Common genetic variants in the *CLDN2* and *PRSS1-PRSS2* loci alter risk for alcohol-related and sporadic pancreatitis

David C Whitcomb[1–3], Jessica LaRusch[1], Alyssa M Krasinskas[4], Lambertus Klei[5], Jill P Smith[6], Randall E Brand[1], John P Neoptolemos[7], Markus M Lerch[8], Matt Tector[9], Bimaljit S Sandhu[10], Nalini M Guda[9], Lidiya Orlichenko[1], Alzheimer's Disease Genetics Consortium[11], Samer Alkaade[12], Stephen T Amann[13], Michelle A Anderson[14], John Baillie[15], Peter A Banks[16], Darwin Conwell[16], Gregory A Coté[17], Peter B Cotton[18], James DiSario[19], Lindsay A Farrer[20,89,90,95], Chris E Forsmark[21], Marianne Johnstone[7], Timothy B Gardner[22], Andres Gelrud[1], William Greenhalf[7], Jonathan L Haines[23,24], Douglas J Hartman[4], Robert A Hawes[18], Christopher Lawrence[18], Michele Lewis[25], Julia Mayerle[8], Richard Mayeux[26,27], Nadine M Melhem[5], Mary E Money[28], Thiruvengadam Muniraj[29], Georgios I Papachristou[1], Margaret A Pericak-Vance[30,31], Joseph Romagnuolo[18], Gerard D Schellenberg[32], Stuart Sherman[17], Peter Simon[8], Vijay P Singh[1], Adam Slivka[1], Donna Stolz[2], Robert Sutton[7], Frank Ulrich Weiss[8], C Mel Wilcox[33], Narcis Octavian Zarnescu[1], Stephen R Wisniewski[34], Michael R O'Connell[1], Michelle L Kienholz[1], Kathryn Roeder[35], M Michael Barmada[3], Dhiraj Yadav[1] & Bernie Devlin[3,5]

© 2012 Nature America, Inc. All rights reserved.

Pancreatitis is a complex, progressively destructive inflammatory disorder. Alcohol was long thought to be the primary causative agent, but genetic contributions have been of interest since the discovery that rare *PRSS1*, *CFTR* and *SPINK1* variants were associated with pancreatitis risk. We now report two associations at genome-wide significance identified and replicated at *PRSS1-PRSS2* ($P < 1 \times 10^{-12}$) and X-linked *CLDN2* ($P < 1 \times 10^{-21}$) through a two-stage genome-wide study (stage 1: 676 cases and 4,507 controls; stage 2: 910 cases and 4,170 controls). The *PRSS1* variant likely affects disease susceptibility by altering expression of the primary trypsinogen gene. The *CLDN2* risk allele is associated with atypical localization of claudin-2 in pancreatic acinar cells. The homozygous (or hemizygous in males) *CLDN2* genotype confers the greatest risk, and its alleles interact with alcohol consumption to amplify risk. These results could partially explain the high frequency of alcohol-related pancreatitis in men (male hemizygote frequency is 0.26, whereas female homozygote frequency is 0.07).

The exocrine pancreas is a simple digestive gland of only two primary cell types, each with a single function (**Supplementary Fig. 1** and **Supplementary Note**). Recurrent acute pancreatic inflammation can but does not always progress to irreversible damage of the gland, including fibrosis, atrophy, pain and exocrine and endocrine insufficiency[1–3], known as chronic pancreatitis. Different genetic and environmental factors produce the same clinical phenotype[4].

We collected biological samples and phenotypic data from 1,000 individuals with recurrent acute pancreatitis and chronic pancreatitis as well as controls in the North American Pancreatitis Study 2 (NAPS2)[5]. The primary environmental risk factor identified was heavy alcohol drinking when symptoms of pancreatitis began, defined on the basis of the assessment of the study physician, termed here alcohol-related pancreatitis.

To further define genetic risk, we conducted a two-stage (discovery and replication) genome-wide association study (GWAS). The final data set for the stage 1 cohort included 676 chronic pancreatitis cases and 4,507 controls of European ancestry (**Supplementary Figs. 2** and **3**) genotyped at 625,739 SNPs (**Table 1** and **Supplementary Table 1**). Associations at genome-wide significance ($P < 5 \times 10^{-8}$) were identified at two loci. The most highly associated SNP fell at Xq23.3, termed the *CLDN2* locus, and the other was located at 7q34, termed the *PRSS1-PRSS2* locus (**Fig. 1**, **Table 2**, **Supplementary Figs. 4** and **5** and **Supplementary Table 2**). *CLDN2* encodes the claudin-2 protein, *PRSS1* encodes cationic trypsinogen, and *PRSS2* encodes anionic trypsinogen.

The stage 2 cohort included 910 cases (331 chronic pancreatitis and 579 recurrent acute pancreatitis; **Table 1** and **Supplementary Table 1**), also genotyped at 625,739 SNPs, and 4,170 controls, most of whom were genotyped previously on the Illumina HumanOmni1_Quad_V1-0_B chip. All subjects were of European ancestry, as determined by genetic analyses. Recurrent acute pancreatitis and chronic pancreatitis were modeled as having common susceptibilities, with chronic pancreatitis occurring over time in the presence of additional disease-modifying factors[6]. It is possible that this assumption reduces power relative to that of a study comprising solely chronic pancreatitis or recurrent acute pancreatitis cases. Our primary focus in stage 2 was on

A full list of authors and affiliations appears at the end of the paper.

Received 10 June; accepted 12 October; published online 11 November 2012; doi:10.1038/ng.2466

Table 1 Characterization of cases used for GWAS

	Alcohol-related pancreatitis[a]	CP	RAP	CP + RAP
Stage 1	All cases	676	–	676
	Yes	264	–	264
	No	411	–	411
	Unknown	1	–	1
Stage 2	All cases	331	579	910
	Yes	70	113	183
	No	256	462	718
	Unknown	5	4	9
Combined	All cases	930	579	1,506
	Yes	334	113	447
	No	667	462	1,129
	Unknown	6	4	10

CP, chronic pancreatitis; RAP, recurrent acute pancreatitis. Information is not included on controls from stage 1 (n = 4,514) or stage 2 (n = 4,053). More complete characterization of samples is provided in **Supplementary Table 1**.
[a]Alcohol-related pancreatitis status was assigned by the study physician at enrollment.

the *PRSS1-PRSS2* and *CLDN2* loci, although we also conducted a joint analysis[7] of data from stages 1 and 2 to uncover any new risk loci. After controlling for ancestry, these data showed significant effects for the *CLDN2* and *PRSS1-PRSS2* loci (**Fig. 1**, **Supplementary Figs. 6** and **7** and **Supplementary Tables 2** and **3**). The quality of the SNP genotypes supported the associations (**Supplementary Fig. 8**). The frequencies of the putative risk alleles at these two loci were 0.57 for the C allele at rs10273639 (*PRSS1-PRSS2* locus), with the minor T allele reducing risk, and 0.26 for the T allele at rs12688220 (*CLDN2* locus). No other locus showed association after accounting for SNP genotype quality (**Supplementary Figs. 6–8**).

PRSS1 gain-of-function mutations, such as the one encoding a p.Arg122His alteration, increase risk for recurrent acute pancreatitis and chronic pancreatitis[8], as does increased copy number[9,10], whereas rare loss-of-function mutations in *PRSS2* are protective[11]. rs10273639 is in the 5′ promoter region of *PRSS1*. Because it is the only highly associated SNP in the locus, we validated its genotypes by independent TaqMan genotyping and also genotyped two SNPs in linkage disequilibrium with it (**Supplementary Table 4**)[12,13]. We screened *PRSS1* for rare variants in 1,138 subjects, 418 with chronic pancreatitis, 350 with recurrent acute pancreatitis and 379 controls. Three known disease-associated variants (encoding p.Ala16Val, p.Asn29Ile and p.Arg122His alterations) were identified in 23 subjects (**Supplementary Table 4**). These gain-of-function variants occurred almost exclusively in cases (22 out of 23), and 2 of them (encoding p.Ala16Val and p.Arg122His alterations) are likely included on the C or risk haplotype of this locus (**Supplementary Table 4**). Nonetheless, with only 19 cases harboring alleles encoding p.Ala16Val and p.Arg122His alterations, these rare alleles cannot account for the association observed at this locus.

We genotyped 69 control pancreas tissue samples from 3 sources at rs10273639 and used cDNA to quantify expression of *PRSS1* and control genes (**Supplementary Table 5**). In all three sets of quantitative PCR data, the slope relating the count of the C allele genotype to *PRSS1* expression was positive. Taken together, the samples provided evidence that alleles at rs10273639 affect expression of *PRSS1* (P = 0.01): expression levels were highest in subjects homozygous for the C allele at rs10273639, intermediate in heterozygotes and lowest in subjects homozygous for the T allele. On the basis of this evidence, we posit that reduced trypsinogen expression protects the pancreas from injury, as has been observed in mouse genetic models[14].

CLDN2 is considered the primary candidate gene within the X-linked locus. Claudin-2 is an attractive candidate because it serves as a highly regulated tight junction protein, forming low-resistance, cation-selective ion and water channels between endothelial cells[15,16], and is normally expressed at low levels in the tight junction between cells of the pancreatic ducts and in pancreatic islets[17,18]. The *CLDN2* promoter includes a nuclear factor (NF)-κB–binding site[19], and *CLDN2* expression is enhanced in other cells under conditions associated with injury or stress[20–22]. Claudin-2 can also be expressed by acinar cells when stressed, as reported in porcine models of acute pancreatitis[23]. Other genes within the *CLDN2* locus include *MORC4*, *RIPPLY1* and *TBC1D8B*. *MORC4* is expressed at low levels in most tissues, including the pancreas, with higher levels detected in the placenta and testes[24]. The MORC4 protein contains a CW four-cysteine zinc-finger motif, nuclear localization signal and nuclear matrix–binding domain, suggesting that it may be a transcription factor[24], but its expression does not seem to correlate with pancreatitis (**Supplementary Fig. 9**). *RIPPLY1* and *TBC1D8B* are not known to be expressed in the pancreas.

To our knowledge, genetic variation in *CLDN2* has not previously been associated with disease in humans. We assessed DNA sequence variants around *CLDN2* and RNA and protein expression for claudin-2 in control tissue classified by histology and genotype (**Supplementary Fig. 10** and **Supplementary Table 6**). Evaluating 1000 Genomes Project data, we did not identify exonic variation that could explain the association signal. Using similar methods to those described for the analysis of *PRSS1* expression, we determined that *CLDN2* expression levels in control tissues did not correlate with genotype at the *CLDN2* risk locus (P = 0.32). Protein blot analysis of protein extracted from the tissue

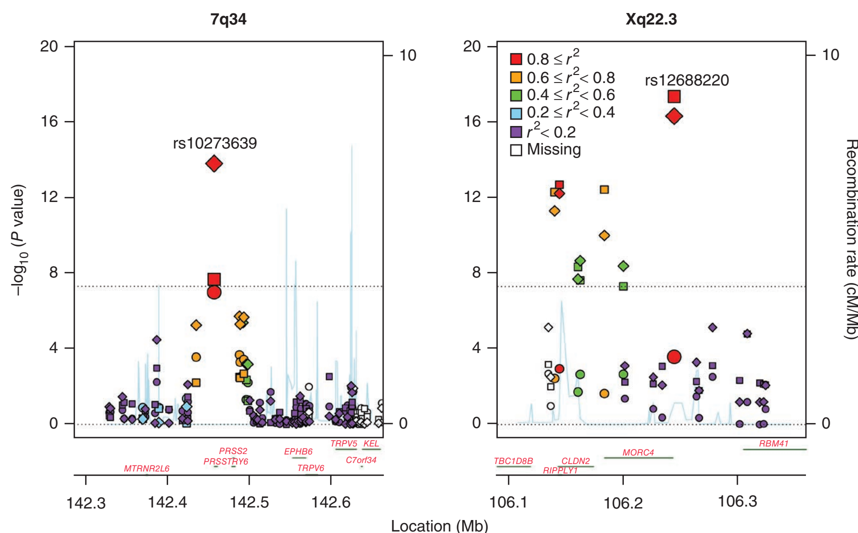

Figure 1 Regional association plots for analysis of chronic and recurrent acute pancreatitis. Manhattan plots showing the $-\log_{10}$ (P value) for the association of each SNP with affected status for all SNPs passing quality control filters and falling within selected regions of the *PRSS1-PRSS2* (left) and *CLDN2* (right) loci. Regions were selected to highlight the most associated SNPs. Squares, stage 1 data; circles, stage 2 data; diamonds, combined stage 1 and 2 data. After accounting for the most highly associated SNP at each locus, no other SNP approached genome-wide-significance for association (indicated by top horizontal dotted line).

© 2012 Nature America, Inc. All rights reserved.
npg

Table 2 Results for lead SNPs at the *PRSS1-PRSS2* and *CLDN2* loci from stage 1, stage 2 and joint analysis

					CP + RAP		CP			CP + RAP			CP + RAP		
					Allele frequency (A1)		Stage 1			Stage 2			Combined		
Chr.	SNP	Position	A1[a]	A2	Cases	Controls	OR	se(OR)	*P*	OR	se(OR)	*P*	OR	se(OR)	*P*
7	rs10273639	142456928	T	C	0.350	0.424	0.712	0.044	3.0×10^{-8}	0.748	0.039	7.5×10^{-8}	0.734	0.029	2.0×10^{-14}
X	rs7057398	106144529	C	T	0.374	0.281	1.493	0.075	1.4×10^{-15}	1.210	0.066	1.8×10^{-5}	1.321	0.049	4.6×10^{-17}
X	rs12688220	106244767	T	C	0.367	0.261	1.612	0.081	2.4×10^{-21}	1.238	0.073	2.3×10^{-6}	1.385	0.054	2.3×10^{-22}

Chr., chromosome; OR, odds ratio; se(OR), standard error of the odds ratio.
[a]A1 is the allele counted for purposes of computing the odds ratio and associated statistics. The model used here includes covariates to control for the two leading eigenvectors for ancestry, as was done in the PLINK analyses, but differs in its treatment of the minor allele count for the *CLDN2* locus, which resides on the X chromosome (Online Methods). Alleles given are refSNP alleles according to dbSNP. All SNPs passing quality control and showing association at $P < 5 \times 10^{-7}$ in stage 1 or 2 or the joint analysis are listed in **Supplementary Table 2**.

with antibody to claudin-2 detected only one protein band of the appropriate size, the expression of which correlated with tissue inflammation, as determined by systematic grading of histology in adjacent tissue (**Fig. 2a** and **Supplementary Fig. 10**). Immunohistochemical staining with antibody to claudin-2 was verified in normal tissue (**Fig. 2b**), with kidney, duodenum and bile ducts serving as additional positive controls (data not shown). We assessed protein localization in 12 GWAS cases who underwent pancreatic surgery (6 with the high-risk genotype at *CLDN2* and 6 without this genotype). Claudin-2 staining of cytoplasmic granules was markedly higher in both duct and acinar cells in chronic pancreatitis cases as compared with controls. (**Fig. 2c**–**e**). Only chronic pancreatitis cases with the high-risk *CLDN2* genotype showed moderate-to-strong claudin-2 staining along the basolateral membrane of acinar cells (**Fig. 2d**,**e** and **Supplementary Table 6**). Claudin-2 was also expressed in macrophages, which could contribute to the pathological inflammatory process[25] (**Fig. 2c**,**f**).

Most studies report excessive alcohol consumption as the major risk factor for adult-onset chronic pancreatitis[26–29]. However, only 3% of individuals who are alcoholics develop chronic pancreatitis[30], suggesting a pancreas-specific risk factor. We compared genotypes on the basis of whether pancreatitis was alcohol related (yes/no)[5,31]. Setting control genotype counts as the baseline category against which case genotypes were compared, the jointly estimated odds ratios for cases with positive alcohol-related pancreatitis were greater for both rs10273639 (*PRSS1-PRSS2* locus) and rs12688220 (*CLDN2* locus) than those estimated for cases with pancreatitis not related to alcohol consumption (**Table 3**). Thus, the effects of both loci seemed to be amplified by alcohol consumption. In a case-only analysis, both loci seemed to interact with alcohol-related pancreatitis (**Table 3**), with the association for the *CLDN2* locus being more prominent ($P = 4 \times 10^{-7}$).

We conclude that a common allele in the *PRSS1-PRSS2* locus is associated with lower *PRSS1* gene expression and that this effect is independent of the previously reported rare gain-of-function *PRSS1* variants that increase susceptibility to both recurrent acute pancreatitis and chronic pancreatitis[8]. For this reason and because risk variants at the *PRSS1-PRSS2* locus exert a similar effect in subjects with recurrent acute pancreatitis and those with chronic pancreatitis, it is reasonable to conjecture that variation at rs10273639 or variation at sites in linkage disequilibrium with it directly affects risk for chronic pancreatitis and recurrent acute pancreatitis through its impact on trypsinogen expression. Variation at the *CLDN2* locus, however, is much more strongly associated with chronic pancreatitis than recurrent acute pancreatitis, suggesting that it probably acts as a disease modifier to accelerate the transition from recurrent acute pancreatitis to chronic pancreatitis. The significant association of the *CLDN2*

© 2012 Nature America, Inc. All rights reserved.

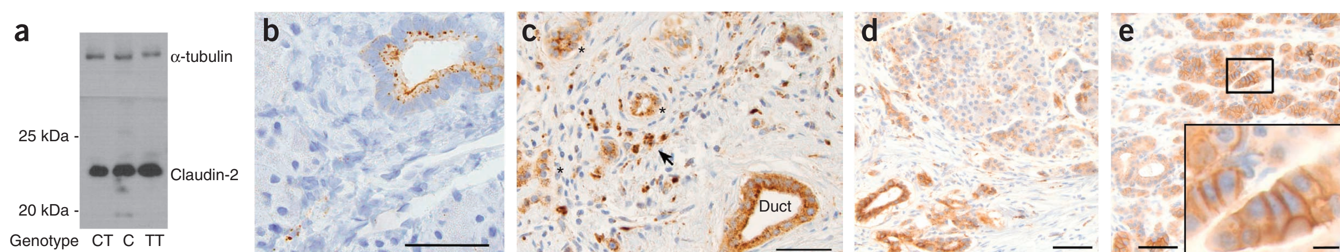

Figure 2 Expression and localization of claudin-2 in the human pancreas relative to rs12688220 genotype. (**a**) Protein blot analysis, using mouse antibody to claudin-2, of three control samples genotyped at rs12688220 (C, male; the TT genotype confers high risk). The antibody reacts with a protein of ~22–23 kDa, consistent with the size of claudin-2. Histology of adjacent tissue showed that samples had inflammation and/or fibrosis. α-tubulin was used as a loading control. Blots from all controls are presented in **Supplementary Figure 10**. (**b**) Staining with mouse antibody to claudin-2 (brown) of normal-appearing control tissue localizing to ducts but not acinar cells. Scale bar, 50 μm. (**c**) Severe chronic pancreatitis in a case with the high-risk T (male) genotype. Claudin-2 staining localizes to the intralobular duct (duct), atrophic acini (*) and cells with the morphological appearance of macrophages (arrow). Scale bar, 50 μm. (**d**) Chronic pancreatitis tissue from a case with the low-risk genotype (C, male) with staining localizing to the duct and granular staining in acinar cells. Scale bar, 100 μm. (**e**) Chronic pancreatitis case with the high-risk genotype showing intense staining of acinar cell basolateral membrane. Scale bar, 100 μm. Inset, enlarged view of the boxed region; scale bar, 10 μm. (**f**) Immunofluorescent staining of control human pancreatic tissue for claudin-2 (red), showing localization to ducts (*) and colocalization with the macrophage marker CD68 (green). Colocalization is shown in yellow (arrows). Nuclei are stained with Hoechst's dye (blue). Scale bar, 100 μm.

Table 3 Allele frequencies for rs10273639 and rs12688220 when data are stratified by controls or cases with or without alcohol-related pancreatitis

Status	Alcohol related	Number of individuals	rs10273639[C] frequency	rs12688220[T] frequency
Control	–	8,029	0.576	0.261
Pancreatitis	No	1,129	0.634	0.322
	Yes	447	0.696	0.427

Using data from cases only in a joint analysis of both SNPs, rs12688220 predicts alcohol-related pancreatitis on the basis of genotype ($\chi^2 = 29.57$; degrees of freedom (DF) = 2; $P = 4 \times 10^{-7}$) or count of risk alleles ($\chi^2 = 13.17$; DF = 1; $P = 3 \times 10^{-4}$). rs10273639 (*PRSS1-PRSS2* locus) is a significant predictor (count of risk alleles: $\chi^2 = 5.68$; DF = 1; $P = 0.017$; genotypes: $\chi^2 = 6.05$; DF = 2; $P = 0.049$), even after accounting for the effects of rs12688220.

locus with alcohol-related disease suggests that the high-risk allele in the *CLDN2* locus may modify risk through a non-trypsin–dependent process. Thus, we have characterized two common genetic risk modifiers for sporadic and alcohol-related chronic pancreatitis.

URLs. The Developmental Studies Hybridoma Bank at the University of Iowa, http://dshb.biology.uiowa.edu/Antibody-list.

METHODS

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

Note : Supplementary information is available in the online version of the paper.

© 2012 Nature America, Inc. All rights reserved. npg

ACKNOWLEDGMENTS

Technical support was provided by K. Stello, S. Das, D. Dwyer, A. Rowland, P.A. Blake, M. Ross, C. McGovern, L. Kish, H. Nawaz, S. Solomon, S. Boggiano, R. Ostroff, M. Goss and J. Timm. Data management was provided by L. Silfies and D. Protivnak. Clinical support was provided by S. Boggaino, M. Hendricks, B. Elinoff, L. McHenry, G. Lehman, J. Watkins, E. Fogel and L. Lazzell-Pannell. Additional samples were provided by F. Burton (deceased; St. Louis University School of Medicine); S. Lo (Cedars-Sinai Medical Center, University of California, Los Angeles); M.T. DeMeo (Rush University Medical Center); W.M. Steinberg (Washington Hospital Center); M.L. Kochman (University of Pennsylvania); B. Etemad (Ochsner Medical Center); and H. Zeh, A.J. Moser and K.K. Lee (University of Pittsburgh). This publication was made possible by grants R56DK061451 (D.C.W.), R01DK054709 (D.C.W.), T32DK063922 (D.C.W.), R01MH057881 (B.D. and K.R.), R01CA117926 (J.P.S.), R01DK077906 (D.C.W. and D.Y.), UL1RR024153 and UL1TR000005 from the US National Institutes of Health (NIH). Its contents are solely the responsibility of the authors and do not necessarily represent the official view of the NIH. This project used the University of Pittsburgh Genomics and Proteomics Core Laboratories (UL1RR024153, UL1RR024153 and UL1TR000005) and the University of Pittsburgh Cancer Institute (UPCI) Clinical Genomics Immunoproteomics and Sequencing Facility (NIH P30CA047904). J.L. was supported by Digestive Disease Training Program T32DK063922 (D.C.W.). N.O.Z. was supported by the American Gastroenterology Association John I Isenberg International Scholar Award. The Liverpool cohort was supported by the National Institute for Health Research (NIHR) Biomedical Research Unit award. Additional support was provided by the National Pancreas Foundation (D.C.W.), the Frieda G. and Saul F. Shapira BRCA-Associated Cancer Research Program (D.C.W.) and the Wayne Fusaro Pancreatic Cancer Research Fund (D.C.W.) and Gift of Life Foundation (J.P.S.).

Support for control cohort 1.
Alzheimer's Disease Genetics Consortium (ADGC): The NIH–National Institute on Aging (NIH-NIA) supported this work through the following grants: ADGC, U01AG032984 and RC2AG036528; National Alzheimer's Coordinating Center (NACC), U01AG016976; National Cell Repository for Alzheimer's Disease (NCRAD), U24AG021886; Banner Sun Health Research Institute, P30AG019610; Boston University, P30AG013846, U01AG10483, R01CA129769, R01MH080295, R01AG017173, R01AG025259 and R01AG33193; Columbia University, P50AG008702 and R37AG015473; Duke University, P30AG028377 and AG05128; Emory University, AG025688; Indiana University, P30AG10133; Johns Hopkins University, P50AG005146 and R01AG020688; Massachusetts General Hospital, P50AG005134; Mayo Clinic, P50AG016574; Mount Sinai School of Medicine, P50AG005138 and P01AG002219; New York University, P30AG08051, MO1RR00096 and UL1RR029893; Northwestern University, P30AG013854; Oregon Health & Science University, P30AG008017 and R01AG026916; Rush University, P30AG010161, R01AG019085, R01AG15819, R01AG17917 and R01AG30146; University of Alabama at Birmingham, P50AG016582 and UL1RR02777; University of Arizona, R01AG031581; University of California, Davis, P30AG010129; University of California, Irvine, P50AG016573, P50AG016575, P50AG016576 and P50AG016577; University of California, Los Angeles, P50AG016570; University of California, San Diego, P50AG005131; University of California, San Francisco, P50AG023501 and P01AG019724; University of Kentucky, P30AG028383 and AG05144; University of Michigan, P50AG008671; University of Pennsylvania, P30AG010124; University of Pittsburgh, P50AG005133 and AG030653; University of Southern California, P50AG005142; University of Texas Southwestern, P30AG012300; University of Miami, R01AG027944, AG010491, AG027944, AG021547 and AG019757; University of Washington, P50AG005136; Vanderbilt University, R01AG019085; and Washington University, P50AG005681 and P01AG03991. The Kathleen Price Bryan Brain Bank at the Duke University Medical Center is funded by National Institute of Neurological Disorders and Stroke (NINDS) grant NS39764 and National Institute of Mental Health (NIMH) grant MH60451 and by GlaxoSmithKline. We thank D.S. Snyder and M. Miller who are ex officio ADGC members.

Support for control group 2.
Control group 2 genotypes were obtained from the Genome-Wide Association Study of Parkinson Disease: Genes and Environment, database of Genotypes and Phenotypes (dbGaP) study phs000196.v2.p1. This NeuroGenetics Research Consortium (NGRC) is a gene-environment study of Parkinson's disease. The principal investigator is H. Payami (New York State Department of Health Wadsworth Center, Albany, New York, USA), and the co-investigators are J. Nutt (Oregon Health & Sciences University, Portland, Oregon, USA); C. Zabetian (University of Washington and Puget Sound Veterans Medical Center, Seattle, Washington, USA); S. Factor (Emory University, Atlanta, Georgia, USA); E. Molho (Albany Medical Center, Albany, New York, USA); and D. Higgins (Albany Medical Center and Albany Veterans Medical Center, Albany, New York, USA). Funding was provided by NIH grant R01NS36960. Genotyping was performed by the Genotyping Center at the Johns Hopkins University Center for Inherited Disease Research (CIDR). Funding for genotyping was provided by NIH grant HHSN268200782096C. The NIH contract for the project is entitled 'High throughput genotyping for studying the genetic contributions to human disease'.

Support for the German cohort.
The German cohort was supported by the Alfried-Krupp-von-Bohlen-und-Hahlbach-Foundation (Graduate Schools of Tumour Biology and Free Radical Biology), the Deutsche Krebshilfe/Dr. Mildred-Scheel-Stiftung (109102), the Deutsche Forschungsgemeinschaft (DFG GRK840-E3/E4, MA 4115/1-2/3 and NI 1297/1-1), the Federal Ministry of Education and Research (BMBF GANI-MED 03152061A and BMBF 0314107) and the European Union (EU-FP-7; EPC-TM and EU-FP7-REGPOT-2010-1).

AUTHOR CONTRIBUTIONS

The writing group comprised A.M.K., B.D., D.C.W., J.L., L.K. and M.L.K. Project design, management and coordination were carried out by B.D., D.C.W., D.Y., J.P.N., M.M.B., M.M.L., N.M.M. and S.R.W. Sample collection and phenotyping were performed by A.G., A.S., B.S.S., C.E.F., C.L., C.M.W., D.C., D.C.W., D.Y., F.U.W., G.A.C., G.I.P., J.B., J.D., J.M., J.P.N., J.P.S., J.R., M.A.A., M.E.M., M.J., M.L., M.M.B., M.M.L., M.T., N.M.G., P.A.B., P.B.C., P.S., R.A.H., R.E.B., R.S., S.A., S.S., S.T.A., T.B.G., T.M. and W.G. Genotyping and expression studies were performed by A.M.K., D.J.H., D.S., G.D.S., J.L., J.L.H., J.P.S., L.A.F., L.O., M.A.P.-V., N.O.Z., R.M. and V.K.S. A.M.K. reviewed immunohistochemical staining. Statistical analysis was carried out by B.D., D.Y., K.R., L.K., M.M.B. and M.R.O.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/doifinder/10.1038/ng.2466.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Etemad, B. & Whitcomb, D.C. Chronic pancreatitis: diagnosis, classification, and new genetic developments. *Gastroenterology* **120**, 682–707 (2001).
2. Chen, J.M. & Ferec, C. Chronic pancreatitis: genetics and pathogenesis. *Annu. Rev. Genomics Hum. Genet.* **10**, 63–87 (2009).
3. Witt, H., Apte, M.V., Keim, V. & Wilson, J.S. Chronic pancreatitis: challenges and advances in pathogenesis, genetics, diagnosis, and therapy. *Gastroenterology* **132**, 1557–1573 (2007).

4. Whitcomb, D.C. What is personalized medicine and what should it replace? *Nat. Rev. Gastroenterol. Hepatol.* **9**, 418–424 (2012).
5. Whitcomb, D.C. *et al.* Multicenter approach to recurrent acute and chronic pancreatitis in the United States: the North American Pancreatitis Study 2 (NAPS2). *Pancreatology* **8**, 520–531 (2008).
6. Yadav, D. & Whitcomb, D.C. The role of alcohol and smoking in pancreatitis. *Nat. Rev. Gastroenterol. Hepatol.* **7**, 131–145 (2010).
7. Skol, A.D., Scott, L.J., Abecasis, G.R. & Boehnke, M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat. Genet.* **38**, 209–213 (2006).
8. Whitcomb, D.C. *et al.* Hereditary pancreatitis is caused by a mutation in the cationic trypsinogen gene. *Nat. Genet.* **14**, 141–145 (1996).
9. Masson, E. *et al.* Trypsinogen copy number mutations in patients with idiopathic chronic pancreatitis. *Clin. Gastroenterol. Hepatol.* **6**, 82–88 (2008).
10. LaRusch, J., Barmada, M.M., Solomon, S. & Whitcomb, D.C. Whole exome sequencing identifies multiple, complex etiologies in an idiopathic hereditary pancreatitis kindred. *JOP* **13**, 258–262 (2012).
11. Witt, H. *et al.* A degradation-sensitive anionic trypsinogen (*PRSS2*) variant protects against chronic pancreatitis. *Nat. Genet.* **38**, 668–673 (2006).
12. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–1073 (2010).
13. Johnson, A.D. *et al.* SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* **24**, 2938–2939 (2008).
14. Dawra, R. *et al.* Intra-acinar trypsinogen activation mediates early stages of pancreatic injury but not inflammation in mice with acute pancreatitis. *Gastroenterology* **141**, 2210–2217 (2011).
15. Van Itallie, C.M. *et al.* The density of small tight junction pores varies among cell types and is increased by expression of claudin-2. *J. Cell Sci.* **121**, 298–305 (2008).
16. Amasheh, S. *et al.* Claudin-2 expression induces cation-selective channels in tight junctions of epithelial cells. *J. Cell Sci.* **115**, 4969–4976 (2002).
17. Lee, J.H. *et al.* Immunohistochemical analysis of claudin expression in pancreatic cystic tumors. *Oncol. Rep.* **25**, 971–978 (2011).
18. Aung, P.P. *et al.* Differential expression of claudin-2 in normal human tissues and gastrointestinal carcinomas. *Virchows Arch.* **448**, 428–434 (2006).
19. Sakaguchi, T. *et al.* Cloning of the human claudin-2 5′-flanking region revealed a TATA-less promoter with conserved binding sites in mouse and human for caudal-related homeodomain proteins and hepatocyte nuclear factor-1α. *J. Biol. Chem.* **277**, 21361–21370 (2002).
20. Mankertz, J. *et al.* TNFα up-regulates claudin-2 expression in epithelial HT-29/B6 cells via phosphatidylinositol-3-kinase signaling. *Cell Tissue Res.* **336**, 67–77 (2009).
21. Suzuki, T., Yoshinaga, N. & Tanabe, S. Interleukin-6 (IL-6) regulates claudin-2 expression and tight junction permeability in intestinal epithelium. *J. Biol. Chem.* **286**, 31263–31271 (2011).
22. Mankertz, J. *et al.* Functional crosstalk between Wnt signaling and Cdx-related transcriptional activation in the regulation of the claudin-2 promoter activity. *Biochem. Biophys. Res. Commun.* **314**, 1001–1007 (2004).
23. Meriläinen, S. *et al.* Acute edematous and necrotic pancreatitis in a porcine model. *Scand. J. Gastroenterol.* **43**, 1259–1268 (2008).
24. Liggins, A.P. *et al.* MORC4, a novel member of the MORC family, is highly expressed in a subset of diffuse large B-cell lymphomas. *Br. J. Haematol.* **138**, 479–486 (2007).
25. Van den Bossche, J. *et al.* Claudin-1, claudin-2 and claudin-11 genes differentially associate with distinct types of anti-inflammatory macrophages *in vitro* and with parasite- and tumor-elicited macrophages *in vivo. Scand. J. Immunol.* **75**, 588–598 (2012).
26. Ammann, R.W., Akovbiantz, A., Largiader, F. & Schueler, G. Course and outcome of chronic pancreatitis. Longitudinal study of a mixed medical-surgical series of 245 patients. *Gastroenterology* **86**, 820–828 (1984).
27. Marks, I.N., Bank, S. & Louw, J.H. Chronic pancreatitis in the Western Cape. *Digestion* **9**, 447–453 (1973).
28. Robles-Díaz, G., Vargas, F., Uscanga, L. & Fernandez-del Castillo, C. Chronic pancreatitis in Mexico City. *Pancreas* **5**, 479–483 (1990).
29. Irving, H.M., Samokhvalov, A.V. & Rehm, J. Alcohol as a risk factor for pancreatitis. A systematic review and meta-analysis. *JOP* **10**, 387–392 (2009).
30. Yadav, D., Eigenbrodt, M.L., Briggs, M.J., Williams, D.K. & Wiseman, E.J. Pancreatitis: prevalence and risk factors among male veterans in a detoxification program. *Pancreas* **34**, 390–398 (2007).
31. Yadav, D. *et al.* Alcohol consumption, cigarette smoking, and the risk of recurrent acute and chronic pancreatitis. *Arch. Intern. Med.* **169**, 1035–1045 (2009).

© 2012 Nature America, Inc. All rights reserved.

Alzheimer's Disease Genetics Consortium: Marilyn S Albert[36], Roger L Albin[37,38], Liana G Apostolova[39], Steven E Arnold[40], Clinton T Baldwin[41], Robert Barber[42], Lisa L Barnes[43,44], Thomas G Beach[45], Gary W Beecham[30,31], Duane Beekly[46], David A Bennett[43,47], Eileen H Bigio[48], Thomas D Bird[49], Deborah Blacker[50,51], Adam Boxer[52], James R Burke[15], Joseph D Buxbaum[53–55], Nigel J Cairns[56], Laura B Cantwell[32], Chuanhai Cao[57], Regina M Carney[58], Steven L Carroll[59], Helena C Chui[60], David G Clark[61], David H Cribbs[62], Elizabeth A Crocco[63], Carlos Cruchaga[64,65], Charles DeCarli[66], F Yesim Demirci[3], Malcolm Dick[67], Dennis W Dickson[68], Ranjan Duara[69], Nilufer Ertekin-Taner[68,70], Kelley M Faber[71], Kenneth B Fallon[59], Martin R Farlow[72], Steven Ferris[73], Tatiana M Foroud[71], Matthew P Frosch[74], Douglas R Galasko[75], Mary Ganguli[5], Marla Gearing[76,77], Daniel H Geschwind[78], Bernardino Ghetti[79], John R Gilbert[30,31], Sid Gilman[37], Jonathan D Glass[80], Alison M Goate[64,65], Neill R Graff-Radford[68,70], Robert C Green[81], John H Growdon[82], Hakon Hakonarson[83], Kara L Hamilton-Nelson[30,31], Ronald L Hamilton[84], Lindy E Harrell[61], Elizabeth Head[85], Lawrence S Honig[27], Christine M Hulette[86], Bradley T Hyman[82], Gregory A Jicha[87], Lee-Way Jin[88], Gyungah Jun[41,89,90], M Ilyas Kamboh[3,91], Anna Karydas[52], Jeffrey A Kaye[92,93], Ronald Kim[94], Edward H Koo[75], Neil W Kowall[95,96], Joel H Kramer[97], Patricia Kramer[92,98], Walter A Kukull[99], Frank M LaFerla[100], James J Lah[80], James B Leverenz[101], Allan I Levey[80], Ge Li[102], Chiao-Feng Lin[32], Andrew P Lieberman[103], Oscar L Lopez[91], Kathryn L Lunetta[90], Constantine G Lyketsos[104], Wendy J Mack[105], Daniel C Marson[61], Eden R Martin[30,31], Frank Martiniuk[106], Deborah C Mash[107], Eliezer Masliah[75,108], Ann C McKee[95,96], Marsel Mesulam[109], Bruce L Miller[52], Carol A Miller[110], Joshua W Miller[88], Thomas J Montine[101], John C Morris[56,111], Jill R Murrell[71,79], Adam C Naj[30,31,130], John M Olichney[66], Joseph E Parisi[112,113], Elaine Peskind[102], Ronald C Petersen[114], Aimee Pierce[62], Wayne W Poon[67], Huntington Potter[57], Joseph F Quinn[92], Ashok Raj[57], Murray Raskind[102], Eric M Reiman[115–117], Barry Reisberg[73,118], Christiane Reitz[26,27], John M Ringman[39], Erik D Roberson[61], Howard J Rosen[52], Roger N Rosenberg[119], Mary Sano[54], Andrew J Saykin[71,120], Julie A Schneider[43,121], Lon S Schneider[60,122], William W Seeley[52], Amanda G Smith[57], Joshua A Sonnen[101], Salvatore Spina[79], Robert A Stern[95], Rudolph E Tanzi[82], John Q Trojanowski[32], Juan C Troncoso[123], Debby W Tsuang[102], Otto Valladares[32], Vivianna M Van Deerlin[32], Linda J Van Eldik[124], Badri N Vardarajan[26,27], Harry V Vinters[39,125], Jean Paul Vonsattel[27], Li-San Wang[32], Sandra Weintraub[109], Kathleen A Welsh-Bohmer[15,126], Jennifer Williamson[27], Randall L Woltjer[127], Clinton B Wright[128], Steven G Younkin[68], Chang-En Yu[129] & Lei Yu[43]

[1]Division of Gastroenterology, Hepatology and Nutrition, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA. [2]Department of Cell Biology and Physiology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA. [3]Department of Human Genetics, University of Pittsburgh Graduate School of Public Health, Pittsburgh, Pennsylvania, USA. [4]Department of Pathology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA. [5]Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA. [6]Division of Gastroenterology and Liver Diseases, Pennsylvania State University, Hershey, Pennsylvania, USA. [7]Department of Molecular and Clinical Cancer Medicine, National Institute for Health Research Liverpool Pancreas Biomedical Research Unit, Institute of Translational Medicine, University of Liverpool, Liverpool, UK. [8]Department of Medicine A, University Medicine Greifswald, Greifswald, Germany. [9]Aurora Health Care System, St. Luke's Medical Center, Milwaukee, Wisconsin, USA. [10]Division of Gastroenterology, Hepatology and Nutrition, Virginia Commonwealth University Medical Center, Richmond, Virginia, USA. [11]A full list of members and affiliations appears at the end of this paper. [12]Department of Medicine, Saint Louis University, St. Louis, Missouri, USA. [13]North Mississippi Medical Center, Tupelo, Mississippi, USA. [14]Division of Gastroenterology, University of Michigan, Ann Arbor, Michigan, USA. [15]Department of Medicine, Duke University Medical Center, Durham, North Carolina, USA. [16]Division of Gastroenterology, Brigham and Women's Hospital, Boston, Massachusetts, USA. [17]Department of Medicine, Indiana University Medical Center, Indianapolis, Indiana, USA. [18]Digestive Disease Center, Medical University of South Carolina, Charleston, South Carolina, USA. [19]Department of Medicine, University of Utah Health Science Center, Salt Lake City, Utah, USA. [20]Department of Medicine/Biomedical Genetics Section, Boston University School of Medicine, Boston, Massachusetts, USA. [21]Department of Medicine, University of Florida, Gainesville, Florida, USA. [22]Department of Medicine, Division of Gastroenterology and Hepatology, Dartmouth-Hitchcock Medical Center, Hanover, New Hampshire, USA. [23]Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, Tennessee, USA. [24]Vanderbilt Center for Human Genetics Research, Vanderbilt University, Nashville, Tennessee, USA. [25]Division of Gastroenterology and Hepatology, Mayo Clinic, Jacksonville, Florida, USA. [26]Gertrude H Sergievsky Center, Department of Neurology, Columbia University, New York, New York, USA. [27]Taub Institute on Alzheimer's Disease and the Aging Brain, Department of Neurology, Columbia University, New York, New York, USA. [28]Meritus Medical Center, Hagerstown, Maryland, USA. [29]Department of Internal Medicine, Yale University and Griffin Hospital, Derby, Connecticut, USA. [30]The John P Hussman Institute for Human Genomics, University of Miami, Miami, Florida, USA. [31]Dr John T Macdonald Foundation Department of Human Genetics, University of Miami, Miami, Florida, USA. [32]Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [33]Department of Medicine, Division of Gastroenterology and Hepatology, University of Alabama at Birmingham, Birmingham, Alabama, USA. [34]Department of Epidemiology, Epidemiology Data Center, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania, USA. [35]Department of Statistics, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA. [36]Department of Neurology, Johns Hopkins University, Baltimore, Maryland, USA. [37]Department of Neurology, University of Michigan, Ann Arbor, Michigan, USA. [38]Geriatric Research, Education and Clinical Center (GRECC), VA Ann Arbor Healthcare System (VAAAHS), Ann Arbor, Michigan, USA. [39]Department of Neurology, University of California, Los Angeles, Los Angeles, California, USA. [40]Department of Psychiatry, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [41]Department of Medicine (Genetics Program), Boston University, Boston, Massachusetts, USA. [42]Department of Pharmacology and Neuroscience, University of North Texas Health Science Center, Fort Worth, Texas, USA. [43]Department of Neurological Sciences, Rush University Medical Center, Chicago, Illinois, USA. [44]Department of Behavioral Sciences, Rush University Medical Center, Chicago, Illinois, USA. [45]Civin Laboratory for Neuropathology, Banner Sun Health Research Institute, Phoenix, Arizona, USA. [46]National Alzheimer's Coordinating Center, University of Washington, Seattle, Washington, USA. [47]Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, Illinois, USA. [48]Department of Pathology, Northwestern University, Chicago, Illinois, USA. [49]Department of Neurology, University of Washington, Seattle, Washington, USA. [50]Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts, USA. [51]Department of Psychiatry, Massachusetts General Hospital/Harvard Medical School, Boston, Massachusetts, USA. [52]Department of Neurology, University of California, San Francisco, San Francisco, California, USA. [53]Department of Neuroscience, Mount Sinai School of Medicine, New York, New York, USA. [54]Department of Psychiatry, Mount Sinai School of Medicine, New York, New York, USA. [55]Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, New York, USA. [56]Department of Pathology and Immunology, Washington University, St. Louis, Missouri, USA. [57]University of South Florida (USF) Health Byrd Alzheimer's Institute, University of South Florida, Tampa, Florida, USA. [58]Department of Psychiatry, Vanderbilt University, Nashville, Tennessee, USA. [59]Department of Pathology, University of Alabama at Birmingham, Birmingham, Alabama, USA. [60]Department of Neurology, University of Southern California, Los Angeles, California, USA. [61]Department of Neurology, University of Alabama at Birmingham, Birmingham, Alabama, USA. [62]Department of Neurology, University of California, Irvine, Irvine, California, USA. [63]Department of Psychiatry and Behavioral Sciences, Miller School of Medicine, University of Miami, Miami, Florida, USA. [64]Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri, USA. [65]Hope Center Program on Protein Aggregation and Neurodegeneration, Washington University School of Medicine, St. Louis, Missouri, USA. [66]Department of Neurology, University of California, Davis, Sacramento, California, USA. [67]Institute for Memory Impairments and Neurological Disorders, University of California, Irvine, Irvine, California, USA. [68]Department of Neuroscience, Mayo Clinic, Jacksonville, Florida, USA. [69]Wien Center for Alzheimer's Disease and Memory Disorders, Mount Sinai Medical Center, Miami Beach, Florida, USA. [70]Department of Neurology, Mayo Clinic, Jacksonville, Florida, USA. [71]Department of Medical and Molecular Genetics, Indiana University, Indianapolis, Indiana, USA. [72]Department of Neurology, Indiana University, Indianapolis, Indiana, USA. [73]Department of Psychiatry, New York University, New York, New York, USA. [74]CS Kubik Laboratory for Neuropathology, Massachusetts General Hospital, Charlestown, Massachusetts, USA. [75]Department of Neurosciences, University of California, San Diego, La Jolla, California, USA. [76]Department of Pathology and Laboratory Medicine, Emory University, Atlanta, Georgia, USA. [77]Emory Alzheimer's Disease Center, Emory University, Atlanta, Georgia, USA. [78]Neurogenetics Program, University of California, Los Angeles, Los Angeles, California, USA. [79]Department of Pathology and Laboratory Medicine, Indiana University, Indianapolis, Indiana, USA. [80]Department of Neurology, Emory University, Atlanta, Georgia, USA. [81]Division of Genetics, Department of Medicine and Partners Center for Personalized Genetic Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA. [82]Department of Neurology, Massachusetts General Hospital/Harvard Medical School, Boston, Massachusetts, USA. [83]Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA. [84]Department of Pathology (Neuropathology), University of Pittsburgh, Pittsburgh, Pennsylvania, USA. [85]Sanders-Brown Center on Aging, Department of Molecular and Biomedical Pharmacology, University of Kentucky, Lexington, Kentucky, USA. [86]Department of Pathology, Duke University, Durham, North Carolina, USA. [87]Sanders-Brown Center on Aging, Department of Neurology, University of Kentucky, Lexington, Kentucky, USA . [88]Department of Pathology and Laboratory Medicine, University of California, Davis, Sacramento, California, USA. [89]Department of Biostatistics, Boston University, Boston, Massachusetts, USA. [90]Department of Ophthalmology, Boston University, Boston, Massachusetts, USA. [91]University of Pittsburgh Alzheimer's Disease Research Center, Pittsburgh, Pennsylvania, USA. [92]Department of Neurology, Oregon Health & Science University, Portland, Oregon, USA. [93]Department of Neurology, Portland Veterans Affairs Medical Center, Portland, Oregon, USA. [94]Department of Pathology and Laboratory Medicine, University of California, Irvine, Irvine, California, USA. [95]Department of Neurology, Boston University, Boston, Massachusetts, USA. [96]Department of Pathology, Boston University, Boston, Massachusetts, USA. [97]Department of Neuropsychology, University of California, San Francisco, San Francisco, California, USA. [98]Department of Molecular & Medical Genetics, Oregon Health & Science University, Portland, Oregon, USA. [99]Department of Epidemiology, University of Washington, Seattle, Washington, USA. [100]Department of Neurobiology and Behavior, University of California, Irvine, Irvine, California, USA. [101]Department of Pathology, University of Washington, Seattle, Washington, USA. [102]Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, Washington, USA. [103]Department of Pathology, University of Michigan, Ann Arbor, Michigan, USA. [104]Department of Psychiatry, Johns Hopkins University, Baltimore, Maryland, USA. [105]Department of Preventive Medicine, University of Southern California, Los Angeles, California, USA. [106]Department of Medicine–Pulmonary, New York University, New York, New York, USA. [107]Department of Neurology, University of Miami, Miami, Florida, USA. [108]Department of Pathology, University of California, San Diego, La Jolla, California, USA. [109]Cognitive Neurology and Alzheimer's Disease Center, Northwestern University, Chicago, Illinois, USA. [110]Department of Pathology, University of Southern California, Los Angeles, California, USA. [111]Department of Neurology, Washington University, St. Louis, Missouri, USA. [112]Department of Anatomic Pathology, Mayo Clinic, Rochester, Minnesota, USA. [113]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota, USA. [114]Department of Neurology, Mayo Clinic, Rochester, Minnesota, USA. [115]Arizona Alzheimer's Consortium, Phoenix, Arizona, USA. [116]Banner Alzheimer's Institute, Phoenix, Arizona, USA. [117]Neurogenomics Division, Translational Genomics Research Institute, Phoenix, Arizona, USA. [118]Alzheimer's Disease Center, New York University, New York, New York, USA. [119]Department of Neurology, University of Texas Southwestern, Dallas, Texas, USA. [120]Department of Radiology and Imaging Sciences, Indiana University, Indianapolis, Indiana, USA. [121]Department of Pathology (Neuropathology), Rush University Medical Center, Chicago, Illinois, USA. [122]Department of Psychiatry, University of Southern California, Los Angeles, California, USA. [123]Department of Pathology, Johns Hopkins University, Baltimore, Maryland, USA. [124]Sanders-Brown Center on Aging, Department of Anatomy and Neurobiology, University of Kentucky, Lexington, Kentucky, USA. [125]Department of Pathology & Laboratory Medicine, University of California, Los Angeles, Los Angeles, California, USA. [126]Department of Psychiatry & Behavioral Sciences, Duke University, Durham, North Carolina, USA. [127]Department of Pathology, Oregon Health & Science University, Portland, Oregon, USA. [128]Evelyn F McKnight Brain Institute, Department of Neurology, Miller School of Medicine, University of Miami, Miami, Florida, USA. [129]Department of Medicine, University of Washington, Seattle, Washington, USA. [130]Present address: Department of Biostatistics and Epidemiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA. Correspondence should be addressed to D.C.W. (whitcomb@pitt.edu) or B.D. (devlinbj@upmc.edu).

© 2012 Nature America, Inc. All rights reserved.

npg

ONLINE METHODS

Subject recruitment. Details of the recruitment of cases and controls are reported in **Supplementary Table 1**. All studies were conducted under institutional review board (IRB)-approved protocols. All cases were prospectively ascertained after signing IRB-approved informed consent forms.

Stage 1 samples. All 758 stage 1 case samples were from the NAPS2 (ref. 5), were diagnosed with chronic pancreatitis and were characterized for alcohol-related pancreatitis (**Table 1**). Chronic pancreatitis occurs in less than 0.05% of the population, such that a convenience sample provides essentially identical power as a same-sized sample of controls selected for the absence of chronic pancreatitis[32]. For controls, we used genotypes from 4,076 cases and controls from ADGC and 493 NAPS2 subjects, all genotyped on the same platform as the chronic pancreatitis samples.

Stage 2 samples. The stage 2 samples consisted of 343 chronic pancreatitis and 627 recurrent acute pancreatitis cases (**Table 1** and **Supplementary Table 1**) as well as 4,191 control subjects (3,986 from NGRC and 205 NAPS2 controls).

Genotypes. All cases and NAPS2 controls were genotyped by the University of Pittsburgh Genomics and Proteomics Core Laboratories using the Illumina HumanOmniExpress BeadChip. Samples were processed and scanned using the manufacturer's recommended protocols with no modifications. ADGC samples[33] were also genotyped using Illumina HumanOmniExpress BeadChips, whereas NGRC samples[34] were genotyped on the Illumina Human1M-Duo DNA Analysis BeadChip.

Quality control for stage 1. Quality control was performed for individuals and then SNPs to determine which samples and SNPs should not be included in the analysis. By assessing sex miscalls on the basis of X-chromosome genotypes using PLINK[35], we excluded 7 chronic pancreatitis cases and 20 controls (10 NAPS2 and 10 ADGC). On the basis of a requirement for ≥ 95% complete genotypes per individual, 40 cases and 27 controls (20 NAPS2 and 7 ADGC) were excluded. By searching for duplicate or highly related samples using GCTA software[36] (defined as those samples with a genetic relationship matrix score (GRM) of >0.4), we excluded 35 cases and 78 controls (2 NAPS2 and 76 ADGC). After applying these quality control filters, 676 cases and 4,507 controls remained for association analysis.

SNP quality control filtering was first performed on NAPS2 and ADGC samples separately. Ancestry was estimated using dacGem[37], taking into consideration data from 9,700 SNPs that had genotype completion rates of ≥99.9% and minor allele frequencies (MAFs) of ≥0.05 that were separated by at least 500 kb. Analysis of genotypes from NAPS2 subjects identified one significant dimension of ancestry and clustered subjects into three groups (**Supplementary Fig. 1**). Groups A and B (shown in **Supplementary Fig. 1**) delineate 764 and 282 subjects, respectively, of European ancestry (self-identified); SNP quality control filtering by MAF and Hardy-Weinberg equilibrium (HWE) were performed on data from these subjects. Of 731,442 SNPs for which data were available, 633,790 passed quality control filters. SNPs were excluded on the basis of map location (3,165), call rate (11,977), MAF of <0.01 (77,300) and departure from HWE ($P < 0.005$) (5,219).

ADGC data were available for 3 subsets of 1,763, 1,110 and 1,266 subjects. In the first subset, data were available for 659,224 SNPs, whereas, in the second and third subsets, data were available for 730,525 SNPs. After quality control filtering as described for the chronic pancreatitis cohort, including harmonization with SNPs passing quality control in the chronic pancreatitis cohort, 604,059, 632,761 and 633,023 SNPs remained in each of the three subsets, respectively. After merging cohorts, 30 related subjects were excluded, leaving 4,046 ADGC subjects. Of the 633,615 unique SNPs in this ADGC cohort, quality control filtering excluded 5 for low MAF and 5,316 for departure from HWE, leaving 628,294 SNPs. Combining the ADGC and chronic pancreatitis cohorts and performing another round of quality control yielded 625,739 SNPs for analysis.

Quality control for stage 2. Quality control for individuals in stage 2 was performed as described for stage 1. These individual-specific quality control filters excluded 60 cases, leaving 331 chronic pancreatitis and 579 recurrent acute pancreatitis cases for analysis; 14 controls were also excluded, leaving 4,177 controls for analysis. We analyzed all SNPs passing quality control in stage 1.

Association analysis. To control confounding due to ancestry, the first ten major eigenvectors from spectral decomposition were used as covariates in stage 1 and stage 2 analyses[38], although only one was significant. We contrasted the genotypes of cases and controls via logistic regression and a log-additive (logit) model using PLINK[35]. Genotypes for any SNPs showing association at $P < 5 \times 10^{-7}$ were manually inspected for valid genotype clustering. SNPs showing poor quality clustering were excluded. We set the overall significance threshold to $P = 5 \times 10^{-8}$, with $P = 5 \times 10^{-7}$ being strongly suggestive of association[7].

To determine whether alcohol status interacts with genetic variation to alter risk of pancreatitis, data from cases were fit to a general linear model in which the counts of alleles or genotypes predicted alcohol etiology (yes/no). The test statistic was obtained as a χ^2 likelihood ratio. In these analyses and all other analyses other than genome-wide association analysis, we modeled the male genotypes as 0 and 2. In the genome-wide association analysis, PLINK sets the count of minor alleles in males as 0 and 1 and includes a sex effect, but the use of 0 and 2 for the encoding for males is a more powerful approach[39,40].

DNA extraction. DNA was obtained using standard methods[41].

Pancreatic tissue processing. Tissue was obtained from two sources (Pitt and Pancreatic Adenocarcinoma Gene-Environment Risk (PAGER) from the University of Pittsburgh and PSU from Pennsylvania State University) and processed in three batches: banked (Pitt) and prospectively collected (PAGER) surgical waste from uninvolved pancreas and normal pancreas specimens from the Gift of Life Program that were not used for transplantation (PSU). PAGER samples were snap frozen, placed in RNAlater solution (Ambion) and stored at −80 °C. PSU pancreas samples were also snap frozen and preserved in formalin or placed in RNAlater solution. RNA was isolated using TRIzol reagent (Invitrogen), and its quality was examined by running samples on 1% agarose gels stained with ethidium bromide. cDNA was transcribed using oligo(dT) primers and the Superscript II reverse transcriptase kit (Invitrogen).

Gene expression. Relative expression of *PRSS1*, *PRSS2*, *CTRC* and 18S RNA was determined by analyzing cDNA using TaqMan-based RT-PCR assays (Applied Biosystems). Raw absolute quantitation results were analyzed and converted to relative expression results by software packages SDS V2.3 and DataAssist V1.0 (Applied Biosystems). Assays were repeated in triplicate or quadruplicate. Three sets of samples were assessed, two from Pitt (n = 10 and 22) and one from PSU (n = 37). PSU results were normalized against the levels of 18S RNA, and Pitt results were normalized against the levels of *CTRC*. From each of these three data sets, mean gene expression per sample was regressed against allele count to obtain an estimated slope, standard error and Z score. We then calculated an overall Z score as a weighted average of the individual Z scores, with weights determined by sample size.

Antibodies. Claudin proteins expression was assessed by protein blot analysis with mouse antibodies to claudin-2 (32-5600, Invitrogen) and claudin-4 (32-9400, Invitrogen), and protein blot analysis with mouse antibody to α-tubulin (AA12.1, The Developmental Studies Hybridoma Bank at the University of Iowa) was used to control for loading. Immunohistochemistry was performed using the antibody to claudin-2 (1:1,000 dilution). Immunoflourescence was performed using the antibody to claudin-2 and goat antibody to human CD68 (sc-7082, Santa Cruz Biotechnology). The secondary antibodies for immunofluorescence were Cy3-conjugated goat antibody to mouse and Cy5-conjugated antibody to goat from Jackson ImmunoResearch.

SDS-PAGE and protein blotting. Protein homogenates for protein blotting were obtained from snap-frozen tissue that was homogenized and sonicated in lysis buffer (50 mM Tris (pH 7.5), 150 mM NaCl, 0.4 mM EDTA, 1 mM EGTA, 1% sodium-deoxycholate, 1% Triton X-100, 0.1% sodium azide, 0.2 mM Na_3VO_4) and complete protease inhibitor mixture (Roche Diagnostics). Protein concentration was determined by bicinchoninic acid (BCA) method using a BCA protein assay kit from Thermo Scientific Pierce. Proteins were

© 2012 Nature America, Inc. All rights reserved.

npg

separated by 12% SDS-PAGE[42] and transferred to polyvinylidene difluoride (PVDF) membranes[43] for protein blotting[44]. Immunodetection of bound antibodies on PVDF membrane was performed using enhanced chemiluminescence (ECL) reagents (Amersham Biosciences). All procedures were carried out according to the manufacturer's instructions.

Immunohistochemistry. Standard automated immunohistochemistry was performed for claudin-2 on formalin-fixed, paraffin-embedded tissue sections that were 5 µm thick. After deparaffinization in xylene and rehydration in ethanol, antigen retrieval was performed using EDTA buffer, pH 8. Dako Autostainer Plus was used; slides were incubated for 30 min with the primary antibodies and then incubated with the secondary reagent (Mach 2 Mouse HRP Polymer from Biocare Medical) for 30 min. The chromogen was developed (Dako DAB+) for 10 min. Cytoplasmic, granular and membranous staining, predominantly in the lateral cell membranes, was graded on an intensity scale of 0–4 (0, no staining; 1, weak staining; 2, moderate staining; 3, strong staining). The staining intensity was very patchy from lobule to lobule in most samples.

Immunofluorescence. Cryostat sections (5 µm thick) of pancreas were washed three times with PBS and then washed three times with 0.5% BSA in PBS. Sections were blocked with 2% BSA solution for 30 min. Slides were incubated for 1 h at room temperature with primary antibody for claudin-2 (1:100 dilution) and goat antibody to human CD68 in 0.5% BSA solution. Slides were washed three times with BSA solution and incubated for 1 h at 20 °C with Cyr5-conjugated secondary antibody to goat (1:500 dilution) and Cy3-conjugated goat secondary antibody to mouse (1:1,000 dilution) in BSA solution. Nuclei were stained with Hoechst's dye (bisbenzimide, 1 mg in 100 ml of water) for 30 s. After three rinses with PBS, coverslips were applied with Gelvatol mounting medium. Fluorescent images were captured with an Olympus Fluoview 1000 confocal microscope (software version 1.7a). The Cy5 signal (CD68) was pseudocolored green, such that colocalization with the red claudin-2 signal appeared yellow.

32. Bacanu, S.A., Devlin, B. & Roeder, K. The power of genomic control. *Am. J. Hum. Genet.* **66**, 1933–1944 (2000).
33. Naj, A.C. *et al.* Common variants at *MS4A4*/*MS4A6E*, *CD2AP*, *CD33* and *EPHA1* are associated with late-onset Alzheimer's disease. *Nat. Genet.* **43**, 436–441 (2011).
34. Hamza, T.H. *et al.* Genome-wide gene-environment study identifies glutamate receptor gene *GRIN2A* as a Parkinson's disease modifier gene via interaction with coffee. *PLoS Genet.* **7**, e1002237 (2011).
35. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
36. Yang, J., Lee, S.H., Goddard, M.E. & Visscher, P.M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
37. Klei, L., Kent, B.P., Melhem, N., Devlin, B. & Roeder, K. GemTools: a fast and efficient approach to estimating genetic ancestry. Preprint at http://arxiv.org/pdf/1104.1162. (2011).
38. Price, A.L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
39. Clayton, D.G. Sex chromosomes and genetic association studies. *Genome Med.* **1**, 110 (2009).
40. Zheng, G., Joo, J., Zhang, C. & Geller, N.L. Testing association for markers on the X chromosome. *Genet. Epidemiol.* **31**, 834–843 (2007).
41. Pfützer, R.H. *et al. SPINK1*/*PSTI* polymorphisms act as disease modifiers in familial and idiopathic chronic pancreatitis. *Gastroenterology* **119**, 615–623 (2000).
42. Laemmli, U.K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685 (1970).
43. Towbin, H., Staehelin, T. & Gordon, J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA* **76**, 4350–4354 (1979).
44. Burnette, W.N. "Western blotting": electrophoretic transfer of proteins from sodium dodecyl sulfate–polyacrylamide gels to unmodified nitrocellulose and radiographic detection with antibody and radioiodinated protein A. *Anal. Biochem.* **112**, 195–203 (1981).

© 2012 Nature America, Inc. All rights reserved.

npg

doi:10.1038/ng.2466