Tissue-specific alternative splicing of TCF7L2

Ludmila Prokunina-Olsson1,2,*, Cullan Welch3, Ola Hansson5, Neeta Adhikari6, Laura J. Scott6, Nicolle Usher3, Maurine Tong1, Andrew Sprau1, Amy Swift1, Lori L. Bonnycastle1, Michael R. Erdos1, Zhi He6, Richa Saxena7, Brennan Harmon8, Olga Kotova5, Eric P. Hoffman8, David Altshuler7, Leif Groop5, Michael Boehnke6, Francis S. Collins1 and Jennifer L. Hall3,4

1Genome Technology Branch, National Human Genome Research Institute, NIH, Bethesda, MD 20892, USA, 2Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, 8717 Grovemont Circle, Bethesda, MD 20892, USA, 3Lillehei Heart Institute, Department of Medicine and 4Developmental Biology Center, University of Minnesota, Minneapolis, MN 55455, USA, 5Lund University Diabetes Centre, Department of Clinical Sciences, Lund University, 20502 Malmö, Sweden, 6Center for Statistical Genetics, Department of Biostatistics, University of Michigan, 1420 Washington Heights, Ann Arbor, MI 48109, USA, 7Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA and 8Children’s National Medical Center, Center for Genetic Medicine Research (CGMR), Washington, DC 20010, USA

Received June 15, 2009; Revised and Accepted July 10, 2009

Common variants in the transcription factor 7-like 2 (TCF7L2) gene have been identified as the strongest genetic risk factors for type 2 diabetes (T2D). However, the mechanisms by which these non-coding variants increase risk for T2D are not well-established. We used 13 expression assays to survey mRNA expression of multiple TCF7L2 splicing forms in up to 380 samples from eight types of human tissue (pancreas, pancreatic islets, colon, liver, monocytes, skeletal muscle, subcutaneous adipose tissue and lymphoblastoid cell lines) and observed a tissue-specific pattern of alternative splicing. We tested whether the expression of TCF7L2 splicing forms was associated with single nucleotide polymorphisms (SNPs), rs7903146 and rs12255372, located within introns 3 and 4 of the gene and most strongly associated with T2D. Expression of two splicing forms was lower in pancreatic islets with increasing counts of T2D-associated alleles of the SNPs: a ubiquitous splicing form (P = 0.018 for rs7903146 and P = 0.020 for rs12255372) and a splicing form found in pancreatic islets, pancreas and colon but not in other tissues tested here (P = 0.009 for rs12255372 and P = 0.053 for rs7903146). Expression of this form in glucose-stimulated pancreatic islets correlated with expression of proinsulin (r² = 0.84–0.90, P < 0.00063). In summary, we identified a tissue-specific pattern of alternative splicing of TCF7L2. After adjustment for multiple tests, no association between expression of TCF7L2 in eight types of human tissue samples and T2D-associated genetic variants remained significant. Alternative splicing of TCF7L2 in pancreatic islets warrants future studies. GenBank Accession Numbers: FJ010164–FJ010174.

INTRODUCTION

Common genetic variants within the transcription factor 7-like 2 gene (TCF7L2, previously known as T-cell factor 4, or TCF-4) have been strongly associated with increased risk of type 2 diabetes (T2D) (1). The association with T2D and impaired function of pancreatic islets was confirmed in several genome-wide association studies (GWAS) (2–4) and follow-up association studies (5–23). TCF7L2 is a ubiquitous protein that belongs to a family of TCF/lymphoid enhancer factor (LEF)) transcription factors (3). Prior to being associated with increased T2D risk, TCF7L2 was known as an
important component of the WNT pathway (24,25) and as a regulator of the proliferative cellular compartment in the intestine (26). Several hypotheses have been proposed to explain how the genetic variants in TCF7L2 could increase T2D risk. TCF7L2 might affect the ability of pancreatic islets to produce insulin in response to stimulation by insulinotropic intestinal hormone Glucagon-Like Peptide 1 (GLP-1) or glucose. This hypothesis was based on the observations that TCF7L2 could directly regulate mRNA expression of proglucagon precursor in-vitro (27) and that the deletion of Tcf7l2 gene in a mouse knock-out model resulted in loss of gut epithelial secretory cells that produce GLP-1 and other hormones (26). In support of this hypothesis, it was demonstrated that carriers of risk variants of TCF7L2 have impaired incretin effect (28) and impaired GLP-1 stimulated insulin secretion (16,29). It is also possible that TCF7L2 has pleiotropic tissue-specific roles and alterations in one or more of these functions could lead to T2D.

The most significant genetic T2D association within TCF7L2 was detected for two intronic single nucleotide polymorphisms (SNPs), rs7903146 and rs12255372, located 50 kb from each other within a 92 kb linkage disequilibrium (LD) block spanning exon 4 and parts of introns 3 and 4 (1–15). All exons of TCF7L2 were sequenced in multiple patients and controls but no coding variation was identified (1) suggesting that T2D risk is associated with a non-coding variation. Genetic variations may affect levels of expression and splicing architecture of mRNA transcripts (30,31) thus suggesting that T2D risk is associated with a non-coding variation. Genetic variations may affect levels of expression and splicing architecture of mRNA transcripts (30,31) thus several studies have tested the relationships between TCF7L2 mRNA expression and genotypes of rs7903146 and rs12255372 in human tissues (28,32–38). However, these studies have been limited by the use of small sets of samples and tissue types. Alternative exons reported for TCF7L2 (39–41) represent a potentially important source of functional molecular variation, yet the tissue specificity of the alternative splicing and its relationship with genetic variation has not been fully established.

Here, we evaluated the splicing diversity and expression of TCF7L2 in a broad range of human tissues. We provide evidence that expression of alternative exons of TCF7L2 is tissue-specific. One splicing form was unique to pancreatic islets, pancreas and colon and not found in other tissues studied. Expression of this form correlated with proinsulin expression in glucose-stimulated pancreatic islets. Our study does not provide strong evidence for association of T2D-associated SNPs rs7903146 and rs12255372 with expression of alternatively spliced forms of TCF7L2 in several tissues but points out a potentially functional biological effect in pancreatic islets. Future studies in larger sets of samples should confirm and explore this finding.

RESULTS

Expression of assay ‘ex7–8’ of TCF7L2 in human tissues and test for association with rs7903146 and rs12255372

First, we measured the mRNA expression of TCF7L2 using an assay that targeted exons 7 and 8 of the gene (assay ‘ex7–8’, Fig. 1), and detected signals in all human tissues tested (Supplementary Material, Table S1). When normalized to expression of two housekeeping genes, β2 microglobulin (B2M) and Glyceraldehyde 3-phosphate dehydrogenase (GAPDH), the highest levels of expression were found in pancreas, colon, brain, small intestine and peripheral blood monocytes. The lowest levels of expression were found in resting and activated peripheral blood T and B cells and lymphoblastoid cell lines (Fig. 2 and Supplementary Material, Table S1). Expression of TCF7L2, measured with the assay ‘ex7–8’ in 311 samples of pancreas, pancreatic islets, colon, liver, monocytes and subcutaneous adipose tissue from individuals without T2D, was not associated with genotypes of the two most strongly T2D-associated SNPs, rs7903146 and rs12255372 (Table 1). On the basis of the number of samples and standard deviation in expression of assays in each tissue tested, we estimated the detectable fold difference between the groups of samples carrying the risk and non-risk alleles (Table 1). In this analysis, we had 80% power (with a type I error rate of 5%) to detect a 1.13-fold difference in expression in colon, a 1.33-fold difference in liver, a 1.37-fold difference in monocytes, 2.06-fold in skeletal muscle, a 2.34-fold difference in pancreatic islets and a 2.87-fold difference in adipose tissue (Table 1).

Splicing diversity of TCF7L2

Since our analysis did not reveal any association between expression of TCF7L2 measured by assay ‘ex7–8’ and SNPs rs7903146 and rs12255372, we hypothesized that these genetic variants might be associated with alternative splicing. Members of the TCF/LEF family of transcription factors use alternative transcription start sites (TSSs) to produce functionally distinct proteins (42–44). To identify all TSS of TCF7L2, we performed 5′ rapid amplification of cDNA ends (5′RACE) in an RNA pool of eight tissues (pancreas, pancreatic islets, colon, liver, monocytes, skeletal muscles, subcutaneous adipose tissue and lymphoblastoid cell lines). We identified several TSS: (1) TSS1, located at −536 bp from the first translation start site and corresponding to RefSeq sequence (NM_030756); (2) TSS2, located at position +83 bp within exon 1; (3) a cluster collectively referred as TSS3 located in the first intron of the gene (+200, +239, +302, +338 bp, Fig. 3 and Supplementary Material, Table S2). Only transcripts initiated from the TSS1 can be translated from a start codon within exon 1 and produce protein isoforms with the β-catenin binding domain encoded by exons 1 and 2. The next in-frame translation start site that can be used by transcripts expressed from the TSS2 and TSS3, is located within exon 3 (Fig. 3), and the resulting proteins would lack the β-catenin binding domain. Both the main (exon 1) and the
alternative translation start sites (exon 3) have weak Kozak consensus sequences, ‘aaaaaaATGC’ and ‘ttcatcATGA’, compared with a classical consensus site ‘gccaccATGG’ (45).

We also performed 3′RACE in an attempt to detect alternative splicing forms of the 3′-UTRs. We did not detect any short PCR fragments corresponding to mRNA transcripts with significantly truncated 3′-UTR (data not shown). We PCR amplified, cloned and sequenced full-length cDNA fragments starting from the first translation start site within exon 1 through the end of the transcript in exon 14 (based on RefSeq NM_030756). We did not attempt to clone full-length cDNAs initiated from the TSS2 because this transcription start is embedded from exon 1 of the TSS1 form and the products would represent a mixture of transcripts. Because of low expression from the TSS3, further study of the transcripts initiated from this TSS was not pursued in detail. Sequencing of full-length cDNA clones reconfirmed several known alternative exons (3a, 4a, 12, 13, 13a and 13b) and minor in-frame inclusions of 3, 12 and 15 bp in exons 4a, 6 and 8, respectively (39–41) (Supplementary Material, Table S2). All sequences for full-length cDNA expression constructs were deposited in the GenBank (FJ010164–FJ010174, Supplementary Material, Table S3).

### Tissue-specific expression of TCF7L2

We designed 13 expression assays (Fig. 4) that targeted the majority of the TCF7L2 splicing forms, we observed in human tissues. Expression of these assays was quantified in samples from pancreas, pancreatic islets, colon, liver, monocytes, skeletal muscles, subcutaneous adipose tissue and lymphoblastoid cell lines. Expression of an alternative exon 13b (assay ‘ex13–13b’) was detected only in pancreatic islets, pancreas and colon and this assay was not used in the analysis that included all tissue samples. The expression of assay ‘ex4–4a’ was low in all tissues and was not studied further. Supplementary Material, Fig. S1 illustrates the diversity in expression of each assay in human tissues. We performed Principal Component Analysis (PCA) on samples from eight tissues using four sets of expression assays (Fig. 5A–D). We observed better discrimination between pancreatic islets and other tissues based on expression of assays for C-terminal
alternative exons 12, 13, 13a and 14 of TCF7L2 (Fig. 5C) than those of assays detecting the N-terminal exons (Fig. 5A and B) or all expression assays together (Fig. 5D and Supplementary Material, Table S4 for relative expression of each splicing form in human tissues).

Test for association between expression of alternatively spliced forms of TCF7L2 and rs7903146 and rs12255372

In each tissue, we used all available covariates (age, sex, BMI, etc.) to identify factors that could affect TCF7L2 expression (Supplementary Material, Tables S5A–G). Since the covariates were not available for all the samples and in the absence of strong and consistent effects of any of these factors, they were not used in the expression analysis. In pancreatic islets of non-T2D individuals, increasing counts of T2D-associated risk alleles of rs12255372 were associated with lower expression of several assays targeting C-terminal alternative exons of TCF7L2 in up to 24 additional samples of pancreatic islets (Supplementary Material, Table S6B). In the joint analysis of up to 50 samples, expression of assays ‘ex13–14’ and ‘ex13–13b’, adjusted for sample set, remained lower in samples with increasing counts of T2D-associated risk alleles: for assay ‘ex13–14’ \( P = 0.018 \) for rs12255372 and \( P = 0.020 \) for rs7903146 (Table 2) and for assay ‘ex13–13b’ \( P = 0.009 \) for rs12255372 and \( P = 0.053 \) for rs7903146 (Table 2 and Fig. 6).

In pancreas (\( n = 46 \)), colon (\( n = 128 \)), monocytes (\( n = 65 \)), liver (\( n = 62 \)) and adipose tissue (\( n = 14 \)), expression of TCF7L2 splicing forms was not associated with rs7903146 or rs12255372 (Supplementary Material, Table S6C–G). In skeletal muscle (\( n = 25 \)), expression of several assays of TCF7L2 was significantly increased in carriers of risk alleles of rs12255372 (\( P = 0.049–0.005 \), Supplementary Material, Table S6H) but the effect was not proportional to the counts of risk alleles.

Test for association of expression of metabolic/endocrine factors with genotypes of rs7903146 and rs12255372

The counts of T2D-associated risk alleles of rs12255372 and rs7903146 were not associated with mRNA expression of several metabolic/endocrine factors important for T2D and expressed in pancreatic islets, pancreas or colon, such as proinsulin, proglucagon, ghrelin/obestatin, somatostatin, pancreatic and duodenum homeobox 1 (PDX1), proliferating cell nuclear antigen (PCNA) and vasoactive intestinal peptide (VIP) (Supplementary Material, Table S6A, C, D).

Test for correlation between expression of proinsulin and TCF7L2 assays in human pancreatic islets

We also correlated the expression of proinsulin and TCF7L2 assays in 23 samples of human pancreatic islets in unstimulated conditions and in 12 samples incubated for 24 h in the media with normal glucose concentration (5.5 mM) and with high glucose concentration (16.7 mM), corresponding to glucose blood level of T2D individuals. Expression of proinsulin most strongly correlated with expression of assay ‘ex13–13b’ both in 5.5 mM glucose condition, \( r^2 = 0.90, P < 0.0001 \) and in 16.7 mM glucose condition, \( r^2 = 0.84, P = 0.00063 \), but not in unstimulated samples (Table 3). Limited by small sample set of pancreatic islets (\( n = 12 \)), we did not address effects of each of SNPs on glucose-stimulated expression.

DISCUSSION

Family-based linkage studies identified a region on chromosome 10q25.3 as a locus associated with T2D and fine-mapping studies within this region identified TCF7L2 as a candidate gene for the association (1,46,47). Recent GWAS and follow-up studies confirmed several genetic variants within the TCF7L2 gene as the strongest risk factors for development of T2D (2–15) and impaired function of pancreatic islets (16–22). As the first step towards understanding the biology of this genetic association, we performed studies on
splicing and mRNA expression of TCF7L2 in human tissues. We show here: (1) the expression of TCF7L2 alternative exons represents a tissue-specific signature; (2) a unique splicing form of TCF7L2 is expressed in pancreatic islets, pancreas and colon but not in other tissues examined here; (3) expression of this splicing form correlates with expression of proinsulin in glucose-stimulated pancreatic islets; (4) expression of alternatively spliced forms of TCF7L2 in eight human tissues examined here is not associated with T2D-associated risk variants rs12255372 and rs7903146.

Non-coding genetic variations have been shown to increase disease risk by altering levels of mRNA expression and splicing architecture of mRNA transcripts (30,31). In this study, we identified several types of potentially functional splicing
vations within TCF7L2: alternative TSSs, alternative exons 3a and 4a encompassing the associated LD block; and alternative C-terminal exons 12, 13, 13a and 13b. Alternative TSSs can generate mRNA transcripts encoding proteins with opposing biological functions such as proliferation versus differentiation as described for LEF/TCF, c-Myc and p73 proteins (48–50). In the case of TCF7L2, the alternative promoters produce protein isoforms that can activate or repress the WNT pathway dependent on the presence of β-catenin binding domain encoded by exons 1 and 2 (51). Alternative exons 3a and 4a are located within a protein region responsible for interaction with a Groucho repressor protein (52). The combinatorial tissue-specific inclusion of alternative C-terminal exons 12, 13, 13a and 13b can activate alternative stop codons within exons 13a, 13b or 14 and create protein isoforms with short, medium and long reading frames (39,41). The function of these protein isoforms is unclear; however, only the long isoforms have binding sites for the C-terminal binding protein (CTBP), a negative regulator of the WNT pathway (53–55). Therefore, all alternative forms of TCF7L2 could be functionally important. However, the splicing variations tested here were not significantly associated with T2D-associated risk variants rs12255372 and rs7903146 in the human tissues tested.

A potentially biologically important observation in this study, though not strongly supported by statistics after adjustment for multiple tests, was a decrease in expression of two expression assays of TCF7L2, ‘ex13–14’ and ‘ex13–13b’, in pancreatic islets of carriers of T2D-associated risk variants rs12255372 and rs7903146. This observation is in line with a notion that depletion of TCF7L2 in pancreatic islets leads to increased apoptosis, decrease in proliferation and glucose-induced insulin secretion (56) found in carriers of risk variants of TCF7L2 (19–21). The assay ‘ex13–14’ detects a splicing form that encodes a protein with long reading frame and a CRARF protein domain formed upon splicing of exons 13 and 14. The CRARF domain is a potent activator of the WNT pathway and is the most conserved part of the TCF7L2 protein in all species (54,57). The assay ‘ex13–13b’ detects the most tissue-specific splicing form of TCF7L2 expressed in pancreatic islets, colon and at a lower level in pancreas but not in multiple samples representing liver, skeletal muscles, adipocytes, monocytes or lymphoblastoid cell lines tested here. Expression of this form in pancreatic islets was ~30-fold higher than in pancreas (Supplementary Material, Fig. S3). In pancreatic islets cultured at normal (5.5 mM) or high glucose (16.7 mM) conditions, mRNA expression of proinsulin most strongly correlated with expression of assay ‘ex13–13b’ (r² = 0.84–0.90, P < 0.00063), but not with other assays of TCF7L2. Perhaps, the splicing form detected by this assay marks a specific glucose-sensing or secretory cell type within pancreatic islets. Limited by a small sample set of glucose-induced islets (n = 12), we did not address effects of these SNPs on correlation with insulin expression. We did not observe association between counts of risk alleles of rs7903146 and rs12255372 and mRNA levels of proinsulin, proglucagon (the precursor of the insulinotropic hormone GLP-1) and several other metabolic/endocrine factors in pancreatic islets, pancreas and colon but this does not exclude the possibility that protein expression and secretion of these factors may be affected in carriers of the TCF7L2 risk variants.

The reports on association of TCF7L2 expression in human pancreatic islets and genotypes of T2D-associated variants within TCF7L2 are few. Lyssenko et al. (28) used the assay ‘ex7–8’ of TCF7L2 in 15 samples of human pancreatic islets and reported a significantly higher expression in carriers of rs7903146 risk genotypes. Osmark et al. (38) addressed the expression of several splicing forms in a set of 17 pancreatic islets but did not observe any association between expression of several TCF7L2 assays and genotypes of rs7903146. For several assays that targeted the same splicing forms (‘ex3a–4’, ‘ex11–14’, ‘ex11–13’, ‘ex11–13a’) we see the same results as reported by Osmark et al. The assays ‘ex13–14’ and ‘ex13–13b’ whose expression showed some evidence for association with T2D-associated variants of TCF7L2 in our study have not been previously used in any tissue. Additionally, we used larger sample sets for each tissue and for the first time studied expression of TCF7L2 in tissues such as pancreas, peripheral blood monocytes and colon.

We recognize the limitations of the approaches used in this study. Human tissue samples consist of multiple cell types that

<table>
<thead>
<tr>
<th>TCF7L2 assay</th>
<th>Unstimulated</th>
<th>Stimulated 5.5 mM glucose, 16.7 mM glucose</th>
<th>Stimulated 5.5 mM glucose, n = 12</th>
<th>P-value</th>
<th>P-value</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>TSS3</td>
<td>−0.296</td>
<td>NS</td>
<td>0.14</td>
<td>NS</td>
<td>0.17</td>
<td>NS</td>
</tr>
<tr>
<td>Ex3a–4</td>
<td>−0.407</td>
<td>NS</td>
<td>0.35</td>
<td>NS</td>
<td>0.29</td>
<td>NS</td>
</tr>
<tr>
<td>Ex11–13</td>
<td>−0.333</td>
<td>NS</td>
<td>0.61</td>
<td>0.035</td>
<td>0.57</td>
<td>NS</td>
</tr>
<tr>
<td>Ex11–13a</td>
<td>−0.481</td>
<td>0.020</td>
<td>0.46</td>
<td>NS</td>
<td>0.29</td>
<td>NS</td>
</tr>
<tr>
<td>Ex11–14</td>
<td>−0.436</td>
<td>0.038</td>
<td>0.63</td>
<td>0.028</td>
<td>0.46</td>
<td>NS</td>
</tr>
<tr>
<td>Ex12–13</td>
<td>−0.021</td>
<td>NS</td>
<td>0.73</td>
<td>0.007</td>
<td>0.54</td>
<td>NS</td>
</tr>
<tr>
<td>Ex13–13a</td>
<td>−0.158</td>
<td>NS</td>
<td>0.69</td>
<td>0.013</td>
<td>0.46</td>
<td>NS</td>
</tr>
<tr>
<td>Ex13–14</td>
<td>−0.292</td>
<td>NS</td>
<td>0.53</td>
<td>0.03</td>
<td>0.44</td>
<td>NS</td>
</tr>
<tr>
<td>Ex13–13b</td>
<td>0.189</td>
<td>NS</td>
<td>0.90</td>
<td>&lt;0.0001</td>
<td>0.84</td>
<td>0.00063</td>
</tr>
</tbody>
</table>

r²—Pearson correlation coefficient, NS—not significant, in bold—significant correlations, P < 0.05; Glucose-stimulated islets were incubated in 5.5 or 16.7 mM glucose media at 37°C for 24 h.
can exhibit diverse pattern of TCF7L2 expression. Tissue biopsies were taken during surgeries preceded by fasting and stress, thus these samples may not represent optimal physiological conditions for studies on T2D-related phenotypes. We attempted to identify and quantify the majority of splicing variation within TCF7L2. However, the analysis did not cover several minor splicing variations and yet unknown splicing forms of TCF7L2 that can exist and contribute to disease. The P-values for association between the counts of T2D-associated alleles and TCF7L2 expression were not adjusted for multiple testing. A Bonferroni correction would be overly stringent given the correlation between the 13 assays within each tissue. Our goal was to identify potential biologically important differences that should be tested in other sets of samples and with additional methods. We understand that with relatively small sample sets representing each tissue type, some of our results might be false-positives or false-negatives. Larger numbers of samples from each tissue type and, particularly, pancreatic islets, would increase the power to detect association between TCF7L2 expression and TCF7L2 SNP variants. Risk variants of TCF7L2 SNPs might be associated with several types of molecular changes relevant for T2D and our study addressed only one of these types, mRNA expression of alternative splicing forms. It is possible that there might be other important molecular phenotypes of TCF7L2 yet to be identified and examined. The C-terminal exons of TCF7L2 are located >100 kb away from the associated LD block. How the associated DNA variants can affect the mRNA expression and splicing at this distance, is not yet understood. A potential mechanism could be related to long-distance co-transcriptional regulation of mRNA expression and splicing (58,59).

In conclusion, our findings provide new evidence of tissue-specific patterns of alternative splicing of TCF7L2. Moreover, we provide the first suggestive evidence that decreased expression of one ubiquitous and one rare tissue-specific alternative splicing forms of the gene in pancreatic islets may be associated with T2D-associated genetic variants of TCF7L2. Future studies should confirm and explore the observed association between genetic variation within TCF7L2, tissue-specific alternative splicing and risk of T2D.

MATERIALS AND METHODS

Tissue samples

Purified human pancreatic islets in the US set were obtained from the Islet Cell Resource Centers (IRB exemption number 3072) and the National Disease Research Interchange (IRB exemption number 3269) under IRB exemption by the National Institutes of Health Office of Human Subjects Research. Pancreatic islets were provided by Washington University, St. Louis, MO, USA; City of Hope National Medical Center, Duarte, CA, USA; University of Miami, Miami, FL, USA and the Joslin Diabetes Center, Boston, MA, USA. Information on age, sex, BMI, islets purity and viability, duration of cold ischemic time (time between pancreas removal and islet preparation), cause of death, T2D status and CMV infection was available for these samples. Pancreatic islets in the Swedish set were provided by the Nordic Network for Clinical Islet Transplantation by the courtesy of Dr Olle Korsgren at Uppsala University, Sweden and Dr Jalal Taneera at the Human Tissue laboratory at Lund University Diabetes Centre, Sweden.

Human pancreas, colon and liver samples were biopsies obtained from the University of Minnesota Tissue Procurement Center, Minneapolis, MN, USA, and Children’s Hospital, Boston, MA, USA. For the majority of these samples clinical reports indicating primary diagnosis, T2D status, age, sex and BMI were available. All tissue biopsies were non-malignant tissues as confirmed by pathology reports. Skeletal muscles biopsies were collected from healthy individuals and patients with Duchenne or Becker Muscular dystrophy at the Research Center for Genetic Medicine, Children’s National Medical Center, Washington, DC, USA. Information about age, sex, BMI and disease status was available for these samples. Samples of fresh blood were from anonymous human blood donors and information about age and sex of the subjects were provided by the NIH Clinical Center Division of Transfusion Medicine. Monocytes were separated from fresh blood with CD14+ coated magnetic MicroBeads with AutoMacs (Miltenyi Biotec). Frozen samples of human subcutaneous adipose tissue were purchased from Zen-Bio, Inc. and information on age, sex, BMI and T2D status was available for these samples. IRB approval for de-identified tissues was obtained from all Institutions involved. All samples with known T2D were excluded (five samples in the whole set of all tissues). 293T (human embryonic kidney), HeLa (human cervical carcinoma), HepG2 (human hepatocellular carcinoma) were purchased from ATCC (the American Type Culture Collection) and lymphoblastoid cell lines from CEPH (Centre d’ Etude du Polymorphisme Humain) individuals were purchased from Coriell cell repositories.

Glucose-stimulated pancreatic islets

Twelve samples from the US set were split in six aliquots each. Three aliquots were incubated in CMRL-1066 cell culture media with 5.5 mM glucose while other three aliquots of the same samples were incubated with CMRL-1066 cell culture media with 16.7 mM glucose for 24 h.

Preparation of DNA and RNA

DNA from tissue samples was prepared with DNeasy Kit (Qiagen) except for adipose tissue where DNA was recovered after RNA preparation according to Trizol protocol (Invitrogen). RNA was prepared with Trizol (Invitrogen) followed by RNeasy kit (Qiagen). RNA from the second set of pancreatic islets RNA was isolated using the AllPrep DNA/RNA Mini Kit (Qiagen). No sign of RNA degradation was observed using agarose gel electrophoresis and Experion DNA 1 K gel chips (Bio-Rad). RNA samples from all additional human tissues and cDNA from purified blood fractions (Multiple Tissue cDNA panels) were purchased from Clontech.

Genotyping

Genotyping was performed with an allelic-discrimination TaqMan assays for rs12255372 and rs7903146 and the Univer-
sal Genotyping Master Mix (Applied Biosystems). Alternatively or as an additional control, genotyping was performed by primer extension method using a Sequenom MALDI-TOF platform (Broad Institute).

**cDNA preparation and qPCR**

cDNA was prepared in a 20 μl volume with 500 ng of DNase-treated total RNA with use of random hexamers and Superscript III reverse transcriptase (Invitrogen) (set 1) or with the Qiagen Omniscript RT kit with 1 μM dT18 oligomer and 3 μM random hexamer primers (set 2 (38)). DNA was diluted with water to 200 μl and an equivalent of 5–10 ng of total RNA was used for each reaction (except for lymphoblastoid cell lines where 2 μg of total RNA was used as a starting amount). All reactions were run in three or four technical replicates. Standard deviations of below 0.2 Ct between the replicates were acceptable for high-expressing assays and 0.5 Ct for low-expressing assays. Endogenous controls B2M (β-2 microglobulin) or GAPDH (Glyceraldehyde-3-phosphate dehydrogenase) were measured from the same cDNA preparations with expression assays Hs00187842_m1 for B2M and 4333764F for GAPDH (Applied Biosystems). All primers for custom-designed assays were created to amplify only from cDNA and TaqMan probes were placed on the junctions between specific exons (Supplementary Material, Table S7 for primer sequences and assay information). All quantitative PCR reactions were performed in a 10 μl volume using the 7900 Real-Time PCR System (ABI) in 384 well plates. Universal Expression Master Mix was used for all expression studies with TaqMan probes according to standard conditions (Applied Biosystems). The reactions with 2× SYBR Green PCR mix (Qiagen) and specific primers were carried out for 15 min at 95°C, followed by 40 cycles of 15 s at 95°C, 15 sec at 59°C and 45 s at 72°C. Post-PCR melting curve analysis was used after each run. Initially, gel-purified PCR fragments were sequenced to ensure the specificity of amplification and correct splicing.

**Data analysis**

Values from three to four technical replicates for each expression assay were averaged and normalized to the expression levels of BM2 and GAPDH run in separate reactions from the same cDNA preparation. Relative differences in expression were calculated according to formula: \( \frac{Ct_{reference} - Ct_{target}}{Ct_{target}} \), where ‘reference’ represents levels of expression (Ct values) of B2M, GAPDH or both genes and ‘target’ represents levels of expression of TCF7L2 assays; negative Ct means that a target is expressed lower than a reference gene; fold difference in expression between a reference and a target can be calculated as \( 2^{-\Delta C_{t}} \). Within each tissue and assay, the log-transformed expression values were tested for normality. The associations between specified variables were tested using univariate linear model with inclusion of covariates where indicated (SPSS 16.0 software). The SNP-TCF7L2 expression or SNP-T2D related gene expression analyses were used according to an additive genetic model. The PCA analysis on mean-centered expression values for selected assays was performed with GenEx software (MultiD). We estimated the fold-difference in gene expression between samples with and without the T2D associated risk genotype that could be detected with 80 or 95% power with a type 1 error rate of 5%. For the power calculations for assay ‘ex7–8’ have been performed with observed standard deviation for this assay in each tissue and the number of samples with and without T2D associated risk genotypes (StatMate software, GraphPad). The power analysis differs from our regression analysis of gene expression and allele count, but is a good approximation when the less frequent genotype homozygote samples are rare.

For each sample of glucose-stimulated pancreatic islets, expression values of technical triplicates were averaged and then normalized to expression of endogenous control, B2M. The normalized values were averaged for three aliquots of the same samples. Therefore, for each sample and experimental condition expression was measured nine times. The Pearson correlation coefficients between expression of TCF7L2 assays and proinsulin were calculated with SPSS 16.0.

**SUPPLEMENTARY MATERIAL**

Supplementary Material is available at HMG online.

**ACKNOWLEDGEMENTS**

A special thank you to Marjorie Carlson and Mona Svärdh for technical assistance, Sara Bowell and Drs Ralph Powell, Amy Skubitz, Stephen Schmechel, and Charlie Moldow for facilitation of the access to tissue samples through the University of Minnesota Tissue Procurement Facility, to Islet Cell Resource Centers and the National Disease Research Interchange Network for human islets. Human islets (Swedish cohort) were provided by the Nordic Network for Clinical Islet Transplantation by the courtesy of Dr Olle Korsgren at Uppsala University, Sweden and Dr Jalal Taneera at the Human Tissue laboratory at Lund University Diabetes Centre, Sweden. LPO was a Wenner-Gren Fellow.

Conflict of Interest statement. None declared.
FUNDING

The project was supported by NIH grant 1R21DK078029-01 (JH), Intramural Research Programs of NHGRI and NCI of NIH. Work at LUDC was funded by grants from the Swedish Research Council, the Wallenberg Foundation and the Novo Nordisk Foundation. Funding to pay the Open Access publication charges for this article was provided by the intramural research program of NCI/NIH.

REFERENCES

Human Molecular Genetics, 2009, Vol. 18, No. 20


