No Association Between CEL-HYB Hybrid Allele and Chronic Pancreatitis in Asian Populations.

Wen-Bin Zou, Arnaud Boulling, Atsushi Masamune, Prachand Issarapu, Emmanuelle Masson, Hao Wu, Xiao-Tian Sun, Liang-Hao Hu, Dai-Zhan Zhou, Lin He, et al.

No Association Between CEL–HYB Hybrid Allele and Chronic Pancreatitis in Asian Populations

Short Title: CEL-HYB Variants and Chronic Pancreatitis

Wen-Bin Zou,[1,2,3,4,*] **Arnaud Boulling,**[2,3,5,*] **Atsushi Masamune,**[6,*] **Prachand Issarapu,**[7,*] Emmanuelle Masson,[2,8] Hao Wu,[1,4] Xiao-Tian Sun,[1,4] Liang-Hao Hu,[1,4] Dai-Zhan Zhou,[9] Lin He,[9] Yann Fichou,[2,3] Eriko Nakano,[6] Shin Hamada,[6] Yoichi Kakuta,[6] Kiyoshi Kume,[6] Hiroyuki Isayama,[10] Sumit Paliwal,[7] K Radha Mani,[7] Seema Bhaskar,[7] David N. Cooper,[11] Claude Férec,[2,3,5,8] Tooru Shimosegawa[6,#], Giriraj R Chandak,[7,12#] Jian-Min Chen,[2,3,5,#] Zhao-Shen Li,[1,4,#] and Zhuan Liao[1,4,#]

***Authors share co-first authorship.**

[#]Authors share co-senior authorship.

[1]Department of Gastroenterology, Changhai Hospital, the Second Military Medical University, Shanghai, China; [2]Institut National de la Santé et de la Recherche Médicale (INSERM), U1078, Brest, France; [3]Etablissement Français du Sang (EFS) – Bretagne, Brest, France; [4]Shanghai Institute of Pancreatic Diseases, Shanghai, China; [5]Faculté de Médecine et des Sciences de la Santé, Université de Bretagne Occidentale (UBO), Brest, France; [6]Division of Gastroenterology, Tohoku University Graduate School of Medicine, Sendai, Japan; [7]Genomic Research on Complex Diseases, CSIR-Centre for Cellular and Molecular Biology (CSIR-CCMB), Hyderabad, India; [8]Laboratoire de Génétique Moléculaire et d'Histocompatibilité, Centre Hospitalier Universitaire (CHU) Brest, Hôpital Morvan, Brest, France; [9]Key Laboratory of Developmental Genetics and Neuropsychiatric Diseases (Ministry of Education), Bio-X Institutes, Shanghai Jiao Tong University, Shanghai, China;

[10]Department of Gastroenterology, Graduate School of Medicine, the University of Tokyo, Tokyo, Japan; [11]Institute of Medical Genetics, School of Medicine, Cardiff University, Cardiff, United Kingdom; [12]Human Genetics Division, Genome Institute of Singapore, Biopolis, Singapore

Grant Support

WBZ is a joint PhD student between Changhai Hospital and INSERM U1078 and received a one-year scholarship (the year 2015) from the China Scholarship Council (No. 201403170271). Support for this study came from the National Natural Science Foundation of China (Grant Nos. 81470884, 81422010 [Z.L.]), the Shanghai Rising-Star Program (Grant No.13QA1404600 [Z.L.]), the HIROMI Medical Research Foundation (A.M.), the Mother and Child Health Foundation (A.M.), the Smoking Research Foundation (A.M.), the Pancreas Research Foundation of Japan (E.N.), the Ministry of Health, Labour and Welfare of Japan (Principal investigators: Yoichi Matsubara and Yoshifumi Takeyama), Council of Scientific and Industrial Research, Ministry of Science and Technology, Government of India (Grant No. BSC0121 [G.R.C.]), the Conseil Régional de Bretagne, the Association des Pancréatites Chroniques Héréditaires, the Association de Transfusion Sanguine et de Biogénétique Gaetan Saleun, and the Institut National de la Santé et de la Recherche Médicale (INSERM), France.

Abbreviations used in this paper: CP, chronic pancreatitis; NAHR, non-allelic homologous recombination; NMD, nonsense-mediated mRNA decay; RT-PCR, reverse transcription-PCR; SNP, single-nucleotide polymorphism; VNTR, variable number tandem repeat; WT, wild-type.

Correspondence to:

Zhuan Liao, MD, Department of Gastroenterology, Changhai Hospital, the Second Military Medical University, 168 Changhai Road, Shanghai 200433, China. e-mail: liaozhuan@smmu.edu.cn.; Tel.: 86-21-31161344; Fax: 0086-21-55621735.

Zhao-Shen Li, MD, Department of Gastroenterology, Changhai Hospital, the Second Military Medical University, 168 Changhai Road, Shanghai 200433, China. e-mail: zhaoshenli@hotmail.com; Tel.: 86-21-31161335; Fax: 0086-21-55621735.

Jian-Min Chen, MD, PhD, INSERM U1078 and EFS – Bretagne, 46 rue Félix Le Dantec, 29218 Brest 29218, France; Jian-Min.Chen@univ-brest.fr; Tel.: +33-2-98449333; Fax: +33-2-98430555

Conflicts of Interest

The authors are unaware of any conflicts of interest.

Author Contributions

JMC, ZSL and ZL designed and directed the overall project, with the assistance of CF. AM and TS directed the Japanese study whereas GRC directed the Indian study. WBZ, AM, PI performed the genetic analysis, with substantial contributions from EM, HW, XTS, LHH, DZZ, LH, EN, SH, YK, KK, HI, SP, KRM and SB. EM established the modified genotyping method. WBZ and AB performed the functional analysis, with contributions from YF. JMC and ZL wrote the manuscript, with substantial contributions from WBZ, AB, EM, AM and GRC. DNC, CF and ZSL critically revised the manuscript with important intellectual input.

AM, GRC, ZSL and ZL recruited study subjects. AM, CF, GRC and ZL obtained the funding. All authors approved the final manuscript.

ABSTRACT

A hybrid allele between the carboxyl ester lipase gene (*CEL*) and its pseudogene, *CELP* (called CEL–HYB), generated by non-allelic homologous recombination between *CEL* intron 10 and *CELP* intron 10′, was found to increase susceptibility to chronic pancreatitis in a case–control study of patients of European ancestry. We attempted to replicate this finding in 3 independent cohorts from China, Japan, and India, but failed to detect the CEL–HYB allele in any of these populations. The CEL–HYB allele might therefore be an ethnic-specific risk factor for chronic pancreatitis. An alternative hybrid allele (CEL–HYB2) was identified in all 3 Asian populations (1.7% combined carrier frequency), but was not associated with chronic pancreatitis.

Keywords: Disease Susceptibility; Human Genetics; Nonsense-Mediated mRNA Decay; Pancreatic Acinar Cells.

The etiology of chronic pancreatitis (CP) is complex and involves a subtle interplay between genetic and environmental factors. Most of the hitherto reported genes/loci affecting disease susceptibility encode proteins of the protease-antiprotease system of the pancreatic acinar cells.[1-7] A hybrid allele (CEL–HYB), involving the *CEL* gene (encoding pancreatic carboxyl ester lipase[8]) and its tandemly linked pseudogene (*CELP*), was recently reported to be significantly overrepresented in CP cases as compared with controls, firstly in a discovery cohort of patients with familial CP and then in three replication cohorts of patients with idiopathic CP.[9] The CEL–HYB allele resulted from non-allelic homologous recombination (NAHR) occurring between *CEL* intron 10 and *CELP* intron 10′; replacement of the eleventh and last exon of *CEL* by *CELP* exon 11′ would yield a premature stop codon within the third "pseudo" 33-bp variable number tandem repeat (VNTR)[9] (Figure 1*A*). The mutant enzyme is more stable than its wild-type (WT) counterpart and induced autophagy in cellular models,[9] suggesting a novel pathogenic mechanism. However, since the patients analyzed in the original study[9] were solely of European ancestry, replication in independent populations of different ethnicity was warranted.[10]

The detection of the CEL–HYB allele is dependent upon a long-range duplex PCR assay.[9] Using French CEL–HYB positive samples, we found that the originally described PCR assay[9] yielded better results under slightly modified conditions (see Supplementary Material and Supplementary Figure 1). Consequently, these modified conditions were employed to screen for the CEL–HYB allele in three Asian populations. We first analyzed a cohort of Han Chinese patients with idiopathic CP. The previously described CEL–HYB-specific 3.2-kb band was detected in 2.4% (19/799) of patients and 1.9% (20/1028) of controls ($P = 0.64$; Table 1). Sequencing of the 3.2-kb band-containing PCR products however revealed that none of these Chinese CEL–HYB alleles corresponded to that expected; instead, they invariably resulted from an NAHR event occurring within a 239-bp sequence tract

affecting the intron 9/exon 10 boundary of *CEL* and the intron 9′/exon 10′ boundary of *CELP* (Supplementary Figure 2). Hereafter, we term the previously reported disease-associated CEL–HYB allele[9] as CEL–HYB1 so as to distinguish it from this newly described allele, CEL–HYB2.

Closer inspection of the 39 Chinese CEL–HYB2 alleles revealed that they could be divided into two subtypes by reference to three single nucleotide polymorphisms (SNPs) present within the substituting *CELP* sequence (Supplementary Figure 2); 38 harbored the G allele of rs10901232A/G, the T allele of rs10901233C/T and the G allele of rs671412A/G (termed CEL–HYB2a), the sole exception harboring the alternative alleles of these three SNPs (termed CEL–HYB2b). We further genotyped two idiopathic CP cohorts from Japan and India. Results from the Japanese cohort were remarkably similar to those from the Chinese cohort whereas in the Indian cohort, only one control carried a CEL–HYB allele, and this corresponded to CEL–HYB2b (Table 1). Note that a variant similar to CEL–HYB2 was mentioned in passing by Fjeld et al.[9] to be present in one European case and four controls, although no sequence details were presented. Sequencing of two corresponding French CEL–HYB positive samples revealed their identity to be CEL–HYB2b. In short, the PCR assay employed efficiently detects all three *CEL* hybrid alleles, the differentiation of which relies upon the sequencing of the respective PCR products.

Both CEL–HYB2a and CEL–HYB2b variants harbor premature stop codons within their chimeric exons 10 (Figure 1*A*); their corresponding mRNAs could therefore be subject to significant degradation by nonsense-mediated mRNA decay (NMD). However, this might not be the case for CEL–HYB1 because mRNAs that harbor a stop codon in the final exon usually escape degradation by NMD.[11] To explore this possibility, we first sought to compare the mRNA expression levels of CEL–HYB2a versus CEL–HYB1 *in vitro* (Supplementary Material and Supplementary Tables 1 and 2). Briefly, we PCR amplified the full-length

genomic CEL–HYB2a and CEL–HYB1 sequences from their corresponding carriers and cloned the resulting PCR products into the pcDNA3.1/V5-His-TOPO vector. Reverse transcription-PCR (RT-PCR) analyses of mRNAs from subsequently transfected HEK293T cells indicated lower CEL–HYB2a mRNA expression as compared to CEL–HYB1 (Figure 1*B*). Further quantitative RT-PCR analyses demonstrated that the mRNA expression of CEL–HYB2a accounted for only 60% of that of CEL–HYB1 (Figure 1*C*). Then, we tested whether the mRNA expression level of CEL–HYB2a could be increased by treatment of the transfected cells with cycloheximide, a known NMD inhibitor[12] and found this to be the case (Figure 1*D*). Hence, we conclude that mRNA expression from the CEL–HYB2a allele is significantly reduced by NMD. This conclusion probably also applies to CEL–HYB2b owing to its high sequence similarity with CEL–HYB2a (Figure 1*A*).

In summary, contrary to our expectation, we failed to identify the CEL–HYB1 allele in any of the three Asian cohorts. Based on the allele frequency of CEL–HYB1 in healthy German and French populations (0.4%),[9] the power to detect at least one CEL–HYB1 carrier among the cases is estimated to be >86%, even for the smallest cohort. Given that most rare variants (defined as a minor allele frequency of <0.5%) are population-specific,[13] CEL–HYB1 may be an ethnicity-specific disease risk factor, although this remains to be confirmed by replication in an independent cohort of European ancestry (e.g., the North American Pancreatitis Study 2[14]). The other unexpected finding was the identification of an alternative CEL–HYB2 allele in all three Asian populations. The significant degradation of CEL–HYB2a mRNA by NMD and the observation that pancreatic exocrine function has been found to be normal in *Cel*-knockout mice[15] are consistent with the apparent lack of any association between CEL–HYB2a and CP.

Figure Legend

Figure 1. (***A***) Key differences between CEL–HYB1 (originally termed CEL–HYB[9]), CEL–HYB2 and the wild-type *CEL* gene in terms of the defining exon 10, intron 10 and exon 11 sequences. The gene structure of *CEL* (in black) and that of the replacement *CELP* sequences (in green) within CEL–HYB1 and CEL–HYB2 were described in accordance with GenBank accession number AF072711.1. In *CEL*, the VNTRs within exon 11 are indicated in purple; nucleotide positions of the exon 10 boundaries and the beginning of exon 11 are numbered by reference to the A of the translational initiation codon ATG as c.1; the translational termination codon is denoted by a vertical blue bar, with the last coding nucleotide and the amino acid position of the translational termination codon being numbered above and below the bar. In CEL–HYB1, CEL–HYB2a and CEL–HYB2b, the presumed premature stop codons are indicated in a similar manner. **(*B*)** A representative gel showing the RT-PCR analyses of HEK293T cells transfected with expression constructs carrying the full-length CEL–HYB1 and CEL–HYB2a genomic sequences. Sanger sequencing of the approximately 2.2-kb CEL–HYB1 and CEL–HYB2a products revealed that all introns were spliced correctly. CV, control vector. (***C***) Relative mRNA expression levels of CEL–HYB2a versus CEL–HYB1 *in vitro* as determined by quantitative RT-PCR analyses. CV, control vector. (***D***) Relative mRNA expression levels of CEL–HYB2a in transfected cells with (grey) and without (black) cycloheximide treatment as determined by quantitative RT-PCR analyses. **, $P < 0.01$.

References

1. Whitcomb DC, et al. Nat Genet 1996;14:141-145.

2. **Le Maréchal C, Masson E,** et al. Nat Genet 2006;38:1372-1374.

3. Witt H, et al. Nat Genet 2000;25:213-216.

4. **Witt H, Sahin-Tóth M,** et al. Nat Genet. 2006;38:668-673.

5. **Rosendahl J, Witt H, Szmola R,** et al. Nat Genet 2008;40:78-82.

6. Masson E, et al. Hum Genet 2008;123:83-91.

7. Witt H, et al. Nat Genet 2013;45:1216-1220.

8. Holmes RS, Cox LA. Cholesterol 2011;2011:781643.

9. Fjeld K, et al. Nat Genet 2015;47:518-522.

10. MacArthur DG, et al. Nature 2014;508:469-476.

11. Karam R, et al. Biochim Biophys Acta 2013;1829:624-633.

12. Pereverzev AP, et al. Sci Rep 2015;5:7729.

13. **Tennessen JA, Bigham AW, O'Connor TD,** et al. Science 2012;337:64-69.

14. Whitcomb DC, et al. Nat Genet 2012;44:1349-1354.

15. Vesterhus M, et al. Pancreatology 2010;10:467-476.

Author names in bold designate shared co-first authorship.

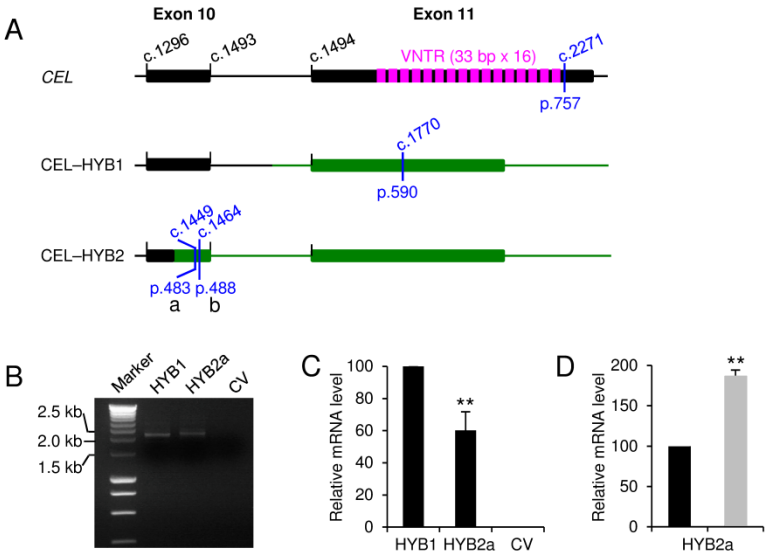

A
Exon 10
Exon 11
c.1296
c.1493
c.1494
VNTR (33 bp x 16)
c.2271
CEL
p.757
c.1770
CEL–HYB1
p.590
c.1449
c.1464
CEL–HYB2
p.483
p.488
a
b
B
Marker
HYB1
HYB2a
CV
2.5 kb
2.0 kb
1.5 kb
C
Relative mRNA level
100
80
60
40
20
0
**
HYB1
HYB2a
CV
D
Relative mRNA level
200
150
100
50
0
**
HYB2a

Table 1. Carrier frequencies of the CEL–HYB2 variant in Chinese, Japanese and Indian subjects with idiopathic chronic pancreatitis*

Populations	**Case**		**Control**		***P* value**[a]
	+/n	%	+/n	%	
Chinese	19/799	2.4	20 (1)/1028	1.9	0.64
Japanese	4[b]/248	1.6	7 (2)/403	1.7	0.85
Indian	0/280	0	1 (1)/225	0.4	0.91
Combined	23/1327	1.7	28 (4)/1656	1.7	0.96

* Number of subjects carrying the less frequent CEL–HYB2 subtype, CEL–HYB2b, is indicated in parentheses wherever applicable.

[a] Two-tailed Fisher's exact test.

[b] One patient is a homozygote.

Supplementary Material

Subjects

Chinese, Japanese and Indian patients with idiopathic chronic pancreatitis (i.e. absence of a positive family history and any known precipitating factor such as alcohol abuse) and healthy controls from their corresponding general populations were recruited by the three leading Asian pancreatitis groups. A clinical diagnosis of chronic pancreatitis was based on two or more of the following criteria: (i) presence of a typical history of recurrent pancreatitis; (ii) radiological findings such as pancreatic calcification and/or pancreatic irregularities revealed by endoscopic retrograde pancreatography or by magnetic resonance imaging of the pancreas; and (iii) pathological sonographic findings. Informed consent was obtained from each patient and the study was approved by the respective ethics committees of the three leading Asian pancreatitis groups.

Screening for the CEL–HYB Allele

Screening for the disease-associated CEL–HYB allele (termed CEL–HYB1 in the present study) was performed in accordance with the previously reported method[9] but with some modified reaction conditions. Specifically, we performed the PCR reactions in total volumes of 10 µl containing 2 × GC buffer, 0.4 µM primer L11F (5'-GTCCCTCACTCATTCTTCTATGGCAAC-3'), 0.2 µM of the primers IAR (5'-TCCAAAGCCCTAGCAGTAACGA-3') and *CELP* VNTR-rev (5'-CTGTGGAGGGCATGGAACT-3'), 0.4 mM each dNTP, 0.5 U *LA Taq* DNA polymerase and 10-50 ng genomic DNA. These conditions work better than those originally described for genotyping the French CEL–HYB1 positive samples (Supplementary Figure 1).

Sequencing the Positive Samples

Samples showing the CEL–HYB-specific 3.2-kb band[9] were subjected to Sanger sequencing using the previously published primers, S10F and S10R.[9] Crossover regions of the CEL–HYB alleles were assigned by evaluating the nucleotide positions that serve to discriminate the *CEL* and *CELP* sequences within the region of interest. To this end, we aligned *CEL* exon 10 and its partial flanking sequences against the corresponding *CELP* sequences, the discriminant nucleotides being clearly indicated (Supplementary Figure 2).

The assessment of significance of the differences between the carrier frequencies of the CEL–HYB alleles in patients and controls was performed by means of two-tailed Fisher's exact test using StatCalc (2 × 2 Tables) from Epi Info™ 7 (http://wwwn.cdc.gov/epiinfo/index.htm). The difference was regarded as being statistically significant when the *P* value was ≤ 0.05.

Construction of Expression Vectors for mRNA Expression Analysis in the Context of the Full-Length CEL–HYB1 and CEL–HYB2a Genomic Sequences

We PCR amplified the full-length (approximately 10-kb) genomic CEL–HYB1 and CEL–HYB2a sequences (going from the 5'-untranslated region of *CEL* to the 3'-untranslated region of the replacement *CELP*) from their corresponding carriers and cloned the resulting PCR products into the pcDNA3.1/V5-His-TOPO vector using previously described methods.[16,17] Specifically, the long-range PCR was performed using 50 ng DNA in a 50 µL reaction mixture with 2.5 U TaKaRa LA Taq DNA polymerase, 8 µL dNTP Mixture (400 µM final), 5 µL 10 × PCR buffer system, and 1 µM each primer of primer pair A1 (Supplementary Table 1). Thermal cycling conditions were: initial denaturation at 94°C for 1 min, followed by 30

cycles of denaturation at 98°C for 10 s, annealing and extension at 68°C for 12 min, and a final extension step at 72°C for 10 min. PCR products of the expected size were purified using the NucleoSpin® Gel kit (Macherey Nagel) after gel electrophoresis. 3'-A overhangs were added to the purified products before being cloned into the pcDNA3.1/V5-His-TOPO vector (Invitrogen) in accordance with the manufacturer's instructions. Transformation was performed using XL10-Gold Ultracompetent Cells (Stratagene). Transformed cells were spread onto LB agar plates with 50 µg/ml ampicillin and incubated at 37°C overnight.

Plasmid constructs containing inserts in the right orientation were selected for simultaneously by two PCRs using the HotStarTaq Master Mix Kit (Qiagen). The first PCR amplifies a 235-bp fragment extending from the T7 promoter/priming site (located within the vector) to the exon 1/intron 1 boundary of *CEL* (i.e. primer pair A2 in Supplementary Table 1). The second PCR amplifies an approximately 1.3-kb fragment going from the beginning of *CEL* exon 10 to the BGH reverse priming site (located within the vector) (i.e. primer pair A3 in Supplementary Table 1). Selected plasmids were then sequenced using primer S10F (see Supplementary Figure 2) to pick up those carrying the desired CEL–HYB1 or CEL–HYB2a variant.

Cell Culture, Transfection and Cycloheximide Treatment

Human embryonic kidney (HEK) 293T cells were cultured in DMEM nutrient mixture with 10% fetal calf serum. 3.5 x 10^5 cells were seeded per well of 6-well plates 24 hours prior to transfection. For conventional RT-PCR analyses (see below), 1 µg pcDNA3.1-CEL–HYB1 or pcDNA3.1-CEL–HYB2a plasmid, which were mixed with 3 µL Lipofectamine 2000 Reagent (Invitrogen), were used for transfection per well. For real-time quantitative RT-PCR analyses (see below), 500 ng pcDNA3.1-CEL–HYB1 or pcDNA3.1-CEL–HYB2a plasmid were mixed with 500 ng pGL3-GP2 minigene for transfection. Forty-eight hours after transfection, total RNA was extracted using the RNeasy Mini Kit (Qiagen). For the NMD inhibition experiment, 50 µg/mL (final concentration) cycloheximide (Sigma) were added to the cells 4 hours before RNA extraction.

Reverse Transcription

Reverse transcription (RT) was performed using the SuperScript II Reverse Transcriptase (Life Technologies) with 1 µg total RNA, 500 µM dNTPs, 5 µM 20mer-oligo(dT) and 10 mM DTT (Eurogentec). The resulting cDNAs were treated with 2U RNAseH (Life Technologies) to degrade the remaining RNA.

Conventional RT-PCR and Sequencing of the Resulting Products

Conventional RT-PCR was performed in a 25-µl reaction mixture containing 0.5 U KAPA HiFi HotStart DNA Polymerase (Kapa Biosystems), 1 µl cDNA and 0.3 µM each primer of primer pair A4 (Supplementary Table 1). Note that this primer pair was designed to amplify a full-length CEL–HYB1 or CEL–HYB2a transcript. The PCR program comprised an initial denaturation step at 95°C for 5 min, followed by 30 cycles of denaturation at 98°C for 20 s, annealing at 72°C for 15 s, and extension at 72°C for 3 min, and a final extension step at 72°C for 5 min.

RT-PCR products were cleaned using ExoSAP-IT® before being sequenced with five primers (Supplementary Table 2) by means of the BigDye Terminator v1.1 Cycle Sequencing Kit (Life Technologies).

Real-Time Quantitative RT-PCR

Real-time quantitative RT-PCR analyses were performed essentially as described elsewhere.[16] In brief, pcDNA3.1-CEL–HYB1 or pcDNA3.1-CEL–HYB2a was co-transfected

with the previously constructed pGL3-GP2 minigene.[16] The target and reference genes were PCR amplified separately; target gene expression was determined using pGL3-GP2 minigene expression as a reference by means of the ΔΔCt method in accordance with the Pfaffl efficiency-corrected calculation model.[18] Real-time quantitative RT-PCR was performed in a 25μL mixture containing 12.5 μL HotStarTaq Master Mix Kit (Qiagen), 1 μL 1:25 diluted cDNA, 0.5 μM SYTO9 (Life Technologies) and 0.3 μM each primer (primer pair A5 was employed for both CEL–HYB1 and CEL–HYB2a whilst primer pair A6 was used for the reference gene; Supplementary Table 1). The PCR program comprised an initial denaturation step at 95°C for 15 min, followed by 40 cycles of denaturation at 94°C for 45 s, annealing at 55°C for 30 s and extension at 72°C for 30 s. Quantitative analysis of each RT-PCR amplification was performed in triplicate on a Lightcycler 480II (Roche, Paris, France) using "relative quantification" and "2nd derivative maxima" options.

The difference between the expression levels of CEL–HYB1 and CEL–HYB2a (results from six independent transfection experiments) and that between expression levels of CEL–HYB2a with and without cycloheximide treatment (results from three independent transfection experiments) were assessed for significance by the Student's t-test. The difference was regarded as being statistically significant when the *P* value was ≤ 0.05.

Supplementary References

16. Boulling A, Chen JM, Callebaut I, *et al*. Is the *SPINK1* p.Asn34Ser missense mutation *per se* the true culprit within its associated haplotype? WebmedCentral GENETICS 2012;3:WMC003084 [http://static.webmedcentral.com/wmcpdf/Article_WMC003084.pdf].
17. **Zou WB, Boulling A**, Masson E, et al. Clarifying the clinical relevance of *SPINK1* intronic variants in chronic pancreatitis. Gut. 2015 Dec 30. pii: gutjnl-2015-311168. doi: 10.1136/gutjnl-2015-311168.
18. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 2001;29:e45.

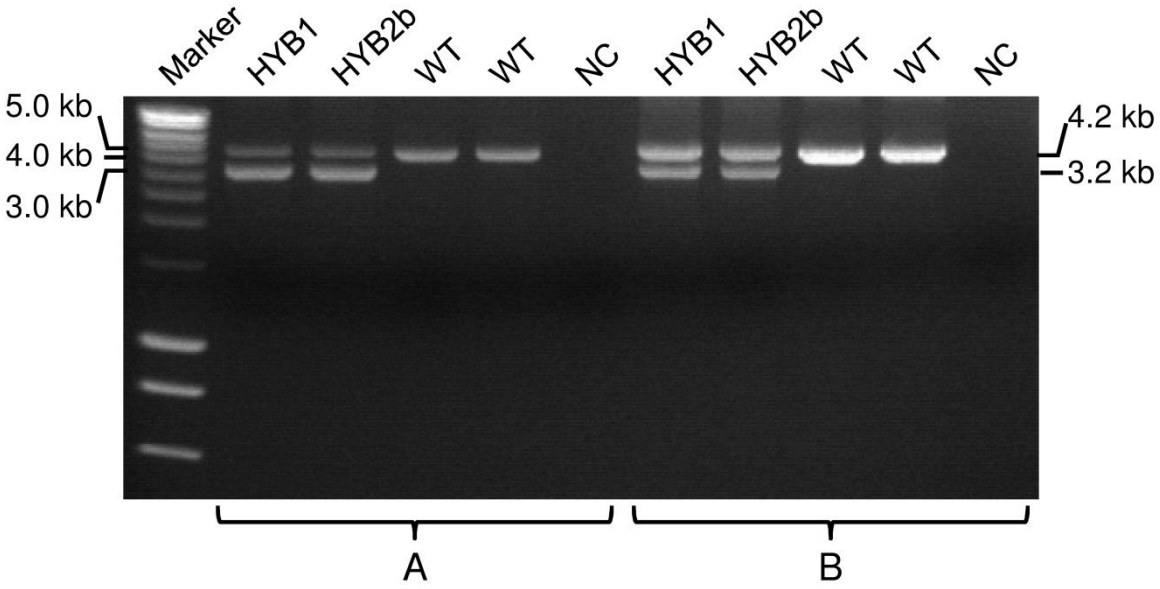
Author names in bold designate shared co-first authorship.

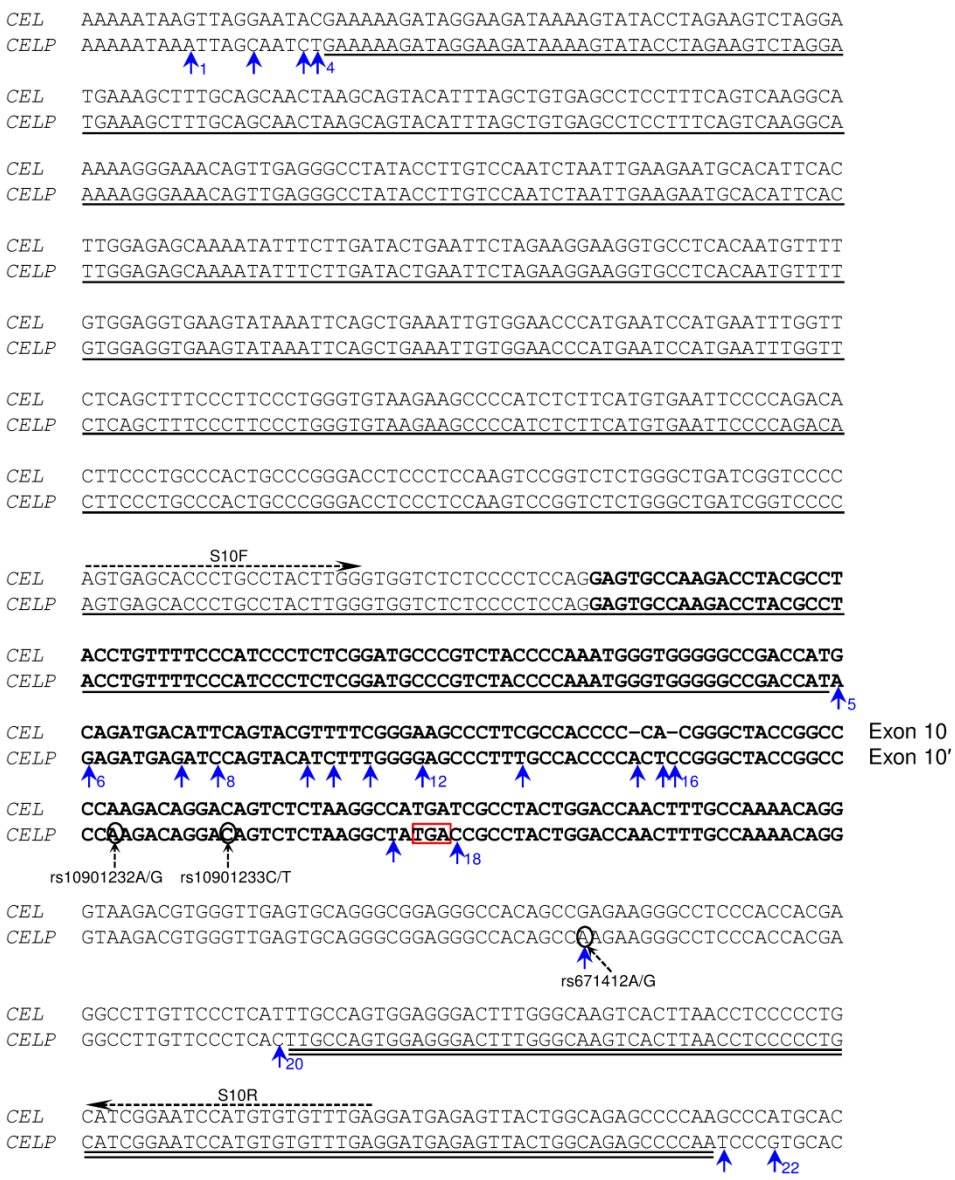
Supplementary Figure Legends

Supplementary Figure 1. Comparison of the originally reported PCR assay (**A**) and the currently used PCR assay (**B**) for detecting the CEL–HYB1 (as well as the CEL–HYB2b) allele (3.2 kb) in French heterozygotes. WT, wild-type (4.2 kb). NC, negative control.

Supplementary Figure 2. Alignment of reference *CEL* and *CELP* sequences in the context of exon 10 (in bold) and partial flanking sequences. The crossover region for the CEL–HYB deletion allele[9] (termed here CEL–HYB1) is doubly underlined whereas that for CEL–HYB2 is singly underlined. The discriminant positions between the two paralogous sequences are indicated by numbered upwardly pointing arrows. The CEL–HYB2a variant harbors the G allele of rs10901232A/G, the T allele of rs10901233C/T and the G allele of rs671412A/G whereas the CEL–HYB2b variant harbors the alternative alleles of these three SNPs (the three affected sites are indicated by circles). The premature translation termination codon in the chimeric exon 10 of the CEL–HYB2b variant is indicated by a box. The two primers used for sequencing the breakpoint junctions of the CEL–HYB alleles, S10F and S10R,[9] are indicated by horizontal dotted arrows. The *CEL* and *CELP* sequences shown in the Figure correspond to

nucleotide positions 12883-13599 and 28652-29371, respectively, from GenBank accession number AF072711.1.




CEL AAAAATAAGTTAGGAATACGAAAAAGATAGGAAGATAAAAGTATACCTAGAAGTCTAGGA
CELP AAAAATAAATTAGCAATCTGAAAAAGATAGGAAGATAAAAGTATACCTAGAAGTCTAGGA
1
4
CEL TGAAAGCTTTGCAGCAACTAAGCAGTACATTTAGCTGTGAGCCTCCTTTCAGTCAAGGCA
CELP TGAAAGCTTTGCAGCAACTAAGCAGTACATTTAGCTGTGAGCCTCCTTTCAGTCAAGGCA
CEL AAAAGGGAAACAGTTGAGGGCCTATACCTTGTCCAATCTAATTGAAGAATGCACATTCAC
CELP AAAAGGGAAACAGTTGAGGGCCTATACCTTGTCCAATCTAATTGAAGAATGCACATTCAC
CEL TTGGAGAGCAAAATATTTCTTGATACTGAATTCTAGAAGGAAGGTGCCTCACAATGTTTT
CELP TTGGAGAGCAAAATATTTCTTGATACTGAATTCTAGAAGGAAGGTGCCTCACAATGTTTT
CEL GTGGAGGTGAAGTATAAATTCAGCTGAAATTGTGGAACCCATGAATCCATGAATTTGGTT
CELP GTGGAGGTGAAGTATAAATTCAGCTGAAATTGTGGAACCCATGAATCCATGAATTTGGTT
CEL CTCAGCTTTCCCTTCCCTGGGTGTAAGAAGCCCCATCTCTTCATGTGAATTCCCCAGACA
CELP CTCAGCTTTCCCTTCCCTGGGTGTAAGAAGCCCCATCTCTTCATGTGAATTCCCCAGACA
CEL CTTCCCTGCCCACTGCCCGGGACCTCCCTCCAAGTCCGGTCTCTGGGCTGATCGGTCCCC
CELP CTTCCCTGCCCACTGCCCGGGACCTCCCTCCAAGTCCGGTCTCTGGGCTGATCGGTCCCC
S10F
CEL AGTGAGCACCCTGCCTACTTGGGTGGTCTCTCCCCTCCAGGAGTGCCAAGACCTACGCCT
CELP AGTGAGCACCCTGCCTACTTGGGTGGTCTCTCCCCTCCAGGAGTGCCAAGACCTACGCCT
CEL ACCTGTTTTCCCATCCCTCTCGGATGCCCGTCTACCCCAAATGGGTGGGGGCCGACCATG
CELP ACCTGTTTTCCCATCCCTCTCGGATGCCCGTCTACCCCAAATGGGTGGGGGCCGACCATA
5
CEL CAGATGACATTCAGTACGTTTTCGGGAAGCCCTTCGCCACCCC-CA-CGGGCTACCGGCC Exon 10
CELP GAGATGAGATCCAGTACATCTTTGGGGAGCCCTTTGCCACCCCACTCCGGGCTACCGGCC Exon 10′
6
8
12
16
CEL CCAAGACAGGACAGTCTCTAAGGCCATGATCGCCTACTGGACCAACTTTGCCAAAACAGG
CELP CCAAGACAGGACAGTCTCTAAGGCTATGACCGCCTACTGGACCAACTTTGCCAAAACAGG
18
rs10901232A/G
rs10901233C/T
CEL GTAAGACGTGGGTTGAGTGCAGGGCGGAGGGCCACAGCCGAGAAGGGCCTCCCACCACGA
CELP GTAAGACGTGGGTTGAGTGCAGGGCGGAGGGCCACAGCCGAGAAGGGCCTCCCACCACGA
rs671412A/G
CEL GGCCTTGTTCCCTCATTTGCCAGTGGAGGGACTTTGGGCAAGTCACTTAACCTCCCCCTG
CELP GGCCTTGTTCCCTCACTTGCCAGTGGAGGGACTTTGGGCAAGTCACTTAACCTCCCCCTG
20
S10R
CEL CATCGGAATCCATGTGTGTTTGAGGATGAGAGTTACTGGCAGAGCCCCAAGCCCATGCAC
CELP CATCGGAATCCATGTGTGTTTGAGGATGAGAGTTACTGGCAGAGCCCCAATCCCGTGCAC
22

Supplementary Table 1. PCR primer pairs used in this study

Primer pair	Sequence (5' > 3')	Location	Amplified sequence	Amplicon size (bp)
A1	Forward: TCCATAAATACCCGAGGCCC	5'-untranslated region of *CEL*	Full-length genomic CEL–HYB sequences	9676 (HYB1) 9678 (HYB2a)
	Reverse: CTTCCTGCAGCTTAGCCTTG	3'-untranslated region of *CELP*		
A2	Forward: TAATACGACTCACTATAGGG	T7 promoter/priming site	5' extremity of the constructs	235
	Reverse: CTCTGCTGGGCTCTTACCTT	Exon 1/intron 1 boundary of genomic *CEL*		
A3	Forward: CAAGACCTACGCCTACCTGT	Exon 10 of genomic *CEL*	3' extremity of the constructs	1334 (HYB1) 1336 (HYB2a)
	Reverse: TAGAAGGCACAGTCGAGG	BGH reverse priming site		
A4	Forward: GGAGACCCAAGCTGGCTAGT	pcDNA3.1/5'-untranslated region	Full length CEL–HYB transcripts	2159 (HYB1) 2161 (HYB2a)
	Reverse: CCCTCTCGGCCTCTTGAG	Between the stop codon and polyadenylation signal of *CELP*		
A5	Forward: GGAGACCCAAGCTGGCTAGT	pcDNA3.1/5'-untranslated region	5' extremity of CEL–HYB cDNA	280
	Reverse: ATGTCCACAGAGTCACCCAG	Exon 2 of *CEL*		
A6[a]	Forward: ACTGTTGGTAAAGCCACCAT	pGL3-control/5'-untranslated region	Exons 8 and 9 of *GP2*	479
	Reverse: TGTATCTTATCATGTCTGCTCGAA	pGL3-control/3'-untranslated region		

[a] Primer sequences were taken from ref. 16.

Supplementary Table 2. Primers used for sequencing the full-length CEL–HYB1 and CEL–HYB2a transcripts

Primer	Sequence (5' > 3')	Location
1	GGAGACCCAAGCTGGCTAGT	pcDNA3.1/5'-untranslated region
2	GGACCTGCCCGTTATGATCT	Exon 4 of *CEL*
3	TGGCTGAGAAGGTGGGTTG	Exon 7 of *CEL*
4	CAAGACCTACGCCTACCTGT	Exon 10 of *CEL*
5	CCCTCTCGGCCTCTTGAG	Between the stop codon and polyadenylation signal of *CELP*