The American Journal of Human Genetics, Volume *108*

Supplemental information

A 584 bp deletion in *CTRB2* inhibits chymotrypsin B2 activity and secretion and confers risk of pancreatic cancer

Ashley Jermusyk, Jun Zhong, Katelyn E. Connelly, Naomi Gordon, Sumeth Perera, Ehssan Abdolalizadeh, Tongwu Zhang, Aidan O'Brien, Jason W. Hoskins, Irene Collins, Daina Eiser, Chen Yuan, PanScan Consortium, PanC4 Consortium, Harvey A. Risch, Eric J. Jacobs, Donghui Li, Mengmeng Du, Rachael Z. Stolzenberg-Solomon, Alison P. Klein, Jill P. Smith, Brian M. Wolpin, Stephen J. Chanock, Jianxin Shi, Gloria M. Petersen, Christopher J. Westlake, and Laufey T. Amundadottir

SUPPLEMENTARY MATERIAL

Consortia authors for the Pancreatic Cancer Cohort Consortium (PanScan) and the Pancreatic Cancer Case Control Consortium (PanC4).

Demetrius Albanes[1], Alan A. Arslan[2], Aurelio Barricarte Gurrea[3], Laura Beane-Freeman[4], Paige M. Bracci[5], Bas Bueno-de-Mesquita[6], Julie Buring[7], Federico Canzian[8], Mengmeng Du[9], Stephen Gallinger S[10], J. Michael Gaziano[11], Graham G. Giles[12], Phyllis J. Goodman[13], Eric Jacobs[14], Mattias Johansson[15], Charles Kooperberg[16], Loic LeMarchand[17], Nuria Malats[18], Rachel E. Neale[19], Salvatore Panico[20], Ulrike Peters[16], Francisco X. Real[21], Xiao-Ou Shu[22], Malin Sund[23], Marc Thornquist[16], Anne Tjønneland[24], Ruth C. Travis[25], Stephen K. Van Den Eeden[26], Kala Visvanathan K[27], Wei Zheng[22], Donghui Li[28], Harvey A. Risch[29], Stephen J. Chanock S[30], Peter Kraft P[31], Brian M. Wolpin[32], Gloria M. Petersen[33], Rachael Z. Stolzenberg-Solomon[30], Alison P. Klein[34], Laufey T. Amundadottir[35]

[1]Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD,USA, [2]Department of Obstetrics and Gynecology, Department of Environmental Medicine, and Department of Population Health, New York University School of Medicine, New York, NY, USA, [3]Instituto de Salud Pública de Navarra, Pamplona, Spain, [4]Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA, [5]Department of Epidemiology and Biostatistics, University of California, San Francisco, CA,USA, [6]Department for Determinants of Chronic Diseases (DCD), National Institute for Public Health and the Environment (RIVM), Utrecht, The Netherlands, [7]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, MA, USA, [8]Genomic Epidemiology Group, German Cancer Research Center (DKFZ), Heidelberg, Germany, [9]Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, USA, [10]Prosserman Centre for Population Health Research, Lunenfeld-Tanenbaum Research Institute, Sinai Health System, Toronto, ON, Canada, [11]Departments of Medicine, Brigham and Women's Hospital, VA Boston, and Harvard Medical School, Boston, MA, USA., [12]Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, VIC, Australia.; Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Parkville, VIC, Australia.; Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, VIC, Australia, [13]SWOG Statistical Center, Fred Hutchinson Cancer Research Center, Seattle, WA or Fred Hutchinson Cancer Research Center, Seattle, WA, USA, [14]Behavioral and Epidemiology Research Group, American Cancer Society, Atlanta, GA, USA, [15]International Agency for Research on Cancer (IARC), 150 cours Albert Thomas, Lyon, France, [16]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA, USA, [17]Cancer Epidemiology Program, University of Hawaii Cancer Center, Honolulu, HI, USA, [18]Genetic and Molecular Epidemiology Group, Centro Nacional de Investigaciones Oncológicas (CNIO), Madrid, Spain, [19]Population Health Department, QIMR Berghofer Medical Research Institute, Brisbane, Australia, [20]Dipartimento di Medicina Clinica e Chirugia, Federico II University, Naples, Italy, [21]Centro de Investigación Biomédica en Red de Cáncer (CIBERONC), 28029 Madrid, Spain., Epithelial Carcinogenesis Group, Spanish National Cancer Research Centre (CNIO), 28029 Madrid, Spain., Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, 08193 Barcelona, Spain. [22]Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, TN, USA, [23]Department of Surgical and Perioperative Sciences, Umeå University, Umeå, Sweden, [24]Danish Cancer Society Research Center and Department of Public Health, University of Copenhagen, Copenhagen, Denmark, [25]Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford OX3 0NR, UK, [26]Kaiser Permanente Division of Research, Oakland, CA, USA, [27]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA, [28]Division of Cancer Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX,USA, [29]Department of Epidemiology and Public Health, Yale School of Public Health, New Haven, CT, USA, [30]Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, DHHS, Rockville, MD, USA, [31]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA, [32]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, USA, [33]Department of Health Sciences Research, Mayo Clinic College of Medicine, Rochester, MN, [34]Department of Oncology, Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins School of Medicine, Baltimore, MD, USA, [35]Laboratory of Translational Genomics, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA.

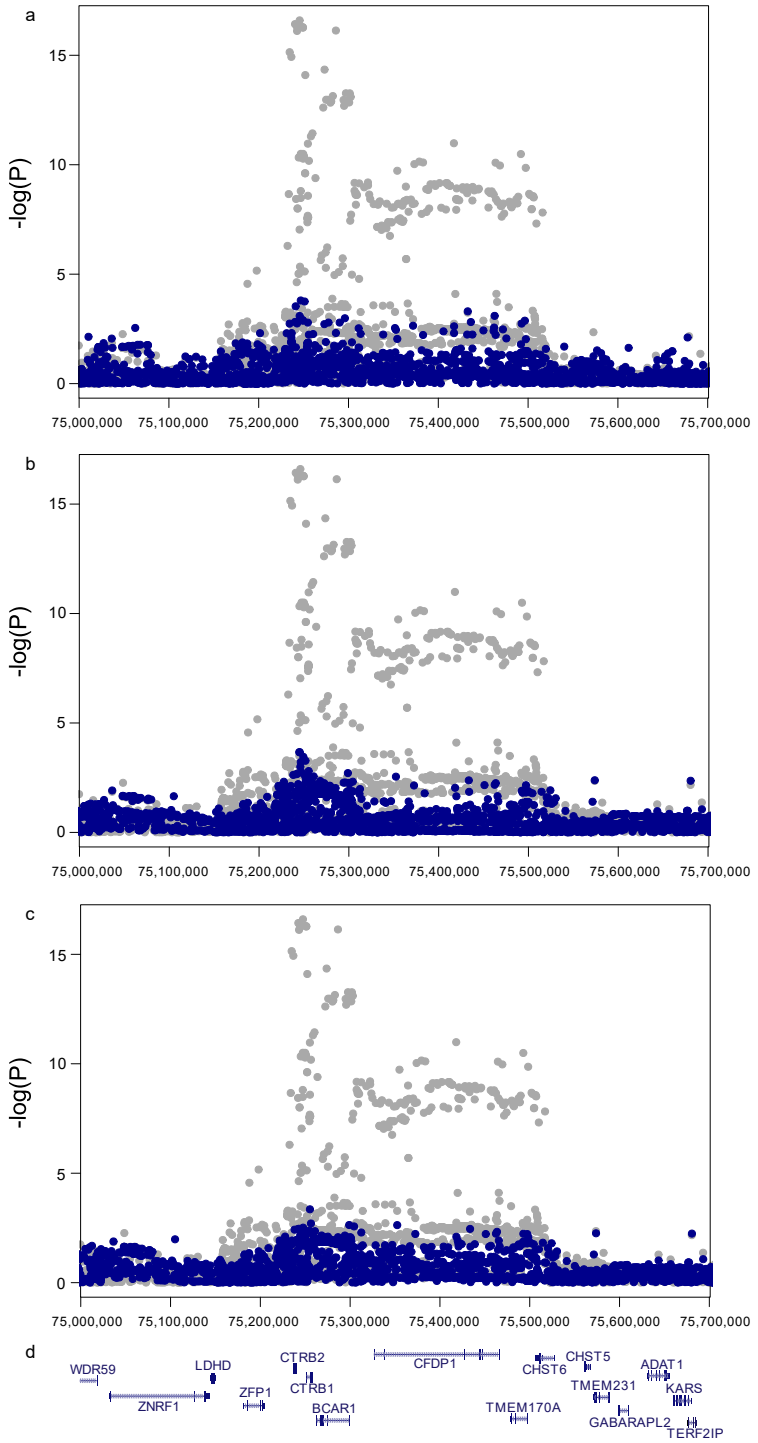

Fig. S1: Results from the association analysis at the 16q23.1 pancreatic cancer risk locus after conditioning the analysis on select variants. The association results from the meta-analysis of PanScan I, II, III, and PanC4 before (gray) and after conditioning (blue) on: (**a**) the original marker SNP rs13337397, (**b**) the most significant tag SNP after imputation rs72802365, and, (**c**) the 584 bp genomic insertion/deletion variant in *CTRB2*. NCBI RefSeq genes within the chr16q23.1 TAD as visualized using UCSC genome browser (Hg19) are shown in (**d**).

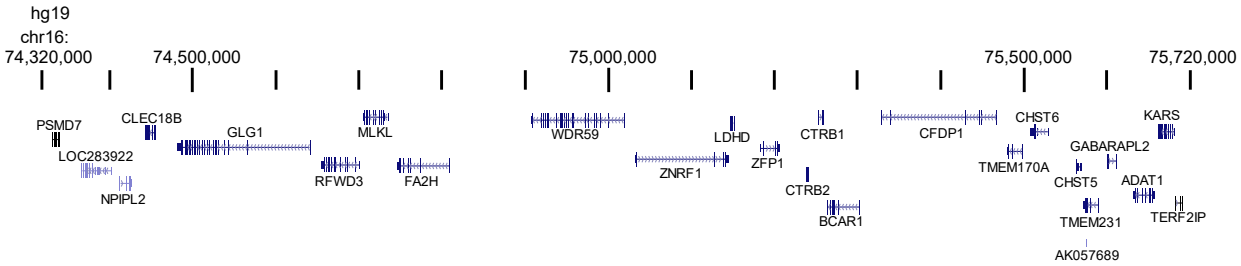

Fig. S2: The genomic region of the topologically associated domain (TAD) at the chr16q23.1 pancreatic cancer risk locus. NCBI RefSeq genes within the chr16q23.1 TAD are visualized using the UCSC genome browser (Hg19).

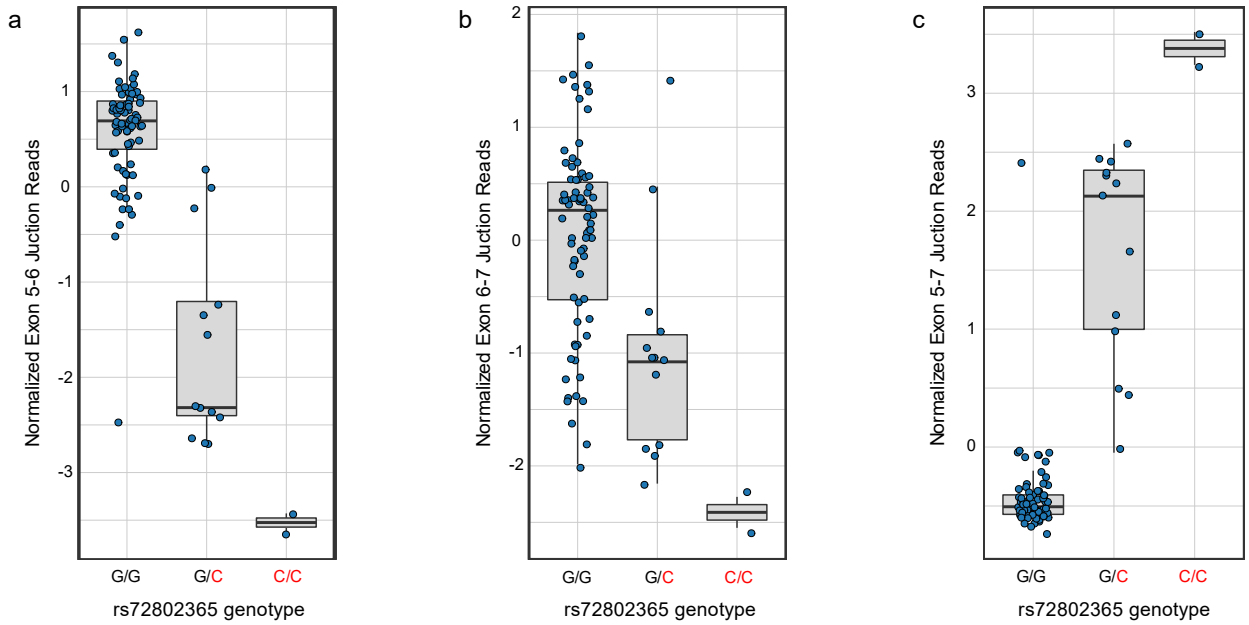

Fig. S3: *CTRB2* splicing QTLs at the chr16q23.1 pancreatic cancer risk locus for rs72802365 in the LTG (histologically normal) pancreatic QTL dataset. Boxplots show normalized *CTRB2* junction reads generated by Leafcutter in LTG pancreas samples for *CTRB2* Exon 5-6 (**a**), Exon 6-7 (**b**), and Exon 5-7 (**c**). The risk increasing allele (C) is indicated in red. Genotype counts are: G/G n=75, G/C n=13, C/C n=2.

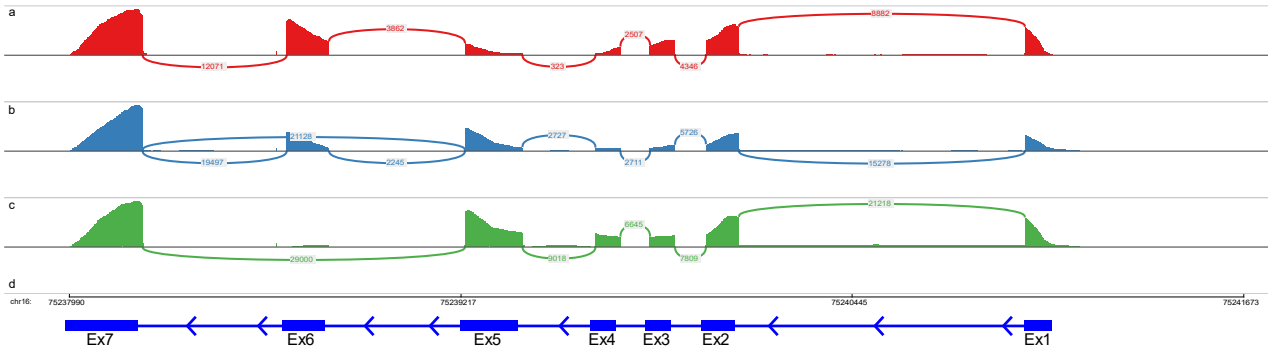

Fig. S4: A Sashimi plot showing *CTRB2* splicing in individuals with different genotypes for the *CTRB2* exon 6 indel variant. Sashimi plot of raw RNA-seq exonic read densities from representative individuals showing junction reads with: (**a**) no copies, (**b**) one copy, and (**c**) two copies, of the *CTRB2* exon 6 genomic deletion as visualized using the IGV browser[1]. Note the difference in RNA-seq read density over exon 6 in samples with 0 (a), 1 (b) and 2 (c) copies of the deletion allele. (**d**), Genomic coordinates (Hg19) and the exon/intron structure of the *CTRB2* gene are shown.

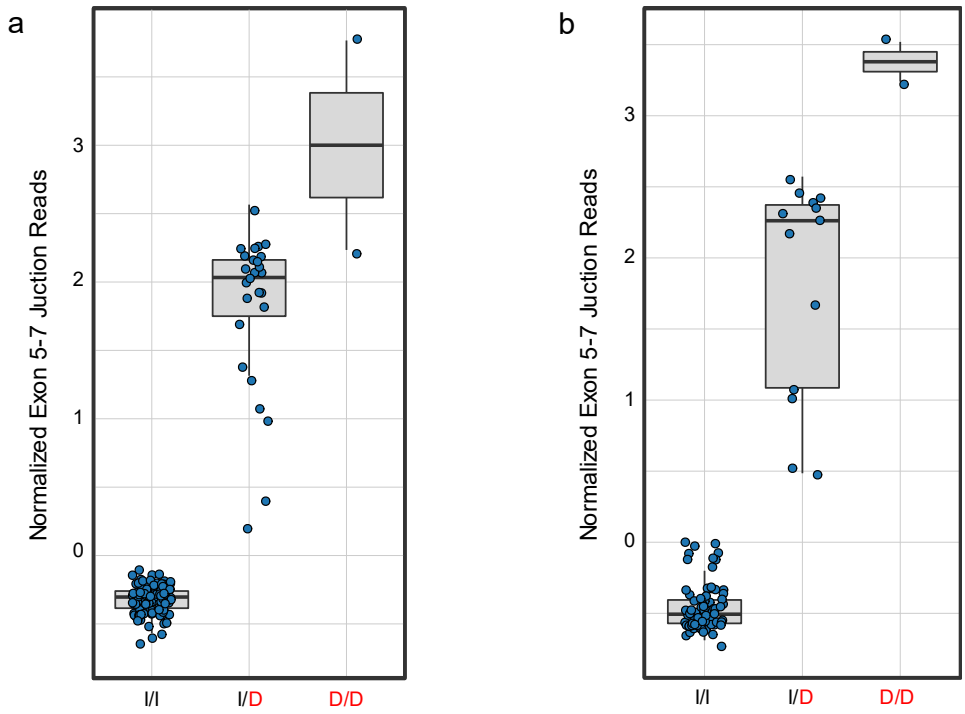

Fig. S5: Splicing QTLs for *CTRB2* and the 584 bp *CTRB2* genomic insertion/deletion variant in the LTG and GTEx histologically normal) pancreatic QTL datasets. The boxplots show normalized *CTRB2* junction reads from Leafcutter in GTEx samples (**a**, I/I n=148, I/D n=24, D/D n=2) and LTG pancreas samples (**b**, I/I n =75, I/D n=13, D/D n=2). I: insertion allele, D: deletion allele. The risk increasing allele (deletion) is indicated in red).

a

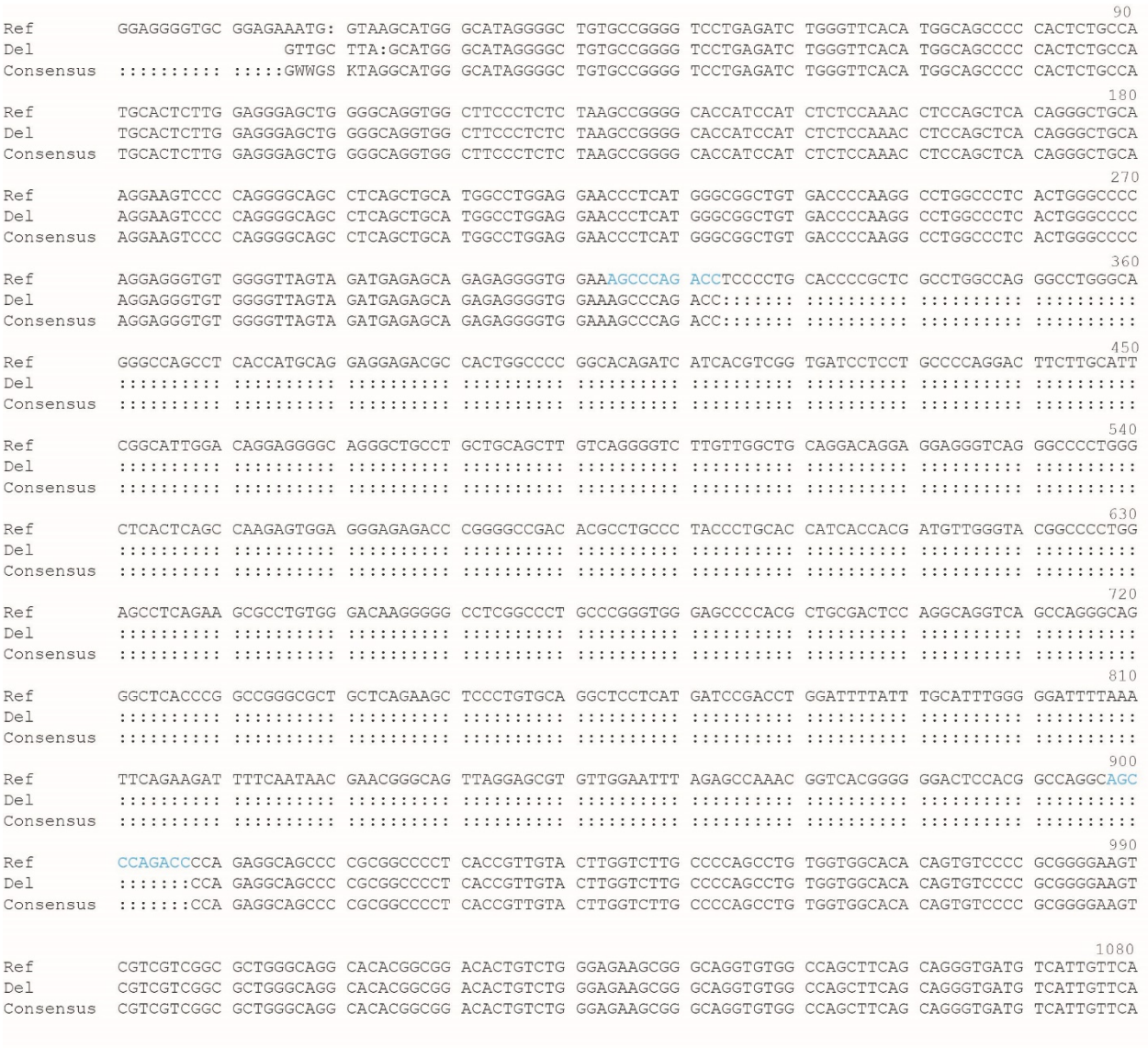

Fig. S6a: Nucleotide sequence of the *CTRB2* exon 6 deletion variant in a 1000G EUR sample. Nucleotide sequence of a 1000 Genomes subject identified to be homozygous for the *CTRB2* Ex6 deletion aligned to the reference sequence reveals a 584 bp deletion. The coordinates for the 584 bp deletion are ambiguous due to a repeated sequence at the beginning and end of the deleted sequence (marked in blue letters) and could range from Hg19: chr16:75,238,611-75,239,194 to chr16:75,238,621-75,239,204. The coordinates in the manuscript are listed as Hg19: chr16:75,238,616-75,239,199 +/- 5bp to reflect this ambiguity.

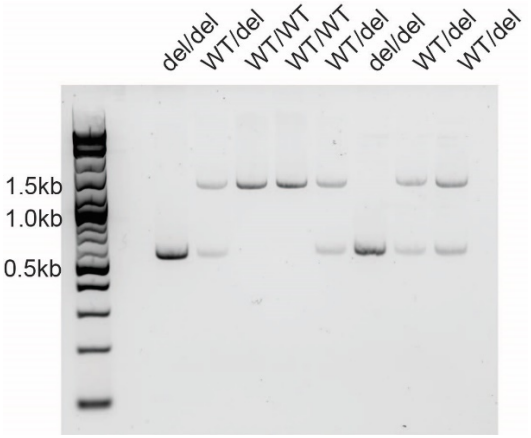

Fig S6b: Nucleotide sequence of the *CTRB2* exon 6 deletion variant in LTG eQTL samples. PCR was performed on genomic DNA from 8 human samples from the LTG eQTL dataset to amplify the *CTRB2* ex5-ex7 region (primers are listed in **Table S3**). The upper band represents the insertion allele with a 1112 bp product; the lower band represents the *CTRB2* exon 6 deletion allele with a 528 bp product (top figure). The two *CTRB2* exon 6 homozygous del/del samples are in lanes 1 and 6. Bands were excised and purified for sequencing (see nucleotide sequence in bottom part of Figure S6b below).

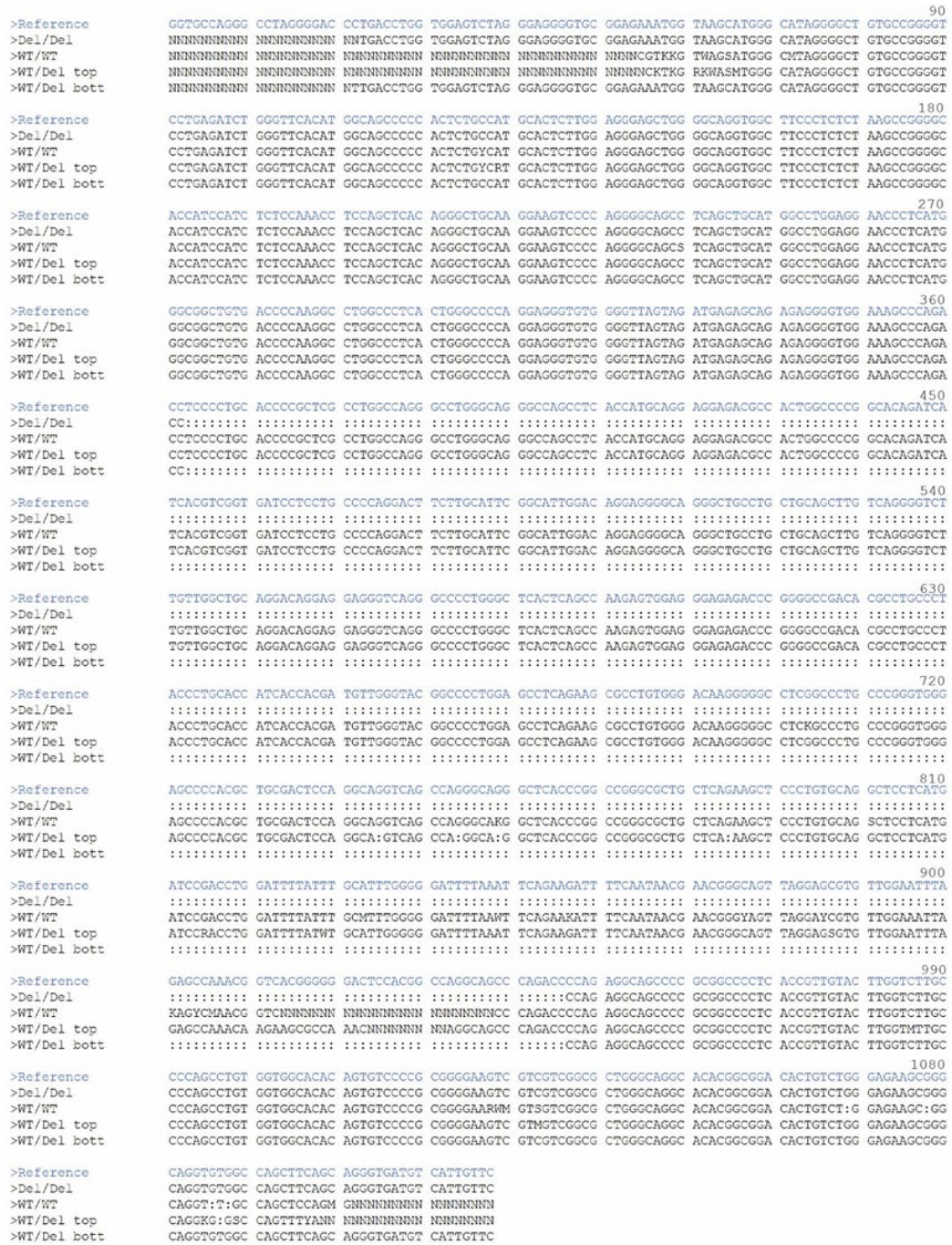

Fig S6b (continued): Bands from the gel above were excised and purified for sequencing. Top line: "Reference" represents the Human Genome Reference sequence (in blue text); "Del/Del" represents one of the homozygous *CTRB2* exon 6 deletion carriers; "WT/WT" represents a sample with no copies of the deletion; "WT/Del top" represents the top band from the PCR for a heterozygous WT/del carrier; "WT/Del bott" represents the lower band from a heterozygous WT/del carrier. Note the missing 584 bps corresponding to the *CTRB2* deletion allele in the Del/Del carrier and the bottom band from the WT/del carrier. "N" indicates missing sequencing from shorter reads, and non-A, T, C, G, N indicates ambiguous results from the sequencing.

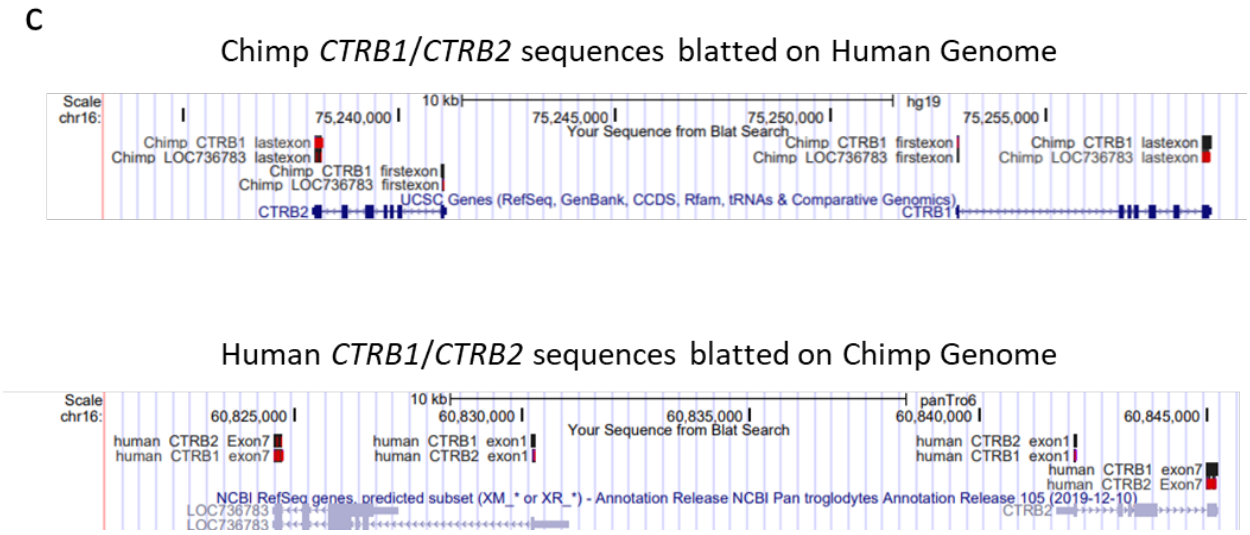

Fig. S6c: Ancestral and derived alleles for the *CTRB1*-exon1/*CTRB2*-exon1 inversion variant. Ancestral and derived alleles for the *CTRB1*/*CTRB2* inversion variant. The nucleotide sequence of human *CTRB1* and *CTRB2* genes on chr16q23.1 is identical apart from exons 1 and 7. To determine if the Human Reference allele is the inverted or non-inverted allele, we compared exons 1 and 7 between human and chimpanzee (*Pan troglodytes*) by BLAT in the UCSC Genome Browser. Note that while the chimp *CTRB1* gene (gene on right side in lower panel) is named *CTRB2*, it is listed as "Also known as *CTRB1*" in NCBI (https://www.ncbi.nlm.nih.gov/gene/736467). Similarly, while the chimp *CTRB2* gene (gene on left side of lower panel) is named LOC736783, it is listed as "Chymotrypsinogen B" and "*CTRB2*" in NCBI (https://www.ncbi.nlm.nih.gov/gene/736783) under General Protein Information. Red tick-marks within the aligned segments of the BLAT track represent substitutions with respect to the reference sequence. The **top panel** shows the BLAT alignment of exons 1 and 7 of chimp *CTRB2* (LOC736783) and *CTRB1* to the hg19 Human Reference genome in UCSC Genome Browser. Chimp *CTRB1* exon 1 aligns best with human *CTRB2* exon 1 and chimp *CTRB2* (LOC736783) exon 1 aligns best to human *CTRB1*. The last exon of chimp *CTRB1* and *CTRB2* (LOC736783) align best to their respective human genes. The **bottom panel** shows the BLAT alignment of human *CTRB2* and *CTRB1* exon 1 and exon 7 to the chimp reference genome. This shows that human *CTRB1* exon 1 aligns best with chimp *CTRB2* (LOC736783), and human *CTRB2* exon 1 aligns best with chimp *CTRB1* (listed as *CTRB2* in chimp browser but is *CTRB1* according to NCBI). Also note the large intron 1 of human *CTRB1* (top panel) and chimp *CTRB2* (LOC736783, bottom panel). The better alignment of human exon 1 to the opposite chimp chymotrypsinogen precursor gene, and the fact that the large intron 1 is part of *CTRB1* in the Human Reference genome but *CTRB2* (LOC736783) in the Chimp Reference genome, indicates that the human alternate allele is ancestral while the human reference allele is the inverted (and derived) allele.

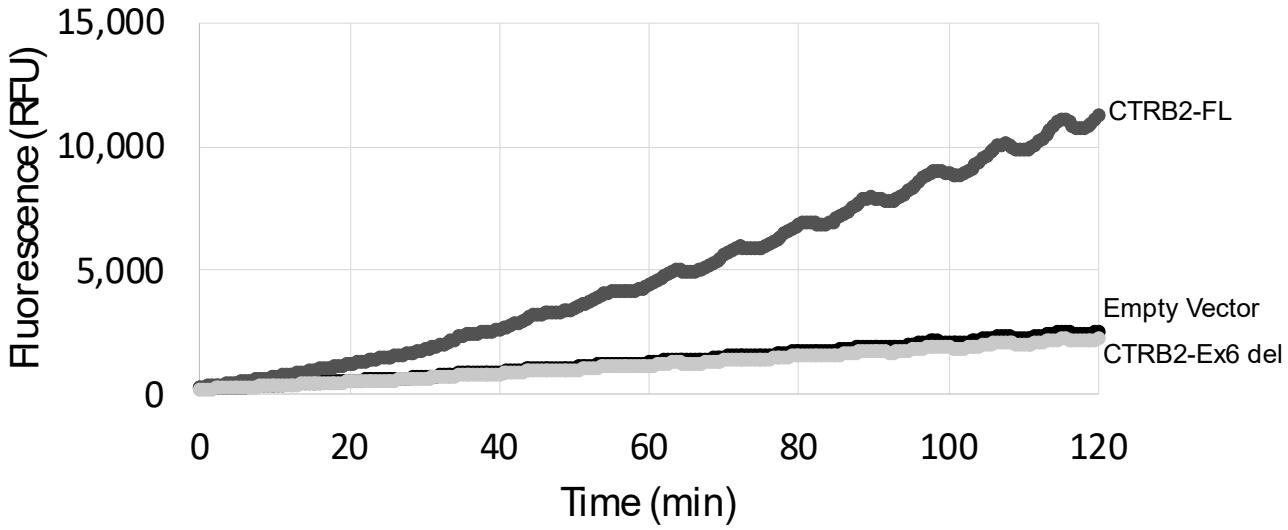

Fig. S7: Effect of the *CTRB2* exon 6 deletion on chymotrypsin reaction kinetics. A representative experiment showing fluorescence (RFU) as a measure of chymotrypsin activity on a synthetic fluorogenic substrate in HEK293T cell lysates after transient transfection with empty vector control (black), FLAG-tagged full-length *CTRB2* (CTRB2-FL, dark gray), and FLAG-tagged *CTRB2* with the exon-6 deletion (CTRB2-Ex6 del, light gray) plasmids. The linear change in fluorescence over a given time period was used to calculate activity, quantification of activity across triplicate experiments is shown in Fig. 3b.

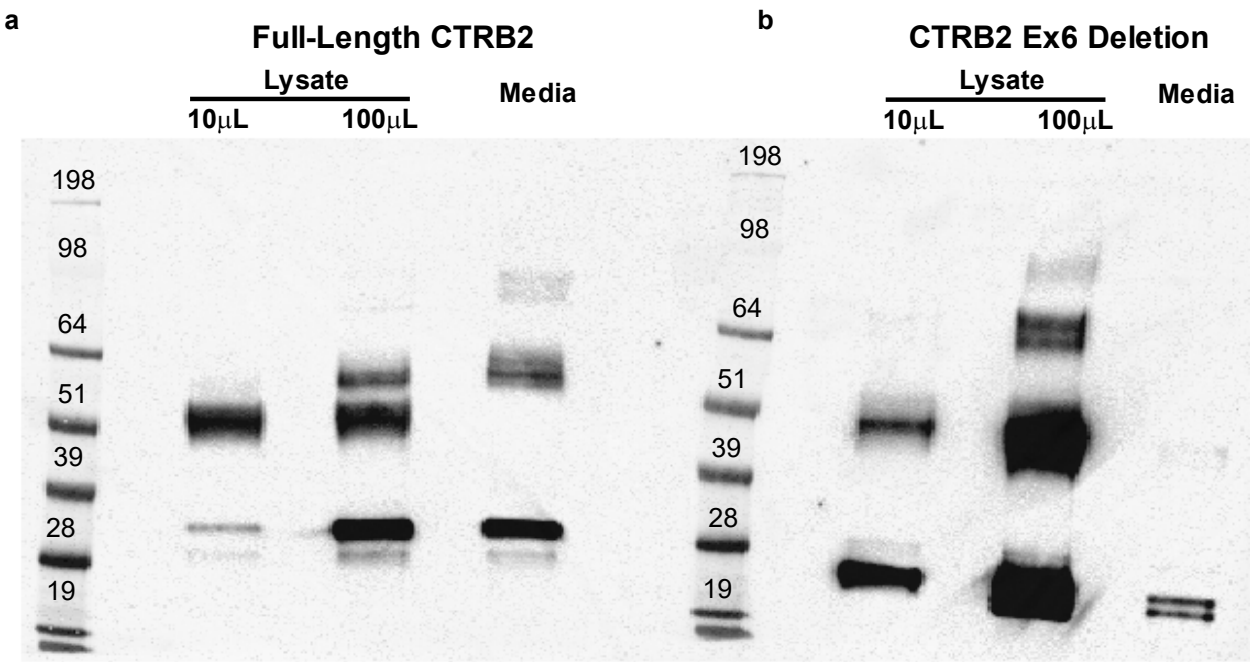

Fig. S8: Full western blot images related to Figure 3c for FLAG-tagged CTRB2 expression constructs. Western blot images corresponding to Fig. 3c-d showing: (**a)** full-length CTRB2 protein (~28 kDa), and (**b)** truncated CTRB2 protein (with exon-6 deletion, ~18 kDa). Immunoprecipitation (IP) for intracellular proteins was performed using either 10μL (first lane) or 100μL (second lane) of lysate; IP for secreted proteins was performed using conditioned media after concentration. Protein secretion was assessed by densitometric quantification of bands on the gel followed by comparing the amount of protein in the cell lysate (first or second lane, band of ~26 kDa) to the media (third lane, bands of ~16 kDa). The trend of less secretion for the truncated as compared to the full length CTRB2 proteins was similar for both lysate-media comparisons. Bands at ~26 kDa in (**a)**, and ~16 kDa in (**b)**, correspond to the full-length and truncated CTRB2 proteins with the 2 kDa signal peptide removed, respectively. Higher bands correspond to the heavy chain of IgG (~50 kDa) from the IP pull-down.

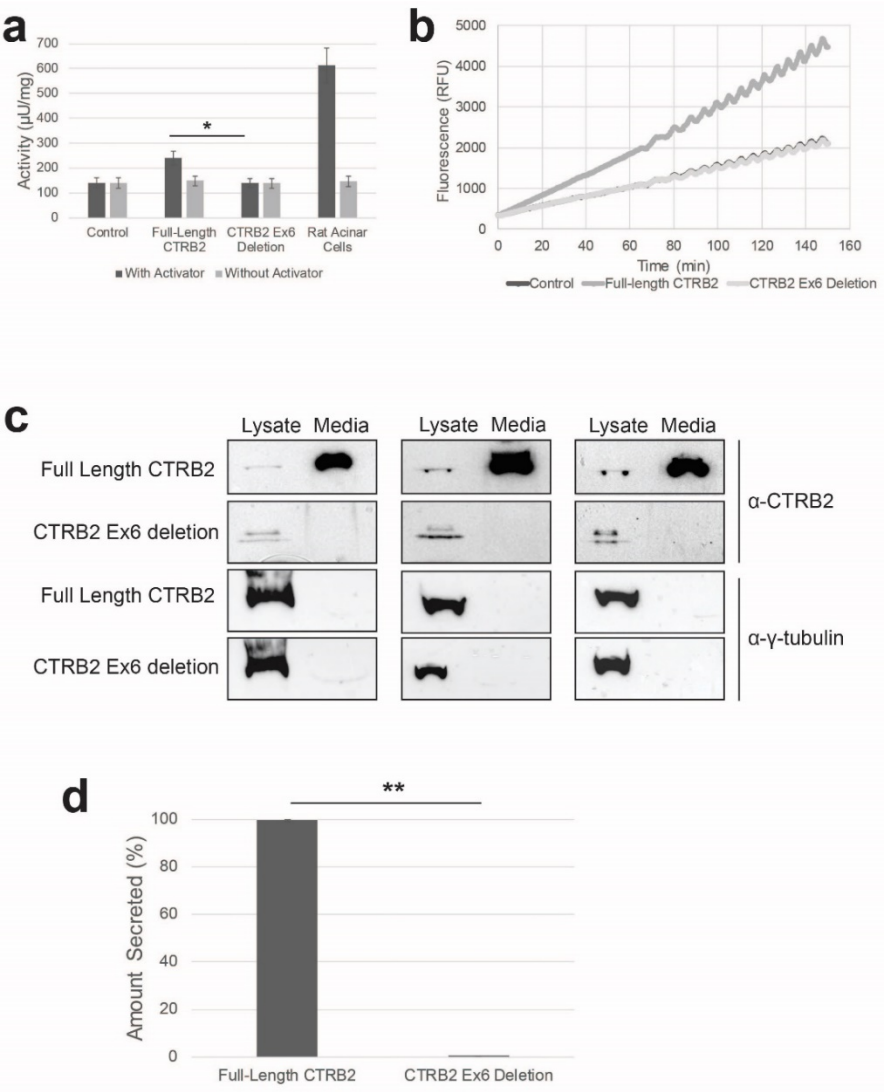

Fig S9. Activity and secretion for untagged CTRB2 expression constructs. (a) Chymotrypsin activity (µg/mg) calculated for the untagged full-length and Ex6 deletion constructs (n=4) and AR42J Rat acinar cells (n=3) (* denotes P <0.05) with (dark grey) and without (light grey) chymotrypsin activator (i.e. trypsin), error bars represent standard error of the mean (SEM). **(b)** Representative chymotrypsin activity time-course measured by fluorescence (RFU) for untagged full-length (dark grey), Ex6 deletion (light grey), and control (darkest grey). The activity of the full length CTRB2 protein is 1.75-fold that of the truncated protein (P=0.023), and the latter is at background levels. (**c**) Amount of CTRB2 protein in cell lysates and media assessed by western blot analysis in HEK293T cells transfected with plasmids expressing full-length and Ex6 deletion CTRB2 (three separate experiments). (**d**) Western blot images in (**c**) were taken using a ChemiDoc Touch Imaging System and bands quantified using ImageLab software. The amount of secreted proteins is summarized from triplicate experiments (** denotes P <0.001, error bars represent SEM). Almost all of the full length CTRB2 protein is detected in the media but almost none of the truncated protein (99.4% vs 0.1%, P=2.7x10^{-6}).

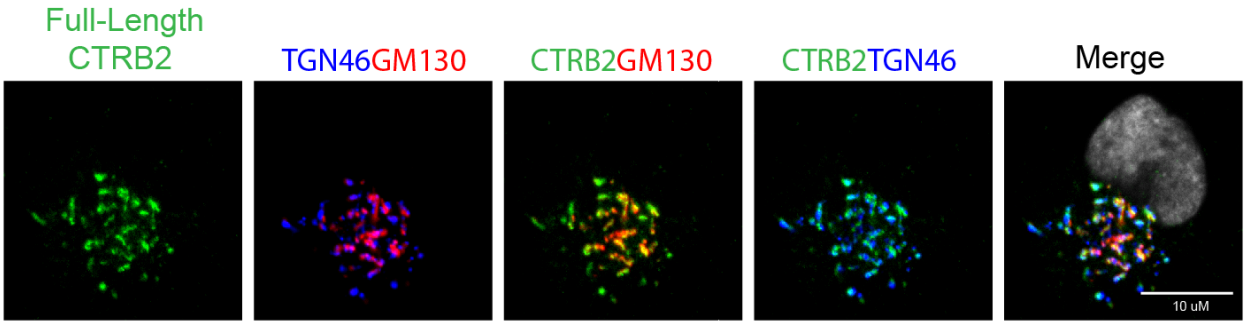

Fig. S10: The full length CTRB2 protein localizes to the Golgi. Normal pancreas-derived HPDE cells transiently expressing GFP-tagged full-length CTRB2 protein (green) immunostained with Golgi markers TGN46 (blue) and GM130 (red) show colocalization of the full length CTRB2 protein with the Golgi. Representative single xy images are shown taken using a spinning disk confocal microscope using and a 63x oil objective and a sCMOS camera. A z-stack was captured and processed with deconvolution software (Slidebook). DAPI stained nuclei shown in grey, scale bar = 10 μM.

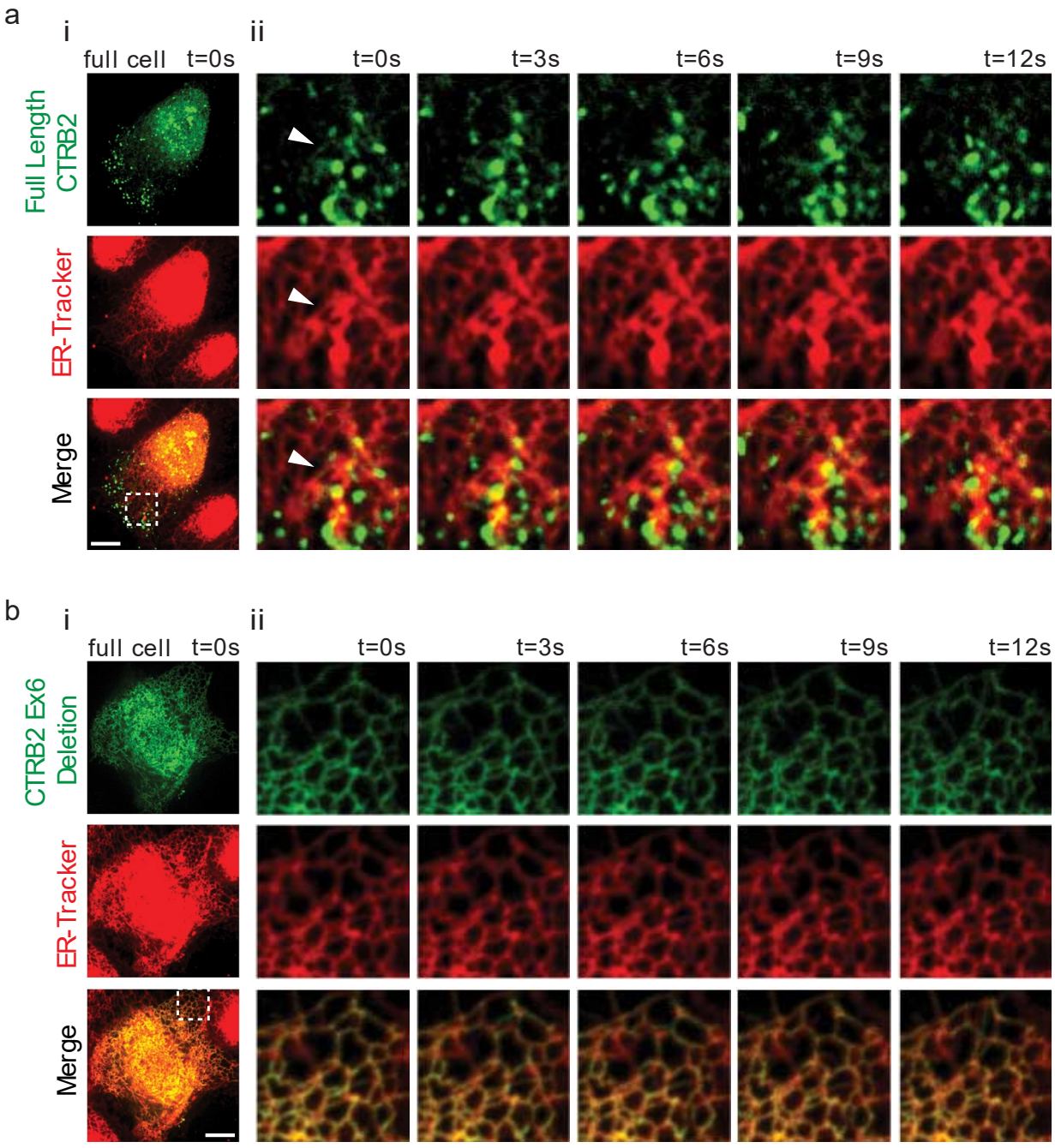

Fig. S11: Live imaging of CTRB2 protein localization in HPDE cells. HPDE pancreatic cells transiently expressing (**a**) GFP-tagged (green) full-length CTRB2 protein, and (**b**) truncated CTRB2 protein (with exon 6 deleted), for 48 hrs were incubated with ER-TrackerTM (red) for 15 min and followed by live spinning disk confocal microscopy with a 63x oil objective every 3 seconds at 37°C in 5% CO_2. (ii) shows an enlarged region (white square) from (i) for each panel. Scale bar = 10um. The full-length CTRB2 protein shows vesicular localization and only little ER overlap (white arrows), whereas the truncated CTRB2 protein completely colocalizes with the ER stain.

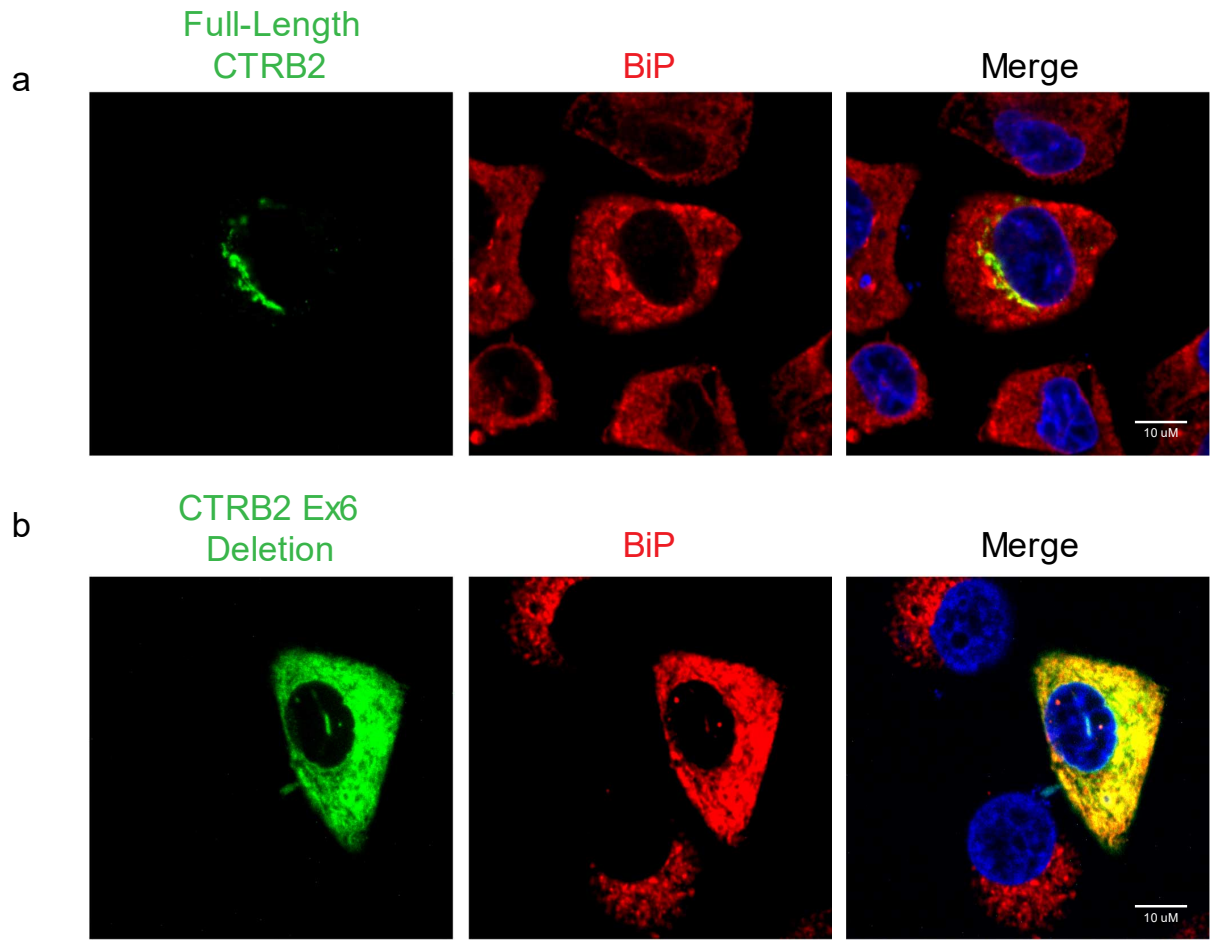

Fig. S12: The truncated CTRB2 protein induces Endoplasmic Reticulum (ER) stress. PANC-1 pancreatic cancer cells transiently expressing GFP-tagged (green) (**a**) full length CTRB2, and (**b**) truncated CTRB2 protein (exon 6 deleted), for 48 hrs were paraformaldehyde fixed and immuno-stained with a marker for ER-stress, BiP (red). Elevated levels of BiP (HSPA5, red) were seen in cells expressing the truncated CTRB2 protein (green in **b**), but not in cells expressing full length CTRB2 (green in **a**). Representative single xy images are shown taken using a spinning disk confocal microscope using a 63x oil objective and a sCMOS camera. A z-stack was captured and processed with Slidebook deconvolution software. Scale bar = 10 μM.

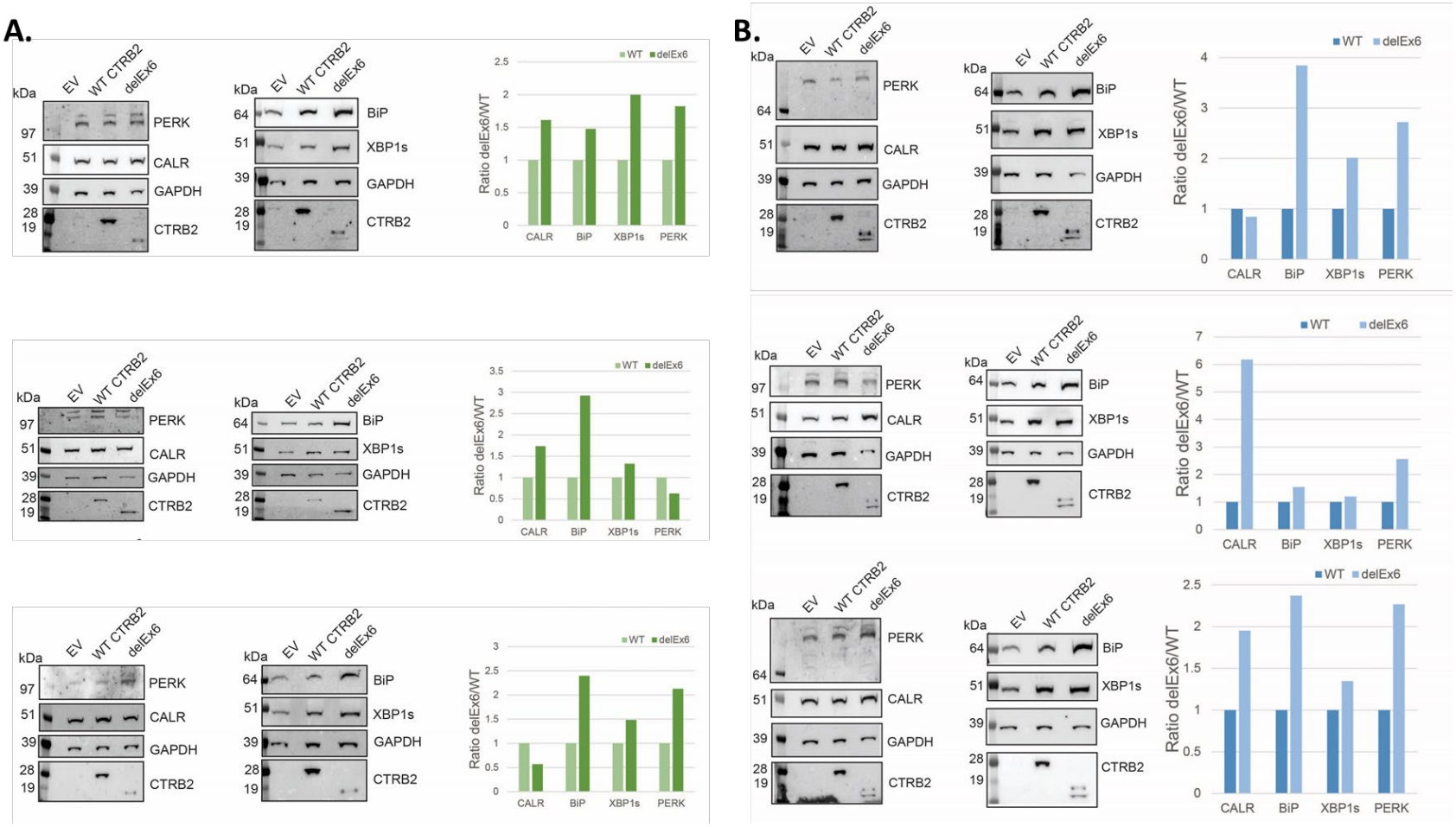

Figure S13: Overexpression of truncated (exon 6 deleted) CTRB2 leads to an increase in ER stress proteins. HEK293T (A) and PANC-1 (B) cells transfected with equal DNA amounts of empty vector, full length CTRB2, and truncated (exon 6 deleted) CTRB2 constructs for 72 hrs. reveal an increase in ER stress proteins in the delEx6 transfections as compared to full length CTRB2. Three biological replicates are shown.

Table S2: Minor allele frequency (MAF) for 16q23.1 GWAS SNPs, *CTRB2* Exon 6 insertion/deletion, and *CTRB2*/*CTRB1* inversion variants found in Europeans via PCR genotyping.

Genotype	EUR	CEU	FIN	GBR	IBS	TSI
rs72802365	0.084	0.089	0.057	0.104	0.130	0.047
Deletion	0.082	0.089	0.068	0.071	0.125	0.061
rs72802342	0.081	0.000	0.068	0.088	0.120	0.047
Inversion*	0.158	0.091	0.104	0.264	0.140	0.160

* The inversion is the minor (derived) allele

Table S3: Primers used for genotyping the *CTRB2* Exon 6 genomic insertion/deletion and the *CTRB2*/*CTRB1* genomic inversion variants.

Structural Variant Tested	Forward Primer (5'-3')	Reverse Primer (5'-3')	Annealing Temperature	Products
Deletion	TGACCTGGTGGAGTCTAGGG	TCAGCATTCTGACCGTGAAC	56.6°C	Reference: 1112bp product/Deletion: 528bp product
	TGACCTGGTGGAGTCTAGGG	TGGCGTCTCCTCCTGCATG	62°C	Reference: 409bp product/Deletion: no product
Inversion	CCTGCTCACTCTACCAAACC	TGGAATTTTGCAACAAGGCG	56°C	Reference (inversion): no product/ Ancestral (noninversion): 1974bp product
	GGGAACGTTTGAGCCTAGAG	TGGAATTTTGCAACAAGGCG	56°C	Reference (inversion): 1728bp/ Ancestral (noninversion): no product

Table S4: Imputation accuracy for two structural variants in the *CTRB2* and *CTRB1* gene region on chr16q23.1.

Genotype (584 bp indel in *CTRB2*)	No. of Individuals		% Accuracy
	Imputed	Genotyped	
Insertion/Insertion	72	70	97%
Insertion/Deletion	15	19	73%
Deletion/Deletion	9	7	78%
Genotype (*CTRB1*/*CTRB2* inversion variant)	**No. of Individuals**		**% Accuracy**
	Imputed	**Genotyped**	
Noninversion/Noninversion	55	52	95%
Noninversion/Inversion	26	29	88%
Inversion/Inversion	14	14	100%

A total of 96 DNA samples from PanScan III were genotyped by PCR for the *CTRB2* exon 6 deletion and *CTRB1/CTRB2* exon 1 inversion (see Methods).

Table S5: Association results for the 15 most likely functional variants at chr16q23.1 (LR < 1:1,000). A Meta-analysis of PanScan I, II, III and PanC4 GWAS data after imputation with the 1000G European reference panel was performed.

SNP	Location (bp)[^]	*P*-value	Likelihood Ratio (LR)*	Posterior Inclusion Probability (PIP)[%]	Odds Ratio OR (95% CI)	r^2 to best GWAS SNP[+]	r^2 to 584 bp *CTRB2* exon 6 deletion variant	EUR MAF[#]	Conditional Analysis	
									P-value conditioned on rs72802365	*P*-value conditioned on deletion
rs72802365	75,246,035	2.51E-17	1.0	0.13	1.36 (1.31-1.40)	1.00	0.69	0.083	1.000	0.007
rs72802352	75,240,883	3.72E-17	1.5	0.15	1.35 (1.31-1.40)	0.99	0.68	0.082	0.589	0.011
rs9936550	75,242,850	4.77E-17	1.9	0.06	1.35 (1.30-1.40)	1.00	0.69	0.082	0.974	0.013
rs111852127	75,249,170	5.11E-17	2.0	0.09	1.35 (1.30-1.40)	1.00	0.69	0.083	0.375	0.011
rs147630228	75,251,551	5.16E-17	2.0	0.09	1.35 (1.30-1.41)	1.00	0.69	0.083	0.628	0.011
rs184458383	75,249,552	5.33E-17	2.1	0.09	1.35 (1.30-1.40)	1.00	0.69	0.083	0.350	0.012
rs111869668	75,249,791	5.33E-17	2.1	0.09	1.35 (1.30-1.41)	1.00	0.69	0.083	0.350	0.012
rs72802395	75,286,484	7.30E-17	2.9	0.09	1.35 (1.30-1.40)	0.99	0.68	0.083	0.552	0.008
rs72802357	75,243,142	7.62E-17	3.0	0.18	1.35 (1.30-1.40)	1.00	0.69	0.082	0.619	0.018
rs72802342	75,234,872	7.13E-16	27.1	NA	1.34 (1.29-1.39)	0.84	0.81	0.081	0.615	0.105
rs371183658	75,234,273	7.78E-16	29.6	NA	1.33 (1.30-1.39)	0.82	0.82	0.081	0.602	0.107
rs55993634	75,236,763	1.17E-15	44.2	NA	1.32 (1.28-1.37)	0.78	0.79	0.089	0.618	0.123
rs72802391	75,273,829	4.46E-15	164.9	NA	1.35 (1.30-1.39)	0.85	0.57	0.075	0.923	0.073
rs8056814	75,252,327	7.92E-15	290.3	NA	1.31 (1.27-1.36)	0.94	0.64	0.088	0.068	0.132
rs68181471	75,252,820	1.14E-14	415.8	NA	1.31 (1.26-1.35)	0.93	0.63	0.091	0.192	0.121

[^] in human GRCh37/hg19 build coordinates

* Likelihood Ratio (LR) is given as 1:LR to the most significant SNP, rs72802365

[%] Posterior Inclusion Probability (PIP) for each SNP using SuSiE's Iterative Bayesian Stepwise Selection. A single credible set was observed at this locus. NA indicates SNPs not included in the credible set

[+] Linkage disequilibrium (r^2) to best GWAS SNP rs72802365

[#] Minor Allele Frequency in the 1000G EUR population

Table S6: Colocalization of expression QTLs (eQTLs) and pancreatic cancer GWAS signals.

Dataset	Tissue	Gene	Probability of Different Causal Variants	Probability of Same Causal Variant
GTEx	Skin - Sun Exposed	*TMEM170A*	1.0	<0.01
GTEx	Skin - Not Sun Exposed	*TMEM170A*	1.0	<0.01
LTG	Pancreas	*TMEM170A*	0.92	0.08
GTEx	Adipose -Visceral	*CFDP1*	1.0	<0.01
GTEx	Adipose - Subcutaneous	*CFDP1*	1.0	<0.01
GTEx	Skin - Sun Exposed	*CFDP1*	1.0	<0.01
GTEx	Whole Blood	*BCAR1*	0.69	0.31
GTEx	Pancreas	*CHST6*	0.87	0.13
TCGA	Pancreas	*CHST6*	0.90	0.08
GTEx	Pancreas	*CTRB1*	0.99	<0.01
GTEx	Pancreas	*CTRB2*	1.0	<0.01

Colocalization was performed using the coloc package in R.

Table S7: Colocalization of *CTRB2* splicing QTLs (sQTLs) and pancreatic cancer GWAS signals.

Dataset	*CTRB2* Splicing	Probability Different Causal Variants	Probability Same Causal Variant
LTG	exon 5-6	0.11	0.89
LTG	exon 6-7	0.54	0.46
LTG	exon 5-7	0.13	0.87
GTEx	exon 5-6	0.09	0.91
GTEx	exon 6-7	<0.01	0.94
GTEx	exon 5-7	0.13	0.87

Colocalization was performed using the coloc package in R.

Table S10: Differential expression of Endoplasmic Reticulum (ER) stress markers in pancreatic tissues from individuals with an increasing number of copies of the *CTRB2* Exon 6 deletion allele.

	GTEx		LTG	
Gene	**logβ**	**P-Value**	**logβ**	**P-Value**
HSPA5 (BiP)	0.335	0.022	-0.103	0.421
EIF2AK3 (PERK)	0.231	0.049	-0.183	0.264
ERN1 (IRE1-alpha)	0.230	0.066	-0.391	0.074
P4HB (PDI)	0.167	0.084	-0.246	0.300
CANX (Calnexin)	0.098	0.211	-0.151	0.118
DDIT3 (CHOP)	0.189	0.218	-0.065	0.681
ERO1A (ERO1-L-alpha)	0.087	0.255	-0.156	0.171
CALR	0.094	0.448	-0.140	0.159
ATF6	0.083	0.537	-0.355	0.049
XBP1	0.042	0.711	-0.375	0.108

Table S11: Colocalization of *CTRB2* splicing QTLs and diabetes GWAS signals.

Diabetes Dataset	sQTL Dataset	*CTRB2* Splicing	Probability Different Causal Variant	Probability Same Causal Variant
T1D - Barrett	GTEx	exon 5-6	<0.01	1.0
T1D - Barrett	GTEx	exon 6-7	<0.01	1.0
T1D - Barrett	GTEx	exon 5-7	<0.01	1.0
T1D - Barrett	LTG	exon 5-6	<0.01	1.0
T1D - Barrett	LTG	exon 6-7	0.36	0.64
T1D - Barrett	LTG	exon 5-7	<0.01	1.0
T2D	GTEx	exon 5-6	0.05	0.96
T2D	GTEx	exon 6-7	0.03	0.98
T2D	GTEx	exon 5-7	0.08	0.92
T2D	LTG	exon 5-6	0.34	0.66
T2D	LTG	exon 6-7	0.51	0.49
T2D	LTG	exon 5-7	0.50	0.50

Colocalization was performed using the coloc package in R.

Supplemental References

1. Robinson, James T.; Thorvaldsdottir, Helga; Winckler, Wendy; Guttman, Mitchell; Lander, Eric S.; Getz, Gad; Mesirov, J.P. (2011). Integrative genomics viewer. Nat. Biotechnol. *29*, 24–26.

Funding

The work conducted at NCI was supported by the Intramural Research Program (IRP) of the Division of Cancer Epidemiology and Genetics, National Cancer Institute, US National Institutes of Health (NIH).

The Melbourne Collaborative Cohort Study cohort recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was further augmented by Australian National Health and Medical Research Council grants 209057, 396414 and 1074383 and by infrastructure provided by Cancer Council Victoria. Cases and their vital status were ascertained through the Victorian Cancer Registry and the Australian Institute of Health and Welfare, including the National Death Index and the Australian Cancer Database.

The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201600018C, HHSN268201600001C, HHSN268201600002C, HHSN268201600003C, and HHSN268201600004C. The authors thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A full listing of WHI investigators can be found at: http://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Long%20List.pdf"

Cancer incidence data for CLUE were provided by the Maryland Cancer Registry, Center for Cancer Surveillance and Control, Department of Health and Mental Hygiene, 201 W. Preston Street, Room 400, Baltimore, MD 21201, http://phpa.dhmh.maryland.gov/cancer, 410-767-

4055. We acknowledge the State of Maryland, the Maryland Cigarette Restitution Fund, and the National Program of Cancer Registries of the Centers for Disease Control and Prevention for the funds that support the collection and availability of the cancer registry data." We thank all the CLUE participants.

The NYU study was funded by NIH R01 CA098661, UM1 CA182934 and center grants P30 CA016087 and P30 ES000260.

The Physicians' Health Study was supported by research grants CA-097193, CA-34944, CA40360, HL-26490, and HL-34595 from the National Institutes of Health, Bethesda, MD USA.

The Women's Health Study was supported by research grants CA-047988, HL-043851, HL080467, and HL-099355 from the National Institutes of Health, Bethesda, MD USA.

Health Professionals Follow-up Study is supported by NIH grant UM1 CA167552 from the National Cancer Institute, Bethesda, MD USA.

Nurses' Health Study is supported by NIH grants UM1 CA186107, and R01 CA49449 from the National Cancer Institute, Bethesda, MD USA.

The PANKRAS II Study in Spain was supported by research grants from Instituto de Salud Carlos III-FEDER, Spain: Fondo de Investigaciones Sanitarias (FIS) ((#PI95/0017, #PI12/00815, #PI13/00082 and #PI15/01573), Red Temática de Investigación Cooperativa en Cáncer (#RD12/0036/0050), and CIBER de Epidemiología (CIBERESP); Ministerio de Ciencia y Tecnología (CICYT SAF 2000-0097); Generalitat de Catalunya (CIRIT - SGR), Spain.

The IARC/Central Europe study was supported by a grant from the US National Cancer Institute at the National Institutes of Health (R03 CA123546-02) and grants from the Ministry of Health of the Czech Republic (NR 9029-4/2006, NR9422-3, NR9998-3, MH CZDRO-MMCI 00209805).

The National Familial Pancreas Tumor Registry at Johns Hopkins University was supported by the NCI Grants P50CA062924 and R01CA97075. Additional support was provided by the Lustgarten Foundation, Susan Wojcicki and Dennis Troper and the Sol Goldman Pancreas Cancer Research Center. The PANC4 GWAS was supported by RO1 CA154823 and federal funds from the National Cancer Institute (NCI), US National Institutes of Health (NIH) under contract number HHSN261200800001E.

The Mayo Clinic Biospecimen Resource for Pancreas Research study is supported by the Mayo Clinic SPORE in Pancreatic Cancer (P50 CA102701).

Funding at Memorial Sloan Kettering was supported by the National Cancer Institute of the National Institutes of Health grant number P30 CA008748.

Research reported in this publication was supported in part by the National Cancer Institute of the National Institutes of Health under Award Numbers U10 CA37429 (CD Blanke), and UM1 CA182883 (CM Tangen/IM Thompson) for SELECT.

The PACIFIC Study was supported by RO1CA102765, Kaiser Permanente and Group Health Cooperative.

The Queensland Pancreatic Cancer Study was supported by a grant from the National Health and Medical Research Council of Australia (NHMRC) (Grant number 442302). RE Neale is supported by a NHMRC Senior Research Fellowship (#1060183).

The UCSF pancreas study was supported by NIH-NCI grants (R01CA1009767, R01CA109767-S1 and R0CA059706) and the Joan Rombauer Pancreatic Cancer Fund. Collection of cancer incidence data was supported by the California Department of Public Health as part of the statewide cancer reporting program; the NCI's SEER Program under contract HSN261201000140C awarded to CPIC; and the CDC's National Program of Cancer Registries, under agreement #U58DP003862-01 awarded to the California Department of Public Health.

The Yale (CT) pancreas cancer study is supported by National Cancer Institute at the U.S. NIH, grant 5R01CA098870. The cooperation of 30 Connecticut hospitals, including Stamford Hospital, in allowing patient access, is gratefully acknowledged. The Connecticut Pancreas Cancer Study was approved by the State of Connecticut Department of Public Health Human Investigation Committee. Certain data used in that study were obtained from the Connecticut Tumor Registry in the Connecticut Department of Public Health. The authors assume full responsibility for analyses and interpretation of these data.

The EPIC-Norfolk study (DOI 10.22025/2019.10.105.00004) has received funding from the Medical Research Council (MR/N003284/1 and MC-UU_12015/1) and Cancer Research UK (C864/A14136). We are grateful to all the participants who have been part of the project and to the many members of the study teams at the University of Cambridge who have enabled this research.