	0.38	0.00	0.16
193	cg19641327	chr5	15834770	*FBXL7*	0.57	0.18	0.18	0.24	0.29	0.21	0.20	0.27
194	cg17276853	chr17	32964596	*TMEM132E*	0.56	0.01	0.18	0.46	0.16	0.41	0.00	0.15
195	cg27089606	chr19	40166804		0.56	0.04	0.45	0.36	0.37	0.49	0.03	0.20
196	cg02285920	chr19	17958956	*JAK3*	0.56	0.04	0.24	0.29	0.25	0.45	0.00	0.00
197	cg03983213	chr7	157475737	*PTPRN2*	0.56	0.02	0.05	0.18	0.16	0.67	0.13	0.28
198	cg03362174	chr2	26893188		0.56	0.08	0.20	0.06	0.00	0.71	0.14	0.00
199	cg20704394	chr7	104992726	*SRPK2*	0.55	0.03	0.05	0.44	0.00	0.29	0.00	0.00
200	cg11318133	chr12	113416268	*OAS2*	0.55	0.05	0.00	0.00	0.03	0.26	0.04	0.00
201	cg05493841	chr12	70083096	*BEST3*	0.55	0.07	0.02	0.36	0.18	0.39	0.03	0.02
202	cg13938774	chr9	117148097	*AKNA*	0.55	0.04	0.05	0.07	0.40	0.52	0.05	0.10
203	cg01254505	chr19	17516470	*BST2*	0.55	0.04	0.14	0.11	0.11	0.50	0.00	0.00
204	cg18454299	chr15	44959064	*PATL2*	0.55	0.03	0.37	0.45	0.40	0.19	0.16	0.15
205	cg21988119	chr19	17958937	*JAK3*	0.55	0.03	0.48	0.35	0.37	0.65	0.00	0.00
206	cg11683232	chr12	65063404	*RASSF3*	0.55	0.03	0.00	0.07	0.03	0.37	0.03	0.00
207	cg20918937	chr17	3443626	*TRPV3*	0.55	0.07	0.25	0.41	0.27	0.11	0.11	0.11
208	cg14003853	chr15	86250370	*AKAP13*	0.55	0.03	0.04	0.47	0.02	0.27	0.02	0.02
209	cg21689664	chr18	7525014		0.54	0.05	0.41	0.08	0.22	0.23	0.03	0.04
210	cg22507304	chr14	77709189	*TMEM63C*	0.54	0.06	0.40	0.36	0.23	0.30	0.00	0.15
211	cg23140245	chr8	121255387	*COL14A1*	0.54	0.02	0.00	0.44	0.17	0.49	0.00	0.12
212	cg05479471	chr14	106733112		0.54	0.06	0.31	0.16	0.31	0.15	0.00	0.18
213	cg15122111	chr3	130569453	*ATP2C1*	0.54	0.03	0.02	0.27	0.02	0.16	0.00	0.00
214	cg15177359	chr17	73840750	*UNC13D*	0.53	0.06	0.23	0.44	0.06	0.53	0.18	0.00
215	cg27613992	chr20	4350779		0.53	0.09	0.37	0.43	0.20	0.31	0.04	0.14
216	cg03378725	chr1	225438290	*DNAH14*	0.53	0.05	0.23	0.27	0.28	0.30	0.15	0.09

217	cg04385966	chr3	183273741	*KLHL6*	0.53	0.03	0.06	0.15	0.17	0.30	0.08	0.11
218	cg26940178	chr1	76082888		0.53	0.04	0.21	0.23	0.03	0.50	0.03	0.04
219	cg02744457	chr11	95374866		0.52	0.03	0.01	0.00	0.25	0.39	0.21	0.06
220	cg11657018	chr5	111115809	*C5orf13*	0.52	0.08	0.02	0.38	0.05	0.44	0.02	0.02
221	cg06596806	chr16	66550178	*TK2*	0.52	0.05	0.35	0.15	0.14	0.38	0.05	0.07
222	cg24680129	chr12	13318625		0.52	0.06	0.12	0.20	0.18	0.58	0.03	0.18
223	cg03368155	chr6	10206281		0.51	0.03	0.11	0.17	0.29	0.31	0.05	0.12
224	cg11883221	chr19	54352786		0.51	0.07	0.00	0.30	0.13	0.23	0.00	0.02
225	cg22999025	chr17	19977025		0.50	0.05	0.05	0.29	0.18	0.45	0.03	0.05
226	cg19769780	chr1	226863138	*ITPKB-IT1; ITPKB*	0.50	0.03	0.39	0.24	0.08	0.26	0.03	0.00
227	cg06483857	chr5	88692528		0.50	0.03	0.02	0.20	0.06	0.27	0.02	0.03
228	cg19408145	chr1	160681530	*CD48*	0.50	0.04	0.35	0.41	0.35	0.33	0.02	0.12

Table A.4: The differentially hypermethylated CpG sites in the islet genome.

	CpG ID	chromosome	position	Gene Name	islet	PBMC	aorta	liver	lung	pancreas	thymus	spleen
1	cg10469980	chr16	85660101	*KIAA0182*	0.25	0.83	0.94	0.83	0.76	0.06	1	0.88
2	cg17546376	chr3	42172097	*TRAK1*	0.23	0.93	0.93	0.65	0.81	0.15	0.96	0.89
3	cg27580905	chr2	39335085	*SOS1*	0.46	0.88	0.65	0.81	0.81	0.37	0.94	0.86
4	cg03628365	chr6	47432121		0.21	0.85	0.84	0.6	0.71	0.17	0.93	0.93
5	cg11802553	chr14	65203330	*PLEKHG3*	0.27	0.9	0.6	0.61	0.86	0.12	0.95	0.77
6	cg00866541	chr16	85395661		0.36	0.89	0.89	0.56	0.67	0.3	0.95	0.88
7	cg02261780	chr19	3097728	*GNA11*	0.43	0.98	0.7	0.79	0.63	0.18	1	0.85
8	cg09934926	chr4	111115815	*ELOVL6*	0.2	0.92	0.93	0.53	0.83	0.24	0.98	0.96
9	cg18603224	chr5	33945764	*SLC45A2*	0.21	0.81	0.72	0.64	0.62	0.12	0.74	0.74
10	cg23520688	chr14	103294740	*TRAF3*	0.26	0.92	0.86	0.84	0.7	0.23	0.97	0.92
11	cg03395247	chr20	42991330	*HNF4A*	0.42	0.84	0.79	0.94	0.67	0.29	0.95	0.65
12	cg06849167	chr11	61449886	*DAGLA*	0.22	0.94	0.79	0.83	0.79	0.09	1	0.86
13	cg24214659	chr12	117534674	*TESC*	0.2	0.95	0.64	0.52	0.7	0.11	0.92	0.82
14	cg25576801	chr10	49932245	*WDFY4*	0.45	0.85	0.92	0.87	0.62	0.37	0.97	0.69
15	cg07971156	chr10	114779609	*TCF7L2*	0.18	0.9	0.95	0.65	0.59	0.1	0.9	0.88
16	cg06000659	chr3	168235595	*EGFEM1P*	0.16	0.72	0.91	0.52	0.93	0.15	0.97	0.88
17	cg06830529	chr9	124960107	*MORN5*	0.32	0.79	0.67	0.69	0.67	0.22	0.96	0.92
18	cg15426815	chr12	7071712	*MIR-200C*	0.31	0.86	0.92	0.97	0.63	0.13	0.88	0.7
19	cg16887521	chr10	76784798	*KAT6B*	0.19	0.88	0.61	0.81	0.54	0.08	0.96	0.83
20	cg00333703	chr7	123384406	*WASL*	0.31	0.9	0.92	0.63	0.83	0.33	0.98	0.89
21	cg00053338	chr5	24714819		0.45	0.88	0.87	0.66	0.76	0.22	0.9	0.78
22	cg11432036	chr2	226335995	*NYAP2*	0.23	0.87	0.81	0.76	0.69	0.15	0.98	0.91
23	cg23008238	chr5	2112459		0.13	0.93	0.8	0.94	0.93	0.18	0.97	0.98
24	cg06635849	chr5	121755088	*SNCAIP*	0.2	0.85	0.95	0.61	0.57	0.23	0.92	0.93

25	cg04340203	chr19	891986	*MED16*	0.2	0.86	0.93	0.85	0.63	0	1	0.7
26	cg18318843	chr8	42354362	*SLC20A2*	0.23	0.91	0.56	0.63	0.76	0.35	0.85	0.88
27	cg00698688	chr19	49055432	*SULT2B1*	0.16	0.86	0.83	0.63	0.54	0	0.67	0.89
28	cg08549388	chr11	73780601	*C2CD3*	0.24	0.93	0.89	1	0.71	0.1	0.96	0.94
29	cg10097015	chr2	55336607		0.2	0.88	0.92	1	0.8	0.07	0.98	0.9
30	cg18630178	chr14	93154484	*RIN3*	0.34	0.98	0.57	0.82	0.86	0.15	0.94	0.92
31	cg09299592	chr22	40765943	*SGSM3*	0.38	0.8	0.65	0.74	0.56	0.36	0.86	0.83
32	cg14058329	chr7	27183946	*HOXA5*	0.37	0.95	0.76	0.68	0.51	0.32	0.96	1
33	cg16117653	chr19	4891192	*ARRDC5*	0.44	0.93	0.73	0.63	0.8	0.3	0.96	1
34	cg09010597	chr17	78945313		0.28	0.83	0.84	0.57	0.86	0.09	0.89	0.69
35	cg02933228	chr11	64611087	*CDC42BPG*	0.41	0.84	0.97	0.88	0.57	0.16	0.83	0.65
36	cg21819409	chr8	120981915	*DEPTOR*	0.22	0.92	0.71	0.6	0.71	0.09	0.96	0.86
37	cg04564944	chr2	45558218		0.21	0.86	0.84	0.67	0.64	0.18	0.98	0.98
38	cg09023515	chr3	185120477	*MAP3K13*	0.31	0.89	0.92	0.88	0.88	0.16	0.96	0.91
39	cg18048027	chr2	130502532		0.38	0.92	0.57	0.57	0.7	0.27	0.97	0.86
40	cg06464468	chr11	65414413	*SIPA1*	0.41	0.9	0.72	0.94	0.87	0.5	0.92	0.94
41	cg17540105	chr6	151634287	*AKAP12*	0.18	0.88	0.86	0.74	0.8	0.13	0.88	0.6
42	cg01300791	chr16	30448949		0.33	0.88	0.69	0.62	0.64	0.17	0.97	0.86
43	cg04772968	chr19	12404112	*ZNF44*	0.44	0.93	0.7	0.51	0.72	0.24	0.97	0.86
44	cg22280402	chr7	143596849	*FAM115A*	0.3	0.84	0.73	0.57	0.79	0.24	0.98	0.89
45	cg27473972	chr6	140864365		0.22	0.84	0.85	0.62	0.8	0.22	1	0.83
46	cg03377039	chr2	214150956	*SPAG16*	0.2	0.86	0.79	0.94	0.96	0.16	0.93	0.91
47	cg01998947	chr7	45500384		0.27	0.91	0.88	0.64	0.85	0.21	0.96	0.86
48	cg09583041	chr15	60245804		0.1	0.94	0.6	0.53	0.51	0.02	1	0.8
49	cg21475076	chr1	1100493		0.17	0.96	0.88	0.83	0.89	0.2	1	0.96
50	cg14527983	chr20	44562104	*PCIF1*	0.45	0.95	0.96	1	0.98	0.21	1	0.89
51	cg14285795	chr17	71179222		0.33	0.9	0.69	0.89	0.82	0.21	0.97	0.76
52	cg03540794	chr5	2112109		0.32	0.97	0.86	0.93	0.96	0.19	0.88	0.87
53	cg12776287	chr18	24125939	*KCTD1*	0.15	0.88	0.65	0.75	0.81	0.08	0.97	0.6

54	cg20331074	chr6	154936706		0.19	0.89	0.77	0.69	0.67	0.13	0.94	0.79
55	cg22165171	chr7	139763574	*PARP12*	0.45	0.91	0.62	0.67	0.63	0.21	0.96	0.79
56	cg05327192	chr11	75133593	*KLHL35*	0.22	0.9	0.85	0.51	0.71	0.26	1	0.98
57	cg06562682	chr19	35607199	*FXYD3*	0.27	0.77	0.83	0.75	0.64	0.2	0.94	0.92
58	cg22846423	chr3	177089912		0.25	0.92	0.93	0.86	0.56	0.16	1	0.91
59	cg10112854	chr16	80839183	*CDYL2*	0.49	0.83	0.57	0.66	0.67	0.53	1	0.81
60	cg16852463	chr1	15523893	*TMEM51*	0.28	0.85	0.82	0.63	0.81	0.15	1	0.95
61	cg12092005	chr14	99532847		0.37	0.88	0.88	0.83	0.79	0.32	1	0.83
62	cg08801866	chr21	46323808	*ITGB2*	0.42	0.93	0.61	0.91	0.74	0.58	0.92	0.96
63	cg18631905	chr8	76723432		0.09	0.83	0.58	0.54	0.58	0	0.76	0.7
64	cg02327604	chr7	808834	*HEATR2*	0.37	0.78	0.57	0.83	0.69	0.37	0.86	0.7
65	cg19147912	chr17	17596386	*RAI1*	0.46	0.88	0.78	0.66	0.59	0.32	0.76	0.74
66	cg16205506	chr3	172266828		0.2	0.92	0.87	0.7	0.58	0.11	0.97	0.96
67	cg23893406	chr1	81918997		0.36	0.91	0.88	0.62	0.75	0.25	0.95	0.94
68	cg10374063	chr13	93496028	*GPC5*	0.12	0.91	0.73	0.86	0.68	0.14	0.93	0.74
69	cg06445944	chr5	67586928	*PIK3R1*	0.15	0.9	0.85	0.51	0.64	0.09	0.98	0.89
70	cg23822445	chr12	107699855		0.26	0.9	0.85	0.72	0.89	0.3	1	0.91
71	cg11929042	chr20	19222240	*SLC24A3*	0.43	0.92	0.82	0.72	0.55	0.25	1	0.75
72	cg09879666	chr3	157218259	*VEPH1*	0.32	0.78	0.88	0.61	0.52	0.13	0.88	0.66
73	cg27104398	chr22	43343739	*PACSIN2*	0.3	0.92	0.92	0.67	0.79	0.09	0.93	0.8
74	cg13917433	chr3	52722445	*GNL3; SNORD19*	0.16	0.93	0.88	0.52	0.83	0.24	0.88	0.89
75	cg01485975	chr5	156486186	*HAVCR1*	0.19	0.88	0.7	0.61	0.71	0.11	0.9	0.91
76	cg01090766	chr2	232660830	*COPS7B*	0.18	0.93	0.53	0.87	0.83	0.17	0.93	0.87
77	cg19521573	chr8	74226253	*RDH10*	0.18	0.9	0.83	0.8	0.62	0.09	0.89	0.82
78	cg19738924	chr18	24126072	*KCTD1*	0.38	0.89	0.67	0.9	0.9	0.16	0.96	0.83
79	cg14435796	chr5	55529207	*ANKRD55*	0.24	0.92	0.58	0.63	0.7	0.18	0.93	0.79
80	cg02275413	chr12	101253527	*ANO4*	0.18	0.87	0.66	0.56	0.6	0.21	0.96	0.78
81	cg09959931	chr4	671550	*MYL5*	0.38	0.85	0.81	0.8	0.88	0.13	1	0.78
82	cg22197141	chr13	103552708		0.28	0.94	0.96	0.6	0.83	0.12	0.96	0.98

83	cg19176928	chr2	139378922		0.27	0.9	0.91	0.64	0.64	0.24	0.98	0.74
84	cg24345224	chr4	39593037	*UGDH-AS1; SMIM14*	0.5	0.93	0.53	0.65	0.73	0.21	0.92	0.94
85	cg05245094	chr16	85669572	*KIAA0182*	0.46	0.82	0.97	0.68	0.91	0.42	0.94	0.88
86	cg05780620	chr2	216139649		0.33	0.82	0.84	0.52	0.52	0.17	0.88	0.73
87	cg14698646	chr17	46684750	*HOXB7*	0.5	0.91	0.77	0.52	0.73	0.11	1	0.9
88	cg13362636	chr2	82289309		0.33	0.86	0.81	0.82	0.87	0.19	0.82	0.83
89	cg07815009	chr8	95220797	*CDH17*	0.4	0.97	0.87	0.56	0.78	0.24	0.92	0.92
90	cg01621034	chr6	26190175	*HIST1H4D*	0.39	0.93	0.85	0.91	0.85	0.2	0.98	1
91	cg18621357	chr12	65204725		0.49	0.89	0.81	0.86	0.91	0.44	1	0.98
92	cg03207666	chr7	27183950	*HOXA5*	0.47	0.89	0.81	0.74	0.53	0.43	0.95	1
93	cg11489916	chr18	48320337		0.19	0.94	0.95	0.95	0.95	0.19	0.98	1
94	cg21151735	chr12	64172919	*TMEM5*	0.5	0.92	0.58	0.68	0.84	0.29	0.92	0.81
95	cg05859179	chr20	11963914		0.3	0.93	0.81	0.54	0.95	0.28	0.98	0.97
96	cg06179152	chr16	1069368		0.16	0.93	0.59	0.71	0.52	0.08	1	0.59
97	cg00823303	chr7	129753942	*KLHDC10*	0.2	0.9	0.83	0.54	0.75	0.14	0.99	0.72
98	cg08885197	chr1	10747107	*CASZ1*	0.21	0.86	0.74	0.63	0.53	0.42	1	0.94
99	cg09097152	chr6	25411727	*LRRC16A*	0.36	0.88	0.97	1	0.8	0.27	0.97	0.88
100	cg06532789	chr5	159008967		0.16	0.83	0.51	0.6	0.61	0.1	0.94	0.84
101	cg11055649	chr6	31868764	*ZBTB12; C2*	0.36	0.97	0.9	0.64	0.76	0.31	0.94	0.96
102	cg21270847	chr1	164604806	*PBX1*	0.35	0.97	0.93	0.91	0.96	0.63	0.89	0.95
103	cg04584649	chr18	10525075	*NAPG*	0.21	0.75	0.64	0.73	0.67	0.14	0.96	0.67
104	cg05923197	chr14	54418804	*BMP4*	0.45	0.97	0.81	0.75	0.51	0.23	0.77	0.76
105	cg22663604	chr1	156030671	*RAB25*	0.25	0.93	0.53	0.83	0.63	0.27	0.87	0.85
106	cg00128386	chr16	85665539	*KIAA0182*	0.46	0.95	0.9	0.82	0.87	0.19	1	0.97
107	cg06518884	chr3	185822021	*ETV5*	0.44	0.88	0.97	0.96	0.71	0.38	0.98	0.91
108	cg16584696	chr10	134875826		0.39	0.87	0.93	0.89	0.89	0.41	0.87	0.93
109	cg02482497	chr1	153610672	*C1orf77*	0.09	0.91	0.87	0.88	0.74	0.08	1	0.81
110	cg17744883	chr5	2279386		0.25	0.84	0.56	1	0.71	0.23	0.85	0.91
111	cg03680996	chr6	10725534	*TMEM14C*	0.46	0.87	0.69	0.95	0.71	0.29	0.94	0.88

112	cg25271413	chr7	115274698		0.25	0.85	0.59	0.67	0.75	0.17	0.91	0.74
113	cg19449588	chr16	11841810		0.34	0.89	0.83	0.55	0.78	0.17	1	0.98
114	cg17624315	chr2	202289200	*TRAK2*	0.23	0.88	0.66	0.53	0.55	0.08	0.98	0.74
115	cg12753366	chr10	696321	*DIP2C; C10orf108*	0.33	0.98	0.65	0.9	0.78	0.31	0.96	0.97
116	cg15026620	chr11	114081012	*ZBTB16*	0.3	0.93	0.67	0.75	0.75	0.33	0.9	0.87
117	cg14945501	chr21	29765810		0.22	0.76	0.69	0.85	0.57	0.21	0.98	0.77
118	cg15655180	chr20	36831960		0.3	0.87	0.65	0.69	0.78	0.14	0.91	0.88
119	cg10321146	chr19	6822469	*VAV1*	0.35	0.9	0.6	0.69	0.56	0.18	0.92	0.93
120	cg23206888	chr2	242833114		0.41	0.96	1	0.98	0.95	0.5	1	0.89
121	cg09479341	chr17	40119957	*CNP*	0.31	0.87	0.68	0.78	0.68	0.08	0.89	0.67
122	cg24145118	chr10	2777041		0.04	0.97	0.83	0.69	0.7	0.06	0.86	0.79
123	cg00611395	chr16	86696049		0.17	0.81	0.53	0.65	0.57	0.06	0.8	0.71
124	cg16279954	chr13	45444906		0.27	0.9	0.95	0.89	0.76	0.07	0.92	0.64
125	cg27420224	chr20	43029777	*HNF4A*	0.32	0.97	0.8	0.55	0.83	0.36	1	0.85
126	cg18794321	chr7	39446233	*POU6F2-AS1; POU6F2*	0.22	0.84	0.8	0.52	0.7	0.14	0.89	0.9
127	cg07954572	chr12	117484521	*TESC*	0.38	0.73	0.73	0.67	0.76	0.24	0.92	0.85
128	cg05177409	chr7	24620338	*MPP6*	0.25	0.85	0.85	1	0.55	0.27	0.99	0.84
129	cg00722323	chr19	31883537		0.36	0.83	0.59	0.7	0.67	0.31	0.87	0.83
130	cg05422457	chr16	73266282		0.22	0.89	0.95	0.52	0.94	0.22	0.97	0.98
131	cg08488569	chr4	1749241		0.5	0.92	0.88	0.54	0.61	0.44	1	0.86
132	cg19455171	chr14	100943011	*WDR25; WDR25*	0.34	0.87	0.78	0.94	0.84	0.36	0.96	0.83
133	cg15707050	chr11	77706483	*INTS4*	0.26	0.87	0.59	0.63	0.84	0.06	1	0.85
134	cg26206456	chr8	992484		0.27	0.87	0.66	0.68	0.62	0.13	0.76	0.79
135	cg00920423	chr15	58505242		0.22	0.92	0.67	0.59	0.76	0.11	0.98	0.79
136	cg06062600	chr10	73977495	*ANAPC16*	0.29	0.92	0.73	0.57	0.6	0.15	0.93	0.73
137	cg04853831	chr3	63178269		0.21	0.87	0.66	0.58	0.88	0.27	0.87	0.88
138	cg04331189	chr2	160142303	*WDSUB1*	0.27	0.9	0.81	0.73	0.78	0.11	0.93	0.87
139	cg25076325	chr11	118135407	*MPZL2*	0.48	0.93	0.73	0.64	0.65	0.56	0.9	0.96
140	cg18090390	chr1	56880606		0.26	0.9	0.67	0.51	0.74	0.11	0.97	0.89

141	cg22162694	chr13	24323408	*MIPEP*	0.26	0.92	0.62	0.62	0.55	0.03	0.97	0.9
142	cg05460314	chr1	56880573		0.2	0.92	0.89	0.52	0.79	0.15	0.95	0.92
143	cg03610849	chr11	129080113		0.16	0.88	0.83	0.67	0.68	0.16	0.91	0.82
144	cg15869931	chr6	141347930		0.24	0.69	0.95	0.78	0.55	0.11	0.92	0.74
145	cg13704328	chr20	10845549		0.31	0.82	0.63	0.86	0.55	0.06	1	0.93
146	cg05444103	chr3	80745928		0.24	0.86	0.83	0.8	0.68	0.14	0.93	0.78
147	cg26456435	chr16	31053788		0.45	0.84	0.7	0.94	1	0.69	0.97	0.79
148	cg19835136	chr19	17650995	*FAM129C*	0.43	0.85	0.53	0.68	0.71	0.65	0.8	0.94
149	cg00386007	chr10	50143337	*WDFY4*	0.39	0.9	0.63	0.75	0.68	0.24	1	0.8
150	cg01951142	chr1	33787084	*A3GALT2*	0.33	0.89	0.8	0.54	0.7	0.17	0.88	0.78
151	cg05283521	chr7	157259343		0.14	0.94	0.94	0.96	0.93	0.42	0.89	0.96
152	cg04086239	chr16	24067174	*PRKCB*	0.46	0.9	0.59	0.67	0.85	0.46	1	0.97
153	cg09609106	chr1	204062352	*SOX13*	0.18	0.74	1	0.52	0.92	0.15	1	0.94
154	cg06169135	chr2	231825892		0.48	0.75	0.83	0.75	0.76	0.43	0.71	0.8
155	cg13007502	chr2	208795771	*PLEKHM3*	0.33	0.97	0.98	0.83	0.77	0.2	0.96	0.98
156	cg07665877	chr4	189456103	*LINC01060*	0.28	0.93	0.87	1	0.68	0.07	0.97	0.97
157	cg08937768	chr4	3342519	*RGS12*	0.5	0.88	0.91	0.67	0.68	0.65	0.87	0.97
158	cg11647944	chr12	5603148	*NTF3*	0.27	0.91	0.75	0.72	0.63	0.14	1	0.94
159	cg20647386	chr8	120981885	*DEPTOR*	0.09	0.93	0.66	0.53	0.7	0.07	0.98	0.85
160	cg05391892	chr2	208795620	*PLEKHM3*	0.2	0.97	0.67	0.72	0.85	0.17	0.96	0.89
161	cg27100471	chr11	794834	*SLC25A22*	0.35	0.98	0.9	0.54	0.53	0.13	0.91	0.89
162	cg25610697	chr9	676996	*KANK1*	0.3	0.93	0.53	0.63	0.67	0.11	0.98	0.98
163	cg08663896	chr17	40064311	*ACLY*	0.26	0.8	0.75	0.78	1	0.26	0.96	0.88
164	cg23613317	chr5	154092109	*LARP1*	0.19	0.92	0.93	0.88	0.59	0.05	0.97	0.82
165	cg10143991	chr6	2962408	*SERPINB6*	0.24	0.78	0.93	0.79	0.64	0.12	0.97	0.85
166	cg14744743	chr18	13871595		0.45	0.72	0.58	0.56	0.66	0.33	0.84	0.72
167	cg18967533	chr19	51472725	*KLK6*	0.46	0.75	0.73	0.69	0.88	0.52	0.82	0.88
168	cg01914045	chr21	31588584	*CLDN8*	0.22	0.78	0.85	0.54	0.55	0.23	0.78	0.79
169	cg23067082	chr12	7073179	*MIR-141*	0.17	0.78	0.8	0.75	0.73	0.09	1	0.73

170	cg19698648	chr7	150759376	*SLC4A2*	0.3	0.89	0.59	0.64	0.69	0.13	0.95	0.9
171	cg00366037	chr10	76781121	*MYST4*	0.21	0.89	0.8	0.81	0.68	0.14	0.97	0.92
172	cg15431338	chr15	25568711		0.25	0.88	0.76	0.66	0.68	0.23	0.98	0.86
173	cg09480190	chr16	85671771	*KIAA0182*	0.43	0.97	0.76	0.88	0.67	0.53	1	0.98
174	cg03235759	chr1	121264025	*EMBP1*	0.18	0.91	0.79	0.58	0.57	0.11	1	0.72
175	cg08329684	chr16	4932620	*PPL*	0.3	0.91	0.74	0.67	0.67	0.33	1	0.95
176	cg02442572	chr7	1303351		0.26	0.93	0.96	0.88	0.87	0.34	1	0.94
177	cg25390025	chr17	64225568	*APOH*	0.13	0.81	0.94	0.55	0.77	0.24	0.98	1
178	cg13553204	chr1	26504548	*CNKSR1*	0.17	0.8	0.9	0.83	0.66	0.1	1	0.81
179	cg03962019	chr1	41807865		0.29	0.77	0.82	0.81	0.67	0.39	1	0.8
180	cg19755928	chr3	128723491	*CCDC48*	0.31	0.92	0.91	1	0.51	0.33	0.96	0.74
181	cg04493700	chr10	126827067	*CTBP2*	0.21	0.92	0.87	0.57	0.61	0.16	0.97	0.8
182	cg17616537	chr17	46628663	*HOXB3*	0.37	0.87	0.94	0.83	0.72	0.1	0.95	0.91
183	cg26806260	chr15	62356512		0.25	0.92	0.75	0.83	0.61	0.15	0.93	0.89
184	cg19794481	chr12	7073240	*MIR-141*	0.24	0.94	0.93	0.92	0.59	0.17	0.95	0.92
185	cg12031809	chr5	132074036	*KIF3A*	0.33	0.77	0.77	0.81	0.54	0.11	0.87	0.81
186	cg01835489	chr12	53299310	*KRT8*	0.29	0.92	0.91	0.53	0.71	0.17	0.97	0.92
187	cg00298399	chr1	17697934	*PADI6*	0.47	0.86	0.69	0.59	0.57	0.39	0.87	0.84
188	cg13354219	chr3	143437525	*SLC9A9*	0.47	0.93	0.85	0.77	0.69	0.26	0.97	0.93
189	cg24154336	chr11	64659044	*MIR192; MIR194-2*	0.44	0.95	0.94	0.51	0.85	0.24	1	0.93
190	cg22843428	chr17	48620231	*EPN3*	0.39	0.87	0.73	0.72	0.75	0.24	0.95	0.89
191	cg03461013	chr2	38374414	*CYP1B1-AS1*	0.43	0.73	0.57	0.84	0.7	0.33	0.94	0.88
192	cg16992440	chr12	53777383	*SP1*	0.35	0.74	0.62	0.66	0.55	0.08	0.86	0.68
193	cg01218903	chr17	74911481	*MGAT5B*	0.42	0.97	0.98	0.89	0.75	0.47	1	0.92
194	cg17785018	chr19	17644601	*FAM129C*	0.42	0.79	0.59	0.58	0.65	0.07	1	0.71
195	cg15404473	chr8	30421394	*RBPMS*	0.37	0.91	0.75	0.94	0.89	0.33	1	1
196	cg23982489	chr12	55979609		0.21	0.89	0.82	0.69	0.79	0.27	0.95	0.93
197	cg24296440	chr4	99946235	*METAP1*	0.33	0.91	0.76	0.87	0.79	0.3	1	0.9
198	cg15289899	chr11	2397255	*CD81*	0.26	0.95	0.99	0.51	0.77	0.14	0.98	0.97

199	cg00131213	chr19	39105043	*MAP4K1*	0.3	0.82	0.68	0.57	0.71	0.1	0.9	0.92
200	cg09627060	chr10	134875916		0.28	0.84	0.93	0.75	0.84	0.21	0.91	0.91
201	cg24702147	chr12	7071814	*MIR-141; MIR-200C*	0.12	0.96	0.52	0.8	0.56	0.12	0.78	0.76
202	cg03255783	chr1	6487540	*ESPN*	0.49	0.78	0.74	0.7	0.73	0.38	0.9	0.77
203	cg04568522	chr9	97273915		0.17	0.84	0.9	0.72	0.65	0.19	0.91	0.75
204	cg14151930	chr7	83652513	*SEMA3A*	0.18	0.92	0.74	0.81	0.65	0.13	0.89	0.83
205	cg01651311	chr11	117704570	*FXYD6-FXYD2*	0.2	0.94	0.69	0.54	0.68	0.09	0.9	0.77
206	cg09978753	chr10	14709476	*FAM107B*	0.47	0.91	0.88	0.9	0.91	0.45	0.98	0.92
207	cg08287262	chr4	146652027	*C4orf51*	0.25	0.91	0.86	0.51	0.72	0.18	0.97	0.9
208	cg11371119	chr6	167321666		0.31	0.92	0.76	0.72	0.74	0.2	1	0.8
209	cg15203905	chr1	2845514		0.32	0.82	0.93	0.69	0.65	0.28	0.89	0.85
210	cg23142048	chr14	103691563		0.36	0.71	0.67	0.85	0.69	0.18	0.91	0.73
211	cg18792131	chr1	1099166		0.17	0.97	0.94	0.91	0.68	0.19	0.94	0.97
212	cg18037117	chr4	186732207	*SORBS2*	0.1	0.9	0.78	0.54	0.68	0.05	0.8	0.83
213	cg21304163	chr19	35607221	*FXYD3*	0.27	0.88	0.97	0.88	0.78	0.12	1	0.96
214	cg22412649	chr13	24736569	*SPATA13; MIR-2276*	0.33	0.88	0.92	0.71	0.81	0.25	1	0.87
215	cg22942897	chr1	3680345	*CCDC27*	0.39	0.84	0.85	0.86	0.88	0.21	0.97	0.95
216	cg22115892	chr4	177198411		0.25	0.96	0.97	0.89	0.92	0.1	0.98	0.95
217	cg00207921	chr1	1098992		0.19	0.79	0.86	0.86	0.62	0.18	0.9	0.84
218	cg09088720	chr5	154091976	*LARP1*	0.11	0.78	0.76	0.75	0.52	0.12	0.9	0.8
219	cg07716832	chr2	46077837	*PRKCE*	0.2	0.88	0.87	0.71	0.71	0.25	0.89	0.92
220	cg19928703	chr13	30143971	*SLC7A1*	0.25	0.87	0.53	0.53	0.67	0.11	0.94	0.88
221	cg06202269	chr7	152272077		0.46	0.86	0.89	0.98	0.94	0.32	0.95	0.88
222	cg06733198	chr10	10504299	*LOC101928322*	0.4	0.91	0.96	0.54	0.8	0.22	0.89	0.96
223	cg13958794	chr19	7483010	*ARHGEF18*	0.37	0.83	0.51	0.84	0.66	0.39	0.98	0.9
224	cg10916196	chr12	43966266		0.31	0.91	0.64	0.79	0.53	0.15	0.81	0.78
225	cg00340024	chr1	92907439		0.39	0.9	0.83	0.85	0.8	0.2	1	0.94
226	cg07335712	chr5	81682829		0.23	0.93	0.71	0.6	0.67	0.13	0.98	0.94
227	cg11975222	chr3	69228976	*FRMD4B*	0.35	0.92	0.61	0.53	0.53	0.39	1	0.88

228	cg14819525	chr20	11113825		0.41	0.94	0.84	1	0.82	0.36	1	0.93
229	cg22154447	chr13	78550069	*EDNRB*	0.42	0.78	0.81	0.56	0.68	0.29	0.88	0.82
230	cg03554522	chr4	170556187	*CLCN3*	0.28	0.9	0.93	1	0.77	0.17	0.98	0.98
231	cg02217206	chr8	94509022	*LINC00535*	0.27	0.92	0.91	0.76	0.73	0.23	0.91	0.79
232	cg06912665	chr5	154092290	*LARP1*	0.21	0.89	0.75	0.68	0.64	0.07	0.95	0.81
233	cg14053169	chr1	92907586		0.3	0.86	0.8	0.74	0.88	0.2	1	0.98
234	cg03895045	chr14	39703160	*MIA2*	0.16	0.92	0.71	1	0.7	0.16	0.93	0.93
235	cg17112518	chr14	69223221		0.21	0.87	0.91	0.6	0.72	0.25	0.91	0.89
236	cg09381559	chr1	201567256		0.15	0.75	0.85	0.68	0.51	0.07	0.87	0.74
237	cg20385489	chr5	95069104	*RHOBTB3*	0.23	0.92	0.95	0.69	0.85	0.12	0.95	0.9
238	cg14840561	chr12	121418741	*HNF1A*	0.13	0.96	1	0.56	0.79	0.11	0.81	1
239	cg04151661	chr14	99603265		0.4	0.94	1	0.86	0.96	0.27	1	0.83
240	cg21924314	chr1	99636927		0.2	0.91	0.75	0.73	0.78	0.15	0.95	0.85
241	cg06876892	chr16	833231	*MSLNL*	0.25	0.92	0.64	0.6	0.55	0.19	0.92	0.69
242	cg05275001	chr11	64497598	*RASGRP2*	0.35	0.91	0.91	0.87	0.76	0.11	0.94	0.65
243	cg22342893	chr3	10150224	*FANCD2OS*	0.49	0.88	0.56	0.92	0.83	0.37	0.91	0.81
244	cg25335876	chr2	16894580		0.21	0.75	0.75	0.86	0.8	0.2	0.87	0.69
245	cg24635930	chr5	38148572		0.23	0.92	0.67	0.76	0.71	0.16	0.9	0.91
246	cg11023565	chr6	47428069		0.28	0.93	0.95	0.56	0.71	0.18	0.94	0.96
247	cg03519907	chr5	154092426	*LARP1*	0.19	0.87	0.95	0.83	0.68	0.08	0.89	0.82
248	cg10427062	chr5	60728124	*ZSWIM6*	0.16	0.9	0.8	0.59	0.59	0.18	0.96	0.76
249	cg17470674	chr5	1702470		0.24	0.96	0.86	0.9	0.86	0.11	0.92	0.81
250	cg23659134	chr10	106035168	*GSTO2*	0.27	0.75	0.65	0.64	0.65	0.16	0.79	0.86
251	cg25567444	chr1	223018067	*DISP1*	0.25	0.92	0.86	0.88	0.87	0.29	0.98	0.96
252	cg10911997	chr11	34653201	*EHF*	0.33	0.85	0.94	0.53	0.81	0.23	0.97	0.93
253	cg23167506	chr7	1303302		0.17	0.78	0.58	0.78	0.63	0.27	0.96	0.91
254	cg26516974	chr6	52475065		0.28	0.75	0.74	0.57	0.52	0.08	0.89	0.71
255	cg05973305	chr5	87168997		0.27	0.85	0.57	0.78	0.81	0.19	0.93	0.85
256	cg05664613	chr19	33209411	*TDRD12*	0.39	0.82	0.9	0.71	0.75	0.29	0.94	0.98

257	cg02624246	chr12	7073264	*MIR-141*	0.23	0.85	0.89	0.91	0.58	0.15	1	0.89
258	cg00985983	chr22	30728510	*SF3A1*	0.37	0.81	0.54	0.73	0.62	0.21	0.98	0.69
259	cg14943002	chr11	47256035	*DDB2*	0.18	0.93	0.91	0.56	0.63	0.18	0.97	0.98
260	cg11797316	chr8	74226661	*RDH10; RDH10-AS1*	0.2	0.9	0.92	0.87	0.69	0.17	0.93	0.96
261	cg08155365	chr1	228319085		0.37	0.92	0.61	0.79	0.64	0.15	0.94	0.91
262	cg08297393	chr2	100937505	*LONRF2*	0.16	0.88	0.93	0.94	0.69	0.19	0.83	0.8
263	cg13779682	chr1	11131308	*EXOSC10*	0.47	0.83	0.89	0.82	0.91	0.28	0.95	0.97
264	cg16347969	chr2	38673971		0.45	0.85	0.89	0.78	0.91	0.37	0.85	0.81
265	cg09834343	chr15	32942910	*SCG5*	0.13	0.92	0.89	0.51	0.82	0.17	0.93	0.73
266	cg21668964	chr3	182360510		0.47	0.88	0.8	1	0.62	0.43	1	0.89
267	cg04951677	chr6	137316259		0.14	0.9	0.66	0.67	0.62	0.09	0.98	0.78
268	cg25726425	chr2	234637574	*UGT1A10; UGT1A6; UGT1A8; UGT1A4; UGT1A3; UGT1A9; UGT1A7; UGT1A5*	0.23	0.91	0.87	0.54	0.76	0.26	0.86	0.97
269	cg21899374	chr8	144915517		0.19	0.91	0.91	0.53	0.72	0.2	0.98	0.89
270	cg13595161	chr10	126713388	*CTBP2*	0.24	0.84	0.53	0.61	0.71	0.16	0.96	0.94
271	cg09217922	chr12	7096338	*LPCAT3*	0.33	0.87	0.89	0.75	0.8	0.33	0.96	0.96
272	cg04670206	chr7	6699522		0.3	0.75	0.51	0.56	0.7	0.22	0.85	0.61
273	cg00617776	chr3	36878751	*TRANK1*	0.45	0.85	0.59	0.83	0.79	0.37	0.88	0.74
274	cg11515851	chr16	81271013	*BCO1*	0.5	0.89	0.92	0.58	0.92	0.49	0.94	0.97
275	cg22874988	chr3	111451090	*PHLDB2*	0.2	0.83	0.96	0.54	0.95	0.12	1	0.97
276	cg07375022	chr13	113243073	*TUBGCP3*	0.35	0.93	0.9	0.71	0.83	0.25	0.97	0.93
277	cg00674569	chr2	189340785	*GULP1*	0.24	0.78	0.62	0.59	0.72	0.13	0.83	0.65
278	cg14266423	chr13	106632013		0.18	0.89	0.98	0.65	0.69	0.08	1	0.92
279	cg01181301	chr22	22098620		0.13	0.91	0.52	0.73	0.65	0.17	1	0.83
280	cg08650811	chr13	77572745	*CLN5*	0.22	0.88	0.89	0.75	0.59	0.12	0.96	0.75
281	cg03275396	chr3	111451616	*PLCXD2*	0.23	0.94	0.94	0.73	0.86	0.11	0.98	0.92
282	cg15489932	chr16	73266628		0.19	0.88	0.78	0.71	0.7	0.12	1	0.82
283	cg04779752	chr7	73245457	*CLDN4*	0.21	0.97	0.75	0.67	0.74	0.29	0.93	0.89

284	cg24361808	chr22	46477834		0.37	0.87	0.51	0.75	0.51	0.4	0.95	0.83
285	cg01274257	chr22	31746540		0.21	0.85	0.56	0.53	0.54	0.08	0.94	0.88
286	cg23270631	chr12	124371406	*DNAH10*	0.16	0.91	0.71	0.69	0.6	0.04	0.88	0.77
287	cg02813126	chr10	115241025		0.24	0.94	0.69	0.55	0.78	0.15	0.97	0.95
288	cg26281550	chr5	171991589		0.24	0.9	0.95	0.6	0.81	0.16	0.97	0.93
289	cg13556554	chr15	52855978	*ARPP19*	0.19	0.92	0.94	0.52	0.86	0.16	0.97	0.9
290	cg06372503	chr3	54401037	*CACNA2D3*	0.21	0.9	0.93	0.77	0.81	0.24	0.88	0.83
291	cg11613427	chr3	52932208	*TMEM110*	0.41	0.92	0.83	0.96	0.85	0.28	0.95	0.91
292	cg26115667	chr14	103294656	*TRAF3*	0.13	0.56	0.65	0.63	0.53	0.14	0.81	0.57
293	cg02639667	chr5	154092321	*LARP1*	0.26	0.98	0.91	0.93	0.57	0.08	0.89	0.8
294	cg01999399	chr3	195537508	*MUC4*	0.49	0.8	0.91	0.61	0.89	0.33	1	0.9
295	cg08656596	chr22	32641159	*SLC5A4*	0.16	0.91	0.86	0.75	0.69	0.03	1	0.97
296	cg12912809	chr4	128887467	*MFSD8; C4orf29*	0.15	0.83	0.63	0.57	0.58	0.1	0.91	0.79
297	cg27219362	chr7	157259285		0.12	0.89	0.54	0.83	0.6	0.1	0.79	0.8
298	cg06398482	chr13	77411461		0.4	0.83	0.9	0.78	0.87	0.56	0.97	0.81
299	cg06942239	chr20	4057157		0.36	0.87	0.74	1	0.69	0.09	1	0.93
300	cg13358134	chr9	96025413	*WNK2*	0.21	0.87	0.97	0.85	0.62	0.12	0.86	0.86
301	cg18039298	chr14	100942739	*WDR25*	0.34	0.95	0.89	0.93	0.96	0.36	0.98	1
302	cg02063944	chr13	45429070		0.2	0.79	0.92	0.52	0.65	0.11	0.89	0.94
303	cg12849046	chr5	72063769		0.27	0.92	0.51	0.6	0.52	0.06	0.91	0.71
304	cg11562836	chr10	134222581	*PWWP2B*	0.3	0.79	1	0.88	0.64	0.07	1	0.83
305	cg24698925	chr15	57873089		0.29	0.9	0.97	0.86	0.75	0.12	1	0.98
306	cg00708814	chr11	30751744		0.35	0.77	0.51	0.85	0.66	0.14	0.95	0.76
307	cg00965566	chr10	17331997		0.46	0.95	0.91	0.76	0.91	0.2	0.98	0.96
308	cg02676406	chr14	51863257		0.16	0.92	0.68	0.84	0.88	0.26	0.97	0.91
309	cg10687709	chr4	11013085		0.39	0.9	0.96	0.55	0.8	0.27	0.96	0.98
310	cg00769408	chr1	209827637		0.35	0.87	0.95	0.53	0.79	0.23	0.93	0.98
311	cg00270764	chr16	28511362	*IL27*	0.41	0.75	0.73	0.73	0.71	0.44	0.95	0.69
312	cg21883012	chr8	38124904	*PPAPDC1B*	0.17	0.92	0.62	0.95	0.83	0.23	0.98	1

313	cg18695263	chr3	196387505	*LRRC33*	0.3	0.92	0.79	0.88	0.85	0.12	0.97	0.88
314	cg14094347	chr9	139131620	*QSOX2*	0.11	0.72	0.82	0.86	0.68	0.13	0.88	0.73
315	cg10191560	chr11	115793792		0.39	0.82	0.9	0.8	0.93	0.2	0.96	0.93
316	cg11856351	chr11	22646355	*FANCF*	0.31	0.88	0.64	0.53	0.79	0.15	0.89	0.93
317	cg14528964	chr2	178938591	*PDE11A*	0.15	0.82	0.55	0.54	0.58	0.09	0.82	0.67
318	cg01303385	chr3	111451470	*PHLDB2*	0.22	0.89	0.79	1	0.82	0.04	0.91	0.96
319	cg09985163	chr19	7187927	*INSR*	0.4	0.93	0.68	0.74	0.73	0.23	0.9	0.76
320	cg00909094	chr3	12586368		0.18	0.89	0.84	0.73	0.98	0.09	0.94	0.97
321	cg21234836	chr11	13018935		0.22	0.93	0.78	0.67	0.82	0.12	0.95	0.87
322	cg17185473	chr5	64907370	*TRIM23*	0.2	0.86	0.81	0.61	0.65	0.13	0.92	0.92
323	cg20785060	chr8	64606039		0.16	0.88	0.62	0.65	0.8	0.11	0.86	0.69
324	cg26940755	chr1	32694147	*EIF3I*	0.46	0.94	0.7	0.59	0.86	0.24	0.94	0.78
325	cg11417701	chr6	401311	*IRF4*	0.42	0.88	0.96	0.76	0.91	0.2	0.94	0.9
326	cg14983771	chr14	61201189	*MNAT1*	0.19	0.92	0.51	0.88	0.87	0.05	0.83	0.93
327	cg06089322	chr14	65203351	*PLEKHG3*	0.34	0.9	0.79	0.83	0.83	0.11	1	0.91
328	cg20810117	chr8	22255186	*SLC39A14*	0.38	0.88	0.62	0.54	0.86	0.33	0.94	0.86
329	cg23682432	chr12	113246173	*RPH3A*	0.33	0.94	0.93	0.81	0.85	0.26	1	0.93
330	cg22603770	chr18	11005527	*PIEZO2*	0.44	0.75	0.95	0.72	0.86	0.24	0.98	0.94
331	cg17006098	chr17	29822786	*RAB11FIP4*	0.27	0.79	0.85	0.73	0.72	0.27	0.79	0.7
332	cg01779847	chr7	150599373		0.28	0.77	0.81	0.82	0.71	0.11	0.94	0.69
333	cg25611750	chr14	101755656		0.34	0.86	0.81	0.58	0.88	0.17	0.97	0.98
334	cg16331818	chr12	6033190	*ANO2*	0.22	0.89	0.65	0.8	0.57	0.11	0.9	0.86
335	cg19023320	chr3	13033587	*IQSEC1*	0.39	0.83	0.94	0.86	0.87	0.33	1	0.85
336	cg14123409	chr17	2831903	*RAP1GAP2*	0.44	0.86	0.92	0.79	0.69	0.77	1	0.97
337	cg24346519	chr9	131268003	*GLE1; MIR1268A*	0.43	0.79	0.57	0.81	0.73	0.25	0.92	0.76
338	cg27304415	chr3	77064282		0.16	0.9	0.8	0.63	0.65	0.13	0.89	0.67
339	cg23919160	chr9	79484243	*PRUNE2*	0.21	0.88	0.93	0.76	0.65	0.1	0.97	0.82
340	cg25451497	chr2	178990729	*RBM45*	0.44	0.91	0.69	0.82	0.55	0.3	0.93	0.9
341	cg07848608	chr1	19836013		0.39	0.92	0.74	0.97	0.86	0.35	1	0.96

342	cg07015239	chr1	235400645	*ARID4B*	0.23	0.94	0.9	0.54	0.7	0.16	0.98	0.95
343	cg12691464	chr10	32231950		0.26	0.91	0.99	0.57	0.66	0.11	0.98	0.95
344	cg03166324	chr1	161059535	*PVRL4*	0.3	0.73	0.83	0.83	0.59	0.05	0.83	0.91
345	cg04084026	chr16	57832309	*KIFC3*	0.32	0.92	0.93	0.58	0.69	0.36	1	0.88
346	cg23425669	chr8	117000211	*LINC00536*	0.19	0.84	0.79	0.69	0.82	0.21	0.95	0.9
347	cg08977118	chr10	80288744		0.32	0.93	0.9	0.59	0.88	0.19	1	0.95
348	cg11738790	chr6	24355819	*KAAG1; DCDC2*	0.17	0.87	0.71	0.76	0.59	0.23	0.95	0.74
349	cg21235987	chr9	133853434		0.33	0.89	0.61	0.54	0.52	0.14	0.96	0.74
350	cg04929665	chr17	26652175	*TMEM97*	0.23	0.95	0.84	0.8	0.68	0.21	0.94	0.91
351	cg11807356	chr11	71736227	*NUMA1*	0.46	0.8	0.65	0.81	0.68	0.19	0.92	0.76
352	cg26936966	chr8	136612125	*KHDRBS3*	0.2	0.77	0.62	0.92	0.74	0.07	0.64	0.67
353	cg00606396	chr5	19531574	*CDH18*	0.29	0.91	0.96	0.89	0.57	0.35	0.98	0.91
354	cg26190835	chr2	202937899		0.16	0.79	0.6	0.58	0.64	0.13	0.91	0.69
355	cg05498681	chr20	35973318	*SRC*	0.34	0.9	0.92	0.52	0.69	0.22	0.97	0.88
356	cg06103218	chr4	2073751	*POLN*	0.27	0.87	0.59	0.52	0.56	0.04	0.96	0.65
357	cg24680492	chr12	3904727		0.34	0.94	0.71	0.53	0.72	0.13	1	0.94
358	cg11131672	chr1	170588581		0.09	0.88	0.89	0.62	0.66	0.1	0.94	0.79
359	cg22115148	chr3	177216338	*LINC00578*	0.31	0.92	0.51	0.91	0.73	0.33	0.92	0.93
360	cg01434302	chr8	3267208	*CSMD1*	0.14	0.86	0.52	0.63	0.62	0.22	0.96	0.78
361	cg02233275	chr1	28961119	*TAF12*	0.34	0.8	0.77	0.57	0.67	0.18	0.94	0.8
362	cg02693498	chr13	24758821	*SPATA13*	0.14	0.93	0.99	0.65	1	0.09	1	0.93
363	cg00759354	chr8	143758944	*PSCA*	0.47	0.85	0.69	0.62	0.52	0.3	0.93	0.88
364	cg21655740	chr8	52945868		0.18	0.89	0.51	0.51	0.61	0.21	0.93	0.87
365	cg22379697	chr20	40104506	*CHD6*	0.2	0.91	0.57	1	0.54	0.02	0.93	0.76
366	cg24491939	chr18	29186408		0.23	0.94	0.87	0.67	0.87	0.14	0.98	0.89
367	cg01552748	chr13	112225526		0.22	0.94	0.67	0.56	0.54	0.04	1	0.82
368	cg07138944	chr7	75929652		0.28	0.89	0.83	0.72	0.72	0.24	0.91	0.92
369	cg05595429	chr1	222759637	*TAF1A*	0.34	0.89	0.66	0.97	0.79	0.32	0.93	0.86
370	cg26150922	chr2	100937072	*LONRF2*	0.2	0.9	0.79	0.91	0.7	0.14	0.98	0.87

371	cg16919295	chr2	23421947		0.29	0.85	0.81	0.54	0.51	0.43	0.93	0.91
372	cg01823510	chr2	242776354		0.34	0.86	0.75	0.78	0.67	0	0.83	0.69
373	cg13793676	chr6	7480452		0.14	0.76	0.61	0.79	0.56	0.05	0.93	0.62
374	cg05802545	chr1	107921855	*NTNG1*	0.4	0.76	0.55	0.58	0.59	0.21	0.8	0.65
375	cg24691398	chr9	36637730	*MELK*	0.47	0.94	0.62	0.91	0.56	0.26	1	0.94
376	cg05415871	chr8	1008188		0.4	0.97	0.64	0.67	0.65	0.28	0.91	0.89
377	cg08713760	chr12	40948140	*MUC19*	0.49	0.82	0.71	0.73	0.88	0.47	0.94	0.81
378	cg26147915	chr3	47426793	*PTPN23*	0.4	0.94	0.74	0.69	0.64	0.2	0.93	0.92
379	cg02960500	chr11	131469924	*NTM*	0.45	0.85	0.56	0.64	0.76	0.2	0.88	0.9
380	cg22455699	chr4	79946748	*LINC01088*	0.37	0.89	0.67	0.86	0.77	0.3	0.98	0.96
381	cg08271443	chr16	4845348	*LOC440335*	0.17	0.76	0.9	1	0.76	0.23	1	0.74
382	cg26289155	chr17	73546144	*LLGL2*	0.42	0.84	0.97	0.92	0.86	0.08	0.89	0.93
383	cg10463011	chr5	168978822		0.41	0.9	0.83	0.52	0.72	0.28	1	0.91
384	cg22802766	chr12	14361288		0.45	0.81	0.87	0.7	0.52	0.21	0.91	0.84
385	cg10692636	chr17	76470423	*DNAH17*	0.38	0.9	0.74	0.6	0.89	0.22	0.89	0.87
386	cg24530480	chr8	1008155		0.43	0.91	0.53	0.76	0.79	0.21	0.9	0.82
387	cg18700339	chr11	128634547	*FLI1*	0.49	0.85	0.82	0.95	0.68	0.47	1	0.98
388	cg16427033	chr10	109023583		0.44	0.97	0.63	0.59	0.71	0.26	0.87	0.86
389	cg04658707	chr21	47453889		0.45	0.94	0.97	0.97	0.62	0.35	0.94	0.96
390	cg22958715	chr14	23855309	*MYH6*	0.35	0.94	0.92	0.82	0.94	0.19	1	0.95
391	cg21190017	chr11	134124575	*THYN1; ACAD8*	0.46	0.8	0.62	0.68	0.74	0.22	0.91	0.94
392	cg07494102	chr11	34662394	*EHF*	0.4	0.94	0.84	0.93	0.67	0.21	0.94	0.92
393	cg07173284	chr16	24320466	*CACNG3*	0.42	0.89	0.67	0.58	0.67	0.21	1	0.91
394	cg13997361	chr17	8703224	*MFSD6L*	0.46	0.89	0.91	0.6	0.64	0.43	0.91	0.65
395	cg01315205	chr5	171534880	*STK10*	0.49	0.77	0.51	0.83	0.57	0.26	0.91	0.6
396	cg19866944	chr1	174966309		0.36	0.91	0.93	0.89	0.9	0.19	0.96	0.92
397	cg06610395	chr12	125350471		0.49	0.78	0.91	0.94	0.51	0.54	0.97	0.87
398	cg22912485	chr8	57896881	*IMPAD1*	0.43	0.92	0.74	0.8	0.91	0.23	0.98	0.9
399	cg27554156	chr12	13248725	*GSG1*	0.25	0.88	0.7	0.8	0.72	0.06	0.94	0.71

400	cg03661901	chr12	133093451	*FBRSL1*	0.45	0.91	0.95	0.89	0.9	0.37	1	0.94
401	cg16659470	chr8	94744382	*RBM12B*	0.37	0.88	0.92	0.57	0.65	0.25	1	0.95
402	cg25415853	chr12	5950373	*ANO2*	0.48	0.93	0.96	0.95	0.93	0.33	0.93	0.91
403	cg07484053	chr13	113243279	*TUBGCP3*	0.31	0.92	0.66	0.76	0.88	0.19	0.96	0.87
404	cg21545428	chr10	2808784		0.33	0.77	0.61	0.58	0.71	0.03	0.92	0.71
405	cg01003257	chr12	105056211	*CHST11*	0.28	0.91	0.78	0.75	0.88	0.12	0.95	0.88
406	cg02423000	chr5	601568		0.38	0.84	0.88	0.72	0.93	0.13	0.92	0.87
407	cg25741597	chr3	147722622		0.46	0.88	0.69	0.74	0.71	0.28	0.98	0.86
408	cg00853742	chr20	39990972	*EMILIN3*	0.39	0.83	0.87	0.87	0.92	0.22	1	0.89
409	cg05382021	chr14	23640996	*SLC7A8*	0.47	0.86	0.92	0.85	0.62	0.48	1	0.95
410	cg11194863	chr13	98945688	*FARP1*	0.28	0.82	0.9	0.83	0.7	0.2	0.97	0.92
411	cg07331478	chr2	221961668		0.44	0.9	0.77	0.82	0.81	0.2	0.94	0.85
412	cg13812733	chr19	55721821	*PTPRH*	0.45	0.89	0.96	0.94	0.8	0.34	0.95	0.91
413	cg08524891	chr7	120365958	*KCND2*	0.34	0.9	0.66	0.8	0.9	0.23	0.94	1
414	cg10340409	chr19	9944858	*PIN1*	0.48	0.88	0.76	0.64	0.76	0.27	1	0.72
415	cg13049624	chr3	1050278		0.41	0.86	0.73	0.65	0.66	0.15	0.92	0.74
416	cg22500255	chr1	28384743	*EYA3*	0.26	0.82	0.66	0.81	0.62	0.22	0.9	0.82
417	cg18463993	chr17	26653167	*TMEM97*	0.2	0.91	0.72	0.64	0.64	0.15	1	0.83
418	cg26748715	chr3	195587468		0.49	0.89	0.95	0.54	0.8	0.38	0.9	0.86
419	cg04782106	chrX	73075985		0.46	0.8	0.54	0.81	0.71	0.32	0.95	0.61
420	cg04456720	chr17	54250143	*ANKFN1*	0.32	0.72	0.59	0.59	0.58	0.36	0.89	0.65
421	cg10095968	chr21	45847903	*TRPM2*	0.49	0.92	0.88	0.96	0.9	0.33	0.92	0.87
422	cg05966608	chr11	117162779	*BACE1-AS; BACE1*	0.44	0.87	0.62	0.88	0.69	0.27	0.96	0.92
423	cg03685774	chr1	159870326	*CCDC19*	0.36	0.93	0.71	0.84	0.68	0.2	0.95	0.9
424	cg11168687	chr4	185726782	*ACSL1*	0.21	0.87	0.89	0.57	0.62	0.3	0.93	0.57
425	cg01479336	chr2	238972915	*SCLY; UBE2F-SCLY*	0.47	0.94	0.91	0.52	0.95	0.16	0.94	0.93

Table A.5: Summery of methylation percentage data derived from targeted next generation sequencing of *MIR-200c/MIR-141* mutual promoter as well as post MIR-141 (Chr12: 6,963,000-6,964,001).

	CpG Methylation (%)							
CpG	**EndoC-βH1 cell line**	**48h high Glucose Exposed EndoC-βH1 cell line**	**Pancreatic Islet**	**PBMCs**	**Lung**	**PNT2**	**LNcaP**	**PC3**
-595	0.0	0.0	2.0	52.8	21.7	0.5	0.8	84.1
-586	0.0	0.0	6.5	74.8	46.1	2.1	0.0	95.0
-524	0.0	0.0	4.1	95.5	60.4	0.0	1.5	95.1
-510	0.0	0.0	4.4	97.6	72.2	1.4	0.0	90.8
-488	0.0	0.0	4.3	100.0	67.3	0.0	0.0	97.5
-477	0.0	1.1	6.7	93.0	64.8	0.0	0.0	96.6
-470	0.0	0.0	4.4	93.0	70.4	0.0	0.0	97.5
-449	0.0	0.0	4.4	100.0	72.2	0.0	0.0	100.0
-437	0.0	0.0	4.4	95.1	68.5	0.0	0.0	97.1
-425	0.0	2.2	6.7	95.5	75.0	1.4	0.0	98.2
-397	0.0	2.2	4.5	97.6	66.0	0.0	0.0	97.4
-359	0.0	0.0	2.5	64.3	17.6	0.0	0.0	75.0
-280	0.0		3.6	97.2	57.6	1.8	0.0	97.4
-270	0.0		3.7	94.4	58.1	1.8	0.0	97.3
-268	0.0		2.0	94.4	61.3	2.0	0.0	92.1
-263	0.0		0.0	96.0	55.2	1.9	0.0	89.5
-250	0.0		6.1	100.0	67.9	0.0	0.0	94.4
-223	0.0		7.4	94.4	41.2	0.0	0.0	100.0
-220	0.0		7.7	90.9	54.3	1.9	0.0	97.2
-215	0.0		5.7	85.3	26.5	1.8	0.0	91.9
-212	0.0		7.4	94.3	47.1	1.8	0.0	97.4
-208	0.0		7.4	84.8	48.5	1.8	0.0	100.0
-198	0.0		9.4	77.8	51.4	3.7	0.0	86.8
-166	0.0		5.6	66.7	52.8	0.0	0.0	82.9
-157	0.0		7.5	91.4	58.3	1.8	0.0	100.0
-135	0.0		7.3	88.6	61.8	1.9	0.0	100.0
-97	0.0	1.1	5.6	69.1	45.7	0.0	0.0	93.1

+53	0.0	0.0	9.0	88.8	69.5	3.9	1.5	90.2
+97	0.0	0.0	4.1	44.6	21.4	1.6	0.0	62.6
+148	0.0	0.0	4.3	72.6	24.6	2.5	0.0	94.9
+155	0.0	2.7	5.8	50.9	22.8	2.9	0.0	82.1
+166	0.0	0.0	4.8	50.1	20.8	2.3	0.0	75.0
+261	1.3	0.5	9.9	89.3	63.5	2.9	0.6	96.2
+295	0.1	0.8	2.6	49.8	23.2	0.1	0.8	93.9
+318	0.0	0.2	4.5	81.2	51.5	0.3	0.2	95.0
+379	1.3	4.5	9.8	93.2	67.4	2.5	1.1	90.8
+381	1.1	2.3	6.9	93.8	66.2	2.5	1.4	96.5
+386	0.3	1.0	5.3	88.1	56.6	2.4	0.6	95.5
+399	0.3	1.0	6.6	89.0	55.3	2.7	0.6	92.1
+403	0.5	0.5	7.2	89.2	59.0	1.9	0.7	92.9
+483	2.1		5.7	95.8	64.5	0.0	0.0	97.0
+497	1.3		5.1	88.8	64.4	0.4	1.2	92.2
+580	2.9		4.5	75.0	45.3	5.5	0.0	71.1

Appendix B

Materials

100 mM Creosol Red (50 mL)

0.4g Creosol red

Disolved in 10 mL double distilled water

Vortex to mix.

60% Sucrose /1 mM Creosol Red (50 mL)

30g sucrose

Dissolved in 50 mL autoclaved de-ionised water

500 μL 100 mM Creosol red

Vortex to mix, then aliquot 1 mL into 1.5 mL Eppendorf tubes.

Store at -20°C.

0.5 M EDTA, pH 8.0 (500 mL)

93.06g EDTA

Dissolved in 500 mL distilled water to form a white solution.

Add 12g Sodium Hydroxide, NaOH, (~95 pellets)

Adjust pH to 8.0

If pH is above 8.0 add drops of HCl, if below add NaOH.

Autoclave.

5M NaCl (200 mL)

58.44g Sodium Chloride

Dissolved in 200 mL distilled water

Autoclave.

10% SDS (50 mL)

5g Sodium Dodecyl Sulphate

Dissolved in 50mls double distilled water.

3 M Sodium acetate (CH_3COONa) solution, pH 5.2 (100 mL)

24.61g CH3COONa (Molecular weight = 82.0343).

Add 80 mL double distilled water. Mix until sodium acetate is completely dissolved.

Adjust the pH to 5.2 with glacial acetic acid

Adjust the volume to 100 mL, Mix

Transfer the solution to autoclavable bottle, Autoclave.

Material	Catalogue Number	Company	Website
Agarose, molecular biology grade	MB1200	MelFord	https://www.melford.co.uk/search.html?keywords=MB1200
Ampicillin Antibiotic	A5354	Sigma	https://www.sigmaaldrich.com/catalog/product/sigma/a5354?lang=en&region=GB
BD™ Vacutainer™ SST™ II Advance Tubes	12927696	Becton Dickinson	https://www.fishersci.co.uk/shop/products/vacutainer-sst-ii-advance-tubes/p-8259011
C1000 Touch™ Thermal Cycler	184-1100	Bio-Rad	https://www.bio-rad.com/en-uk/product/c1000-touch-thermal-cycler?ID=LGTW9415
Cell Culture Flask, 550 ML, 175 CM^2	660975	Greiner Bio-One International	https://shop.gbo.com/en/england/products/bioscience/cell-culture-products/advanced-tc-cell-culture-vessels/advanced-tc-flasks/660975.html?_ga=2.55505518.415049627.1532703946-1554377241.1532703946
cell-free DNA collection tubes	07785666001	Roche	https://sequencing.roche.com/en/products-solutions/by-category/sample-collection/cell-free-dna-collection-tube/ordering.html
Cell-Free DNA Urine Preserve	230216	Streck	https://www.streck.com/collection/cell-free-dna-urine-preserve/
CMRL-1066 growth medium	C0422	Sigma-Aldrich	https://www.sigmaaldrich.com/catalog/product/sigma/c0422?lang=en&region=GB
ddPCR Supermix for Probes	186-3027	Bio-Rad	https://www.bio-rad.com/en-uk/sku/1863027-ddpcr-supermix-for-probes?ID=1863027
ddPCR™ Droplet Reader Oil	186-3004	Bio-Rad	https://www.bio-rad.com/en-uk/sku/1863004-ddpcr-droplet-reader-oil?ID=1863004
DMEM/F-12	11320033	Thermo Fisher Scientific	https://www.thermofisher.com/order/catalog/product/11320033?SID=srch-srp-11320033
DNA ladder (50 bp)	N3236	BioLabs	https://international.neb.com/products/n3236-50-bp-dna-ladder#Product%20Information
DNA LoBind Tubes	0030108035	Eppendorf	https://online-shop.eppendorf.co.uk/UK-en/Laboratory-Consumables-44512/Tubes-44515/DNA-LoBind-Tubes-PF-56252.html?_ga=2.175604714.99066778.1532706480-6596190.1532706480
DG8™ Cartridges for QX200™/QX100™ Droplet Generator	186-4008	Bio-Rad	https://www.bio-rad.com/en-uk/sku/1864008-dg8-cartridges-for-qx200-qx100-droplet-generator?ID=1864008
Dulbecco's modified Eagles Medium	11965092	Thermo Fisher Scientific	https://www.thermofisher.com/order/catalog/product/11965092?SID=srch-hj-11965092
E4 XLS Multichannel pipette	8479899499	Rainin	https://www.shoprainin.com/Products/Pipettes-and-Tips/Pipettes/Multichannel-Electronic-Pipettes/E4-XLS%2B-with-LTS/E4-Multi-Pipette-E8-300XLS%2B/p/17013796

Eppendorf® LoBind microcentrifuge tubes (volume 1.5 mL)	Z666548-250EA	Sigma	https://www.sigmaaldrich.com/catalog/product/sigma/z666548?lang=en&region=GB
Ethylenediaminetetraacetic acid	ED-100G	Sigma-Aldrich	https://www.sigmaaldrich.com/catalog/product/sial/ed?lang=en&region=GB
EZ DNA Methylation Kit	D5001	Zymo Research	https://www.zymoresearch.eu/ez-dna-methylation-kit#format=4&reactions=42
EZ DNA Methylation-Lightning™ Kit	D5030	Zymo Research	https://www.zymoresearch.eu/ez-dna-methylation-lightning-kit#format=4&reactions=42
Fetal Bovine Serum, certified, heat inactivated	10082147	Thermo Fisher Scientific	https://www.thermofisher.com/order/catalog/product/10082147?SID=srch-hj-10082147
Gaskets	186-3009	Bio-Rad	https://www.bio-rad.com/en-uk/sku/1863009-dg8-gaskets-for-qx200-qx100-droplet-generator?ID=1863009
Horse Serum	16050122	Thermo Fisher Scientific	https://www.thermofisher.com/order/catalog/product/16050122?SID=srch-hj-16050122
HotStar Taq DNA Polymerase	203203	Qiagen	https://www.qiagen.com/gb/shop/pcr/end-point-pcr-enzymes-and-kits/regular-pcr/hotstartaq-dna-polymerase/#orderinginformation
Hypermethylated and Hypomethylated human DNA set	D5014	Zymo Research	https://www.zymoresearch.eu/human-methylated-non-methylated-dna-set#methylation_type=271
L- Glutamine	17-605C	Fisher Scientifics	https://www.fishersci.com/shop/products/lonza-biowhittaker-l-glutamine-200mm-l-glutamine-200mm-50ml/bw17605c
Lonza™ art™ RPMI 1640 with L-Glutamine	12072-F	Fisher Scientifics	https://www.fishersci.fi/shop/products/lonza-biowhittaker-rpmi-1640-l-glutamine-7/11605220
MedoriGreen	MG04	Nippon Genetics	https://www.nippongenetics.eu/en/product/midori-green-advance/
MinElute Gel Extraction Kit	28604	Qiagen	https://www.qiagen.com/kr/shop/sample-technologies/dna/dna-clean-up/minelute-gel-extraction-kit/#orderinginformation
miRNeasy Serum/Plasma Kit	217184	Qiagen	https://www.qiagen.com/kr/shop/sample-technologies/rna/mirna/mirneasy-serumplasma-kit/#orderinginformation
Newport Green DCFTM fluorescent indicator	N7991	Thermo Fisher Scientific	https://www.thermofisher.com/order/catalog/product/N7991?SID=srch-hj-N7991
Nuclease-Free Water	129114	Qiagen	https://www.qiagen.com/kr/shop/lab-basics/buffers-and-reagents/nuclease-free-water/#orderinginformation
PCR Plate Heat Seal, foil, pierceable	181-4040	Bio-Rad	https://www.bio-rad.com/en-uk/sku/1814040-pcr-plate-heat-seal-foil-pierceable?ID=1814040
pGEM-T Easy vector system I	A1360	Promega	https://www.promega.co.uk/Products/PCR/PCR-Cloning/pGEM-T-Easy-Vector-Systems/?fq=A1360&catNum=A1360
PX1™ PCR Plate Sealer	181-4000	Bio-Rad	https://www.bio-rad.com/en-uk/sku/1814000-px1-pcr-plate-sealer?ID=1814000
QIAamp Circulating Nucleic Acid Kit	55114	Qiagen	https://www.qiagen.com/kr/shop/sample-technologies/dna/genomic-dna/qiaamp-circulating-nucleic-acid-kit/#orderinginformation

QiaAmp DNA Blood Mini Kit	51106	Qiagen	https://www.qiagen.com/kr/shop/sample-technologies/dna/genomic-dna/qiaamp-dna-blood-mini-kit/#orderinginformation
QIAprep Spin miniprep kit	27104	Qiagen	https://www.qiagen.com/kr/shop/sample-technologies/dna/plasmid-dna/qiaprep-spin-miniprep-kit/#orderinginformation
QuantiFluor® dsDNA system	E2670	Promega	https://www.promega.co.uk/Products/DNA-Purification-Quantitation/DNA-and-RNA-Quantitation/QuantiFluor-dsDNA-System/?fq=E2670&catNum=E2670
QuantiFluor® RNA System	E3310	Promega	https://www.promega.co.uk/Products/DNA-Purification-Quantitation/DNA-and-RNA-Quantitation/QuantiFluor-RNA-System/?fq=E3310&catNum=E3310
QuantiFluor® ssDNA system	E3190	Promega	https://www.promega.co.uk/Products/DNA-Purification-Quantitation/DNA-and-RNA-Quantitation/QuantiFluor-ssDNA-System/?fq=E3190&catNum=E3190
Quantus™ Fluorometer	E6150	Promega	https://www.promega.co.uk/Products/Fluorometers-Luminometers-Multimode-Readers/Fluorometers/Quantus-Fluorometer/?fq=E6150&catNum=E6150
QX200 Droplet Generation Oil for Probes	186-3005	Bio-Rad	https://www.bio-rad.com/en-uk/sku/1863005-droplet-generation-oil-for-probes?ID=1863005
QX200™ ddPCR™ EvaGreen Supermix	186-4033	Bio-Rad	https://www.bio-rad.com/en-uk/sku/1864033-qx200-ddpcr-evagreen-supermix?ID=1864033
QX200™ Droplet Generation Oil for EvaGreen	1864005	Bio-Rad	https://www.bio-rad.com/en-uk/sku/1864005-qx200-droplet-generation-oil-for-evagreen?ID=1864005
QX200™ Droplet Generator	186-4002	Bio-Rad	https://www.bio-rad.com/en-uk/sku/1864002-qx200-droplet-generator?ID=1864002
QX200™ Droplet reader	186-4003	Bio-Rad	https://www.bio-rad.com/en-uk/sku/1864003-qx200-droplet-reader?ID=1864003
REPLI-g Mini Kit	150023	Qiagen	https://www.qiagen.com/gb/shop/sample-technologies/dna/genomic-dna/repli-g-mini-kit/#orderinginformation
Semi-skirted and PCR clean 96-well PCR plate	0030128.575	Eppendorf	https://online-shop.eppendorf.co.uk/UK-en/eshopproduct/view/0030128575?_ga=2.167142262.99066778.1532706480-6596190.1532706480
OmniPure Tris/Borate/EDTA	EM8830	Merck Millipore	http://www.merckmillipore.com/GB/en/product/OmniPur-10X-TBE-Buffer-Calbiochem,EMD_BIO-8820-OP
Universal cDNA Synthesis Kit II	203301	Exiqon	https://www.qiagen.com/gb/shop/pcr/primer-sets/mircury-lna-rt-kit/#orderinginformation
QIAGEN Proteinase K (10 mL)	19133	Qiagen	https://www.qiagen.com/gb/shop/lab-basics/enzymes/qiagen-proteinase-k/#orderinginformation