

Article

LncRNA-Associated Genetic Etiologies Are Shared between Type 2 Diabetes and Cancers in the UAE Population

Roberta Giordo ¹, Rida Gulsha ¹, Sarah Kalla ¹, George A. Calin ^{2,3} and Leonard Lipovich ^{1,*}

¹ College of Medicine, Mohammed Bin Rashid University of Medicine and Health Sciences, Dubai 505055, United Arab Emirates; roberta.giordo@mbru.ac.ae (R.G.); rida.gulsha@students.mbru.ac.ae (R.G.); sarah.kalla@students.mbru.ac.ae (S.K.)

² Department of Translational Molecular Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA; gcalin@mdanderson.org

³ Center for RNA Interference and Non-Coding RNAs, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA

* Correspondence: leonard.lipovich@mbru.ac.ae

Simple Summary: The genome-wide association study (GWAS) approach to common human disease relies on single nucleotide polymorphisms (SNPs), the most common type of genetic variation in the human genome, and distinguishes “risk” and “healthy” SNP alleles. In parallel with increasing insights into the non-coding genome, emerging studies reveal that most disease-associated SNPs reside within a non-coding sequence, including lncRNA genes. These developments lay the foundation for deciphering the aetiology of complex diseases, including type 2 diabetes (T2D), and its association with an increased risk of certain cancers. Here, deploying a customized annotation pipeline on GWAS datasets, we successfully identified, and characterized, six genetic variants significantly associated with both T2D and cancer in lncRNA or genes and other non-coding regions. These variants suggest evidential proof of a shared genetic architecture between the two diseases, help to functionally explain the casual association of diabetes with cancer, and comprise a potential shortlist of candidate drug targets.

Abstract: Numerous epidemiological studies place patients with T2D at a higher risk for cancer. Many risk factors, such as obesity, ageing, poor diet and low physical activity, are shared between T2D and cancer; however, the biological mechanisms linking the two diseases remain largely unknown. The advent of genome wide association studies (GWAS) revealed large numbers of genetic variants associated with both T2D and cancer. Most significant disease-associated variants reside in non-coding regions of the genome. Several studies show that single nucleotide polymorphisms (SNPs) at or near long non-coding RNA (lncRNA) genes may impact the susceptibility to T2D and cancer. Therefore, the identification of genetic variants predisposing individuals to both T2D and cancer may help explain the increased risk of cancer in T2D patients. We aim to investigate whether lncRNA genetic variants with significant diabetes and cancer associations overlap in the UAE population. We first performed an annotation-based analysis of UAE T2D GWAS, confirming the high prevalence of variants at or near non-coding RNA genes. We then explored whether these T2D SNPs in lncRNAs were relevant to cancer. We highlighted six non-coding genetic variants, jointly reaching statistical significance in T2D and cancer, implicating a shared genetic architecture between the two diseases in the UAE population.

Keywords: T2D; cancer; SNP; lncRNA; UCSC Genome Browser; GWAS



Citation: Giordo, R.; Gulsha, R.; Kalla, S.; Calin, G.A.; Lipovich, L. LncRNA-Associated Genetic Etiologies Are Shared between Type 2 Diabetes and Cancers in the UAE Population. *Cancers* **2022**, *14*, 3313. <https://doi.org/10.3390/cancers14143313>

Academic Editors: Margaritis Avgeris, Andreas Scorilas and Mark Gorrell

Received: 16 March 2022

Accepted: 6 May 2022

Published: 7 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

One of the major revelations of the Human Genome Project was that a mere 1.5% of our genome encodes proteins, while the remaining 98.5% is non-coding [1]. The HGP was succeeded by major post-genomic consortia, including ENCODE (Encyclopedia of

DNA Elements) [2], which built an official GENCODE gene catalog [3], and FANTOM (Functional ANnoTation Of the Mammalian genome), which enabled the community to expand the human gene catalog with the addition of tens of thousands of validated and annotated non-coding RNA (ncRNA) genes [2,4]. The current GENCODE catalog contains 60,000 human genes, of which ~20,000 encode proteins, whereas most of the other 40,000 are comprised of several types of non-protein-coding RNA (ncRNA) genes. Long non-coding RNA (lncRNA) genes, generally defined as giving rise to RNA transcripts of more than 200 nucleotides without apparent protein-coding potential, are hence a prevalent class of human genes responsible for the most abundant and frequent type of transcripts in humans [4,5]. They have been increasingly recognized in recent years as integrally taking part in a wide range of biological and physiological processes, including transcriptional and post-transcriptional regulation of gene expression, protein translation and stability, cellular differentiation, cell lineage choice, organogenesis, and other key facets of normal development as well as disease [6]. A massive and growing amount of evidence from genomic epidemiology and population genetics has definitively indicated the association of lncRNAs with many common human diseases (e.g., cancer, diabetes, cardiovascular disease, etc.), but the underlying molecular mechanisms of this association are still poorly known [7,8]. For some lncRNAs, the connection with human diseases was made by differential gene expression analyses or functional studies as well as model organisms. A universe of other lncRNAs has been discovered from genome-wide association studies (GWAS), an approach used to statistically associate specific genetic variants with human diseases [9,10]. GWAS typically focus on associations between single-nucleotide polymorphisms (SNPs) and common diseases, but can equally be applied to any other genetic variants and any other organisms [9]. Single-nucleotide polymorphisms (SNPs) are single base-pair differences between the genomes of different individuals that occur every 500–1000 throughout the 3.3 billion bases long genome [11]. While many SNPs appear not to be associated with specific phenotypes, other SNPs' alleles can be clearly characterized as a disease (or risk) allele and a healthy (non-risk) allele upon comparative statistical evaluation of the incidence of the disease or trait in question in individuals homozygous for either allele and in heterozygotes. Classically, it was thought that disease-associated alleles of SNPs did not directly cause disease, and that they were simply genetically-linked to (in linkage disequilibrium, LD, with) the true causative variant, typically at or controlling a protein-coding gene in the vicinity that was causing the disease. However, now that we know that 82% of the genome is functional [2], that dogma is being reassessed: these simple variants in the genome may directly functionally contribute to the risk of disease, rather than merely be in LD with a coding functional variant elsewhere. The emerging post-ENCODE model thus stipulates that disease-associated SNPs may be functional and contributing to the pathogenesis of the disease. Indeed, contrary to early expectations, the vast majority (~93%) of disease- and trait-associated SNPs emerging from GWAS lie within a non-coding sequence, which includes intergenic and intronic regions, promoter regions, small ncRNA as well as lncRNA genes, antisense transcriptional units, and enhancer or insulator regions, and therefore they are likely to influence gene regulation [12–14]. Accordingly, in parallel with increased insights into the non-coding genome, emerging studies are characterizing SNPs that are located within non-coding RNA regions and are associated with various complex diseases, including diabetes and cancer [14]. The prevalence of diabetes, especially type 2 diabetes (T2D), and cancer has increased significantly in recent years, with a huge impact on health worldwide [15]. Epidemiologic evidence indicates that T2D and cancer often coexist in the same patients, and many risk factors, such as obesity, a sedentary lifestyle, smoking, and ageing, are common for both diseases [16]. Moreover, it is commonly understood that, while both genetic and environmental (including nutritional) factors contribute to T2D, the additive effects of common disease-associated genetic variants—which differ in different populations and parts of the world—are central to the etiology of this disorder and its relationship to cancer; however, the functional basis of the biological link between these diseases still needs to be better understood and emphasized.

Interestingly, the United Arab Emirates (UAE) population has one of the world's highest prevalence rates of diabetes (<https://diabetesatlas.org/>, accessed on 5 January 2022), and the burden of cancer is ranked as the second leading cause of non-communicable diseases (NCD)-related mortality in the country [17]. With the aim to highlight the importance of non-coding variants and lncRNA genes as a causative factor in both T2D and cancer, in this work we survey the major published UAE T2D GWAS to search for, and identify, putative new lncRNA-associated genetic etiologies that are shared between T2D and cancer in the UAE population.

2. Methods

2.1. Information Sources and Search Strategy

For this analysis, the input datasets were obtained from public GWAS of type 2 diabetes (T2D). The systematic literature search in PubMed was performed using the terms, "Type 2 Diabetes" and "GWAS" and/or "Genetic Variations" and "UAE" and/or "United Arab Emirates." The two most recent datasets [18,19] were included in the analysis. Manual annotation of all published significant T2D-associated genetic variants outside of protein-coding genes was conducted by using the public web-based bioinformatic tool, UCSC Genome Browser [20] (<https://genome.ucsc.edu/>, accessed on 10 March 2022). The location of the genetic variant (for example, in an exon or intron of a lncRNA gene), the epigenetic and expression profile of the region containing the variant, in thousands of human samples and tissue types surveyed across the datasets (Epigenome Roadmap, GTEX, FANTOM, mRNA/EST, others) in the Browser, the evolutionary conservation of the region, all ENCODE Consortium experimental data from all ENCODE data tracks in the Browser spanning the region and judged in our manual annotation as pertinent to its function, and all other data necessary to determine the relevance and potential function of the SNP were all analyzed in the UCSC Genome Browser. The SSTAR functionality of the FANTOM5 website https://fantom.gsc.riken.jp/5/sstar/Main_Page, accessed on 10 March 2022 [21] was also used to identify the top 10 tissue types, primary cell cultures, and/or cell lines where the gene/s associated with SNPs have the highest expression. For genes with multiple promoters in SSTAR, we have aggregated data across the promoters.

2.2. Eligibility Criteria

The tables referenced in our study were obtained from, and correspond to the main and supplementary tables of, the two major selected GWAS datasets [18,19]. The typical data set input table format, per line, included a reference SNP number ("rs" followed by a number) alongside an associated gene, a "mapped gene", or a "reported gene." An rs number is a universal SNP ID, that allows for searching of the SNP in any database. We manually analyzed all SNPs that were located in lncRNA genes (as evident from the gene name) or non-coding regions, and also all SNPs that were genomically complex. The latter was defined as one or more of the following: the presence of discrepancy between the mapped and reported gene names, or more than one gene (protein-coding and/or non) associated to the SNPs in the input data, or an alphanumeric gene name that was not corresponding to a protein family name/function, for example, gene names such as those that began with "AC" followed by numbers, or those displaying a mixture of letters and numbers without a gene-family root, and those that began with "Linc" (long intergenic non-coding RNA) and "Loc" (Locus, of presumably unknown function), which usually refer to lncRNA genes, and genes of unknown function many of which may be non-coding, respectively, were considered as indicators of SNPs potentially located in lncRNAs. All the SNPs in the dataset that met one or more of these criteria were then highlighted and annotated.

2.3. Data Collection and SNP Annotation

A total of 1284 SNPs were extracted from the GWAS input datasets. This number was then narrowed to 1132 by removing all the redundant SNPs. Next, based on the approach described in “Eligibility criteria” above, 370 SNPs were highlighted as potentially located in lncRNA genes (flow chart). Finally, the most promising SNPs (64 total) were thoroughly analyzed and annotated. The UCSC Genome Browser [20,22] was used to analyze and manually annotate every GWAS SNP. Upon selecting the GRCh37/hg19 human genome assembly, the UCSC Genome Browser was configured incorporating all pertinent browser tracks. The options “pack” and “show” for the selected tracks were chosen. From the default tracks that are displayed under the “Genes and Gene Predictions” Category, the “UCSC genes,” “NCBI Ref Seq” and “GENCODE” were chosen. For the “Phenotype and Literature” category, the “GWAS catalog” and “SNPedia” tracks were chosen. Further, for the “mRNA and EST” category, “Human ESTs” and “Human mRNAs” were selected whereas for the “Expression” category, the “Gtex GeneV8” was set to “full”. We used the ENCODE Regulation track of the UCSC Genome Browser to examine the H3K4Me1 and H3K27Ac (enhancer) signatures, and we viewed the DNase I cluster track, to determine whether the SNPs were located in areas with chromatin signatures consistent with enhancers in ENCODE Tier 1 cell types and in open-chromatin regions based on DNaseI-seq of 125 ENCODE cell types, respectively. All other settings that are not mentioned were set to “hide”. To find the SNPs in the genome, the SNP ID was copied into the search box of the UCSC Genome Browser. Then all the SNPs located in non-coding regions were annotated as well as assessed for relevance to cancer or T2D. The relationship to cancer was annotated both in the case of a direct correlation to a specific cancer, and in non-cancer-specific contexts but still cancer-related, including any mention in the associated literature of phenomena such as disabling of a tumor suppressor gene, apoptotic factors, and mitotic activity regulation. Similarly, we qualitatively assessed the relationship to diabetes. Lastly, when the highlighted SNP was surrounded by a large number of other SNPs that are also significantly associated with the same disease or trait or with other diseases or conditions in GWAS data in the NHGRI/EBI track of the Browser (a phenomenon we term a “SNP cloud”), as shown in Figure 1, the publications and gene names linked to all those nearby SNPs were analyzed for verifying any potential relevance to both T2D and cancer.

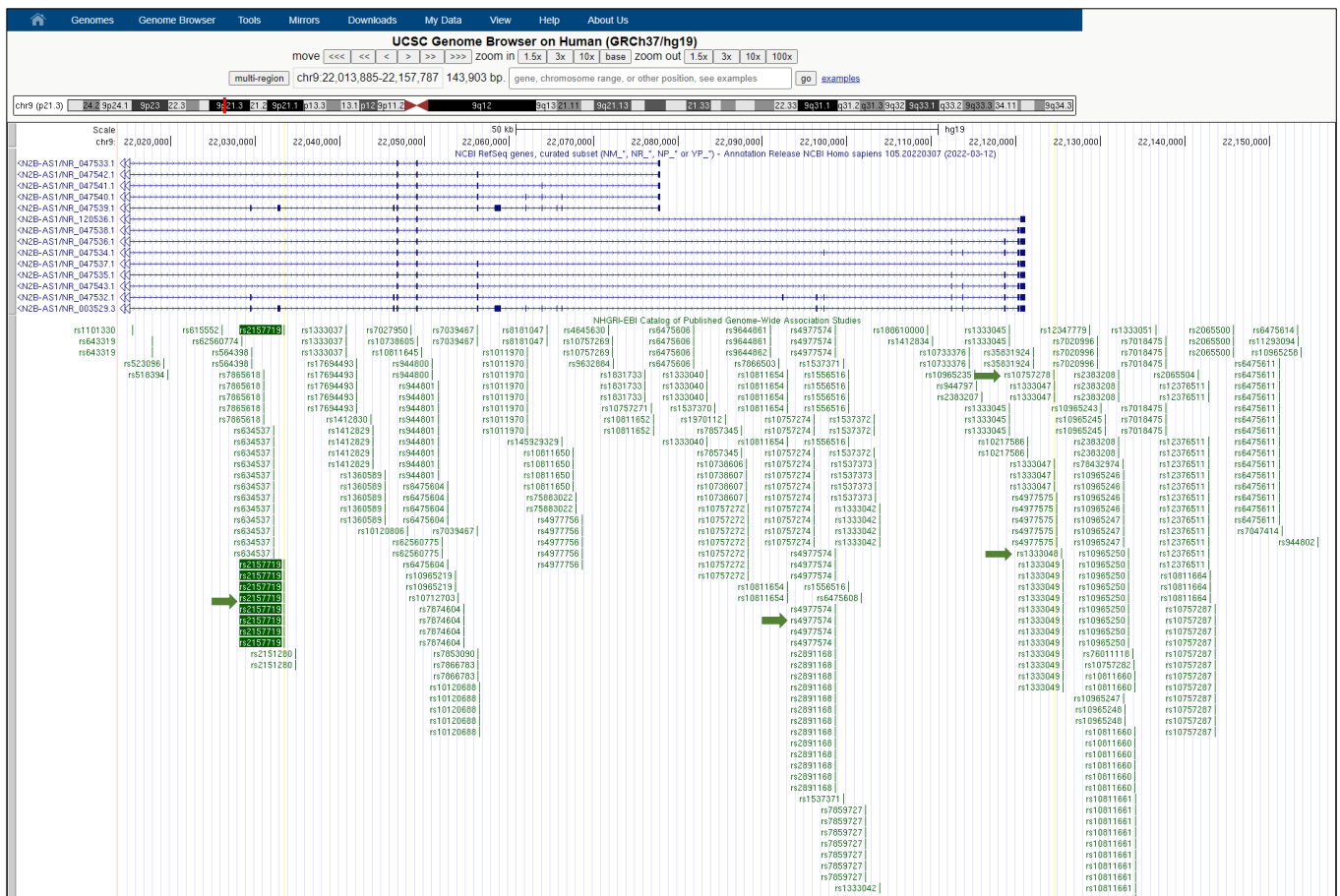


Figure 1. UCSC Genome Browser view of the human CDKN2B-AS1 gene. CDKN2B-AS1 (CDKN2B antisense RNA 1) is a lncRNA gene, also known as ANRIL. The SNP rs2157719 is intronic to CDKN2B-AS1 and it is surrounded by a large number of other significantly disease-associated SNPs in close proximity on the NHGRI-EBI GWAS Catalog track of the UCSC Genome Browser (a “SNP cloud”). Green arrows indicate the SNP rs2157719 associated with T2D and cancer, and the SNPs rs1333048, rs4977574, and rs10757278 associated with cancer.

3. Results

3.1. Customized Annotation Pipeline SNPs Selection

In this study an in-depth re-annotation of GWAS studies [18,19], in order to identify all statistically significant genetic variants residing in non-coding regions of the genome or in lncRNA genes, and that are of joint relevance to T2D and cancer in the UAE population was performed. The input dataset was obtained from public GWAS of T2D in the UAE population. From the few studies performed in this region and field, the two most recently published datasets in the UAE population [18,19], as these were applicable to our project, were selected. A total of 1284 SNPs associated with T2D and obesity were extracted from the main and supplementary tables of the GWAS input datasets [18,19]. The total number was then reduced to 1132 SNPs because we identified and excluded 152 redundant SNPs within and between the input datasets. This number was further narrowed to 370 by removing all the SNPs (762) associated with protein-coding genes. Finally, 67 SNPs of the 370 marked as potentially located in lncRNA genes or in non-coding regions (Figure 2), were selected for further analysis. These SNPs were confirmed (except for those otherwise mentioned below) for residing in non-coding regions. It is widely known that genetic variants are prone to mis-annotation and there is a substantial bias toward protein-coding genes in SNP annotations and in the related literature [23,24]; therefore, an unbiased reannotation was essential, and we used the UCSC Genome Browser to implement such a

reannotation. In this regard, out of the 64 selected SNPs, 13 were classified as “corrections” because of discrepancies between the previous annotation and our interpretation of the UCSC Browser results. These discrepancies were mainly related to the name of the nearest SNP-associated gene and/or the actual location in the genome. Furthermore, five SNPs, previously annotated in non-coding regions of the genome, were instead confirmed as located in protein-coding regions; regardless, we proceeded to annotate these SNPs for any relevance to T2D and cancer. The remaining 59 SNPs were confirmed as residing in non-coding regions (Figure 2). Finally, six SNPs jointly associated with both cancer and T2D (rs1495741, rs1061810, rs2521501, rs8042680, rs7526425, rs2157719), implicating a genetic link between the two diseases, were recognized. Summary of the six SNPs is showed in Table 1 and in Figure 3 which illustrates the six SNPs, the chromosomal region where the SNPs are located and the nearest gene.

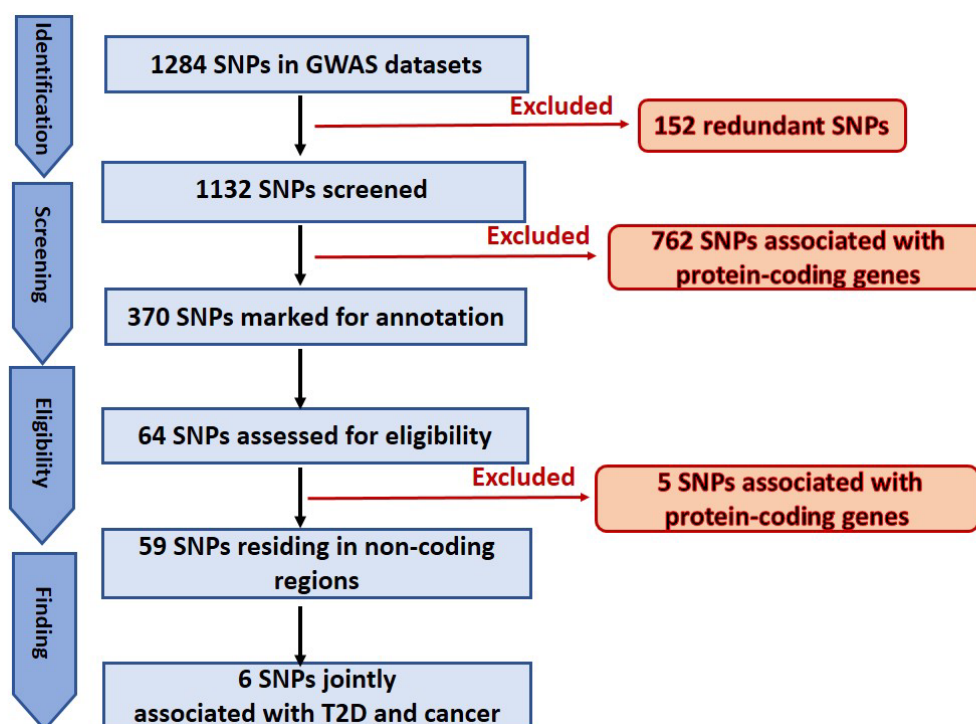


Figure 2. Schematic representation of SNPs selection.

Table 1. Summary of the six SNPs jointly associated with T2D and cancer.

SNP ID	Chromosomal Region	Nearest Gene	Nearby Genes	Type of Cancer	References
rs1495741	8p22	NAT2	NAT2, PSD3	HCC, LC, ESCC, AML, BC	[25–29]
rs1061810	11p11.2	HSD17B12	HSD17B12, AC087521.2, AC087521.4	BC, OC, MM	[30–32]
rs2521501	15q26.1	FES	FURIN, FES	APL, SARC	[33,34]
rs8042680	15q26.1	PRC1, PRC1-AS1	PRC1, PRC1-AS1	HCC	[35]
rs7526425	1q32.3	RD3	AC105275.1, SLC30A1, RD3	APL, NB	[36]
rs2157719	9p21.3	CDKN2B	CDKN2A, CDKN2B	OC, PC, MM, HNSCC	[37–40]

HCC: HepatoCellular Carcinoma; LC: Lung Cancer; ESCC: Esophageal Squamous-Cell Carcinoma; AML: Acute Myeloid Leukemia; BC: Breast Cancer; OC: Ovarian Cancer; MM: Malignant Melanoma; APL: Acute Promyelocytic Leukemia; SARC: Sarcoma; NB: Neuroblastoma; PC: Pancreatic Cancer; HNSCC: Head and Neck Squamous Cell Carcinoma.

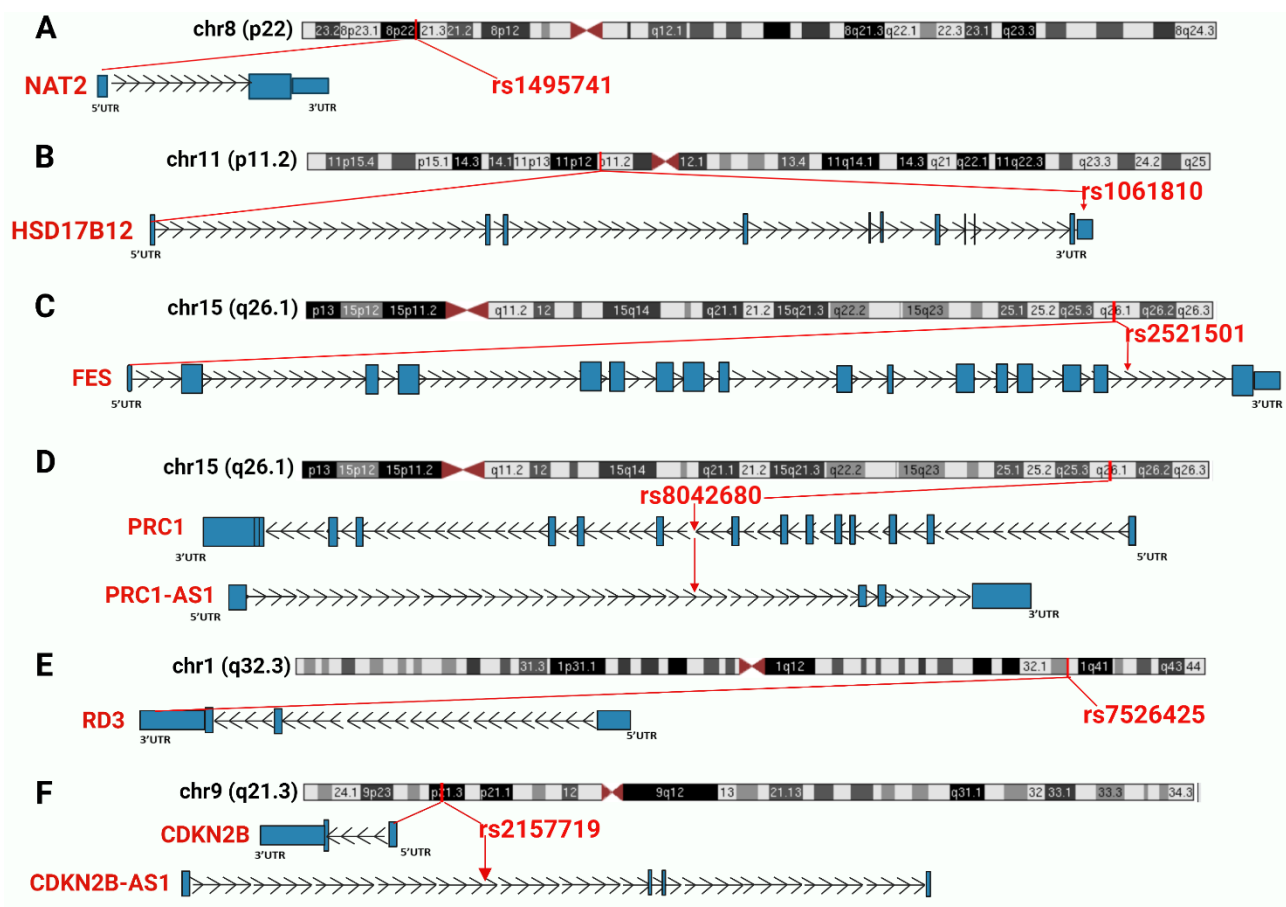


Figure 3. Summary of the chromosomal regions and the nearest genes of the six SNPs jointly associated with T2D and cancer. SNPs and nearest genes are marked in red. Introns are represented as lines with arrows indicating the direction of transcription, while coding exons are represented by blocks. (A) NAT2 gene and rs1495741; (B) HSD17B12 gene and rs1061810; (C): FES gene and rs2521501; (D) PRC1 gene, PRC1-AS1 antisense lncRNA and rs8042680; (E) RD3 gene and rs7526425; (F): CDKN2B gene, CDKN2B-AS1 antisense lncRNA and rs2157719.

3.2. Characteristics of SNPs Associated with T2D and Cancer

3.2.1. rs1495741

This SNP was previously reported as associated with the NAT2 (N-acetyltransferase 2) and PSD3 (Pleckstrin and Sec7 Domain Containing 3) genes. According to the NCBI Ref Seq track in the UCSC Genome Browser, NAT2 encodes an enzyme that functions to both activate and deactivate arylamine and hydrazine drugs and carcinogens [41,42] (Table 1 and Supplementary Figure S1). Polymorphisms in this gene are also associated with higher incidences of cancer (e.g., lung cancer, esophageal squamous cell carcinoma, acute myeloid leukemia, and breast cancer) and drug toxicity [26–29]. Moreover, a recent study combining a GWAS meta-analysis of 2764 individuals with direct, reference measures of insulin sensitivity with functional validation both in vitro and in vivo, identified NAT2 as a novel insulin sensitivity locus [43]. A second arylamine N-acetyltransferase gene (NAT1) is located near NAT2 (Supplementary Figure S1). PSD3 is a protein-coding gene associated with hepatocellular carcinoma (HCC), one of the most common types of primary liver cancer that often occurs in people with chronic liver diseases [25]. Risk factors generally include those which cause chronic liver disease, such as viral hepatitis B and C [44] but also metabolic disorders such as T2D and obesity [45]. The presence of two associated genes suggested that this SNP could reside in a genomically complex locus [46]; complex loci often, though not always, contain lncRNA genes. In this case, our analysis confirmed rs1495741 as an intergenic variant located between PSD3 and NAT2, at approximately

14 Kb from the 3' end of the NAT2, and approximately 100 Kb from the 3' end of PSD3. NAT2 is thus the nearest gene to the SNP (Table 1 and Figure 3A). The genomic region within boundaries of NAT2 and PSD3 genes is well-conserved across primates, with the highest level of conservation in the chimpanzee, as expected. Partial conservation of this region is observed in pig, mouse and rat; however, the genomic region around the SNP (100 bp) is not conserved in gorilla, mouse and rat, whereas in the pig it is only partially conserved. We located NAT2 in the FANTOM5 SSTAR database and reviewed both of the annotated transcription start sites (TSSs, promoters). The highest expression was in liver and hepatocytes from different donors, small intestine, colon, and fetal duodenum. The GTEx RNA-seq track of the UCSC Browser, which shows median gene expression levels in 52 tissues and 2 cell lines, indicates that NAT2 has the highest median expression in liver, whereas PSD3 has the highest median expression in the brain. For rs1495741, our review of the ENCODE Regulation track of the UCSC Genome Browser showed a lack of DNase I hypersensitive site signatures overlapping the SNP. The H3K4Me1 and H3K27Ac ChIP-seq signals were also minimal and arose from different cell types, suggesting a lack of enhancer signatures overlapping this SNP. No evidence of any transcription factor binding sites overlapping this SNP was observed in ENCODE ChIP-Seq data for 161 transcription factors. The only transcription factor binding sites near this SNP was GATA3. Regarding SNP-disease association as evidenced by the NHGRI-EBI GWAS Catalog track of the UCSC Genome Browser, rs1495741 is reported in 14 published papers where it is mainly associated with triglycerides, cholesterol levels and metabolic disorders. Moreover, two GWAS studies associate this variant with bladder cancer risk. Therefore, summarily, while this is not a non-coding RNA variant, it is a non-coding (intergenic) variant at a complex locus clearly associated with diabetes in the UAE population as well as with cancer.

3.2.2. rs1061810

The variant, rs1061810, was previously reported associated with two GENCODE Transcript annotations, ENST00000530450.1_3 (AC087521.2), and ENST00000637427.1_3 (AC087521.4) described as lncRNA and the protein-coding HSD17B12 (hydroxysteroid 17-beta dehydrogenase 12). Our analysis found this SNP exonic to the protein-coding gene HSD17B12, specifically it falls in the 3' UTR (Figure 3B); whereas, relative to the two non-coding genes at this locus, rs1061810 is intronic to AC087521.2, and exonic to AC087521.4, which is antisense to HSD17B12 (Table 1, Figures 3B and S2). Based on the NCBI Ref Seq track in the UCSC Genome Browser, HSD17B12 encodes for the enzyme 17 beta-hydroxysteroid dehydrogenase (17beta-HSD) that converts estrone into estradiol (E2) in ovarian tissue, but it is also involved in fatty acid elongation. The fatty Acyl-CoA biosynthesis and metabolism of steroid hormone pathways are both related to this gene and they are both linked with cancer. In fact, estrogen is well-known as a proliferative hormone and a driver of oncogenic and proliferative gene networks in estrogen receptor positive breast cancer, where it serves as a nuclear hormone; indeed, upon binding its receptor, it leads to the internalization of the receptor which in turn serves as a transcription factor, transitioning to the nucleus where it binds the promoters of the oncogenes that it activates (and of the tumor suppressors that it represses) [47,48]. Fatty acids are key players in cellular processes (cellular bioenergetics, membrane biosynthesis and intracellular signaling) involved in cancer development and progression [49]. In this regard, a study in human breast carcinoma suggested that involvement of HSD17B12 in the growth of carcinoma cells is not necessarily linked to the peripheral E2 biosynthesis but rather to the synthesis of very long chain fatty acids (VLCFAs), such as arachidonic acid, which contributes to breast carcinoma progression [30]. Furthermore, HSD17B12 is a marker of poor prognosis in ovarian carcinoma [31] and it is also associated with cutaneous melanoma [32] (Table 1 and Supplementary Figure S2). Rs1061810 is associated with T2D in two GWAS studies in the European population [50,51], highlighting the importance of this locus in T2D susceptibility in diverse populations and not only in the Middle East. The conservation of the locus where the variant rs1061810 resides is high across primates, and partial in the mouse, pig

and dog. The conservation level is low in rat. We located HSD17B12 in the FANTOM5 SSTAR database and reviewed all five TSSs (FANTOM5-annotated clustered promoters). The top ten cell lines where HSD17B12 is mostly expressed are: mast cells, the gall bladder carcinoma cells, smooth muscle cells from the aorta, mesenchymal stem cells, macrophages, adipocytes from two different donors, melanocytes, fibroblasts from aortic adventitial and fibroblasts from skin. For rs1061810, our review of the ENCODE Regulation track of the UCSC Genome Browser showed a lack of DNase I hypersensitive site signatures overlapping the SNP. A strong signal was observed in the RNA-seq subtrack in all nine ENCODE Tier 1 cell lines. The signal was weak in the H3K4Me1 and H3K27Ac ENCODE histone modification ChIP-seq data for ENCODE Tier 1 cell types. No evidence of any transcription factor binding sites overlapping or near this SNP was observed in ENCODE ChIP-Seq data for 161 transcription factors.

3.2.3. rs2521501

Previously mapped to the *FURIN-FES* locus, rs2521501 is intronic to *FES* (Figure 3C), whereas *FURIN* is the second nearest gene (approximately 10 kb) to this SNP (Table 1 and Supplementary Figure S3). According to the NCBI Ref Seq track in the UCSC Genome Browser, the proto-oncogene *FES* (Feline Sarcoma) encodes the human cellular counterpart of a feline sarcoma retrovirus protein with transforming capabilities. The gene product has tyrosine-specific protein kinase activity, which is required for the maintenance of cellular transformation. Its chromosomal location is linked it to a specific translocation event identified in patients with acute promyelocytic leukemia [33]; sarcoma is another type of cancer associated with *FES* [34]. The *FURIN* gene encodes for a calcium-dependent serine endoprotease which is expressed in many tissues, and it is involved in various physiological and pathophysiological processes ranging from embryonic development to carcinogenesis [52]. This endoprotease has rocketed to prominence in the past two years as a consequence of its role as a SARS-CoV-2/COVID-19 co-factor (which cleaves the S-protein) [53]. Several reports suggested that *FURIN* inhibition can suppress the tumorigenic properties of various cancer cell types, while other studies reported instead that *FURIN* inhibition may lead to a more aggressive phenotype of cancer cells [54]; however, despite these controversies, it is well established that *FURIN* plays a key role in cancer [54]. The intronic region of the *FES* gene where the SNP is located (approximately 1400 bp) is well conserved only in some primates, such as the chimpanzee, gorilla, rhesus monkey, crab-eating macaque, and green monkey. In rat, mouse and pig the conservation level is very low with a partial conservation only in the region around the SNP (70 bp). We located *FES* in the FANTOM5 SSTAR database and reviewed all four TSSs (FANTOM5-annotated clustered promoters). The highest expression was in different cell lines of CD14+ monocytes (from different donors), CD14+CD16– monocytes (different cell lines from different donors), biphenotypic B myelomonocytic leukemia cells, acute myeloid leukemia (FAB M5) cells, and eosinophils (different cell lines from different donors). The GTEx track indicates *FURIN* has the highest median expression in the liver, followed by the pancreas (organs directly relevant to T2D and obesity pathogenesis), whereas *FES* has the highest median expression in the spleen followed by the lungs. In several GWAS studies, as evidenced by the NHGRI-EBI GWAS Catalog track of the UCSC Genome Browser, rs2521501 is associated with diastolic and systolic blood pressure as well as with the interaction of blood pressure with alcohol consumption and/or cigarette smoking. The ENCODE Regulation track of the UCSC Genome Browser shows a lack of DNase I hypersensitive site signatures overlapping the SNP. A low level of expression was observed in the RNA-seq data from the ENCODE Tier 1 nine cell lines, consistent with the intronic localization of the SNP, as the signal was greater in the gene's exons as expected. The signal was also weak in the H3K4Me1 and H3K27Ac ENCODE histone modification ChIP-seq data for ENCODE Tier 1 cell types. For rs2521501, there was no evidence of any transcription factor binding sites overlapping, or near this SNP in ENCODE ChIP-Seq data for 161 transcription factors.

3.2.4. rs8042680

PRC1 and PRC1-AS1 were previously reported as the closest genes to rs8042680; we found that this SNP is intronic to both genes (Table 1, Figures 3D and S4). PRC1 (protein regulator of cytokinesis 1) encodes for a protein involved in cytokinesis and microtubules organization. PRC1 is overexpressed in human hepatocellular carcinoma cells and it is associated with the increased chemoresistance of these cells [55]. PRC1-AS1 is an antisense lncRNA and, similarly to PRC1, is associated with hepatocellular carcinoma [35]. The intronic region where the SNP is located is conserved only in primates with the exception of marmoset and squirrel monkey, while in rat, mouse and pig it is not conserved. We were unable to find PRC1-AS1 antisense lncRNA in the FANTOM5 SSTAR database, and therefore visually reviewed that lncRNA gene in FANTOM CAGE data via the graphical browser (testis, tongue, mesothelioma, chondrocyte, and mesenchymal stem cells were the top expressors). We located PRC1 in the SSTAR database and reviewed all four clustered TSSs that have precomputed FANTOM5 expression data. The highest expression was in reticulocytes, hepatoblastoma, osteosarcoma, bone marrow, and CD14+ monocytes (the latter were in the top 15 expressors for the antisense too). With regards to tissue expression, PRC1 has the highest median expression in fibroblasts and EBV-transformed lymphoblastoid cell lines, while PRC1-AS1 has the highest median expression in testis. For rs8042680, our review of the ENCODE Regulation track of the UCSC Genome Browser shows a lack of DNase I hypersensitive site signatures overlapping the SNP. The H3K4Me1 and H3K27Ac ChIP-seq signals were also minimal and arose from different cell types, suggesting a lack of enhancer signatures overlapping this SNP. No evidence of any transcription factor binding sites overlapping or near this SNP was observed in ENCODE ChIP-Seq data for 161 transcription factors. As evidenced by the NHGRI-EBI GWAS Catalog track of the UCSC Genome Browser, two papers reported the association of rs8042680 with T2D. Summarily, hence, this is another SNP that is associated both with T2D in the UAE population and with cancer.

3.2.5. rs7526425

From the input datasets' published annotations, rs7526425 was reported as mapped near three genes: AC105275.1, SLC30A1, and RD3. The SNP is not found in the "NHGRI-EBI Catalogue of Published GWAS" track of the UCSC Genome Browser (Table 1 and Supplementary Figure S5). RD3 is the closest gene to the SNP (Figure 3E); this gene encodes a retinal protein that is associated with promyelocytic leukemia-gene product (PML) bodies in the nucleus. Moreover, RD3 plays a regulatory role in neuroblastoma progression and its loss is associated with aggressive neuroblastoma and poor clinical outcomes [36]. The gene AC105275.1 encodes a lncRNA and is antisense to RD3. The gene SLC30A1 encodes a zinc transporter protein, which is involved in maintaining cellular zinc homeostasis in mammalian cells. This gene is therefore relevant to diabetes, since zinc is an essential co-factor for insulin metabolism in the pancreatic β -cell [56]. Moving on to the tissue expression of the genes, RD3 has the highest median expression in the pituitary and SLC30A1 has the highest median expression in the liver; both organs are directly relevant to diabetes pathogenesis and etiology. The genomic region around the SNP (100 bp) is highly conserved across most primates, as expected and is well-conserved in rat, mouse, rabbit, pig and dog. We found RD3 in the FANTOM5 SSTAR database and reviewed all three TSSs (FANTOM5-annotated clustered promoters). The highest expression was in the pineal gland, retinoblastoma cell line Y79, fetal eye, lung carcinoma cells, and medulloblastoma cells. Finally, the UCSC Browser shows a large SNP cloud surrounding the SNP. This is hence yet another SNP with a dual diabetes and cancer association, and is a UAE T2D susceptibility risk variant per the input datasets. A compelling case can be made for a non-coding regulatory region encompassing this SNP (Supplementary Figure S5) as being even more relevant than RD3 as an explanation for the SNP's functional significance: the SNP falls directly into a prominent ENCODE DNase I hypersensitive region that also corresponds to a transcription factor binding (ChIP-seq) consensus signature across 13

of the 125 ENCODE cell types directly overlapping the SNP. In addition, there was a strong H3K4Me1 signal and weak H3K27Ac signal in the ENCODE histone modification ChIP-seq data for ENCODE Tier 1 cell types, also directly overlapping the SNP. The two enhancer histone modifications occurred only in GM12878 cells, which were also one of the 13 DNase I hypersensitive cell types in the region. Summarily, there is evidence for open chromatin and putative enhancer modifications directly encompassing this SNP in GM12878. For rs7526425, there are three ENCODE TFBS ChIP-seq hits directly overlapping and encompassing the SNP: RELA, POU2F2, and MEF2A. While all three contain consensus recognition sites for the respective TFs, none of the consensus sequences overlaps the SNP (whereas the rest of the ChIP-seq-detected binding site contains the SNP). Therefore, the SNP is in TFBSs, but is unlikely to affect binding affinity, due to its location outside of the consensus. Interestingly, all three binding signals include GM12878 cells, where the DNase I and enhancer signatures also occurred. The regulatory significance of the region containing this variant in GM12878 is also supported by the multiple Hi-C interactions found ~130 bp away with multiple nearby genomic regions in the UCSC Genome Browser's Rao et al. 2014 Hi-C track.

3.2.6. rs2157719

As previously reported, rs2157719 is within the CDKN2B-CDKN2A gene cluster. The closest protein-coding gene to this SNP is CDKN2B (cyclin dependent kinase inhibitor 2B) (Figure 3F) which is adjacent to the tumor suppressor gene CDKN2A (Cyclin-Dependent Kinase 4 Inhibitor A) in a region that is frequently mutated and deleted in a wide variety of tumors (Table 1 and Figure 2). Both genes are associated with several cancer types, including ovarian cancer, pancreatic cancer, melanoma and, head and neck squamous cell carcinoma [37–40]. Rs2157719 is also intronic to the non-coding RNA, CDKN2B antisense RNA 1 (CDKN2B-AS1), also known as ANRIL (Figure 3E), which is linked to the progression of diabetic nephropathy, one of the most common T2D complications [57]; however, the prominence of CDKN2B-AS1 in the literature to date is mostly due to its role in cancer, where it can enhance cell proliferation, cell cycle progression, and inhibit apoptosis and senescence. Moreover, CDKN2B-AS1 is overexpressed in many cancer types and is a well-recognized prognostic and diagnostic biomarker in cancer [58–60]. A close examination of the “SNP cloud” surrounding rs2157719 (Figure 1) reveals that several SNPs in this region are significantly associated with other cancer risks; while consistent with the linkage disequilibrium expectation that a genetic region containing multiple SNPs is associated with the risk, this is surprising because of the diversity of these cancers. For instance, rs1333048 (Figure 1) is a biomarker of breast cancer susceptibility, and it is associated with the risk of toxicity after chemotherapeutic drug (cisplatin) treatment in lung cancer patients [61,62]. Moreover, rs4977574 is associated with kidney cancer development in the Ukrainian population [63] while rs10757278 is associated with ANRIL expression levels and cisplatin resistance in cancer [58,61] (Figure 1). The GTEx track of the UCSC Genome Browser shows that CDKN2B-AS1 has the highest median expression in the colon and small intestine. From GWAS studies reflected in the NHGRI-EBI track of the UCSC Browser, rs2157719 itself is associated with diverse cancers; the SNP has been referenced in several papers as related to glaucoma and glioma [64–67]. The intronic region where the SNP is located is highly conserved across most primates, as expected. Moreover, the region is partially conserved in rat, mouse, rabbit, pig and dog. We located both CDKN2B and CDKN2B-AS antisense lncRNA in the FANTOM5 SSTAR database. For CDKN2B we reviewed all two TSSs (FANTOM5-annotated clustered promoters) and found that the highest expression for this gene was in lens epithelial cells, preadipocytes, mesenchymal cells, cardiac fibroblasts, adipocytes, differentiated osteoblast, and melanoma cells. Meanwhile, for the CDKN2B-AS antisense lncRNA, we reviewed all three promoters and found the highest expression in carcinosarcoma cells, gastrointestinal carcinoma cells, osteosarcoma cells, lens epithelial cells and gall bladder carcinoma cells. For rs2157719, there was a DNase I hypersensitive region in 8 of the 125 ENCODE cell types directly overlapping the

SNP. In addition, there was a strong signal in the transcription track showing the transcription levels assayed by RNA-seq on nine cell lines, with the highest signal for the HeLa-S3 cells. There was a modest H3K4Me1 signal (highest in the Epidermal Keratinocyte Cells NHEK), and weak H3K27Ac signal in the ENCODE histone modification ChIP-seq data for ENCODE Tier 1 cell types also directly overlapping the SNP. Summarily, there is evidence for open chromatin and putative enhancer modifications directly encompassing this SNP in HeLa-S3 and NHEK cells. No evidence of any transcription factor binding sites overlapping or near this SNP was observed in ENCODE ChIP-Seq data for 161 transcription factors.

4. Discussion

Traditionally, proteins and small molecules, and the molecular pathways and regulatory networks containing them, have given rise to the majority of drug targets. For this reason, GWAS projects, despite being an inherently unbiased approach, have been historically biased toward the identification of protein-coding regions and genes carrying disease risks. Meanwhile, data on non-coding SNPs has either not been pursued or has been consistently misinterpreted in favor of distant protein-coding variants, while ignoring the nearby or overlapping lncRNA genes or non-coding regulatory regions that can be more likely causal determinants of the disease phenotype. Simply because 98.5% of the human genome is non-coding, and also because two-thirds of the approximately 60,000 human genes are lncRNA genes, most significant disease-associated variants are in non-protein-coding regions of the genome. Therefore, most GWAS studies that have limited their analysis to previously known, hence primarily protein-coding risk variants, despite having genotyped patients and controls on a whole-genome level, contain frequent and abundant mis-annotations, and encompass numerous lncRNA genes and non-coding regions in raw data, which are often subsequently not interpreted [68,69]. An unbiased reannotation of disease-associated SNPs is therefore essential and moreover, re-annotation from previous research findings can potentially result in a different knowledge outcome [70]. For example, in our work, we have uncovered previously-unknown dual diabetes–cancer associations of specific non-coding variants based on GWAS signals specifically from the UAE population. In this study, we performed an integrated re-annotation of recent UAE population specific T2D datasets, we highlighted numerous SNPs in non-coding regions and several in lncRNA genes, and we computationally defined the relevance to cancer of a subset of those regions. Previous studies, mainly observational epidemiological studies, have established associations between T2D and cancer in terms of shared common risk factors such as, hyperglycemia, hyperinsulinemia, obesity, a lack of physical activity and diet, but the biological mechanisms supporting these associations have remained poorly known [71]. Similarly, genetics studies, such as those based on the identification of the same genetic variant and risk allele that independently predisposes to both type 2 diabetes and cancer, are still insufficient to explain the increased risk of cancer in people with T2D [71]. Here, we addressed the knowledge gap in functionally and genetically understanding the diabetes–cancer connection: we successfully discovered six significant disease-associated genetic variants that signify new biological links between T2D and cancer. All six SNPs reside in non-coding regions: specifically, two SNPs (rs1495741, and rs7526425) are intergenic, three SNPs (rs1061810, rs8042680, and rs2157719) reside in lncRNA genes, and one SNP (rs2521501) is intronic to the FES gene. All six SNPs show a clear and strong association with both T2D and cancer; this is demonstrated not only by the GWAS studies reported in the NHGRI-EBI GWAS Catalog track, but also by the direct involvement, in both diseases, of the nearby, or overlapping, SNPs-associated genes. For instance, the variant rs1495741 is directly correlated with bladder cancer [72], elevated triglyceride levels, T2D risk [73], and liver injury [74]. Additionally, both the proximal and the distal gene to the SNP, NAT2 and PSD3, respectively, are strongly correlated with both cancer and T2D [26,43,75–78]. Another SNP significantly associated with T2D and cancer is rs2157719; this variant is indeed directly correlated to glaucoma and glioma [64–67] and resides in a genomic region that is frequently mutated and deleted in a wide variety of tumors. Moreover, it is intronic

to the lncRNA CDKN2B-AS1, which is aberrantly expressed in various malignancies and it is also implicated in numerous non-malignant diseases, including diabetes and its complications [79]. Several SNPs are associated with genes that play a key role in T2D and, according to the GTEx RNA-seq track of the UCSC Genome Browser, have the highest median expression in the liver; this is the case of NAT2 (rs1495741), FURIN (rs2521501), and SLC30A1 (rs2157719). NAT2 is a cellular enzyme involved in the metabolism of a variety of different compounds, including carcinogens. Deficiency of this gene causes mitochondrial dysfunction with decreased cellular respiration and ATP generation, suggesting that NAT2 may mediate insulin resistance and mitochondrial dysfunction by binding key regulators of energy balance [80]. FURIN is a membrane-bound protease broadly involved in the maintenance of cellular homeostasis. A recent study showed that individuals with high plasma furin levels concentration have an elevated risk of diabetes mellitus and premature mortality [81]. Moreover, furin protein upregulation results in worse outcomes in diabetic patients with SARS-CoV-2 infection [82]. Finally, SLC30A1 plays a key role in maintaining the cellular zinc levels within a physiological range [83]. Zinc is a second messenger that controls many processes associated with insulin signaling [84,85]. Furthermore, all these genes are mostly expressed in the liver, an organ that plays a key role in the regulation of glucose metabolism [86]. T2D is an excellent setting for RNA therapeutics because much of the pathogenesis occurs in the liver, and RNA drugs injected into the bloodstream naturally go to the liver but not to any other organs [87]. RNA therapeutics target mRNA as well as non-coding RNAs, including lncRNAs, with small interfering RNAs (siRNA), antisense oligonucleotides (ASO), modified-backbone oligonucleotides (MBO), and other RNAi-based drugs [88]. RNA drugs are highly sequence-specific, as each drug has only one target, can distinguish the risk and the non-risk alleles and hence can be used in patients that have the causal disease-risk variant. Therefore, RNA-based drugs have fewer side effects and are very stable. For these reasons, RNA drugs are destined to change the way disease is treated in a more targeted and personalized manner [87,88]. For example, the milestones achieved with the development and approval of Inclisiran as a synthetic siRNA drug against an RNA target, specific to the liver, illustrate the potential for the development of similar drugs targeting diabetes and cancer, deployed against other, including lncRNA, targets. Many diabetes genes are indeed expressed in the liver, which is easily targeted by this approach, and the GWAS-empowered discovery of non-coding, liver-expressed diabetes and cancer candidates can generate a target list for future development of Inclisiran-class drugs.

Limitations of our study include how the GWAS data had been analyzed by the teams that provided the input datasets, and the incomplete availability of biological data types as well as experimental datasets from specific cell and tissue types in the UCSC Genome Browser. We focused on published GWAS datasets of SNPs deemed jointly relevant to T2D and cancer, analyzing these variants' relevance to the non-coding regions of the genome, specifically lncRNA; however, GWAS approaches are inherently prone to false positives and false negatives with regards to the association of a variant with a phenotype: there is a possibility of SNPs in the datasets that are unrelated to the diseases (false positive) or SNPs that are related that were not included at all (false negative). This could be due to ascertainment biases, weak association in a common disease/common variant context where a single SNP fails to reach significance, a population-specific lack of genotype-phenotype association that exists in other population, a smaller study population and, thus, sample sizes utilized by the authors of these published papers on GWAS, or other factors. Additionally, most GWAS studies, including those that performed high-throughput genotyping rather than de-novo sequencing of multiple individual genomes, are based on previously known variants, resulting in the exclusion of newer variants that may have a stronger association or may be located nearer to the actual functional feature in the genome. The impact of this limitation leads to the exclusion of relevant and possibly targetable SNPs (false negative) and the inclusion of SNPs that are unrelated (false positives) to T2D. The UCSC Genome Browser may not contain expression data for certain low-abundance transcripts that lack cDNA/EST and ENCODE RNA-seq representation. In addition, the UCSC Genome Browser lacks

representation of certain datasets that could be essential to our study and that have been published but are not reflected in the Browser's tracks. While our work should facilitate the development of targeted therapies, a limitation is that the identified genes would need to be the direct functional causes of the phenotype (pending functional validation experiments) in order to be targeted; the specific risk alleles of the SNPs should be targeted in a sequence-specific fashion if causative; and the expression of the target would have to be confined to tissues and organs where it directly underlines the disease phenotype, so as to avoid off target effects when sequence-based, such as RNAi, therapies are deployed. Despite possibly being limited to the specific populations of the UAE, this work provides genetics-based evidence of novel non-coding biological links between T2D and cancer, which feature specific putative regulatory regions and lncRNA genes. This work lays the foundation to further explore the shared genetic architecture between these two common diseases, as well as to expand the study to multiple populations [89]. Functional studies will be needed in order to understand the functional implications of these SNPs, and to provide new insights about the mechanisms leading to the progression of the two diseases.

5. Conclusions

In conclusion, we have gathered genetic-based evidential proof of a biological link between T2D and cancer. Our major findings include the successful discovery of six significant disease-associated genetic variants, identified in the UAE population, that have established a previously unknown biological link between T2D and cancer. The novelty of our study is in leveraging upon UAE-origin GWAS T2D data (given the relative lack of metabolic-disease GWAS in the Gulf and Middle East region) to identify variants of joint relevance in cancer and diabetes. These variants suggest specific functional links between those two common diseases, links that help to better explain their casually-known co-occurrence. These variants resided in lncRNA genes and at or near protein-coding genes. Genetic heterogeneity is central in common diseases (such as diabetes and cancers) that have a partly-genetic etiology; therefore, it is crucial to understand the genetic basis of a disease in a specific population, in addition to considering risk variants common to multiple populations. They provide an early foundation for a shortlist of possible drug targets, including those personalized toward variants occurring in the Gulf and Middle East/North Africa regions but not globally.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/cancers14143313/s1>, Figure S1: UCSC Genome Browser view of the SNP rs1495741; chr8 (p22):18,272,853-18,272,908; Figure S2: UCSC Genome Browser view of the SNP rs1061810; chr11 (p11.2):43,877,906-43,877,961; Figure S3: UCSC Genome Browser view of the SNP rs2521501; chr15 (q26.1): 91,437,360-91,437,415; Figure S4: UCSC Genome Browser view of the SNP rs8042680; chr15 (q26.1):91,521,309-91,521,364; Figure S5: UCSC Genome Browser view of the SNP rs7526425; chr1: (q32.3) 211,700,630-211,700,685.

Author Contributions: L.L., R.G. (Roberta Giordo) and G.A.C. contributed to the study concept and design. L.L., R.G. (Roberta Giordo), R.G. (Rida Gulsha) and S.K. contributed to data analysis and/or interpretation. R.G. (Roberta Giordo), R.G. and S.K. contributed to the original draft preparation. L.L., R.G. (Roberta Giordo) and G.A.C. contributed to the review and edit of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Mohammed Bin Rashid University Internal Research Grant to Leonard Lipovich; the Mohammed Bin Rashid University Postdoctoral Fellowship to Roberta Giordo; and by the Al Jalila Foundation. We also gratefully acknowledge support from the Mohammed Bin Rashid University for Article processing fees/page charges for this manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gregory, T.R. Synergy between sequence and size in large-scale genomics. *Nat. Rev. Genet.* **2005**, *6*, 699–708. [[CrossRef](#)] [[PubMed](#)]
2. Consortium, E.P. An integrated encyclopedia of DNA elements in the human genome. *Nature* **2012**, *489*, 57. [[CrossRef](#)] [[PubMed](#)]
3. Derrien, T.; Guigó, R.; Johnson, R. The long non-coding RNAs: A new (p) layer in the “dark matter”. *Front. Genet.* **2012**, *2*, 107. [[CrossRef](#)] [[PubMed](#)]
4. Abugessaisa, I.; Ramilowski, J.A.; Lizio, M.; Severin, J.; Hasegawa, A.; Harshbarger, J.; Kondo, A.; Noguchi, S.; Yip, C.W.; Ooi, J.L.C. FANTOM enters 20th year: Expansion of transcriptomic atlases and functional annotation of non-coding RNAs. *Nucleic Acids Res.* **2021**, *49*, D892–D898. [[CrossRef](#)]
5. Jain, S.; Thakkar, N.; Chhatai, J.; Pal Bhadra, M.; Bhadra, U. Long non-coding RNA: Functional agent for disease traits. *RNA Biol.* **2017**, *14*, 522–535. [[CrossRef](#)]
6. Schmitz, S.U.; Grote, P.; Herrmann, B.G. Mechanisms of long noncoding RNA function in development and disease. *Cell. Mol. Life Sci.* **2016**, *73*, 2491–2509. [[CrossRef](#)]
7. Esteller, M. Non-coding RNAs in human disease. *Nat. Rev. Genet.* **2011**, *12*, 861–874. [[CrossRef](#)]
8. DiStefano, J.K. The emerging role of long noncoding RNAs in human disease. *Dis. Gene Identif.* **2018**, *1706*, 91–110.
9. Manolio, T.A.; Collins, F.S.; Cox, N.J.; Goldstein, D.B.; Hindorff, L.A.; Hunter, D.J.; McCarthy, M.I.; Ramos, E.M.; Cardon, L.R.; Chakravarti, A. Finding the missing heritability of complex diseases. *Nature* **2009**, *461*, 747–753. [[CrossRef](#)]
10. Sparber, P.; Filatova, A.; Khantemirova, M.; Skoblov, M. The role of long non-coding RNAs in the pathogenesis of hereditary diseases. *BMC Med. Genom.* **2019**, *12*, 63–78. [[CrossRef](#)]
11. Consortium, G.P. A global reference for human genetic variation. *Nature* **2015**, *526*, 68. [[CrossRef](#)] [[PubMed](#)]
12. Maurano, M.T.; Humbert, R.; Rynes, E.; Thurman, R.E.; Haugen, E.; Wang, H.; Reynolds, A.P.; Sandstrom, R.; Qu, H.; Brody, J. Systematic localization of common disease-associated variation in regulatory DNA. *Science* **2012**, *337*, 1190–1195. [[CrossRef](#)] [[PubMed](#)]
13. Giral, H.; Landmesser, U.; Kratzer, A. Into the wild: GWAS exploration of non-coding RNAs. *Front. Cardiovasc. Med.* **2018**, *5*, 181. [[CrossRef](#)] [[PubMed](#)]
14. Edwards, S.L.; Beesley, J.; French, J.D.; Dunning, A.M. Beyond GWASs: Illuminating the dark road from association to function. *Am. J. Hum. Genet.* **2013**, *93*, 779–797. [[CrossRef](#)]
15. Wang, M.; Yang, Y.; Liao, Z. Diabetes and cancer: Epidemiological and biological links. *World J. Diabetes* **2020**, *11*, 227. [[CrossRef](#)]
16. Giovannucci, E.; Harlan, D.M.; Archer, M.C.; Bergenstal, R.M.; Gapstur, S.M.; Habel, L.A.; Pollak, M.; Regensteiner, J.G.; Yee, D. Diabetes and cancer: A consensus report. *Diabetes Care* **2010**, *33*, 1674–1685. [[CrossRef](#)]
17. Radwan, H.; Hasan, H.; Ballout, R.A.; Rizk, R. The epidemiology of cancer in the United Arab Emirates: A systematic review. *Medicine* **2018**, *97*, e13618. [[CrossRef](#)]
18. Hachim, M.Y.; Aljaibeji, H.; Hamoudi, R.A.; Hachim, I.Y.; Elemam, N.M.; Mohammed, A.K.; Salehi, A.; Taneera, J.; Sulaiman, N. An integrative phenotype–genotype approach using phenotypic characteristics from the UAE national diabetes study identifies HSD17B12 as a candidate gene for obesity and type 2 diabetes. *Genes* **2020**, *11*, 461. [[CrossRef](#)]
19. Osman, W.; Hassoun, A.; Jelinek, H.F.; Almahmeed, W.; Afandi, B.; Tay, G.K.; Alsafar, H. Genetics of type 2 diabetes and coronary artery disease and their associations with twelve cardiometabolic traits in the United Arab Emirates population. *Gene* **2020**, *750*, 144722. [[CrossRef](#)]
20. Lee, B.T.; Barber, G.P.; Benet-Pagès, A.; Casper, J.; Clawson, H.; Diekhans, M.; Fischer, C.; Gonzalez, J.N.; Hinrichs, A.S.; Lee, C.M. The UCSC Genome Browser database: 2022 update. *Nucleic Acids Res.* **2022**, *50*, D1115–D1122. [[CrossRef](#)]
21. Abugessaisa, I.; Shimoji, H.; Sahin, S.; Kondo, A.; Harshbarger, J.; Lizio, M.; Hayashizaki, Y.; Carninci, P.; Forrest, A.; Kasukawa, T. FANTOM5 transcriptome catalog of cellular states based on Semantic MediaWiki. *Database* **2016**, *2016*. [[CrossRef](#)] [[PubMed](#)]
22. Karolchik, D.; Baertsch, R.; Diekhans, M.; Furey, T.S.; Hinrichs, A.; Lu, Y.; Roskin, K.M.; Schwartz, M.; Sugnet, C.W.; Thomas, D.J. The UCSC genome browser database. *Nucleic Acids Res.* **2003**, *31*, 51–54. [[CrossRef](#)] [[PubMed](#)]
23. Platt, A.; Vilhjálmsón, B.J.; Nordborg, M. Conditions under which genome-wide association studies will be positively misleading. *Genetics* **2010**, *186*, 1045–1052. [[CrossRef](#)] [[PubMed](#)]
24. Manning, A.K.; Goustin, A.S.; Kleinbrink, E.L.; Thepsuwan, P.; Cai, J.; Ju, D.; Leong, A.; Udler, M.S.; Brown, J.B.; Goodarzi, M.O. A long non-coding RNA, LOC157273, is an effector transcript at the chromosome 8p23.1-PPP1R3B metabolic traits and type 2 diabetes risk locus. *Front. Genet.* **2020**, *11*, 615. [[CrossRef](#)] [[PubMed](#)]
25. Llovet, J.; Kelley, R.; Villanueva, A.; Singal, A.; Pikarsky, E.; Roayaie, S.; Lencioni, R.; Koike, K.; Zucman-Rossi, J.; Finn, R. Hepatocellular carcinoma nature reviews. *Dis. Primers* **2021**, *7*, 6. [[CrossRef](#)] [[PubMed](#)]
26. Liu, C.; Cui, W.; Cong, L.; Wang, L.; Ruan, X.; Jia, J.; Liu, Y.; Jia, X.; Zhang, X. Association between NAT2 polymorphisms and lung cancer susceptibility. *Medicine* **2015**, *94*, e1947. [[CrossRef](#)] [[PubMed](#)]
27. Matejčić, M.; Vogelsang, M.; Wang, Y.; Parker, I.M. NAT1 and NAT2 genetic polymorphisms and environmental exposure as risk factors for oesophageal squamous cell carcinoma: A case-control study. *BMC Cancer* **2015**, *15*, 150.
28. Zou, Y.; Dong, S.; Xu, S.; Gong, Q.; Chen, J. Genetic polymorphisms of NAT2 and risk of acute myeloid leukemia: A case-control study. *Medicine* **2017**, *96*, e7499. [[CrossRef](#)]

29. Kocabaş, N.A.; Şardaş, S.; Cholerton, S.; Daly, A.K.; Karakaya, A.E. N-acetyltransferase (NAT2) polymorphism and breast cancer susceptibility: A lack of association in a case-control study of Turkish population. *Int. J. Toxicol.* **2004**, *23*, 25–31. [[CrossRef](#)]
30. Nagasaki, S.; Suzuki, T.; Miki, Y.; Akahira, J.-i.; Kitada, K.; Ishida, T.; Handa, H.; Ohuchi, N.; Sasano, H. 17 β -Hydroxysteroid dehydrogenase type 12 in human breast carcinoma: A prognostic factor via potential regulation of fatty acid synthesis. *Cancer Res.* **2009**, *69*, 1392–1399. [[CrossRef](#)]
31. Szajnik, M.; Szczepanski, M.J.; Elishaev, E.; Visus, C.; Lenzner, D.; Zabel, M.; Glura, M.; DeLeo, A.B.; Whiteside, T.L. 17 β hydroxysteroid dehydrogenase type 12 (HSD17B12) is a marker of poor prognosis in ovarian carcinoma. *Gynecol. Oncol.* **2012**, *127*, 587–594. [[CrossRef](#)] [[PubMed](#)]
32. Dai, W.; Liu, H.; Xu, X.; Ge, J.; Luo, S.; Zhu, D.; Amos, C.I.; Fang, S.; Lee, J.E.; Li, X. Genetic variants in ELOVL2 and HSD17B12 predict melanoma-specific survival. *Int. J. Cancer* **2019**, *145*, 2619–2628. [[CrossRef](#)] [[PubMed](#)]
33. O’Leary, N.A.; Wright, M.W.; Brister, J.R.; Ciufu, S.; Haddad, D.; McVeigh, R.; Rajput, B.; Robbertse, B.; Smith-White, B.; Ako-Adjei, D. Reference sequence (RefSeq) database at NCBI: Current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **2016**, *44*, D733–D745. [[CrossRef](#)] [[PubMed](#)]
34. Zhao, Y.; Wang, Z.; Wang, Q.; Sun, L.; Li, M.; Ren, C.; Xue, H.; Li, Z.; Zhang, K.; Hao, D. Overexpression of FES might inhibit cell proliferation, migration, and invasion of osteosarcoma cells. *Cancer Cell Int.* **2020**, *20*, 102. [[CrossRef](#)]
35. Samudh, N.; Shrilall, C.; Arbuthnot, P.; Bloom, K.; Ely, A. Diversity of Dysregulated Long Non-Coding RNAs in HBV-Related Hepatocellular Carcinoma. *Front. Immunol.* **2022**, *13*, 129. [[CrossRef](#)]
36. Khan, F.H.; Pandian, V.; Ramraj, S.K.; Aravindan, S.; Natarajan, M.; Azadi, S.; Herman, T.S.; Aravindan, N. RD3 loss dictates high-risk aggressive neuroblastoma and poor clinical outcomes. *Oncotarget* **2015**, *6*, 36522. [[CrossRef](#)]
37. Xia, L.; Zhang, W.; Gao, L. Clinical and prognostic effects of CDKN2A, CDKN2B and CDH13 promoter methylation in ovarian cancer: A study using meta-analysis and TCGA data. *Biomarkers* **2019**, *24*, 700–711. [[CrossRef](#)]
38. Goldstein, A.M.; Chan, M.; Harland, M.; Gillanders, E.M.; Hayward, N.K.; Avril, M.-F.; Azizi, E.; Bianchi-Scarra, G.; Bishop, D.T.; Bressac-de Paillerets, B. High-risk melanoma susceptibility genes and pancreatic cancer, neural system tumors, and uveal melanoma across GenoMEL. *Cancer Res.* **2006**, *66*, 9818–9828. [[CrossRef](#)]
39. Chen, W.S.; Bindra, R.S.; Mo, A.; Hayman, T.; Husain, Z.; Contessa, J.N.; Gaffney, S.G.; Townsend, J.P.; Yu, J.B. CDKN2A copy number loss is an independent prognostic factor in HPV-negative head and neck squamous cell carcinoma. *Front. Oncol.* **2018**, *8*, 95. [[CrossRef](#)]
40. Zhao, R.; Choi, B.Y.; Lee, M.-H.; Bode, A.M.; Dong, Z. Implications of genetic and epigenetic alterations of CDKN2A (p16INK4a) in cancer. *EBioMedicine* **2016**, *8*, 30–39. [[CrossRef](#)]
41. Butcher, N.; Boukouvala, S.; Sim, E.; Minchin, R. Pharmacogenetics of the arylamine N-acetyltransferases. *Pharm. J.* **2002**, *2*, 30–42. [[CrossRef](#)] [[PubMed](#)]
42. Sim, E.; Abuhammad, A.; Ryan, A. Arylamine N-acetyltransferases: From drug metabolism and pharmacogenetics to drug discovery. *Br. J. Pharmacol.* **2014**, *171*, 2705–2725. [[CrossRef](#)] [[PubMed](#)]
43. Knowles, J.W.; Xie, W.; Zhang, Z.; Chennemsetty, I.; Assimes, T.L.; Paananen, J.; Hansson, O.; Pankow, J.; Goodarzi, M.O.; Carcamo-Orive, I. Identification and validation of N-acetyltransferase 2 as an insulin sensitivity gene. *J. Clin. Investig.* **2015**, *125*, 1739–1751. [[CrossRef](#)] [[PubMed](#)]
44. Tarao, K.; Nozaki, A.; Ikeda, T.; Sato, A.; Komatsu, H.; Komatsu, T.; Taguri, M.; Tanaka, K. Real impact of liver cirrhosis on the development of hepatocellular carcinoma in various liver diseases—Meta-analytic assessment. *Cancer Med.* **2019**, *8*, 1054–1065. [[CrossRef](#)]
45. Wang, C.; Wang, X.; Gong, G.; Ben, Q.; Qiu, W.; Chen, Y.; Li, G.; Wang, L. Increased risk of hepatocellular carcinoma in patients with diabetes mellitus: A systematic review and meta-analysis of cohort studies. *Int. J. Cancer* **2012**, *130*, 1639–1648. [[CrossRef](#)]
46. Engström, P.G.; Suzuki, H.; Ninomiya, N.; Akalin, A.; Sessa, L.; Lavorgna, G.; Brozzi, A.; Luzi, L.; Tan, S.L.; Yang, L. Complex loci in human and mouse genomes. *PLoS Genet.* **2006**, *2*, e47. [[CrossRef](#)] [[PubMed](#)]
47. Dai, X.; Xiang, L.; Li, T.; Bai, Z. Cancer hallmarks, biomarkers and breast cancer molecular subtypes. *J. Cancer* **2016**, *7*, 1281. [[CrossRef](#)]
48. Lin, C.-Y.; Kleinbrink, E.L.; Dachet, F.; Cai, J.; Ju, D.; Goldstone, A.; Wood, E.J.; Liu, K.; Jia, H.; Goustin, A.-S. Primate-specific oestrogen-responsive long non-coding RNAs regulate proliferation and viability of human breast cancer cells. *Open Biol.* **2016**, *6*, 150262. [[CrossRef](#)]
49. Koundouros, N.; Poulogiannis, G. Reprogramming of fatty acid metabolism in cancer. *Br. J. Cancer* **2020**, *122*, 4–22. [[CrossRef](#)]
50. Mahajan, A.; Taliun, D.; Thurner, M.; Robertson, N.R.; Torres, J.M.; Rayner, N.W.; Payne, A.J.; Steinthorsdottir, V.; Scott, R.A.; Grarup, N. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat. Genet.* **2018**, *50*, 1505–1513. [[CrossRef](#)]
51. Scott, R.A.; Scott, L.J.; Mägi, R.; Marullo, L.; Gaulton, K.J.; Kaakinen, M.; Pervjakova, N.; Pers, T.H.; Johnson, A.D.; Eicher, J.D. An expanded genome-wide association study of type 2 diabetes in Europeans. *Diabetes* **2017**, *66*, 2888–2902. [[CrossRef](#)] [[PubMed](#)]
52. Khatib, A.-M.; Sfaki, F. FURIN (furin (paired basic amino acid cleaving enzyme)). *Atlas Genet. Oncol.* **2012**, 638. [[CrossRef](#)]
53. Johnson, B.A.; Xie, X.; Kalveram, B.; Lokugamage, K.G.; Muruato, A.; Zou, J.; Zhang, X.; Juelich, T.; Smith, J.K.; Zhang, L. Furin cleavage site is key to SARS-CoV-2 pathogenesis. *BioRxiv* **2020**. Available online: <https://www.biorxiv.org/content/10.1101/2020.08.26.268854v1> (accessed on 1 January 2020).

54. He, Z.; Khatib, A.-M.; Creemers, J.W. The proprotein convertase furin in cancer: More than an oncogene. *Oncogene* **2022**, *41*, 1252–1262. [[CrossRef](#)]
55. Liu, X.; Li, Y.; Meng, L.; Liu, X.-Y.; Peng, A.; Chen, Y.; Liu, C.; Chen, H.; Sun, S.; Miao, X. Reducing protein regulator of cytokinesis 1 as a prospective therapy for hepatocellular carcinoma. *Cell Death Dis.* **2018**, *9*, 534. [[CrossRef](#)]
56. Cai, Y.; Kirschke, C.P.; Huang, L. SLC30A family expression in the pancreatic islets of humans and mice: Cellular localization in the β -cells. *J. Mol. Histol.* **2018**, *49*, 133–145. [[CrossRef](#)]
57. Chang, J.; Yu, Y.; Fang, Z.; He, H.; Wang, D.; Teng, J.; Yang, L. Long non-coding RNA CDKN2B-AS1 regulates high glucose-induced human mesangial cell injury via regulating the miR-15b-5p/WNT2B axis. *Diabetol. Metab. Syndr.* **2020**, *12*, 109. [[CrossRef](#)]
58. Zhang, Z.; Gu, M.; Gu, Z.; Lou, Y.-R. Role of Long Non-Coding RNA Polymorphisms in Cancer Chemotherapeutic Response. *J. Pers. Med.* **2021**, *11*, 513. [[CrossRef](#)]
59. Nie, F.-q.; Sun, M.; Yang, J.-s.; Xie, M.; Xu, T.-p.; Xia, R.; Liu, Y.-w.; Liu, X.-h.; Zhang, E.-b.; Lu, K.-h. Long noncoding RNA ANRIL promotes non-small cell lung cancer cell proliferation and inhibits apoptosis by silencing KLF2 and P21 expression. *Mol. Cancer Ther.* **2015**, *14*, 268–277. [[CrossRef](#)]
60. Lin, L.; Gu, Z.-T.; Chen, W.-H.; Cao, K.-J. Increased expression of the long non-coding RNA ANRIL promotes lung cancer cell metastasis and correlates with poor prognosis. *Diagn. Pathol.* **2015**, *10*, 14. [[CrossRef](#)]
61. Huang, X.; Zhang, W.; Shao, Z. Association between long non-coding RNA polymorphisms and cancer risk: A meta-analysis. *Biosci. Rep.* **2018**, *38*, BSR20180365. [[CrossRef](#)] [[PubMed](#)]
62. Gong, W.-j.; Peng, J.-b.; Yin, J.-y.; Li, X.-p.; Zheng, W.; Xiao, L.; Tan, L.-m.; Xiao, D.; Chen, Y.-x.; Li, X. Association between well-characterized lung cancer lncRNA polymorphisms and platinum-based chemotherapy toxicity in Chinese patients with lung cancer. *Acta Pharmacol. Sin.* **2017**, *38*, 581–590. [[CrossRef](#)] [[PubMed](#)]
63. Volkogon, A.; Harbuzova, V.Y.; Ataman, A. Analysis of ANRIL gene polymorphism rs4977574 association with kidney cancer development in Ukrainian population. *Med. Perspect.* **2020**, *25*, 60–65. [[CrossRef](#)]
64. Pasquale, L.R.; Loomis, S.J.; Kang, J.H.; Yaspan, B.L.; Abdrabou, W.; Budenz, D.L.; Chen, T.C.; DelBono, E.; Friedman, D.S.; Gaasterland, D. CDKN2B-AS1 genotype–glaucoma feature correlations in primary open-angle glaucoma patients from the United States. *Am. J. Ophthalmol.* **2013**, *155*, 342–353.e5. [[CrossRef](#)] [[PubMed](#)]
65. Thakur, N.; Kupani, M.; Mannan, R.; Pruthi, A.; Mehrotra, S. Genetic association between CDKN2B/CDKN2B-AS1 gene polymorphisms with primary glaucoma in a North Indian cohort: An original study and an updated meta-analysis. *BMC Med. Genom.* **2021**, *14*, 1. [[CrossRef](#)] [[PubMed](#)]
66. Chen, Y.D.; Zhang, N.; Qiu, X.G.; Yuan, J.; Yang, M. LncRNA CDKN2BAS rs2157719 genetic variant contributes to medulloblastoma predisposition. *J. Gene Med.* **2018**, *20*, e3000. [[CrossRef](#)] [[PubMed](#)]
67. Dahlin, A.M.; Wibom, C.; Andersson, U.; Hougaard, D.M.; Bybjerg-Grauholm, J.; Deltour, I.; Hultman, C.M.; Kähler, A.K.; Karlsson, R.; Hjalmar, U. Genetic variants in the 9p21. 3 locus associated with glioma risk in children, adolescents, and young adults: A case–control study. *Cancer Epidemiol. Prev. Biomark.* **2019**, *28*, 1252–1258. [[CrossRef](#)]
68. Tam, V.; Patel, N.; Turcotte, M.; Bossé, Y.; Paré, G.; Meyre, D. Benefits and limitations of genome-wide association studies. *Nat. Rev. Genet.* **2019**, *20*, 467–484. [[CrossRef](#)]
69. Frayling, T. Genome-wide association studies: The good, the bad and the ugly. *Clin. Med.* **2014**, *14*, 428. [[CrossRef](#)]
70. van den Berg, B.H.; McCarthy, F.M.; Lamont, S.J.; Burgess, S.C. Re-annotation is an essential step in systems biology modeling of functional genomics data. *PLoS ONE* **2010**, *5*, e10642. [[CrossRef](#)]
71. Vincent, E.E.; Yaghoobkar, H. Using genetics to decipher the link between type 2 diabetes and cancer: Shared aetiology or downstream consequence? *Diabetologia* **2020**, *63*, 1706–1717. [[CrossRef](#)] [[PubMed](#)]
72. Ma, C.; Gu, L.; Yang, M.; Zhang, Z.; Zeng, S.; Song, R.; Xu, C.; Sun, Y. rs1495741 as a tag single nucleotide polymorphism of N-acetyltransferase 2 acetylator phenotype associates bladder cancer risk and interacts with smoking: A systematic review and meta-analysis. *Medicine* **2016**, *95*, e4417. [[CrossRef](#)] [[PubMed](#)]
73. Ahmad, S.; Mora, S.; Ridker, P.M.; Hu, F.B.; Chasman, D.I. Gene-based elevated triglycerides and type 2 diabetes mellitus risk in the women’s genome health study. *Arterioscler. Thromb. Vasc. Biol.* **2019**, *39*, 97–106. [[CrossRef](#)] [[PubMed](#)]
74. Chamorro, J.G.; Castagnino, J.P.; Musella, R.M.; Nogueras, M.; Frías, A.; Visca, M.; Aidar, O.; Costa, L.; de Larrañaga, G.F. tagSNP rs1495741 as a useful molecular marker to predict antituberculosis drug-induced hepatotoxicity. *Pharm. Genom.* **2016**, *26*, 357–361. [[CrossRef](#)]
75. Wu, H.; Wang, X.; Zhang, L.; Mo, N.; Lv, Z. Association between N-acetyltransferase 2 polymorphism and bladder cancer risk: Results from studies of the past decade and a meta-analysis. *Clin. Genitourin. Cancer* **2016**, *14*, 122–129. [[CrossRef](#)]
76. Al-Shaqha, W.M.; Alkharfy, K.M.; Al-Daghri, N.M.; Mohammed, A.K. N-acetyltransferase 1 and 2 polymorphisms and risk of diabetes mellitus type 2 in a Saudi population. *Ann. Saudi Med.* **2015**, *35*, 214–221. [[CrossRef](#)]
77. Jin, L.; Zheng, D.; Bhandari, A.; Chen, D.; Xia, E.; Guan, Y.; Wen, J.; Wang, O. PSD3 is an oncogene that promotes proliferation, migration, invasion, and G1/S transition while inhibits apoptotic in papillary thyroid cancer. *J. Cancer* **2021**, *12*, 5413. [[CrossRef](#)]
78. Gong, S.; Xu, C.; Wang, L.; Liu, Y.; Owusu, D.; Bailey, B.A.; Li, Y.; Wang, K. Genetic association analysis of polymorphisms in PSD3 gene with obesity, type 2 diabetes, and HDL cholesterol. *Diabetes Res. Clin. Pract.* **2017**, *126*, 105–114. [[CrossRef](#)]
79. Song, C.; Qi, Y.; Zhang, J.; Guo, C.; Yuan, C. CDKN2B-AS1: An indispensable long non-coding RNA in multiple diseases. *Curr. Pharm. Des.* **2020**, *26*, 5335–5346. [[CrossRef](#)]

80. Chennamsetty, I.; Coronado, M.; Contrepolis, K.; Keller, M.P.; Carcamo-Orive, I.; Sandin, J.; Fajardo, G.; Whittle, A.J.; Fathzadeh, M.; Snyder, M. Nat1 deficiency is associated with mitochondrial dysfunction and exercise intolerance in mice. *Cell Rep.* **2016**, *17*, 527–540. [[CrossRef](#)]
81. Fernandez, C.; Rysä, J.; Almgren, P.; Nilsson, J.; Engström, G.; Orho-Melander, M.; Ruskoaho, H.; Melander, O. Plasma levels of the proprotein convertase furin and incidence of diabetes and mortality. *J. Intern. Med.* **2018**, *284*, 377–387. [[CrossRef](#)] [[PubMed](#)]
82. Adu-Agyeiwaah, Y.; Grant, M.B.; Obukhov, A.G. The potential role of osteopontin and furin in worsening disease outcomes in COVID-19 patients with pre-existing diabetes. *Cells* **2020**, *9*, 2528. [[CrossRef](#)] [[PubMed](#)]
83. Ryu, M.-S.; Lichten, L.A.; Liuzzi, J.P.; Cousins, R.J. Zinc transporters ZnT1 (Slc30a1), Zip8 (Slc39a8), and Zip10 (Slc39a10) in mouse red blood cells are differentially regulated during erythroid development and by dietary zinc deficiency. *J. Nutr.* **2008**, *138*, 2076–2083. [[CrossRef](#)] [[PubMed](#)]
84. Yamasaki, S.; Sakata-Sogawa, K.; Hasegawa, A.; Suzuki, T.; Kabu, K.; Sato, E.; Kurosaki, T.; Yamashita, S.; Tokunaga, M.; Nishida, K. Zinc is a novel intracellular second messenger. *J. Cell Biol.* **2007**, *177*, 637–645. [[CrossRef](#)]
85. Norouzi, S.; Adulcikas, J.; Sohal, S.S.; Myers, S. Zinc transporters and insulin resistance: Therapeutic implications for type 2 diabetes and metabolic disease. *J. Biomed. Sci.* **2017**, *24*, 87. [[CrossRef](#)]
86. Cotrozzi, G.; Relli, P.; Buzzelli, G. Role of the liver in the regulation of glucose metabolism in diabetes and chronic liver disease. *Ann. Ital. Med. Interna Organo Uff. Della Soc. Ital. Med. Interna* **1997**, *12*, 84–91.
87. Damase, T.R.; Sukhovshin, R.; Boada, C.; Taraballi, F.; Pettigrew, R.L.; Cooke, J.P. The limitless future of RNA therapeutics. *Front. Bioeng. Biotechnol.* **2021**, *9*, 161. [[CrossRef](#)]
88. Feng, R.; Patil, S.; Zhao, X.; Miao, Z.; Qian, A. RNA Therapeutics-Research and Clinical Advancements. *Front. Mol. Biosci.* **2021**, *8*, 913. [[CrossRef](#)]
89. Gaulton, K.J.; Ferreira, T.; Lee, Y.; Raimondo, A.; Mägi, R.; Reschen, M.E.; Mahajan, A.; Locke, A.; William Rayner, N.; Robertson, N. Genetic fine mapping and genomic annotation defines causal mechanisms at type 2 diabetes susceptibility loci. *Nat. Genet.* **2015**, *47*, 1415–1425. [[CrossRef](#)]