Article

Integrated Analysis of Tissue-Specific Promoter Methylation and Gene Expression Profile in Complex Diseases

Kibaick Lee [1],[†], Sanghoon Moon [1],[†], Mi-Jin Park [1], In-Uk Koh [1], Nak-Hyeon Choi [1], Ho-Yeong Yu [1], Young Jin Kim [1], Jinhwa Kong [1], Hee Gyung Kang [2] , Song Cheol Kim [3],* and Bong-Jo Kim [1],*

[1] Division of Genome Research, Center for Genome Science, Korea National Institute of Health, Chungcheongbuk-do 28519, Korea; kibaicklee@korea.kr (K.L.); moon.sanghoon@daum.net (S.M.); pmjbibi@korea.kr (M.-J.P.); kohinuk@korea.kr (I.-U.K.); knihep@korea.kr (N.-H.C.); hoyeong6428@korea.kr (H.-Y.Y.); inthistime@korea.kr (Y.J.K.); kjh594@korea.kr (J.K.)

[2] Department of Pediatrics, Seoul National University College of Medicine, Seoul 03080, Korea; kanghg@snu.ac.kr

[3] Department of Surgery, Asan Medical Center, AMIST, University of Ulsan College of Medicine, Seoul 05505, Korea

* Correspondence: drksc@amc.seoul.kr (S.C.K.); kbj6181@korea.kr (B.-J.K.)

† These authors contributed equally to this work.

check for
updates

Abstract: This study investigated whether the promoter region of DNA methylation positively or negatively regulates tissue-specific genes (TSGs) and if it correlates with disease pathophysiology. We assessed tissue specificity metrics in five human tissues, using sequencing-based approaches, including 52 whole genome bisulfite sequencing (WGBS), 52 RNA-seq, and 144 chromatin immunoprecipitation sequencing (ChIP-seq) data. A correlation analysis was performed between the gene expression and DNA methylation levels of the TSG promoter region. The TSG enrichment analyses were conducted in the gene–disease association network (DisGeNET). The epigenomic association analyses of CpGs in enriched TSG promoters were performed using 1986 Infinium MethylationEPIC array data. A correlation analysis showed significant associations between the promoter methylation and 449 TSGs' expression. A disease enrichment analysis showed that diabetes- and obesity-related diseases were high-ranked. In an epigenomic association analysis based on obesity, 62 CpGs showed statistical significance. Among them, three obesity-related CpGs were newly identified and replicated with statistical significance in independent data. In particular, a CpG (cg17075888 of *PDK4*), considered as potential therapeutic targets, were associated with complex diseases, including obesity and type 2 diabetes. The methylation changes in a substantial number of the TSG promoters showed a significant association with metabolic diseases. Collectively, our findings provided strong evidence of the relationship between tissue-specific patterns of epigenetic changes and metabolic diseases.

Keywords: tissue-specific expression; DNA methylation; type 2 diabetes; obesity

1. Introduction

The human body consists of more than 200 types of cells in various tissues that perform specific functions in different biological processes [1]. Tissue-specific genes (TSGs) are specifically expressed in their respective tissues, in contrast to housekeeping genes. The discovery of TSGs has broadened our understanding of the functions of various tissues [2–4], as well as their biological mechanisms [5], since the aberrant expression of TSGs may be implicated in various diseases [6].

Epigenetic events, including DNA methylation and histone modification, are key regulators of gene expression and phenotype [7]. Methylation changes in the promoter region generally act as silencers of downstream genes [8–10]. Reportedly, gene body methylation positively regulates gene expression [11–13]. Since the epigenetic level differs across human tissues, it is important to explore epigenetic factors to understand the mechanisms underlying TSGs.

To date, several studies have been conducted to examine the relationship between gene expression and DNA methylation. However, most of these studies had limitations in investigating the relationship between the epigenetic changes and gene expression levels. First, since array-based studies relied on predefined markers and/or hybridization efficiencies, their resolution was not sufficient to cover the entire genome. A systematic TSG analysis by massive mRNA-sequencing [14] would be required for this purpose. Second, cohort studies mainly used DNA samples from peripheral blood. Although blood DNA is an excellent material for epigenetic studies, and biological phenomena in the blood environment closely reflect those in target cells or tissues, explaining tissue specificity using only blood DNA remains difficult. Finally, previous studies focused on either a single target (single gene or single tissue) or performed case-control analyses. However, it remained challenging to investigate the correlations between DNA methylation and gene expression based on the inter-tissue and inter-individual differences. Most recently, Blake and colleagues applied a systematic analysis to gene expression and DNA methylation patterns in multiple tissues across species, utilizing a genome sequencing-based profiling to overcome the above-mentioned limitations [15]. Moreover, some functional studies were conducted on this subject [15–17].

In the present study, we performed an integrated analysis of matched samples, employing whole genome bisulfite sequencing (WGBS), mRNA-seq, and ChIP-seq (chromatin immunoprecipitation followed by sequencing), to investigate the relationship between the methylation of the promoter region and expression of the gene, based on tissue specificity. We also analyzed the gene expression pattern influenced by epigenetic regulation, according to the direction of correlation (positive or negative), and investigated the relevance of the disease by a network analysis and epigenetic association analysis. Consequently, we identified metabolic disease-associated CpG markers within the promoter of TSGs, involved in controlling gene expression in practice.

2. Results

2.1. Identification of TSGs by Gene Expression Patterns

All mRNA-seq data were uniformly pre-processed by the International Human Epigenome Consortium (IHEC) standard mRNA-seq pipeline (see Materials and Methods). In brief, all FASTQ files were mapped to human genome assembly GRCh37, and read counts of each gene were counted based on the GENCODE v19 gene model. A downstream analysis was performed on 11,111 protein-coding genes. We then profiled gene expression. We detected no batch effect in any group (Figure S1). Next, a pairwise correlation analysis was performed based on the transcripts per million (TPM) value. While samples from the same tissue were highly correlated, those from different tissues showed weaker correlations (mean Pearson's correlation coefficients were 0.956 and 0.721, respectively; Figure S2).

We identified 677 TSGs, with Tau ≥ 0.8 (threshold (the number of TSGs for adipocyte, fibroblast, islet, kidney, and skeletal muscle cells was 248, 62, 89, 241, and 37, respectively; Table S1)). Each TSG was highly expressed in specific tissues, while 677 non-specific genes, which were randomly sampled with Tau < 0.8, showed a ubiquitous expression in all tissues (Figure 1a). We re-analyzed the expression data from the Genotype-Tissue Expression (GTEx) version 7 [18], using the same metric to confirm tissue specificity. The distribution of Tau for both housekeeping [19] and non-housekeeping genes was comparable in both sets (Figure 1b,c). The number of genes common with the TSGs in the GTEx set was 160; the total number of genes is shown in Figure 1d.
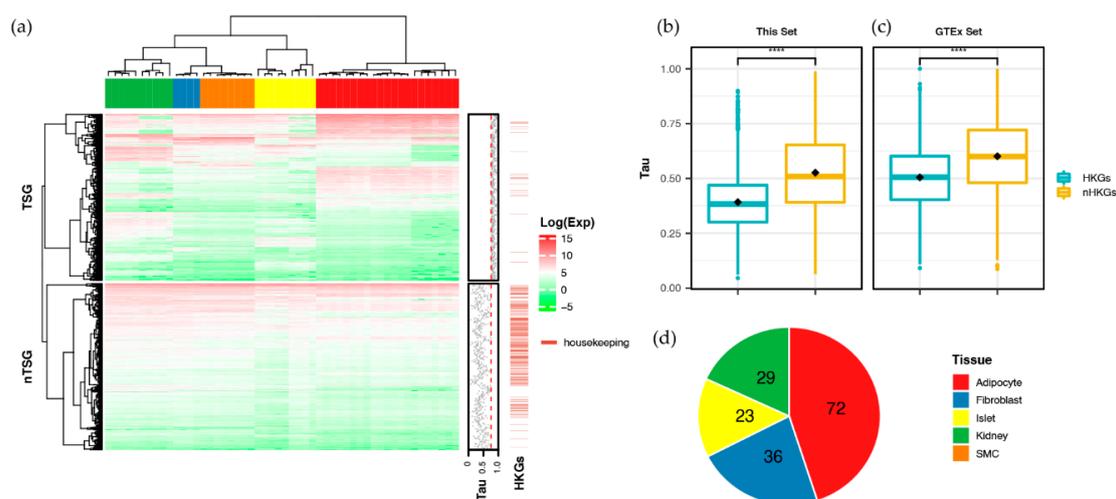
Figure 1. Gene expression revealed a tissue-specific pattern. (**a**) The gene expression pattern of protein-coding genes. (**b,c**) The distribution of Tau revealed whether the gene is housekeeping or not: our set (**b**) and the Genotype-Tissue Expression (GTEx) set (**c**). (**d**) A summary of all the common tissue-specific genes (TSGs) in the two data sets. nTSG, not TSG; HKGs, housekeeping genes; nHKGs, non-housekeeping genes.

2.2. Selection of Promoter Region Based on Methylation Levels

We assessed the methylation patterns of the promoter region for 11,111 protein-coding genes. The methylation level within the proximal promoter was profiled instead of that within the typical promoter (region from the transcription start site (TSS) to 2 kb upstream) due to its functional and regulative importance in transcriptional processes [20]. Furthermore, there are many transcription factor binding sites (TFBS) in the proximal promoter regions that are responsible for transcriptional regulation [21]; therefore, this region may have open chromatin and result in a low methylation level. We calculated the whole genome-scale methylation level of a single-base pair resolution using Python script in the whole genome bisulfite sequence MAPping program (BSMAP), and then extracted the methylation level of the promoter region based on the gene model of GENCODE v19. We obtained the average of all the promoter methylation levels. Our results revealed that the methylation level of proximal promoters was lower than that of a typical promoter. The mean methylation level of proximal and typical promoters was 9.3% and 23%, respectively (Figure S3).

2.3. Correlation Analysis Between Gene Expression and Promoter Methylation

To identify the relationship between gene expression and promoter methylation based on tissue specificity, we performed a correlation analysis of each gene. The correlation analysis in five tissues showed 449 significant expression–methylation correlations at the threshold (with an absolute value of Pearson's correlation coefficient ≥ 0.3) out of a total of 677 TSGs. The number of genes that negatively and positively correlated are 337 and 112, respectively (Table 1).

For the 337 negatively correlated genes, we performed a supervised hierarchical clustering analysis based on the methylation pattern of the proximal promoter region. We first divided the samples and methylation levels of the proximal promoters into five clusters based on their tissue identity (rows) and tissue specificity (columns). Thereafter, we clustered each proximal promoter using K-means clustering into five clusters. The proximal promoter of genes specific to a specific tissue tended to be mostly hypo-methylated in this tissue compared to other tissues (Figure 2a). In contrast, the 112 positively correlated genes showed the opposite pattern (Figure 2b).

Table 1. Number of correlated genes in five different tissues.

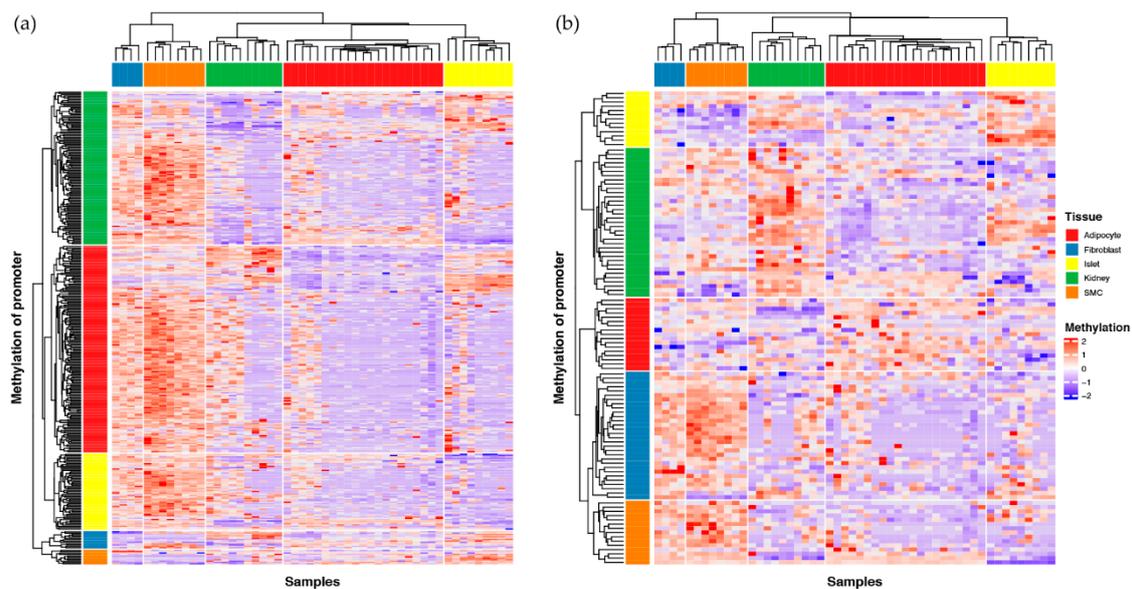| Tissue	Negatively Correlated Genes	Positively Correlated Genes
Adipocyte	162	19
Fibroblast	19	16
Islet	51	14
Kidney	91	49
SMC [1]	14	14
Total	337	112

[1] Skeletal muscle cell.

Figure 2. The proximal promoter methylation pattern according to the correlation between the gene expression and methylation levels. (**a**) The proximal promoter methylation patterns of 337 genes that were negatively correlated with gene expression ($\tau > 0.8$, Pearson's correlation coefficient < -0.3). (**b**) The proximal promoter methylation patterns of 112 genes that correlated positively with gene expression ($\tau > 0.8$, Pearson's correlation coefficient > 0.3).

2.4. Gene Set Enrichment Analysis for the Genes Affected by Methylation Perturbation

To identify the functional consequences underlying negatively and positively correlated genes, we performed a gene set enrichment analysis (GSEA) using EnrichR [22]. The pathway enrichment analysis with the Kyoto Encyclopedia of Genes and Genomes (KEGG) indicated the relatively high-rank of insulin- and adipocyte-related pathways for negatively correlated genes. The top 50 related pathways included "Insulin resistance" ($p = 4.88 \times 10^{-4}$), "Insulin signaling pathway" ($p = 5.33 \times 10^{-4}$), "Adipocytokine signaling pathway" ($p = 1.075 \times 10^{-3}$), and "peroxisome proliferator-activated receptor (PPAR) signaling pathway" ($p = 8.83 \times 10^{-3}$ (Table S2)). However, the pathways related to diverse biological process control and signal transduction, such as "axon guidance", "Apoptosis", and "neural crest cell development", were enriched in the positively correlated gene set.

We next investigated the disease association of the negatively correlated genes using the gene–disease association network (DisGeNET) [23] in EnrichR and the Cytoscape plugin [24] (Figure 3). Disease enrichment analysis showed diabetes- and obesity-related diseases to be high-ranked. The top 50 diseases included "Obesity" ($p = 1.24 \times 10^{-13}$), "Diabetes" ($p = 2.68 \times 10^{-10}$), "Diabetes Mellitus, Non-Insulin-Dependent" ($p = 3.79 \times 10^{-10}$), and "Diabetes Mellitus" ($p = 6.43 \times 10^{-9}$ (Table S3)).

We also analyzed the chromatin status to confirm that the positively correlated genes contributed to metabolic diseases. Approximately, 21% (23/112) of positively correlated genes were actively

transcribed in the respective tissue (the red-colored rectangles from the fourth to eighth column in Figure S4).
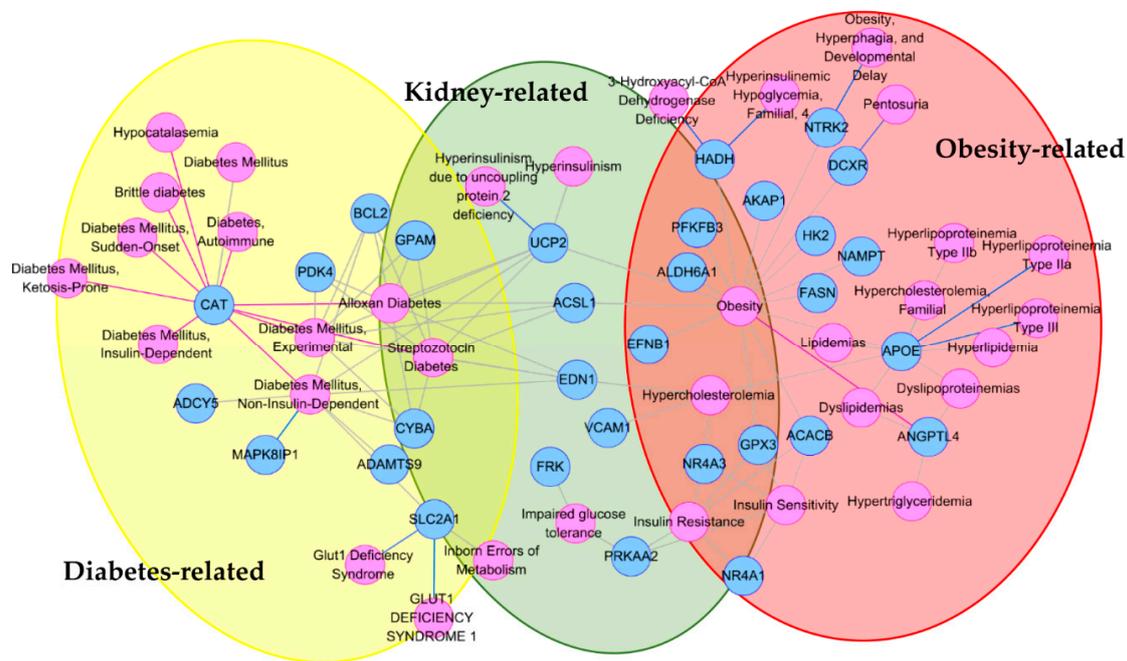

Figure 3. The gene–disease network for metabolic diseases. Pink nodes indicate disease names. Blue nodes represent the subsets of genes identified in our results. Red, yellow, and green circles correspond to obesity-, diabetes-, and kidney-related diseases, respectively.

2.5. Validation of the Epigenetic Markers Associated With Obesity

To confirm whether the TSG promoter methylation markers are associated with diseases, we conducted epigenomic association analyses, adjusted for age and gender, between differentially methylated probes (DMPs) on adipose-specific genes and obesity using 200 obese and 250 controls (Figure 4). A recent functional study suggested that the methylation of the first intron has both a positive and negative correlation with gene expression [25]. Thus, we expanded each range of the promoter regions from 2000 base pairs upstream to the first intron. In total, 4443 CpGs located in the expanded promoter region of adipocyte-specific genes were selected for the association analyses. Of the 4443 CpGs, 62 satisfied the significance threshold with the Bonferroni correction ($p = 1.13 \times 10^{-5}$, 0.05/4,443) in the discovery stage. Of the 62 CpGs, three CpGs were newly identified, and these satisfied the statistical significance criteria with a nominal *p*-value of 0.05 in the subsequent replication stage, using 759 independent data (Table 2, and Table S4). We further performed an annotation analysis of the epigenome-wide association study (EWAS (EWAS catalog and EWAS Atlas)) and the GWAS (GWAS catalog within ± 100 kb of CpGs). The EWAS annotation results showed that a CpG (cg27589809) was previously reported in lipid-related traits such as triglycerides, high-density lipoproteins, and the total cholesterol. Moreover, three CpGs were colocalized with the GWAS risk loci associated with obesity-related traits or metabolite levels within ± 100 kb (Table 2).

We performed epigenomic association analyses, adjusted for age, gender, and body mass index BMI, between four CpGs and Type 2 Diabetes (T2D) using 1534 data (Table 2). Interestingly, cg17075888 on the *PDK4* gene also showed statistical significance for T2D ($p = 2.30 \times 10^{-13}$). In addition, a re-analysis of the *PDK4* gene expression on omental adipose tissue using the Gene Expression Omnibus (GEO) expression data set (GDS3688) showed that the *PDK4* gene expression of obese was higher than that of controls (Figure S5).

Table 2. Obesity-associated methylation markers newly identified in the current study.

CpG ID	Chrom-osome (hg19)	Position	HGNC Symbol	Discovery ($n = 450$) p-Value	Replication ($n = 759$) p-Value	Replication ($n = 759$) p-Value (adj. T2D)	T2DM ($n = 1534$) p-Value (adj. BMI)	EWAS Catalog [1] ($p < 0.05$)	EWAS Atlas [2] ($p < 0.05$)	GWAS Catalog [3] (Disease Associated SNP within ± 100 kb at the CpG Site)
cg27589809 *	3	50650410	*CISH*	1.69×10^{-8}	1.61×10^{-2}	1.48×10^{-2}	8.31×10^{-1}	Triglycerides, Phospholipids to total lipids ratio, High-density lipoprotein cholesterol, Age, smoking, HIV infection	smoking	Waist-to-hip ratio adjusted for BMI, Height, Eosinophil counts
cg17075888 *	7	95225339	*PDK4*	4.75×10^{-6}	1.10×10^{-8}	1.02×10^{-4}	2.30×10^{-13}			Body Mass Index (rs6465468, reported gene: *ASB4*), Metabolite levels
cg20560869	12	52447054	*NR4A1*	1.44×10^{-6}	3.45×10^{-3}	1.56×10^{-2}	5.89×10^{-1}			Metabolite levels, Lung function, neutrophil eosinophil counts, Red blood cell count, Mean corpuscular hemoglobin, Interleukin-13 levels

[1] EWAS Catalog (http://www.ewascatalog.org/), [2] EWAS Atlas (https://bigd.big.ac.cn/ewas), [3] GWAS Catalog (https://www.ebi.ac.uk/gwas/). * These CpG markers are shown in Figure S6.

3. Discussion

In this study, we aimed to address whether the differences in methylation in the promoter region affect the expression of the tissue-specific genes potentially associated with diseases.

We profiled the DNA methylation and gene expression of five different types of tissues using matched samples, and investigated the TSGs affected by DNA methylation perturbation. Excluding the non-TSGs that were expressed either high or low across tissues (lower cluster in Figure 1a), we identified 677 genes (TSGs) specifically expressed in their respective tissues (upper cluster in Figure 1a). Concerning the GTEx, we identified the shared genes expressed in a specific tissue. Approximately 20% (160/677) were shared by both the data sets (Figure 1d); this percentage might reflect the differences in the data size between our set ($n = 52$) and the GTEx ($n = 1,997$). Of the 677 TSGs, 449 were affected by methylation perturbation in the proximal promoter and showed a dominant, negative regulation; approximately 75% (337 of 449) of the TSGs were negatively correlated, whereas 25% (112 of 449) were positively correlated with the proximal promoter methylation pattern (Table 1 and Table S5). This regulation pattern supported the findings of a recent cancer genome study [26]. Figure 2 shows a heatmap of the proximal promoter methylation patterns for 337 negative (Figure 2a) and 112 positive correlations (Figure 2b) of TSGs using supervised hierarchical clustering.

To explore the functional role of TSGs, we conducted a GSEA. The TSGs that negatively correlated with the promoter methylation were mostly enriched in metabolic diseases, such as diabetes, obesity-related pathways, or ontologies (Table S2). For example, *FASN* is a gene encoding for a fatty acid synthase and is known to play a key role in the regulation of obesity [27]; an increased *FASN* expression was associated with an impaired insulin sensitivity [28]. *FASN* might also play an important role in the development of obesity-related T2D [29]. The gene–disease network in DisGeNET showed the TSGs negatively regulated by promoter methylation to be tightly interconnected, and their roles to be distinctly enriched in T2D, obesity, and kidney function (Figure 3). Particularly, hub genes, such as *UCP2*, *CAT*, and *APOE*, having ≥ 6 edges (diseases) in the network, showed an interesting functional impact. Of note, *UCP2* is reportedly associated with insulin secretion, fatty acid metabolism, and glucose sensing [30–32].

Contrary to negative correlation, most TSGs that positively correlated with promoter methylation were not functionally enriched in the GSEA, although some of them, such as *IGFBP2*, had associations with diabetes, insulin resistance, insulin sensitivity, and obesity [33]. We speculated that the gene set might have been too small to identify a functional enrichment. Based on our results, we suggested the possibility that positively correlated genes are associated with chromatin states and alternative epigenetic regulatory mechanisms. Figure S4 represents the specific tissues and chromatin states of five tissues for each positively correlated gene. Most of the red color matching the tissue of the third column indicated the promoter of each gene to be an active promoter. According to the state annotation of the "Expanded 18-state model" in the Roadmap Epigenomics Project [34], red-related colors indicated active TSSs, flanking TSSs, and flanking TSS upstream states. These active states indicated enriched H3K4me3 and H3K27ac histone marks. Since H3K4me3 and H3K27ac are tightly related to chromatin openness [35], the existence of these marks implies that the region is being actively transcribed. Moreover, a previous study had suggested that the establishment of DNA methylation during early development is possibly mediated by histone H3K4me3 [36].

We confirmed that the promoter CpG status of TSG controlling gene expression is highly associated with metabolic diseases. From the epigenetic analysis using approximately 2000 participants, we identified 62 obesity-related DMPs. The disease association of 3, out of 62, DMPs was newly identified in the current study. Although there was no direct link with the EWAS, most of the genes with three embedded DMPs colocalized with the GWAS risk locus, suggesting a potential functional relevance, or a direct functional evidence related to obesity or obesity risk factors (Table 2).

A novel methylation marker, cg17075888 (*PDK4*) also seemed intriguing, since this gene was extensively researched for potential therapeutic targets; a *PDK4* inhibitor has been reported as a potential drug target for metabolic diseases, such as T2D [37]. Pyruvate dehydrogenase kinase 4,

encoded by *PDK4*, is an enzyme located in the matrix of the mitochondria, which catalyzes the oxidative decarboxylation of pyruvate to form acetyl-CoA [38,39]. Therefore, *PDK4* plays a key role in fatty acid metabolism, glucose metabolism, and the tricarboxylic acid (TCA) cycle [39–41]. Changes in the activity of the *PDK* family, including *PDK4*, mediate the inhibitory effect of fatty acids on glucose metabolism; thus, *PDK4* plays a vital role in obesity and metabolic diseases [42]. *PDK4* is also reported as one of the primary target genes of peroxisome proliferator-activated receptor γ (*PPARγ*) [43,44]. *PPARγ* is mainly located in adipose tissue and is involved in various functions such as adipocyte differentiation, lipid metabolism, insulin sensitivity, and fatty acid storage [44–46]. There is additional information on the relationship between *PDK4* and obesity. For example, the GEO expression data set (GDS3688) showed that the *PDK4* gene expression of obese was higher than that of controls (Figure S5). Several studies regarding the promoter methylation of *PDK4* have indicated the *PDK4* methylation levels to negatively correlate with BMI (body mass index) in skeletal muscle samples, and weight loss to be associated with methylation changes of *PDK4* promoters [47–52]. Furthermore, an intergenic variant, rs6465468 nearby the *ASB4* gene associated with BMI, is colocalized with *PDK4* (about 43.5 kb downstream of *PDK4* (Table 2 and Figure S6a)) [53]. However, to date, no previous study has identified which of the obesity-associated methylation markers directly affected the target gene expression. This study firstly suggests that a negative correlation between the hypo-methylated CpG (cg17075888) on the *PDK4* gene and its expression was highly associated with obesity.

Other novel methylation markers, cg27589809 (*CISH*), and cg20560869 (*NR4A1*), are associated with obesity. Cytokine inducible *SH2* containing protein, encoded by *CISH*, is a suppressor in the JAK/STAT pathway and is known to be involved in the regulation of activity of multiple important cytokines, including insulin, leptin, growth hormone, IL-6, prolactin, and interferons. In particular, leptin plays an important role in obesity-related factors, such as body fat storage, and the regulation of body weight [32]. A recent epigenome-wide association study (EWAS) has reported CpGs (cg23005227 and cg21585138) associated with *CISH* in African-Americans (Figure S6b) [42]. These two CpGs were located in exon 3 of *CISH* that were in 10 kb downstream of cg27589809 [54], suggesting that cg27589809 has a more causal effect on the regulation of the target gene expression than those of previously reported two markers. Nuclear receptor subfamily 4 group A member 1, encoded by *NR4A1*, is known to be a mediator of fasting-induced *PPARγ2* regulation in white adipose tissue (WAT) [55]. *NR4A1* is reported to have an association with chronic low-grade inflammation [56]. Obesity is associated with chronic low-grade inflammation, which results in insulin resistance, T2D, vascular disease, chronic renal failure, several cancers, and endocrine and behavioral abnormalities [57].

Considering that we only assessed the tissue specificity of 52 samples of 5 tissues, in future studies, more extensive data sets are required to enhance the statistical power of the analyses. As mentioned above, a relatively small-sized data set limited TSG detection, since the assessment of tissue specificity is fundamentally calculated based on the averaging expression value of tissue samples. To assess the tissue specificity accurately, it would be important to ensure both the tissue diversity and sample size. As one of the first integrative analyses of whole genome-scale DNA methylation and RNA sequencing from matched samples of human tissues [15,34,58], this study could also provide additional evidence of the metabolic disease-related gene–disease network and the EWAS results, considering the epigenetic regulation of gene expression.

4. Materials and Methods

4.1. Analysis Scheme

We generated WGBS, mRNA-seq, and histone modification ChIP-seq data from 9 pancreatic islets, 21 adipose tissues, 10 kidney tissues, 8 SMCs (skeletal muscle cells), and 4 fibroblasts isolated from humans (Tables S6 and S7). First, we profiled the gene expression from the RNA-seq data sets; the tissue specificity of each gene was evaluated by the Tau method [59]. We then defined the genes with a Tau value greater than a certain threshold as TSGs. Thereafter, we calculated the methylation

level of each promoter region (from 200 base pairs upstream to 50 base pairs downstream of the TSS (transcription start site)) using WGBS data sets. We next conducted a correlation analysis between the gene expression levels and promoter methylation for each gene. Each gene was classified according to the tendency of Pearson's correlation coefficient and by the tissue specificity. Genes of each group were subjected to an enrichment analysis using Enrich to predict the gene ontology, relationship with a disease, and biological pathways associated with chronic diseases. Besides, DisGeNET, a database of gene–disease associations, was analyzed to re-construct the disease and gene networks. Finally, we identified obesity-associated CpGs through an obesity EWAS analysis using about 2000 DNA methylation samples (Table S8). The analysis scheme of this study is illustrated in Figure 4.
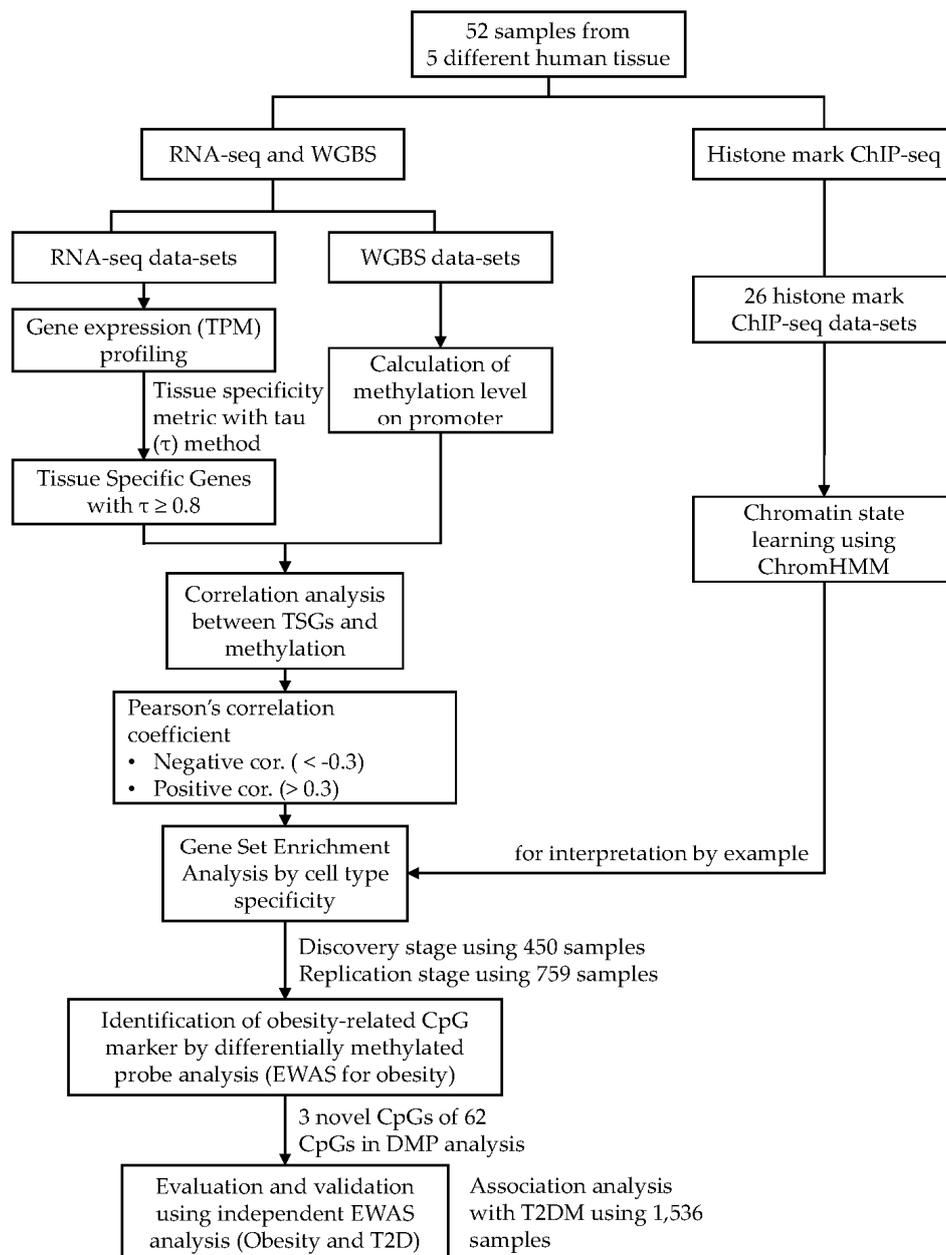
Figure 4. Schematic workflow.

4.2. Study Samples

Islet, adipocyte, and kidney samples were obtained from Seoul Asan Medical Center and Seoul National University Hospital. All tissues were collected from patients of pancreatectomy, obesity, or nephrectomy. The isolation of each tissue was performed according to the traditional cell isolation protocol, with some modification, depending on the tissue. In brief, the traditional protocol of mincing and digesting the tissues was performed, followed by a centrifugation to collect the cells [60]. Islets were digested by liberase or collagenase. After centrifugation, the adipocytes were isolated from the supernatant, whereas the kidney cells were present in the pellet. The kidney cells were separated into mesangial cells and podocytes using magnetic-activated cell sorting (MACS). All samples were aliquoted into three vials: two were reserved for DNA or RNA extraction, and one for a chromatin-fixation with formaldehyde for next-generation sequencing.

DNA samples for the Infinium EPIC array (Illumina, CA, USA) were recruited from the Health Examinees Study (HEXA). We conducted the association analysis using 450 samples (200 obese, BMI: 31.6 ± 1.6 kg/m^2 and 250 control, BMI: 21.4 ± 1.1 kg/m^2) in this cohort.

4.3. Ethics Approval and Consent to Participate

All of the islet, adipocyte, and kidney samples were collected after obtaining informed consent from the participants, in accordance with national laws and institutional ethical requirements. All procedures were approved by the Institutional Review Boards (IRB) of Seoul Asan Medical Center (S2013-0170-0008, approved on 6 August 2014) and Seoul National University Hospital (H-1302-095-466, approved on 19 June 2014).

Fibroblasts and skeletal muscle cells were obtained from the Wonkwang University Biobank of Korea, with written informed consent, under the Korea National Institute for Bioethics Policy (KoNIBP) and the corporation's IRB approval (P01-201606-31-001, approved on 7 June 2016).

All DNA samples for the DNA methylation array were from the National Biobank of Korea, and were obtained with written informed consent, following the KoNIBP and corporation's IRB (P01-201703-31-004, approved on 13 March 2017 and LASIRB-20180222-001/002, approved on 22 February 2018).

This study was approved by the institutional review board at the Korea National Institute of Health (2017-03-02-R-A, approved on 5 September 2017).

4.4. Processing of Sequencing Data

Poly A-captured mRNA libraries were sequenced on a HiSeq 2500 platform (Illumina, San Diego, CA, USA) with 100-bp paired-end reads. Raw sequencing data were trimmed off adaptor sequences by Trimmomatic (ver. 0.36) [61]. All trimmed reads were mapped to the GRCh37/hg19 reference genome using open-source software, STAR (version 2.5.3a, https://github.com/alexdobin/STAR, 18 Mar 2017) [62] in accordance with the ENCODE RNA-seq for long RNAs [58]. The expression levels from the aligned RNA-seq data were estimated by open-source software, RSEM (version 1.3.0, https://github.com/deweylab/RSEM, 10 Dec 2017) [63] based on the GENCODE release 19 gene annotation. Genes that expressed transcripts per million (TPM) > 0 in at least 1 of the samples were selected.

The sodium bisulfite converted DNA libraries were sequenced by the Illumina HiSeq 2500 and HiSeq X Ten systems, with 100-bp paired-end reads. Raw reads were trimmed off adapter sequences using Trimmomatic (ver. 0.36). All trimmed reads were aligned with the GRCh37/hg19 human reference genome using BSMAP (ver. 2.87) [64], with the option for reporting the unique mapped read. Duplicated reads were discarded using Picard (ver. 2.5.0) [65]. Methylation calling was conducted using Python script in the BSMAP tool for a sequencing depth of 10 or more.

Histone modification libraries were sequenced in an Illumina HiSeq 2500 system, with 100-bp paired-end reads. For each sample, 6 histone modification marks and inputs (control) were sequenced. All ChIP-seq reads were trimmed by Trimmomatic and mapped to the GRCh37/hg19 human reference

genome using the Burrows-Wheeler Alignment Tool with maximal exact matches (BWA-MEM) algorithm on a BWA mapper (ver. 0.7.15-r1140) [66] with the default setting. Unmapped and duplicated reads were marked and eliminated by Picard and SAMtools (ver. 0.1.19-96b5f2294a) [67]. To investigate the chromatin states for the promoter region of the histone modification samples, we conducted a ChromHMM (ver. 1.14) analysis [68,69]. Then, we assigned chromatin states of the samples using the "Expanded 18-state model" generated by Roadmap Epigenomic Consortium [34].

For the batch adjustment, we used the ComBat method of sva package in R [70,71].

4.5. Assessment of Tissue Specificity

We used τ (Tau) as a measure of the tissue specificity [59]. For all genes with a TPM > 1, the mean TPM of each of the five tissues was applied to the Tau formula:

$$\tau = \frac{\sum_{i=1}^{n}(1 - \hat{x}_i)}{n - 1}; \ \hat{x}_i = \frac{x_i}{\max\limits_{1 \le i \le n} x_i}$$

Tau ranges from 0 (expressed ubiquitously) to 1 (tissue-specific). We calculated the Tau values from the TPM of each gene in the sample and discarded those with a TPM of less than 1 in all samples. A recent study has shown Tau to be the best metric, among 9 methods, for measuring tissue specificity [59]. We finally selected the genes with Tau ≥ 0.8, and the highest expressed gene in a certain tissue was assigned as specific to that tissue (Pseudo-code in Supplementary Notes).

4.6. Correlation Analysis

We performed two types of correlation analyses. First, we investigated whether each assay showed tissue specificity, using a pairwise correlation test. For each expression and methylation matrix, the pairwise correlation coefficients across all samples were calculated by the cor function in R [72], using a log-transformed TPM value and the mean methylation level, respectively. A correlation plot was generated using the corrplot function of the corrplot package [73]. We then performed a correlation analysis to identify the strength of the relationship between the gene expression and promoter methylation level of the proximal promoter for each gene. We calculated the Pearson's correlation coefficients using the cor function in R.

4.7. Clustering Heatmap Analysis

We conducted clustering and constructed heat maps using the ComplexHeatmap package [74] in R. The clustering distance used the Pearson method and all other options used default values. We used log-transformed TPM values for Figure 1a, and the methylation value scaled by row (gene) for Figure 2, to draw the heat map.

4.8. Functional Annotation and Visualization

For each gene, positively or negatively correlated with the promoter methylation, we performed a KEGG pathway analysis, DisGeNET, and gene ontology enrichment test on EnrichR (http://amp.pharm.mssm.edu/Enrichr). We downloaded *p*-value-sorted text files and then extracted the top signals.

We conducted a network analysis for the negatively correlated genes using the DisGeNET Cytoscape Plugin on Cytoscape [75]. We ran the tool under the options "CURATED" for source, "ANY" for association type, "Nutritional and Metabolic Diseases", and disease class.

4.9. Identification of Obesity-Related CpG Marker

We conducted a differentially methylated probe (DMP) analysis for 168 adipocyte-specific genes in the HEXA cohort, as an independent cohort. In the DMP analysis for CpGs in the promoter region (extended from 2 kb upstream of the TSS to the first intron) of adipocyte-specific genes, we identified DMPs satisfying the Bonferroni corrected *p*-value ($p = 1.13 \times 10^{-5}$). The collected DMPs were validated

using another independent data set (of 759) as a replication set. Finally, to examine the relationship between these DMPs and T2D, we conducted a T2D association test using 1536 data sets with an adjusted BMI. We further confirmed whether the DMPs were reported as obesity-related traits in the EWAS catalog and EWAS Atlas, and also checked the disease-associated SNPs around these DMPs in the GWAS catalog. Next, we conducted an in silico validation using the Gene Expression Omnibus (GEO) expression data set on omental adipose tissue (GDS3688) [76].

4.10. Data Availability

The sequencing data analyzed during the current study are available in the repository: the European Genome-phenome Archive (EGA) database (https://www.ebi.ac.uk/ega/home) under the study accession number EGAS00001001774. In addition, the processed bigwig data are available in the IHEC data portal (https://epigenomesportal.ca/ihec/).

5. Conclusions

In this study, we analyzed the differences in the correlation tendencies between the gene expression and promoter methylation related to metabolic diseases. Correlation analysis between the promoter methylation and gene expression in different tissue types revealed a majority of metabolic diseases-related TSGs, including T2D and obesity. With the additional confirmation of the epigenetic changes in the EWAS, we comprehensively suggested novel epigenetic regulatory features of a potential drug target for T2D, namely *PDK4*.

Supplementary Materials: Supplementary materials can be found at http://www.mdpi.com/1422-0067/21/14/5056/s1. Table S1. Six-hundred and seventy-seven tissue-specific genes; Table S2. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway in EnrichR (gene set enrichment analysis (GSEA [$p < 0.05$])); Table S3. Top 100 gene–disease association network (DisGeNet) diseases in EnrichR (GSEA); Table S4. Replication of obesity-associated methylation markers newly identified in the current study; Table S5. Four-hundred and fourty-nine tissue-specific genes affected by proximal promoter methylation; Table S6. Sample information (number of mapped reads); Table S7. Characteristics of 52 subjects for sequencing data; Table S8. General characteristics of the subjects for an analysis on obesity-related differentially methylated CpGs (* body mass inded (BMI) > 30 kg/m^2, † BMI > 27 kg/m^2). Figure S1. Adjustment of batch effects. (a) PCA plot of 52 RNA-seq data after batch correction. (b) PCA plot of 52 whole genome bisulfite sequencing (WGBS) data after batch correction. (c) Distribution of log-transformed transcripts per million (TPM) value of RNA-seq data; Figure S2. Pairwise correlation analysis of RNA-seq. This analysis was performed using 11,111 genes with a TPM of at least 1 of all the samples of the protein code based on the GENCODE v19 gene model. Samples of intra-tissue revealed a high correlation, but samples of inter-tissue revealed weak correlations; Figure S3. Comparing typical and proximal promoters, where 2000 bp stands for the region from 2000 bp upstream of the transcription state site (TSS) to TSS, 250 bp stands for the region from 250 bp upstream of the TSS to TSS; Figure S4. Chromatin states for positively correlated genes with proximal promoter methylation. Red-, green-, yellow-, and grey-related colored states represent an active transcription, elongation, enhancer, and repressive transcription states, respectively; Figure S5. *PDK4* expression of a geo data set (GDS3688); Figure S6. Examples of obesity-associated methylation markers and related genes. The identified markers were obesity-associated markers that we identified. (a) cg17075888, located in the promoter of *PDK4*, (b) cg27589809, located in the promoter of cg27589809 (*CISH*).

Author Contributions: Conceptualization, K.L. and S.M.; methodology, K.L., and Y.J.K.; formal analysis, K.L., and M.-J.P.; investigation, N.-H.C., H.-Y.Y., J.K., and M.-J.P.; resources, S.C.K. and H.G.K.; data curation, K.L., N.-H.C., H.-Y.Y., and M.-J.P.; writing—original draft preparation, K.L.; writing—review and editing, I.-U.K. and S.M.; visualization, K.L. and M.-J.P.; supervision, B.-J.K.; project administration, B.-J.K. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

Abbreviations

TSG	Tissue-specific gene
WGBS	Whole Genome Bisulfite Sequencing
ChIP-seq	Chromatin Immunoprecipitation Followed by Sequencing
IHEC	International Human Epigenome Consortium
TPM	Transcripts per Million
GTEx	The Genotype-Tissue Expression
GSEA	Gene Set Enrichment Analysis
KEGG	Kyoto Encyclopedia of Genes and Genomes
TSS	Transcription Start Site
TFBS	Transcription Factor Binding Site
DisGeNET	Gene–disease Association Network
CKD	Chronic Kidney Disease

References

1. Winter, E.E.; Goodstadt, L.; Ponting, C.P. Elevated rates of protein secretion, evolution, and disease among tissue-specific genes. *Genome Res.* **2004**, *14*, 54–61. [CrossRef] [PubMed]

2. Schulze Blasum, B.; Schröter, R.; Neugebauer, U.; Hofschröer, V.; Pavenstädt, H.; Ciarimboli, G.; Schlatter, E.; Edemir, B. The kidney-specific expression of genes can be modulated by the extracellular osmolality. *FASEB J.* **2016**, *30*, 3588–3597. [CrossRef] [PubMed]

3. Passaro, A.; Miselli, M.A.; Sanz, J.M.; Dalla Nora, E.; Morieri, M.L.; Colonna, R.; Pišot, R.; Zuliani, G. Gene expression regional differences in human subcutaneous adipose tissue. *BMC Genom.* **2017**, *18*, 202. [CrossRef] [PubMed]

4. Sartorelli, V.; Kurabayashi, M.; Kedes, L. Muscle-specific gene expression. A comparison of cardiac and skeletal muscle transcription strategies. *Circ. Res.* **1993**, *72*, 925–931. [CrossRef] [PubMed]

5. Brenner, R.; Thomas, T.O.; Becker, M.N.; Atkinson, N.S. Tissue-specific expression of a Ca(2+)-activated K+ channel is controlled by multiple upstream regulatory elements. *J. Neurosci.* **1996**, *16*, 1827–1835. [CrossRef]

6. Lage, K.; Hansen, N.T.; Karlberg, E.O.; Eklund, A.C.; Roque, F.S.; Donahoe, P.K.; Szallasi, Z.; Jensen, T.S.; Brunak, S. A large-scale analysis of tissue-specific pathology and gene expression of human disease genes and complexes. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 20870–20875. [CrossRef]

7. Wahl, S.; Drong, A.; Lehne, B.; Loh, M.; Scott, W.R.; Kunze, S.; Tsai, P.-C.; Ried, J.S.; Zhang, W.; Yang, Y.; et al. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature* **2017**, *541*, 81–86. [CrossRef]

8. Jaenisch, R.; Bird, A. Epigenetic regulation of gene expression: How the genome integrates intrinsic and environmental signals. *Nat. Genet.* **2003**, *33*, 245–254. [CrossRef]

9. Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **2002**, *16*, 6–21. [CrossRef]

10. Luo, C.; Hajkova, P.; Ecker, J.R. Dynamic DNA methylation: In the right place at the right time. *Science* **2018**, *361*, 1336–1340. [CrossRef]

11. Bender, C.M.; Gonzalgo, M.L.; Gonzales, F.A.; Nguyen, C.T.; Robertson, K.D.; Jones, P.A. Roles of cell division and gene transcription in the methylation of CpG islands. *Mol. Cell. Biol.* **1999**, *19*, 6690–6698. [CrossRef] [PubMed]

12. Salem, C.E.; Markl, I.D.C.; Bender, C.M.; Gonzales, F.A.; Jones, P.A.; Liang, G. PAX6 methylation and ectopic expression in human tumor cells. *Int. J. Cancer* **2000**, *87*, 179–185. [CrossRef]

13. Kulis, M.; Heath, S.; Bibikova, M.; Queirós, A.C.; Navarro, A.; Clot, G.; Martínez-Trillos, A.; Castellano, G.; Brun-Heath, I.; Pinyol, M.; et al. Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia. *Nat. Genet.* **2012**, *44*, 1236–1242. [CrossRef]

14. Yang, R.Y.; Quan, J.; Sodaei, R.; Aguet, F.; Segrè, A.V.; Allen, J.A.; Lanz, T.A.; Reinhart, V.; Crawford, M.; Hasson, S.; et al. A systematic survey of human tissue-specific gene expression and splicing reveals new opportunities for therapeutic target identification and evaluation. *bioRxiv* **2018**, 311563. [CrossRef]

15. Blake, L.E.; Roux, J.; Hernando-Herraez, I.; Banovich, N.E.; Perez, R.G.; Hsiao, C.J.; Eres, I.; Cuevas, C.; Marques-Bonet, T.; Gilad, Y. A comparison of gene expression and DNA methylation patterns across tissues and species. *Genome Res.* **2020**, *30*, 250–262. [CrossRef]

16. Dezso, Z.; Nikolsky, Y.; Sviridov, E.; Shi, W.; Serebriyskaya, T.; Dosymbekov, D.; Bugrim, A.; Rakhmatulin, E.; Brennan, R.J.; Guryanov, A.; et al. A comprehensive functional analysis of tissue specificity of human gene expression. *BMC Biol.* **2008**, *6*, 49. [CrossRef]

17. Pai, A.A.; Bell, J.T.; Marioni, J.C.; Pritchard, J.K.; Gilad, Y. A genome-wide study of DNA methylation patterns and gene expression levels in multiple human and chimpanzee tissues. *PLoS Genet.* **2011**, *7*, e1001316. [CrossRef]

18. Lonsdale, J.; Thomas, J.; Salvatore, M.; Phillips, R.; Lo, E.; Shad, S.; Hasz, R.; Walters, G.; Garcia, F.; Young, N.; et al. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **2013**, *45*, 580–585. [CrossRef] [PubMed]

19. Eisenberg, E.; Levanon, E.Y. Human housekeeping genes, revisited. *Trends Genet.* **2013**, *29*, 569–574. [CrossRef]

20. Yanai, I.; Benjamin, H.; Shmoish, M.; Chalifa-Caspi, V.; Shklar, M.; Ophir, R.; Bar-Even, A.; Horn-Saban, S.; Safran, M.; Domany, E.; et al. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics* **2005**, *21*, 650–659. [CrossRef]

21. Zhao, X.; Xuan, Z.; Zhang, M.Q. Boosting with stumps for predicting transcription start sites. *Genome Biol.* **2007**, *8*, R17. [CrossRef] [PubMed]

22. Kuleshov, M.V.; Jones, M.R.; Rouillard, A.D.; Fernandez, N.F.; Duan, Q.; Wang, Z.; Koplev, S.; Jenkins, S.L.; Jagodnik, K.M.; Lachmann, A.; et al. Enrichr: A comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **2016**, *44*, W90–W97. [CrossRef] [PubMed]

23. Queralt-Rosinach, N.; Piñero, J.; Bravo, À.; Sanz, F.; Furlong, L.I. DisGeNET-RDF: Harnessing the innovative power of the Semantic Web to explore the genetic basis of diseases. *Bioinformatics* **2016**, *32*, 2236–2238. [CrossRef]

24. Bauer-Mehren, A.; Bundschus, M.; Rautschka, M.; Mayer, M.A.; Sanz, F.; Furlong, L.I. Gene-disease network analysis reveals functional modules in mendelian, complex and environmental diseases. *PLoS ONE* **2011**, *6*, e20284. [CrossRef] [PubMed]

25. Anastasiadi, D.; Esteve-Codina, A.; Piferrer, F. Consistent inverse correlation between DNA methylation of the first intron and gene expression across tissues and species. *Epigenetics Chromatin* **2018**, *11*, 1–17. [CrossRef] [PubMed]

26. Spainhour, J.C.G.; Lim, H.S.; Yi, S.V.; Qiu, P. Correlation patterns between DNA methylation and gene expression in the cancer genome atlas. *Cancer Inform.* **2019**, *18*, 1176935119828776. [CrossRef] [PubMed]

27. Diraison, F.; Dusserre, E.; Vidal, H.; Sothier, M.; Beylot, M. Increased hepatic lipogenesis but decreased expression of lipogenic gene in adipose tissue in human obesity. *Am. J. Physiol. Endocrinol. Metab.* **2002**, *282*, E46–E51. [CrossRef]

28. Berndt, J.; Kovacs, P.; Ruschke, K.; Klöting, N.; Fasshauer, M.; Schön, M.R.; Körner, A.; Stumvoll, M.; Blüher, M. Fatty acid synthase gene expression in human adipose tissue: Association with obesity and type 2 diabetes. *Diabetologia* **2007**, *50*, 1472–1480. [CrossRef]

29. Shen, J.; Zhu, B. Integrated analysis of the gene expression profile and DNA methylation profile of obese patients with type 2 diabetes. *Mol. Med. Rep.* **2018**, *17*, 7636–7644. [CrossRef]

30. Zhang, C.-Y.; Baffy, G.; Perret, P.; Krauss, S.; Peroni, O.; Grujic, D.; Hagen, T.; Vidal-Puig, A.J.; Boss, O.; Kim, Y.-B.; et al. Uncoupling protein-2 negatively regulates insulin secretion and is a major link between obesity, β cell dysfunction, and type 2 diabetes. *Cell* **2001**, *105*, 745–755. [CrossRef]

31. de Souza, B.M.; Assmann, T.S.; Kliemann, L.M.; Gross, J.L.; Canani, L.H.; Crispim, D. The role of uncoupling protein 2 (UCP2) on the development of type 2 diabetes mellitus and its chronic complications. *Arq. Bras. Endocrinol. Metabol.* **2011**, *55*, 239–248. [CrossRef] [PubMed]

32. Babon, J.J.; Nicola, N.A. The biology and mechanism of action of suppressor of cytokine signaling 3. *Growth Factors* **2012**, *30*, 207–219. [CrossRef]

33. Hoeflich, A.; Russo, V.C. Physiology and pathophysiology of IGFBP-1 and IGFBP-2—Consensus and dissent on metabolic control and malignant potential. *Best Pract. Res. Clin. Endocrinol. Metab.* **2015**, *29*, 685–700. [CrossRef]

34. Roadmap Epigenomics Consortium; Kundaje, A.; Meuleman, W.; Ernst, J.; Bilenky, M.; Yen, A.; Heravi-Moussavi, A.; Kheradpour, P.; Zhang, Z.; Wang, J.; et al. Integrative analysis of 111 reference human epigenomes. *Nature* **2015**, *518*, 317–329. [CrossRef] [PubMed]

35. Zhou, V.W.; Goren, A.; Bernstein, B.E. Charting histone modifications and the functional organization of mammalian genomes. *Nat. Rev. Genet.* **2011**, *12*, 7–18. [CrossRef] [PubMed]

36. Cedar, H.; Bergman, Y. Linking DNA methylation and histone modification: Patterns and paradigms. *Nat. Rev. Genet.* **2009**, *10*, 295–304. [CrossRef] [PubMed]

37. Lee, D.; Pagire, H.S.; Pagire, S.H.; Bae, E.J.; Dighe, M.; Kim, M.; Lee, K.M.; Jang, Y.K.; Jaladi, A.K.; Jung, K.Y.; et al. Discovery of novel pyruvate dehydrogenase kinase 4 inhibitors for potential oral treatment of metabolic diseases. *J. Med. Chem.* **2019**, *62*, 575–588. [CrossRef]

38. Gudi, R.; Bowker-Kinley, M.M.; Kedishvili, N.Y.; Zhao, Y.; Popov, K.M. Diversity of the pyruvate dehydrogenase kinase gene family in humans. *J. Biol. Chem.* **1995**, *270*, 28989–28994. [CrossRef]

39. Harris, R.A.; Bowker-Kinley, M.M.; Huang, B.; Wu, P. Regulation of the activity of the pyruvate dehydrogenase complex. *Adv. Enzyme Regul.* **2002**, *42*, 249–259. [CrossRef]

40. Zhang, M.; Zhao, Y.; Li, Z.; Wang, C. Pyruvate dehydrogenase kinase 4 mediates lipogenesis and contributes to the pathogenesis of nonalcoholic steatohepatitis. *Biochem. Biophys. Res. Commun.* **2018**, *495*, 582–586. [CrossRef]

41. Zhang, S.; Hulver, M.W.; McMillan, R.P.; Cline, M.A.; Gilbert, E.R. The pivotal role of pyruvate dehydrogenase kinases in metabolic flexibility. *Nutr. Metab.* **2014**, *11*, 1–9. [CrossRef] [PubMed]

42. Rosa, G.; Di Rocco, P.; Manco, M.; Greco, A.V.; Castagneto, M.; Vidal, H.; Mingrone, G. Reduced PDK4 expression associates with increased insulin sensitivity in postobese patients. *Obes. Res.* **2003**, *11*, 176–182. [CrossRef] [PubMed]

43. Degenhardt, T.; Saramäki, A.; Malinen, M.; Rieck, M.; Väisänen, S.; Huotari, A.; Herzig, K.H.; Müller, R.; Carlberg, C. Three members of the human pyruvate dehydrogenase kinase gene family are direct targets of the peroxisome proliferator-activated receptor β/δ. *J. Mol. Biol.* **2007**, *372*, 341–355. [CrossRef]

44. Lefterova, M.I.; Zhang, Y.; Steger, D.J.; Schupp, M.; Schug, J.; Cristancho, A.; Feng, D.; Zhuo, D.; Stoeckert, C.J.; Liu, X.S.; et al. PPARγ and C/EBP factors orchestrate adipocyte biology via adjacent binding on a genome-wide scale. *Genes Dev.* **2008**, *22*, 2941–2952. [CrossRef] [PubMed]

45. Moller, D.E.; Berger, J.P. Role of PPARs in the regulation of obesity-related insulin sensitivity and inflammation. *Int. J. Obes.* **2003**, *27*, S17–S21. [CrossRef] [PubMed]

46. Ferre, P. The biology of peroxisome proliferator-activated receptors: Relationship with lipid metabolism and insulin sensitivity. *Diabetes* **2004**, *53*, S43–S50. [CrossRef]

47. Van Otterdijk, S.D.; Binder, A.M.; Szarc Vel Szic, K.; Schwald, J.; Michels, K.B. DNA methylation of candidate genes in peripheral blood from patients with type 2 diabetes or the metabolic syndrome. *PLoS ONE* **2017**, *12*, e0180955. [CrossRef] [PubMed]

48. Sala, P.; De Miranda Torrinhas, R.S.M.; Fonseca, D.C.; Ravacci, G.R.; Waitzberg, D.L.; Giannella-Neto, D. Tissue-specific methylation profile in obese patients with type 2 diabetes before and after Roux-en-Y gastric bypass. *Diabetol. Metab. Syndr.* **2017**, *9*, 1–15. [CrossRef]

49. Davegårdh, C.; García-Calzón, S.; Bacos, K.; Ling, C. DNA methylation in the pathogenesis of type 2 diabetes in humans. *Mol. Metab.* **2018**, *14*, 12–25. [CrossRef]

50. Kulkarni, S.S.; Salehzadeh, F.; Fritz, T.; Zierath, J.R.; Krook, A.; Osler, M.E. Mitochondrial regulators of fatty acid metabolism reflect metabolic dysfunction in type 2 diabetes mellitus. *Metabolism* **2012**, *61*, 175–185. [CrossRef]

51. Dziewulska, A.; Dobosz, A.M.; Dobrzyn, A. High-throughput approaches onto uncover (Epi)genomic architecture of type 2 diabetes. *Genes* **2018**, *9*, 374. [CrossRef] [PubMed]

52. Barres, R.; Kirchner, H.; Rasmussen, M.; Yan, J.; Kantor, F.R.; Krook, A.; Näslund, E.; Zierath, J.R. Weight loss after gastric bypass surgery in human obesity remodels promoter methylation. *Cell Rep.* **2013**, *3*, 1020–1027. [CrossRef]

53. Hoffmann, T.J.; Choquet, H.; Yin, J.; Banda, Y.; Kvale, M.N.; Glymour, M.; Schaefer, C.; Risch, N.; Jorgenson, E. A large multiethnic genome-wide association study of adult body mass index identifies novel loci. *Genetics* **2018**, *210*, 499–515. [CrossRef]

54. Wang, X.; Pan, Y.; Zhu, H.; Hao, G.; Huang, Y.; Barnes, V.; Shi, H.; Snieder, H.; Pankow, J.; North, K.; et al. An epigenome-wide study of obesity in African American youth and young adults: Novel findings, replication in neutrophils, and relationship with gene expression. *Clin. Epigenetics* **2018**, *10*, 3. [CrossRef]

55. Duszka, K.; Bogner-Strauss, J.G.; Hackl, H.; Rieder, D.; Neuhold, C.; Prokesch, A.; Trajanoski, Z.; Krogsdam, A.M. Nr4a1 is required for fasting-induced down-regulation of Pparγ2 in white adipose tissue. *Mol. Endocrinol.* **2013**, *27*, 135–149. [CrossRef] [PubMed]

56. Huang, Q.; Xue, J.; Zou, R.; Cai, L.; Chen, J.; Sun, L.; Dai, Z.; Yang, F.; Xu, Y. NR4A1 is associated with chronic Low-Grade inflammation in patients with type 2 diabetes. *Exp. Ther. Med.* **2014**, *8*, 1648–1654. [CrossRef] [PubMed]

57. Margioris, A.N.; Dermitzaki, E.; Venihaki, M.; Tsatsanis, C. Chronic low-grade inflammation. In *Diet, Immunity and Inflammation*; Elsevier: Amsterdam, The Netherlands, 2013; pp. 105–120. ISBN 9780857090379.

58. Dunham, I.; Kundaje, A.; Aldred, S.F.; Collins, P.J.; Davis, C.A.; Doyle, F.; Epstein, C.B.; Frietze, S.; Harrow, J.; Kaul, R.; et al. An integrated encyclopedia of DNA elements in the human genome. *Nature* **2012**, *489*, 57–74.

59. Kryuchkova-Mostacci, N.; Robinson-Rechavi, M. A benchmark of gene expression tissue-specificity metrics. *Brief. Bioinform.* **2017**, *18*, 205–214. [CrossRef]

60. Orr, J.S.; Kennedy, A.J.; Hasty, A.H. Isolation of adipose tissue immune cells. *J. Vis. Exp.* **2013**, *75*, e50707. [CrossRef] [PubMed]

61. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [CrossRef] [PubMed]

62. Dobin, A.; Davis, C.A.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T.R. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **2013**, *29*, 15–21. [CrossRef]

63. Li, B.; Dewey, C.N. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* **2011**, *12*, 323. [CrossRef] [PubMed]

64. Xi, Y.; Li, W. BSMAP: Whole genome bisulfite sequence MAPping program. *BMC Bioinform.* **2009**, *10*, 1–9.

65. Picard Tool. Available online: http://broadinstitute.github.io/picard (accessed on 27 April 2017).

66. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* **2013**, arXiv:1303.3997.

67. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. The sequence alignment/map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [CrossRef] [PubMed]

68. Ernst, J.; Kellis, M. ChromHMM: Automating chromatin-state discovery and characterization. *Nat. Methods* **2012**, *9*, 215–216. [CrossRef] [PubMed]

69. Ernst, J.; Kellis, M. Chromatin-state discovery and genome annotation with ChromHMM. *Nat. Protoc.* **2017**, *12*, 2478–2492. [CrossRef]

70. Johnson, W.E.; Li, C.; Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **2007**, *8*, 118–127. [CrossRef]

71. Leek, J.T.; Johnson, W.E.; Parker, H.S.; Jaffe, A.E.; Storey, J.D. The SVA package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* **2012**, *28*, 882–883. [CrossRef] [PubMed]

72. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2017.

73. Wei, T.; Simko, V.; Levy, M.; Xie, Y.; Jin, Y.; Zemla, J. Package "Corrplot": Visualization of a Correlation Matrix (Version 0.84). Available online: https://github.com/taiyun/corrplot (accessed on 25 March 2019).

74. Gu, Z.; Eils, R.; Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **2016**, *32*, 2847–2849. [CrossRef]

75. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504. [CrossRef] [PubMed]

76. Aguilera, C.M.; Gomez-Llorente, C.; Tofe, I.; Gil-Campos, M.; Cañete, R.; Gil, Á. Genome-wide expression in visceral adipose tissue from obese prepubertal children. *Int. J. Mol. Sci.* **2015**, *16*, 7723–7737. [CrossRef] [PubMed]