HHS Public Access

Author manuscript

Pancreatology. Author manuscript; available in PMC 2023 June 01.

Published in final edited form as:

Pancreatology. 2022 June ; 22(5): 564–571. doi:10.1016/j.pan.2022.04.015.

Variants in the pancreatic CUB and zona pellucida-like domains 1 (*CUZD1*) gene in early-onset chronic pancreatitis - a possible new susceptibility gene

Agnieszka Magdalena Rygiel[1,#,*], **Lara Sophie Unger**[2,3,#], **Franziska Lena Sörgel**[3,#], **Emmanuelle Masson**[4,5], **Ryotaro Matsumoto**[6], **Maren Ewers**[3], **Jian-Min Chen**[4], **Peter Bugert**[7], **Louis Buscail**[8], **Tomasz Gambin**[1,9], **Grzegorz Oracz**[10], **Maria Winiewska-Szajewska**[11], **Agnieszka Mianowska**[10], **Jarosław Poznanski**[11], **Joanna Kosińska**[12], **Piotr Stawinski**[12], **Rafał Płoski**[12], **Dorota Koziel**[13], **Stanisław Gluszek**[13], **Helmut Laumen**[3,14], **Fredrik Lindgren**[15], **J. Matthias Löhr**[16], **Anna Orekhova**[2], **Vinciane Rebours**[17], **Jonas Rosendahl**[14], **Andrea Párniczky**[18,19,20], **Péter Hegyi**[19,21], **Akira Sasaki**[6], **Fumiya Kataoka**[6], **Yu Tanaka**[6], **Shin Hamada**[6], **Miklós Sahin-Tóth**[2,22], **Eszter Hegyi**[2,19,#], **Claude Férec**[4,5,#], **Atsushi Masamune**[6,#], **Heiko Witt**[3,#,*]

[1]Department of Medical Genetics, Institute of Mother and Child, Warsaw, Poland

[2]Center for Exocrine Disorders, Department of Molecular and Cell Biology, Boston University, Henry M. Goldman School of Dental Medicine, Boston, Massachusetts 02118, United States

[3]Pediatric Nutritional Medicine & Else Kröner-Fresenius-Centre for Nutritional Medicine (EKFZ), Technical University Munich (TUM), Freising, Germany

[4]Univ Brest, Inserm, EFS, UMR 1078, GGB, F-29200 Brest, France

[5]Service de Génétique Médicale et de Biologie de la Reproduction, CHRU Brest, F-29200 Brest, France

[6]Division of Gastroenterology, Tohoku University Graduate School of Medicine, Sendai, Japan

[7]Institute of Transfusion Medicine and Immunology, Medical Faculty Mannheim, Heidelberg University, German Red Cross Blood Service of Baden-Württemberg, Mannheim, Germany

[8]Department of Gastroenterology and Pancreatology, CHU Rangueil and University of Toulouse, Toulouse, France

[*] **Corresponding authors:** Agnieszka M. Rygiel, PhD, Department of Medical Genetics, Institute of Mother and Child, Kasprzaka 17a, 01-211 Warsaw, Poland, Phone: +48 22 32 77 177, Fax: +48 22 32 77 200, agnieszka.rygiel@imid.med.pl, Heiko Witt, M.D., Pediatric Nutritional Medicine & Else Kröner-Fresenius-Centre for Nutritional Medicine (EKFZ), Technical University Munich (TUM), Gregor-Mendel-Str. 2, 85354 Freising, Germany, Phone: +49 8161 71 2466, Fax: +49 8161 71 2465, heiko.witt@tum.de.

[#]These authors contributed equally and share first or last authorship, respectively

Author contributions:

A.M.R and H.W. conceived the study.

C.F., E.H., A.M.R., M.S.-T. and H.W. directed the study.

M.E., T.G., E.M., J.P., P.S., L.S.U., and M.W. performed genotyping, genetic data collection and bioinformatical analysis.

R.M., A.S., F.K., Y.T., S.H. and A.M. recruited study subjects, collected clinical data, and performed genotyping.

E.H., A.O., F.L.S. and L.S.U. performed functional analyses of *CUZD1* variants.

A.M.R, M.S.-T. and H.W. drafted and revised the manuscript with substantial help from C.F., E.H. and J.R..

All other co-authors recruited study subjects, collected clinical data and/or provided genomic DNA samples. All authors approved the final manuscript and contributed critical revisions to its intellectual content.

Disclosure statement: The authors report no conflicts of interest.

[9]Institute of Computer Science, Warsaw University of Technology, Warsaw, Poland

[10]Department of Gastroenterology, Hepatology, Feeding Disorders and Pediatrics, The Children's Memorial Health Institute, Warsaw, Poland

[11]Department of Biophysics, Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw, Poland

[12]Department of Medical Genetics, Medical University of Warsaw, Warsaw, Poland

[13]Collegium Medicum, Jan Kochanowski University of Kielce, Poland

[14]Department of Internal Medicine I, Martin Luther University, Halle, Germany

[15]Department of Pediatric, Karolinska University Hospital, Stockholm, Sweden

[16]Department of Clinical Science, Intervention and Technology, Karolinska Institutet, Stockholm, Sweden

[17]Pancreatology and Digestive Oncology Department, Beaujon Hospital, Clichy, APHP, Université de Paris, Paris, France

[18]Heim Pál National Pediatric Institute, Budapest, Hungary

[19]Institute for Translational Medicine, Medical School, University of Pécs, Pécs, Hungary

[20]Center for Translational Medicine, Semmelweis University, Budapest, Hungary

[21]Division of Pancreatic Diseases, Heart and Vascular Center, Semmelweis University, Budapest, Hungary

[22]Department of Surgery, University of California Los Angeles, Los Angeles, California 90095, United States

Abstract

Objective—Non-alcoholic chronic pancreatitis (NACP) frequently develops in the setting of genetic susceptibility associated with alterations in genes that are highly expressed in the pancreas. However, the genetic basis of NACP remains unresolved in a significant number of patients warranting a search for further risk genes.

Design—We analyzed *CUZD1*, which encodes the CUB and zona pellucida-like domains 1 protein that is found in high levels in pancreatic acinar cells. We sequenced the coding region in 1,163 European patients and 2,018 European controls. In addition, we analyzed 297 patients and 1,070 controls from Japan. We analyzed secretion of wild-type and mutant CUZD1 from transfected cells using Western blotting.

Results—In the European cohort, we detected 30 non-synonymous variants. Using different prediction tools (SIFT, CADD, PROVEAN, PredictSNP) or the combination of these tools, we found accumulation of predicted deleterious variants in patients (p-value range 0.002–0.013; OR range 3.1–5.2). No association was found in the Japanese cohort, in which 13 non-synonymous variants were detected. Functional studies revealed >50% reduced secretion of 7 variants, however, these variants were not significantly enriched in European CP patients.

Conclusion—Our data indicate that *CUZD1* might be a novel susceptibility gene for chronic pancreatitis. How these variants predispose to pancreatitis remains to be elucidated.

Keywords

CUZD1; zymogen granule; chronic pancreatitis; genetics

Introduction

Chronic pancreatitis (CP) is a recurrent or continuing inflammation of the pancreas that results in irreversible morphological changes and impairment of exocrine and endocrine function in some patients [1]. CP develops as a consequence of interactions between genetic and environmental factors. In hereditary and idiopathic non-alcoholic CP (NACP), genetic studies firmly established that mutations in the genes *PRSS1* (cationic trypsinogen), *SPINK1* (serine protease inhibitor Kazal type 1), and *CTRC* (chymotrypsinogen C) drive uncontrolled trypsin activity and are central pathogenic factors [2]. Moreover, genetic risk in CP may be conferred by mutations causing misfolding of digestive enzymes encoded by *PRSS1*, *CPA1* (carboxypeptidase A1), and *CEL* (carboxyl-ester lipase) leading to endoplasmic reticulum stress [3]. In addition to these acinar genes, mutations in ductally expressed genes such as *CFTR* and *TRPV6* have been linked to NACP [4–6]. A common characteristic of CP risk genes described to date is that they are highly expressed in the pancreas, in most cases within the acinar cell compartment. Importantly, however, a considerable fraction of patients with NACP do not carry a mutation in the known susceptibility genes, suggesting the involvement of yet undiscovered genetic risk factors.

CUZD1 codes for the CUB and zona pellucida-like domains 1 protein, which was first cloned from the gravid mouse uterus [7] and later from the non-pregnant rat uterus as estrogen-regulated gene 1 (*ERG1*) and the tamoxifen-induced gene *UO-44* [8,9]. CUZD1 contains an N-terminal signal sequence, two consecutive CUB domains, a zona pellucida (ZP) domain, and a C-terminal transmembrane domain (TMD), suggesting that CUZD1 is a type 1 membrane protein (Figure 1). Indeed, in the mouse pancreas CUZD1 was found to localize to the zymogen granule membrane and was designated as integral membrane-associated protein-1 (Itmap1) [10]. However, CUZD1 also contains a cleavage site between the ZP and TMD domains for the Golgi-protease furin (RSKR; amino acids 519–522), raising the possibility that CUZD1 is secreted like many other ZP proteins [11]. *CUZD1* is highly expressed in the acinar cells of the mouse pancreas [10] and cloning of the human orthologue revealed almost exclusive pancreatic expression [12]. Thus, based on its relatively specific and abundant expression, *CUZD1* is a strong candidate as a pancreatitis susceptibility gene. This assumption is further supported by the observation that *Cuzd1* deficient mice demonstrated increased severity of experimentally induced acute pancreatitis [10]. Therefore, we investigated the association of *CUZD1* and chronic pancreatitis.

Methods

Study population.

We investigated 1,163 unrelated European patients with early-onset NACP (age at onset ≤ 20 years). The patients originated from France (n = 536; 246 female), Germany (n = 464; 256 female), Poland (n = 132; 69 female), Sweden (n = 27; 16 female), and Hungary (n = 4; 4 female). In addition, we analyzed 297 patients from Japan (147 female). In total, we enrolled 2,018 European control subjects from France (n = 570), Germany (n = 980), Poland (n = 323), Hungary (n = 135), and Sweden (n=10). The variant frequencies in the Japanese population (n = 1,070) were obtained from the Tohoku Medical Megabank database [13].

The diagnosis of CP was based on two or more of the following findings: presence of a typical history of recurrent pancreatitis, pancreatic calcifications and/or pancreatic ductal irregularities revealed by endoscopic retrograde cholangiopancreatography or by magnetic resonance imaging or CT of the pancreas or pathological sonographic findings with/without decreased pancreatic function.

The institutional review boards of all participating study centres approved this study. All participants gave written informed consent.

CUZD1 genotyping.

Whole exome sequencing (WES) was performed for one Polish family with CP on HiSeq 2500 platform (Ilumina) using Nextera Rapid Capture Exome kit. In the rest of the cases, we analyzed the coding exons and the exon-intron boundaries of *CUZD1* by PCR amplification and Sanger sequencing. Oligonucleotide sequences, PCR and cycle sequencing conditions are described in the Supplementary material. We confirmed all detected variants in a second independent PCR. SIFT (Sorting Tolerant from Intolerant) and CADD (Combined Annotation Dependent Depletion) predictions were performed using the CADD v1.6 webpage (https://cadd.gs.washington.edu/snv). For PROVEAN (Protein Variation Effect Analyzer) predictions the JCVI webpage (http://provean.jcvi.org/index.php) was used [14]. In addition, mutations were analyzed by the PredictSNP tool (https://loschmidt.chemi.muni.cz/predictsnp1/) [15].

Plasmid construction and mutagenesis.

The *CUZD1* coding DNA was custom synthesized (GenScript, Piscataway, NJ) and cloned into the pcDNA3.1(−) vector using XhoI and HindIII restriction sites. Missense mutations were introduced by overlap extension PCR and cloned into the full-length pcDNA3.1(−) *CUZD1* expression plasmid. The C-terminally truncated *CUZD1* constructs RSKR-Stop and No-TMD were generated by PCR mutagenesis using pcDNA3.1(−) *CUZD1* plasmid as template. In the RSKR-Stop construct, translation was terminated after the RSKR furin cleavage site (amino acids 519–522) while in the No-TMD construct, the C-terminal transmembrane domain was removed by truncation after amino acid 568. All constructs were confirmed by Sanger sequencing.

Cell culture and transfection.

Human embryonic kidney (HEK) 293T cells were maintained at 37 °C in Dulbecco's Modified Eagle Medium (#D6429, Sigma-Aldrich) supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin. For functional analyses of *CUZD1* variants, 5×10^5 cells were seeded in 2 mL into 6-well plates 24 h before transfection with 4 μg plasmid using 10 μL Lipofectamine 2000 Reagent (Thermo Fisher Scientific). After 16 h incubation, the medium containing the DNA-liposome complexes was removed, the cells were washed with 1 mL phosphate buffered saline (PBS) and 2 mL Opti-MEM® Reduced Serum Medium (Thermo Fisher Scientific) supplemented with 1% penicillin/streptomycin was added to each well. Conditioned medium was collected 48 h after medium change. To prepare cell lysates, cells were scraped and collected in 1 mL phosphate-buffered saline, sedimented by centrifugation, and the cell pellet was re-suspended in 200 μL Reporter Lysis Buffer (Promega) with 2 μL Halt Protease and Phosphatase Inhibitor Cocktail (100x, Thermo Fisher Scientific). After 15 min incubation on ice, insoluble cell debris was removed by centrifugation.

Western blot analysis of CUZD1 secretion.

For Western blot analyses we used two different protocols. One of them is described in the Supplementary material. Proteins in the conditioned medium were precipitated using trichloroacetic acid (360 μL medium + 40 μL 100% trichloroacetic acid). After 30 min incubation on ice, precipitated proteins were pelleted by centrifugation at 20,000 g and 4 °C for 15 min. The supernatant was aspirated, the protein pellet was washed with 100 μL ice-cold acetone and centrifuged again for 5 min at 20,000 g and 4 °C. The supernatant was removed and the pellet was allowed to air-dry. Proteins were resuspended in 115 μL 1x reducing Laemmli sample buffer (62.5 mM Tris-HCl (pH 6.8), 10% glycerol, 2% sodium dodecyl sulphate (SDS), 2.5% 2-mercaptoethanol, 0.005% bromophenol blue) and boiled at 95 °C for 5 min. For enzymatic removal of N-linked oligosaccharides, the protein pellet was resuspended in 60 μL H_2O and 14 μL resuspended protein was combined with 2 μL 10x GlycoBuffer 2 and 2 μL PNGase F (#P0704, New England Biolabs) in a total volume of 20 μL. After incubation at 37 °C for 2 h, the deglycosylation reaction was terminated by adding 7 μL 4x reducing Laemmli sample buffer and boiling the sample at 95 °C for 5 min. Aliquots of 16 μL (corresponding to 50 μL conditioned medium) were electrophoresed on a 10% SDS-polyacrylamide gel (SDS-PAGE). Proteins were blotted on a nitrocellulose membrane at 360 mA for 30 min using a wet blotting system. The membrane was blocked with 5% bovine serum albumin (BSA) in PBS for 1 h at room temperature prior to overnight incubation at 4 °C with primary antibodies diluted in blocking solution supplemented with 0.1% Tween 20. The rabbit anti-CUZD1 polyclonal antibody (#HPA068479 (Lot: R100090), Sigma-Aldrich) was used at 1:800 dilution, the mouse anti-CUZD1 monoclonal antibody (#sc-514578, Santa Cruz Biotechnology) was used at 1:1,000 dilution and the goat anti-β-actin polyclonal antibody (ab8229 (Lot: GR3272170–1), Abcam) was used at 1:2,500 dilution. Secondary antibodies conjugated to IRDye® 680RD (anti-rabbit IgG, #926–68073; anti-mouse IgG, #926–68072; LI-COR Biosciences) and IRDye® 800CW (anti-goat IgG, #926–32214, LI-COR Biosciences), respectively, were used at 1:10,000 dilution in PBS-T for 2 h at room temperature. Fluorescence was captured using a Sapphire™ Biomolecular Imager (Azure Biosystems) and were saved as TIF files. Images

were analyzed densitometrically with LI-COR Image Studio Lite software version 5.2.5 by encircling each red band manually. Analyzed areas were identical in size within each blot and median local background correction was performed by the software.

Statistical analysis.

Differences in the frequency of variants between cases and controls were determined with two-tailed Fisher's exact test. A *p*-value of less than 0.05 was considered significant. The statistical analyses were conducted using the SPSS software (Chicago, IL; v 22) and GraphPad Prism (v 4.03).

Results

Genotyping.

We performed whole exome sequencing (WES) on a Polish index patient and her affected sister and identified in both cases a heterozygous *CUZD1* variant, p.C229S, which is not reported in the gnomAD database (https://gnomad.broadinstitute.org/) and which is predicted to destabilize the CUB2 domain by disrupting the p.C207-p.C229 disulfide bond. Independently, we examined German NACP patients and controls for *CUZD1* variants in a candidate gene approach using Sanger sequencing and found several mutations. Thus, we broadened our analyses to other European cohorts and to a Japanese cohort.

In the European cohort, we found 30 non-synonymous variants (Table 1). We discovered one of these variants, p.C462Y, in three unrelated patients. All other alterations, with the exception of p.I285T, were only found once or twice in patients. Due to their rarity, association of the individual variants with CP did not achieve statistical significance with the exception of p.C462Y (P=0.049). None of the European subjects were homozygous or compound heterozygous for a *CUZD1* variant.

Fifteen variants were predicted as deleterious by SIFT. Including p.Q481Pfs*, which cannot be predicted by any applied tool, 17/1,163 (1.5%) patients but only 8/2,018 (0.4%) controls carried a deleterious variant (P=0.002; OR 3.7; 95% CI 1.6–8.7). Using other prediction tools such as CADD, PROVEAN and PredictSNP, we also found a significant enrichment of predicted deleterious variants in the patient cohort (Table 1). When combining the prediction tools, we also found an association of variants that all tools classified as deleterious. Using a strict CADD score (PHRED score ≥25), 9/1163 (0.8%) of patients but only 3/2,018 (0.2%) of controls showed a deleterious variant (P=0.012; OR=5.2; 95% CI 1.4–19.4). Using a less stringent CADD score (PHRED ≥20), 14 (1.2%) patients but only 6 (0.3%) of the control group had a deleterious variant (P=0.004; OR=4.1; 95% CI 1.6–10.7) (Table 1).

We analyzed the German cohort for trans-heterozygosity for known pancreatitis risk genes (*CEL*, *CFTR*, *CPA1*, *CTRC*, *PNLIP*, *PRSS1*, *SPINK1*, and *TRPV6*). One patient with p.I203R was homozygous for *SPINK1* p.N34S and the patient with p.Q227H was heterozygous for *CFTR* p.S1235R. Among carriers of the p.R123H variant, one patient was heterozygous for *SPINK1* p.N34S, and another was heterozygous for *TRPV6* p.L299Q. In the case of the more common p.I285T variant, three patients were trans-heterozygous: one patient for *SPINK1* c.194+2T>C, one for *CFTR* p.F508del, and one for *TRPV6*

p.L299Q and the *CEL-HYB1* allele. Notably, all 4 variants in which trans-heterozygosity was observed, were predicted to be innocuous by at least two prediction tools and showed normal secretion.

In the Japanese cohort, we found 13 different non-synonymous variants (Table 2). With exception of p.C207Y and p.R231H, none of these alterations were found in European patients. Three variants (p.S31R, p.E43D and p.V246M) were always found together, indicating a complete linkage disequilibrium, consistent with the SNiPA data (LD r^2 <1.0) (https://snipa.helmholtz-muenchen.de/snipa3/) [16]. Therefore, we analyzed their functional effect by generating a triple mutant. Applying different prediction tools, the distribution of predicted deleterious variants between cases and controls was not significantly different (*p*-values ranging from 0.38 to 1.0) (Table 2).

Experimental analysis of CUZD1 variants.

To analyze the functional consequences of *CUZD1* mutations, we studied secretion of all *CUZD1* variants in transiently transfected HEK 293T cells. First, we tested whether CUZD1 is secreted to the cell culture medium by using Western blot to detect CUZD1 in conditioned media and cell lysates. We found that binding of our commercial antibody (sc-514578, Santa Cruz Biotechnology) was blocked by N-glycosylation of CUZD1, which had to be removed before SDS-PAGE by treatment with PNGase F (Supplementary Figure 1B). De-glycosylated CUZD1 was then readily detected in the growth medium as a strong, broad band of around 55 kDa and as a sharper band of the same size in cell lysates (Supplementary Figure 1C). An additional band of 34 kDa was always prominent in cell lysates and probably represents non-specific antibody binding (Supplementary Figure 1C). Previous work on mouse ZP proteins demonstrated that truncation of the protein immediately after the furin cleavage site resulted in decreased secretion with increased intracellular retention, while constructs truncated just before the TMD were secreted efficiently [11]. We tested the effect of these C-terminal deletions on CUZD1 and found that truncation after the furin site (RSKR-Stop) elevated intracellular levels while truncation before the TMD (No-TMD) increased secretion, relative to wild-type CUZD1 (Supplementary Figure 1A and 1C). Taken together, our data indicate that CUZD1 is a secretory protein, which behaves similarly to the well-characterized mammalian egg coat ZP proteins.

Analysis of the CUZD1 variants revealed that 7 variants exhibited some form of secretion defect (Figure 2; for the complete uncropped Western blot images see Supplementary Figure 2). As expected, the two nonsense variants p.R464X and p.Q481Pfs* showed no secretion. The secretion of p.C207Y was strongly decreased and the secretion of p.G95R, p.C229S, p.L322P and p.G391D was markedly diminished. The variants p.Y255C, p.R355H and p.R511C were secreted at reduced levels compared to wild-type CUZD1 (secretion level approximately 50–62% of wild-type). Importantly, variants with impaired secretion such as p.C207Y, p.C229S and p.L322P were detectable in cell lysates (Supplementary Figure 3). Variant p.C462Y showed reduced secretion, which can be attributed to the fact that the variant lies in a region against which the antibody was generated (HPA068479, amino acids 425–496). Staining with another antibody (sc-514578) showed normal secretion. As expected, p.R464X and p.Q481Pfs* could not be detected (Supplementary Figure 2 and 4A).

Consistent with the predictions, all variants considered to have no effect on protein function with at least three prediction tools showed normal secretion (Table 1).

Interestingly, variant p.Y255C showed a relative secretion defect with slightly reduced extracellular and significantly elevated intracellular levels, relative to wild-type CUZD1 (Supplementary Figure 3). This observation indicates that this variant becomes intracellularly retained and degraded to a varying extent, in all likelihood due to misfolding.

Three variants showed faster electrophoretic mobility: p.A24E, the triple variant (p.S31R_p.E43D_p.V246M) and p.N148H (Figure 2). This is probably due to the loss of glycosylation sites in all three cases. N-linked glycosylation of asparagine residues occurs at the sequons Asn-X-Thr and Asn-X-Ser (X denotes any amino acid except proline) [17]. CUZD1 contains five predicted N-glycosylation sites at positions 29, 57, 67, 394 and 419 (https://www.uniprot.org/uniprot/Q86UP6). The lower molecular weight of the triple variant p.S31R_p.E43D_p.V246M can therefore be explained by the loss of N-glycosylation at position 29 due to the replacement of serine by arginine at position 31. After removal of N-linked glycans by PNGase F, the molecular weight no longer differed (Supplementary Figure 4B). The lower molecular weight of p.N148H can also be explained by the loss of a glycosylation site, since asparagine at position 148 is followed by isoleucine and serine, making it a suitable site for N-glycosylation. Alanine at position 24 is the last amino acid of the signal peptide. According to the −3,−1 rule, signal peptidases prefer small and uncharged amino acids at positions −3 (P3) and −1 (P1) relative to the cleaved peptide bond (P1-P1') [18]. The exchange of alanine by the charged amino acid glutamate in p.A24E (P1 position) likely disrupts the cleavage site. In this case, CUZD1 is probably cleaved after alanine at position 30 because together with glycine at position 28 the "−3,−1-rule" would be fulfilled. Since mature CUZD1 would lack six amino acids, including the N-glycosylation site at position 29, this would also explain the lower molecular weight of p.A24E and why this difference disappears after PNGase F treatment (Supplementary Figure 4B).

Discussion

In this study, we investigated the association of *CUZD1* variants with NACP in two independent cohorts. *CUZD1* emerged as a potential pancreatitis gene because of its high, relatively pancreas-specific expression, which has been a common feature of most risk genes identified to date. Our results revealed that most of the non-synonymous variants detected in *CUZD1* are rare in all populations studied. Using different prediction tools such as SIFT, CADD, PROVEAN and PredictSNP, we observed a significant association of predicted deleterious variants in the European cohort, both with the individual tools and with a combination of the tools. In contrast, we found no association in the Japanese cohort, suggesting that *CUZD1* may not be a global risk gene for NACP.

A limitation of our study is that the composition of the European patient and control cohort is not harmonized. Although we cannot rule out that ethnic differences influence the outcome of the study, it is worth noting that two putative deleterious mutations were found in different populations. p.R355H was detected in a German and a French patient

and p.R511C in a French and a Swedish patient. None of the two variants were found in controls.

Functional analyses showed the predicted damaging effect in 7 variants as a secretion defect (>50% secretion reduction relative to wild-type). In contrast, variants predicted as tolerated by at least three prediction tools showed normal secretion. Remarkably, p.C462Y, which was found three times in patients, but not in our controls and only once in more than 120,000 subjects reported in gnomAD, showed normal secretion and intracellular levels. Thus, our functional experiments may not have captured the pathological mechanism of *CUZD1* mutations in NACP.

Our finding that CUZD1 may predispose to NACP is consistent with previously reported mouse experiments demonstrating increased severity of experimental pancreatitis in *Cuzd1*-deficient mice [10]. Paradoxically, however, the *Cuzd1*-deleted mice also showed significantly reduced intra-pancreatic trypsin activation, which raises the possibility that the observed effects were related to the mixed genetic background of the *Cuzd1*-deficient mice. Recent studies showed that even sub-strains of the same mouse line can respond with different severity to experimentally induced pancreatitis and intra-acinar trypsin activation also shows strain-dependent variations [19]. In any event, a mechanistic explanation for the pathogenic effect of human *CUZD1* variants is not readily apparent from the published mouse studies.

Variants in acinar risk genes identified to date exert their pathogenic effect predominantly via one of two well-defined pathological pathways: the trypsin-dependent and the misfolding-dependent pathway [2,3]. Although data are not shown, we found no effect of recombinant CUZD1 (Biovendor, RD172487100) on trypsin activity or trypsinogen autoactivation. Similarly, co-transfection of HEK 293T cells with PRSS1 and CUZD1 did not alter trypsinogen secretion. Therefore, it appears that CUZD1 has no direct interaction with trypsinogen or trypsin and *CUZD1* variants are unlikely to exert their effect through a trypsin-dependent mechanism. Our functional analyses showed that some CUZD1 variants exhibited reduced secretion, probably caused by CUZD1 misfolding with subsequent intracellular retention. Previous studies with CPA1 variants and some PRSS1 variants demonstrated that misfolding exerts its pathogenic effect through the induction of endoplasmic reticulum stress, which can result in acinar cell death with consequent inflammation [3]. We evaluated ER stress markers in HEK 293T cells expressing misfolding CUZD1 variants but found no elevation of the chaperone BiP or increased splicing of XBP1 (data not shown). However, these negative results should be interpreted with caution, since the heterologous expression of CUZD1 was relatively low and may not reflect the conditions in human acinar cells.

It is also conceivable that the mechanism of action of CUZD1 variants is novel and related to a diminished CUZD1 function. Western blot analyses showed that CUZD1 is secreted, consistent with a previous publication [20]. In this report, CUZD1 was detected by Western blot analysis in pancreatic tissue extract and pancreatic juice. The biological role of CUZD1 remains undefined but the protein contains multiple functional domains that may give us clues. The CUB domains are frequently found in various extracellular and plasma membrane

proteins, including a number of proteases [21,22]. These proteins are often developmentally regulated and may be involved in a variety of functions in which CUB-domain mediated protein-protein interactions are important. The ZP domain is also present in a variety of proteins, which often form extracellular fibrils or filaments where the ZP domain acts as a polymerization module [11,23]. A newly discovered function of ZP domain proteins is the control of cell shape and polarization in differentiated epithelia [23]. Finally, more recent mouse experiments demonstrated that CUZD1 is essential for mammary gland development and mammary overexpression of *Cuzd1* may cause breast tumors [24,25]. In this role, CUZD1 appears to act as a cytoplasmic protein that can complex phosphorylated Stat5 and translocate into the cell nucleus. It remains unclear how the secretory protein CUZD1 may acquire a cytoplasmic localization. Taken together, the wide array of functions of CUB and ZP domain-containing proteins, and the newly discovered roles for CUZD1, suggest that partial loss of CUZD1 in the human pancreas may have pathological consequences, including pancreatitis.

In conclusion, our observations suggest that *CUZD1* represents a novel risk gene for NACP. Missense *CUZD1* variants predicted to be deleterious were significantly overrepresented in European patients with NACP. Although our functional data suggest that reduced secretion due to protein misfolding may be involved, the precise disease mechanism of pathogenic *CUZD1* variants remains uncertain.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Financial support: The authors thank all study participants for providing clinical data and blood samples. This work was supported by the National Science Centre, Poland grant no. 2015/19/B/NZ5/02224 (to AMR), the Else Kröner-Fresenius-Foundation (EKFS) 2017_A108 – EKFZ-Witt (to HW), the National Institutes of Health (NIH) grant R01 DK058088 (to MST), the DFG RO 3929/5-1 and RO 3929/6-1 (to JR), Hungarian National Research Grant FK138929 (to AP), Eötvös Loránd Research Network award 460051 (to EH), AMED 21ek0109550h0001 (to AM), the Association Gaétan-Saleün, the Association des Pancréatites Chroniques Héréditaires, and the Institut National de la Santé et de la Recherche Médicale (INSERM), France.

List of Abbreviations:

CUZD1	CUB and zona pellucida-like domains 1 protein
CP	chronic pancreatitis
NACP	non-alcoholic chronic pancreatitis

References

[1]. Witt H, Apte MV, Keim V, Wilson JS. Chronic pancreatitis: challenges and advances in pathogenesis, genetics, diagnosis, and therapy. Gastroenterology 2007; 132: 1557–1573. [PubMed: 17466744]

[2]. Hegyi E, Sahin-Tóth M. Genetic risk in chronic pancreatitis: The trypsin-dependent pathway. Dig Dis Sci 2017; 62: 1692–1701. [PubMed: 28536777]

[3]. Sahin-Tóth M.Genetic risk in chronic pancreatitis: the misfolding-dependent pathway. Curr Opin Gastroenterol 2017; 33: 390–395. [PubMed: 28650851]

[4]. Sharer N, Schwarz M, Malone G, Howarth A, Painter J, Super M, et al. Mutations of the cystic fibrosis gene in patients with chronic pancreatitis. N Engl J Med 1998; 339: 645–652. [PubMed: 9725921]

[5]. Masamune A, Kotani H, Sörgel FL, Chen JM, Hamada S, Sakaguchi R, et al. Variants that affect function of calcium channel TRPV6 are associated with early-onset chronic pancreatitis. Gastroenterology 2020; 158: 1626–1641. [PubMed: 31930989]

[6]. Oracz G, Zaród M, Ewers M, Laumen H, Gambin T, Kamiński P, et al. Loss of function TRPV6 variants are associated with chronic pancreatitis in nonalcoholic early-onset Polish and German patients. Pancreatology 2021, 21:1434–1442. [PubMed: 34538581]

[7]. Kasik JW. A cDNA cloned from pregnant mouse uterus exhibits temporo-spatial expression and predicts a novel protein. Biochem J 1998; 330: 947–950. [PubMed: 9480914]

[8]. Chen D, Xu X, Zhu LJ, Angervo M, Li Q, Bagchi MK, et al. Cloning and uterus/oviduct-specific expression of a novel estrogen-regulated gene (ERG1). J Biol Chem 1999; 274: 32215–32224. [PubMed: 10542259]

[9]. Huynh H, Ng CY, Lim KB, Ong CK, Ong CS, Tran E, et al. Induction of UO-44 gene expression by tamoxifen in the rat uterus and ovary. Endocrinology 2001; 142: 2985–2995. [PubMed: 11416020]

[10]. Imamura T, Asada M, Vogt SK, Rudnick DA, Lowe ME, Muglia LJ. Protection from pancreatitis by the zymogen granule membrane protein integral membrane-associated protein-1. J Biol Chem 2002; 277:50725–50733. [PubMed: 12401800]

[11]. Jovine L, Darie CC, Litscher ES, Wassarman PM. Zona pellucida domain proteins. Annu Rev Biochem 2005; 74: 83–114. [PubMed: 15952882]

[12]. Leong CT, Ng CY, Ong CK, Ng CP, Ma ZS, Nguyen TH, et al. Molecular cloning, characterization and isolation of novel spliced variants of the human ortholog of a rat estrogen-regulated membrane-associated protein, UO-44. Oncogene 2004; 23: 5707–5718. [PubMed: 15184879]

[13]. Nagasaki M, Yasuda J, Katsuoka F, Nariai N, Kojima K, Kawai Y, et al. Rare variant discovery by deep whole-genome sequencing of 1,070 Japanese individuals. Nat Commun 2015; 6: 8018. [PubMed: 26292667]

[14]. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the Functional Effect of Amino Acid Substitutions and Indels. PLoS One 2012; 7: e46688.

[15]. Bendl J, Stourac J, Salanda O, Pavelka A, Wieben ED, Zendulka J, et al. PredictSNP: robust and accurate consensus classifier for prediction of disease-related mutations. PLoS Comput Biol 2014; 10 : e1003440.

[16]. Arnold M, Raffler J, Pfeufer A, Suhre K, Kastenmüller G. SNiPA: an interactive, genetic variant-centered annotation browser. Bioinformatics 2015; 31: 1334–1336. [PubMed: 25431330]

[17]. Kornfeld R, Kornfeld S. Assembly of asparagine-linked oligosaccharides. Annu Rev Biochem 1985; 54: 631–64. [PubMed: 3896128]

[18]. von Heijne G.Patterns of amino acids near signal-sequence cleavage sites. Eur J Biochem 1983; 133: 17–21. [PubMed: 6852022]

[19]. Ulmasov B, Oshima K, Rodriguez MG, Cox RD, Neuschwander-Tetri BA. Differences in the degree of cerulein-induced chronic pancreatitis in C57BL/6 mouse substrains lead to new insights in identification of potential risk factors in the development of chronic pancreatitis. Am J Pathol 2013; 183: 692–708. [PubMed: 23845568]

[20]. Farkona S, Soosaipillai A, Filippou P, Korbakis D, Serra S, Rückert F, et al. Generation of monoclonal antibodies and development of an immunofluorometric assay for the detection of CUZD1 in tissues and biological fluids. Clin Biochem 2017; 50: 1168–74. [PubMed: 28746862]

[21]. Bork P, Beckmann G. The CUB domain. A widespread module in developmentally regulated proteins. J Mol Biol 1993; 231: 539–545. [PubMed: 8510165]

[22]. Gaboriaud C, Gregory-Pauron L, Teillet F, Thielens NM, Bally I, Arlaud GJ. Structure and properties of the Ca^{2+}-binding CUB domain, a widespread ligand-recognition unit involved in major biological functions. Biochem J 2011; 439: 185–193. [PubMed: 21954942]

[23]. Plaza S, Chanut-Delalande H, Fernandes I, Wassarman PM, Payre F. From A to Z: apical structures and zona pellucida-domain proteins. Trends Cell Biol 2010; 20: 524–532. [PubMed: 20598543]

[24]. Mapes J, Li Q, Kannan A, Anandan L, Laws M, Lydon JP, et al. CUZD1 is a critical mediator of the JAK/STAT5 signaling pathway that controls mammary gland development during pregnancy. PLoS Genet 2017; 13: E1006654.

[25]. Mapes J, Anandan L, Li Q, Neff A, Clevenger CV, Bagchi IC, et al. Aberrantly high expression of the CUB and zona pellucida-like domain-containing protein 1 (CUZD1) in mammary epithelium leads to breast tumorigenesis. J Biol Chem 2018; 293: 2850–2864. [PubMed: 29321207]

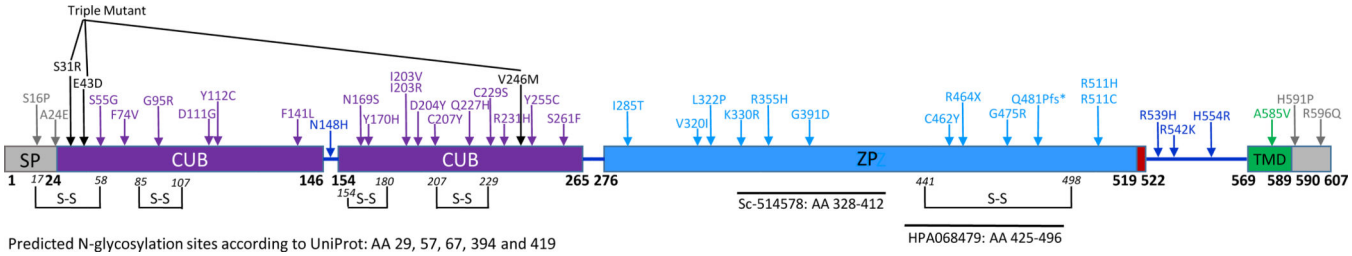

Figure 1. Domain structure of CUZD1 and localization of the detected variants

AA: amino acid

SP: signal peptide (AA 1–24); CUB: complement C1r/C1s, Uegf, Bmp1 domain (AA 25–146 and AA 154–265); ZP: zona pellucida domain (AA 276–519); TMD: transmembrane domain (AA 569–589). The red box indicates the furin cleavage site (RSKR) (AA 519–522). S-S depicts predicted disulfide bonds according to UniProt. Fragments used for the generation of the two commercial antibodies (sc-514578 and HPA068479) are also shown.

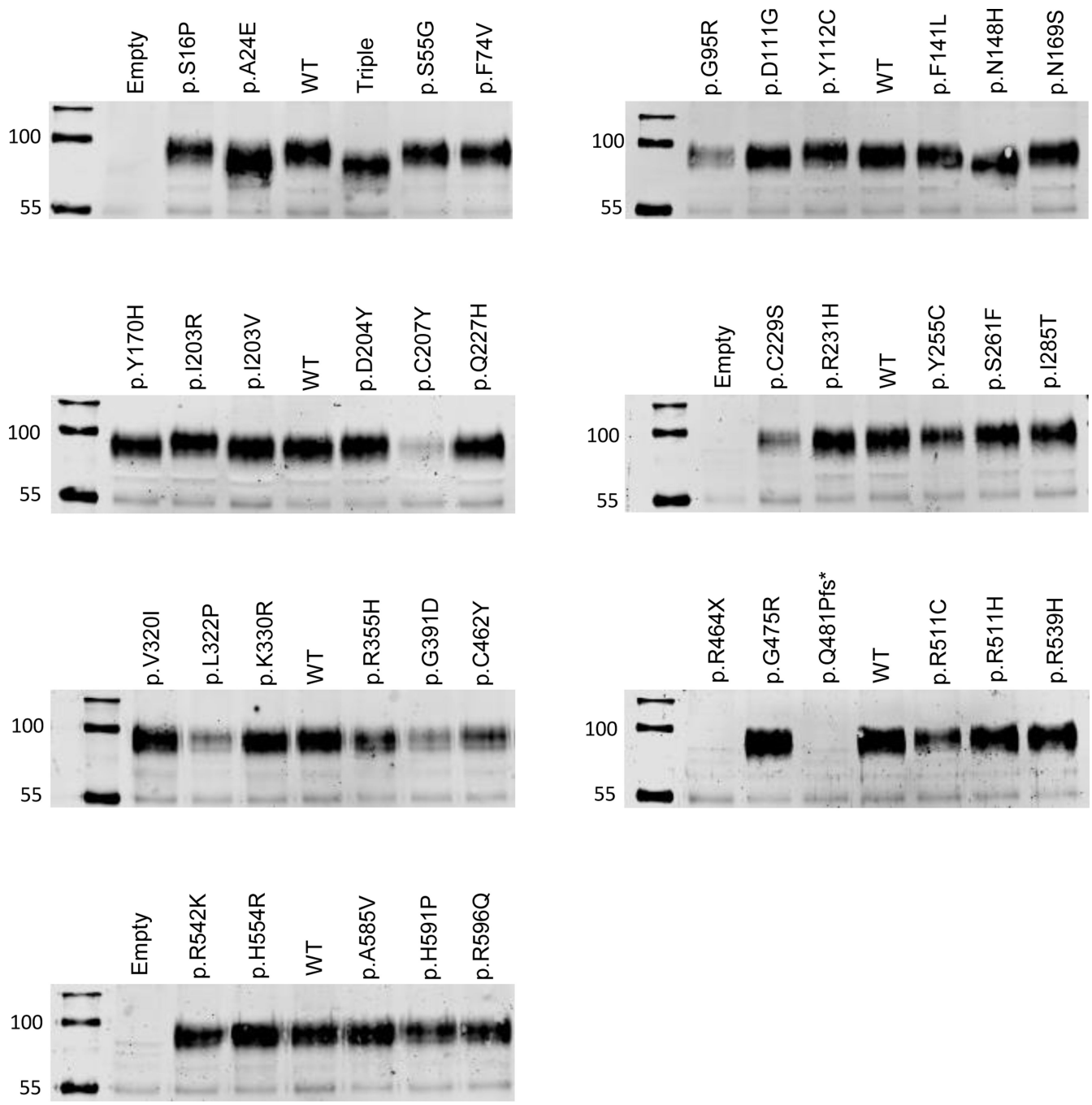

Figure 2. Effect of CUZD1 missense and nonsense variants on CUZD1 secretion

Wild-type and mutated CUZD1 constructs were expressed in HEK 293T cells. CUZD1 secretion was analyzed by Western blotting using the rabbit anti-CUZD1 polyclonal antibody HPA068479 (two biological replicates). Representative Western blots are shown. Uncropped Western blot images can be found in Supplementary Figure 2. The numbers 55 and 100 indicate the size of the protein standard in kDa.

Empty: pcDNA3.1(−) vector without insert; WT: wild-type; Triple: p.S31R_p.E43D_p.V246M.

Table 1.

Non-synonymous *CUZD1* variants in European subjects with non-alcoholic CP and controls

Exon	Nucleotide change	Amino acid change	rs number	SIFT	CADD	PROVEAN	PredictSNP	Secretion (% of WT)	CP (n=1,163) n (%)	Controls (n=2,018) n (%)	*p*-value	OR	95% CI
1	c.46T>C	p.S16P	*rs1003444066*	0.33	8.6	Neutral	Neutral	79	0	1 (0.05)	1.0		
1	c.71C>A	p.A24E	*rs565610081*	0.15	11.1	Neutral	Neutral	128	1 (0.09)	0	0.37		
2	**c.220T>G**	**p.F74V**	**-**	**0**	**23.8**	**Deleterious**	**Deleterious**	**83**	**1 (0.09)**	**0**	**0.37**		
3	**c.283G>A**	**p.G95R**	***rs764182845***	**0.01**	**21.5**	**Deleterious**	**Deleterious**	**49**	**0**	**1 (0.05)**	**1.0**		
3	c.335A>G	p.Y112C	*rs148941738*	0.14	11.9	Neutral	Neutral	78	0	1 (0.05)	1.0		
3	c.421T>C	p.F141L	-	0.13	23.7	Neutral	Neutral	77	0	1 (0.05)	1.0		
3	c.442A>C	p.N148H	-	0.1	7.4	Neutral	Neutral	97	0	1 (0.05)	1.0		
4	c.506A>G	p.N169S	*rs141623717*	0.17	22.7	Neutral	Neutral	111	2 (0.17)	3 (0.15)	1.0		
4	**c.508T>C**	**p.Y170H**	**-**	**0**	**26.1**	**Deleterious**	**Deleterious**	**112**	**0**	**1 (0.05)**	**1.0**		
5	**c.608T>G**	**p.I203R**	***rs148411775***	**0.02**	**16.3**	**Neutral**	**Neutral**	**89**	**2 (0.17)** [&]	**1 (0.05)**	**0.28**		
5	**c.610G>T**	**p.D204Y**	***rs747068473***	**0**	**23.7**	**Deleterious**	**Deleterious**	**91**	**0**	**1 (0.05)**	**1.0**		
5	**c.620G>A**	**p.C207Y**	**-**	**0**	**23.8**	**Deleterious**	**Deleterious**	**18**	**1 (0.09)**	**1 (0.05)**	**1.0**		
5	**c.681A>T**	**p.Q227H**	***rs200891112***	**0.04**	**23.4**	**Neutral**	**Neutral**	**110**	**1 (0.09)** [§]	**0**	**0.37**		
5	**c.686G>C**	**p.C229S**	**-**	**0**	**25.8**	**Deleterious**	**Deleterious**	**41**	**1 (0.09)**	**0**	**0.37**		
5	c.692G>A	p.R231H	*rs144483251*	0.51	17.3	Neutral	Neutral	94	3 (0.26)[$]	10 (0.5)	0.4		
5	**c.764A>G**	**p.Y255C**	**-**	**0**	**27.1**	**Deleterious**	**Deleterious**	**62**	**1 (0.09)**	**0**	**0.37**		
5	**c.782C>T**	**p.S261F**	**-**	**0.44**	**23.8**	**Deleterious**	**Deleterious**	**98**	**0**	**1 (0.05)**	**1.0**		
6	c.854T>C	p.I285T	*rs36212072*	0.21	22.8	Neutral	Neutral	88	13 (1.11)[%]	27 (1.35)	0.63		
6	c.958G>A	p.V320I	*rs201994934*	1	11.0	Neutral	Neutral	95	2 (0.17)	1 (0.05)	0.28		
6	**c.965T>C**	**p.L322P**	***rs762907320***	**0**	**25.4**	**Deleterious**	**Deleterious**	**31**	**1 (0.09)**	**2 (0.1)**	**1.0**		
6	**c.989A>G**	**p.K330R**	***rs759790848***	**0.06**	**24.9**	**Neutral**	**Neutral**	**95**	**0**	**1 (0.05)**	**1.0**		
7	**c.1064G>A**	**p.R355H**	***rs188322992***	**0.01**	**23.7**	**Deleterious**	**Deleterious**	**61**	**2 (0.17)**	**0**	**0.13**		
7	**c.1172G>A**	**p.G391D**	**-**	**0.05**	**23.9**	**Deleterious**	**Deleterious**	**27**	**1 (0.09)**	**0**	**0.37**		
8	**c.1385G>A**	**p.C462Y**	***rs1242843548***	**0**	**29.1**	**Deleterious**	**Deleterious**	**52** [*]	**3 (0.26)**	**0**	**0.049**	**∞**	**NaN–∞**
8	c.1423G>C	p.G475R	*rs1035904361*	0.56	14.4	Neutral	Neutral	105	1 (0.09)	0	0.37		
8	**c.1442insC**	**p.Q481Pfs** [*]	***rs770045215***	**-**	**-**	**-**	**-**	**1**	**1 (0.09)**	**0**	**0.37**		

Exon	Nucleotide change	Amino acid change	rs number	SIFT	CADD	PROVEAN	PredictSNP	Secretion (% of WT)	CP (n=1,163) n (%)	Controls (n=2,018) n (%)	*p*-value	OR	95% CI
8	**c.1531C>T**	**p.R511C**	***rs373802557***	**0**	**32.0**	**Deleterious**	**Deleterious**	**50**	**2 (0.17)**	**0**	**0.13**		
9	c.1661A>G	p.H554R	*rs61730786*	0.48	0.1	Neutral	Neutral	123	1 (0.05)	0	0.37		
9	c.1754C>T	p.A585V	*rs948370801*	0.48	7.6	Neutral	Neutral	103	0	2 (0.1)	0.54		
9	**c.1772A>C**	**p.H591P**	***rs1238211350***	**0.01**	**20.7**	**Neutral**	**Deleterious**	**72**	**0**	**1 (0.05)**	**1.0**		
Subjects carrying predicted deleterious *CUZD1* variants (%)													
	SIFT								17 (1.5)	8 (0.4)	0.002	3.7	1.6–8.7
	CADD score ≥25								9 (0.8)	3 (0.2)	0.012	5.2	1.4–19.4
	PROVEAN								14 (1.2)	7 (0.4)	0.006	3.5	1.4–8.7
	PredictSNP								14 (1.2)	8 (0.4)	0.013	3.1	1.3–7.3
	All (SIFT-CADD25-PROVEAN-PredictSNP)								9 (0.8)	3 (0.2)	0.012	5.2	1.4–19.4
	All (SIFT-CADD20-PROVEAN-PredictSNP)								14 (1.2)	6 (0.3)	0.004	4.1	1.6–10.7
Secretion <50%									5 (0.4)	4 (0.2)	0.29	2.4	0.6–8.9

CUZD1 variants with a deleterious prediction or less than 50% secretion are in bold. *P*-values were determined by two-tailed Fisher's exact test.

Percent secretion are means of two densitometric measurements from two independent Western blots.

*Reduced secretion of p.C462Y is probably due to weaker binding of antibody HPA068479 and could not be confirmed with a second antibody (sc-514578).

&One patient was homozygous for *SPINK1* p.N34S

$This patient was heterozygous for *CFTR* p.S1235R

$One patient was heterozygous for *SPINK1* p.N34S and one patient heterozygous for *TRPV6* p.L299Q

%One patient each was heterozygous for *SPINK1* c.194+2T>C, one for *CFTR* p.F508del and one for *TRPV6* p.L299Q and the *CEL-HYB1* hybrid allele

Table 2.

Non-synonymous *CUZD1* variants in Japanese subjects with non-alcoholic CP and controls

Exon	Nucleotide change	Amino acid change	rs number	SIFT	CADD	PROVEAN	PredictSNP	Secretion (% of WT)	CP (n=297) n (%)	Controls (n=1,070) n (%)	*p*-value
2	c.93C>G	p.S31R *	*rs34246902*	0.25	13.0	Neutral	Neutral	88 *	30 (10.1)	87 (8.13)	0.29
2	c.129G>T	p.E43D *	*rs35299879*	0.63	9.2	Neutral	Neutral	88 *	30 (10.1)	87 (8.13)	0.29
2	**c.163A>G**	**p.S55G**	***rs141881176***	**0.02**	**22.1**	**Neutral**	**Neutral**	**97**	**7 (2.36)**	**18 (1.68)**	**0.46**
3	c.332A>G	p.D111G	*rs143659470*	0.1	20.9	Neutral	Neutral	95	0	1 (0.09)	1.0
5	c.607A>G	p.I203V	*rs762712325*	1	0.1	Neutral	Neutral	93	0	3 (0.28)	0.36
5	**c.620G>A**	**p.C207Y**	**-**	**0**	**23.8**	**Deleterious**	**Deleterious**	**18**	**1 (0.34)**	**0**	**0.22**
5	c.692G>A	p.R231H	*rs144483251*	0.51	17.3	Neutral	Neutral	94	1 (0.34)	1 (0.09)	0.39
5	c.736G>A	p.V246M *	*rs117062933*	0.11	22.1	Neutral	Neutral	88 *	30 (10.1)	87 (8.13)	0.29
8	**c.1390C>T**	**p.R464X**	***rs200662445***	**-**	**35.0**	**-**	**-**	**2**	**0**	**2 (0.19)**	**1.0**
8	**c.1532G>A**	**p.R511H**	***rs201005450***	**0**	**31.0**	**Deleterious**	**Deleterious**	**84**	**0**	**2 (0.19)**	**1.0**
8	**c.1616G>A**	**p.R539H**	***rs34962077***	**0.05**	**22.9**	**Neutral**	**Neutral**	**89**	**1 (0.34)**	**0**	**0.22**
8	c.1625G>A	p.R542K	*rs144406058*	0.32	15.0	Neutral	Neutral	96	3 (1.01)	3 (0.28)	0.12
9	c.1787G>A	p.R596Q	*rs200173778*	0.37	11.8	Neutral	Neutral	106	0	1 (0.09)	1.0
Subjects carrying predicted deleterious *CUZD1* variants (%)											
	SIFT								9 (3.0)	22 (2.1)	0.38
	CADD score ≥25								0	4 (0.4)	0.58
	PROVEAN								1 (0.3)	4 (0.4)	1.0
	PredictSNP								1 (0.3)	4 (0.4)	1.0
	All (SIFT-CADD25-PROVEAN-PredictSNP)								0	4 (0.4)	0.58
	All (SIFT-CADD20-PROVEAN-PredictSNP)								1 (0.3)	2 (0.2)	0.52
Secretion <50%									1 (0.3)	2 (0.2)	0.52

* These three variants are in complete linkage disequilibrium. Two controls were homozygous. Secretion value of the triple mutant.

CUZD1 variants with a deleterious prediction or less than 50% secretion are in bold. *P-values* were determined by two-tailed Fisher's exact test.

Percent secretion are means of two densitometric measurements from two independent Western blots.