
Dayeh, Tasnim; Olsson, Anders H; Volkov, Petr; Almgren, Peter; Rönn, Tina; Ling, Charlotte

Published in:
Diabetologia

DOI:
10.1007/s00125-012-2815-7

2013

Link to publication

Citation for published version (APA):
Identification of CpG-SNPs associated with type 2 diabetes and differential DNA methylation in human pancreatic islets

T. A. Dayeh · A. H. Olsson · P. Volkov · P. Almgren · T. Rönn · C. Ling

Received: 14 November 2012 / Accepted: 6 December 2012 / Published online: 6 March 2013 © The Author(s) 2013. This article is published with open access at Springerlink.com

Abstract
Aims/hypothesis To date, the molecular function of most of the reported type 2 diabetes-associated loci remains unknown. The introduction or removal of cytosine–phosphate–guanine (CpG) dinucleotides, which are possible sites of DNA methylation, has been suggested as a potential mechanism through which single-nucleotide polymorphisms (SNPs) can affect gene function via epigenetics. The aim of this study was to examine if any of 40 SNPs previously associated with type 2 diabetes introduce or remove a CpG site and if these CpG-SNPs are associated with differential DNA methylation in pancreatic islets of 84 human donors.

Methods DNA methylation was analysed using pyrosequencing.

Results We found that 19 of 40 (48%) type 2 diabetes-associated SNPs introduce or remove a CpG site. Successful DNA methylation data were generated for 16 of these 19 CpG-SNP loci, representing the candidate genes TCF7L2, KCNQ1, CDKN2A, ADCY5, WFS1, CDKAL1, DUSP9, DUSP8, HHEX, IRS1, PPARG, SLC30A8, SRR, CHCHD9, TSPAN8 and DUSP8. All analysed CpG-SNPs were associated with differential DNA methylation of the CpG-SNP site in human islets. Moreover, six CpG-SNPs, representing TCF7L2, KCNQ1, CDKN2A, ADCY5, WFS1 and HMG2, were also associated with DNA methylation of surrounding CpG sites. Some of the type 2 diabetes CpG-SNP sites that exhibit differential DNA methylation were further associated with gene expression, alternative splicing events determined by splice index, and hormone secretion in the human islets. The 19 type 2 diabetes-associated CpG-SNPs are in strong linkage disequilibrium ($r^2>0.8$) with a total of 295 SNPs, including 91 CpG-SNPs.

Conclusions/interpretation Our results suggest that the introduction or removal of a CpG site may be a molecular mechanism through which some of the type 2 diabetes SNPs affect gene function via differential DNA methylation and consequently contributes to the phenotype of the disease.

Keywords Alternative splicing · CpG-SNP · DNA methylation · Epigenetics · Glucagon · Human pancreatic islets · Insulin secretion · SNP · Splice index · Type 2 diabetes

Abbreviations
CpG Cytosine–phosphate–guanine
GWAS Genome-wide association study
LD Linkage disequilibrium
SI Stimulation index
SNAP SNP annotation and proxy search
SNP Single-nucleotide polymorphism

Introduction
Type 2 diabetes is a heterogeneous disease caused by a complex interplay between multiple genetic, epigenetic and environmental factors [1]. Attempts at elucidating the genetic mechanisms behind this disease have identified approximately 40 single-nucleotide polymorphisms (SNPs)
that show genome-wide associations with type 2 diabetes [2]. However, evidence for how these SNPs predispose to type 2 diabetes is still scarce and is only available for a few variants that mainly affect insulin secretion and sensitivity [3]. Understanding the molecular mechanisms through which these SNPs exert an effect on their target genes is not an easy task, especially as most of them are non-coding SNPs [2]. To further understand the pathogenesis of the disease, additional studies are required that link these SNPs with their molecular and biological functions.

Recent studies from our group and others have shown that epigenetic modifications such as DNA methylation and histone modifications may affect the pathogenesis of type 2 diabetes [4–9]. In differentiated mammalian cells, DNA methylation mainly takes place at a cytosine in a cytosine–phosphate–guanine (CpG) dinucleotide [10] and increased DNA methylation has been associated with transcriptional silencing [11]. In a previous study, we showed that genetic, epigenetic and non-genetic factors integrate to influence the expression of the candidate gene for type 2 diabetes in human skeletal muscle [12]. We identified a SNP near NDUFB6 that introduces a CpG site and this CpG-SNP was associated with increased DNA methylation, decreased gene expression, and insulin resistance in the muscle of elderly, but not young, individuals [12]. Moreover, a recent study has suggested that CpG-SNPs may affect the expression of target genes by interfering with the binding of certain proteins [13]. However, to our knowledge the role of CpG-SNPs associated with type 2 diabetes has not been examined in human pancreatic islets.

Therefore, the aim of this study was to examine if 40 SNPs previously associated with type 2 diabetes [2] introduce or remove possible DNA methylation sites – CpG sites – and if these CpG-SNPs are associated with differential DNA methylation and subsequently affect gene expression, alternative splicing events and hormone secretion in human pancreatic islets.

Methods

Human pancreatic islet donors Pancreatic islets from 84 deceased non-diabetic donors were provided by the Nordic Network for Islet Transplantation, Uppsala University. Clinical characteristics of these islet donors are given in Table 1. The donor before death or her/his relatives on admission to intensive care unit gave their consent to donate organs and the local ethics committees approved the protocols. Islet culture and nucleic acid purification were performed as described previously [7].

SNP selection and identification of CpG-SNPs The selection criterion for SNPs was an association with type 2 diabetes identified by major genome-wide association studies (GWASs) as reported by McCarthy [2]. To determine if the selected SNPs were CpG-SNPs, the DNA sequences of the variants were retrieved from the dbSNP database (www.ncbi.nlm.nih.gov/snp/, on or after 03 July 2011). A SNP was considered to be a CpG-SNP if it introduces or removes a CpG site (Table 2). As a result, 19 SNPs fulfilled our criteria of being CpG-SNPs that show genome-wide association with type 2 diabetes. DNA sequences of these 19 CpG-SNPs and surrounding regions were then retrieved from University of California, Santa Cruz Human Genome Browser (UCSC) – hg18 assembly.

A flow chart of the study design and SNP selection can be seen in Fig. 1.

DNA methylation assay design To analyse DNA methylation of the 19 selected CpG-SNP sites, pyrosequencing PCR and sequencing primers were designed for the forward strand using the PyroMark Assay design Software 2.0 (Qiagen, Hilden, Germany) (electronic supplementary material [ESM] Table 1). The reverse or forward primer was biotinylated at its 5′ end. When possible, CpG sites surrounding the CpG-SNP of interest were also analysed. Problems caused by sequence characteristics that resulted in amplicons that were too long or too many CpG sites in the primer sequences resulted in failure of the assay design for rs5219 (KCNJ11) (Fig. 1).

DNA methylation analysis Bisulfite conversion of 500 ng genomic DNA from human pancreatic islets was performed using the EpiTect Bisulfite Kit (Qiagen) according to the manufacturer’s instructions. Bisulfite-converted DNA was amplified by PCR using the PyroMark PCR kit (Qiagen) and the designed primers (ESM Table 1). Pyrosequencing was performed using the PSQ HS96A (Biotage, Uppsala, Sweden) and PyroMark Gold CDT kit (Qiagen) or using PyroMark Q96ID and the PyroMark Gold Q96 kit (Qiagen), all according to the manufacturer’s instructions. Data analysis was performed using the Pyro Q-CpG 1.0.9 and PSQ HS96A 1.2 software (Biotage) or the PyroMark Q96 (version 2.5.8) software (Qiagen). Only data that passed a quality check or had medium quality were included in the analysis and as a result the following assay was excluded:

Table 1 Characteristics of the human pancreatic islet donors

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number (women/men)</td>
<td>84 (32/52)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>57.1±10.3</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>25.9±3.6</td>
</tr>
<tr>
<td>HbA1c (%)</td>
<td>5.6±0.4</td>
</tr>
<tr>
<td>HbA1c (mmol/mol)</td>
<td>48.1±5</td>
</tr>
</tbody>
</table>

Data are presented as mean ± SD

HbA1c was measured with the Mono-S method
### Table 2 SNPs associated with type 2 diabetes identified by GWAS

<table>
<thead>
<tr>
<th>SNP ID</th>
<th>Candidate gene</th>
<th>Risk allele</th>
<th>Chr.</th>
<th>SNP position (build 37.1)</th>
<th>SNP location</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs9543526</td>
<td>DUSP9&lt;sup&gt;a&lt;/sup&gt;</td>
<td>A</td>
<td>X</td>
<td>15289992</td>
<td>Intergenic</td>
<td>Voight et al (2010) [20]</td>
</tr>
<tr>
<td>rs64398</td>
<td>CDKN2A&lt;sup&gt;a&lt;/sup&gt;</td>
<td>A</td>
<td>9</td>
<td>22029547</td>
<td>Intergenic</td>
<td>Zeggini et al (2007) [21]</td>
</tr>
<tr>
<td>rs11708067</td>
<td>ADCY5&lt;sup&gt;a&lt;/sup&gt;</td>
<td>A</td>
<td>3</td>
<td>123065778</td>
<td>Intron</td>
<td>Dupuis et al (2010) [22]</td>
</tr>
<tr>
<td>rs1801214</td>
<td>WFS1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>T</td>
<td>4</td>
<td>6303022</td>
<td>Cds-synon</td>
<td>Voight et al (2010) [20]</td>
</tr>
<tr>
<td>rs2334499</td>
<td>DUSP8&lt;sup&gt;a&lt;/sup&gt;</td>
<td>A</td>
<td>11</td>
<td>1696849</td>
<td>Intergenic</td>
<td>Kong et al (2009) [23]</td>
</tr>
<tr>
<td>rs758326</td>
<td>IRS1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>A</td>
<td>2</td>
<td>227020653</td>
<td>Intergenic</td>
<td>Voight et al (2010) [20]</td>
</tr>
<tr>
<td>rs5219</td>
<td>KCNJ11&lt;sup&gt;b&lt;/sup&gt;</td>
<td>T</td>
<td>11</td>
<td>17409572</td>
<td>Cds-nonsynon</td>
<td>Scott et al (2007) [24]</td>
</tr>
<tr>
<td>rs1801282</td>
<td>PPPARG&lt;sup&gt;a&lt;/sup&gt;</td>
<td>C</td>
<td>3</td>
<td>12393125</td>
<td>Cds-nonsynon</td>
<td>Scott et al (2007) [24]</td>
</tr>
<tr>
<td>rs13292136</td>
<td>CHCHD9&lt;sup&gt;a&lt;/sup&gt;</td>
<td>C</td>
<td>9</td>
<td>81952128</td>
<td>Intergenic</td>
<td>Voight et al (2010) [20]</td>
</tr>
<tr>
<td>rs7901695</td>
<td>TCF7L2&lt;sup&gt;a&lt;/sup&gt;</td>
<td>G</td>
<td>10</td>
<td>114754088</td>
<td>Intron</td>
<td>Zeggini et al (2007) [21]</td>
</tr>
<tr>
<td>rs7754840</td>
<td>CDKAL1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>C</td>
<td>6</td>
<td>20661250</td>
<td>Intron</td>
<td>Scott et al (2007) [24]</td>
</tr>
<tr>
<td>rs391300</td>
<td>SRR&lt;sup&gt;a&lt;/sup&gt;</td>
<td>G</td>
<td>17</td>
<td>2216258</td>
<td>Intron</td>
<td>Tsai et al (2010) [25]</td>
</tr>
<tr>
<td>rs5015480</td>
<td>HHEX&lt;sup&gt;a&lt;/sup&gt;</td>
<td>G</td>
<td>10</td>
<td>94665559</td>
<td>Intergenic</td>
<td>Voight et al (2010) [20]</td>
</tr>
<tr>
<td>rs13266634</td>
<td>SLC30A8&lt;sup&gt;a&lt;/sup&gt;</td>
<td>G</td>
<td>8</td>
<td>118184783</td>
<td>Cds-nonsynon</td>
<td>Sladek et al (2007) [26]</td>
</tr>
<tr>
<td>rs4457053</td>
<td>ZBED3&lt;sup&gt;a&lt;/sup&gt;</td>
<td>G</td>
<td>5</td>
<td>76424949</td>
<td>Intergenic</td>
<td>Voight et al (2010) [20]</td>
</tr>
<tr>
<td>rs7961581</td>
<td>TSPAN8&lt;sup&gt;a&lt;/sup&gt;</td>
<td>C</td>
<td>12</td>
<td>71663102</td>
<td>Intron</td>
<td>Zeggini et al (2008) [27]</td>
</tr>
<tr>
<td>rs1531343</td>
<td>HMGA2&lt;sup&gt;a&lt;/sup&gt;</td>
<td>C</td>
<td>12</td>
<td>66174894</td>
<td>Intergenic</td>
<td>Voight et al (2010) [20]</td>
</tr>
<tr>
<td>rs2237895</td>
<td>KCNQ1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>C</td>
<td>11</td>
<td>2857194</td>
<td>Intron</td>
<td>Yasuda et al (2008) [28]</td>
</tr>
<tr>
<td>rs12779790</td>
<td>CDC123&lt;sup&gt;a&lt;/sup&gt;</td>
<td>G</td>
<td>10</td>
<td>12328010</td>
<td>Intron</td>
<td>Zeggini et al (2008) [27]</td>
</tr>
<tr>
<td>rs243021</td>
<td>BCL11A&lt;sup&gt;a&lt;/sup&gt;</td>
<td>A</td>
<td>2</td>
<td>60584819</td>
<td>Intergenic</td>
<td>Voight et al (2010) [20]</td>
</tr>
<tr>
<td>rs10830963</td>
<td>MTNR1B&lt;sup&gt;a&lt;/sup&gt;</td>
<td>G</td>
<td>11</td>
<td>92708710</td>
<td>Intron</td>
<td>Dupuis et al (2010) [22]</td>
</tr>
<tr>
<td>rs7578597</td>
<td>THADA&lt;sup&gt;a&lt;/sup&gt;</td>
<td>T</td>
<td>2</td>
<td>43732823</td>
<td>Missense</td>
<td>Zeggini et al (2008) [27]</td>
</tr>
<tr>
<td>rs1758499</td>
<td>PTPRD&lt;sup&gt;a&lt;/sup&gt;</td>
<td>T</td>
<td>9</td>
<td>8879118</td>
<td>Intron</td>
<td>Tsai et al (2010) [25]</td>
</tr>
<tr>
<td>rs10923931</td>
<td>NOTCH2&lt;sup&gt;a&lt;/sup&gt;</td>
<td>T</td>
<td>1</td>
<td>120517959</td>
<td>Intron</td>
<td>Zeggini et al (2008) [27]</td>
</tr>
<tr>
<td>rs7593730</td>
<td>RBMS1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>C</td>
<td>2</td>
<td>161171454</td>
<td>Intron</td>
<td>Qi et al (2010) [29]</td>
</tr>
<tr>
<td>rs4607103</td>
<td>ADAMTS9&lt;sup&gt;a&lt;/sup&gt;</td>
<td>C</td>
<td>3</td>
<td>64711904</td>
<td>Intergenic</td>
<td>Zeggini et al (2008) [27]</td>
</tr>
<tr>
<td>rs1470579</td>
<td>IGF2BP2&lt;sup&gt;a&lt;/sup&gt;</td>
<td>C</td>
<td>3</td>
<td>185529080</td>
<td>Intron</td>
<td>Voight et al (2010) [20]</td>
</tr>
<tr>
<td>rs864745</td>
<td>JAZF1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>T</td>
<td>7</td>
<td>28180556</td>
<td>Intron</td>
<td>Zeggini et al (2008) [27]</td>
</tr>
<tr>
<td>rs972283</td>
<td>KLF14&lt;sup&gt;a&lt;/sup&gt;</td>
<td>G</td>
<td>7</td>
<td>130466854</td>
<td>Intergenic</td>
<td>Voight et al (2010) [20]</td>
</tr>
<tr>
<td>rs896854</td>
<td>TP53INP1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>T</td>
<td>8</td>
<td>95960511</td>
<td>Intron</td>
<td>Voight et al (2010) [20]</td>
</tr>
<tr>
<td>rs1552224</td>
<td>CENTD2&lt;sup&gt;a&lt;/sup&gt;</td>
<td>A</td>
<td>11</td>
<td>72433098</td>
<td>Intron</td>
<td>Voight et al (2010) [20]</td>
</tr>
<tr>
<td>rs7957197</td>
<td>HNF1A&lt;sup&gt;a&lt;/sup&gt;</td>
<td>T</td>
<td>12</td>
<td>121460686</td>
<td>Intron</td>
<td>Voight et al (2010) [20]</td>
</tr>
<tr>
<td>rs11634397</td>
<td>ZFAND6&lt;sup&gt;a&lt;/sup&gt;</td>
<td>G</td>
<td>15</td>
<td>80432222</td>
<td>Intergenic</td>
<td>Voight et al (2010) [20]</td>
</tr>
<tr>
<td>rs8042680</td>
<td>PRC1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>A</td>
<td>15</td>
<td>91521337</td>
<td>Intron</td>
<td>Voight et al (2010) [20]</td>
</tr>
<tr>
<td>rs9939609</td>
<td>FTO&lt;sup&gt;a&lt;/sup&gt;</td>
<td>A</td>
<td>16</td>
<td>53820527</td>
<td>Intron</td>
<td>Frayling et al (2007) [30]</td>
</tr>
<tr>
<td>rs4607517</td>
<td>GCK&lt;sup&gt;a&lt;/sup&gt;</td>
<td>A</td>
<td>7</td>
<td>44235668</td>
<td>Intergenic</td>
<td>Dupuis et al (2010) [22]</td>
</tr>
<tr>
<td>rs2191349</td>
<td>DGKB&lt;sup&gt;a&lt;/sup&gt;</td>
<td>T</td>
<td>7</td>
<td>15064309</td>
<td>Intergenic</td>
<td>Dupuis et al (2010) [22]</td>
</tr>
<tr>
<td>rs780094</td>
<td>GCKR&lt;sup&gt;a&lt;/sup&gt;</td>
<td>C</td>
<td>2</td>
<td>27741237</td>
<td>Intron</td>
<td>Dupuis et al (2010) [22]</td>
</tr>
<tr>
<td>rs340874</td>
<td>PROX1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>C</td>
<td>1</td>
<td>214159256</td>
<td>Intergenic</td>
<td>Dupuis et al (2010) [22]</td>
</tr>
<tr>
<td>rs4430796</td>
<td>HNF1B&lt;sup&gt;a&lt;/sup&gt;</td>
<td>G</td>
<td>17</td>
<td>36098040</td>
<td>Intron</td>
<td>Voight et al (2010) [20]</td>
</tr>
</tbody>
</table>

* Analysed CpG-SNPs
* Assay design problems
* Failed assay

Chr., chromosome; Cds-synon., coding-synonymous; Cds-nonsynon., coding-nonsynonymous

---

1038 Diabetologia (2013) 56:1036–1046
rs12779790 (CDC123) (Fig. 1). Moreover, the assay for rs4457053 (ZBED3) failed PCR amplification and hence was also excluded.

Genotyping analysis Islet donors were genotyped for the 16 CpG-SNPs for which we retrieved DNA methylation data. The following 12 SNPs were genotyped using the HumanOmniExpressBeadChip and iScan system (Illumina, San Diego, CA, USA) according to the manufacturer’s instructions: rs5015480, rs7901695, rs391300, rs11708067, rs7754840, rs13266634, rs564398, rs5945326, rs7961581, rs2237895, rs1801282 and rs2334499. Four SNPs (rs13292136, rs7578326, rs1531343 and rs1801214) were not included on the Illumina chip and were therefore genotyped using TaqMan allelic discrimination assays on the ABI7900 platform (assay IDs: C__11871167_10, C__29270128_10, C__1692218_10 and C__8841076_10; Applied Biosystems, Foster City, CA, USA). Furthermore, three CpG-SNPs, rs4712523, rs3821943 and rs12778642, which are in strong linkage disequilibrium (LD) with the three type 2 diabetes CpG-SNPs rs7754840, rs1801214 and rs5015480, respectively, were genotyped using the HumanOmniExpressBeadChip and iScan system (Illumina).

Genotype frequencies were in Hardy–Weinberg equilibrium. The genotype data analysed with either the Illumina chip or TaqMan assays showed 100% concordance with the pyrosequencing genotype data.

Associations between CpG-SNPs and DNA methylation The effects of the SNPs on DNA methylation of the CpG-SNP sites and surrounding CpG sites were analysed using the Kruskal–Wallis H test. The association between CpG-SNPs and DNA methylation was based on an additive genetic model and p values were adjusted for multiple testing using Bonferroni correction and significance threshold <0.05 after Bonferroni correction was considered significant. Statistical calculations were performed using IBM SPSS statistics 20 (SPSS, Chicago, IL, USA).

Analysis of mRNA expression in human islets Expression of mRNA was analysed in pancreatic islets from the human donors using the GeneChip Human Gene 1.0 ST array from Affymetrix (Santa Clara, CA, USA) as described previously by Olsson et al [14]. The array data were summarised and normalised using the robust multi-array analysis method and the software ‘Expression Console’ (Affymetrix), and were corrected for batch effects within the array data using COMBAT [15].

Associations between type 2 diabetes CpG-SNPs that exhibit differential DNA methylation and mRNA expression in human islets Associations between type 2 diabetes CpG-SNPs that exhibit differential DNA methylation and mRNA expression in the human islets were analysed using linear regression assuming an additive model. In the linear model, mRNA expression values were used as quantitative traits, SNP genotypes were encoded as 0, 1 or 2 according to their minor allele, and sex, age, BMI, islet purity and islet culture days were included as covariates. The analysis was restricted to genes located within a maximum distance of 500 kb from a CpG-SNP site. The analysis was performed by using the Matrix expression quantitative trait loci (eQTL) R package [16].

Associations between type 2 diabetes CpG-SNPs that exhibit differential DNA methylation and alternative mRNA splicing The Gene Array Analyzer (GAA) [17] was used together with mRNA expression data from the GeneChip Human Gene 1.0 ST array (Affymetrix) to test if type 2 diabetes CpG-SNPs that exhibit differential DNA methylation are associated with alternative splicing events in human islets. Each exon on the GeneChip Human Gene 1.0 ST array is covered by at least two probes, which is why the expression data from these arrays can be used to analyse alternative splicing events, as previously described [17]. The
SNP annotation and proxy (SNAP) search [19] was used to identify SNPs in LD with type 2 diabetes CpG-SNPs. The search was based on genotype data from the 1000 Genomes project with the Centre d’Etude du Polymorphisme Humain (Utah residents with northern and western European ancestry) (CEU) population panel. An r² threshold >0.8 from pairwise LD calculations and a distance limit of 500 kb from the query SNP were set as inclusion criteria for LD SNPs. To determine if the identified LD SNPs were CpG-SNPs, the DNA sequences were retrieved from the dbSNP database (www.ncbi.nlm.nih.gov/snp/). A SNP was considered to be a CpG-SNP if it introduces or removes a CpG site.

Results

Identification of CpG-SNPs associated with type 2 diabetes Of the 40 type 2 diabetes genome-wide-associated SNPs [2, 20–30], 19 (48%) of the risk alleles introduce or remove a CpG site and are considered CpG-SNPs (Table 2). Of these 19 CpG-SNPs associated with type 2 diabetes, ten SNPs are intergenic, five are located in introns and four are coding SNPs (Table 2).

CpG-SNPs associated with type 2 diabetes affect DNA methylation in human pancreatic islets The impact of these 19 type 2 diabetes-associated CpG-SNPs on DNA methylation in pancreatic islets from 84 human donors was examined (Fig. 1 and Table 2). Successful DNA methylation data were generated for 16 of the 19 type 2 diabetes-associated CpG-SNP loci representing the candidate genes CDKAL1, CHCHD9, DUSP8, IRS1, SRR, TSPAN8, CDKN2A, HHEX, PPARG, TCF7L2, DUSP9, SLC30A8, HMG2, KCNQ1, ADCY5 and WFS1 in human islets (Fig. 1). DNA methylation data were also generated for a total of 22 CpG sites located near the CDKN2A, HHEX, PPARG, TCF7L2, DUSP9, SLC30A8, HMG2, KCNQ1, ADCY5 and WFS1 type 2 diabetes-risk SNPs. As a result of either design or technical problems, DNA methylation data could not be generated for the remaining three type 2 diabetes CpG-SNP loci representing KCNJ11, ZBED3 and CDC123 (Fig. 1).

We then tested if the type 2 diabetes-associated CpG-SNPs are associated with differential DNA methylation of the CpG-SNP sites as well as the surrounding CpG sites. All 16 type 2 diabetes CpG-SNPs that generated successful data were associated with differential DNA methylation of the CpG-SNP site in the human islets (p≤2.3×10⁻⁵, Fig. 2a-p and ESM Table 2). In addition, six CpG-SNP loci representing CDKN2A, TCF7L2, HMG2, KCNQ1, ADCY5 and WFS1 were associated with differential DNA methylation of surrounding CpG sites (p<0.05, Fig. 2a-p and ESM Table 2).

Associations between type 2 diabetes CpG-SNPs that exhibit differential DNA methylation and mRNA expression in human islets Increased DNA methylation of certain genome regions may silence gene transcription [7, 11]. We therefore tested if carriers of CpG-SNP genotypes associated with increased DNA methylation have decreased mRNA expression of genes located in regions 500 kb up- and downstream of the CpG-SNPs using microarray expression data from the human islets. We found that the CpG-SNP genotype carriers of rs391300 and rs5945326 that have increased DNA methylation also have decreased mRNA expression of RPA1 and AVPR2, respectively, in human islets (Table 3).

Associations between type 2 diabetes CpG-SNPs that exhibit differential DNA methylation and alternative splicing events in human islets There are further data supporting a role for DNA methylation in alternative splicing events [31]. We hence tested if the type 2 diabetes CpG-SNPs that exhibit differential DNA methylation and are located in gene introns or exons are associated with differentially expressed exons assessed by splice index using Gene Array Analyzer (http://GAA.mpi-bn.mpg.de) [17]. We found that type 2
Fig. 2 Histograms showing the mean degree of DNA methylation for each genotype of 16 type 2 diabetes-associated CpG-SNPs and surrounding CpG sites in human pancreatic islets of 84 donors (a–p). *p < 0.05 for associations between the CpG-SNP and DNA methylation after Bonferroni correction (n=16). Data are presented as mean ± SD.

Table 3 Associations between type 2 diabetes CpG-SNPs that exhibit differential DNA methylation and mRNA expression in human pancreatic islets

<table>
<thead>
<tr>
<th>Chr.</th>
<th>SNP ID</th>
<th>Gene</th>
<th>SNP location</th>
<th>Genotype</th>
<th>n</th>
<th>Gene expression</th>
<th>Mean ± SD</th>
<th>β (SEM)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>rs391300</td>
<td>RPA1</td>
<td>~413.4 kb down</td>
<td>A A</td>
<td>12</td>
<td>574.3±54.3</td>
<td>574.3±54.3</td>
<td>-0.056 (0.027)</td>
<td>0.040</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A G</td>
<td>47</td>
<td>562.1±57.4</td>
<td>562.1±57.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G G</td>
<td>22</td>
<td>537.1±60.5</td>
<td>537.1±60.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>rs5945326</td>
<td>AVPR2</td>
<td>~268.1 kb up</td>
<td>A A</td>
<td>50</td>
<td>66.8±9.0</td>
<td>66.8±9.0</td>
<td>-0.052 (0.024)</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A G</td>
<td>9</td>
<td>64.1±5.1</td>
<td>64.1±5.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>G G</td>
<td>22</td>
<td>62.3±8.7</td>
<td>62.3±8.7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data are expressed as mean ± SD of unlogged gene expression values

β (SEM) and p values are based on a linear regression of log2 values of gene expression adjusted for age, sex, BMI, islet purity and islet culture days

Chr., chromosome
diabetes CpG-SNPs located in SLC30A8, WFS1, CDKAL1 and TCF7L2 were associated with alternative splicing events in the respective gene in the human islets (Table 4).

**Associations between type 2 diabetes CpG-SNPs that exhibit differential DNA methylation and measures of islet function**

We further examined if the type 2 diabetes CpG-SNPs that exhibit differential DNA methylation were associated with measures of islet function. As type 2 diabetes is characterised by decreased glucose-stimulated insulin secretion and elevated glucagon levels [32], we tested if the risk alleles of the type 2 diabetes CpG-SNPs were associated with lower insulin content and/or secretion as well as elevated glucagon secretion from the human islets. The CpG-SNPs rs11708067 (ADCY5) and rs5015480 (HHEX) were associated with insulin secretion, while rs11708067 (ADCY5) and rs2237895 (KCNQ1) were associated with glucagon secretion (Table 5). Moreover, the type 2 diabetes CpG-SNP rs564398 (CDKN2A) was associated with insulin content in the human islets (Table 5).

**Identification of CpG-SNPs in LD with the 19 type 2 diabetes-associated CpG-SNPs**

We next performed a systematic analysis to examine if the 19 type 2 diabetes CpG-SNPs are in LD with any nearby SNPs and if any of these LD SNPs are also CpG-SNPs using SNAP search [19] together with genotype data from the 1000 Genomes project. We identified a total of 295 SNPs, of which 91 are CpG-SNPs.

**Table 4** Associations between CpG-SNPs that exhibit differential DNA methylation and alternative splicing events, assessed by splice index in human pancreatic islets

<table>
<thead>
<tr>
<th>SNP ID</th>
<th>Gene</th>
<th>Affymetrix transcript cluster ID</th>
<th>Affymetrix probe set ID</th>
<th>Exon annotation for probe set</th>
<th>AA vs Aa</th>
<th>AA vs aa</th>
<th>Aa vs aa</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs13266634</td>
<td>SLC30A8</td>
<td>8148003</td>
<td>8148008</td>
<td>Exon 3</td>
<td>0.066</td>
<td>0.39</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.34</td>
<td>0.031</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.085</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rs1801214</td>
<td>WFS1</td>
<td>8093906</td>
<td>8093909</td>
<td>Exon 2</td>
<td>−0.029</td>
<td>0.70</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.23</td>
<td>0.004</td>
<td>0.011</td>
</tr>
<tr>
<td>rs7754840</td>
<td>CDKAL1</td>
<td>8117140</td>
<td>8117152</td>
<td>Exon 10</td>
<td>−0.039</td>
<td>0.40</td>
<td>−0.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.035</td>
<td>0.012</td>
<td>0.10</td>
</tr>
<tr>
<td>rs7901695</td>
<td>TCF7L2</td>
<td>7930537</td>
<td>7930538</td>
<td>Exon 1</td>
<td>−0.38</td>
<td>0.045</td>
<td>−0.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.56</td>
<td>0.033</td>
<td>0.96</td>
</tr>
</tbody>
</table>

AA, homozygotes for major allele; Aa, heterozygotes; aa, homozygotes for minor allele; p value, based on t test of gene-level-normalised intensity of exon

**Table 5** Associations between type 2 diabetes CpG-SNPs that exhibit differential DNA methylation and measures of islet function including glucose-stimulated insulin secretion (SI), glucagon secretion and insulin content in human pancreatic islets

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>SNP ID</th>
<th>Candidate gene</th>
<th>Genotype</th>
<th>n</th>
<th>Mean ± SD</th>
<th>β (SEM)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insulin secretion (SI)</td>
<td>rs11708067</td>
<td>ADCY5</td>
<td>AA</td>
<td>48</td>
<td>7.1±7.8</td>
<td>0.18 (0.08)</td>
<td>0.031</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AG</td>
<td>28</td>
<td>13.2±20.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GG</td>
<td>2</td>
<td>16.2±17.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glucagon (pg islet⁻¹h⁻¹, 8.3 mmol/l glucose)³</td>
<td>rs11708067</td>
<td>ADCY5</td>
<td>AA</td>
<td>3</td>
<td>7.0±2.6</td>
<td>−0.64 (0.15)</td>
<td>0.027</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AG</td>
<td>5</td>
<td>3.2±2.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glucagon (pg islet⁻¹h⁻¹, 16.7 mmol/l glucose)³</td>
<td>rs2237895</td>
<td>KCNQ1</td>
<td>AA</td>
<td>17</td>
<td>8.6±10.6</td>
<td>0.26 (0.12)</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AC</td>
<td>19</td>
<td>23.6±26.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CC</td>
<td>8</td>
<td>35.2±42.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insulin content (ng/ng DNA)</td>
<td>rs564398</td>
<td>CDKN2A</td>
<td>AA</td>
<td>24</td>
<td>4.5±2.5</td>
<td>0.10 (0.04)</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AG</td>
<td>36</td>
<td>6.9±3.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GG</td>
<td>11</td>
<td>6.5±4.1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data are expressed as mean ± SD. β (SEM) and p values are from linear regression adjusted for age, sex and BMI as covariates

³ Data available from a subset of the cohort
SNPs in strong LD ($r^2 > 0.8$) with the 19 type 2 diabetes CpG-SNPs (Table 6). The 295 LD SNPs can be found in ESM Table 3. We further tested if CpG-SNPs that are in LD with type 2 diabetes CpG-SNPs also give rise to differential DNA methylation in the CpG-SNP site. Three LD CpG-SNPs were selected for this analysis based on their location near a type 2 diabetes CpG-SNP with a large number of LD CpG-SNPs. DNA methylation was subsequently analysed in the three CpG-SNPs rs4712523, rs3821943 and rs12778642, which are in strong LD with the type 2 diabetes CpG-SNPs rs7754840 (CDKAL1), rs1801214 (WFS1) and rs5015480 (HHEX), respectively. The three analysed LD CpG-SNPs were associated with differential DNA methylation in the CpG-SNP site in human islets (Fig. 3a–c, ESM Table 4) and the differences in methylation were similar to the data seen in respective type 2 diabetes CpG-SNPs (Fig. 2).

**Discussion**

Approximately 40 type 2 diabetes-associated loci have been identified by GWASs [2]. Furthermore, a number of studies that have aimed to unravel the physiological mechanisms through which these loci predispose to diabetes have shown that most of the identified loci act through impaired insulin secretion while some affect insulin sensitivity [3, 33]. However, the molecular mechanisms through which these SNPs influence gene function and disease pathology are largely unknown. In this study, we found that 19 of 40 SNPs associated with type 2 diabetes introduce or remove CpG sites, which are potential sites for DNA methylation. Many of these CpG-SNPs have previously been associated with impaired insulin secretion, hence the focus on these CpG-SNPs in human pancreatic islets [2]. Interestingly, we show that all analysed type 2 diabetes CpG-SNPs are associated with differential DNA methylation at the SNP site. Moreover, some of the type 2 diabetes CpG-SNPs that exhibit differential DNA methylation are associated with gene expression, alternative splicing events and hormone secretion in the human islets. The 19 type 2 diabetes CpG-SNPs are in strong LD with 91 additional CpG-SNPs that also seem to exhibit differential DNA methylation at the SNP site. Our data suggest that interactions between genetic and epigenetic factors may affect type 2 diabetes as several risk SNPs that influence islet function, such as TCF7L2, HHEX, CDKN2A, SLC30A8, CDKAL1,
ADCY5 and WFS1 [2, 3], are CpG-SNPs that exhibit differential DNA methylation in human islets.

An important characteristic of CpG sites is their mutability, which makes them hotspots for human genetic disease [34]. Indeed, more than 30% of all point mutations in humans are a consequence of DNA methylation [35]. While hydrolytic deamination of cytosines produces uracil, which is a foreign base in the DNA sequence that is identified and fixed by the DNA repair system, methylated cytosines are hydrolytically deaminated to thymines, which are not identified as foreign and hence less likely to be repaired, resulting in a C→T transition [34], which is the most common point mutation in humans [34, 36]. Indeed, of the approximately 2.3 million C/T or G/A autosomal SNPs identified by the second-phase HapMap database, 34% are located within a CpG site [37] and some have also been associated with disease [12, 13, 36, 38]. However, to our knowledge no previous study has examined if the 40 type 2 diabetes-associated SNPs are CpG-SNPs.

CpG-SNPs may provide a possible molecular mechanism through which SNPs can influence a phenotype. One suggested scenario is that CpG-SNPs can affect local DNA methylation which, in turn, affects the expression of a gene by allowing or preventing the binding of CpG methyl-binding proteins [13]. Another scenario is that CpG-SNPs may affect transcriptional splicing via differential DNA methylation. Indeed, DNA methylation has been shown to play a role in the regulation of splicing and it has been suggested that exon DNA methylation may aid the splicing in distinguishing exons from introns [31]. Interestingly, recent studies of TCF7L2 variants have suggested that SNPs located in introns can affect gene function through splicing [39]. Intragenic DNA methylation has also been suggested to play a direct role in reducing or enhancing transcription elongation efficiency via alternative promoters located within gene bodies [40] or indirectly by regulating non-coding transcripts that then affect the transcription of the associated genes [41, 42]. Another possible mechanism is by affecting the recombination rate [37]. Our data suggest that differential DNA methylation in CpG-SNP sites of human islets may influence gene expression of nearby genes as well as alternative splicing when the CpG-SNP is located within a gene. This may subsequently affect islet function. Two of the analysed type 2 diabetes CpG-SNPs that show differential DNA methylation (rs391300 and rs5945326) are associated with differential gene expression and 50% of the analysed type 2 diabetes CpG-SNPs located in introns and/or exons affect alternative splicing events in the human islets. Moreover, several type 2 diabetes CpG-SNPs are associated with measures of islet function. It should be taken into account that our study included 84 donors of human islets and although this is one of the largest human islet cohorts existing, for some of the type 2 diabetes CpG-SNPs there are only a few or no donors homozygous for the rare variant, which may explain the lack of power to identify associations between more CpG-SNPs and gene expression or hormone measures. Other reasons for the lack of associations between some of the type 2 diabetes CpG-SNPs and the analysed phenotypes could be that some of the type 2 diabetes CpG-SNPs do not affect gene expression or hormone secretion but possibly cause downstream modifications of mRNA or protein translation, or other phenotypes such as genome stability and recombination. Future studies are needed to fully elucidate the role of differential DNA methylation in CpG-SNPs.

SNPs within TCF7L2 confer the highest genetic risk of developing type 2 diabetes [21, 24, 43–45]. While rs7903146 within intron 4 of TCF7L2 is the most studied SNP, there are multiple SNPs in strong LD with rs7903146, including rs7901695, that also show strong associations with type 2 diabetes [46, 47]. The aim of our study was to examine if known type 2 diabetes SNPs introduce or remove CpG sites. The SNP rs7903146 is not a CpG-SNP. However, other TCF7L2 SNPs associated with type 2 diabetes, including rs7901695, delete CpG sites and we hence included rs7901695 data in our study. While rs7901695 and rs7903146 are in full LD ($r^2=1$) in the islet donors included in our study, these two SNPs are in strong, but not full, LD in the CEU population based on genotype data from the 1000 genome project ($r^2=0.892$).

The causal variants remain to be identified for a number of the identified type 2 diabetes SNPs [2]. It should further be noted that many of the type 2 diabetes SNPs have been connected to a gene name based on location and/or biological relevance, rather than showing association with expression of that particular gene. Several of the type 2 diabetes SNPs may hence be associated with expression of genes other than the one to which they have been assigned. Indeed, we found that rs391300 and rs5945326 are associated with mRNA expression of RPA1 and AVPR2, respectively, in the human islets. RPA1 encodes replication protein A1, which is involved in DNA replication, and AVPR2 encodes the vasopressin receptor, type 2, which belongs to the seven-transmembrane-domain G-protein-coupled receptor (GPCR) superfamily that stimulates adenylate cyclase. However, whether these two genes affect islet function remains to be tested.

We found that a CpG-SNP may not only be associated with differential DNA methylation of the CpG-SNP site itself, but may also be associated with DNA methylation of surrounding CpG sites. These results are in agreement with another study that has reported CpG-SNPs to affect DNA methylation of other, nearby, CpG sites and not only within the CpG-SNP site [48]. The observed correlation in DNA methylation between the CpG-SNP site and the adjacent CpG sites may be due to the presence of two active sites on DNMT3a, which allow for the simultaneous methylation of two CpG sites separated by one DNA helical turn [49].
We also found 91 CpG-SNPs out of a total of 295 SNPs in strong LD with the 19 type 2 diabetes-associated CpG-SNPs. The LD CpG-SNPs seem to affect the DNA methylation pattern in a similar fashion to the type 2 diabetes CpG-SNP. Genome regions with a large number of CpG-SNPs in LD are therefore likely to affect the chromatin structure and subsequently gene function. 

Together, this study demonstrates how genetic (SNPs) and epigenetic (DNA methylation) mechanisms can interact and our data provide a possible molecular mechanism through which type 2 diabetes-associated SNPs that introduce or remove CpG sites may affect gene function through differential DNA methylation and consequently the phenotype of the disease.

Acknowledgements  We thank the Nordic Network for Clinical Islet Transplantation (JDRF award 31-2008-413), the tissue isolation teams and the Human Tissue Laboratory of the EXODIAB/Lund University Diabetes Centre for providing human pancreatic islets.

Funding  This investigation was supported by grants from the Swedish Research Council, Region Skåne, Knut and Alice Wallenberg Foundation, Novo Nordisk Foundation and Söderberg, Diabetes Fonden, Pålsson, Linné grant (B31 5631/2006).

Duality of interest  The authors declare that there is no duality of interest associated with this manuscript.

Contribution statement  TAD and CL designed the study and experiments. AHO, PV, PA and TR contributed to the design of the study and the experiments. TAD, AHO and TR performed the research. All authors analysed, interpreted and discussed the data. TAD and CL wrote the manuscript. All authors reviewed and revised the manuscript critically. All authors approved the final version of the manuscript. TAD and CL are the guarantors of this work and, as such, had full access to all of the study data and take responsibility for the integrity of the data.

Open Access  This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References