

# Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <http://orca.cf.ac.uk/94705/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Boulling, Arnaud, Abrantes, Amandine, Masson, Emmanuelle, Cooper, David Neil, Robaszkiewicz, Michel, Chen, Jian-Min and Férec, Claude 2016. Discovery and functional annotation of PRSS1 promoter variants in chronic pancreatitis. *Human Mutation* 37 (11) , pp. 1149-1152.  
10.1002/humu.23053 file

Publishers page: <http://dx.doi.org/10.1002/humu.23053> <<http://dx.doi.org/10.1002/humu.23053>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.





## Discovery and Functional Annotation of Novel *PRSS1* Promoter Variants in Chronic Pancreatitis

Journal:	<i>Human Mutation</i>
Manuscript ID	humu-2016-0116
Wiley - Manuscript type:	Brief Report
Date Submitted by the Author:	10-Mar-2016
Complete List of Authors:	Boulling, Arnaud; Institut National de la Santé et de la Recherche Médicale (INSERM), U613, Abrantes, Amandine; Genetique Genomique Fonctionnelle et Biotechnologies Masson, Emmanuelle; Genetique Genomique Fonctionnelle et Biotechnologies Cooper, David; Cardiff University, Institute of Medical Genetics, College of Medicine; Robaszkiewicz, Michel; Service de Gastro-Entérologie, Centre Hospitalier Régional Universitaire (CHRU) Brest, Chen, Jian-Min; INSERM U1078 and EFS-Bretagne, Ferec, Claude; Université Brest, ; INSERM U613,
Key Words:	chronic pancreatitis, precision medicine, promoter variant, reporter gene assay

SCHOLARONE™  
Manuscripts

## Brief Report

### **Discovery and Functional Annotation of Novel *PRSSI* Promoter Variants in Chronic Pancreatitis**

Arnaud Boulling,<sup>1,2,3#</sup> Amandine Abrantes,<sup>2,4,5#</sup> Emmanuelle Masson,<sup>1,4</sup> David N. Cooper,<sup>6</sup>  
Michel Robaszkiewicz,<sup>5</sup> Jian-Min Chen,<sup>1,2,3\*</sup> and Claude Férec<sup>1,2,3,4</sup>

<sup>1</sup>Institut National de la Santé et de la Recherche Médicale (INSERM), U1078, Brest, France;  
<sup>2</sup>Faculté de Médecine et des Sciences de la Santé, Université de Bretagne Occidentale (UBO),  
Brest, France; <sup>3</sup>Etablissement Français du sang (EFS) – Bretagne, Brest, France; <sup>4</sup>Laboratoire  
de Génétique Moléculaire et d'Histocompatibilité, Centre Hospitalier Régional Universitaire  
(CHRU) Brest, Hôpital Morvan, Brest, France; <sup>5</sup>Service de Gastro-Entérologie, Centre  
Hospitalier Régional Universitaire (CHRU) Brest, Hôpital Cavale Blanche Morvan, Brest,  
France; <sup>6</sup>Institute of Medical Genetics, School of Medicine, Cardiff University, Cardiff,  
United Kingdom

#These authors contributed equally to the work.

\*Correspondence to: Jian-Min Chen, INSERM U1078 and EFS – Bretagne, 46 rue Félix Le  
Dantec, 29218 Brest, France. Tel: +33-2-98449333; Fax: +33-2-98430555; E-mail: Jian-  
Min.Chen@univ-brest

Contract grant sponsors: The Conseil Régional de Bretagne, the Association des Pancréatites  
Chroniques Héritaires, the Association de Transfusion Sanguine et de Biogénétique Gaetan  
Saleun, and the Institut National de la Santé et de la Recherche Médicale (INSERM), France.

1  
2  
3 **ABSTRACT:** In the new age of precision medicine, interpretation of the clinical relevance of  
4 disease-associated variants, particularly those rare ones, is most challenging. Herein we  
5 illustrate the fundamental importance of functional analysis in this issue, in the context of rare  
6 *PRSSI* (encoding cationic trypsinogen) promoter variants. Gain-of-function *PRSSI* missense  
7 and copy number variants are known to cause or predispose to chronic pancreatitis (CP). A  
8 *PRSSI* promoter variant, c.-30\_-28delTCC, was identified in a single individual with  
9 idiopathic CP nearly 20 years ago. It has long been speculated to increase trypsinogen  
10 expression, thereby predisposing to CP. We performed functional analysis of this variant  
11 together with three new rare variants discovered from the resequencing of the *PRSSI*  
12 promoter in a French CP cohort. Contrary to our expectation, c.-30\_-28delTCC resulted in  
13 reduced rather than increased gene expression, suggesting that it may act as a protective factor  
14 rather than a risk factor for CP.  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33

34 **KEY WORDS:** chronic pancreatitis; precision medicine; promoter variant; reporter gene  
35 assay  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 Chronic pancreatitis (CP) is a progressive inflammation of the pancreas that gives rise to  
4 irreversible morphological changes and impairment of both exocrine and endocrine functions  
5 [Braganza et al. 2011]. The etiology of CP is complex, with contributions from several  
6 [Braganza et al. 2011]. The etiology of CP is complex, with contributions from several  
7 different genetic factors having been recognized over the past two decades (see [Fjeld et al.  
8 2015] and references therein). In particular, findings from three genes, *PRSSI* (encoding  
9 cationic trypsinogen; OMIM #276000), *SPINK1* (encoding pancreatic secretory trypsin  
10 inhibitor; OMIM #167790) and *CTRC* (encoding chymotrypsin C, which specifically  
11 degrades all human trypsinogen/trypsin isoforms [Szmola and Sahin-Tóth 2007]; OMIM  
12 #601405), have firmly established the importance of a balance between activation and  
13 inactivation of trypsinogen within the pancreatic acinar cells. Specifically, gain-of-function  
14 variants, including missense and copy number variants in the *PRSSI* gene [Le Maréchal et al.  
15 2006; Whitcomb et al. 1996] as well as loss-of-function variants in the *SPINK1* [Witt et al.  
16 2000] and *CTRC* [Masson et al. 2008b; Rosendahl et al. 2008] genes, cause or predispose to  
17 CP.

18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34 A large number of genetic variants have so far been reported in each of the  
35 aforementioned three CP genes (see Human Gene Mutation Database (HGMD;  
36 <http://www.hgmd.org>)). Some types of variant are of obvious clinical relevance by virtue of  
37 their mutational type (e.g., *PRSSI* duplication and triplication copy number mutations; splice  
38 site mutations or large deletions in *SPINK1* or *CTRC*). However, the clinical relevance of  
39 other types of variant has often to be ascertained by functional annotation. In this regard,  
40 missense mutations in all the three CP genes have been subject to systematic functional  
41 characterization (*PRSSI*, [Schnúr et al. 2014]; *SPINK1*, [Boulling et al. 2007; Boulling et al.  
42 2012; Király et al. 2007]; and *CTRC*, [Beer et al. 2013]). By contrast, in terms of regulatory  
43 variants, only those in the *SPINK1* gene, including promoter [Boulling et al. 2011; Derikx et al.  
44 2015b; Hegyi et al. 2016] and intronic [Zou et al. 2015; Zou et al. 2016] variants, have  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 been studied systematically. These functional analyses not only served to improve our  
4  
5 understanding of the underlying pathophysiological mechanisms but more importantly, in the  
6  
7 new age of precision medicine [Aronson and Rehm 2015; Oetting et al. 2016], have helped to  
8  
9 establish the clinical relevance or otherwise of CP-associated variants.  
10

11  
12 No regulatory variants had been functionally characterized in *PRSSI*, the first identified  
13  
14 CP gene, until the recent identification of a *PRSSI* promoter variant (i.e., rs4726576:C>A or  
15  
16 c.-204C>A (Fig. 1A)) that is in perfect linkage disequilibrium with the ‘CP-protective’  
17  
18 rs10273639:C>T (i.e., c.-408C>T (Fig. 1A)) variant [Boulling et al. 2015]. This finding was  
19  
20 made during the resequencing of the promoter region of *PRSSI* in 287 French Caucasian  
21  
22 individuals. Given that no other study has analyzed the *PRSSI* promoter in a systematic  
23  
24 manner, we extended the resequencing to 626 French Caucasian individuals, including 242  
25  
26 patients with idiopathic CP and 384 healthy controls. This yielded three additional single  
27  
28 nucleotide substitution variants, each being found only once in either patients or controls  
29  
30 (Table 1). All three variants were confirmed by independent PCR and sequencing. The three  
31  
32 respective carriers were not found to carry missense or splice site variants in the *PRSSI* gene.  
33  
34 None of the three variants are listed in either *Genetic risk factors in chronic pancreatitis*  
35  
36 (<http://www.pancreasgenetics.org/index.php>) or ExAC (Exome Aggregation Consortium;  
37  
38 <http://exac.broadinstitute.org/>). Informed consent was obtained from each patient and the  
39  
40 study was approved by the Ethics Committee of Brest University.  
41  
42  
43  
44

45  
46 Nearly 20 years ago, a three base-pair deletion (c.-30\_-28delTCC) in the *PRSSI* gene was  
47  
48 identified in a single individual with idiopathic CP [Férec et al. 1999]. This c.-30\_-28delTCC  
49  
50 variant was the first CP-associated regulatory variant reported in the context of the  
51  
52 aforementioned CP genes. This variant has not so far been reported in subsequent studies and  
53  
54 is also absent from ExAC. Thus, all three newly found variants (Table 1) and the previously  
55  
56 known c.-30\_-28delTCC represent rare variants. It should also be noted that these rare  
57  
58  
59  
60



1  
2  
3 variants, together with the functional rs4726576 [Boulling et al. 2015], are clustered within  
4  
5 the proximal promoter region of *PRSSI* (Fig. 1A). A similar distribution pattern has also been  
6  
7 noted for variants identified in the *SPINK1* promoter region in both CP patients and controls  
8  
9 [Boulling et al. 2011; Hegyi et al. 2016].  
10

11 The c.-30\_-28delTCC variant is currently registered in OMIM (Online Mendelian  
12 Inheritance in Man; <http://www.omim.org/>) as one of 12 selected examples of pancreatitis-  
13 associated *PRSSI* variants (as of March 10, 2016). This variant has long been speculated to  
14 increase the level of *PRSSI* transcription, thereby predisposing to CP [Férec et al. 1999]. This  
15 idea became more plausible when *PRSSI* duplication and triplication copy number mutations  
16 were identified as a novel cause of CP [Le Maréchal et al. 2006; Masson et al. 2008a]. *In vitro*  
17 functional analysis often represents the only practical means to determine the pathogenicity of  
18 patient-derived sequence variants, particularly when they are rare [MacArthur et al. 2014].  
19  
20 The recent establishment of a reporter gene assay system has now made possible the  
21 functional characterization of *PRSSI* promoter variants [Boulling et al. 2015]. Functional  
22 analysis of both c.-30\_-28delTCC and the newly found CP-associated c.184G>A was  
23 therefore performed as previously described [Boulling et al. 2015]. For the purpose of  
24 comparison, we also included the two variants found in controls, c.173C>T and c.147>T. In  
25 brief, we introduced all four variants separately into the pGL3-(A) luciferase vector (see Fig.  
26  
27 1C in [Boulling et al. 2015]), transfected the corresponding constructs into rat pancreatic  
28 acinar AR42J cells treated with dexamethasone, and determined the relative reporter gene  
29 activities against the pGL3-(A) luciferase vector.  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48

49 Two variants, c.184G>A and c.173C>T, showed no significant effect on reporter gene  
50 expression. However, the other two, c.-147C>T and c.-30\_-28delTCC, served to reduce  
51 reporter gene expression as compared with the wild-type promoter sequence (Fig. 1B).  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000

1  
2  
3 expression. In other words, c.-30\_-28delTCC might reasonably be expected to act as a  
4  
5 protective factor for CP rather than a risk factor. c.-30\_-28delTCC is located within the core  
6  
7 promoter region of the *PRSSI* gene. Within the core promoter, an optimal distance between  
8  
9 the TATA box and transcription start site (TSS) is often required for optimal transcription of  
10  
11 mammalian genes [Ponjavic et al. 2006]. We surmise that c.-30\_-28delTCC may negatively  
12  
13 affect *PRSSI* transcription by altering the optimal TATA–TSS distance. A precondition for  
14  
15 this postulate is however the presence of a TATA box upstream of the variant-affected site. A  
16  
17 search of the *PRSSI* core promoter region using the SCAN tool from the JASPAR database  
18  
19 (<http://jaspar.genereg.net/>) with the TATA-Box matrix model under default conditions yielded  
20  
21 a TATA box motif, GTATAAAGACGAGTC (underlined is the core TATA box), at positions  
22  
23 c.-46\_-32.  
24  
25  
26

27 We previously characterized the common c-408C>T and c.204C>A variants and showed  
28  
29 that the latter was of functional significance [Boulling et al. 2015]. In the present study, we  
30  
31 functionally characterized three rare variants located within the proximal promoter region of  
32  
33 the *PRSSI* gene (i.e., c.-184G>A, c.-173C>T and c.-147C>T; Fig. 1B). The availability of  
34  
35 these five functionally characterized variants prompted us to explore a possible correlation  
36  
37 between the variants' effect on reporter gene expression and their potential ability to disrupt  
38  
39 or create binding sites for transcription factors known to be expressed in the exocrine  
40  
41 pancreas. To this end, we first screened a 41 bp tract comprising 20 bp flanking each affected  
42  
43 nucleotide on both sides, in the context of both the reference and variant alleles, for  
44  
45 transcription factor binding sites, using the SCAN tool from the JASPAR database with all  
46  
47 available human JASPAR matrix models (species 9606) (<http://jaspar.genereg.net/>) under  
48  
49 default conditions. We then checked all transcription factors corresponding to the predicted  
50  
51 binding sites for expression in the exocrine pancreas by means of The Human Protein Atlas  
52  
53 database (<http://www.proteinatlas.org/>). However, we were unable to draw any definite  
54  
55  
56  
57  
58  
59  
60



1  
2  
3 conclusions from this attempt because of (i) the degenerate nature of transcription factor  
4 binding sites and (ii) inadequate information on the spatial and temporal expression patterns  
5 of the relevant transcription factors.  
6  
7

8  
9 In summary, we have discovered three new rare variants in the proximal promoter region  
10 of *PRSSI* through the resequencing of French CP patients and controls. All three newly  
11 identified variants were functionally characterized together with the previously described c.-  
12 30\_-28delTCC. This analysis revealed that c.-30\_-28delTCC resulted in reduced rather than  
13 increased gene expression, suggesting that c.-30\_-28delTCC may act as a protective factor  
14 rather than a predisposing factor for CP as originally postulated [Férec et al. 1999] . As for the  
15 three newly found rare variants, only c.147C>T, which was detected in a solitary control,  
16 exhibited a significant effect on gene expression (Fig. 1B); its negative effect on gene  
17 expression suggests that this variant may also confer protection against CP. Further, all three  
18 functional variants, viz. c.204C>A [Boulling et al. 2015] and c.-147C>T and c.-30\_-  
19 28delTCC, tested here (Fig. 1B), represent loss-of-function variants in terms of *PRSSI*  
20 expression. It may be that it is much more likely for a variant to disrupt a functional  
21 transcription factor binding site than it is for it to create a new transcription factor binding site  
22 that confers an opposing function. Irrespective of the underlying biological mechanisms, our  
23 findings illustrate the fundamental importance of functional analysis in allowing us to confirm  
24 or reject the clinical relevance of disease-associated rare variants.  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46

47 *Disclosure statement.* The authors are unaware of any conflict of interest.  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## References

- Aronson SJ, Rehm HL. 2015. Building the foundation for genomics in precision medicine. *Nature* 526:336-342.
- Beer S, Zhou J, Szabó A, Keiles S, Chandak GR, Witt H, Sahin-Tóth M. 2013. Comprehensive functional analysis of chymotrypsin C (CTRC) variants reveals distinct loss-of-function mechanisms associated with pancreatitis risk. *Gut* 62:1616-1624.
- Boulling A, Le Maréchal C, Trouvé P, Raguénès O, Chen JM, Férec C. 2007. Functional analysis of pancreatitis-associated missense mutations in the pancreatic secretory trypsin inhibitor (*SPINK1*) gene. *Eur J Hum Genet* 15:936-942.
- Boulling A, Witt H, Chandak GR, Masson E, Paliwal S, Bhaskar S, Reddy DN, Cooper DN, Chen JM, Férec C. 2011. Assessing the pathological relevance of *SPINK1* promoter variants. *Eur J Hum Genet* 19:1066-1073.
- Boulling A, Keiles S, Masson E, Chen JM, Férec C. 2012. Functional analysis of eight missense mutations in the *SPINK1* gene. *Pancreas* 41:329-330.
- Boulling A, Sato M, Masson E, Génin E, Chen JM, Férec C. 2015. Identification of a functional *PRSS1* promoter variant in linkage disequilibrium with the chronic pancreatitis-protecting rs10273639. *Gut* 64:1837-1838.
- Braganza JM, Lee SH, McCloy RF, McMahon MJ. 2011. Chronic pancreatitis. *Lancet* 377:1184-1197.
- Derikx MH, Kovacs P, Scholz M, Masson E, Chen JM, Ruffert C, Lichtner P, Te Morsche RH, Cavestro GM, Férec C, Drenth JP, Witt H, et al. 2015a. Polymorphisms at *PRSS1-PRSS2* and *CLDN2-MORC4* loci associate with alcoholic and non-alcoholic chronic pancreatitis in a European replication study. *Gut* 64:1426-1433.
- Derikx MH, Geisz A, Kereszturi É, Sahin-Tóth M. 2015b. Functional significance of *SPINK1* promoter variants in chronic pancreatitis. *Am J Physiol Gastrointest Liver Physiol* 308:G779-784.
- Férec C, Raguénès O, Salomon R, Roche C, Bernard JP, Guillot M, Quéré I, Faure C, Mercier B, Audrézet MP, Guillausseau PJ, Dupont C, et al. 1999. Mutations in the cationic trypsinogen gene and evidence for genetic heterogeneity in hereditary pancreatitis. *J Med Genet* 36:228-232.
- Fjeld K, Weiss FU, Lasher D, Rosendahl J, Chen JM, Johansson BB, Kirsten H, Ruffert C, Masson E, Steine SJ, Bugert P, Cnop M, et al. 2015. A recombined allele of the lipase gene *CEL* and its pseudogene *CELP* confers susceptibility to chronic pancreatitis. *Nat Genet* 47:518-522.
- Hegy E, Geisz A, Sahin-Tóth M, Derikx MH, Nemeth BC, Balazs A, Hritz I, Izbeki F, Halasz A, Parniczky A, Takacs T, Kelemen D, et al. 2016. *SPINK1* Promoter Variants in Chronic Pancreatitis. *Pancreas* 45:148-153.
- Király O, Wartmann T, Sahin-Tóth M. 2007. Missense mutations in pancreatic secretory trypsin inhibitor (*SPINK1*) cause intracellular retention and degradation. *Gut* 56:1433-1438.
- Le Maréchal C, Masson E, Chen JM, Morel F, Ruzsiewski P, Levy P, Férec C. 2006. Hereditary pancreatitis caused by triplication of the trypsinogen locus. *Nat Genet* 38:1372-1374.
- MacArthur DG, Manolio TA, Dimmock DP, Rehm HL, Shendure J, Abecasis GR, Adams DR, Altman RB, Antonarakis SE, Ashley EA, Barrett JC, Biesecker LG, et al. 2014. Guidelines for investigating causality of sequence variants in human disease. *Nature* 508:469-476.
- Masson E, Le Maréchal C, Chandak GR, Lamoril J, Bezieau S, Mahurkar S, Bhaskar S, Reddy DN, Chen JM, Férec C. 2008a. Trypsinogen copy number mutations in patients with idiopathic chronic pancreatitis. *Clin Gastroenterol Hepatol* 6:82-88.
- Masson E, Chen JM, Scotet V, Le Maréchal C, Férec C. 2008b. Association of rare chymotrypsinogen C (*CTRC*) gene variations in patients with idiopathic chronic pancreatitis. *Hum Genet* 123:83-91.
- Oetting WS, Brenner SE, Brookes AJ, Greenblatt MS, Hart RK, Karchin R, Sunyaev SR, Taschner PE. 2016. Pathogenicity Interpretation in the Age of Precision Medicine: The 2015 Annual Scientific Meeting of the Human Genome Variation Society. *Hum Mutat*.

- 1  
2  
3 Ponjavic J, Lenhard B, Kai C, Kawai J, Carninci P, Hayashizaki Y, Sandelin A. 2006. Transcriptional and  
4 structural impact of TATA-initiation site spacing in mammalian core promoters. *Genome Biol*  
5 7:R78.
- 6 Rosendahl J, Witt H, Szmola R, Bhatia E, Ozsvári B, Landt O, Schulz HU, Gress TM, Pfutzer R, Lohr M,  
7 Kovacs P, Bluher M, et al. 2008. Chymotrypsin C (CTRC) variants that diminish activity or  
8 secretion are associated with chronic pancreatitis. *Nat Genet* 40:78-82.
- 9 Schnúr A, Beer S, Witt H, Hegyi P, Sahin-Tóth M. 2014. Functional effects of 13 rare *PRSS1* variants  
10 presumed to cause chronic pancreatitis. *Gut* 63:337-343.
- 11 Szmola R, Sahin-Tóth M. 2007. Chymotrypsin C (caldecrin) promotes degradation of human cationic  
12 trypsin: identity with Rinderknecht's enzyme Y. *Proc Natl Acad Sci USA* 104:11227-11232.
- 13 Whitcomb DC, Gorry MC, Preston RA, Furey W, Sossenheimer MJ, Ulrich CD, Martin SP, Gates LK, Jr.,  
14 Amann ST, Toskes PP, Liddle R, McGrath K, et al. 1996. Hereditary pancreatitis is caused by a  
15 mutation in the cationic trypsinogen gene. *Nat Genet* 14:141-145.
- 16 Whitcomb DC, LaRusch J, Krasinskas AM, Klei L, Smith JP, Brand RE, Neoptolemos JP, Lerch MM,  
17 Tector M, Sandhu BS, Guda NM, Orlichenko L, et al. 2012. Common genetic variants in the  
18 *CLDN2* and *PRSS1-PRSS2* loci alter risk for alcohol-related and sporadic pancreatitis. *Nat*  
19 *Genet* 44:1349-1354.
- 20 Witt H, Luck W, Hennies HC, Classen M, Kage A, Lass U, Landt O, Becker M. 2000. Mutations in the  
21 gene encoding the serine protease inhibitor, Kazal type 1 are associated with chronic  
22 pancreatitis. *Nat Genet* 25:213-216.
- 23 Zou WB, Boulling A, Masson E, Cooper DN, Liao Z, Li ZS, Férec C, Chen JM. 2015. Clarifying the clinical  
24 relevance of *SPINK1* intronic variants in chronic pancreatitis. *Gut*.
- 25 Zou WB, Masson E, Boulling A, Cooper DN, Li ZS, Liao Z, Férec C, Chen JM. 2016. Digging deeper into  
26 the intronic sequences of the *SPINK1* gene. *Gut*.
- 27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

**Figure Legend****Figure 1.** Discovery and functional characterization of *PRSSI* promoter variants. **A:**

Representation of the *PRSSI* promoter region that was resequenced (i.e., extending from c.-530 to c.-1). The three newly detected rare variants (i.e., c.-184G>A, c.-173C>T and c.-147C>T) and a previously described rare variant (i.e., c.-30\_-28delTCC [Férec et al. 1999]), that were functionally characterized in this study, are highlighted in bold. The ‘CP-protective’ rs10273639 (i.e., c.-408C>T [Derikx et al. 2015a; Whitcomb et al. 2012]) and the functional rs4726576 (i.e., c.-204C>A [Boulling et al. 2015]) variants, are shown in normal font; these two common polymorphisms were found to be in perfect linkage disequilibrium in French Caucasians [Boulling et al. 2015]. The promoter region contained within the pGL3-(A) reporter gene vector [Boulling et al. 2015] (used for constructing the variant reporter gene vectors in this study) and the proximal promoter of the *PRSSI* gene are indicated by long and short dotted lines, respectively. ×, translational initiation codon of *PRSSI*. TSS, transcription start site. **B:** Relative promoter activities of the variant constructs in relation to the pGL3-(A) construct (wild-type), in AR42J cells treated with dexamethasone. Results represent means ± SD of 3 to 6 independent experiments. \* $P < 0.05$ ; \*\*\* $P < 0.0001$ .

**Table 1. Variants Detected in Resequencing the *PRSSI* Promoter Region**

Variant <sup>a</sup>	Patients	Controls	GenBank accession number
c.-184G>A	1/242	0/384	Pending
c.-173C>T	0/242	1/384	Pending
c.-147C>T	0/242	1/384	Pending

<sup>a</sup>Variant nomenclature followed HGVS recommendations

(<http://www.hgvs.org/mutnomen/recs.html>), with the A of the translational initiation codon ATG of the *PRSSI* gene being given as c.1. GenBank accession numbers NC\_000007.14 and NM\_002769.4 were used as reference sequences.

For Peer Review

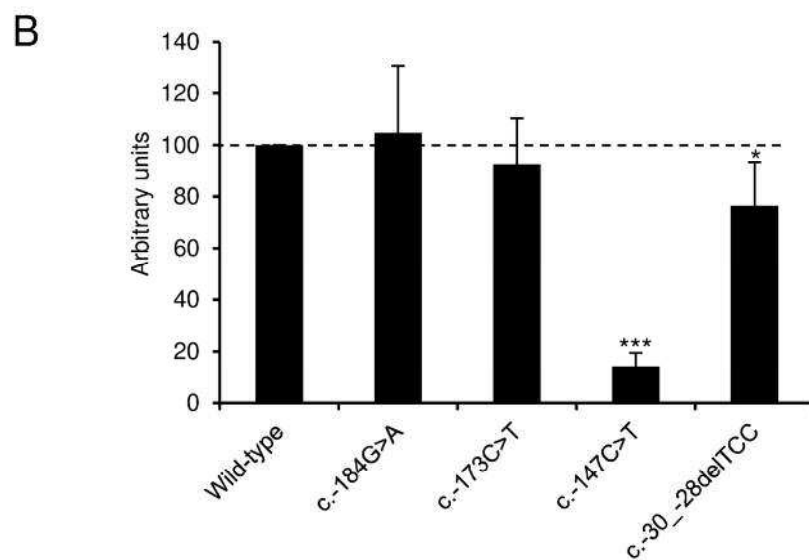
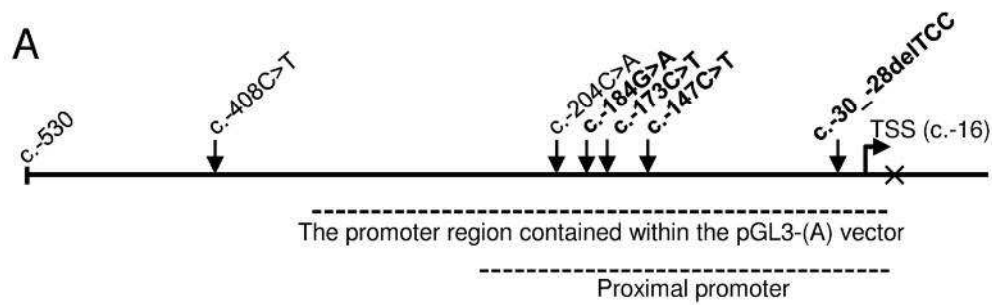


Figure 1  
110x102mm (600 x 600 DPI)