A recombined allele of the lipase gene CEL and its pseudogene CELP confers susceptibility to chronic pancreatitis

Karianne Fjeld1,2, Frank Ulrich Weiss3,25, Denise Lasher4,5,25, Jonas Rosendahl6,25, Jian-Min Chen7–9,25, Bente B Johansson1,2, Holger Kirsten10–12, Claudia Ruffert6,13,14, Emmanuelle Masson7,15, Solrun J Steine1,16, Peter Bugert17, Miriam Cnop18,19, Robert Grützmann20, Julia Mayerle3, Joachim Mössner6, Monika Ringdal1,21, Hans-Ulrich Schulz21, Matthias Sendler3, Peter Simon3, Paweł Sztromwasser1,2,22, Janniche Torsvik1,2, Stefan Johansson1,2 & Anders Molven1,16,24

Carboxyl ester lipase is a digestive pancreatic enzyme encoded by the CEL gene1. Mutations in CEL cause maturity-onset diabetes of the young as well as pancreatic exocrine dysfunction2. Here we describe a hybrid allele (CEL-HYB) originating from a crossover between CEL and its neighboring pseudogene, CELP. In a discovery series of familial chronic pancreatitis cases, we observed CEL-HYB in 14.1% (10/71) of cases compared to 1.0% (5/478) of controls (odds ratio (OR) = 15.5; 95% confidence interval (CI) = 5.1–46.9; P = 1.3 × 10^-6 by two-tailed Fisher’s exact test). In three replication studies of nonalcoholic chronic pancreatitis, we identified CEL-HYB in a total of 3.7% (42/1,122) cases and 0.7% (30/4,152) controls (OR = 5.2; 95% CI = 3.2–8.5; P = 1.2 × 10^-11; formal meta-analysis). The allele was also enriched in alcoholic chronic pancreatitis. Expression of CEL-HYB in cellular models showed reduced lipolytic activity, impaired secretion, prominent intracellular accumulation and induced autophagy. These findings implicate a new pathway distinct from the protease-antiprotease system of pancreatic acinar cells in chronic pancreatitis.

Carboxyl ester lipase (encoded by CEL, MIM114840) is mainly expressed in the acinar cells of the pancreas and the lactating mammary gland1,3. The enzyme is activated by bile salts in the duodenum and participates in the hydrolysis and absorption of cholesterol and lipid-soluble vitamins1. The human CEL locus on chromosome 9q34.3 also includes CELP, a tandemly arranged CEL pseudogene. The main difference between the two genes is that CELP lacks exons 2–7 of CEL and contains a stop codon in its second exon (i.e., exon 8; Fig. 1a); otherwise, the two genomic sequences are highly similar4,5.

The last of the 11 CEL exons contains a variable number of tandem repeat (VNTR) region consisting of nearly identical 33-bp segments3,4. The segment number fluctuates between 7 and 23, with 16 repeats being the most common in all populations studied so far6–10. We have previously reported that single-base deletions in the CEL VNTR cause maturity-onset diabetes of the young, type 8 (MODY8, MIM609812), an autosomal dominantly inherited pancreatic disease characterized by diabetes and exocrine dysfunction accompanied by morphological changes of the gland6,11. Therefore, we surmised that the highly polymorphic CEL gene might be a candidate for influencing the risk of chronic pancreatitis, a complex, long-standing and progressive inflammatory disease resulting in gradual replacement of the pancreatic parenchyma by fibrous tissue. The clinical picture includes abdominal pain as well as exocrine and endocrine pancreatic insufficiency manifesting as malabsorption and diabetes, respectively.

1KG Jebsen Center for Diabetes Research, Department of Clinical Science, University of Bergen, Bergen, Norway. 2Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital, Bergen, Norway. 3Department of Medicine A, University Medicine Greifswald, Greifswald, Germany. 4Pediatric Nutritional Medicine, Technische Universität München (TUM), Freising, Germany. 5Else Kröner-Fresenius-Zentrum für Ernährungsmedizin (EKFZ), TUM, Freising, Germany. 6Department for Internal Medicine, Neurology and Dermatology, Division of Gastroenterology, University of Leipzig, Leipzig, Germany. 7Institut National de la Santé et de la Recherche Médicale (INSERM), U1078, Brest, France. 8Etablissement Français du Sang (EFS)-Brest, Brest, France. 9Faculté de Médecine et des Sciences de la Santé, Université de Bretagne Occidentale, Brest, France. 10Fraunhofer Institute for Cell Therapy and Immunology (IZI), Leipzig, Germany. 11Leipzig Research Center for Civilization Diseases (LIFE-Leipzig), Universität Leipzig, Leipzig, Germany. 12Institute for Medical Informatics, Statistics and Epidemiology (IMISE), Universität Leipzig, Leipzig, Germany. 13Department of Internal Medicine, Neurology and Dermatology, Division of Endocrinology, University of Leipzig, Leipzig, Germany. 14Integrated Research and Treatment Centre (IFB) Adiposity Diseases, University of Leipzig, Leipzig, Germany. 15Laboratoire de Génétique Moléculaire et d’Histocompatibilité, Centre Hospitalier Universitaire (CHU) Brest, Hôpital Morvan, Brest, France. 16Gade Laboratory for Pathology, Department of Clinical Medicine, University of Bergen, Bergen, Norway. 17Institute of Transfusion Medicine and Immunology, Medical Faculty Mannheim, Heidelberg University, German Red Cross Blood Service of Baden-Württemberg-Hessen, Mannheim, Germany. 18Université Libre de Bruxelles (ULB) Center for Diabetes Research, ULB, Brussels, Belgium. 19Division of Endocrinology, Erasmus Hospital, Brussels, Belgium. 20Department of Surgery, Universitätsklinikum Dresden, Dresden, Germany. 21Department of Surgery, Otto-von-Guericke University Magdeburg, Magdeburg, Germany. 22Computational Biology Unit, Department of Informatics, University of Bergen, Bergen, Norway. 23Department of Pediatrics, Haukeland University Hospital, Bergen, Norway. 24Department of Pathology, Haukeland University Hospital, Bergen, Norway. 25These authors contributed equally to this work. Correspondence should be addressed to A.M. (anders.molven@gades.uib.no).

Received 23 September 2014; accepted 19 February 2015; published online 16 March 2015; doi:10.1038/ng.3249
Figure 1 Copy-number variants of the human CEL gene. The large red and blue boxes without numbers correspond to the complete CEL and CELP genes, respectively. Smaller boxes represent the individual exons and are numbered 1–11 (CEL) and 1′ and 11′ (CELP). The stop codon in exon 1′ of CELP is indicated. (a) Structure of the CEL-CELP gene locus. Gray shading marks the CEL exon 2–exon 7 region missing in CELP. The size of the locus is shown below the figure; exon and intron sizes are not drawn to scale. (b) Schematic structure of the duplication hybrid allele and the deletion hybrid allele (CEL-HYB) identified in this study. The breakpoint regions (not drawn to scale) are indicated by asterisks. (c) Proposed mechanism of the generation of CEL-HYB by nonallelic homologous recombination between CEL and CELP. The X symbolizes the crossover event in the exon 10–exon 11 region.

The constellation of a gene and a nearby homologous sequence is well known to enhance the probability of genomic rearrangements, and it can sometimes lead to the formation of new fusion genes that have altered functional properties and are associated with disease. Copy-number variants of CEL have been noted, both by others and by us.9,10,13 Fine mapping of the CEL locus in a subject known to carry an extra copy of CEL exon 11 identified the presence of a duplication hybrid allele (Fig. 1b). This allele apparently resulted from nonallelic homologous recombination (Fig. 1c), with the crossover occurring within an identical region of 522 bp in introns 9 of CEL and CELP (positions 12780–13301 in CEL and 28549–29070 in CELP; GenBank AF072711.1). The portion consisting of the CELP-CEL fusion sequence was, however, unlikely to encode a functional protein by virtue of its structural similarity to CELP (Fig. 1b).

In contrast, the reciprocal recombination product of the duplication hybrid allele, namely a deletion hybrid allele, would potentially encode a new CEL-CELP fusion protein, on the basis of its predicted gene structure (Fig. 1b,c). We developed a long-range PCR assay to detect this allele (Supplementary Fig. 1), and we found one positive subject among 190 Norwegian blood donors. Subsequent DNA sequencing showed that the identified deletion hybrid allele (denoted CEL-HYB) was not exactly reciprocal to the aforementioned duplication allele; the crossover occurred instead in a 548-bp identical region covering intron 10 and the adjacent exon boundaries (Supplementary Fig. 2a). This might indicate that different recombination events have taken place in the intron 9–exon 11 region of the CEL-CELP locus.

We decided to investigate the role of CEL-HYB in the genetic etiology of chronic pancreatitis. Screening a discovery cohort of 71 German subjects with familial chronic pancreatitis (Table 1) identified a carrier frequency of 14.1% (10/71) compared with 1.0% (5/478) in controls (OR = 15.5; 95% CI = 5.1–46.9; P = 1.3 × 10−6). To replicate this finding, we screened three independent cohorts (two German, one French) of subjects with nonalcoholic chronic pancreatitis as well as controls from each contributing laboratory. In each cohort, CEL-HYB was significantly overrepresented among individuals with chronic pancreatitis (Table 1). For the three replication cohorts combined, CEL-HYB was present in 3.7% (42/1,122) of the case group and 0.7% (30/4,152) of the control group. Formal meta-analysis of the replication cohorts was highly significant (OR = 5.2; 95% CI = 3.2–8.5; P = 1.2 × 10−11). Among the 1,122 case subjects, there were 57 individuals carrying rare PRSS1 variants known to associate with hereditary chronic pancreatitis. All PRSS1 carriers tested negative for CEL-HYB. If these carriers were excluded from the study, the combined CEL-HYB frequency for cases increased to 3.9%, and in a meta-analysis of the three replication cohorts, the OR was 5.5 (95% CI = 3.4–8.9; P = 2.6 × 10−12).

We also examined the frequency of CEL-HYB in individuals with alcohol-related chronic pancreatitis. Again, we found a significant enrichment of the CEL-HYB variant with a carrier frequency of 1.8% in the case group versus 0.8% in the control group (OR = 2.3; 95% CI = 1.2–4.4; P = 0.016) (Table 1). All CEL-HYB-positive individuals detected in the five cohorts studied were heterozygous carriers of the allele, except for one homozygous individual among the French patients (nonalcoholic chronic pancreatitis cohort 3).

In the discovery cohort, family members from two of the CEL-HYB-positive subjects were available for screening. Their pedigrees showed an autosomal-dominant inheritance pattern of pancreatitis, although the CEL-HYB allele did not fully cosegregate with disease (Supplementary Fig. 3). Furthermore, when analyzing the CEL-HYB variant in 228 parents of subjects from the German chronic pancreatitis cohort 2, we detected six heterozygous parents. All of them had transmitted the risk allele to their affected children. Two parents had a diagnosis of pancreatitis. These observations underline the substantial risk associated with CEL-HYB.

To exclude the possibility that alterations in exons 1–10 of CEL-HYB could explain the disease association, we performed DNA sequencing of 12 CEL-HYB–positive individuals. We observed no predicted amino

Table 1 Heterozygous carrier frequencies of the CEL-HYB variant in German and French subjects with chronic pancreatitis

| Subjectsa | Case (+/n) | Control (+/n) | OR | 95% CI | P valueb |
|-----------|-----------|--------------|----|--------|----------|
| Discovery cohort | 10/71 | 5/478 | 15.5 | 5.1–46.9 | 1.3 × 10−6 |
| Nonalcoholic CP cohort 1 | 11/260 | 4.2 | 5/569 | 0.9 | 5.0 | 1.7–14.5 | 0.002 |
| Nonalcoholic CP cohort 2 | 22/510 | 4.3 | 16/2,362 | 0.7 | 6.6 | 3.4–12.7 | 1.3 × 10−8 |
| Nonalcoholic CP cohort 3 | 9/352 | 2.6 | 9/1,221 | 0.7 | 3.5 | 1.4–9.0 | 0.009 |
| Meta-analysisc | – | – | – | – | 5.2 | 3.2–8.5 | 1.2 × 10−11 |
| Alcoholic CP cohort | 15/853 | 1.8 | 26/3,409 | 0.8 | 2.3 | 1.2–4.4 | 0.016 |

CP, chronic pancreatitis; OR, odds ratio; CI, confidence interval.

aCohorts are described in Online Methods. bTwo-tailed Fisher’s exact test. cFixed effective model (in replication cohorts only).

© 2015 Nature America, Inc. All rights reserved.
acid changes in exons 1–9. In one subject with chronic pancreatitis, however, we observed a premature stop codon in exon 10 of the CEL-HYB reading frame. The corresponding mRNA might not be translated, owing to nonsense-mediated RNA decay, and, if translated, the protein would not include the VNTR. We then screened all initially positive subjects in the case and control groups for this minor CEL-HYB variant and detected it in one case subject and four control subjects (not included in the carrier frequencies in Table 1).

To test whether the CEL-HYB risk variant might be identifiable indirectly through linkage disequilibrium (LD) with surrounding SNPs, we performed a detailed examination of genome-wide association study (GWAS) data in a subset of 787 individuals with known CEL-HYB status (25 carriers and 762 noncarriers). Proxy SNP analysis within the CEL-CELP LD region did not identify variants in strong LD with CEL-HYB. However, we found some distant correlating variants peaking at about 255 kb upstream of the CEL-HYB LD region did not identify variants in strong LD with CEL-HYB. However, we found some distant correlating variants peaking at about 255 kb upstream of the CEL-HYB locus (Fig. 2). We identified no other SNPs correlating with the CEL-HYB variant at \( r^2 \geq 0.2 \) on chromosome 9 outside the region shown in Figure 2. The best \( r^2 \) was around 0.6 for a few imputed, highly correlated SNPs with low minor-allele frequencies (Supplementary Table 1). Moreover, to evaluate the origin of CEL-HYB, we sequenced the breakpoint region between exons 10 and 11 in 74 carriers of the allele. We observed five CEL-HYB–associated haplotypes (Supplementary Fig. 2b), a result indicating that the variant has arisen through multiple independent recombination events. Thus, it is unlikely that a standard GWAS will be able to thoroughly tag and assess the CEL-HYB risk variant.

We also asked whether SNPs or rare variants within or adjacent to CEL exons 1–10 could be associated with chronic pancreatitis outside the context of the CEL-HYB allele. Sequencing of 364 subjects with nonalcoholic chronic pancreatitis cases and 238 control subjects identified no variants that were overrepresented in the chronic pancreatitis cohort (Supplementary Table 2).

The CEL-HYB allele was predicted to encode a chimeric CEL protein of 589 amino acids containing a functional enzyme region derived from CEL but with a shorter, different VNTR originating from CELP (Fig. 3a,b and Supplementary Fig. 4a). To examine the functional consequences of CEL-HYB, we stably expressed the CEL-HYB protein in human embryonic kidney (HEK) 293 cells, which lack detectable endogenous expression of CEL mRNA or protein\(^{14,15}\). We found that secretion of the CEL-HYB protein was reduced relative to the wild-type protein (CEL-WT) (Fig. 3c). We observed a similar result in transiently transfected mouse acinar (266-6) cells (Supplementary Fig. 4b).

Next, we investigated whether CEL-HYB might encode an overactive lipase. We measured carboxy ester lipase activity in conditioned medium from the CEL-expressing HEK293 cells, normalizing
specific activity to the amount of CEL in the growth medium. When stimulated by bile salts, the CEL-HYB enzyme reached only about 40% of CEL-WT activity (Fig. 3d). An artificial truncated variant of CEL (CEL-TRUNC) lacking the whole VNTR showed about 90% of wild-type enzyme activity, a finding also reported by others. We therefore propose that the impaired activity of CEL-HYB may relate to effects of the new protein C terminus rather than to the lack of a normal VNTR. Further, our results strongly suggest that increased lipase activity is not involved in the disease mechanism.

Finally, we examined the intracellular fate of the protein. When we inhibited protein synthesis with cycloheximide, CEL-HYB accumulated in the cells to a much higher degree than CEL-WT. After 120-min exposure to cycloheximide, 50% of CEL-HYB protein was detected in the cell lysate compared with 9% for CEL-WT (Fig. 4a,b). This intracellular retention of CEL-HYB as well as its impaired secretion (Fig. 4c) could indicate that the pathological effect involves injury inside the cell. We hypothesized that intracellular CEL-HYB might induce compensatory autophagy for removal of defective and excessive secretory proteins. When grown in either rich or starvation medium, CEL-HYB–expressing cells (Fig. 4c) may relate to effects of the new protein C terminus rather than to the lack of a normal VNTR. Further, our results strongly suggest that increased lipase activity is not involved in the disease mechanism.

Finally, we note that classical molecular genetic techniques (long-range PCR, cloning and Sanger DNA sequencing) were needed to map and analyze the genomic rearrangements of the CEL-CELP locus. Our detailed SNP analysis of the CEL-CELP region (Fig. 2) indicates that current high-throughput methodologies, such as GWAS of common variants and exome sequencing, would be unlikely to uncover CEL-HYB and its role in pancreatitis. When genetic risk loci involve complex and highly repetitive regions, these techniques have some

Figure 4 In intracellular properties of CEL-HYB. (a) Reduced cellular clearance of CEL-HYB. Top, western blotting for CEL and β-actin. Samples are stably transfected HEK293 cells expressing CEL-WT or CEL-HYB, treated with the protein-synthesis inhibitor cycloheximide (1 μg/ml) and examined at the indicated time points. Bottom, quantification of bands, with CEL level normalized to β-actin. CEL expression at time 0 was arbitrarily set to 1.0. (b) Immunostaining and confocal imaging cells expressing CEL-HYB or CEL-WT, before and after 120 min of cycloheximide treatment. Blue, 4′,6-diamidino-2-phenylindole (DAPI) staining; green, CEL staining. Scale bar, 10 μm. (c) Autophagy induced by CEL-HYB expression. Top, western blotting for CEL, β-actin and the autophagy marker protein LC3. Samples are stably transfected HEK293 cells incubated in rich (fed) or starvation medium with or without bafilomycin A1 (BafA1) for 3 h. Bottom, quantification of bands, with LC3-II normalized to β-actin. Throughout figure, data are shown as means of three independent experiments ± s.e.m.; *P = 0.01–0.05 by two-tailed unpaired t-test.
inherent limitations, as recently illustrated by the failure to detect mutations in the VNTR of MUC1 by massively parallel sequencing. The identification of CEL-HYB as a new pancreatitis susceptibility gene thus nicely illustrates how in-depth analysis of the more complex regions of the genome can play an important part in elucidating the genetic etiology of human disorders.

METHODS

Methods and any associated references are available in the online version of the paper.

ACKNOWLEDGMENTS

The authors thank all study participants and the members of the Gesellschaft für Padiatrische Gastroenterologie und Ernährung (GPGE) for providing clinical data and blood samples. The confocal imaging was performed at the Molecular Imaging Center, Department of Biomedicine, University of Bergen. We are also grateful to G. Bierey for commenting on the manuscript. This work was supported by grants and fellowships to P.R.N. and A.M. from the Translational Fund of Bergen Medical Research Foundation, KG Jebsen Foundation, University of Bergen, Research Council of Norway and Western Norway Regional Health Authority (Helse Vest), and to P.R.N. from the European Research Council. Work performed by the German Federal Ministry of Education and Research (BMBF GANI-MED 03152061A and BMBF 0314107), European Union Framework Programme 7 (T-BAL, a truncated form of the enzyme, expressed in Escherichia coli), the German Federal Ministry of Education and Research Council of Norway and Western Norway Regional Health Authority (Helse Vest), and to P.R.N. from the European Research Council. We are also grateful to all study participants and the members of the Gesellschaft für Padiatrische Gastroenterologie und Ernährung (GPGE) for providing clinical data and blood samples.
ONLINE METHODS

Subjects. The medical ethical review committees of all participating study centers approved this study. Informed consent was obtained from all subjects. For the discovery phase, we screened 71 unrelated German subjects with a diagnosis of familial nonalcoholic chronic pancreatitis, recruited by the University of Greifswald. These subjects had unknown genetic etiology because they had previously tested negative for well-established pancreatitis risk genes. For the replication studies, we investigated 1,122 unrelated non-alcoholic subjects with chronic pancreatitis from the University of Leipzig (cohort 1; \( n = 260 \)), Technische Universität München (cohort 2; \( n = 510 \)) and University of Bretagne, Brest (cohort 3; \( n = 352 \)). All subjects in the replication cohorts had been screened for rare PRSSI variants associated with hereditary chronic pancreatitis as well as for the SPINK1 risk variant encoding aPn34Ser. The three German subject cohorts were compared to exclude any duplicate samples. In addition, we investigated 853 unrelated German subjects with alcoholic chronic pancreatitis, from Greifswald (\( n = 214 \)), Leipzig (\( n = 296 \)) and Munich (\( n = 343 \)). Controls were healthy blood donors (Greifswald, \( n = 478 \), Leipzig, \( n = 569 \); Munich, \( n = 2,362 \); and Brest, \( n = 1,221 \)). Norwegian blood donors (\( n = 190 \)) were used for the initial identification of CEL-HYB.

The diagnosis of chronic pancreatitis was based on the presence of a typical history of recurrent pancreatitis (two or more acute pancreatitis attacks) and/or other typical findings upon imaging, such as pancreatic calcifications. Familial chronic pancreatitis (discovery cohort) was diagnosed when two or more family members in more than one generation suffered from pancreatitis. The subjects in the replication cohorts were classified as having nonalcoholic chronic pancreatitis when common risk factors such as alcohol abuse or other predisposing disorders were absent. Subjects who reported a history of regular alcohol consumption of at least 60 g per day (women) or 80 g per day (men) were included in the alcoholic chronic pancreatitis cohort.

Reference sequences. The following reference sequences have been used for primer design, sequence verification and SNP analysis of the CEL-CELP locus: GenBank AF072711.1 (Homo sapiens carboxyl ester lipase gene, complete cds; and carboxyl ester lipase pseudogene, complete sequence); Ensembl ENSG00000170835 (CEL) and ENSG00000170827 (CELP).

Characterization of the CEL duplication hybrid allele. The presence of a CEL duplication was originally inferred by the DNA fragment analysis method of Torvik et al.\(^{12}\). Its allelic structure was worked out by PCR walking with \( LA \) Taq polymerase (TakaRa Bio), as described by the manufacturer. For mapping of the CEL-CELP recombinant site, PCR products were verified by agarose gel electrophoresis, cloned into the pCR2.1-TOPO vector (Invitrogen) and analyzed by sequencing.

Screening for the CEL deletion hybrid allele. A long-range, duplex PCR assay was developed for the postulated deletion hybrid allele (CEL-HYB) (Supplementary Fig. 1). We performed the reactions in total volumes of 10 µl containing 1× GC buffer, 0.4 µM of each of the primers L11F, IAR and CELP-VNTR-R (Supplementary Table 3), 1 M betaine solution, 0.4 µM of each dNTP, 0.5 U \( LA \) Taq polymerase and 10–50 ng genomic DNA. Amplification started with a denaturation step at 94 °C for 1 min; this was followed by 14 cycles of 94 °C for 20 s and 60 °C for 6 min; 16 cycles of 94 °C for 20 s and 62 °C for 6 min; a final elongation step of 72 °C for 10 min; and cooling to 4 °C. After visualization on a 1% agarose gel, all CEL-HYB–positive samples were finally verified by Sanger DNA sequencing.

For screening of CEL-HYB in larger samples, we developed a high-capacity assay on the LightCycler 480 platform (Roche Diagnostics). We performed PCR with forward and reverse primers (Supplementary Table 3) at a ratio of 5:1 (0.5 µmol/l and 0.1 µmol/l, respectively), 0.75 U OneTaq polymerase (New England BioLabs) and 0.4 µM of each dNTP, in a total volume of 25 µl. The amplification cycling conditions were as follows: initial denaturation for 12 min at 95 °C; 40 cycles of 20 s denaturation at 95 °C, 40 s of annealing at 60 °C and 90 s of primer extension at 72 °C; and a final extension for 2 min at 72 °C. The amplified PCR products were then tested by melting-curve analysis on a LightCycler 480. We used a SimpleProbe (TIB MOLBIOL) that allowed for discrimination between the CEL and CELP sequences by creating different melting points (Supplementary Table 3; representative image of screening results from the LightCycler assay in Supplementary Fig. 3).

Because the LightCycler assay was based on discrimination of SNPs in the breakpoint region, some false positives were detected. All samples positive in the LightCycler assay were therefore subjected to the long-range PCR method above and sequenced for final verification. This approach was used when screening case and control subjects recruited by the Leipzig and Munich groups. The other cohorts underwent primary screening directly by long-range PCR. The genotyping concordance for the long-range PCR assay was 100% for 533 duplicates tested.

Sanger DNA sequencing. Sequencing was performed by standard protocols on an ABI 3730 capillary sequencer (Applied Biosystems) with M13 forward or reverse primers, or CEL- and CELP-specific primers (Supplementary Table 3).

When sequencing PCR products directly, they were treated with ExoSAP (USB Corporation) as described by the manufacturer, and 2 µl of the treated PCR products was used as a template. Sequencing results were analyzed by FinchTV (see URLs) or SeqScape (Applied Biosystems) software.

SNP mapping of the CEL-CELP region. Analysis of linkage disequilibrium between the CEL-HYB variant and SNPs on chromosome 9 was done in subjects with pancreatitis, with available genetic data from Illumina Omni Human Express 700K SNP chips. Genotypes of 2,824 subjects were called with Illumina GenomeStudio. Quality control was done with PLINK. Thereby, the data were filtered for missingness per marker ≤ 0.05, missingness per individual ≤ 0.05, Hardy-Weinberg equilibrium \( P \) value ≤ 10\(^{-6} \), cryptic relatedness filter \( P_L\text{HAT} ≤ 0.185 \), autosomal heterozygosity within 3 s.d. and population stratification not more extreme than ± 3 s.d. After quality control, we used 2,712 subjects (2,065 males, 630 females and 17 of unknown sex) and 32,138 SNPs from chromosome 9 for imputation. Imputation was done with SHAPEIT v2 and IMPUTE 2.3.0, with application of the 1000 Genomes reference phase 1, release 3 (see URLs). Blocks used in imputing were chosen to include the CEL-CEL-P locus 22 Mb in a single block. For post-imputation quality control, we filtered for SNPs with info score <0.3.

For all further analyses, the cohort was limited to individuals with known CEL-HYB status, i.e., 787 individuals. Among these, 25 (3.2%) were carriers of the CEL-HYB allele. To identify proxies for CEL-HYB, we followed several strategies. First, we calculated Pearson’s correlation of the CEL-HYB variant with allelic doses from imputed and directly measured genotypes (‘gene dose’ analysis). Additionally, we calculated haplotype-based \( r^2 \) of directly typed SNPs and best-guess imputed genotypes. Here, we set all imputed genotypes with posterior probability smaller than 90% to missing. Confidence intervals of haplotype-based \( r^2 \) were calculated as recommended\(^{36} \), with jackknife methodology applied to estimate standard errors. Aggressive two- and three-marker tagging was done with Tagger as implemented in Haploviev 4.2 on measured and best-guess imputed genotypes.

**CEL-HYB plasmid construction.** Human CEL wild-type (CEL-WT) cDNA was inserted into the pcDNA3.1/V5-His vector (Invitrogen). For the enzyme activity experiment, cDNA of a truncated variant of human CEL (CEL-TRUNC), lacking the VNTR after nucleotide 1686, was inserted into the pcDNA3.1/V5-His vector. Both CEL variants were cloned in frame with the V5-His tag of the pcDNA3.1/V5-His vector as previously described\(^{14,15} \).

We used the CEL-WT/V5/His plasmid as template for creating the CEL-HYB construct. BamHI and XhoI restriction sites flanked the VNTR sequence of the wild-type plasmid. A synthetic VNTR sequence of the CEL-HYB allele, flanked by the same restriction sites, was ordered (DNA 2.0). The VNTR of the template was then replaced by the CEL-HYB VNTR to create a CEL-HYB/V5/His plasmid. The C substitution (rs77696629) observed in exon 10 of all CEL-HYB–positive subjects (Supplementary Fig. 2a) was introduced with the QuikChange II XL Site-Directed Mutagenesis Kit (Stratagene). The final structure of the CEL-HYB construct was confirmed by DNA sequencing. Before use in cellular experiments, the plasmid was tested in vitro by expression in the TNT T7 Quick-Coupled Transcription/Translation System (Promega) (Supplementary Fig. 4a).
Cell cultures and transfection. HEK293 cells (Clontech) were cultured in α-minimal essential medium (Sigma) supplemented with 10% FBS (Invitrogen). Hank’s Balanced Salt Solution (PAA, GE Healthcare) was used in starvation experiments. The 266-6 mouse pancreatic acinar cell line (ATCC) was grown in Dulbecco’s Modified Eagle’s Medium (ATCC) containing 10% FBS (Invitrogen) and 1× Antibiotic-Antimycotic. Both cell lines were maintained at 37 °C in a 5% CO₂-humidified atmosphere.

Transient transfection of CEL-WT-V5/His, CEL-HYB-V5/His and the pcDNA3.1/V5-His empty vector in 266-6 cells was performed in six-well plates with Metafectene Pro (Biontex) with a lipid (µl) to DNA (µg) ratio of 1:1. HEK293 cells stably transfected with CEL-WT-V5/His, CEL-TRUNC-V5/His or the pcDNA3.1/V5-His vector were available in the laboratory14,15. For stable transfection of the CEL-HYB-V5/His plasmid, HEK293 cells were transfected with the FuGENE 6 Transfection Reagent (Roche Diagnostics) with a 3:1 ratio of transfection reagent (µl) to DNA (µg). Cells were then grown under the selection of 500 µg/ml geneticin 418 (Invitrogen) for 18 d. Protein blotting and immunostaining confirmed expression of the CEL-HYB protein. All stable cell lines were maintained under constant genetin selection.

Preparation of cell medium and lysate. After transient transfection (48 h), 1 ml conditioned medium was isolated and centrifuged for 5 min at 13,000 r.p.m. and 4 °C. The supernatant was collected and analyzed. Total cell lysate was prepared by washing the cells in PBS before addition of ice-cold RIPA buffer (Millipore) with 1× complete Mini, EDTA-free protease-inhibitor cocktail (Roche Diagnostics). RIPA buffer was added at 150 µl per well and incubated for 30 min at 4 °C. The lysate was centrifuged for 15 min at 13,000 r.p.m. and 4 °C, and the supernatant was further analyzed as total cell lysate. Stably transfected cells were plated in six-well plates (3 × 10⁵ cells per well) and grown for 48 h before isolation of cell medium and lysate, as described above. Quantification of total protein in cell lysates was measured with the BCA Protein Assay (Pierce).

Western blotting. Lysates of 3 or 5 µg total protein were subjected to SDS-PAGE and western blotting. For analysis of CEL protein secretion, the volume of medium loaded on the SDS gel was the same as for the corresponding lysate. The proteins were separated by 10% SDS-PAGE at 180 V for 1.5–2 h and transferred to PVDF membranes by standard methods. The membranes were incubated with primary antibodies at 4 °C overnight, washed and incubated with secondary antibodies for 1 h at room temperature. The blots were developed with the Enhanced Chemiluminescence Plus Western Blotting Detection Kit (GE Healthcare) and further analyzed on a LAS-1000 imager (Fujifilm).

CEL activity assay. Aliquots of conditioned medium from stably transfected HEK293 cells expressing CEL-WT, CEL-TRUNC or CEL-HYB were used. HEK293 cells stably transfected with empty vector were included as a negative control. The cells were grown to 80% confluency in normal growth medium. Subsequently, the medium was substituted with serum-free medium, and the cells were incubated for additional 18 h. Conditioned medium was collected and centrifuged at 300 g and 4 °C to remove detached cells and was kept on ice for immediate use.

CEL enzyme activity was measured with 4-nitrophenyl valerate (4-npv; Sigma) as a substrate, and hydrolysis of 4-npv was monitored spectrophotometrically in the presence of sodium taurocholate (Sigma) as an activator. The assay was performed at 37 °C in a buffer containing 100 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 mM 4-npv (sulbolized in ethanol), and sodium taurocholate at concentrations varying between 0.02 and 20 mM. Reactions were run in triplicate in a 96-well microplate containing 115 µl buffer and 10 µl conditioned medium with an Infinite M200 Pro microplate reader (Tecan) with iControl Tecan Software continuously monitoring the release of para-nitrophenol at 405 nm. The molar extinction coefficient of 4-npv was 11,500 M⁻¹ cm⁻¹. Enzyme activity was normalized against the amount of CEL in the conditioned medium. CEL protein concentration was determined by SDS-PAGE and immunoblotting with titration of recombinant mouse CEL (568–CE, R&D Systems) as a standard.

Treatments of stably transfected cells. HEK293 cells were seeded in triplicate in six-well plates (3 × 10⁵ cells per well) coated with poly-L-lysine (Sigma), and they were cultured for 48 h before treatment. For investigating the rate of cellular CEL clearance during inhibition of protein synthesis, the cells were treated with 1 µg/ml cycloheximide (Sigma). Cell lysates were harvested for western blot analysis after 0, 30, 60 and 90 min of treatment. For autophagy analysis, cells were incubated in rich or starvation medium with or without bafilomycin A1 (100 nM) for 3 h; this was followed by harvesting of cell lysates for western blotting.

Immunofluorescence and confocal microscopy. Cells were seeded in 12-well plates (1.5 × 10⁵ cells per well) and grown on poly-L-lysine–coated coverslips for 48 h. Upon cycloheximide treatment, as described above, the cells were fixed for 30 min with 3% paraformaldehyde before being immunostained as described previously37. After antibody staining, the coverslips were inverted onto objective glasses in a drop of ProLong Gold antifade mounting reagent (Molecular Probes) with DAPI (Invitrogen). The samples were examined with a Leica TCS SP5 confocal microscope (Leica Microsystems) with a 63×/NA 1.4 Plan-Apochromat oil-immersion objective. The samples were examined with a Leica TCS SP5 confocal microscope (Leica Microsystems) with a 63×/NA 1.4 Plan-Apochromat oil-immersion objective, –1.2 Airy unit pinhole aperture and the appropriate filter combinations. Images were acquired with Diode 405 and 488 Argon lasers and processed with Adobe Photoshop CS5 image software (Adobe Systems). This experiment was repeated three times.

Antibodies. Protein expression in HEK293 and 266-6 cells was measured by western blotting with the mouse anti-V5 primary antibody (R960–25, Invitrogen) or a rabbit polyclonal antibody (Vank0) for CEL detection14,15. The Vanko antibody was raised against recombinant CEL lacking the VNTR region and was a generous gift from O. Hernell (Umeå University, Sweden). For detection of the autophagosome marker LC3, a rabbit polyclonal LC3B antibody (NB600-1384, Novus Biologicals) was used38. An antibody detecting β-actin (sc-1615, Santa Cruz Biotechnology) was included to control for loading39. Immunofluorescence staining of CEL was performed with the mouse monoclonal antibody Ax201, also kindly provided by O. Hernell15, AlexaFluor488-conjugated goat anti-mouse (A-11017, Invitrogen) was used as secondary antibody.

Statistical analyses. The significance of differences between allele frequencies in affected and control individuals was tested by two-tailed Fisher’s exact test, and OR and 95% CI were calculated with two-by-two table analysis in Stata/IC 13.0 (StataCorp). A fixed-effect model was used for the formal meta-analysis, which was performed by the method of Mantel and Haenszel, as implemented with the Metan version 9 module in Stata/IC 13.0, with each study site treated as a separate group. For the functional protein analyses, significance of change was analyzed with a two-tailed unpaired t-test. The difference was regarded as significant when the P value was ≤0.05. Results are given as mean ± s.e.m.

36. Scholz, M. & Hasenclever, D. Comparison of estimators for measures of linkage disequilibrium. Int. J. Biostat. 6 (Article 1) (2010).
37. Sannierud, R., Marie, M., Hansen, B.B. & Saraste, J. Use of polarized PC12 cells to monitor protein localization in the early biosynthetic pathway. Methods Mol. Biol. 457, 253–265 (2008).
38. Castino, R., Fiorentino, I., Cagnin, M., Giovia, A. & Isidoro, C. Chelation of lysosomal iron protects dopaminergic SH-SY5Y neuroblastoma cells from hydrogen peroxide toxicity by precluding autophagy and Akt dephosphorylation. Toxicol. Sci. 123, 523–541 (2011).
39. Hou, X. et al. Proteome of the porosome complex in human airway epithelia: interaction with the cystic fibrosis transmembrane conductance regulator (CFTR). J. Proteomics 96, 82–91 (2014).