



# Functional and association analysis of an Amerindian-derived population-specific p.(Thr280Met) variant in *RBPJL*, a component of the PTF1 complex

Anup K. Nair<sup>1</sup> · Jeff R. Sutherland<sup>1</sup> · Michael Traurig<sup>1</sup> · Paolo Piaggi<sup>1</sup> · Peng Chen<sup>1</sup> · Sayuko Kobes<sup>1</sup> · Robert L. Hanson<sup>1</sup> · Clifton Bogardus<sup>1</sup> · Leslie J. Baier<sup>1</sup>

Received: 28 July 2017 / Revised: 13 October 2017 / Accepted: 18 November 2017 / Published online: 4 January 2018

© The Author(s) 2018. This article is published with open access

## Abstract

PTF1 complex is critical for pancreatic development and maintenance of adult exocrine pancreas. As a part of our ongoing studies to identify genetic variation that contributes to type 2 diabetes (T2D) in American Indians, we analyzed variation in genes that form this complex, namely *PTF1A*, *RBPJ*, and its paralogue *RBPJL*. A c.839C>T (p.(Thr280Met)) variant (rs200998587:C>T, risk allele frequency = 0.03) in *RBPJL*, identified only in Amerindian-derived populations, associated with T2D (OR = 1.60[1.21–2.13] per Met allele,  $P = 0.001$ ) and age of diabetes onset (HR = 1.40[1.14–1.72],  $P = 0.001$ ). Knockdown of *Rbpjl* in mouse pancreatic acinar cells resulted in a significant decrease in the mRNA expression of genes encoding exocrine enzymes including *Ctrb*. *CTRB1/2* is an established T2D locus where the protective allele associates with increased GLP-1-stimulated insulin secretion and higher expression of *CTRB1/2*. In vitro studies show that cells expressing the Met280 allele had lower RBPJL protein levels than cells expressing the Thr280 allele, despite having comparable levels of RNA, suggesting that the Met280 RBPJL is less stable. Additionally, luciferase assays in HEK293 cells which examined two different RBPJL responsive promoters, including the promoter for *CTRB1*, also identified reduced transactivation by the Met280 RBPJL. Similarly, overexpression of both Met280 and Thr280 RBPJL in mouse pancreatic acinar cells identified a significant impairment in the expression of *Cel* when transactivated by the Met280 RBPJL. In summary, we identified a functional, Amerindian-derived population-specific c.839C>T (p.(Thr280Met)) variant in the pancreas specific *RBPJL* that may modify T2D risk by regulating exocrine enzyme expression.

## Introduction

The trimeric pancreas transcription factor 1 (PTF1) complex consists of a dimer of PTF1A and a class A basic helix loop helix protein bound to either RBPJ or its paralogue RBPJL [1]. The expression of both *PTF1A* and *RBPJL* is mostly

limited to the pancreas (as observed in the GTEx Portal,[2]). In adults, RBPJL is expressed in both exocrine (acinar) and endocrine (islets) cells of the pancreas, whereas PTF1A expression is restricted to acinar cells [1, 3]. In acinar cells, RBPJL and PTF1A regulate the expression of digestive enzymes and maintain the differentiated state of adult exocrine pancreas whereas the role of RBPJL in islets is not defined [1, 4–6]. In contrast, RBPJ is widely expressed [2] and can either bind to PTF1A or the Notch IntraCellular Domain (NICD), which mediates Notch signaling. The binding of RBPJ to either PTF1A or NICD is mutually exclusive, and this process plays a key role in pancreatic development [1, 6].

Variants in *PTF1A* are a known cause of pancreatic and cerebellar agenesis [7]. A hypomorphic variant in *PTF1A* has also been identified in subjects with neonatal diabetes and exocrine pancreatic insufficiency without neurodevelopmental delay [8]. Other genes that are related to neonatal diabetes have been shown to contain variation that

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1038/s41431-017-0062-6>) contains supplementary material, which is available to authorized users.

---

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

---

✉ Leslie J. Baier  
lbaier@phx.niddk.nih.gov

<sup>1</sup> Phoenix Epidemiology and Clinical Research Branch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, 445N 5th Street, Phoenix, AZ 85004, USA

associates with either maturity onset diabetes of the young (MODY) and/or adult-onset T2D [9–11]. The role of exocrine pancreas in the pathogenesis of T2D has not been extensively studied. However, studies have shown a relationship between a decreased pancreatic exocrine function and T2D which is mediated by an incretin effect [12, 13]. Based on these observations, we sought to determine whether variations in *RBPJL*, *RBPJ* or *PTF1A*, that encode the protein components of the PTF1 complex and is important for the maintenance of adult exocrine pancreas, may affect risk of T2D in American Indians.

## Materials and methods

### Subjects with genotypic data

Subjects for the association analyses were 7659 American Indians from a community in AZ, USA, who had participated in a longitudinal study of T2D, where T2D status was determined (1997 ADA criteria) at outpatient biennial exams using a 75 g OGTT. Among this population-based sample, 3719 individuals were full-heritage Pima Indians (57% female, 46.5% had T2D, mean age =  $40.2 \pm 16.8$  years) and the remaining 3940 individuals had some degree of American Indian heritage (on average a person's heritage was 2/3rd American Indian, 54.7% female, 21.4% had T2D, mean age =  $28.3 \pm 14.1$  years). All participants gave written informed consent. The study was approved by the institutional review board of the National Institute of Diabetes and Digestive and Kidney Diseases.

Genome-wide genotypic data for the 7659 American Indians had previously been generated using a custom Axiom array by Affymetrix, Santa Clara, CA. The custom array had been designed to tag ( $r^2 \geq 0.85$  within a window of 300 kb) all common non-coding variation with a minor allele frequency (MAF)  $\geq 0.05$  and missense variation with a MAF  $\geq 0.01$  detected in whole-genome sequence data of 266 full-heritage Pima Indians from different nuclear families.

### Association analysis

Gene regions analyzed in this study, defined as 25 kb upstream and downstream of the gene, were *PTF1A* (chr10:23456460–23508181, RefSeq: NM\_178161.2), *RBPJ* (chr4:26296332–26461752, RefSeq: NM\_005349.3) and *RBPJL* (chr20:43910483–43971464, RefSeq: NM\_014276.3). All positions throughout are given with respect to human genome build 37 (hg19). Using whole genome sequencing data from 335 American Indians, we determined that genotypic data from our array provided coverage for all common non-coding variation (MAF  $\geq$

0.05) and missense variation (MAF  $\geq 0.01$ ) for these regions (assuming tags have a  $r^2 \geq 0.8$ ) to be as follows: *PTF1A*, 26 out of 30 variants captured by 10 tags; *RBPJ*, 207 out of 212 variants captured by 17 tags; and *RBPJL*, 162 out of 170 variants captured by 17 tags. Variants analyzed in the current study passed genotyping quality control metrics (call rate  $\geq 90\%$ , discrepancy rate  $\leq 2$  pairs among 100 blind duplicates, and lack of deviation from Hardy-Weinberg equilibrium [ $P$ -value  $> 10^{-4}$ ]) and has been submitted to ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>).

Statistical analysis was performed using SAS (version 9.2, SAS institute, Cary, NC, USA). T2D association was analyzed by logistic regression using an additive model. Generalized estimating equations were used to account for sibship. The OR and  $P$ -values were adjusted for age, sex, birth-year and first five principal components (to account for genetic admixture) derived using the genotypic data from the custom Axiom array as described previously [14].  $P$ -values were further controlled for inflation by a genomic control method; the mean  $\chi^2$  statistic over all markers on the Axiom array was taken as the inflation factor ( $\lambda = 1.38$ ). A variant was considered to be associated with T2D if the  $P$ -value was  $\leq 0.001$  (experiment-wide Bonferroni-correction accounting for 44 variants). Association with age of diabetes onset was assessed using a Cox proportional hazards model with sex, birth-year and first five principal components as covariates.

### Subjects used to assess replication of rs200998587: C>T variant in RBPJL with T2D

#### Urban American Indians (N = 2993)

Cross-sectional data from four groups of urban-dwelling American Indians living in or near Phoenix, AZ was used to assess replication of the rs200998587:C>T variant with T2D. These individuals were  $\geq 50\%$  American Indian and were characterized for T2D using the ADA 2010 criteria. These subjects were recruited under a protocol which does not allow publication of specific tribal identity; thus, this data set is described as urban-dwelling American Indians I–IV.

#### SIGMA T2D Consortium exome chip data

SIGMA T2D Consortium exome sequencing data were also used to assess the T2D association of the rs200998587:C>T variant using the publicly available T2D Knowledge Portal [15].

### Expression vectors, Promoter constructs, and small interfering RNA (siRNA)

Expression vectors containing full-length untagged *RBPJL*, myc-DDK tagged *RBPJL*, and *PTF1A* were obtained from

Origene (Rockville, MD, USA). Site-directed mutagenesis was used to create the 280Met (c.839T) substitution in *RBPJL* using the QuikChange Lightning site-directed mutagenesis kit (Agilent technologies, Santa Clara, CA, USA). *RBPJL* responsive promoter–reporter constructs were created by cloning the promoter regions ( $\approx -1500$  bp to  $+50$  bp from the transcription start sites) of two known *RBPJL* responsive genes, *CTRB1* and *CPA1*, into the pGL3-basic vector upstream of the luciferase gene (pGL3B\_CPA1 and pGL3B\_CTRB1). In-vivo binding of Ptf1a and Rbpjl to *Cpa1* and *Ctrb* promoters in rat pancreas has been previously reported [1]. The promoter fragments were confirmed to contain an E-box (PTF1A binding site) and a TC-box (*RBPJL* binding site). All constructs were sequenced to confirm the integrity of the insert. Ambion silencer-select siRNA for *Rbpjl* and control siRNA were from Thermo Fisher Scientific (Waltham, MA, USA).

### Cell culture and transfections

HEK-293 and a mouse pancreatic cell line (266-6) derived from a pancreatic acinar cell tumor and known to express digestive enzymes were obtained from American Type Culture Collection and maintained per their instructions. For *Rbpjl* knockdown, 266-6 cells were grown in six-well plates and transfected with equal amounts (150 pm) of *Rbpjl* siRNA or control siRNA by reverse transfection using RNAimax (Thermo Fisher). Cells were harvested 72 h post transfection. All other transfections were done using Lipofectamine 3000 (Thermo Fisher) per manufacturer's instruction. For the *RBPJL* responsive promoter–reporter assays, HEK-293 cells were seeded in 12-well plates at an appropriate density so the cells would reach 70–80% confluency by the next day. For transfections, 600ng of the *CPA1* or *CTRB1* promoter–reporter construct was co-transfected with either 200ng of pCMV6\_*RBPJL*\_myc-DDK or pCMV6\_*PTF1A* or both. Equal amounts of DNA were used for all conditions by using pCMV6 empty vector wherever necessary. Co-transfection with the pGL4.74 Renilla luciferase vector (50 ng) was used to monitor transfection efficiency. For protein expression analysis, HEK-293 cells grown in six-well plates were transfected with 1  $\mu$ g of pCMV6 expression vector carrying either myc-DDK tagged Thr280 *RBPJL* or myc-DDK tagged Met280 *RBPJL*. For assessing the effect of *RBPJL* p.(Thr280Met) variant on *Cel* expression, a known *Rbpjl* target gene, 266-6 cells grown in six-well plates were transfected with 1  $\mu$ g of plasmid containing either untagged Thr280 *RBPJL* or untagged Met280 *RBPJL* along with *PTF1A*. Similar PTF1 complex binding sites were seen during bioinformatic analysis of promoters ( $+1$  to  $-300$  bp from putative transcription start sites) from human *CEL* gene and mouse *Cel* gene.

### Luciferase assay

Luciferase assays were conducted using the dual-luciferase assay kit (Promega, Madison, WI, USA) per manufacturer's instructions. Cells were harvested 48 h after transfection and were freeze-thawed once before the assay. Two different *RBPJL* responsive promoter–reporter constructs were used to assess the activity of Thr280 and Met280 *RBPJL*. Transfections were done in triplicate and the assays were repeated four times. Luciferase activity for each well (transfection) was read twice and average activity was recorded. For statistical analysis, fold increase in relative luciferase activity between PTF1A alone and PTF1A in combination with either Thr280 or Met280 *RBPJL* was calculated for each sample. Average fold increase from the triplicates in each independent assay was calculated and the difference in activity was compared using an unpaired *t*-test ( $N = 4$ ).

### Real-time polymerase chain reaction (PCR)

Total RNA was isolated from harvested cells using RNA easy plus mini kit (Qiagen, Germantown, MD, USA) and quantified by nanodrop (Thermo Fisher). Total RNA (2  $\mu$ g) was converted to cDNA using the high capacity RNA to cDNA kit (Thermo Fisher). SYBR-green chemistry was used for real-time PCR using the PowerUp SYBR-Green master mix (Thermo Fisher). Primers were designed either in exon–intron junctions or in different exons separated by large introns to avoid amplification of residual genomic DNA. All primer pairs were validated by PCR and melt-curve analysis to ensure a single product and absence of primer dimer. A  $C_t$  (threshold cycle) validation was done to ensure similar amplification efficiency of the target primers and the endogenous control primers. Real-time PCR assays were done in triplicate and data were normalized to the expression of *Tbp* (for experiments in 266-6 cells) and *RPL10* (for experiments in HEK-293 cells). *B2m* and *Gapdh* were used as additional controls. Results were compared using an unpaired *t*-test.

### Western blot

Transfected cells were washed once with  $1 \times$  PBS, harvested using trypsin-EDTA and divided into two equal aliquots. One aliquot was used for RNA isolation and subsequent real-time PCR. The other aliquot was washed three times with ice cold PBS and lysed with RIPA lysis buffer supplemented with  $1 \times$  Halt Protease inhibitor cocktail (Thermo Fisher) and 5 mM EDTA. After 30-min incubation in ice, the lysates were sonicated at low amplitude in ice to reduce viscosity, and cellular debris was removed by high-speed centrifugation at 4 °C. Total

**Table 1** Association of missense variants in *RBPJL* with type 2 diabetes in American Indians

Variant information		Type 2 diabetes association					
Variant ID	Alleles <sup>a</sup>	Risk allele, freq <sup>b</sup>	Variant <sup>c</sup>	Full-Pima ( <i>n</i> = 3719)	Non-Full-Pima ( <i>n</i> = 3940)	Combined ( <i>N</i> = 7659)	<i>P</i> -value <sup>de</sup>
				OR (95%CI) <sup>d</sup>	OR (95%CI) <sup>d</sup>	OR (95%CI) <sup>d</sup>	<i>P</i> -value <sup>de</sup>
rs20098587 <sup>f</sup>	C/T	T, 0.03	c.839C>T (p.(Thr280Met))	1.49 (1.02–2.17)	1.78 (1.22–2.60)	1.60 (1.21–2.13)	<b>0.001</b>
chr20:43945378 <sup>g</sup>	G/A	A, 0.015	c.1333G>A (p.(Asp445Asn))	1.08 (0.69–1.69)	2.74 (1.46–5.13)	1.43 (0.97–2.12)	0.07

<sup>a</sup> Alleles shown are reference allele/variant allele

<sup>b</sup> Risk Allele and frequency of the risk allele in combined samples

<sup>c</sup> Based on *RBPJL* refseq: NM\_014276.3

<sup>d</sup> OR and *P*-values were adjusted for age, sex, birth-year and first five principal components from a Pima Indian specific GWAS. OR are given per copy of the risk allele

<sup>e</sup> *P*-value shown are after controlling for inflation by genomic control method

<sup>f</sup> 7227 subjects (94.3%) were successfully genotyped for the Thr280Met variant. Among these, 439 subjects were carriers of the risk allele. The prevalence of T2D among carriers of the Met allele was 41% (180/439) while among non-carriers was 32.8% (2232/6788)

<sup>g</sup> Position is based on human genome build 37 (hg19)

protein was quantified using Pierce BCA protein assay kit (Thermo Fisher) and 25 µg was separated using SDS–PAGE and transferred to a nitrocellulose membrane for protein detection. Mouse anti-DDK antibody to detect myc-DDK tagged RBPJL was from Origene. Monoclonal mouse anti-β-actin antibody was from Sigma-Aldrich (St. Louis, MO, USA). HRP-conjugated secondary antibodies were from Origene and Sigma-Aldrich. Blots were developed by chemiluminescence using Clarity ECL western blotting substrate (Biorad, Hercules, CA, USA) and ChemiDoc touch imaging system (Biorad). Blots were quantified using Image Lab software (version 5.2.1, Biorad).

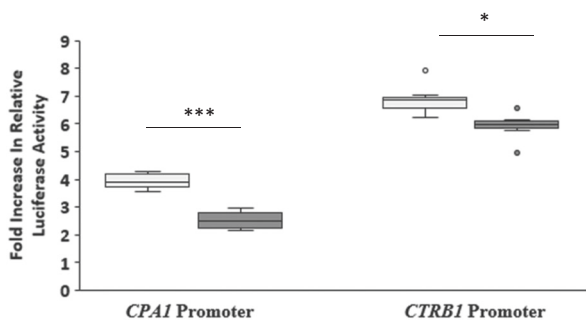
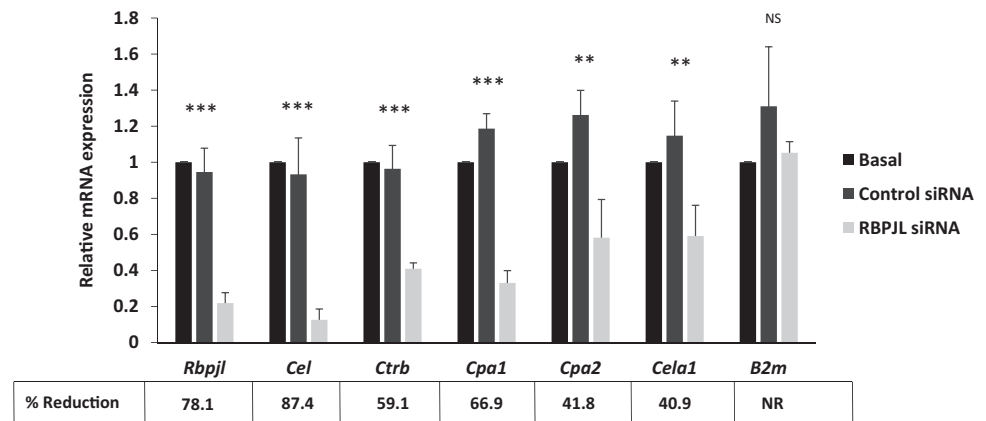
## Results

### Association analysis of variants in PTF1 complex genes with T2D

Forty-four tag variants that pass the genotyping quality control metrics and capture non-coding (MAF ≥ 0.05) and coding variation (MAF ≥ 0.01) across the three genes that code for the PTF1 complex, namely *RBPJL*, *RBPJ*, and *PTF1A*, and their associations with T2D in 7659 subjects are shown in electronic supplementary material (ESM) Tables 1–3. Among these variants, only rs76474829:C>G (hg19 chr20:g.43914005C>G) upstream of *RBPJL* achieved a significant association with T2D after an experiment-wide Bonferroni correction (OR = 1.57[1.19–2.07], *P*-value = 0.001, ESM Table 1). Since rs76474829:C>G tags ( $r^2 = 0.85$ ) a missense variant, rs200998587:C>T (hg19 chr20:g.43942756C>T, c.839C>T, p.(Thr280Met)) in *RBPJL*, the c.839C>T (p.(Thr280Met)) variant was directly genotyped to test for an association with T2D. The c.839C>T (p.(Thr280Met)) variant did associate with T2D (OR per Met allele = 1.60[1.21–2.13], *P*-value = 0.001 with genomic control, covariates-age, sex, birth year and principal components 1–5 from genome-wide association study, Table 1), where association was observed in both the full-heritage and non-full-heritage Pima Indians which make up this population-based sample (Table 1). This variant was also associated with age of diabetes onset when analyzed using a Cox proportional hazard model (HR per Met-allele = 1.40[1.14–1.72], *P* value = 0.001).

In addition to the c.839C>T (p.(Thr280Met)), a novel missense variant [hg19 chr20:g.43945378G>A, c.1333G>A, p.(Asp445Asn)] was detected in *RBPJL*; however, analysis of this variant showed no association with T2D (Table 1, ESM Table 1) in our data set. Analysis of previously obtained whole genome sequencing data from 335 American Indians identified two additional missense variants (one in *RBPJL* and one in *PTF1A*), which were not analyzed in the current study due to their rare frequency

**Fig. 1** Confirming the role of RBPJL in regulating the expression of pancreatic exocrine enzymes. 266-6 cells (mouse pancreatic acinar cells) were transfected with either control siRNA or siRNA targeting mouse *Rbpjl*. Results shown are mean  $\pm$  SD from four independent experiments. % reduction is the percent reduction in expression of target genes in cells treated with *Rbpjl* siRNA compared to basal (untreated cells). No effect was seen on the expression of *B2m* a housekeeping gene not regulated by *Rbpjl*. Statistical association was determined using an unpaired *t*-test. \*\*\* $P < 0.001$ , \*\* $P < 0.01$ , NS not significant, NR no reduction. Gene expression (mRNA) levels were normalized to the expression of *Tbp*



**Fig. 2** RBPJL Met280 variant has lower transactivation activity. Luciferase assays using two different RBPJL responsive promoters (*CTRB1* and *CPA1* promoters) to identify the difference in transactivation potential between RBPJL Thr280 and Met280 was done in HEK293 cells. Shown is the fold increase in relative luciferase activity with RBPJL Thr280 (light gray box) or Met280 (dark gray box) normalized to relative luciferase activity with PTF1A alone while using *CPA1* promoter–reporter construct and *CTRB1* promoter–reporter construct. Data are represented as a box plot from 12 data points recorded from four experiments done in triplicates. The horizontal line within the box represents the median. Fold increase in luciferase activity from each experiment was averaged from the triplicates and the difference in activity between Thr280 and Met280 was compared using an unpaired *t*-test ( $N = 4$ ) \*\*\* $P = 0.0003$ , \* $P = 0.01$  (MAF  $< 0.002$ ) and their prediction to be benign according to online prediction tools (ESM Table 4).

### Confirming the role of RBPJL in regulating the expression of pancreatic exocrine enzymes

A human cDNA expression panel was screened to confirm that *RBPJL* is expressed predominantly in the pancreas

(ESM Fig. 1). To establish the role of RBPJL as a regulator of pancreatic exocrine enzyme expression, siRNA was used to achieve a 78.1% relative reduction of *Rbpjl* expression in 266-6 mouse pancreatic acinar cells. Comparing the expression of known *Rbpjl* target genes, *Cel*, *Ctrb*, *Cpa1*, *Cpa2*, and *Cela1*, between untreated and *Rbpjl* siRNA treated cells showed a highly significant reduction in the expression of all five target genes; whereas *B2m*, which is not regulated by *Rbpjl*, showed no difference in expression between the untreated and *Rbpjl* knockdown cells (Fig. 1). Expression of these genes was not reduced in cells treated with a control siRNA.

### RBPJL Met280 variant has lower transactivation activity

The effect of the RBPJL c.839C>T (p.(Thr280Met)) variant on trans-activating the promoters of two known target genes, *CPA1* and *CTRB1* was assessed in Human Embryonic Kidney cells-293 (HEK-293). First, we examined the transactivation of *CPA1* and *CTRB1* promoters by normal (Thr280) RBPJL alone or in combination with PTF1A. HEK293 cells which express RBPJL but not RBPJL or PTF1A were co-transfected with either the *CPA1* or *CTRB1* promoter–reporter vector and either (1) RBPJL expression vector alone, (2) PTF1A expression vector alone, or (3) both RBPJL and PTF1A expression vectors. In agreement with previous reports [1], our data showed that co-transfection of the promoter–reporter constructs and the vector expressing RBPJL was not able to significantly activate either the *CPA1* or *CTRB1* promoter (ESM Fig. 2).

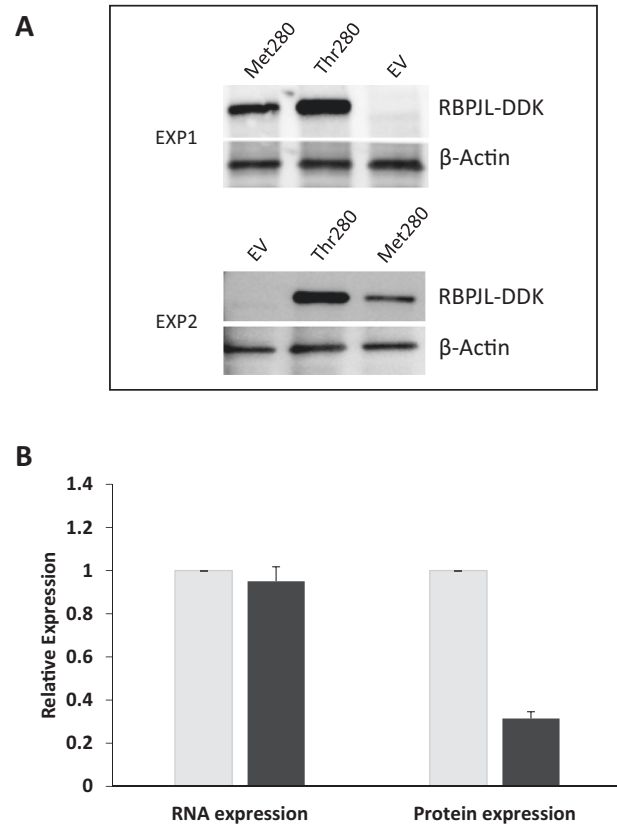
In contrast, an increase in *CPAI* and *CTRB1* promoter activity was observed when co-transfected with PTF1A alone (ESM Fig. 2), presumably due to endogenous expression of RBPJ in HEK293 cells. A 3.2-fold and 6.3-fold increase in *CPAI* and *CTRB1* promoter activity, respectively, was observed when both RBPJL and PTF1A expression vectors were co-transfected with the promoter–reporter constructs compared to PTF1A alone (ESM Fig. 2). Therefore, this condition was used to compare the transactivation function between Thr280 RBPJL and Met280 RBPJL. The *CPAI* promoter–reporter construct had a 36% decrease in luciferase activity when co-transfected with Met280 RBPJL as compared to Thr280 RBPJL ( $P$ -value = 0.0003, Fig. 2). A similar result was obtained using the *CTRB1* promoter–reporter construct (15% decrease in activity of the Met280 RBPJL;  $P$ -value = 0.01, Fig. 2).

### RBPJL Met280 variant affects protein expression

To determine whether the c.839C>T (p.(Thr280Met)) variation affects RBPJL protein levels, vectors expressing either Thr280 or Met280 RBPJL were transiently transfected into HEK293 cells. Both RNA and protein were extracted and subjected to real-time PCR and western blot analysis. Our western blot data showed a reduced amount of Met280 RBPJL protein compared to Thr280 RBPJL (Fig. 3a), whereas no difference was seen in the relative mRNA expression levels measured from the same cells (Fig. 3b). Quantification of the western blot data identified a 68% relative decrease in Met280 RBPJL protein as compared to Thr280 RBPJL after normalizing with  $\beta$ -actin (Fig. 3b).

### Functional assessment of RBPJL p.(Thr280Met) variant in pancreatic acinar cells

To assess whether the c.839C>T (p.(Thr280Met)) variant in RBPJL functionally impacts expression of target genes, Met280 RBPJL or Thr280 RBPJL was co-transfected along with PTF1A, in mouse pancreatic acinar cells (266-6). An average 56.2-fold relative increase in *Cel* RNA expression was observed when Thr280 RBPJL was overexpressed along with PTF1A as compared to overexpression of PTF1A alone (Fig. 4). In contrast, overexpression of Met280 RBPJL along with PTF1A only led to an average 11.6-fold relative increase in the expression of *Cel* RNA compared to PTF1A alone (Fig. 4). This represents a 4.8-fold decrease in activity of the Met280 RBPJL compared to Thr280 RBPJL ( $P$ -value < 0.0001). Consistent with our earlier experiments, the average mRNA expression of Thr280 *RBPJL* and Met280 *RBPJL* was not significantly different. As a control, *Gapdh* (a gene not regulated by RBPJL) mRNA levels were also shown to be comparable between the three conditions

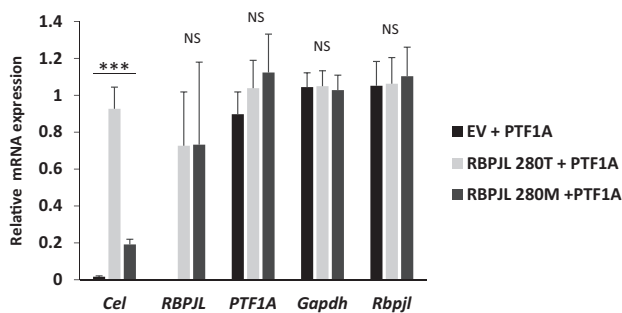


**Fig. 3** RBPJL Met280 variant affects protein expression. HEK293 cells were transfected with either RBPJL Thr280, Met280 or pCMV6 empty vector (EV). Forty-eight hours after the transfection, cells were harvested and divided into two equal aliquots. One aliquot was used for RNA isolation and subsequent real-time PCR. The other aliquot was used for protein extraction and western blotting. **a** Western blots were repeated twice and data (blots) from both experiments are shown. EV is pCMV6 empty vector used as a negative control. The blots were cropped for conciseness of the presentation. An expanded image is shown in electronic supplementary figure 4. **b** Relative quantification of RNA expression (real-time PCR) and protein expression (western blots) are shown as mean  $\pm$  sd from the two experiments. Gene expression (mRNA) levels were normalized to the expression of *RPL10*. Light gray, results for RBPJL Thr280 and dark gray, results for RBPJL Met280

(Fig. 4). Expression of *PTF1A* and mouse *Rbpjl* was also not significantly different, suggesting that the difference in *Cel* RNA expression is due to reduced quantity and/or activity of the Met280 RBPJL (Fig. 4).

### Association analysis of the functional p.(Thr280Met) variant in independent data sets

To determine whether the functional Amerindian-derived population-specific c.839C>T (p.(Thr280Met)) variant in *RBPJL* has a discernable effect on T2D in other studies, we further analyzed this variant in additional American data sets. In a cross-sectional study of urban-dwelling American Indians informative for T2D status living near Phoenix, AZ,



**Fig. 4** Functional assessment of *RBPJL* p.(Thr280Met) in mouse pancreatic acinar cells. The data show the effect of *RBPJL* Thr280 (light gray) and Met280 (dark gray) on the expression of *Cel*, a *Rbpjl* target gene, in mouse pancreatic acinar cell line (266-6). As a control, the effect on the expression of *Gapdh*, which is not a target gene, was also assessed. Expression levels of *RBPJL*, *PTF1A* and mouse *Rbpjl* are also shown. Data shown are mean  $\pm$  sd from six different experiments, except for mouse *Rbpjl* which is the mean of four experiments. \*\*\* $P$ -value  $< 0.0001$ , NS not significant, EV pCMV6 empty vector. Gene expression (mRNA) levels were normalized to the expression of *Tbp*

the Met280 allele had a directionally consistent effect on T2D association ( $N = 2993$ , Met allele frequency = 0.04, OR = 1.19[0.87–1.62], ESM Table 5), however, the  $P$ -value did not reach statistical significance ( $P$ -value = 0.27). This data set consisted of subjects belonging to different tribal affiliations. Among these groups, the Met280 allele had an association with T2D in 843 individuals (group I) who had  $\geq 50\%$  American Indian ancestry (frequency of the Met allele = 0.03;  $P$ -value = 0.002; OR = 2.77[1.45–5.31]) (ESM Table 5).

Because this variant in *RBPJL* is unique to American population in the 1000 Genomes database (Phase 3), it was also assessed for replication of the diabetes association using the SIGMA T2D Consortium exome chip data, which includes cohorts of Mexican descent [15]. Data from the SIGMA T2D Consortium exome chip from the publicly available T2D Knowledge Portal [15] indicated that the c.839C>T (p.(Thr280Met)) variant in *RBPJL* was less common in the SIGMA cohort of Mexican Americans (frequency of Met allele = 0.01) than observed in American Indians, and the association with T2D was directionally consistent but not statistically significant ( $N = 8214$ , OR = 1.16;  $P$ -value = 0.33). A meta-analysis of the Pima Indian, all Urban American Indian, and SIGMA data sets using the inverse-variance method resulted in a summary OR of 1.32 per Met allele ( $P$ -value = 0.0015) with little evidence for heterogeneity in effect ( $P$ -value for heterogeneity = 0.23).

## Discussion

We evaluated whether common variation in any of the genes that encode proteins forming the PTF1 complex

contribute to T2D in American Indians. Our analysis identified a c.839C>T (p.(Thr280Met)) missense variant (rs200998587:C>T) in the pancreas specific gene *RBPJL* that associates with T2D in a population-based sample of Pima Indians. There are four alternative transcripts for *RBPJL*, of which three encode protein products. The p.(Thr280Met) is positioned in the beta-trefoil DNA-binding domain (exon 8) of all three protein coding isoforms (ESM Fig. 3). The p.(Thr280) in *RBPJL* is highly conserved across many species including mice, and the p.(Thr280Met) variation is predicted to be deleterious, damaging and probably damaging by Provean, SIFT, and Polyphen-2 online prediction programs respectively (ESM Table 4). Our Western blot and real-Time PCR analyses provided evidence that the 280Met is less stable than the 280Thr *RBPJL* (Fig. 3a, b). Our in vitro studies demonstrated reduced transactivation by *RBPJL* in cells expressing the 280Met as compared to the 280Thr variant. These experiments included analyzing transactivation on two different promoters, as well as, direct measurement of the effect of *RBPJL* c.839C>T (p.(Thr280Met)) on the expression of *Cel*, a known target gene (Figs. 2, 4). In comparison, the *RBPJL* c.1333G>A (p.(Asp445Asn)) variant, which was not associated with T2D, did not have a significant effect on *RBPJL* function or expression (data not shown). These results led us to conclude that the *RBPJL* c.839C>T (p.(Thr280Met)) is a functional variant and the Met280 *RBPJL* has reduced activity.

The *RBPJL* c.839C>T (p.(Thr280Met)) variant is a low frequency variant (Met allele frequency = 0.03). This variant associated with T2D in our study (OR = 1.60 [1.21–2.13] per Met-allele,  $P$ -value = 0.001), where the association was consistent when considering either the full-heritage or the non-full-heritage Pima Indians from this sample separately. In an independent study of urban-dwelling American Indians, the *RBPJL* c.839C>T (p.(Thr280Met)) variant associated with T2D in one out of four groups and the combined analysis of the four groups identified a directionally consistent effect on T2D risk, whereas in the publicly available T2D Knowledge Portal, this variant again had a directionally consistent effect on T2D risk but was not statistically significant. Therefore, we cannot say that this variant is a reproducible predictor of T2D, however our studies indicate a role for this variant in modifying T2D risk by its effect on pancreatic exocrine enzyme expression.

*RBPJL* is predominately expressed in pancreatic acinar cells, and knockout models have shown a critical role for *Rbpjl* in the development and maintenance of the exocrine pancreas via regulation of exocrine enzymes [5]. Knock-out mouse models also provide evidence that *Rbpjl* plays a role in glucose homeostasis, where female mice with a targeted *Rbpjl* deletion develop impaired glucose tolerance (<http://www.informatics.jax.org/a>

allele/allgenoviews/MGI:5636933). Our *Rbpjl* knockdown studies in mouse pancreatic acinar cells provide further support for the role of *Rbpjl* in the regulation of exocrine enzyme expression. Therefore, it is possible that a reduction in RBPJL level and/or function could lead to a decrease in pancreatic enzymes in the gut causing improper breakdown of proteins, fats, and carbohydrates. This may in-turn lead to impaired incretin (GLP-1 and GIP1) synthesis and secretion [16–18] and reduced incretin mediated insulin secretion which is a known risk factor for T2D [19].

Two of the target genes of RBPJL have been previously implicated in diabetes. Previous publications as well as our current study have shown that *Rbpjl* regulates the expression of *Cel*, which encodes carboxyl-ester lipase, and *Ctrb* which encodes the pancreatic enzyme chymotrypsinogen B [5]. Specific variants in *CEL* cause MODY [20, 21]; however, these *CEL* variants result in protein misfolding and aggregation [22, 23]; therefore, it is unknown whether impaired expression of *CEL*, because of reduced RBPJL, would similarly affect diabetes risk. Two different genome-wide association studies have shown that a variant, rs7202877:T>G (hg19 chr16: g.75247245T>G), at the *CTRB1/2* locus associates with T2D, and increased GLP-1-stimulated insulin secretion, respectively [24, 25]. The G allele, which is protective for T2D and increases GLP-1-stimulated insulin secretion, was also shown to increase expression of *CTRB1* and *CTRB2* in normal pancreatic biopsy samples and pancreatic samples from cadavers [25]. Interestingly, in American Indians, this SNP has a modest but directionally consistent association with T2D (OR = 0.76 (0.60–0.97) per G allele, *P*-value = 0.03), suggesting that the pancreatic enzyme mediated GLP-1 secretion and subsequent stimulation of insulin secretion may play a role in T2D pathogenesis in American Indians. Therefore, our data are consistent with a mechanism whereby RBPJL binds to the promoter of *CTRB1* and other exocrine enzymes and regulate their expression, and variation that diminishes RBPJL may decrease expression of *CTRB1* (and other exocrine enzymes) leading to an increased risk of T2D via reduced incretin effect.

In conclusion, this study identified a functional missense variant in *RBPJL* that associates with T2D in American Indians. Although the variant itself is not a very strong predictor of risk for diabetes, our data show that the variant impacts transactivation by RBPJL, which could potentially lead to a reduction in exocrine enzymes and/or reduced GLP-1-stimulated insulin secretion. Further studies to explore the role of pancreatic exocrine enzymes in the pathogenesis of T2D and the role of RBPJL in this process are warranted.

**Acknowledgements** The authors thank the volunteers who participated in the study and also thank the staff of Phoenix Epidemiology

and Clinical Research Branch who helped with the study. This work was supported by the intramural research program of the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. If you remix, transform, or build upon this article or a part thereof, you must distribute your contributions under the same license as the original. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/4.0/>.

## References

1. Beres TM, Masui T, Swift GH, Shi L, Henke RM, MacDonald RJ. PTF1 is an organ-specific and Notch-independent basic helix-loop-helix complex containing the mammalian Suppressor of Hairless (RBP-J) or its paralogue, RBP-L. *Mol Cell Biol*. 2006;26:117–30.
2. GTEX consortium. GTEX Portal. <https://www.gtexportal.org/home/>. Accessed 8th Aug 2016.
3. Danielsson A, Pontén F, Fagerberg L, et al. The human pancreas proteome defined by transcriptomics and antibody-based profiling. *PLoS One*. 2014;9:e115421.
4. Masui T, Swift GH, Hale MA, Meredith DM, Johnson JE, MacDonald RJ. Transcriptional autoregulation controls pancreatic Ptf1a expression during development and adulthood. *Mol Cell Biol*. 2008;28:5458–68.
5. Masui T, Swift GH, Deering T, et al. Replacement of Rbpj with Rbpjl in the PTF1 complex controls the final maturation of pancreatic acinar cells. *Gastroenterology*. 2010;139:270–80.
6. Masui T, Long Q, Beres TM, Magnuson MA, MacDonald RJ. Early pancreatic development requires the vertebrate Suppressor of Hairless (RBPJ) in the PTF1 bHLH complex. *Genes Dev*. 2007;21:2629–43.
7. Sellick GS, Barker KT, Stolte-Dijkstra I, et al. Mutations in PTF1A cause pancreatic and cerebellar agenesis. *Nat Genet*. 2004;36:1301–5.
8. Houghton JA, Swift GH, Shaw-Smith C, et al. Isolated pancreatic aplasia due to a hypomorphic PTF1A mutation. *Diabetes*. 2016;65:2810–5.
9. Gloyn AL, Weedon MN, Owen KR, et al. Large-scale association studies of variants in genes encoding the pancreatic beta-cell KATP channel subunits Kir6.2 (KCNJ11) and SUR1 (ABCC8) confirm that the KCNJ11 E23K variant is associated with type 2 diabetes. *Diabetes*. 2003;52:568–72.
10. Bowman P, Flanagan SE, Edghill EL, et al. Heterozygous ABCC8 mutations are a cause of MODY. *Diabetologia*. 2012;55:123–7.



11. Baier LJ, Muller YL, Remedi MS, et al. ABCC8 R1420H loss-of-function variant in a Southwest American Indian Community: association with increased birth weight and doubled risk of type 2 diabetes. *Diabetes*. 2015;64:4322–32.
12. Lozinska L, Weström B, Prykhodko O, et al. Decreased insulin secretion and glucose clearance in exocrine pancreas-insufficient pigs. *Exp Physiol*. 2016;101:100–12.
13. Hayden MR, Patel K, Habibi J, Gupta D, et al. Attenuation of endocrine-exocrine pancreatic communication in type 2 diabetes: pancreatic extracellular matrix ultrastructural abnormalities. *J Cardiometab Syndr*. 2008;3:234–43.
14. Piaggi P, Masindova I, Muller YL, et al. A genome-wide association study using a custom genotyping array identifies variants in GPR158 associated with reduced energy expenditure in American Indians. *Diabetes*. 2017;66:2284–95.
15. T2D-GENES Consortium, GoT2D Consortium, DIAGRAM Consortium. Type 2 Diabetes Knowledge Portal. <http://www.type2diabetesgenetics.org>. Accessed 28 June 2017.
16. Ebert R, Creutzfeldt W. Reversal of impaired GIP and insulin secretion in patients with pancreatogenic steatorrhea following enzyme substitution. *Diabetologia*. 1980;19:198–204.
17. Beglinger S, Drewe J, Schirra J, Göke B, D'Amato M, Beglinger C. Role of fat hydrolysis in regulating glucagon-like Peptide-1 secretion. *J Clin Endocrinol Metab*. 2010;95:879–86.
18. Kuo P, Stevens JE, Russo A, et al. Gastric emptying, incretin hormone secretion, and postprandial glycemia in cystic fibrosis—effects of pancreatic enzyme supplementation. *J Clin Endocrinol Metab*. 2011;96:E851–855.
19. Holst JJ, Knop FK, Vilsbøll T, Krarup T, Madsbad S. Loss of incretin effect is a specific, important, and early characteristic of type 2 diabetes. *Diabetes Care*. 2011;34(Suppl 2):S251–257.
20. Raeder H, Johansson S, Holm PI, et al. Mutations in the CEL VNTR cause a syndrome of diabetes and pancreatic exocrine dysfunction. *Nat Genet*. 2006;38:54–62.
21. Torsvik J, Johansson S, Johansen A, et al. Mutations in the VNTR of the carboxyl-ester lipase gene (CEL) are a rare cause of monogenic diabetes. *Hum Genet*. 2010;127:55–64.
22. Vesterhus M, Raeder H, Kurpad AJ, et al. Pancreatic function in carboxyl-ester lipase knockout mice. *Pancreatology*. 2010;10:467–76.
23. Johansson BB, Torsvik J, Bjørkhaug L, et al. Diabetes and pancreatic exocrine dysfunction due to mutations in the carboxyl ester lipase gene-maturity onset diabetes of the young (CEL-MODY): a protein misfolding disease. *J Biol Chem*. 2011;286:34593–605.
24. Morris AP, Voight BF, Teslovich TM, et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet*. 2012;44:981–90.
25. 't Hart LM, Fritsche A, Nijpels G, et al. The CTRB1/2 locus affects diabetes susceptibility and treatment via the incretin pathway. *Diabetes*. 2013;62:3275–81.