

Supplementary Table 1. Performance of gene expression prediction models built

Prediction performance (R^2)	All	Protein	lncRNAs	miRNAs
≥ 0.01	8479	6788	1690	1
≥ 0.04	5623	4495	1127	1
≥ 0.09	3350	2656	693	1
≥ 0.16	2010	1602	408	0

Protein: Protein coding genes; lncRNAs: long non-coding RNAs; miRNAs: microRNAs.

Supplementary Table 4. The functions of *RCCD1*, and *CFDPI* and their reported links with human cancers

Region	Gene name	Functions	Cancer	References (PMID)
15q26.1	<i>RCCD1</i>	The <i>RCCD1</i> (<i>RCC1</i> domain-containing protein 1) protein plays an important role in regulating alpha-tubulin deacetylation and cytoskeletal microtubule stability, thereby promoting cell migration and TGF-beta-induced epithelial to mesenchymal transition (EMT).	Lung cancer	31582214
				28455245
			Breast cancer	28362817
			Ovarian cancer	27432226
16q23.1	<i>CFDPI</i>	Gene <i>CFDPI</i> (Craniofacial development protein 1) is 139,815 bp long with 7 exons and 6 introns and maps to chromosome 16 in the band 16q22.2-q22.3. <i>CFDPI</i> gene expression has been detected in a wide range of cancer tissues.	Esophageal adenocarcinoma	31891614
			Angiotropic melanoma	23275074

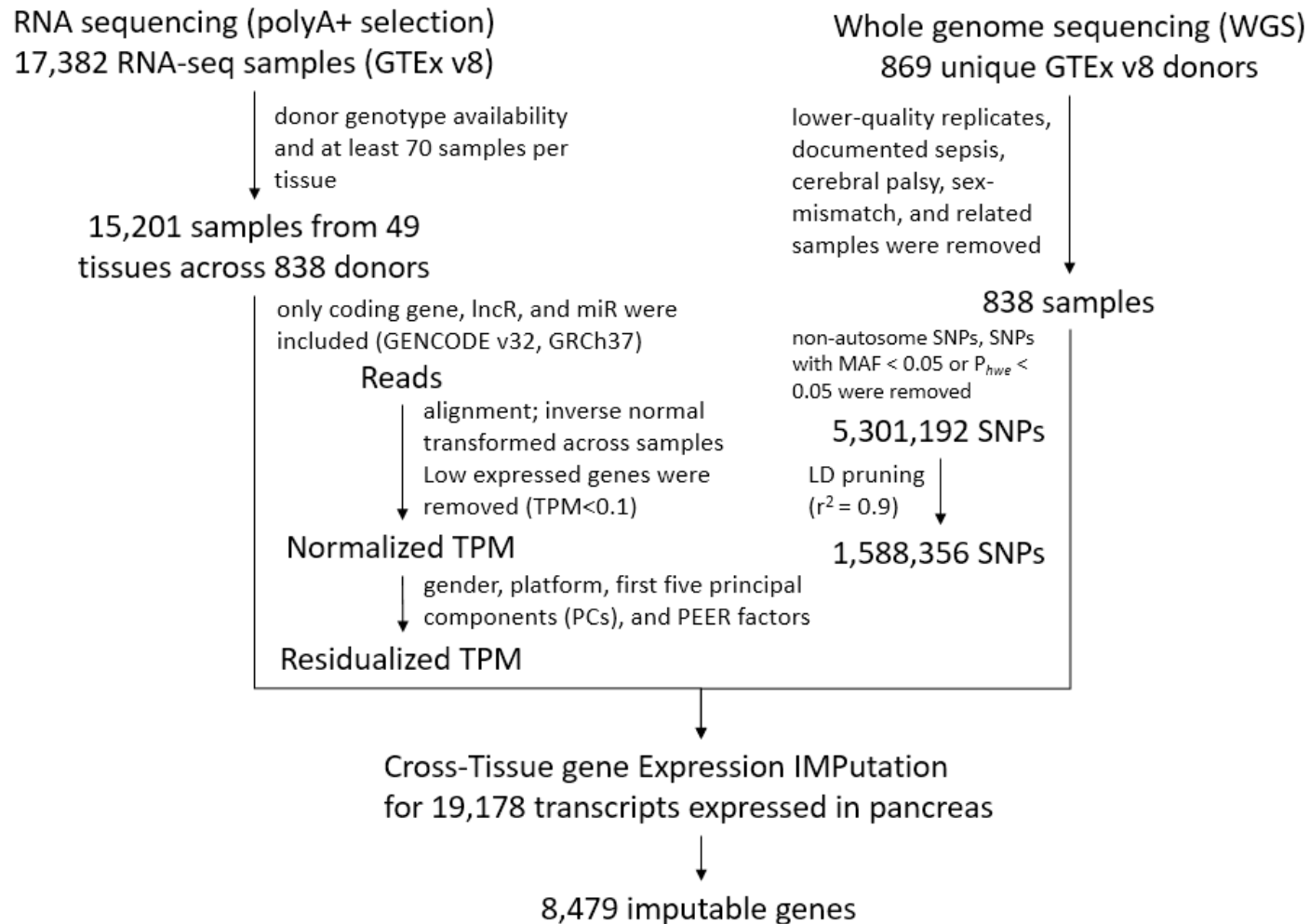
Supplementary Table 5. Associations of genes identified in Zhong et al. with pancreatic cancer risk in the present study

Previous TWAS (Zhong et al.)						The present study		
Region	Gene name	Approach	Training tissue	Z	P value after FDR	OR	95% CI	P value
1p36.12	<i>CELA3B</i>	FUSION	GTEEx pancreas	-3.98	6.89×10^{-5}	1.02	0.09-11.49	0.99
		FUSION	Combined pancreas	-4.62	3.80×10^{-6}			
		MetaXcan	GTEEx pancreas	-4.43	9.38×10^{-6}			
		MetaXcan	Combined pancreas	-4.29	1.83×10^{-5}			
5p15.33	<i>TERT</i>	SMulTiXcan	Cross tissue	-8.24 to 4.20	5.80×10^{-18}	No prediction model with $R^2 \geq 0.01$ built		
5p15.33	<i>CLPTMIL</i>	SMulTiXcan	Cross tissue	-8.33 to 0.91	1.48×10^{-16}	No prediction model with $R^2 \geq 0.01$ built		
5p15.33	<i>ZDHHC11B</i>	SMulTiXcan	Cross tissue	-1.6 to 3.13	3.18×10^{-6}	1.21	0.94-1.56	0.13
7p14.1	<i>INHBA</i>	MetaXcan	GTEEx pancreas	-5.11	3.20×10^{-7}	0.56	0.47-0.68	8.44×10^{-7}
		SMulTiXcan	Cross tissue	-5.11 to -0.72	4.10×10^{-6}			
9p31.1	<i>SMC2</i>	FUSION	Combined pancreas	4.95	7.52×10^{-7}	2.80	2.04-3.83	7.71×10^{-8}
		MetaXcan	Combined pancreas	4.93	8.19×10^{-7}			
		SMulTiXcan	Cross tissue	-3.34 to 5.35	8.50×10^{-6}			
9p31.1	<i>SMC2-AS1</i>	SMulTiXcan	Cross tissue	-4.9 to 4.8	1.39×10^{-5}	No prediction model with $R^2 \geq 0.01$ built		
9q34.2	<i>ABO</i>	FUSION	LTG pancreas	9.38	6.71×10^{-21}	1.25	1.17-1.33	1.38×10^{-8}
		FUSION	GTEEx pancreas	6.96	3.44×10^{-12}			
		FUSION	Combined pancreas	7.55	4.34×10^{-14}			
		MetaXcan	LTG pancreas	10.72	8.07×10^{-27}			

		MetaXcan	GTE _x pancreas	7.08	1.41×10^{-12}			
		MetaXcan	Combined pancreas	7.65	2.05×10^{-14}			
10q23.31	<i>RP11-80H5.9</i>	SMulTiXcan	Cross tissue	-2.21 to 4.4	8.23×10^{-6}	No prediction model with $R^2 \geq 0.01$ built		
12q13.13	<i>SMUG1</i>	FUSION	GTE _x pancreas	-4.04	5.40×10^{-5}	0.90	0.84-0.96	3.04×10^{-3}
13q12.2	<i>PDX1</i>	MetaXcan	GTE _x pancreas	-7.18	6.85×10^{-13}	0.50	0.42-0.60	5.29×10^{-10}
		SMulTiXcan	Cross tissue	-7.18 to -6.59	4.87×10^{-12}			
13q22.1	<i>KLF5</i>	FUSION	GTE _x pancreas	4.91	9.17×10^{-7}	No prediction model with $R^2 \geq 0.01$ built		
		FUSION	Combined pancreas	5.92	3.15×10^{-9}			
14q32.33	<i>BTBD6</i>	FUSION	GTE _x pancreas	4.06	4.98×10^{-5}	1.52	1.14-2.02	4.56×10^{-3}
		FUSION	Combined pancreas	4.00	6.30×10^{-5}			
15q23	<i>HEXA</i>	FUSION	Combined pancreas	-4.02	5.68×10^{-5}	No prediction model with $R^2 \geq 0.01$ built		
15q26.1	<i>RCCD1</i>	FUSION	LTG pancreas	-3.98	6.94×10^{-5}	0.85	0.79-0.91	2.48×10^{-4}
		FUSION	GTE _x pancreas	-4.04	5.38×10^{-5}			
		FUSION	Combined pancreas	-3.99	6.52×10^{-5}			
16q23.1	<i>WDR59</i>	FUSION	Combined pancreas	-4.70	2.54×10^{-6}	No prediction model with $R^2 \geq 0.01$ built		
16q23.1	<i>CFDPI</i>	FUSION	LTG pancreas	6.07	1.26×10^{-9}	1.51	1.33-1.72	1.29×10^{-7}
		FUSION	Combined pancreas	5.76	8.47×10^{-9}			
		MetaXcan	Combined pancreas	5.58	2.40×10^{-8}			
		SMulTiXcan	Cross tissue	2.50 to 6.89	2.02×10^{-8}			

16q23.1	<i>BCAR1</i>	SMulTiXcan	Cross tissue	-5.60 to 6.49	1.94×10^{-7}	No prediction model with $R^2 \geq 0.01$ built		
16q23.1	<i>TMEM170A</i>	SMulTiXcan	Cross tissue	-3.69 to 2.86	1.21×10^{-5}	No prediction model with $R^2 \geq 0.01$ built		
17q12	<i>PNMT</i>	FUSION	LTG pancreas	4.86	1.20×10^{-6}	No prediction model with $R^2 \geq 0.01$ built		
17q12	<i>CDK12</i>	FUSION	GTE _x pancreas	-4.05	5.15×10^{-5}	No prediction model with $R^2 \geq 0.01$ built		
17q12	<i>PGAP3</i>	FUSION	LTG pancreas	3.91	9.11×10^{-5}	1.19	1.10-1.29	3.32×10^{-4}
		FUSION	GTE _x pancreas	3.98	6.96×10^{-5}			
		MetaXcan	GTE _x pancreas	4.11	3.03×10^{-5}			
		MetaXcan	Combined pancreas	4.17	2.98×10^{-5}			
17q22	<i>SUPT4H1</i>	FUSION	GTE _x pancreas	4.12	3.72×10^{-5}	1.19	0.97-1.46	0.099
		MetaXcan	GTE _x pancreas	4.11	3.90×10^{-5}			
18q11.22	<i>RP11-888D10.3</i>	MetaXcan	GTE _x pancreas	-4.07	4.67×10^{-5}	No prediction model with $R^2 \geq 0.01$ built		
19p13.11	<i>PGPEP1</i>	MetaXcan	Combined pancreas	-4.13	3.67×10^{-5}	0.81	0.73-0.90	1.46×10^{-4}

Supplementary Figure 1. The flowchart of quality control and prediction model training in the reference dataset. GTEx v8 was used as the reference dataset to build the prediction model. More details can be found in the preprint of the v8 release (1). Totally 838 samples and 49 tissues were included in the model training. The weights were jointly estimated by (penalized) regressing the residual of gene expression on genotypes.



Reference

1. Aguet F, Barbeira, A. N., Bonazzola, R., Brown, A., Castel, S. E., Jo, B., et al. The GTEx Consortium atlas of genetic regulatory effects across human tissues. BioRxiv2019.